



---

Publicly Accessible Penn Dissertations

---

Fall 12-22-2010

# The Tumor Suppressor Protein P53 and the Antagonizing Oncoprotein E6 from Human Papillomavirus Type 16

Kimberly A. Malecka

University of Pennsylvania, [kamalecka@gmail.com](mailto:kamalecka@gmail.com)

Follow this and additional works at: <http://repository.upenn.edu/edissertations>

 Part of the [Biochemistry Commons](#)

---

## Recommended Citation

Malecka, Kimberly A., "The Tumor Suppressor Protein P53 and the Antagonizing Oncoprotein E6 from Human Papillomavirus Type 16" (2010). *Publicly Accessible Penn Dissertations*. 277.  
<http://repository.upenn.edu/edissertations/277>

This paper is posted at Scholarly Commons. <http://repository.upenn.edu/edissertations/277>  
For more information, please contact [libraryrepository@pobox.upenn.edu](mailto:libraryrepository@pobox.upenn.edu).

---

# The Tumor Suppressor Protein P53 and the Antagonizing Oncoprotein E6 from Human Papillomavirus Type 16

## **Abstract**

The tumor suppressor protein p53 transactivates genes involved in cell cycle arrest and apoptosis in response to DNA damage, cellular stress and some oncogenic proteins. The wild type polypeptide chain of p53 has four distinct domains, including the sequence specific DNA binding core domain and the C terminal tetramerization domain. p53 must retain its ability to oligomerize and bind DNA targets in vivo to fulfill its function. Using X-ray crystallography and a crosslinking strategy, the structure of a wild type tetrameric p53 core domain bound to its consensus sequence was solved. This structure gives insight into DNA bend, core domain-DNA binding cooperativity and surface residue conservation within the tetramer.

Many cancers result from cells harboring mutant forms of p53 that are unable to function due to their inability to oligomerize, DNA binding defects or, in the case of human papillomavirus infection, p53 degradation.

High risk forms of the human papillomavirus (HPV) infect the basal layers of stratified epithelia and express two oncoproteins, called E6 and E7, which can lead to cell cycle disruption and cervical cancer. E6 mediates its cell transformation, in part, by forming a complex with the cellular E3 ligase E6-Association Protein (E6AP) to target p53 for degradation by the ubiquitin- proteasome pathway. A high throughput solution screen was designed to search for small molecule inhibitors of the E6 / E6AP interaction. Of the 80,000 compounds that were screened, 30 inhibitors with IC50 values in the low-micromolar to mid-nanomolar range were confirmed using secondary assays. Two compounds were shown to specifically block p53 degradation and promote apoptosis in cell lines infected with high risk forms of HPV. These HPV-E6 inhibitors provide a framework for developing HPV inhibitors with possible therapeutic applications.

## **Degree Type**

Dissertation

## **Degree Name**

Doctor of Philosophy (PhD)

## **Graduate Group**

Chemistry

## **First Advisor**

Ronen Marmorstein

## **Keywords**

p53, human papillomavirus, cell cycle, cervical cancer, E6

## **Subject Categories**

Biochemistry

THE TUMOR SUPPRESSOR PROTEIN p53 AND  
THE ANTAGONIZING ONCOPROTEIN E6 FROM HUMAN PAPILLOMAVIRUS TYPE 16

Kimberly Ann Malecka

A DISSERTATION

In

Chemistry

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy

2010

---

Dr. Ronen Marmorstein, Ph.D., Professor, The Wistar Institute

Supervisor of Dissertation

---

Dr. Gary Molander, Ph.D., Professor, Department of Chemistry

Graduate Group Chairperson

Dissertation Committee

Dr. David W. Christianson, Ph.D., Professor, Department of Chemistry, Committee Chair

Dr. Ivan Dmochowski, Ph.D., Professor, Department of Chemistry

Dr. Emmanuel Skordalakes, Ph.D., Department of Chemistry

THE TUMOR SUPPRESSOR PROTEIN p53 AND  
THE ANTAGONIZING ONCOPROTEIN E6 FROM HUMAN PAPILLOMAVIRUS TYPE 16

COPYRIGHT

2010

Kimberly Ann Malecka

I dedicate this thesis to  
Marge, Jim, Carolyn, Christine & John

I could never have achieved this without all of you.

*“Why, anybody can have a brain. ...Back where I come from, we have universities, seats of great learning, where men go to become great thinkers. And when they come out, they think deep thoughts and with no more brains than you have.”*

*Fleming V. (1939) The Wizard of Oz*

## ACKNOWLEDGEMENTS

First, I need to thank my wonderful advisor and mentor, Dr. Ronen Marmorstein. He provided me with excellent projects, guidance, coworkers, and experiences that I will carry with me long after I have left his laboratory. Without his endless confidence, enthusiasm, and experience, I would never have accomplished any of the following work. I have had the opportunity to be a member of his laboratory through happy and sad times over the past five years and have gotten to know his personality both inside and outside the laboratory. I cannot say where my career will lead me, but I will strive to be the type of scientist and leader that he has proven to be.

I would like to thank past and present members of my thesis committee: Dr. David Christianson, Dr. Ivan Dmochowski, Dr. Emmanuel Skordalakes, Dr. Eric Meggers, and Dr. Michael Therien. I have appreciated their patience and discussions not only at our annual gatherings, but throughout the entire year. I also want to express my gratitude towards Dr. David Schultz, the head of the Wistar Institute's Small Molecule Screening Laboratory, for patiently working with me over the past two years.

I must also thank Dr. Hao Zhu, Dr. Kevin Larade, and Dr. H. Franklin Bunn for my first laboratory experiences at Boston's Brigham and Women's Hospital. Without their guidance and enthusiasm for science, I am not sure I would have found the confidence to apply to graduate school.

I am indebted to both Dr. Xin Liu and Dr. William Ho. Dr. Xin Liu, in addition to guiding me throughout my first few months in the Marmorstein laboratory, offered me his previous

work and wisdom for working with the exceptionally difficult human papillomavirus type 16 E6 protein. In frustration, he mentioned to me once that I should abandon the project all together, but I persevered due, in large part, because of the strong scientific example he set in the laboratory. Dr. William Ho and his wife, Dr. Mary Ho, worked endlessly to teach me the finer points of the p53 project. I am grateful that Will entrusted the continuation of his project to me; I hope I did him proud.

To all the Marmorstein laboratory members, both past and present: I thank you for all your support, conversations, and laughter throughout the past five years. In particular, I thank Dr. Michael Brent, Dr. Jasna Maksimoska, Dr. Brandi Sanders, Dr. Ruchi Anand, Dr. Santosh Hodawadekar, and Daniela Fera. I also would like to thank Trey Shroeder, Sarah Johnson and Brittany Riggle for their help on the E6 project. I wish everyone the best in their future endeavors.

I am thankful to the Department of Chemistry at the University of Pennsylvania and the Structural Biology Training grant for each partially funding my graduate career and the Wistar Institute for providing a wonderful, collaborative work environment.

Lastly, I would like to try to express my gratitude towards my friends and family. I have met so many wonderful friends here at Penn and will continue to miss their conversations and presence as life takes us our separate ways. I also thank my family members for always acting interested in my somewhat specialized and difficult to understand research. In particular, I thank my sister Christine, my fabulous grandmother Marge, and my future husband John. Finally, I realize I can never thank my parents, Jim and Carolyn, enough, but I will try: thank you

for your endless support and for teaching me that an education is one of the most important things you can earn in your life.



## ABSTRACT

### THE TUMOR SUPPRESSOR PROTEIN p53 AND THE ANTAGONIZING ONCOPROTEIN E6 FROM HUMAN PAPILLOMAVIRUS TYPE 16

Kimberly Ann Malecka

Ronen Marmorstein

The tumor suppressor protein p53 transactivates genes involved in cell cycle arrest and apoptosis in response to DNA damage, cellular stress and some oncogenic proteins. The wild type polypeptide chain of p53 has four distinct domains, including the sequence specific DNA binding core domain and the C terminal tetramerization domain. p53 must retain its ability to oligomerize and bind DNA targets *in vivo* to fulfill its function. Using X-ray crystallography and a crosslinking strategy, the structure of a wild type tetrameric p53 core domain bound to its consensus sequence was solved. This structure gives insight into DNA bend, core domain-DNA binding cooperativity and surface residue conservation within the tetramer.

Many cancers result from cells harboring mutant forms of p53 that are unable to function due to their inability to oligomerize, DNA binding defects or, in the case of human papillomavirus infection, p53 degradation.

High risk forms of the human papillomavirus (HPV) infect the basal layers of stratified epithelia and express two oncoproteins, called E6 and E7, which can lead to cell cycle disruption and cervical cancer. E6 mediates its cell transformation, in part, by forming a complex with the cellular E3 ligase E6-Association Protein (E6AP) to target p53 for degradation by the ubiquitin-

proteasome pathway. A high throughput solution screen was designed to search for small molecule inhibitors of the E6 / E6AP interaction. Of the 80,000 compounds that were screened, 30 inhibitors with  $IC_{50}$  values in the low-micromolar to mid-nanomolar range were confirmed using secondary assays. Two compounds were shown to specifically block p53 degradation and promote apoptosis in cell lines infected with high risk forms of HPV. These HPV-E6 inhibitors provide a framework for developing HPV inhibitors with possible therapeutic applications.

# TABLE OF CONTENTS

<b>Dedication</b> .....	<b>.iii</b>
<b>Acknowledgements</b> .....	<b>iv</b>
<b>Abstract</b> .....	<b>vii</b>
<b>Table of Contents</b> .....	<b>ix</b>
<b>List of Figures</b> .....	<b>.xiii</b>
<b>List of Tables</b> .....	<b>.xvii</b>
<b>Chapter 1: Introduction</b> .....	<b>1</b>
1.1. A brief history of p53.....	2
1.1.1. <i>The Cell Cycle</i> .....	3
1.1.2. <i>Simian virus SV40 and p53's discovery</i> .....	6
1.2. Functions of p53 in healthy cells.....	10
1.2.1. <i>Reasons for p53 activation</i> .....	10
1.2.2. <i>Cell Cycle Arrest and Senescence</i> .....	13
1.2.3. <i>Apoptosis</i> .....	17
1.3. Cellular transformation and cancer.....	21
1.3.1. <i>General loss of p53</i> .....	22
1.3.2. <i>Loss of p53 in DNA tumor viruses</i> .....	26
1.4. Overview of dissertation.....	29

<b>Chapter 2: Crystal Structure of a p53 core tetramer bound to DNA. . . . .</b>	<b>30</b>
2.1. Introduction. . . . .	31
2.1.1. <i>p53 polypeptide chain and its consensus sequence for DNA binding. . . . .</i>	31
2.1.2. <i>Previously solved p53DBD structures: 1TSR, 2OCJ, 1HU8. . . . .</i>	37
2.1.3. <i>Convertible Nucleoside Approach and Crosslinking. . . . .</i>	43
2.1.4. <i>Previously solved p53DBD structures: 2GEQ, 2AC0, 2ATA, 2AHI, 2ADY. . . . .</i>	48
2.2. Materials and Methods. . . . .	57
2.2.1. <i>Growth and purification of mouse p53DBD. . . . .</i>	57
2.2.2. <i>Synthesis and purification of DNA containing unnatural nucleotides. . . . .</i>	59
2.2.2.1. <i>Base crosslinking DNA strands. . . . .</i>	59
2.2.2.2. <i>Backbone crosslinking DNA strands. . . . .</i>	61
2.2.3. <i>Purification of base crosslinked tetramer. . . . .</i>	63
2.2.4. <i>Crystallization, data collection and refinement of base crosslinked tetramers. .66</i>	
2.2.5. <i>Crystallization of backbone crosslinked and mixed crosslinking tetramers. . . . .</i>	69
2.3. Results and Discussion. . . . .	72
2.3.1. <i>Crystal structure of p53DBD tetramer bound to DNA. . . . .</i>	72
2.3.2. <i>Structure of each p53DBD within the tetramer and DNA contacts. . . . .</i>	78
2.3.3. <i>Structures of each p53DBD dimer and tethers. . . . .</i>	81
2.3.4. <i>Structure of the p53DBD tetramer and dimer-dimer contacts. . . . .</i>	84
2.3.5. <i>Comparison with other p53DBD tetramers. . . . .</i>	90
2.3.6. <i>Multiple tetramerization modes for p53 on DNA. . . . .</i>	94
2.3.7. <i>p53 DNA binding cooperativity. . . . .</i>	95
2.3.8. <i>p53DBD conservation and implication for function. . . . .</i>	96
2.3.9. <i>Base crosslinking and mixed crosslinking results. . . . .</i>	101
2.4. New p53 Tetramer. . . . .	106
2.4.1. <i>Self Assembled Tetramer (Chen et al, 2010) . . . . .</i>	106
2.5. p53 Tetramerization and Conclusions. . . . .	109
2.6. Future Directions. . . . .	110

<b>Chapter 3: Identification and characterization of small molecule HPV E6 inhibitors. . . . .</b>	<b>111</b>
3.1. Introduction. . . . .	111
3.1.1. Human papillomaviruses. . . . .	112
3.1.2. E6 and E7 oncoproteins. . . . .	116
3.1.3. Structural studies of E6. . . . .	119
3.1.4. Previous studies to inhibit E6 Function. . . . .	125
3.2. Materials and Methods. . . . .	128
3.2.1. Growth and purification of E6M-GST, MBP-E6AP, GST-E6AP, and E6M. . . . .	128
3.2.2. Pulldown of E6M and MBP-E6AP. . . . .	130
3.2.3. E6M complex purification and crystallization . . . . .	131
3.2.4. Enzyme-linked immunosorbance assay (ELISA) . . . . .	131
3.2.5. High throughput assay screening / data processing. . . . .	132
3.2.6. Fluorescent thermal stability assay. . . . .	133
3.2.7. In vitro p53 degradation assay. . . . .	134
3.2.8. Cell maintenance. . . . .	135
3.2.9. Cell assays: toxicity, p53 levels, APO-one. . . . .	135
3.3. High Throughput Screening Results and Discussion. . . . .	137
3.3.1. Development of high throughput screen and inhibitor identification. . . . .	137
3.3.2. Characterization of HPV-E6 inhibitors in vitro. . . . .	142
3.3.3. Characterization of HPV-E6 inhibitors in cells. . . . .	150
3.3.4. Discussion. . . . .	156
3.4. Crystallization Results and Discussion. . . . .	158
3.4.1. E6M + MBP-E6AP. . . . .	158
3.4.2. E6M + DSHS 00884. . . . .	162
3.5. Future Work. . . . .	168
APPENDIX: LC-MS Data for RH02007 and DSHS 00884. . . . .	169

<b>REFERENCES, Chapter 1</b> .....	<b>171</b>
<b>REFERENCES, Chapter 2</b> .....	<b>178</b>
<b>REFERENCES, Chapter 3</b> .....	<b>182</b>

## List of Figures

Figure 1.1: Schematic view of the eukaryotic cell cycle. . . . .	5
Figure 1.2: Overview of p53 signaling. . . . .	9
Figure 1.3: Activation of p53 and its downstream events. . . . .	12
Figure 1.4: MDM2 and p53 interaction. . . . .	16
Figure 1.5: The intrinsic and extrinsic apoptosis pathways. . . . .	19
Figure 1.6: Cancer derived p53 mutations. . . . .	23
Figure 2.1: The p53 polypeptide. . . . .	32
Figure 2.2: Tetramerization domain of p53, 1C26. . . . .	35
Figure 2.3: Pentamers of full p53 consensus sites. . . . .	36
Figure 2.4: Structure of human p53DBD bound to DNA, 1TSR. . . . .	38
Figure 2.5: Structure of human p53DBD, 2OCJ. . . . .	40
Figure 2.6: Structure of mouse p53DBD, 1HU8. . . . .	41
Figure 2.7: Noncrystallographic trimer of 1HU8. . . . .	42
Figure 2.8: Asymmetric unit of 1TSR. . . . .	44
Figure 2.9: Schematic crosslinking representation. . . . .	47
Figure 2.10: Structure of mouse p53DBD dimer bound to DNA, 2GEQ. . . . .	49
Figure 2.11: p53DBD cancer derived mutation histogram. . . . .	50
Figure 2.12: Electron density map in the area of the crosslink, 2GEQ. . . . .	52
Figure 2.13: Proposed p53DBD tetramer bound to DNA model. . . . .	53
Figure 2.14: Structure of human p53DBD tetramer bound to DNA, 2AC0. . . . .	55
Figure 2.15: DNA of 2AC0. . . . .	56
Figure 2.16: Gel filtration of p53DBD. . . . .	58

Figure 2.17: Scheme for creating DNA capable of base crosslinking. . . . .	60
Figure 2.18: Scheme for creating DNA capable of backbone crosslinking. . . . .	62
Figure 2.19: Purification of p53DBD tetramer by base crosslinking. . . . .	64
Figure 2.20: Samples of p53 tetramer by backbone crosslinking. . . . .	65
Figure 2.21: Crystals of p53DBD tetramers with base crosslinking. . . . .	67
Figure 2.22: Crystals of backbone crosslinked tetramers. . . . .	70
Figure 2.23: Crystals of mixed crosslinked tetramers. . . . .	71
Figure 2.24: DNA used for crystallization and view of crosslink. . . . .	74
Figure 2.25: MALDI chromatogram for DNA used in p53DBD tetramer crystallization. . . . .	75
Figure 2.26: Crystal packing of DNA 5' and 3' ends in structures 1 and 2. . . . .	76
Figure 2.27: Crystal packing of both structures 1 and 2. . . . .	77
Figure 2.28: Structure of the tetrameric p53DBD / DNA complex. . . . .	79
Figure 2.29: The L1 loops from several previously solved p53DBD structures. . . . .	80
Figure 2.30: View of the electron density at the crosslink. . . . .	83
Figure 2.31: p53DBD dimer-dimer formation. . . . .	85
Figure 2.32: Interactions at the dimer-dimer interface for structure 1. . . . .	86
Figure 2.33: Interactions at dimer-dimer interface for structure 2. . . . .	87
Figure 2.34: Superposition of the p53DBD tetramer with 2AC0. . . . .	92
Figure 2.35: Comparison of the DNA from the p53DBD tetramer and 2AC0. . . . .	93
Figure 2.36: Surface map showing the most highly conserved regions of p53DBD. . . . .	98
Figure 2.37: Surface map showing the dimer-dimer interface p53DBD tetramerization. . . . .	99
Figure 2.38: Surface map showing the back of the dimer-dimer interface. . . . .	100
Figure 2.39: Close up view of Cys274 and other options for crosslinking. . . . .	102
Figure 2.40: DNA bearing two backbone linkers. . . . .	104



Figure 2.41: DNA with mixed crosslinking. . . . .	105
Figure 2.42: 3KMD vs. 3EXJ tetramerization. . . . .	108
Figure 3.1: Genome of the human papillomavirus type 16. . . . .	114
Figure 3.2: Infection of epithelia by HPVs. . . . .	115
Figure 3.3: The concerted effort of E6 and E7 to cause cell cycle disruption. . . . .	118
Figure 3.4: Schematic representation of the full length HPV16 E6 protein. . . . .	120
Figure 3.5: NMR structure of the HPV16 C terminal domain. . . . .	123
Figure 3.6: Full length model of HPV16 E6. . . . .	124
Figure 3.7: Scheme for the high throughput screening assay. . . . .	139
Figure 3.8: Z factor graph. . . . .	140
Figure 3.9: Effect of inhibitors on E6 thermal stability. . . . .	145
Figure 3.10: p53 degradation assay results. . . . .	146
Figure 3.11: Effect of E6 inhibitors on cell toxicity. . . . .	152
Figure 3.12: Effect of E6 inhibitors on p53 degradation in SiHa (HPV16) infected cells. . . . .	153
Figure 3.13: Effect of E6 inhibitors on p53 degradation in HeLa (HPV18) infected cells . . . . .	154
Figure 3.14: Effect of E6 inhibitors on apoptosis in HeLa (HPV18) infected cells. . . . .	155
Figure 3.15: Pulldown of E6AP by E6M. . . . .	159
Figure 3.16: Purification of E6M + MBP-E6AP. . . . .	160
Figure 3.17: Crystals of the E6M + MBP-E6AP complex. . . . .	161
Figure 3.18: Thermostability assay curves for E6M + DMSO, DSHS 00884. . . . .	164
Figure 3.19: Concentration of the E6M + DSHS00884 complex. . . . .	165
Figure 3.20: Crystals of the E6M + DSHS 00884 complex. . . . .	166
Figure 3.21: Full gel of E6M + DSHS00884 concentration. . . . .	167
Figure A.1: LC-MS results for RH02007. . . . .	169

Figure A.2: LC-MS results for DSHS 00884. .... 170

## List of Tables

Table 1.1: Small molecule modulators of p53 and their drug trial status. ....	24
Table 1.2: Three DNA tumor viruses and their similar proteins. ....	28
Table 2.1: Crystallographic statistics for base crosslinked tetramers. ....	68
Table 3.1: Screening summary. ....	141
Table 3.2: Secondary assay results summary for inhibitor compounds. ....	147

## **CHAPTER 1**

### **Introduction**

## 1.1 A Brief History of p53

In 1979, several studies(DeLeo et al, 1979; Lane & Crawford, 1979; Linzer & Levine, 1979; Rotter et al, 1980) were published that discussed the identification a new cellular protein. Unbeknownst to the researchers, their discoveries would spur the advancement of an entirely new research field that would come to encompass cell cycle regulation, apoptosis, cellular senescence and, perhaps most interestingly, cancer. The 53 kilodalton (kDa) protein they discovered was coined p53(DeLeo et al, 1979).

Today, the role of p53 as a tumor suppressor protein and its nickname of “guardian of the genome”(Lane, 1992) are well known, but in the time soon after its discovery, researchers were convinced that p53 was actually an oncoprotein. Some early experiments even supported this hypothesis(Eliyahu et al, 1984; Jenkins et al, 1985; Jenkins et al, 1984a; Jenkins et al, 1984b), but its true function as a tumor suppressor and its importance in the cell has slowly been unraveled over the past thirty years.

### 1.1.1 The Cell Cycle

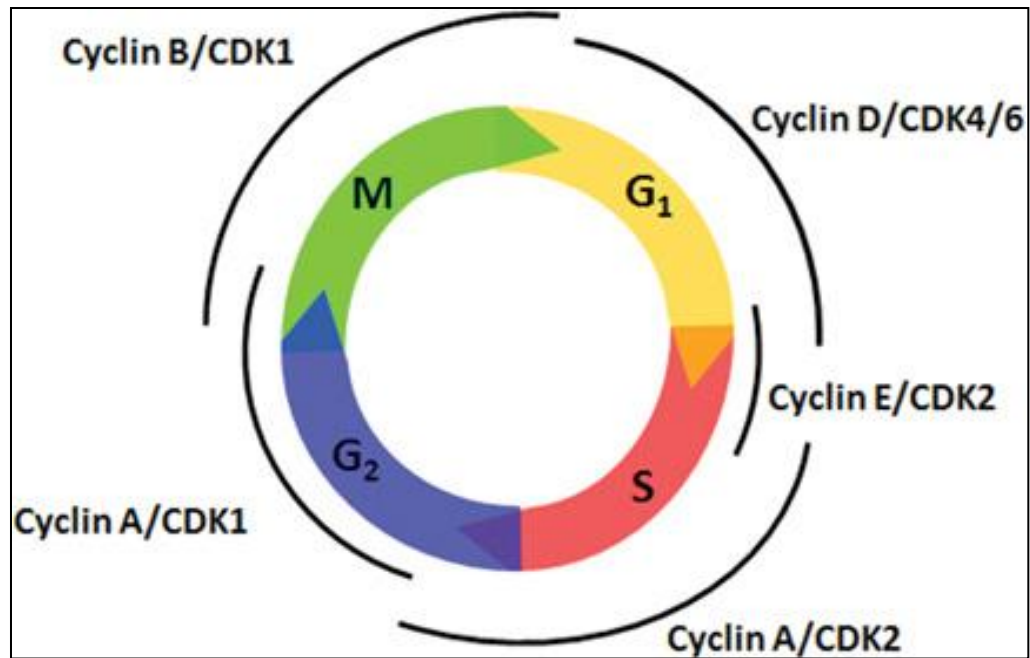
One of the cell's greatest successes is its ability to accurately reproduce itself. Mitosis, the multi-step process of parent cell splitting into its two daughter cells, is only one part of a much bigger and highly regulated cycle that ensures fidelity upon cellular replication (Lodish et al, 2004).

Broadly, the eukaryotic cell cycle is broken into four parts:  $G_1$ , S,  $G_2$  and M (Figure 1.1). During the S phase, the entire DNA molecule, which is broken into individual chromosomes, carried within the cell is replicated. During the M (mitosis) phase, the cell proceeds through the highly organized four phases (prophase, metaphase, anaphase, telophase) to accurately transport one entire set of chromosomes to each daughter cell. The S and M phases are separated by growth phases ( $G_1$  and  $G_2$ ), during which mRNA is transcribed and translated to make proteins and prepare the cell for division (Alberts et al, 2002; Lodish et al, 2004).

The eukaryotic cell cycle is regulated primarily by heterodimeric protein complexes comprised of both cyclins (regulatory subunit) and cyclin-dependent kinases (CDK, catalytic subunit). Only when bound by a cyclin and thus directed to a substrate can a CDK phosphorylate its targets. As cyclin concentrations change depending on where in the cycle the cell is, so do phosphorylation signals, which are used by the cell to exit one phase and enter another (Lodish et al, 2004). This regulation is illustrated by the following basic example: Near the end of the  $G_1$  phase, both  $G_1$  and S phase cyclin-CDK complexes are present in the cell, but the S phase complexes are bound by protein inhibitors. When the cell is ready to enter the S phase,  $G_1$  cyclin-CDK complexes are directed to the inhibitors and disrupt their interaction to the S phase cyclin-CDK complexes by phosphorylation. This signal leads to the inhibitors'

polyubiquitination and subsequent degradation. Now the S-phase cyclin-CDK complexes are free to phosphorylate the DNA replication machinery and begin the most important part of the S phase(Lodish et al, 2004).

Since the cycle is so highly regulated, it should be no surprise that the cell has developed strategies to protect itself when things go awry. The exact mechanism of this protection, in addition to how it related to cancer and oncogenic viruses, began to be more fully understood when p53 finally came on the scene.



**Figure 1.1: Schematic view of the eukaryotic cell cycle.** The phases of G<sub>1</sub>, S, G<sub>2</sub>, and M are shown as well as the major cyclin/CDK complexes that govern the progression of each phase. The figure is adapted from (Suryadinata et al, 2010).



### 1.1.2 *Simian Virus (SV40) and p53's discovery*

The simian vacuolating virus 40 (SV40) is intricately linked to the discovery of p53. The virus, which belongs to the polyomavirus genus, has a double stranded DNA genome and is associated with tumors(Lowe et al, 2007). In its host species of asian macaques, the virus does not cause any noticeable symptoms(Lowe et al, 2007). However, when fetal hamsters were intracerebrally inoculated with the virus, four out of the seven resulting animals developed ependymomas(Kirchstein & Gerber, 1962). Another study by Eddy et al. showed that sarcomas could be induced in hamsters after inoculation with SV40(Eddy et al, 1961).

Focus turned to one particular SV40 protein, known initially as protein A and currently as large T antigen. It was known to be expressed early in viral infection and have a role in viral DNA replication. Researchers wanted to study the individual protein but purification proved difficult because it was constantly pulling down other proteins that could have been either viral or cellular in origin(Tegtmeyer et al, 1975).

Through a variety of experiments in 1979 and 1980(DeLeo et al, 1979; Lane & Crawford, 1979; Linzer & Levine, 1979; Rotter et al, 1980), it was finally concluded that large T antigen was pulling down a cellular protein of 53 kDa, which Old and colleagues dubbed p53. Exactly what p53's role in the cell was had yet to be determined, however the fact that it was present in these tumor cells at higher levels than healthy cells lead researchers to believe it was an oncoprotein. Lane et al. even speculated in their paper that "it may have a crucial role in the modulation of the transformed state."

Four years later, three experiments started to suggest that p53 was no oncoprotein. Maltzmann and Czyzyk showed that irradiating mouse cells with ultraviolet light, which caused

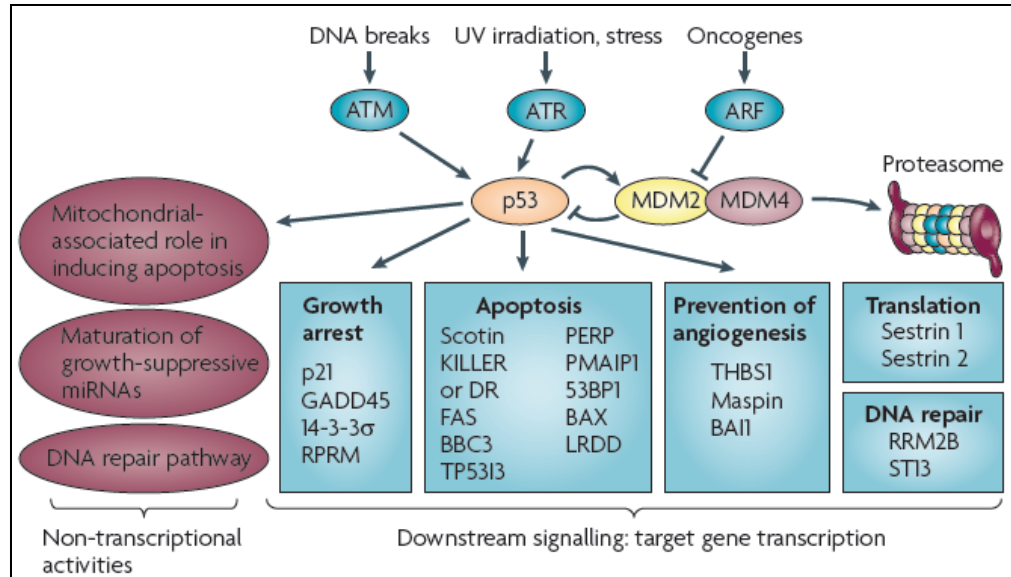
significant DNA damage, resulted in increased p53 levels(Maltzmann & Czyzyk, 1984). Different forms of p53 could be found in cells depending on whether the cell was dividing or not(Milner, 1984) and finally, Mowat et al. noticed differences in the *TP53* (human p53 gene) gene of leukemia cells versus healthy cells(Mowat et al, 1985).

This work was followed up by several studies that discovered p53 gene mutations in tumors and those affected with Li-Fraumeni Syndrome (FLS), which causes its sufferers to have a high susceptibility of childhood and adult cancers(Hollstein et al, 1991; Malkin et al, 1990; Srivastava et al, 1990). Eliyahu et al. and Finlay et al. were also able to suppress tumor formation when cells were overexpressing wild type p53(Eliyahu et al, 1989; Finlay et al, 1989). Moshe Oren summed up all the current p53 research in 1992(Oren, 1992) by declaring p53 as the “ultimate tumor suppressor.”

Today, the results of all these experiments are explained by the framework of p53’s known functions(Lowe et al, 2007) (Figure 1.2). In response to DNA damage, cell stress and certain oncogenic proteins, p53 transactivates genes involved in bringing about one of three outcomes: 1. temporary cell cycle arrest, 2. permanent cell cycle arrest in the form of senescence, 3. programmed cell death, also known as apoptosis(Lowe et al, 2007). These processes are essential for removing defective cells from the actively dividing cell population.

Consider a cell harboring detrimental mutations in its genome. It is in the cell’s best interest to either fix these mutations (cell cycle arrest) or remove itself (senescence, apoptosis) so these mutations are not carried on to daughter cells; p53 ensures that these processes occur. If p53 is inactive and unable to fulfill this function, then this defective cell will continue through the cell cycle, divide and exponentially add defective cells to the population. And, since nothing

is protecting these cells from gathering more deleterious mutations, they can easily go down the path of cellular transformation.



**Figure 1.2: Overview of p53 signaling.** The diagram shows simplistically what proteins are activated in response to different stresses, how those proteins report to p53 and what the response of p53. This figure is adapted from (Brown et al, 2009).

## 1.2 Functions of p53 in healthy cells

### 1.2.1 Reasons for p53 activation

When cultured, mouse embryonic fibroblasts (MEFs) will only actively divide for a short period of time compared with their human counterparts (Parrinello et al, 2003; Sherr & DePinho, 2000). Telomere shortening is a common reason for cell cycle arrest, particularly because it leads to chromosomal aberrations that may bring about cellular transformation. However, MEFs have long telomeres and, in some cases, express telomerase. So what else could be causing the cells to stop dividing? Parrinello et al. assert that high oxygen concentrations (20%) stress the MEFs, cause DNA damage and result in the cultured cells arresting their cell cycles. This work went on to show that when the oxygen was dropped to the more physiological concentration of 3%, the cells did not enter senescence (Parrinello et al, 2003).

Reasons for p53 activation and subsequent entrance into cell cycle arrest, senescence or apoptosis come in all forms (Figure 1.3). One of the most common reasons is DNA damage. Gamma or UV irradiation, alkylation of the bases, depurination of the DNA, telomere shortening, substitution of the ribose sugar and even mismatching of DNA bases during replication are all reasons for the cell to stop itself (Levine et al, 2006). Interestingly, all these problems have their own sets of guardians (proteins, pathways) that will detect and report to the cell the problems; the protein they all report to is p53. The same is true for other types of cellular stresses such as hypoxia, H<sub>2</sub>O<sub>2</sub> oxidative stress (Chen et al, 1998), glucose starvation, low ribosome biogenesis, damage to the cellular spindle and temperature shocks, among other

causes(Levine et al, 2006). p53 will also respond if oncogenic proteins, such as RAS(Di Micco et al, 2007; Ferbeyre et al, 2002), E2F(Di Micco et al, 2007; Dimri et al, 2000),

MOS(Di Micco et al, 2007), RUNX1(Wolyniec et al, 2009; Wotton et al, 2004) and RUNX1-ETO(Wolyniec et al, 2009), become active. Some other obscure stresses on the cell that invoke a p53 response are chemical inhibition of histone deacetylases (HDACs)(Munro et al, 2004) and depletion of chaperone proteins, such as Hsp70(Sherman et al, 2007). All of these situations result in reports to p53 that something is not going well and the cell should seriously reconsider dividing.

The fact that p53 is absent or nonfunctional in a large quantity of human cancers implies its importance and begs the question: why isn't p53's role in the cell backed up? Why did Mother Nature not offer us an alternative pathway should the first one become totally mucked up? The only answer that makes sense is that having one central protein receive all the information is far more efficient than several proteins receiving partial information. The situation is simplified and leads to the most effective response(Levine et al, 2006).



### 1.2.2 Cell Cycle Arrest and Senescence

After receiving the news that there are problems in the cell, p53 responds by both increasing in concentration and binding to specific sequences of DNA to activate the transcription of genes that are involved in either cell cycle arrest/senescence or apoptosis(Levine et al, 2006; Zuckerman et al, 2009). If the problem is minor and can be fixed, such as a few mismatched base pairs that were missed during editing in the S phase, then the cycle is stopped long enough for repair and it re-enters the cell cycle(Campisi & d'Adda di Fagnana, 2007). If the problem is more serious, like an irreparable DNA lesion, then arrest is maintained in the form of senescence or the cell undergoes apoptosis(Zuckerman et al, 2009).

p53's response to induce cell cycle arrest is well illustrated through the following example. In normal cells, pRb is bound to a protein known as E2F. Near the end of the G<sub>1</sub> phase, CyclinE/Cdk2 targets pRb for phosphorylation, which disrupts its interaction with E2F and is one of the main signals for the cell to enter the S phase(Zamanian & La Thangue, 1993). In cultured MEFs that have been stressed, it was shown that p53 stimulates the transcription of p21<sup>WAF1</sup>, which is an inhibitor of CyclinE/Cdk2 function. Hence, transcription of p21<sup>WAF1</sup> keeps the cell in the G<sub>1</sub> phase until the stress is removed(Smogorzewska & de Lange, 2002).

p53 induces the expression of other inhibitors of cell cycle cyclin/Cdk complexes, such as Gadd45 and 14-3-3. Gadd45 inhibits Cdc2-cyclinB complex formation(Jin et al, 2000; Zhan et al, 1999) while 14-3-3 binds to and sequesters Cdc2-cyclin B and cdc4525c to the cytoplasm(Chan et al, 1999; Peng et al, 1997). p53 can also work by repressing gene expression to stall the cell cycle. In particular, it represses Cdk4(Spurgers et al, 2006), cyclinE2(Spurgers et al, 2006), cdc25a, and cdc25c(Rother et al, 2007; St Clair et al, 2004).



In contrast to cell cycle arrest, senescence is when the cell leaves the cell cycle for an extended period of time or even permanently. Several pathways that bring about senescence by way of p53 are well understood and three pathways, ARF, PAI-1 and PML, have been characterized.

In a healthy cell, p53 levels are typically low and are turned over as fast as every 6 – 20 minutes (Levine et al, 2006). The main regulator of p53 levels is an E3 ligase protein known as Mdm2, which works in a negative feedback loop with p53 (Figure 1.4) (Lee & Gu, 2010). p53 stimulates the transcription of Mdm2, but Mdm2 binds to the N terminal transactivation domain of p53 to fulfill two functions: 1. blocking p53's ability to interact with transcriptional machinery to successfully transactivate downstream genes and, 2. marking p53 for degradation by the ubiquitin-proteasome pathway (Lee & Gu, 2010). In this way, as p53 levels go up so do Mdm2 levels, which then lead to greater p53 degradation and less Mdm2 transcription. The two proteins keep each other in check. This regulatory protein's role is so important in the cell that a knockout of the Mdm2's gene is embryonic lethal in mice (Luna et al, 1995).

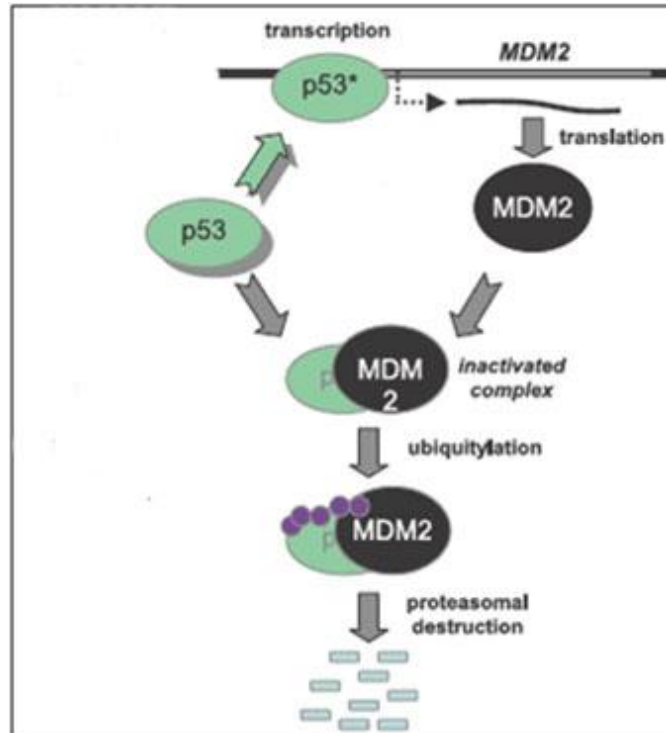
When the tumor suppressor protein ARF (p14ARF in humans, p19ARF in mouse) is upregulated, p53 is activated. This occurs because ARF binds MDM2 (HDM2 in humans) and inhibits its binding to p53, which in return increases p53 levels (Gil & Peters, 2006; Haupt, 2003). Higher protein levels of ARF are commonly seen in murine cells that have entered senescence and it appears to be required for both telomere and non-telomere-induced senescence in MEFs (Sharpless et al, 2004).

The expression of plasminogen activator-inhibitor-1 (PAI-1) is regulated by p53 binding to its mRNA at the 3'-UTR, which leads to increased PAI-1 levels (Shetty et al, 2008). PAI-1 is

used as a common marker of senescent cells(Mu & Higgins, 1995; Serrano et al, 1997) and, interestingly, once PAI-1 is downregulated, cells exit senescence(Kortlever et al, 2006).

Finally, the promyelocytic leukemia (PML) tumor suppressor protein responds to the upregulation of the oncogene RAS by binding to CBP/p300 and marking p53 for acetylation at Lys382. p53 is now stabilized as a result of the acetylation and induces other genes that lead the cell towards senescence(Ferbeyre et al, 2000; Pearson et al, 2000). Pearson et al. were also able to show that PML -/- fibroblasts lose their ability to enter senescence upon Ras activation(Pearson et al, 2000)

What specifically causes p53 to choose cell cycle arrest or senescence versus apoptosis is still widely unknown(Levine et al, 2006; Zuckerman et al, 2009). The response in each cell is greatly affected by the intensity and duration of the stress signal(Espinosa, 2008). Researchers have also realized that the status of p53 at the time of the signal also affects the outcome, where they define p53's status as conformation, localization, activity, and stability status (CLASS)(Espinosa, 2008). This remains an active area of research at this time(Campisi & d'Adda di Fagagna, 2007; Zuckerman et al, 2009).



**Figure 1.4: MDM2 and p53 interaction.** The negative feedback loop involving p53 and its regulator protein MDM2 is depicted. This figure was adapted from (Hardcastle, 2007).

### 1.2.3 Apoptosis

The term apoptosis comes from the Greek words *apo* (from) and *ptosis* (falling) and was first described by Carl Vogt in 1842 (Kerr, 1965). As opposed to necrosis, which is a chaotic and often messy cell death, apoptosis is the very organized destruction of a cell and is governed by a strict set of signals (Alberts et al, 2002).

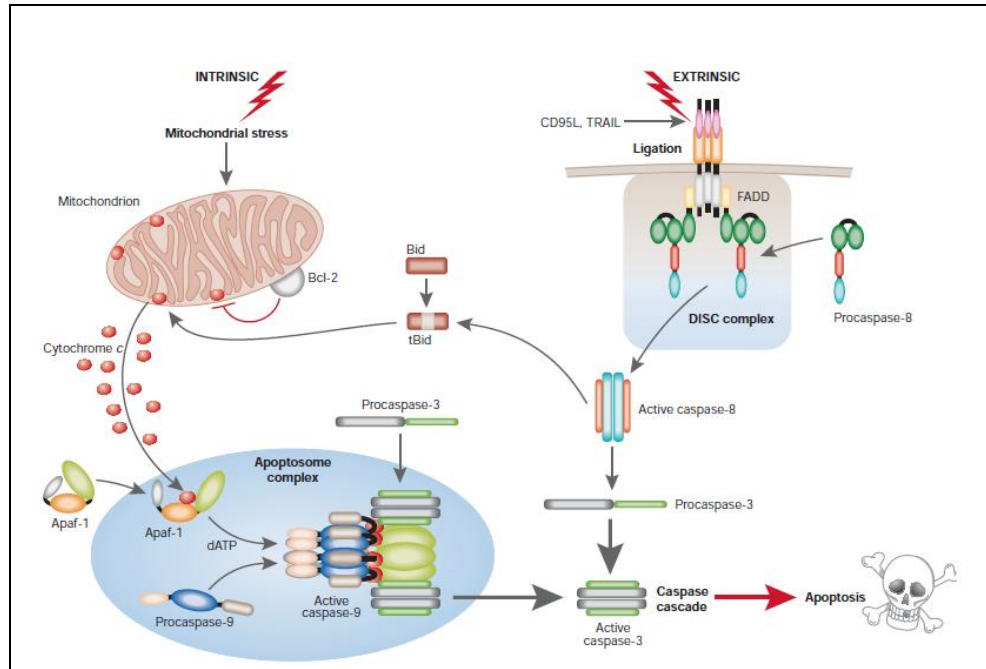
Apoptosis can be induced either extrinsically (signals begin outside the cell) or intrinsically (signals are sent by the cell itself) (Figure 1.5). Regardless of how the cell decides to apoptose, both pathways lead to the activation of proteases that are held in an inactive conformation, known as procaspases, until an apoptotic signal reaches it and uninhibits it, thus making an active caspase (Alberts et al, 2002). Initiator caspases (8 and 9) are the first to be uninhibited, which then target many effector caspases (3, 6 and 7) for activation and lead to all sorts of signals such as permeabilization of the mitochondria, destruction of the cellular cytoskeleton, and cleavage of the genomic DNA. During this time, the dying cell also displays markers on its surface that tell the surrounding cells about its death. Eventually, the surrounding cells can internalize pieces of the dying cell (phagocytosis), such as DNA nucleotides and amino acids, to use in their own functions (Alberts et al, 2002).

Extrinsic apoptosis begins by the binding of a ligand to a cell surface death receptor, such as Fas or tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ) receptor 1, by Fas ligand or TNF- $\alpha$  related apoptosis-inducing ligand (TRAIL) (MacFarlane & Williams, 2004). Upon binding, effector proteins, such as the Fas-associated death domain (FADD), are recruited to the receptor to activate caspase 8 and begin the caspase cascade (Alberts et al, 2002; MacFarlane & Williams, 2004).

Interestingly, p53 has been implicated in inducing the expression of Fas, TRAIL receptor 2 (Haupt et al, 2003), TRAIL death ligand (TNFSF10) and the Fas death ligand (FasL) (Kuribayashi et al, 2008; Maecker et al, 2000), which suggests that p53 can influence apoptotic signaling. Other genes involved in the apoptotic pathway, such as apaf-1, caspase 8 and caspase 6 also seem to be influenced by p53 activation (Haupt et al, 2003).

Intrinsic signals for apoptosis are initiated by cell stress and act first upon the mitochondria. Bcl2 family members, which are comprised of both pro- and anti-apoptotic proteins, coordinate to create pores in the mitochondrial membranes and release cytochrome c, which then coordinates with apoptotic protease-activating factor 1 (Apaf-1) to activate caspase 9 and the subsequent caspase cascade (MacFarlane & Williams, 2004).

To help out this intrinsic pathway, p53 induces the expression of OKL38, which in turn localizes to the mitochondria and helps release cytochrome c. Yao et al. showed that inhibition of OKL38 expression resulted in an increase of tumorigenesis while an overexpression induced apoptosis in tumor cell lines (Yao et al, 2008). Also working at the mitochondrial outer membrane, where it inserts itself and brings about pore formation, is the protein Bax (Vaseva & Moll, 2009). p53 induces the expression of Bax, as well as the pro-apoptotic genes *bid*, *puma* and *nox*, while repressing expression of anti-apoptotic proteins (Vaseva & Moll, 2009).



**Figure 1.5: The intrinsic and extrinsic apoptosis pathways.** The two signal pathways that propagate apoptosis is shown; both intrinsic (left) and extrinsic (right) signals are depicted in greater detail than described in the text. This figure was adapted from (MacFarlane & Williams, 2004).

Intriguingly, p53 has also been shown to bring about apoptosis in a transcription-independent fashion, most notably by localization to the mitochondria. Marchenko et al. showed that following a p53-dependent apoptotic signal, a fraction of the p53 population rapidly headed for the mitochondria and resulted in cytochrome c release (Marchenko et al, 2000). In a follow up work by the same laboratory, Moll and colleagues evidenced that p53 formed complexes with both Bcl-X<sub>L</sub> and Bcl2, facilitating cytochrome c release in irradiated thymocytes (Mihara et al, 2003).

Clearly, p53's role in cell cycle arrest, senescence, and apoptosis is varied and necessary. Through both transcriptional-dependent and independent pathways, the protein mediates its function to augment the cell cycle and protect the cellular population from unwanted problems.

### 1.3 Cellular transformation and cancer

According to the American Cancer Society, cancer is defined by “the uncontrolled growth and spread of abnormal cells.”(Society, 2010) All cancers arise due to mutations or malfunctions of the genes involved in cell cycle control(Society, 2010). Only 5% of cancers are hereditary so the remaining cases are due to the cellular damage that occurs during a person’s lifetime. In fact, 78% of cancer cases are diagnosed in those 55 years and older(Society, 2010). Due to recent advances, the 5 year survival rate for all patients diagnosed between 1999 and 2005 is up to 68%(Society, 2010). Research in novel cancer therapies is targeting new pathways, proteins, and cells from different angles all the time.

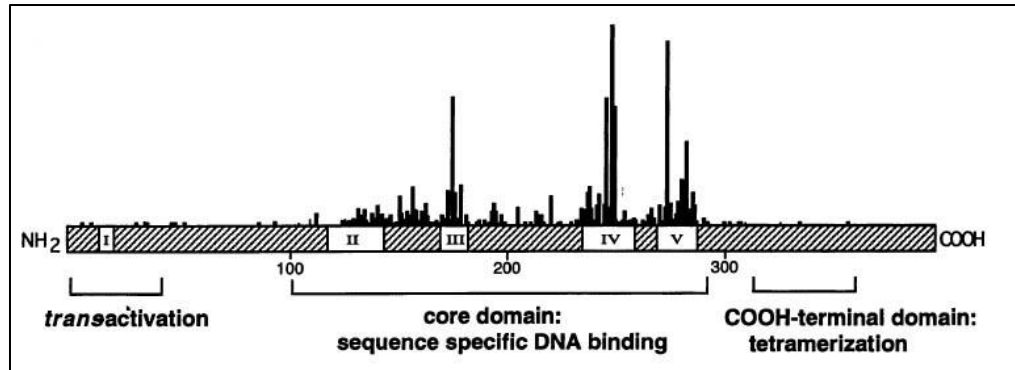
In a review by Lane and colleagues, it is claimed that nearly 11 million people have tumors harboring *TP53* genes which code for inactive protein while another 11 million have tumors in which other proteins involved in the p53 pathway are nonfunctional(Brown et al, 2009). Loss of p53 function appears to be one of the main risk factors for developing cancer over a person’s lifetime and p53 can be inactivated by mutation, gene loss or, as we will see, viral infection.



### 1.3.1 *General loss of p53 function*

Typically, loss of p53 function happens in one of two ways: 1. Point mutations in the p53 gene or 2. Abrogation of protein(s) function that is/are involved in the p53 signaling pathways(Brown et al, 2009).

The full length, wild type, human p53 protein is 393 amino acids and is comprised of four domains: the N terminal transactivation domain, the sequence specific DNA binding domain (p53DBD), the tetramerization domain and the C terminal regulatory domain. Interestingly, 50% of all the tumors bear mutations in the p53 gene(Wang & Sun, 2010) and 95% of these mutations are found in the DNA binding domain of the protein (Figure 1.6), which underscores the importance of p53's transactivating function to the cell. What is even more shocking is that some mutants of p53 are thought to actually gain functions that promote tumorigenesis (review(Brosch & Rotter, 2009)). Many researchers are focused on developing drugs that can "re-activate" p53's wild type properties and turn a mutant p53 into a fully functional tumor suppressor. Several candidates are currently going through preclinical and Phase 1 drug trials (Table 1.1).



**Figure 1.6: Cancer derived p53 mutations.** The full length protein of p53 is shown schematically and three domains are defined. The relative number of times and placement of a particular point mutation is depicted as a histogram above the schematic. The majority of tumor-derived mutations are found in the core/DNA binding domain. This figure was adapted from (Cho et al, 1994).

Molecule	Mechanism of action	Stage in clinical testing
<i>Reactivate mutant p53</i>		
PRIMA-1	Protein folding <sup>49</sup>	Phase I (APR-246)
CP-31398	Protein folding <sup>46</sup>	Preclinical
PhiKan083	Protein thermal stability <sup>69</sup>	Preclinical
<i>Activate wild-type p53</i>		
Nutlin	MDM2 binding <sup>8</sup>	Phase I
MI-219	MDM2 binding <sup>48</sup>	Phase I
Tenovin-6	SIRT1 and SIRT2 inhibition <sup>50</sup>	Preclinical
RITA	p53 binding <sup>51</sup>	Preclinical
Leptomycin B	CRM1 binding <sup>72,135</sup>	Phase I (Elactocin; withdrawn <sup>147</sup> )
Actinomycin D	RPL11 and RPL5 release <sup>118</sup>	Approved (Dactinomycin)
<i>Cyclotherapy (temporal combination of p53 activator and mitotic inhibitor)</i>		
Nutlin*	BI-2536 (PLK1 inhibitor <sup>52</sup> )*	Phase I/Phase I <sup>§</sup>
Nutlin*	VX680 (Aurora inhibitor <sup>9</sup> )*	Phase I/Phase I <sup>§</sup>
Tenovin-6*	Taxol (Tubulin binding <sup>52</sup> )*	Preclinical/approved <sup>§</sup>
Actinomycin D*	Taxol <sup>9</sup>	Approved/approved <sup>§</sup>

**Table 1.1: Small molecule modulators of p53 and their drug trial status.** Several known small molecules capable of stabilizing p53 levels through various mechanisms and their current stage in drug trials are listed. This table was adapted from (Brown et al, 2009).

In some cancers, the p53 gene is lost all together. Gene therapy is actively being attempted to reintroduce a wild type p53 gene to these lacking cells. For example, Ad-p53 is a replicative-defective adenovirus that bears wild type p53 to the cell with reasonable efficiency and low toxicity. Marketed under the name Gendicine or Advexin, the technique is currently in phase 1 clinical trials in the United States and seems to be particularly useful against head, neck, and lung cancers at inducing growth arrest and apoptosis.(Brooks & Gu, 2003; Pearson et al, 2004; Roth, 2006)

Obviously cancers can also exist when p53 is still present in its wild type form but unable to function due to members of its pathway being functionally defective(Wang & Sun, 2010). The best example of this is when Mdm-2 turns oncogenic by way of over-actively degrading p53, which occurs in nearly 7% of all human cancers(de Rozières et al, 2000; Luna et al, 1995; Momand et al, 1998). Small molecules, including nutlins, benzodiazepinediones (BDA) and some spiro-oxindoles derivatives have been developed to inhibit Mdm2 degradation and bring about p53 function (Review (Wang & Sun, 2010)). This type of therapy comes with some controversy because these compounds could incorrectly activate p53 in healthy cells or over-activate Mdm2's interaction with other known partner proteins that could lead to adverse effects(Ambrosini et al, 2007; LaRusch et al, 2007; Lau et al, 2008).

### 1.3.2 *Loss of p53 in tumor viruses*

Cancerous tumors can also arise due to viral infection because the virus has found a way to circumvent p53 function. The fact that one of large T antigen's jobs in SV40 infection was to bind p53, which explained why the two proteins were consistently purified together, and sequester it from service is neither surprising nor unique.

Once inside a cell, the goal of a viral particle is to replicate itself. For this to occur, the cell needs to be forced into DNA replication, which is a tightly regulated cellular process. It should come as no surprise that when a virus forces the cell into the S phase, it also has a back-up plan to deal with the inevitable cellular p53 response. Since cell cycle arrest, senescence or apoptosis are not conducive to viral replication, many viruses have evolved ways to circumvent this problem. For SV40, it was large T antigen's binding to p53. For others, like the adenovirus(Berk, 2005) and human papillomavirus(Huibregste et al, 1993; Scheffner, 1998; Scheffner et al, 1993), it is degradation of p53. From here, the cell cycle is completely dismantled, p53 is out of commission and cellular transformation is a very common result.

As already described, SV40 is a polyomavirus that has been shown to cause tumorigenesis in hamsters(Eddy et al, 1961; Kirchstein & Gerber, 1962). After the polio vaccine was accidentally contaminated with SV40 and administered to patients between 1955 and 1963, a huge uprising of concern followed. Interestingly, no firm data exists that SV40 can cause cancer in humans with any frequency or severity as it is seen in animals(Poulin & DeCaprio, 2006). However, Merkel cell polyomavirus (MCV) has recently been described and is associated with human skin cancers(Feng et al, 2008). In addition, both the adenovirus family(Huebner et al,

1962; Rabson et al, 1964) and the human papillomavirus family(Huibregste et al, 1993; Scheffner et al, 1993) have been linked to animal and human cancers.

These small, DNA tumor viruses have three interesting things in common: they all have a protein to push the cell in the S phase, they all have a protein that will take care of p53's response and they all integrate their DNA into the host's chromosome (Review(Levine, 2009)). Large T antigen, E1A and E7 (from SV40, human adenovirus, and human papillomavirus, respectively) were all shown to bind and inactivate pRb, with the result being premature entrance into the S phase. Obviously p53 takes offense to this and responds accordingly. In return, large T antigen, E1b-55K and E6 (from SV40, human adenovirus and human papillomavirus, respectively) were all shown to bind and inactive p53(Levine, 2009) (Table 1.2).

Cancer is an unfortunate by-product of these viral infections and not the main goal. Bear in mind that host death is not beneficial to viruses either so, especially in the case of human papillomaviruses, transformation is a slow process, oftentimes taking many years(Scheffner, 1998). Similarly, when p53 is mutated or simply missing from the cell, cancer is the inevitable outcome. These situations only serve to highlight the importance of cell cycle arrest, senescence and apoptosis to every person's life.

Virus	Tumor suppressor gene	
	<i>Rb</i>	<i>p53</i>
SV40	T-antigen	T-antigen
Adenovirus	E1A	E1B-SSKd
Papilloma virus	E7	E6

**Table 1.2: Three DNA tumor viruses and their similar proteins.** The three DNA tumor viruses that are known to cause cancer in both humans and animals are listed. All three viruses have evolved proteins that will interact with pRB to push the cell in to the S phase and a protein to inactive p53. The proteins, along with their names and functions are listed. This figure was adapted from (Levine, 2009).

## 1.4 Overview of dissertation

This dissertation will focus on p53 activity in terms of wild type function and loss of function due to human papillomavirus infection. Chapter 2 will explore how the wild type p53 DNA binding domain binds to its consensus sequence as a tetramer, which is essential for p53 transactivation of downstream genes in cell cycle arrest, senescence and apoptosis. Previous work with the domain has identified why certain mutations render p53 inactive and offer insights into p53 function. The tetramer structure is the next step and represents the first time all four DNA binding domains have been seen bound to a full consensus sequence together and depicts one way the protein binds to DNA *in vivo*. Chapter 3 will focus primarily on human papillomavirus type 16 E6, its interaction with the E6 Associating Protein (E6AP) and how these two proteins bring about degradation of p53 during infection. Specifically, the project sought to identify small molecules capable of disrupting the E6 / E6AP interaction, which would inhibit p53 degradation and bring about senescence or apoptosis in cell lines harboring an integrated human papillomavirus genome.



## **CHAPTER 2**

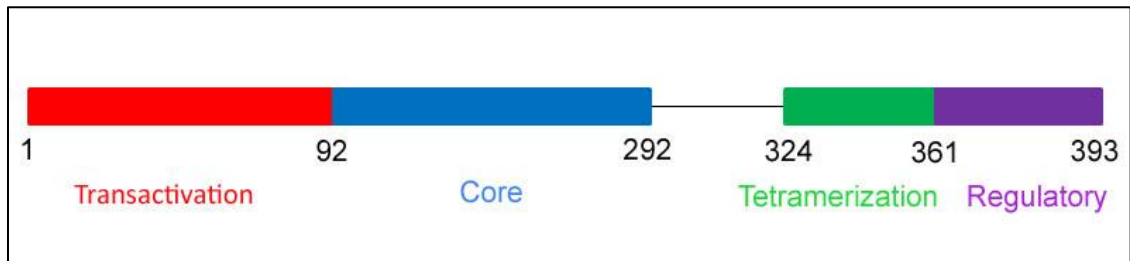
### **Crystal Structure of a p53 core tetramer bound to DNA**

## 2.1. Introduction

### 2.1.1. *p53* polypeptide and its consensus sequence for DNA binding

The full length human p53 protein is 393 amino acids and comprised of four domains: the N terminal transactivation domain, the sequence specific DNA binding (DBD) or core domain, the tetramerization domain, and the C terminal regulatory domain (Figure 2.1). Although dimers have been shown to mediate transactivation of downstream genes in cells (Pietenpol et al, 1994), the functional quaternary structure of p53 *in vivo* is a tetramer.

The N terminal transactivation domain, which interacts with the cellular transcriptional machinery to promote gene transcription, has tricky designation (Viadiu, 2008). The full domain is residues 1 – 92 and can be broken into two subdomains based on amino acid sequence. Residues 1 – 64, which have several aspartic and glutamic acids, are considered the transactivation domain while residues 65 – 92 are the obviously named proline-rich domain. However, the full domain can also be broken into two different subdomains. Residues 1 – 42 are called transcriptional activation domain 1 (AD1) and residues 43 – 92 are transcriptional activation domain 2 (AD2). Interestingly, AD1 appears to promote transcription of cell cycle arrest genes while AD2 promotes transcription of genes involved in apoptosis (Harms & Chen, 2006).



**Figure 2.1: The p53 polypeptide.** The full length human p53 protein is shown schematically with each domain represented as colored blocks.

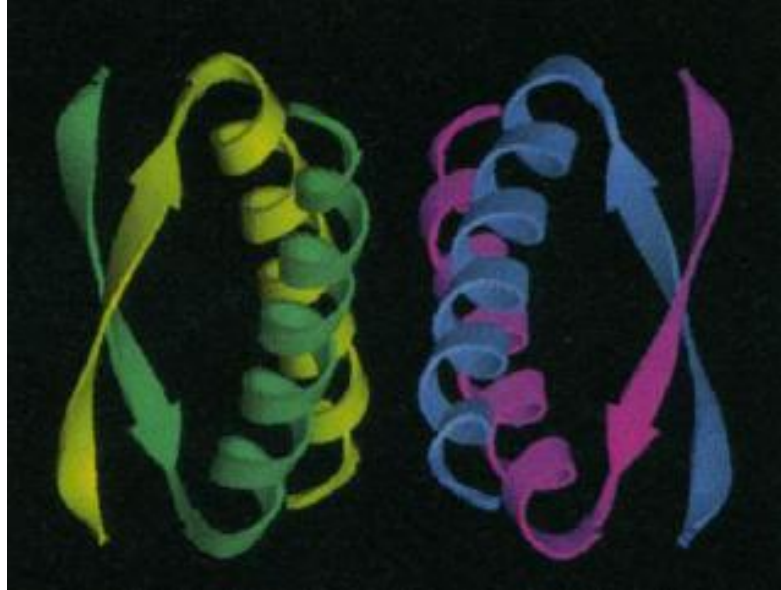
The tetramerization domain runs from amino acids 324 – 361(Jeffrey et al, 1995; Viadiu, 2008) and has a simple structure(Jeffrey et al, 1995). Each domain consists of a  $\beta$  strand (residues 326 – 333) and an  $\alpha$  helix (residues 335 – 354). The crystal structure solved by Jeffrey et al. shows that four tetramerization domains come together as a dimer of dimers. A dimer is formed by two  $\beta$  strands, which come together to form an anti-parallel  $\beta$  sheet, while the two  $\alpha$  helices form an anti-parallel  $\alpha$  helical interface. Two of these dimers come together via the alpha helices to create the tetramer(Jeffrey et al, 1995) (Figure 2.2).

The regulatory domain is composed of many lysines and arginines to create a very basic region of the protein. This is the area where most of the post-translational modifications will occur, such as methylation, phosphorylation, acetylation, and ubiquitination(Viadiu, 2008).

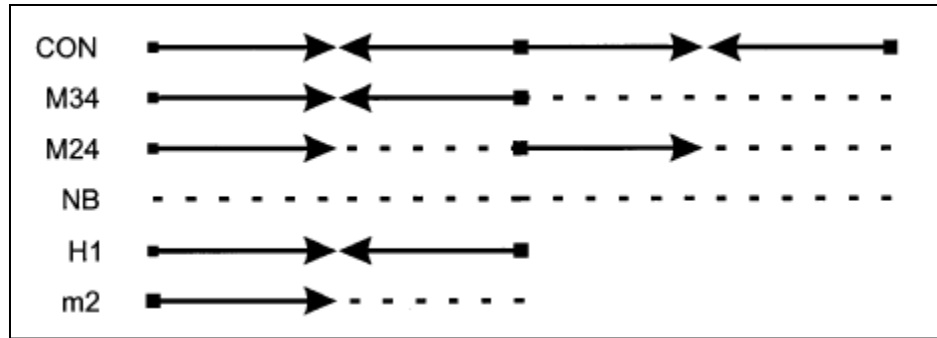
Arguably the most interesting domain of p53 is the DNA binding (DBD), or core domain. This sequence specific binding region of the protein is composed of amino acids 93 – 292 and harbors the vast majority of tumor-derived p53 mutations(Cho et al, 1994; Viadiu, 2008). This region has been shown to bind the WAF1 element with a  $K_d$  of  $83 \pm 1.4$  nM. The domain also binds the DNA cooperatively, with only tetrameric species present after DNA/domain incubation(Balagurumoorthy et al, 1995).

In 1992, Vogelstein and colleagues set out to determine the sequence of DNA to which the DBD typically binds. Using fragmented human genomic DNA, they identified a consensus sequence with “striking internal symmetry”(El-Diery et al, 1992). The full sequence consists of two sets of ten base pairs that follow the motif of 5'-Pu Pu Pu C (A/T) | (T/A) G Py Py Py -3'. The ten base pairs can be placed directly next to each other (icosamers) or separated by up to 13 base pairs(El-Diery et al, 1992).

A follow up work by McLure et al. determined that isolated DBDs would first bind to consecutive, but not alternating, quarter sites (Figure 2.3). This work, as well as Weinberg et al., greatly supported the conclusion that tetrameric p53 binds DNA cooperatively (Balagurumoorthy et al, 1995) by showing that tetramers have at least 50 fold greater binding affinity than dimers alone (McLure & Lee, 1998) and that a tetramer/DNA complex had an approximate  $t_{1/2}$  of 15 minutes while dimer/DNA complexes had an approximate  $t_{1/2}$  of 30 seconds (Weinberg et al, 2004).



**Figure 2.2: Tetramerization domain of p53, 1C26.** Individual tetramerization domains of p53 are shown in yellow, green, blue and purple. The yellow/green and blue/purple pairs represent dimers where the  $\beta$  strands have come together to form an anti-parallel  $\beta$  sheet while the  $\alpha$  helices form an anti-parallel  $\alpha$  helical interface. The  $\alpha$  helices then come together to form the full tetramer, which is a dimer of dimers. This figure is adapted from (Jeffrey et al, 1995).



**Figure 2.3: Pentamers of full p53 consensus sites.** Arrows represent pentamers following the consensus sequence requirements while dashes represent a random DNA sequence. All were tested for DNA binding. CON (positive control), M34 and H1 were all found to bind full length p53 *in vitro*, while M34 and NB (negative control), and m2 showed no appreciable binding. This figure is adapted from (McLure & Lee, 1998).

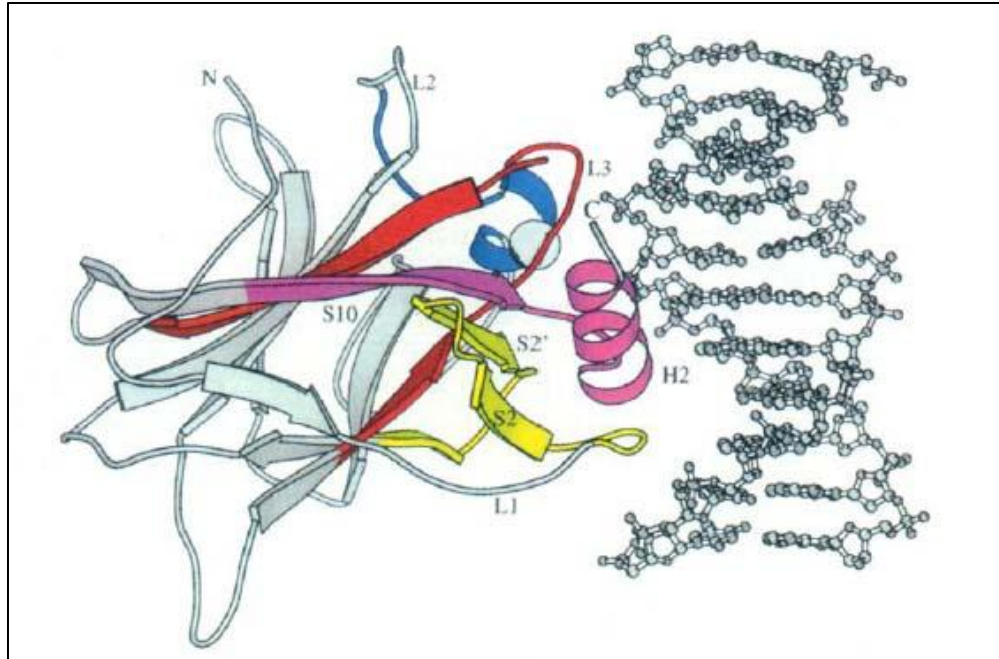
### 2.1.2. *Previously solved p53DBD structures: 1TSR, 2OCJ, 1HU8*

The core domain of p53 has been extensively studied by X-ray crystallography. The first structure, which was published in 1994 by the Pavletich laboratory, consisted of one human p53 core domain bound sequence-specifically to its quarter site(Cho et al, 1994). Following this landmark paper, the structure of both the human(Wang et al, 2007) and mouse(Zhao et al, 2001) p53DBDs in the absence of DNA were also solved.

The first p53DBD structure, which has the protein databank (PDB) code of 1TSR, provided the structural information of the domain bound to DNA, in addition to explaining why several tumor-derived p53 mutations were deleterious to protein function. The central part of the domain consists of a 4 stranded and 5 stranded anti-parallel  $\beta$  sheet that packs together as a sandwich. The strands are highly twisted leading to two very different ends of the sandwich; one end is compact with short loops while the other splays to the loop-sheet-helix motif, which is responsible for DNA binding and contains the tetrahedrally bound  $Zn^{2+}$  ion (Figure 2.4)(Cho et al, 1994).

Careful analysis of this structure highlights cancerous mutations that fell into two categories: 1. domain stabilization (ex. Arg175, Gy252, Arg249, Arg282) and  $Zn^{2+}$  ligation (Cys176, His179, Cys238, and Cys242) , 2. DNA binding (ex. Arg248, Arg273). Domain stabilization mutants comprise 21.7% of tumor-derived p53 mutations, while DNA binding mutants are 18.4%. This work underscores the importance of p53's ability to fold properly and bind DNA to successfully carry out its function(Cho et al, 1994).

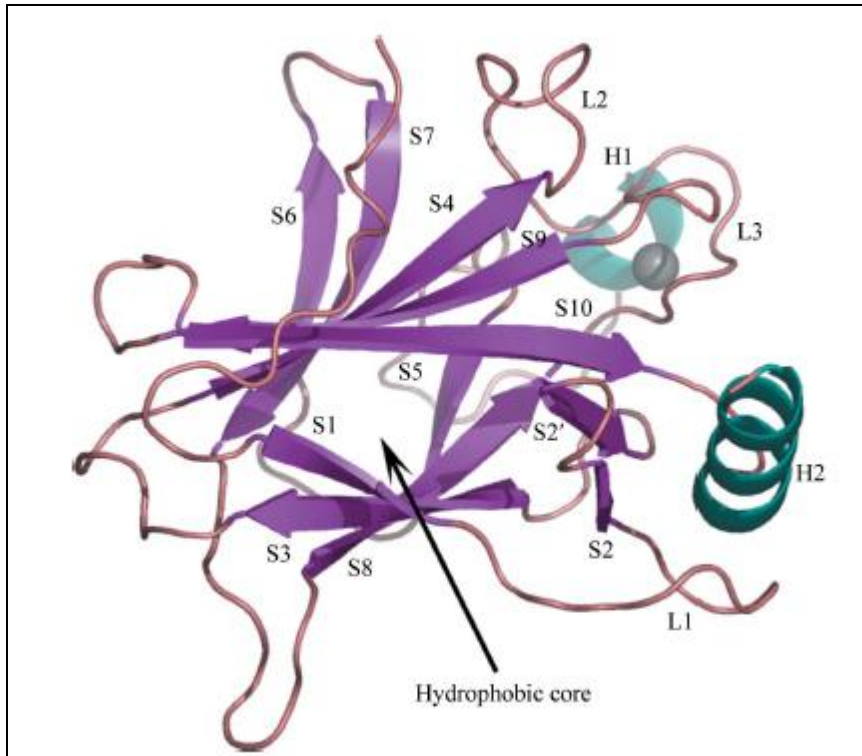




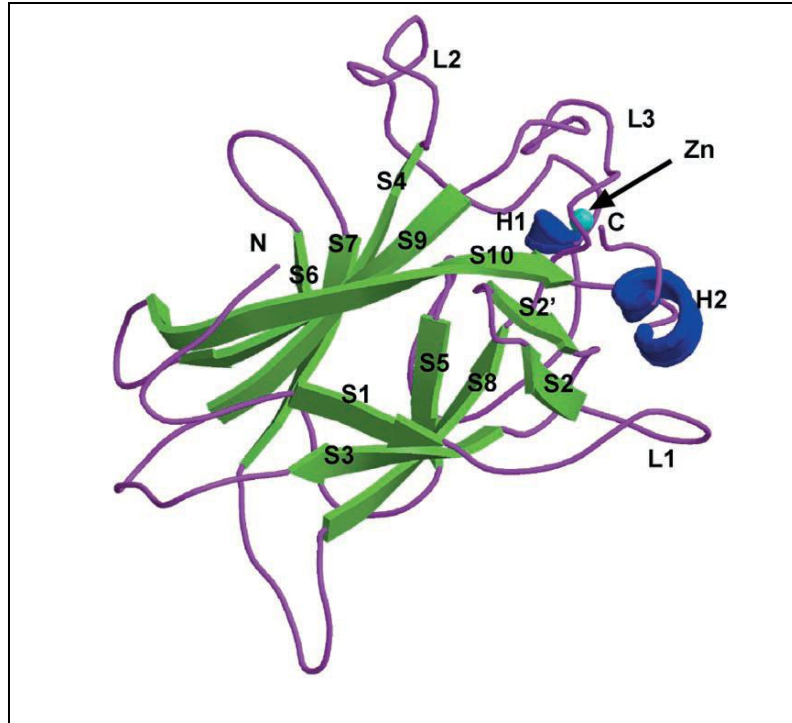
**Figure 2.4: Structure of human p53DBD bound to DNA, 1TSR.** The core domain of human p53 bound to DNA is represented above. The colored areas highlight the loop-sheet-helix motif and highly conserved regions of the domain. Red: S8, L3 loop, S9; Blue: H1 helix; Yellow: S3, L1 loop, S2', and S2; Pink: S10 and H2 helix. This figure is adapted from (Cho et al, 1994).

Seven years later, the structure of the mouse p53DBD in the absence of DNA was solved (1HU8) and, in 2007, the human p53DBD without DNA structure (2OCJ) was solved (Wang et al, 2007; Zhao et al, 2001). The sequence identity between mouse and human p53 core domains is 85% meaning that structurally these domains should be nearly identical (Gasteiger et al, 2003). Both structures show that when not bound to DNA, the domain has largely the same conformation ( $C_{\alpha}$  root mean square deviation (RMSD) of  $\leq 0.87 \text{ \AA}$ ) leading to the conclusion that DNA binding does not lead to a gross structural rearrangement of the domain (Wang et al, 2007; Zhao et al, 2001) (Figure 2.5) (Figure 2.6). One area of the mouse structure that does show a difference is the L1 loop, which contains several amino acids involved in DNA binding and would need to undergo a structural change when DNA was present (Zhao et al, 2001).

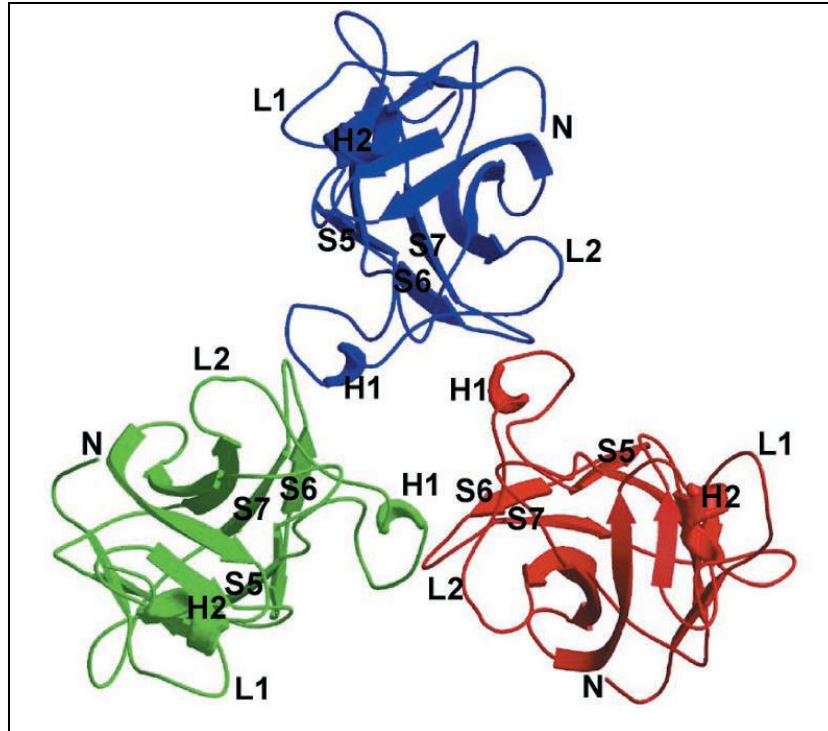
Zhao et al. also report that the domain crystallized as a noncrystallographic trimer, where each subunit-subunit interaction is nearly identical (Figure 2.7). The dimers created in this unit cell were incompatible with DNA binding and the authors postulate that this represents a physiological, but inactive form of p53 (Zhao et al, 2001). It is interesting to note that similar contacts are seen between two p53DBDs in the unit cell of 1TSR, where three DBDs are present but only one is bound sequence specifically (Cho et al, 1994).



**Figure 2.5: Structure of human p53DBD, 2OCJ.** The human p53DBD in the absence of DNA is shown. The  $C_{\alpha}$  RMSD with 1TSR is  $\leq 0.87 \text{ \AA}$ , indicating that the domain does not undergo any significant structural rearrangement upon binding DNA. This figure is adapted from (Wang et al, 2007).



**Figure 2.6: Structure of the mouse p53DBD, 1HU8.** The mouse p53DBD in the absence of DNA is shown. The  $C_{\alpha}$  RMSD with 1TSR is  $\leq 0.87 \text{ \AA}$ , indicating that the domain is structurally the same as human p53DBD and does not undergo any significant structural rearrangement upon DNA binding. This figure is adapted from (Zhao et al, 2001).

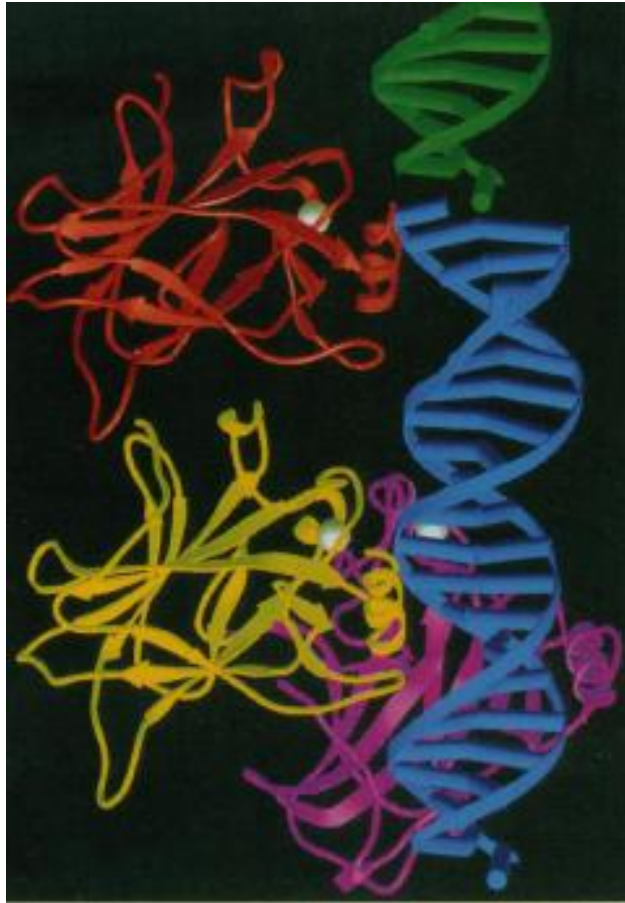


**Figure 2.7: Noncrystallographic trimer of 1HU8.** The asymmetric unit for 1HU8 crystal packing is shown. Each subunit is individually colored as red, blue or green with the key structural aspects labeled. This figure is adapted from (Zhao et al, 2001).

### 2.1.3. *Convertible Nucleoside Approach and Crosslinking*

Twelve years followed the landmark paper of Cho et al. before the first higher order structure of p53 bound to DNA was published (Ho et al, 2006). Efforts to crystallizing these structures face difficulties, which are highlighted well by looking at the unit cell of 1TSR (Cho et al, 1994). The DNA used in that study is a half site, meaning that two p53DBDs could bind sequence specifically. While three p53DBDs are present in the unit cell, only one is bound sequence-specifically (Figure 2.8) (Cho et al, 1994). It is possible that harsh crystallization conditions, strong interactions between the domains and competing core domain dimers that are unable to bind DNA (Zhao et al, 2001) could hinder crystallization. A method was needed to ensure that the domains would remain bound to the DNA properly from purification through data collection.

The solution employed by Ho et al. was the convertible nucleoside strategy and subsequent crosslinking of the p53DBD protein to the DNA (Ho et al, 2006). A quick look at the 1TSR structure shows that Cys277 points in the major groove and makes a base-specific contact with a cytosine in the purine region of the consensus sequence (Cho et al, 1994). If the cytosine is replaced with a similar nucleoside bearing a sulfhydryl, Cys277 is perfectly placed for creating a disulfide bond (crosslink) between the protein and the DNA (Verdine & Norman, 2003).



**Figure 2.8: Asymmetric unit of 1TSR.** The asymmetric unit of 1TSR is shown. Three p53DBDs are present and colored individually (red, yellow and purple). The DNA half site used for crystallization is colored in blue. The yellow domain is bound sequence specifically to its quarter site while the red domain is weakly bound at a nonconsensus site. The purple domain makes no contact with the DNA at all, but does make protein-protein interactions and stabilizes crystal packing. This figure is adapted from (Cho et al, 1994).

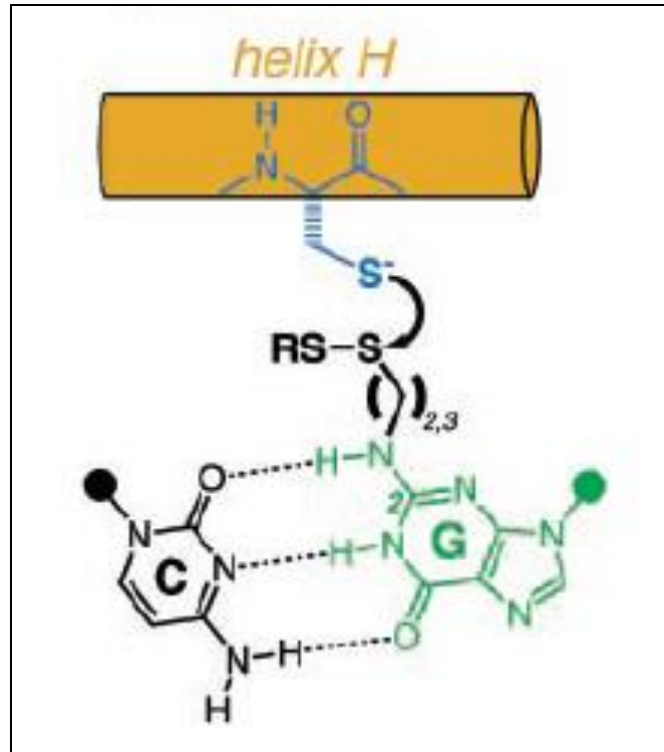
Introducing a crosslink between protein and DNA is especially useful under certain situations. Typically for complex crystallization, the  $K_d$  should be greater than 10 nM, however many physiological interactions are considerably weaker. This problem has left gaps in the structural record, but recent advances with crosslinking have started to fill in the holes(Verdine & Norman, 2003).

Photo-crosslinking is the most widely used technique to covalently link DNA to a protein(Verdine & Norman, 2003). Nucleosides like 5-bromouracil, 5-iodouracil(Willis et al, 1993), or 8-azido-dA(Liu et al, 1994) can be incorporated into DNA then activated to form crosslinks with protein by subjecting them to the proper wavelength of light. Unfortunately, yields are typically low (10 – 20%) and high energy intermediates can lead to biologically irrelevant structures that make this technique essentially useless for structural studies(Verdine & Norman, 2003).

A far better technique for crystallography is incorporating a convertible nucleoside into the DNA, then employing site-directed mutagenesis to introduce a cysteine residue in the protein (Figure 2.9). The technique has some drawbacks, mainly the need for a researcher to have some knowledge of the interaction before beginning. However, if that is known, convertible nucleosides for guanosine, cytosine, adenosine(MacMillan & Verdine, 1991; Xu et al, 1992) are available (Glen Research, TriLink Biotechnologies) and easily added into an growing DNA sequence with no to changes standard protocols and without destroying the hydrogen bonding network in the double stranded DNA(Verdine & Norman, 2003). Typically, yields for this type of crosslinking exceed 80% and purification of crosslinked products is straightforward since the disulfide can be stably formed at pH 8 – 8.5. Also, while disulfides are stable, the bond is also



kinetically labile at pH 8.5 and will resist forming constrained conformers at this pH (Verdine & Norman, 2003). This technique has been successfully used to solve the structure of HIV reverse transcriptase bound to an RNA primer(Huang et al, 1998), the protein GCN4 bound to DNA(Ellenberger et al, 1992), and a p53DBD dimer bound to DNA(Ho et al, 2006). If base crosslinking is not feasible, chemistry has also recently been published demonstrating that crosslinkers can also be added to the backbone of DNA (Banerjee et al, 2006).



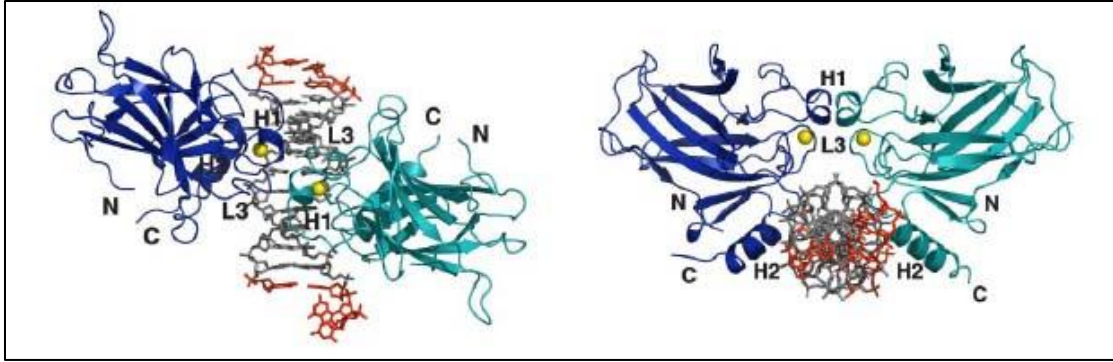
**Figure 2.9: Schematic crosslinking representation.** The protein, represented as the gold cyclinder, bears a cysteine residue that comes close to the DNA. The DNA, represented as a cytosine (black) and modified guanosine (green), has a sulfur-bearing linker on the guanosine. The two should come reasonably close together when bound correctly and allow a disulfide to form between the linker on the guanosine and the cysteine from the protein. This figure is adapted from (Verdine & Norman, 2003).

#### 2.1.4. *Previously solved p53DBD structures: 2GEQ, 2AC0, 2ADY, 2AHI, 2ATA*

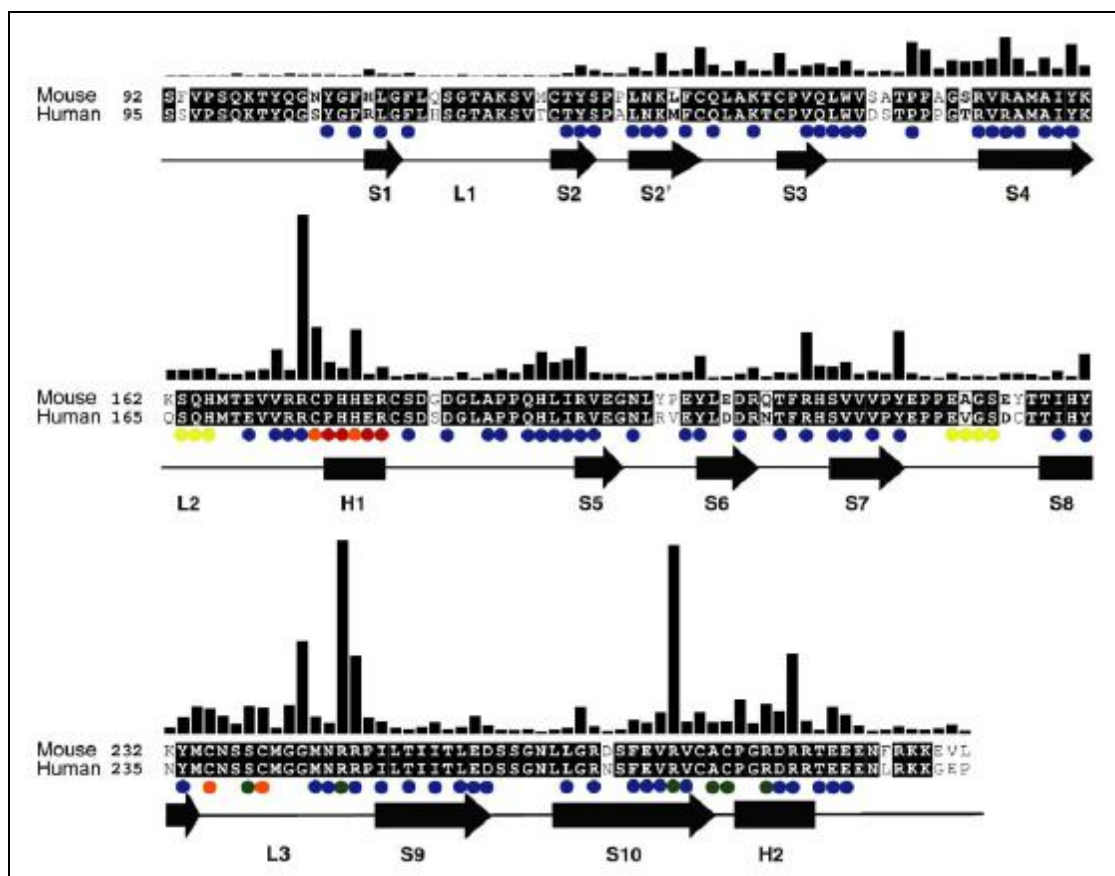
In their 2006 paper, Ho et al. reported the structure of mouse p53DBD dimer bound to a DNA half site after creating a crosslink between the each p53DBD and its quarter site. The structure, in addition to highlighting the residues necessary for p53DBD dimerization, also identified a third hotspot for cancerous mutations within domain and created a model tetrameric p53DBD bound to a full consensus site (Ho et al, 2006).

Each p53DBD in the structure (PDB code 2GEQ) binds to its quarter site in a nearly symmetrical fashion across the DNA (Figure 2.10). C<sub>α</sub> RMSDs between each subunit of the dimer and the original 1TSR p53DBD is  $\leq 0.748 \text{ \AA}$ , implying high structural similarity between all the domains. The DNA contacts of each subunit with its pentamer sequence are also essentially the same as reported by Cho et al. In this structure, the DNA half site also has a 20° bend, which is not seen in 1TSR (Ho et al, 2006).

Dimerization occurs via the H1 helix and L3 loop. Several conserved residues make important contacts with opposing subunits to create the dimerization interface. Interestingly, mutations in these residues are found in tumor-derived p53 genes, underscoring the importance of dimerization to p53 function. Along with domain stabilization, DNA contacts, and Zn<sup>2+</sup> ligation, dimerization residues now represent a third hotspot for p53 mutations that lead to inactive proteins that are associated with human cancers (Figure 2.11) (Ho et al, 2006).



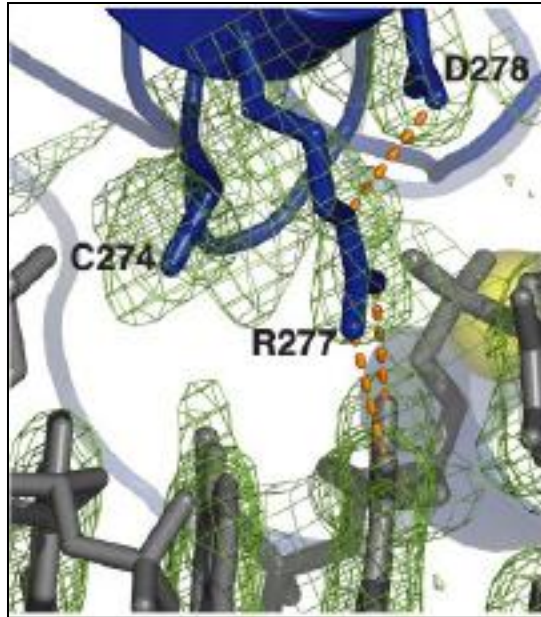
**Figure 2.10: Structure of mouse p53DBD dimer bound to DNA, 2GEQ.** Two representations of the mouse p53DBD dimer are shown. One subunit is represented in blue while the other is represented in green. The consensus sequence half site is shown in grey while flanking bases necessary for crystallization are in red. The  $Zn^{2+}$  ions are represented as yellow spheres. The right-hand view is a  $90^\circ$  turn into the page from the left-hand view.



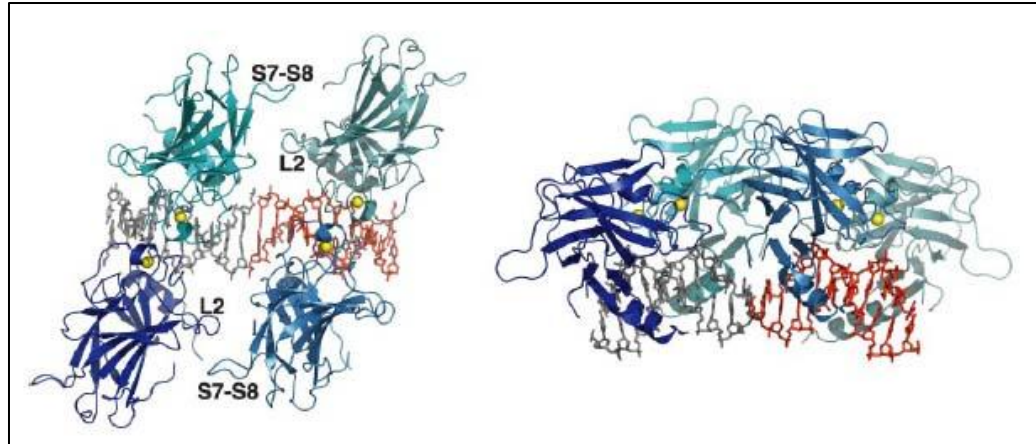
**Figure 2.11: p53DBD cancer-derived mutation histogram.** The core domain amino acid sequences and known secondary structure for both mouse and human proteins are shown. The black bar above each amino acid represents how often the particular residue is found mutated in human cancers. The colored circles below certain residues represent the amino acid's importance: blue = domain stability, orange = Zn<sup>2+</sup> ligation, green = DNA contact, red = dimerization, yellow = residues proposed to be involved in tetramerization. This figure is adapted from (Ho et al, 2006).

Logically, the question of whether the crosslink disrupted natural binding of each p53DBD subunit to the DNA must be addressed. As stated above, the overall structure of each domain and its contacts to the DNA remain unchanged as compared with previous uncrosslinked structures. Also, a  $2F_o - F_c$  map contoured to  $1.25 \sigma$  shows discontinuous density between the Cys274 (Cys277 in human numbering) and the crosslinked cytosine (Figure 2.12). The distance between the sulfur of the Cys274 and N<sub>4</sub> of the cytosine is approximately 4 Å while the full linker is 7.5 Å (Huang et al, 1998). These observations lead to the conclusion that the linker remains flexible, in a disordered conformation and not significantly perturbing the natural binding of the domains to the DNA (Ho et al, 2006).

The final part of Ho et al.'s assessment of their structure was to model a p53 tetramer bound to its full consensus site (Figure 2.13). Dimer-dimer interactions were predicted to involve the S7-S8 loop of one subunit with the L2 loop of another (Ho et al, 2006). In addition, the DNA was also predicted to bend by as much as 40° (Ho et al, 2006), which is a conclusion that is also supported by several solution studies (Nagaich et al, 1997a; Nagaich et al, 1999; Pan & Nussinov, 2007).



**Figure 2.12: Electron density map in the area of the crosslink, 2GEQ.** A close up view of the area surrounding Cys274 (Cys277 in human numbering) is shown. The electron density map is represented as green mesh and contoured to  $1.25 \sigma$ . Connected density between the sulfur of Cys274 and N<sub>4</sub> of the cytosine is not seen. This figure is adapted from (Ho et al, 2006).



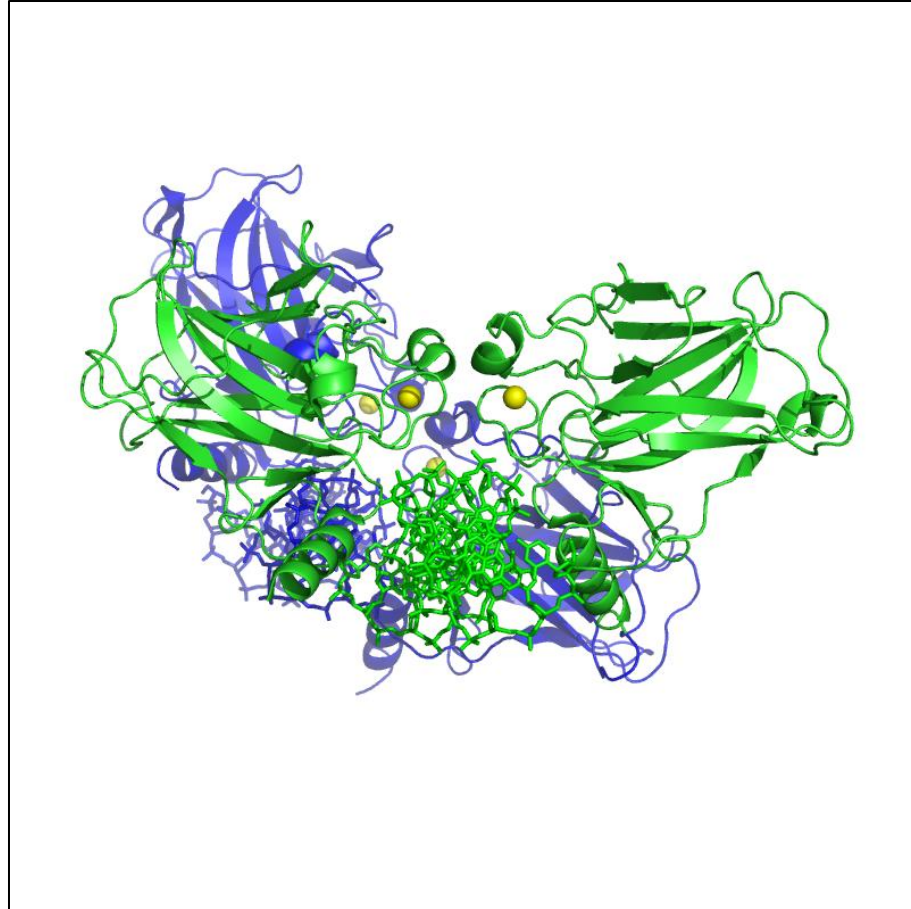
**Figure 2.13: Proposed p53DBD tetramer bound to DNA model.** The structure of the mouse p53DBD dimer bound to DNA (2GEQ) is shown in the left panel on the left-hand side of the tetramer with gray DNA. Following the bend in the DNA, the next dimer, with the exact same structure, was modeled next to it to extrapolate what the full tetramer would look like. The right-hand picture is obtained by rotating the left-hand picture 90° into the page, then a 90° rotation to the right. This figure is adapted from (Ho et al, 2006).



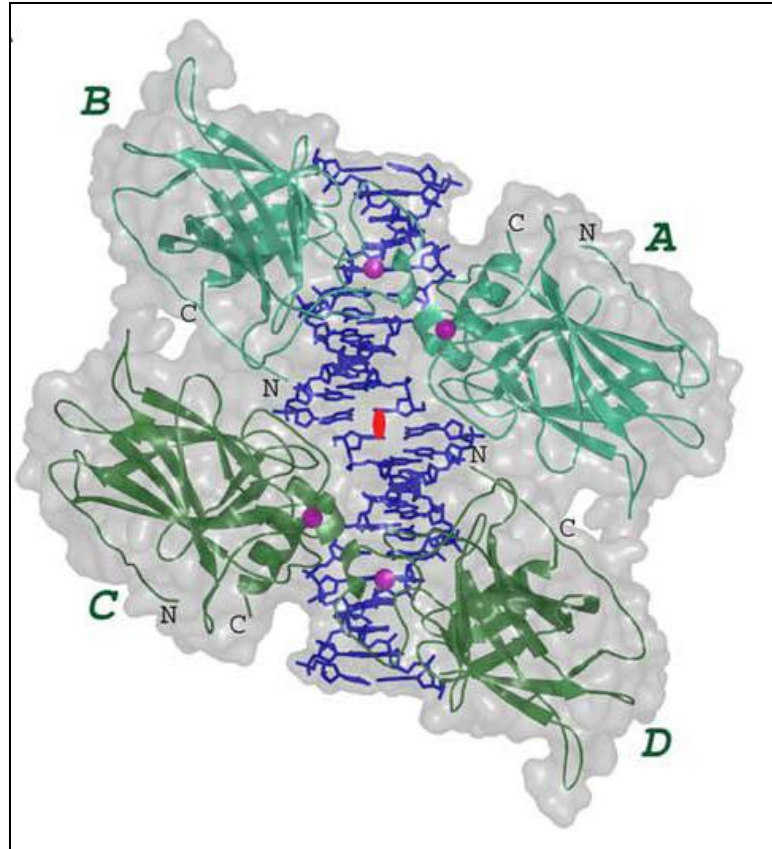
Shortly after the publication of the kinetically trapped dimer structure, Shakked and colleagues published four structures (2AC0, 2ATA, 2AHI, 2ADY) of uncrosslinked human p53DBD dimers bound to DNA that pack together in their crystal lattices as tetramers (Kitayner et al, 2006). Again, the subunits in each structure have high structural similarity to the previously solved domains and the DNA contacts were essentially as previously reported. The dimerization also occurs largely as reported by Ho et al. with a few minor differences (Kitayner et al, 2006).

Each tetramer reported its own unique dimer-dimer interfaces (Figure 2.14). Interestingly, Ser166, Glu198, Gln167 (Ser163, Glu195, and Gln164 in mouse numbering) are involved in dimer-dimer contacts in each tetramer, but mediate different contacts (Kitayner et al, 2006).

Examination of the DNA duplexes shows that they are slightly out of register with each other in the Kitayner et al. tetramers (Kitayner et al, 2006) (Figure 2.15), meaning that a p53DBD tetramer bound to a full consensus site still had not been solved and remained an open area of interest to the p53 community.



**Figure 2.14: Structure of human p53DBD tetramer bound to DNA, 2AC0.** The human p53DBD tetramer is represented as two colored dimers. The front dimer and DNA are colored in green while the back dimer and DNA are colored in blue. Zn<sup>2+</sup> ions are represented as yellow spheres. The view is looking down the DNA axis of the green dimer. The rotation between the green dimer and the blue dimer is  $\sim 33^\circ$ . Four similar structures were published (2AC0, 2ADY, 2ATA, and 2AHI).



**Figure 2.15: DNA of 2AC0.** The tetramer from Figure 2.14 is represented after a 90° rotation out of the page. The A-B dimer is shown in light green while the C-D dimer is shown in dark green. The half sites bound by each dimer are shown in blue. The DNA axis for the A-B dimer is out of register from the C-D dimer. Another picture of this can be seen in [Figure 2.35](#), as well. This figure is adapted from (Kitayner et al, 2006).

## 2.2. Materials and Methods

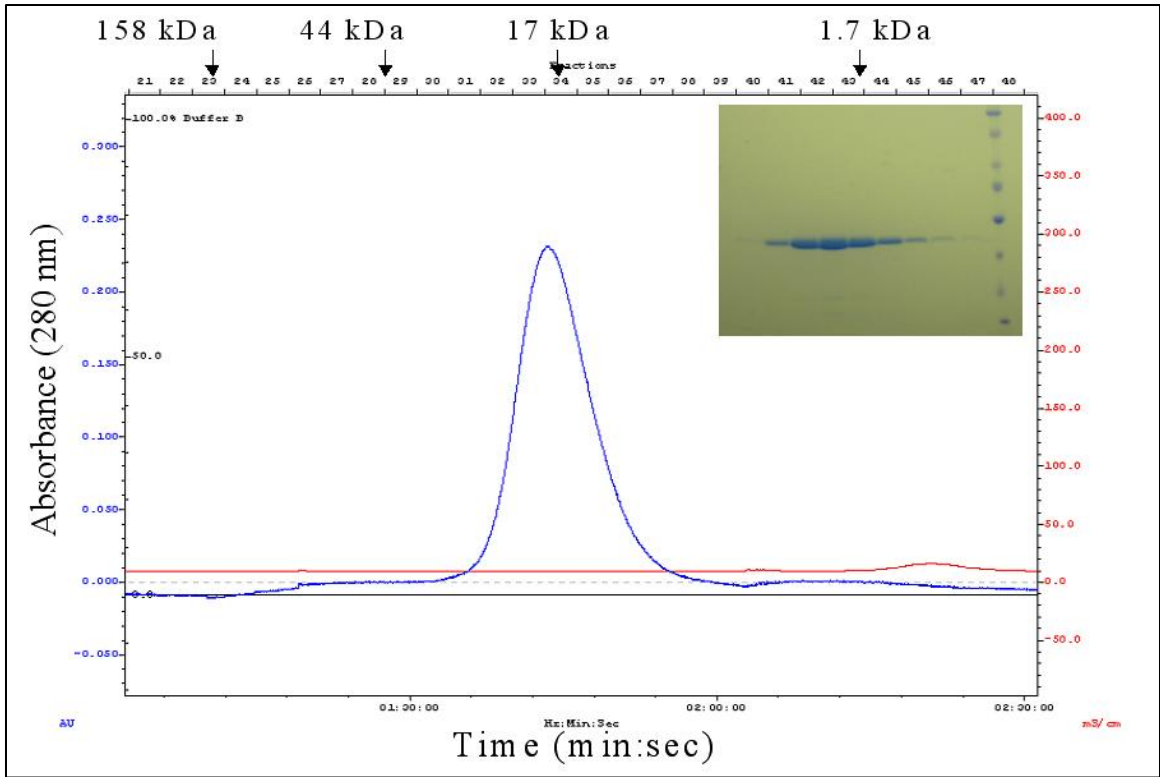
### 2.2.1. Growth and purification of mouse p53DBD

Residues 98 – 292 of mouse p53 were cloned into the pRSETa vector (Invitrogen) such that it had no tag. DE3 *Escherichia coli* (*E. coli*) cells were transformed with the plasmid, grown in the presence of 100 µg/mL ampicillin, and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 10 µM Zn(OAc)<sub>2</sub> overnight at 15°C.

Cell pellets were resuspended in 20 mM citrate pH 6.1, 100 mM NaCl, 10 µM Zn(OAc)<sub>2</sub>, 1 mM tris(2-carboxyethyl) phosphine hydrochloride (TCEP), and 10 µg/mL phenylmethanesulfonyl fluoride (PMSF). Cells were lysed by sonication at 4 °C and the insoluble fraction was separated by centrifugation at 18,000 rpm (SS-34 rotor, Sorvall) for 30 minutes. The soluble fraction was passed over pre-equilibrated SP sepharose resin (GE Healthcare), washed and eluted on a NaCl gradient of 100 mM – 1 M.

The cleanest fractions were pooled, concentrated and run over a Superdex 75 preparative gel filtration column (Amersham Biosciences), where p53DBD eluted as a monomer (Figure 2.16). Peak fractions from the column were pooled and concentrated to at least 5 mg/mL before flash freezing on dry ice and long term storage at - 80°C.

The p53DBD point mutant T281C was prepared by standard mutagenesis and purified in the same manner.



**Figure 2.16: Gel filtration of p53DBD.** Representative superdex 75 gel filtration chromatogram for p53DBD purification is shown. The protein peaks just before the 17 kDa standard, which fits for a monomeric, 22 kDa protein. The 15% SDS-PAGE gel insert shows that the six fractions of the peak are pure p53DBD.

## 2.2.2. *Synthesis and purification of DNA containing unnatural nucleotides*

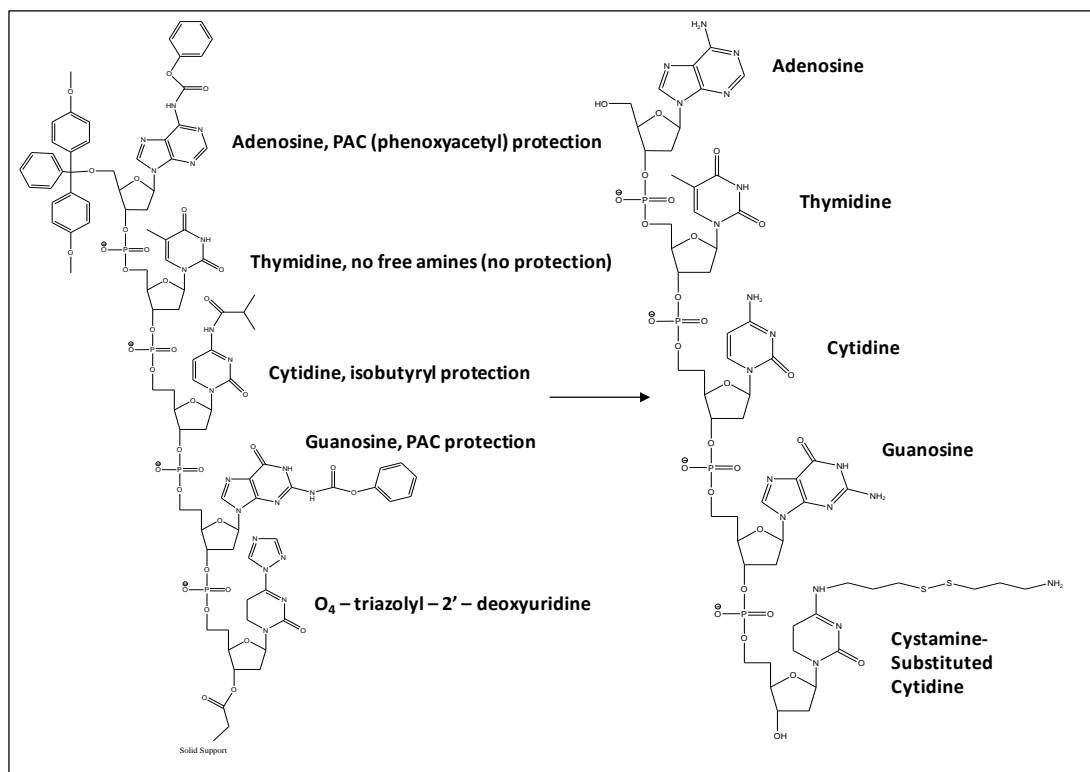
### 2.2.2.1. *Base crosslinking DNA strands*

DNA oligonucleotides containing the unnatural  $O_4$ -triazolyl-2'-deoxyuridine base were ordered from TriLink Biotechnologies (San Diego, CA) or synthesized on the Applied Biosystems DNA Synthesizer 394. DNA strands, capped with a 5'-dimethoxy-trityl group (trityl), were shipped on the resin with standard protecting groups on all bases following the synthesis.

Resin was incubated with 1 mL of 0.5 M cystamine (Sigma), pH 12 and placed on a shaker for two hours to cleave DNA strands from the resin. Samples were spun at maximum speed for 10 minutes to pellet the resin and the supernatant was removed. Cleaved DNA strands were placed in a 55°C heat block overnight for complete deprotection of all nucleotides and cystamine replacement of the triazolyl group.

Following dialysis with Spectrapor 2000 MWCO tubing (Spectrum Laboratories) into water and lyophilization, DNA strands were resuspended in 500  $\mu$ L deionized water. Further purification was performed by reverse-phase HPLC to separate the failure sequences via the 5'-trityl, which was subsequently removed from correct sequences after treatment with trifluoroacetic acid (TFA).

Strands were dialyzed into water once more, lyophilized and dissolved in  $\sim$  40  $\mu$ L of water. DNA concentrations were determined by UV absorbance and strand composition was confirmed by matrix assisted laser desorption ionization (MALDI) time of flight (TOF) mass spectrometry (Figure 2.17).



**Figure 2.17: Scheme for creating DNA capable of base crosslinking.** The left-hand side shows DNA as synthesized, complete with all protecting groups. The right-hand side shows the DNA when fully deprotected, purified, and ready for use.

#### 2.2.2.2. *Backbone crosslinking DNA strands*

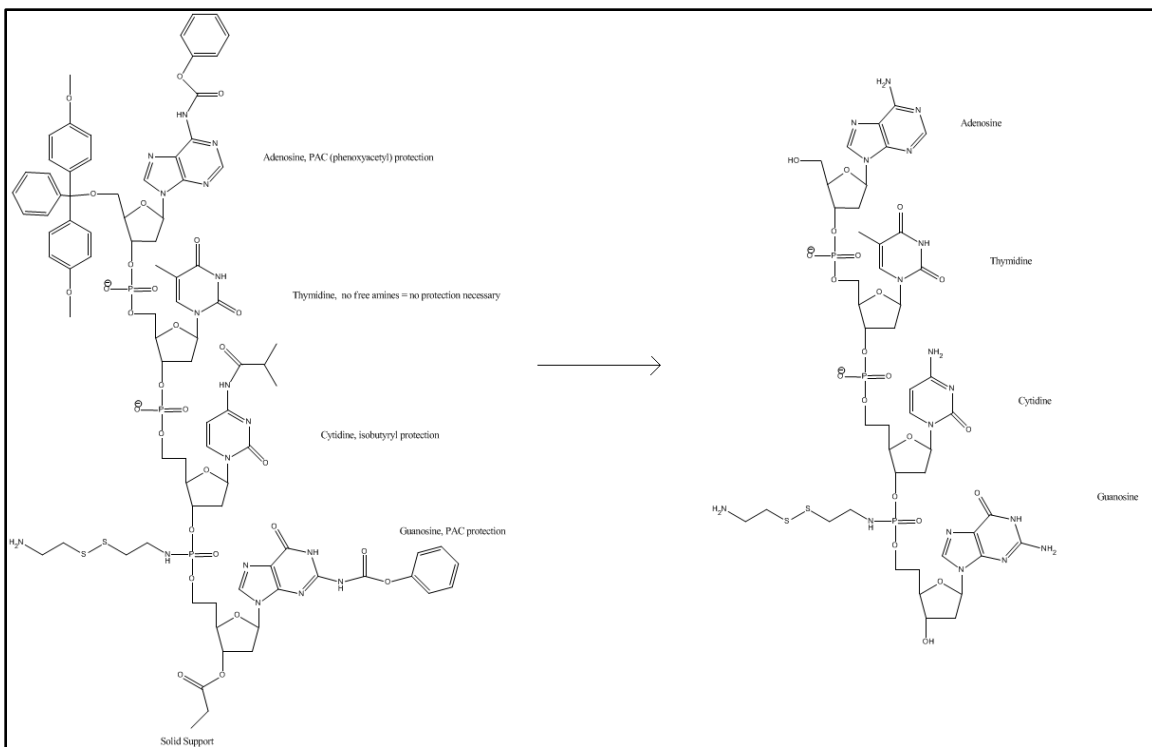
Cystamine free base was prepared from cystamine-HCl (Sigma) by neutralization with sodium hydroxide and subsequent extraction with dichloromethane.

DNA oligonucleotides were alternately synthesized on the Applied Biosystems DNA Synthesizer 394 and by hand. All reagents and phosphoramidites were ordered from Glen Research (Sterling, VA) unless otherwise stated. Backbone modification occurred between two nucleotides on the growing strand so automated synthesis was stopped at the base 3' to the phosphate group modification after trityl removal. Columns were removed from the synthesizer, flushed with 50:50 acetonitrile:pyridine, and incubated with H-phosphonate plus 1-adamante carbonylchloride (activator) for one hour. Columns were then washed, and DNA backbones were oxidized with cystamine free base dissolved in carbon tetrachloride (Sigma) for one hour. Failure sequences were then capped with isopropyl phosphate and columns were washed with 50:50 acetonitrile:pyridine followed by 100% acetonitrile (Banerjee et al, 2006). Columns were placed back on the synthesizer, where the remaining nucleotides were added to the growing strands with the standard Applied Biosystems 1  $\mu$ M DNA synthesis protocol (Figure 2.18).

To make DNA with only backbone crosslinking: Following synthesis, DNA strands were cleaved from the resin and deprotected in ammonium hydroxide before HPLC purification.

To make DNA with mixed crosslinking: Following synthesis, purification of DNA was carried out the same as base crosslinking strands.

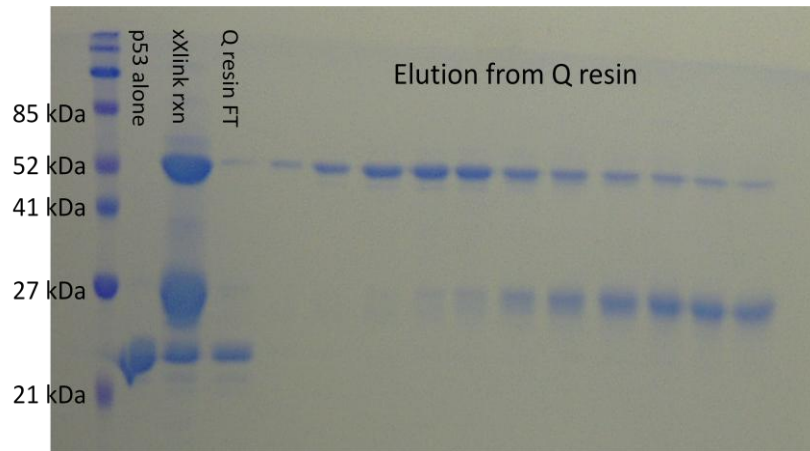




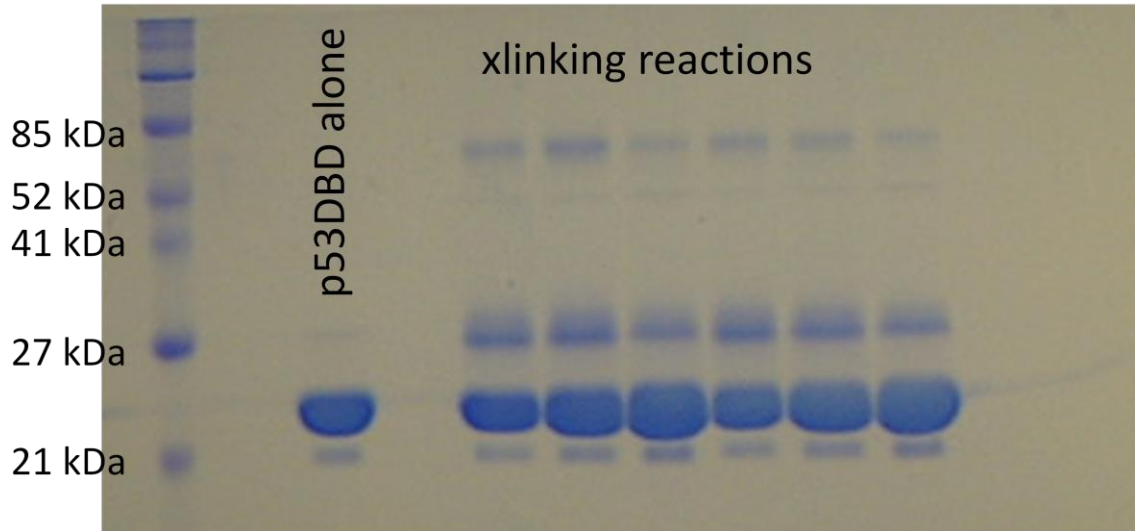
**Figure 2.18: Scheme for creating DNA capable of backbone crosslinking.** The left-hand side shows DNA as was synthesized, complete with all protecting groups. The right-hand side shows the DNA when fully deprotected, purified, and ready for use.

### 2.2.3. *Purification of crosslinked tetramers*

Crosslinking was performed the same for all types of modified strands: base crosslinking with p53DBD, backbone crosslinking with p53DBD T281C, mixed crosslinking with p53DBD T281C (Figure 2.19, Figure 2.20). At room temperature in 20 mM citrate, pH 6.1, 100 mM NaCl, and 0.5 mM TCEP, a 4:1 molar ratio of p53DBD to DNA was added to excess buffer and allowed to sit for three hours at room temperature. The crosslinking reaction was loaded onto Q resin (GE Healthcare) and higher order oligomers were eluted with a gradient of 100 mM to 500 mM NaCl. Tetrameric complexes eluted first since its negative charges were the least available for binding with the resin. Denaturing but non-reducing gels were run to assay purity. Tetramer purity was determined by the presence of only one band corresponding to one strand of DNA with two bound p53DBDs. These fractions were pooled and concentrated to 5  $\mu$ M as determined by a standard curve of known protein-DNA concentrations and frozen at  $-80^{\circ}\text{C}$  until use.



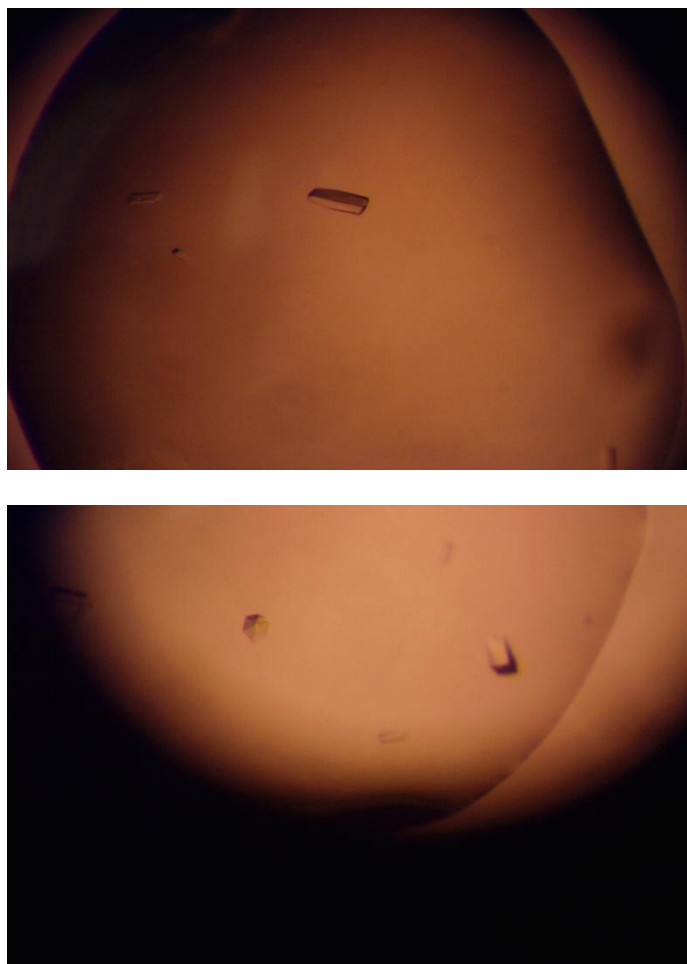
**Figure 2.19: Purification of p53DBD tetramer by base crosslinking.** Following incubation of a 4:1 molar ratio p53DBD:DNA, samples were run over pre-equilibrated Q resin and eluted on a NaCl gradient of 100 – 500 mM. The first three lanes of this 15% denaturing, but non-reducing gel show p53DBD alone, crosslinking reaction before loading onto Q resin, and flow through from the Q resin. The crosslinking reaction clearly shows three bands: lowest band is p53DBD alone (22 kDa), middle band is one p53DBD + one DNA strand (22 kDa + 7 kDa = 29 kDa), the highest band is two p53DBD + one DNA strand (7 kDa + 44 kDa = 51 kDa). The next several lanes all represent fractions of the elution. Fractions that only showed upper band were concluded to contain DNA crosslinked to four p53DBD subunits. These fractions were pooled and concentrated for crystallization.



**Figure 2.20: Samples of p53 tetramer by backbone crosslinking.** Following incubation of a 4:1 molar ratio p53DBD T281C:DNA, samples were run over pre-equilibrated Q resin and eluted on an NaCl gradient of 100 – 500 mM. The first lane of this 15% denaturing, but non-reducing gel shows p53DBD T281C alone (22 kDa). The next several lanes show crosslinking reactions under various buffer conditions and pH with the same molecular weights as described in [Figure 2.19](#). The crosslinking reactions clearly show three bands: lowest band is p53DBD T281C alone (22 kDa), middle band is one p53DBD T281C + one DNA strand (22 kDa + 7 kDa = 29 kDa), the highest band is two p53DBD T281C + one DNA strand (7 kDa + 44 kDa = 51 kDa).

#### 2.2.4. *Crystallization, data collection and refinement of base crosslinked tetramers*

Crystals of base crosslinking tetramers were grown in 200 mM lithium citrate, 20% PEG 3350 by hanging drop vapor diffusion. Small, rectangular crystals appeared in two days and reached a maximum size of 160 x 30 x 30  $\mu\text{m}$  after one week ([Figure 2.21](#)). Crystals were frozen in the well condition with 15 - 20% glycerol. Data collection was carried out on the 23-IDB beamline at the Advanced Photon Source synchrotron at Argonne National Laboratories. Data was indexed and scaled using HKL2000 (Otwinowski & Minor, 1997). During indexing, it was found that tetramers packed one of two ways, which yielded two different unit cells and subsequently two tetramer structures. The structures were solved by molecular replacement using PHASER with 2GEQ as a search model. For structure 1, the resulting solution was the p53DBD dimer bound to its half site in an asymmetric unit cell with a full tetramer containing a crystallographic two-fold axis. For structure 2, the resulting solution was the p53DBD monomer bound to its quarter site in an asymmetric unit. These models were refined in CNS using simulated annealing, minimization and individual B-factor refinement. For structure 1, medium NCS restraints were initially used for the two p53DBD domains in the asymmetric unit. The NCS restraints were gradually lowered and were eventually dropped completely. Between refinement cycles, the model was manually rebuilt using the program Coot (Emsley & Cowtan, 2004). Data collection and refinement statistics for structures 1 and 2 are summarized in [Table 2.1](#).



**Figure 2.21: Crystals of p53DBD tetramers with base crosslinking.** These crystals were grown in 200 mM lithium citrate, 20% PEG 3350 by hanging drop vapor diffusion and were harvested, frozen and collected upon as described in the text.

Table 1 Crystallographic statistics		
Structure	1	2
Space group	C2	C2
<i>Cell dimensions</i>		
a, b and c (Å)	114.75, 68.016, 75.162	109.41, 68.10, 34.42
α, β and γ (deg)	90.00, 111.12 and 90.00	90.00, 104.17 and 90.00
Resolution (Å)	2.00	2.20
R <sub>sym</sub> <sup>a</sup>	0.083 (0.465)	0.100 (0.277)
<I>/σ	16.5 (3.6)	14.7 (5.6)
Completeness	99.1% (99.4%)	99.6% (97.3%)
Redundancy	4.9 (4.9)	4.9 (4.4)
<i>Refinement</i>		
Resolution range (Å)	70.186–2.00	57.354–2.20
Reflections	177 707	60 952
R <sub>work</sub> <sup>b</sup> /R <sub>free</sub> <sup>c</sup>	0.225/0.260	0.205/0.250
Asymmetric unit	Dimer + half site	Monomer + quarter site
<i>Number of atoms</i>		
Protein	3034	1510
DNA	464	204
Zinc	2	1
Water	577	161
Average B factor (Å <sup>2</sup> )	37.67	33.91
<i>RMSDs</i>		
Bond lengths (Å)	0.00547	0.0108
Bond angles (deg)	1.37	1.54

**Table 1. Crystallographic statistics for base crosslinked tetramers.** Tetramers packed in two different ways to form crystals. These differences in packing will be discussed in Section 2.3.1. Crystallographic statistics are provided for both crystal packing forms.

<sup>a</sup>:  $R_{\text{sym}} = \frac{\sum_h \sum_l |I_{hj} - I_h|}{\sum_h \sum_{h,i} I_{hi}}$  where  $I_h$  is the mean intensity of symmetry-related reflections.

<sup>b</sup>:  $R_{\text{work}} = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}$ ,  $F_o$  is observed structure factor amplitudes;  $F_c$  is calculated structure factor amplitudes.

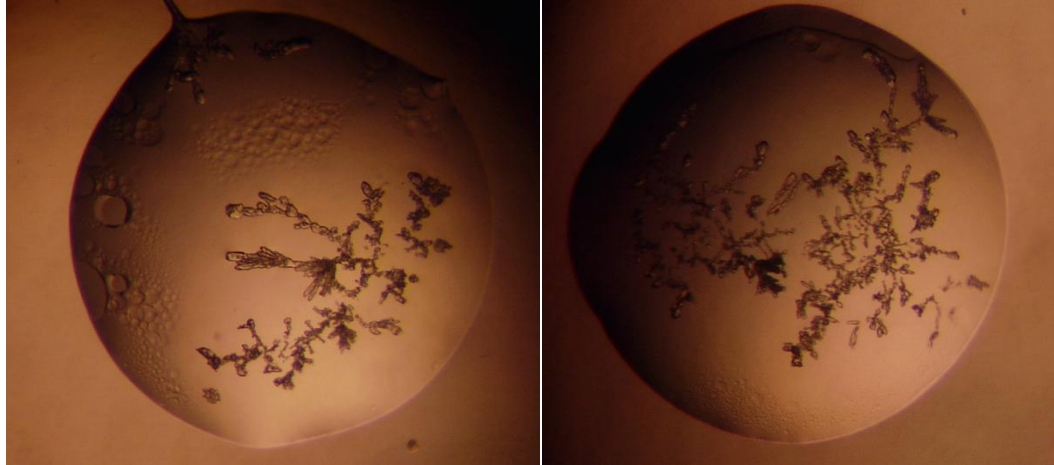
<sup>c</sup>: R<sub>free</sub> is calculated from the withheld 5% of data.

This table is adapted from (Malecka et al, 2009).

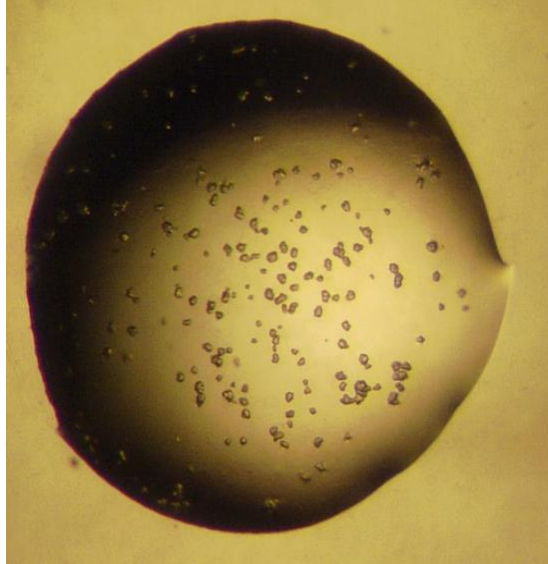
### 2.2.5. *Crystallization of backbone crosslinked and mixed crosslinking tetramers*

Crystals of backbone crosslinked tetramers (all four subunits crosslinked via the DNA backbone) were grown in 200 mM potassium sulfate, 20% PEG 3350 (Figure 2.22). Crystals of tetramers with mixed crosslinking (two subunits crosslinked via Cys274, two subunits via the DNA backbone) were grown in 200 mM zinc acetate dehydrate, 20% PEG 3350 (Figure 2.23) by hanging drop vapor diffusion. The former crystals appeared after one night while the shower of small crystals ( $\leq 30 \mu\text{M}^3$ ) in the latter took at least one month to grow. The latter crystals were frozen in 200 mM zinc acetate, 20% PEG 3350, 20 – 25% glycerol, but yielded no diffraction.





**Figure 2.22: Crystals of backbone crosslinked tetramers.** These crystals were prepared with all four individual subunits crosslinked to the backbone of the consensus sequence DNA and grown in 200 mM potassium sulfate, 20% PEG 3350 by hanging drop vapor diffusion. They appeared overnight. See [Figure 2.40](#) for the DNA sequence.



**Figure 2.23: Crystals of mixed crosslinked tetramers.** These crystals were prepared by mixing base crosslinking and backbone crosslinking. Two subunits are crosslinked via Cys274 while two subunits are crosslinked to the backbone of the DNA. They were grown in 200 mM zinc acetate dehydrate, 20% PEG 3350 by hanging drop vapor diffusion and took at least 1 month to grow.

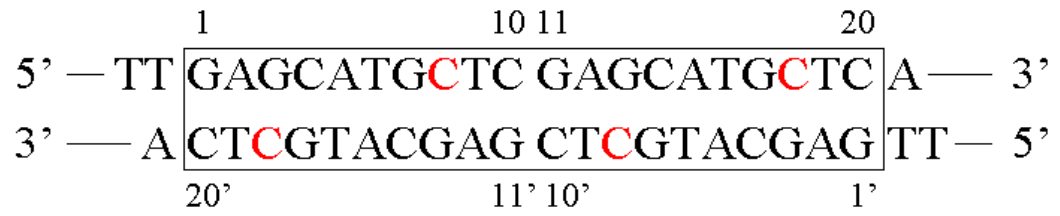
See [Figure 2.41](#) for DNA sequence.

## 2.3. Results and Discussion

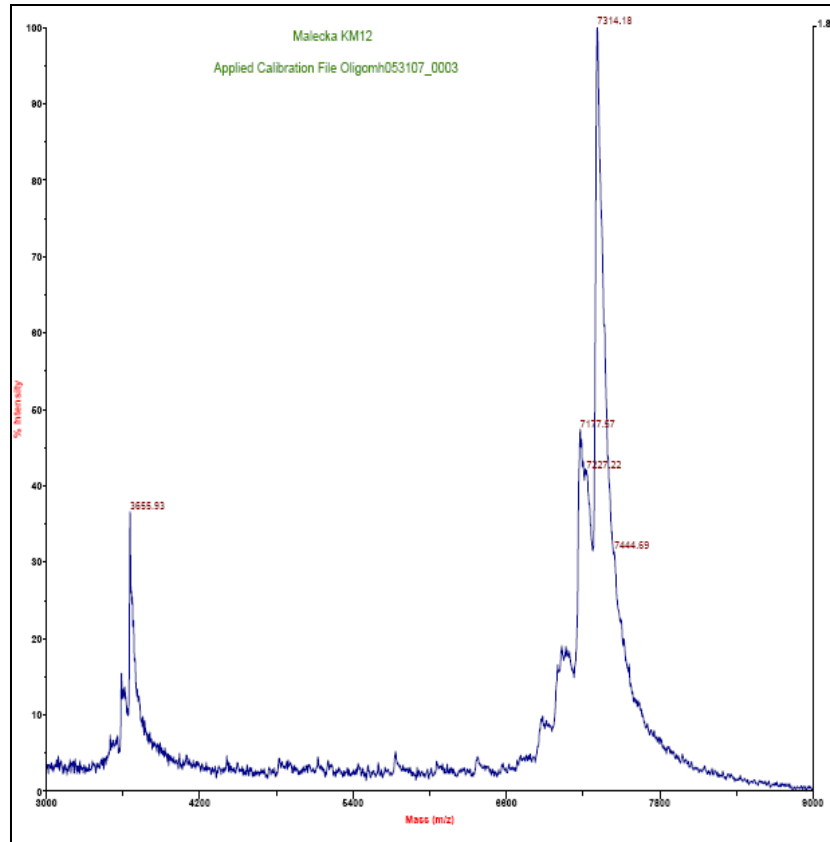
### 2.3.1. *Crystal Structure of p53DBD tetramer bound to DNA*

Two tetrameric p53DBD/DNA complex structures, referred to as structure 1 (PDB code 3EXJ) and structure 2 (PDB code 3EXL), were determined to resolutions of 2.00 and 2.20 Å respectively (Table 2.1). Both structures contain the same DNA duplex (Figure 2.24, Figure 2.25) and the same crosslinks, but structure 1 shows ordered density for the entire 23 basepairs while only the central 20 base pair consensus sequence can be modeled in structure 2. Both protein/DNA complexes crystallized in spacegroup C2, but the cell parameters and asymmetric units of these complexes differ, with structure 1 containing a p53DBD dimer bound to a DNA half site and structure 2 containing a p53DBD subunit bound to a DNA quarter site. As a result of these differences, the crystal packing contacts show significant differences between the two structures. The most significant difference lies along the direction of the DNA helical axis. The DNA in structure 1 forms end-end Hoogsteen base-pairs involving the overhanging 5' thymine base of one DNA duplex with the Watson-Crick base-paired adenine of another duplex (Figure 2.26). In contrast, the DNA duplex in structure 2 does not show electron density for the overhanging 5' thymine base or the adjacent adenine-thymine base pair, which are presumably looped out of the DNA helix. As a result of this, two adjacent DNA duplexes of structure 2 stack end-to-end forming a pseudo continuous helix with the terminal DNA base-pairs from adjacent DNA duplexes about 3.6 Å apart (Figure 2.26). The shorter ordered DNA duplex within structure 2 relative to structure 1 results in more extensive protein contacts between adjacent p53DBD tetramer/DNA complexes within the crystal lattice of structure 2 relative to structure 1 (Figure 2.27). Despite these crystal packing differences between structure 1 and 2, the overall

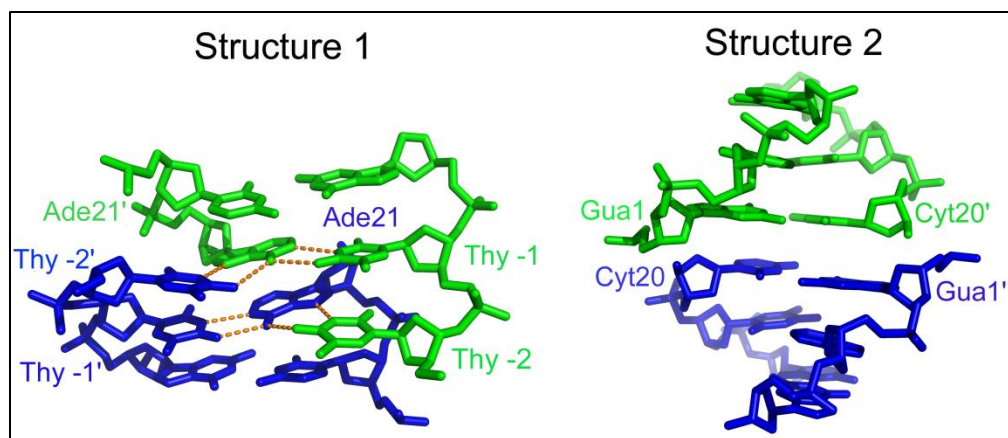
structures of the p53DBD tetramer/DNA complexes are essentially superimposable in the two crystal lattices (Figure 2.27), arguing that the solved tetramer structures reported here are biologically relevant (Rhodes, 2000, pgs 34 - 35). Unless otherwise stated, discussion will focus on the higher resolution structure 1.



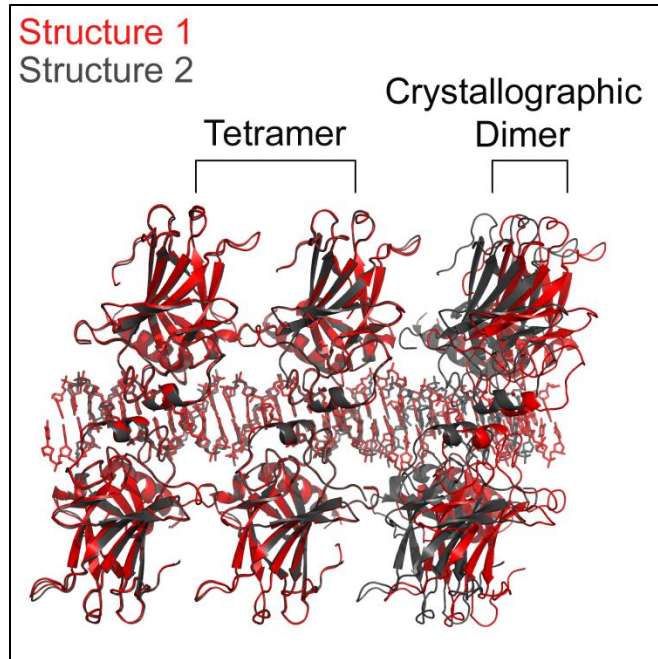
**Figure 2.24: DNA used for crystallization and view of crosslink.** The consensus sequence is boxed and crosslinked bases are in red. This figure is adapted from (Malecka et al, 2009).



**Figure 2.25: MALDI chromatogram for DNA used in p53DBD tetramer crystallization.** The major observed peak is 7314.1 Da while the calculated mass of the DNA duplex is 7310.1 Da.



**Figure 2.26: Crystal packing of DNA 5' and 3' ends in structures 1 and 2.** View of DNA base pairing between crystallographic DNA duplexes in structures 1 and 2 with involved bases numbered as in Figure 2.25. One duplex is colored green and the crystallographic duplex is colored in blue. Hydrogen bonds are depicted in orange. This figure is adapted from (Malecka et al, 2009).



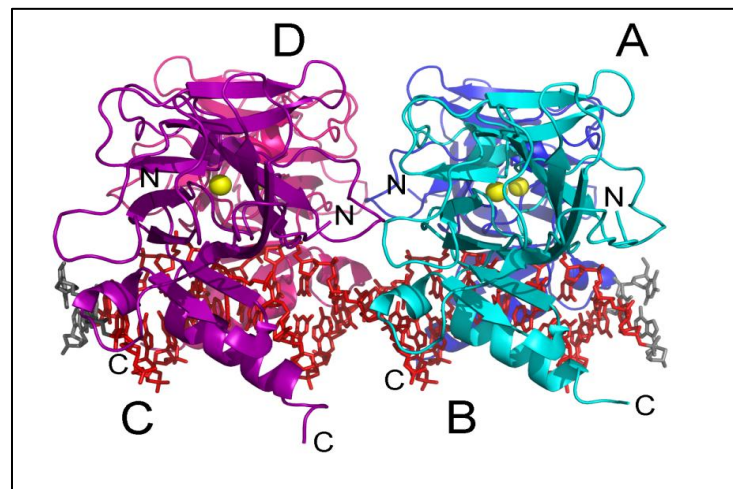
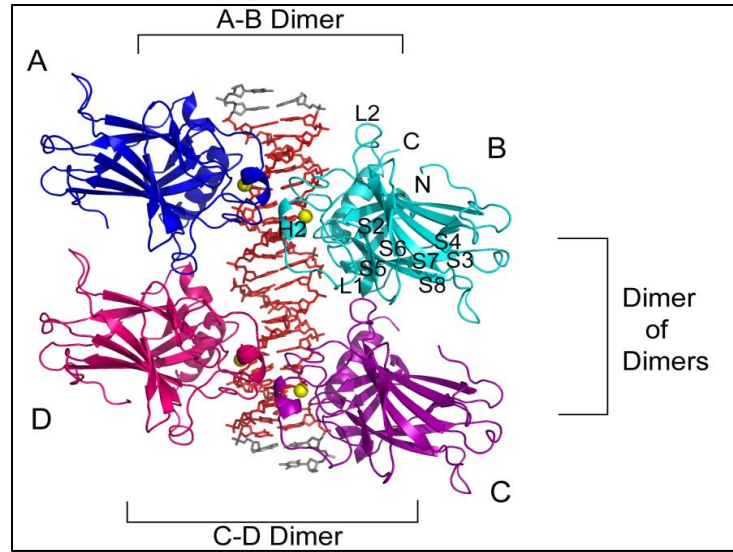
**Figure 2.27: Crystal packing of both structures 1 and 2.** View of tetramer crystal packing in both structures 1 and 2. Structure 1 is colored in red and structure 2 is colored in gray. A tetramer is shown for both structures with a dimer from the adjacent crystallographic tetramer shown to the right. A clear gap is seen between structure 1 and its crystallographic dimer neighbor while no such gap is seen for structure 2. This figure is adapted from (Malecka et al, 2009).



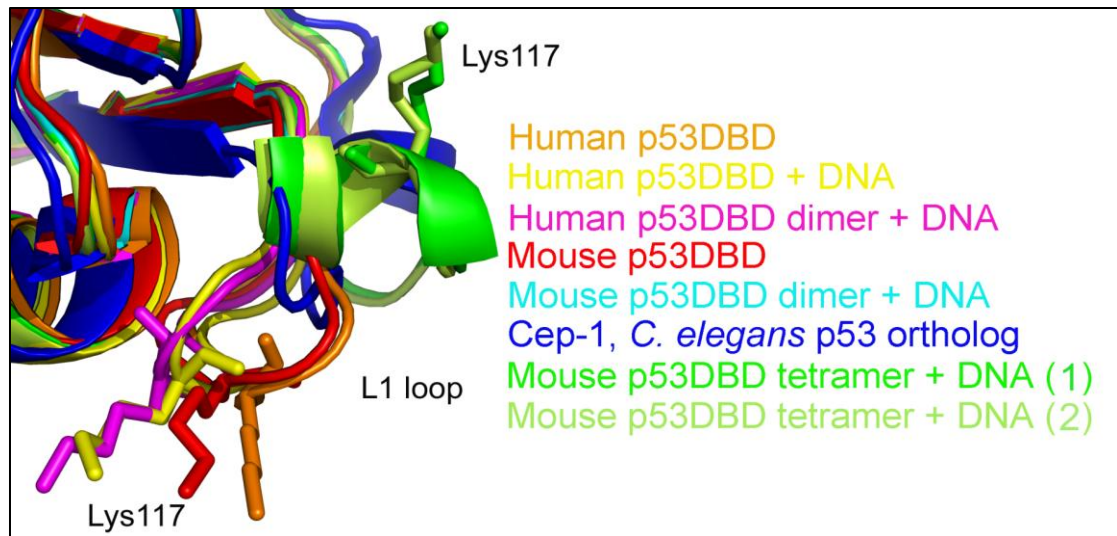
### 2.3.2. Structure of each p53DBD within the tetramer and DNA contacts

Within the tetramer, the structure of each of the p53DBD subunits is nearly identical to each other ( $C\alpha$  rmsd  $\leq 0.105$  Å) and to other p53DBD domain structures that have been determined ( $C\alpha$  rmsd  $\leq 0.559$  Å)(Cho et al, 1994; Wang et al, 2007; Zhao et al, 2001). Briefly, each domain is comprised of two anti-parallel  $\beta$  sheets that form an immunoglobulin-like  $\beta$  sandwich. Each strand is twisted resulting in two very different ends of the sandwich; one end is compact with short loops while the other end splays to the loop-sheet-helix motif, which contains a tetrahedrally bound  $Zn^{2+}$  ion, the L2 loop, and the H2  $\alpha$  helix (Figure 2.28).

The major difference between each tetramer subunit and the previous structures lies within the L1 loop, residues 109 – 119 (Figure 2.29). Both Cho et al. (human p53DBD monomer bound to DNA) and Kitayner et al. (human p53DBD dimer of dimers bound to a discontinuous DNA duplex) stress the importance of the L1 loop for DNA recognition since Lys120 (Lys117 in mouse) at the loop's tip points into the major groove to make a base-specific DNA contact(Cho et al, 1994; Kitayner et al, 2006). In contrast, here Lys117 and the L1 loop in each subunit has moved nearly 15 Å away from the DNA and adopts some  $\alpha$  helical structure.



**Figure 2.28: Structure of the tetrameric p53DBD/DNA complex.** Upper and lower panels show the overall view of the tetramer in cartoon representation. Subunits A (blue) and B (cyan) represent one dimer and subunits C (purple) and D (pink) represent the other. Zn<sup>2+</sup> ions are shown as yellow spheres. The DNA consensus sequence is rendered in red and other bases are rendered in gray. The lower panel is a 90° turn into the page followed by a 90° turn to the right of the upper panel. This figure is adapted from (Malecka et al, 2009).



**Figure 2.29: The L1 loops from several previously solved p53DBD structures.** A close up view of the L1 loops from human (2OCJ, orange), and mouse (1HU8, red) p53BD; human p53DBD (1TSR, yellow), human p53DBD dimer (2AC0, magenta), mouse p53DBD dimer (2GEQ, cyan), bound to DNA; Cep-1 p53 ortholog (1T4W, blue), mouse p53DBD tetramer bound to DNA structure 1 (green) and structure 2 (light green). This figure is adapted from (Malecka et al, 2009).

The DNA contacts made by each subunit to the DNA are essentially as reported by Cho et al. (Cho et al, 1994). Three minor variations, which are also observed in the other p53DBD/DNA structures, are seen in each subunit. First, Lys117 (Lys120 in human p53) of the L1 loop has moved away from the DNA and towards the tetramerization interface as previously described in subunits B and D of structure 1 and all subunits of structure 2 (Figures 2.29, Figure 2.32, and Figure 2.33). In subunits A and C of structure 1, this residue is disordered. Second, Arg280 (Arg283 in human p53), part of the H2 helix, is too far to contact the phosphate backbone. Third, Arg245 (Arg248 in human p53), the most frequently mutated residue in human cancers, adopts two different conformations in structure 1 and its DNA contact changes accordingly. In structure 2 and subunits A and C of structure 1, Arg245 sits in the minor groove and contacts the DNA as reported in Cho et al. However, in subunits B and D of structure 1, the same residue adopts a 90° bend and only contacts the backbone of one DNA strand. Taken together, analysis of the protein/DNA contacts within the tetrameric p53DBD bound to DNA reveals that the basic residues 117, 245 and 280 have variable roles in DNA recognition.

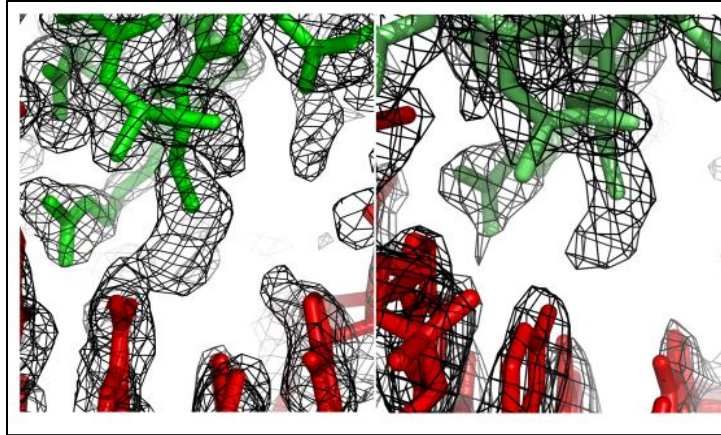
### 2.3.3. Structures of each p53DBD dimer and tethers

Each dimer (A-B and C-D, Figure 2.28) within the tetramer comes together as previously described by Shakked and coworkers for 2AC0 (Kitayner et al, 2006). Key residues involved are Arg178, Pro174, His175, Gly241 and Met243. Also, the dimers show little structural variation from the previously reported mouse p53DBD dimer ( $C\alpha$  rmsd 0.776 Å (Ho et al, 2006)).

Each p53DBD subunit is tethered to its quarter site via the side chain sulfur of Cys274 to the modified sulfur-bearing linker attached to the N4 atom of the cytosine 3' to the invariant

guanine (Figure 2.24). The extended tether has a length of 7.5 Å (Huang et al, 1998). Figure 2.30 shows simulated annealing composite omit maps for each p53DBD/DNA complex structure around residue Cys274. Structure 2 clearly shows a break in density, strongly indicating that the linking atoms are not rigid. Structure 1 shows connected density between the atoms but it lacks defined structure also supporting the tether's flexibility. The distance between the crosslinked atoms is less than 4.1 Å and nearly all the DNA contacts previously reported are intact suggesting that the tether is flexible and does not perturb the p53/DNA interface.

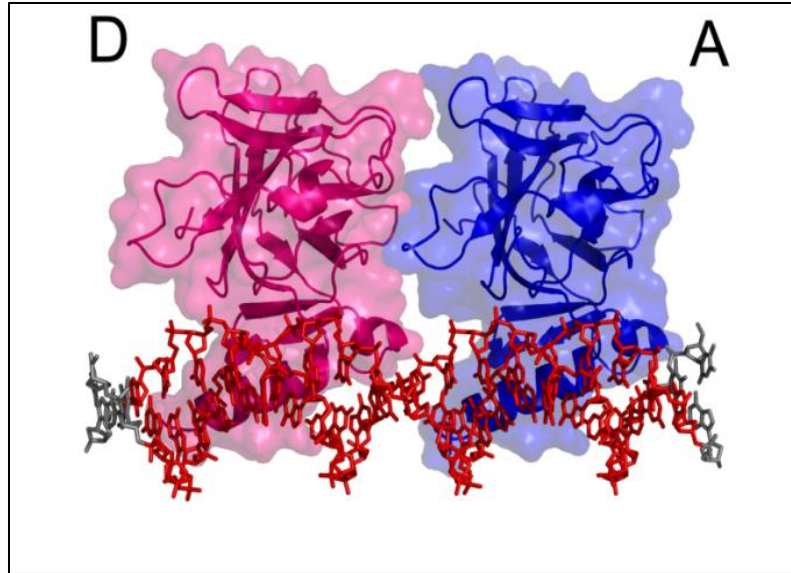
Taken together, the similarity of p53DBD-DNA interactions reported here to the DNA interactions with the monomeric p53DBD (Cho et al, 1994; Wang et al, 2007; Zhao et al, 2001) and dimeric p53DBD (Ho et al, 2006; Kitayner et al, 2006), and the lack of rigid tether density in the structures reported here, these data support the conclusion that the tether does not significantly disrupt the natural binding properties of the p53DBD to its consensus site.



**Figure 2.30: View of the electron density at the crosslink.** Simulated annealing composite omit  $2F_o-F_c$  map contoured at  $1.0 \sigma$  (left) and  $1.3 \sigma$  (right) at the area of crosslinking. DNA consensus sequence atoms are colored in red. Protein atoms are colored in green (structure 1, left) or light green (structure 2, right). This figure was adapted from (Malecka et al, 2009).

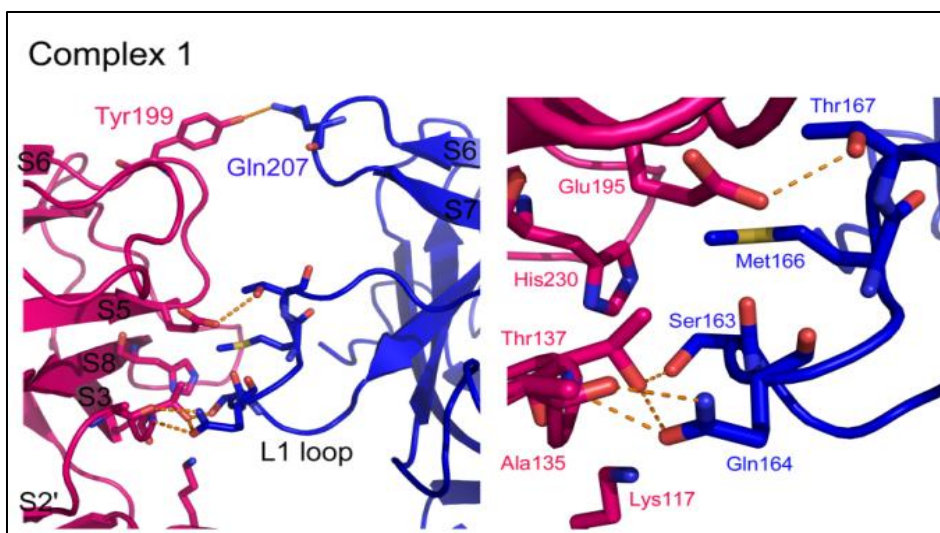
#### 2.3.4. *Structure of the p53DBD tetramer and dimer-dimer contacts*

The p53DBD tetramer binds to its DNA consensus sequence as a dimer-of-dimers as previously proposed (Figure 2.28). Each dimer (A-B and C-D) binds to its respective half site with no relative twist and without significant DNA bending (Figure 2.28, Figure 2.31). Each dimer-dimer interface buries an average of 665 Å<sup>2</sup> of surface area and shows nearly the same interactions in each structure. The major interface involves the L2 loop from subunits A and C tucking against the S2'/S3, S5/S6 and S7/S8 loops from subunits B and D (Figure 2.31) and is stabilized by both van der Waals forces and a network of hydrogen bonds (Figures 2.32, Figure 2.33). The backbone nitrogen of D-Ala135 (S2'/S3) donates a hydrogen bond to A-Gln164 OE1 (L2), which in turn accepts a hydrogen bond from the side chain hydroxyl of D-Thr137 (S3). This hydroxyl also accepts a hydrogen bond from the side chain of A-Ser163 (L2). D-Gln164 NE2 also donates a hydrogen bond to the backbone oxygen of D-Ala135, while the side chain of A-Thr167 (L2) donates a hydrogen bond to OE1 of D-Glu195 (S5/S6). Completing this major contact region are two van der Waals interactions: D-His230 CE1 (S8) with A-Ser163 CB and D-His230 CD2 with A-Met166 CE (L2). In structure 2 (Figure 2.33), the van der Waals contact involving A-Met166 CE is lost due to a different side chain conformation, but an additional hydrogen bond is present between D-His230 ND1 and the side chain of A-Ser163.



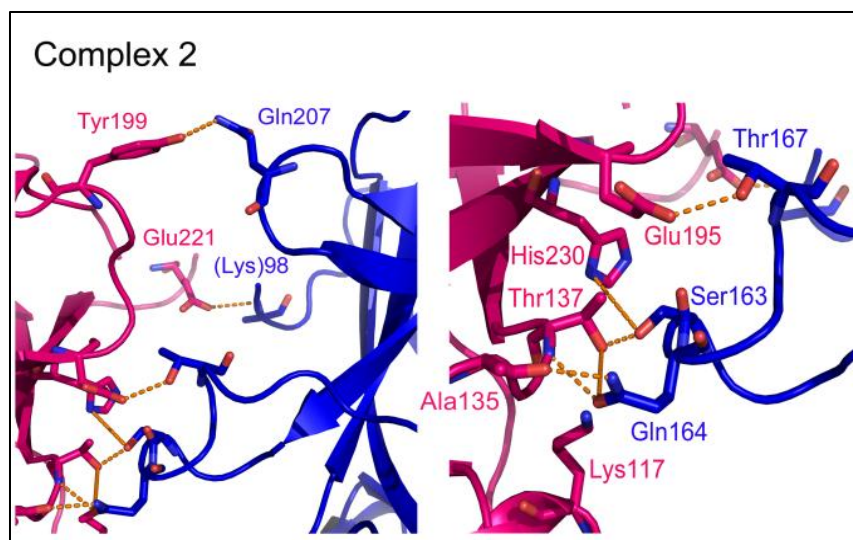
**Figure 2.31: p53DBD dimer-dimer formation.** View of subunit A (blue) and D (magenta) in surface representation for structure 1. Subunits B and C have been removed for clarity. This figure was adapted from (Malecka et al, 2009).





**Figure 2.32: Interactions at dimer-dimer interface for structure 1.** Overall view (left) and close up view (right) of the dimer-dimer interaction between subunits A and D for structure 1.

Hydrogen bonds are depicted as orange dashed lines. The same interactions are seen between subunits B and C. This figure was adapted from (Malecka et al, 2009).



**Figure 2.33: Interactions at dimer-dimer interface for structure 2.** Overall view (left) and close up view (right) of the dimer-dimer interaction between subunits A and D for structure 2.

Hydrogen bonds are depicted as orange dashed lines. The same interactions are seen between subunits B and C. This figure was adapted from (Malecka et al, 2009).

Both structures 1 and 2 show that D-Tyr199 (S5/S6) donates a hydrogen bond to AGln207 (S6/S7) near the top of each interface. Structure 2 has one additional hydrogen bond between D-Glu221 (S7/S8) and the backbone nitrogen of A-Lys98 (modeled as alanine in the structure). Overall, both structures have very similar dimer-dimer interfaces and involve nearly the same residues.

A sequence alignment of all p53DBD sequences available (Gasteiger et al, 2003) underscores the importance of these dimer-dimer interacting residues. Glu195 and Ala135 are strictly conserved, while Thr137 is strictly conserved in all species except chicken. Also, Gln207, Gln164 and Glu221 only show conservative mutations. This described interface is also consistent with previous studies implicating specific residues for dimer-dimer interactions (Cho et al, 1994; Ho et al, 2006; McClure & Lee, 1998; Pan & Nussinov, 2007; Pan & Nussinov, 2008) The most often cited tetramerization residues implicated for such interactions include Ser93, Ser96, Thr167, Ser163, Gln164 from one domain and Glu195, Thr137, and Glu221 from the other. All except for residues Ser93, Ser96 (disordered in these structures) and Gly196, each of the implicated dimer-dimer contact residues are observed to make contacts at the dimer-dimer interface of the structures reported here.

A curious area of the tetramerization interface involves the enigmatic Lys117 of the L1 loop. As previously stated, the L1 loop of each subunit has moved far from where other p53DBD and p53DBD/DNA structures have it placed (Figure 2.28, Figure 2.32, and Figure 2.33). Due to this movement, Lys117 points up towards residues D-Thr137 and A-Ser163 but does not make any specific contacts with the opposite dimer. Several possibilities exist to explain this conformational change. Careful analysis around Cys274 shows that the tether can only loop in

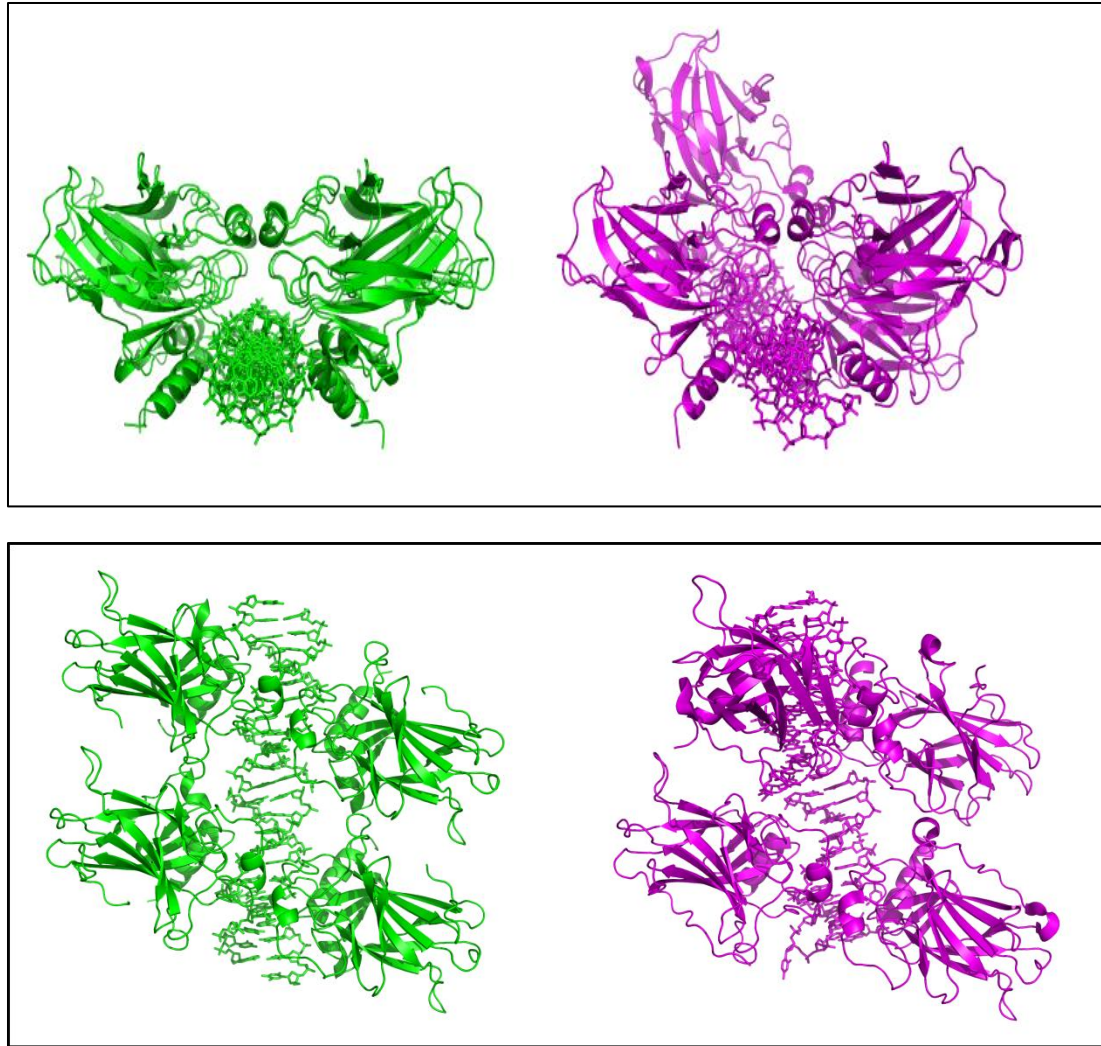
one direction, which is towards the L1 loop. In fact, the tether is likely blocking the DNA bases that Lys117 has previously been shown to contact. However, the tether is only 7.5 Å in length and the L1 loop has moved approximately 15 Å away, which is much too far to attribute only to steric hindrance by the tether. Alternatively, the current position of the L1 loop be stabilized by crystal contacts from a symmetry-related subunit or the presence of a citrate molecule, which sits close to the new conformation. Perhaps the most intriguing possibility is that the shift in position is induced by tetramerization itself. Lys117, which is strictly conserved across all species but shows low cancerous sensitivity (Ho et al, 2006), and although this residue is shown to make base specific contacts in several reported experimental p53/DNA complexes (Cho et al, 1994; Kitayner et al, 2006; Pan & Nussinov, 2008), but it also disordered in others (Ho et al, 2006; Kitayner et al, 2006; Pan & Nussinov, 2008). Within structure 1, Lys117 is only ordered in subunits B and D, the two p53DBD subunits in the crystal lattice that allow Lys117 to come close to the tetramerization interface. Lys117 is ordered in structure 2. Zupnick and Prives studied a K120A (Lys117 in mouse p53DBD) mutation for its effects on p53 function. They found that the mutant had strong DNA binding *in vitro*, but a weaker affinity *in vivo*. Overexpression in cells resulted in normal p53 function, however upon expression at normal cellular levels, the mutant showed defects in transactivation and apoptosis (Zupnick & Prives, 2006). It is interesting to speculate that Lys117 might alternate between DNA base contact and tetramerization contact, thus explaining its ambiguity. Clearly, additional studies will have to be carried out to resolve the role of Lys117 in the DNA binding properties of p53.

### 2.3.5. *Comparison with other p53DBD tetramers*

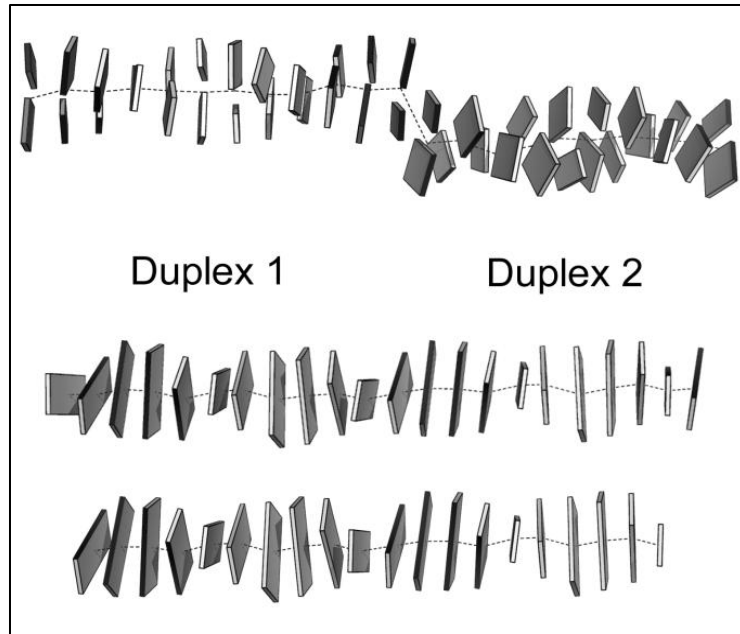
Two p53DBD tetramer complexes with DNA have been made available recently and vary in their structures. Ho et al. proposed a model p53DBD tetramer bound to DNA, which was generated by translating their crystallographic p53DBD dimer down the DNA to the next half site (Ho et al, 2006). This model predicted that the dimers come close at the L2 loop and S7/S8 regions, which is consistent with what we see in our current structures. A 40° DNA bend was also proposed in this model as well as models from Nagaich et al. (Nagaich et al, 1999) and Pan et al. (Pan & Nussinov, 2007). Interestingly, the present structures reveal that the DNA is relatively straight. Previous studies suggested that a straight p53DBD tetramer/DNA complex would not be possible due to steric clashes. To alleviate these clashes, it was proposed that the DNA would have to bend by as much as 50° (Balagurumoorthy et al, 1995; Nagaich et al, 1997a; Nagaich et al, 1999; Nagaich et al, 1997b; Pan & Nussinov, 2007). In contrast to these proposals, the two p53DBD tetramer/DNA structures reported here clearly show that a stable tetrameric complex can bind to straight, B-form DNA and maintain well-ordered dimer-dimer interfaces.

Shakked and coworkers reported on the structure of a p53DBD tetramer bound to discontinuous DNA where two p53DBD dimers bound to DNA came together end-to-end within their crystal lattices (Kitayner et al, 2006). These dimers are twisted 33° relative to each other in each structure (Figure 2.34). While, the two DNA half sites appear to be out of register for each (Figure 2.35), they can be considered approximate p53DBD tetramers and provide insight into alternative dimer-dimer contacts. Each tetramer shows its own set of limited dimer-dimer contacts. 2ACO, 2AHI, and 2ATA are quite similar in that one dimer-dimer interface is devoid of interaction and the other has limited hydrogen bonds and water-mediated interactions. 2ADY is

unique in that both interfaces have the same five hydrogen bond interactions. Interestingly, as in the p53DBD tetramer/DNA complexes reported here, Ser166, Gln167, and Thr170 (Ser163, Gln164 and Thr167 in mouse p53DBD) are involved in hydrogen bonds in the p53DBD tetramer/DNA complexes reported by Shakked and coworkers but these residues mediate different hydrogen bonds within the two complexes.



**Figure 2.34: Superposition of the p53DBD tetramer with 2AC0.** The upper panel is a view down the helical axis of 2AC0 after superposition of the p53DBD tetramer/DNA complex. The upper panel is the same except both structures have been rotated 90° into the page. 2AC0 is rendered in magenta and structure 1 is rendered in green. Each p53DBD tetramer/DNA complex is shown separately for clarity. The 33° twist of the back dimer relative to the front is easily seen for 2AC0. This figure is adapted from (Malecka et al, 2009).



**Figure 2.35: Comparison of the DNA from the p53DBD tetramer and 2AC0.** DNA from 2AC0 (upper), structure 1 (middle) and structure 2 (lower) as Calladine-Drew schematics. Duplex 1 and 2 of 2AC0 are labeled and the central helical axes are shown as dotted lines. Duplex 1 is lower and rotated toward the viewer relative to duplex 1. All schematics were made with 3DNA(Lu & Olson, 2003) and rendered in PyMOL(DeLano, 2002). This figure was adapted from (Malecka et al, 2009).



### 2.3.6. *Multiple tetramerizations modes for p53 on DNA*

After comparison of the p53DBD tetramer/DNA models with the two complexes reported here, we hypothesize that dimer-dimer contacts within the p53DBD tetramer on DNA can vary *in vivo*. This hypothesis is supported by several lines of evidence. First, p53 response elements are varied across the genome due to different base-pair spacing between half sites and response elements which fulfill the consensus sequence requirement (El-Diery et al, 1992). Second, the current structure shows that the DNA need not be bent for four p53DBD domains to stably bind to it. Given that solution studies and molecular dynamics suggest that bent DNA complexes are possible (Nagaich et al, 1999) and the inherent flexibility of the DNA sequences has been studied (McNamara et al, 1990; Nagaich et al, 1994; Zhurkin et al, 1991), it is likely that other energy minima exist besides the one captured crystallographically here. Also, while several of the residues at the interface involved are conserved, cancerous mutations of the residues are low compared with DNA contacts or monomer-dimer contact residues within the p53DBD (Ho et al, 2006). Taken together, these observations suggest that perhaps the tetramerization interface is not fixed, but able to accommodate different contacts depending on different cellular conditions.

### 2.3.7. *p53 DNA binding cooperativity*

It has been shown that the predominantly monomeric p53DBD will bind to its consensus sequence DNA in a highly cooperative manner (Balagurumorthy et al, 1995) and the basis for this cooperativity has been the subject of much speculation. McClure et al. studied how p53DBD dimers and tetramers could bind DNA oligomers by varying the placement of the consensus quarter sites. Results from this work showed that two p53DBD dimers will bind a full consensus site for a far longer period of time ( $t_{1/2} \sim 15$  minutes) than one dimer alone ( $t_{1/2} \sim 30$  seconds) (McClure & Lee, 1998). A similar experiment was performed with wild-type p53 (residues 94 – 360, tetrameric) and an L344A mutant construct that could only form dimers. The results showed that tetramers had a six-fold greater affinity for DNA over dimers (Weinberg et al, 2004). Correlating this data with the two structures reported here suggests that two dimers are more stably bound to DNA due to the numerous dimer-dimer contacts that are observed in the structures reported here. One dimer bound to DNA benefits from both stabilizing DNA and dimerization contacts (Klein et al, 2001), but according to McClure et al and supported by Weinberg et al, will still diffuse away from the DNA in a short period of time. With the added dimerization contacts, the complex is more stable, as evidenced by the increase in  $t_{1/2}$  and affinity.

Other suggestions have been reported for p53DBD cooperativity but are not supported by crystal structures (McClure & Lee, 1998). No conformational change is seen within the p53DBD upon tetramerization or DNA binding. While it is possible that other areas of the protein may undergo structural rearrangement during DNA binding, cooperativity is seen at the level of p53DBD binding to DNA alone, thus implicating determinants for DNA-binding cooperativity

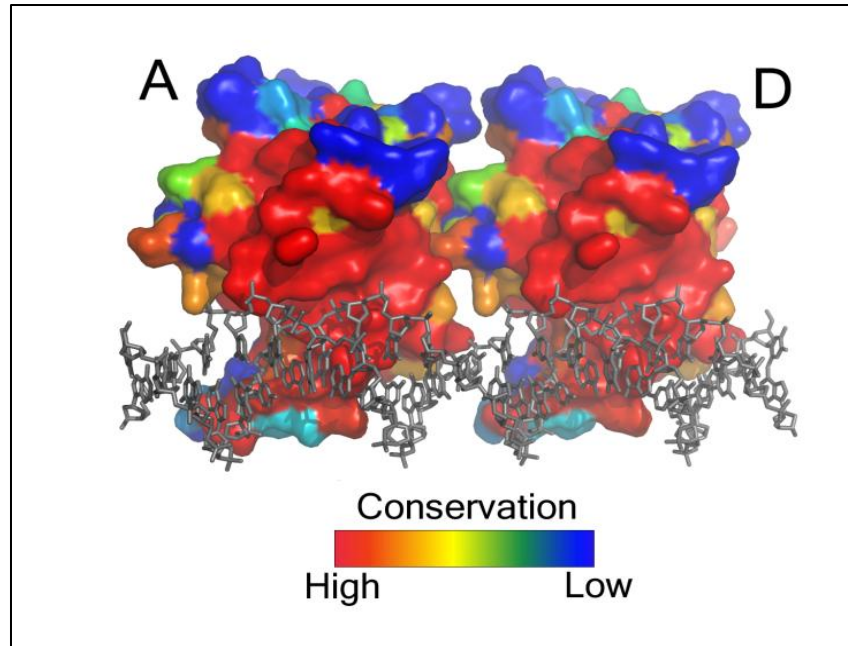
within the DNA-binding domain of p53. Also, cooperative DNA binding by p53 is largely influenced by its tetramerization domain *in vivo*. Taking these observations together with the p53DBD tetramer/DNA complexes reported here suggests the DNA-induced interactions that are observed between the two p53DBD dimers plays a major role in cooperative p53DBD binding to DNA.

### 2.3.8. *p53DBD conservation and implication for function*

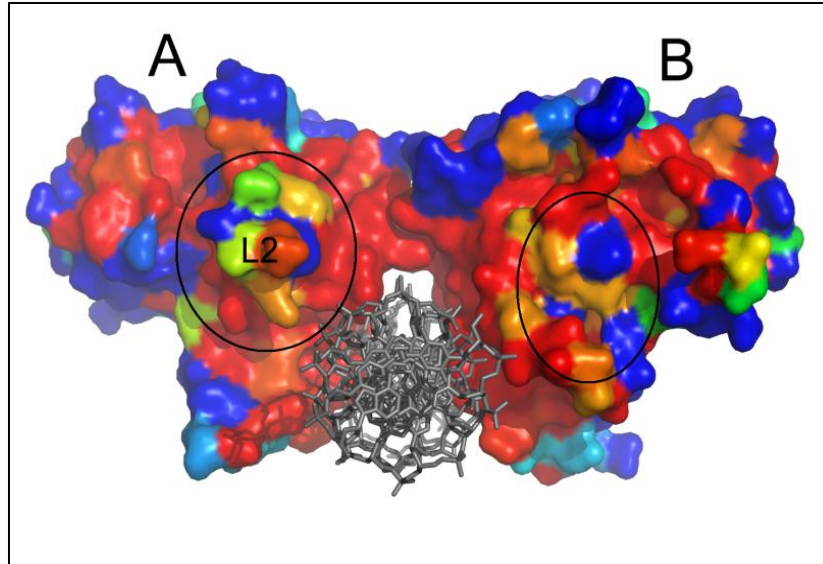
A surface conservation map of the p53DBD in the context of the p53DBD tetramer bound to DNA reveals several areas of high conservation and by inference functional importance. The highest degree of conservation maps to the DNA binding region of the monomer and the monomer-dimer interface (Figure 2.36) highlighting the relative importance of these regions. In contrast, the dimer-dimer interface shows far less conservation (Figure 2.37) arguing for a less significant role of p53DBD dimer-dimer contacts for DNA recognition. Subunit A shows a long groove of conservation that is set back from the L2 loop. A mirrored pattern on subunit B would suggest a strict dimer-dimer interaction, but this is not seen and further supports the model that multiple tetramerization modes can accommodate p53 binding to DNA.

Intriguingly, a comparison of subunit A and B shows that one side of each subunit is more conserved than the other (Figure 2.37). For subunits B and D, this highly conserved surface would face the solvent rather than the opposite dimer. It is possible that these residues participate in binding other domains of p53 or partner proteins of p53. This could also be the

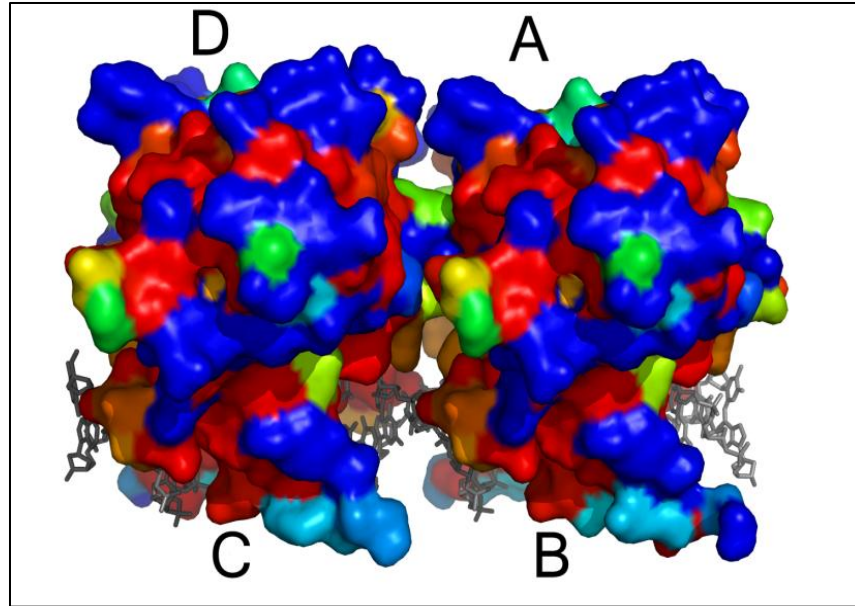
case for the ring of conserved amino acids that map around the back surface of each subunit (Figure 2.38). Taken together, the sequence conservation on the surface to the p53DBD tetramer bound to DNA suggests other protein binding and perhaps regulatory surfaces.



**Figure 2.36: Surface map showing the most highly conserved regions of p53DBD.** A sequence alignment was performed with all known sequences of p53 using ClustalW(Gasteiger et al, 2003) and mapped with ESPrpt(Gouet et al, 1999). View of subunits A and D with strictly conserved residues colored in red and nonconserved residues colored in blue. The DNA is rendered in gray. Subunits B and C have been removed for clarity. This figure was adapted from (Malecka et al, 2009).



**Figure 2.37: Surface map showing the dimer-dimer interface p53DBD tetramerization.** The same techniques were used here as in Figure 2.37. This is a view along the DNA helical axis of subunits A and B. Subunits C and D have been removed for clarity. The dimer-dimer interface is circled with a black line. This figure was adapted from (Malecka et al, 2009).



**Figure 2.38: Surface map showing the back of the dimer-dimer interface.** The same techniques were used here as in Figure 2.37. This is the same orientation as Figure 2.37, but focusing on the back of the subunits. The band of conserved residues is notable. This figure was adapted from (Malecka et al, 2009).

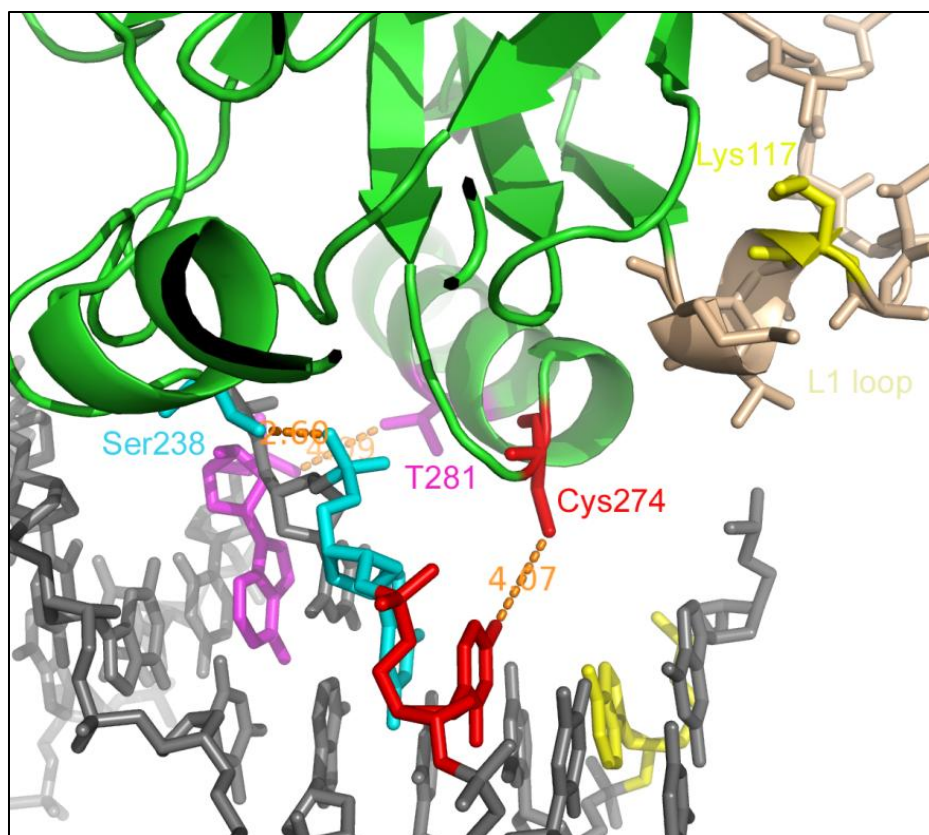
### 2.3.9. Backbone crosslinking and mixed crosslinking results

As previously discussed, the L1 loop is an intriguing area of the p53DBD (Figure 2.29). Some structures, such as the original human p53DBD(Cho et al, 1994) bound to DNA and some of Kitayner et al's tetramers(Kitayner et al, 2006), show Lys120 pointing into the major groove and making a specific contact. Other structures, such as the mouse p53DBD dimer structure(Ho et al, 2006), the mouse tetramer structures reported here, and some of Kitayner et al's structures(Kitayner et al, 2006), Lys120 or the entire L1 loop is disordered.

In 3EXJ and 3EXL, the distance between the sulfur of Cys274 and N<sub>4</sub> of the cytosine is ~ 4 Å and the entire linker is 7.5 Å so the rest of the linker must loop out (Figure 2.30). Careful analysis of the area surrounding Cys274 shows that the extra linker is potentially covering the purines Lys120 has been reported to contact. However, the L1 loop has moved nearly 15 Å away in response, which is much too far to attribute to the linker alone.

Curious to know if this movement was due to tetramerization or linker, we set about moving the crosslink away from the L1 loop. Base crosslinking was not possible anywhere else in the structure, but backbone crosslinking, which had recently been described by Verdine and colleagues(Banerjee et al, 2006), was a good option. Thr281 came within 4 Å of the phosphate backbone between the invariant cytosine and 3' adenosine so, by use of site-directed mutagenesis, became Cys281 (Figure 2.39).





**Figure 2.39: Close up view of Cys274 and other options for crosslinking.** A close up view of the amino acids and DNA bases surrounding Cys274 and the L1 loop is shown. Cys274 and the cytosine bearing the linker are shown in red. The L1 loop is rendered in beige with Lys117 (Lys120 human numbering) colored in yellow. The base Lys120 has been shown to contact in 1TSR is also colored in yellow. Thr281 and the phosphate group it contacts is colored in magenta. Distances are shown as dashed orange lines and labeled.

Two different crosslinking strategies, both using the same DNA sequence that previously crystallized with the p53DBD tetramer, were attempted. The first attempt had two backbone linkers per DNA strand (Figure 2.40). DNA yields were preciously low, but some tetramer was able to be purified and a few crystallization conditions were tried. One condition yielded crystals that were difficult to reproduce.

The second type of crosslinking, referred to as mixed crosslinking, incorporated one base linker and one backbone linker (Figure 2.41). The L1 loops of subunits B and D are present at the tetramerization interface, while subunits A and C have their L1 loops facing outwards (Figure 2.28). From this, we decided that only subunits B and D needed backbone crosslinking while subunits A and C could still have base crosslinking and not perturb the L1 loop at the tetramerization interface. This type of DNA strand was easier to synthesize and lead to higher yields. The p53DBD/DNA complex was purified and crystallization was attempted. One condition yielded crystals after at least 1 month. Unfortunately, these crystals did not diffract at the home source.

## Original:

5' – TT GAGCATGCTC GAGCATGCTC A – 3'  
3' – A CTCGTACGAG CTCGTACGAG TT – 5'

**Figure 2.40: DNA bearing two backbone linkers.** This DNA corresponds to the first attempt to move the linker away from the L1 loop. The phosphate group between the two magenta bases bears the cystamine linker. This DNA duplex sequence is the same as the duplex used to crystallize the p53DBD/tetramer by base crosslinking.

## Original:

5' – TT GAGCATGCTC GAGCATGCTC A – 3'

3' – A CTCGTACGAG CTCGTACGAG TT – 5'

**Figure 2.41: DNA with mixed crosslinking.** This DNA corresponds to the second attempt to move the linker away from the L1 loop. The phosphate group between the two magenta bases and the red cytosine bear a cystamine linker. This DNA duplex is the same in sequence as the duplex used to crystallize the p53DBD/tetramer by base crosslinking.

## 2.4. New p53 Tetramers

### 2.4.1. *Self Assembled Tetramer (Chen et al, 2010)*

Since the publication of the tetramer reported in this text, a human tetrameric p53DBD/DNA structure has been reported(Chen et al, 2010). This self-assembled complex was crystallized after an exhaustive search of many DNA consensus sequences with varying flanking bases of different lengths.

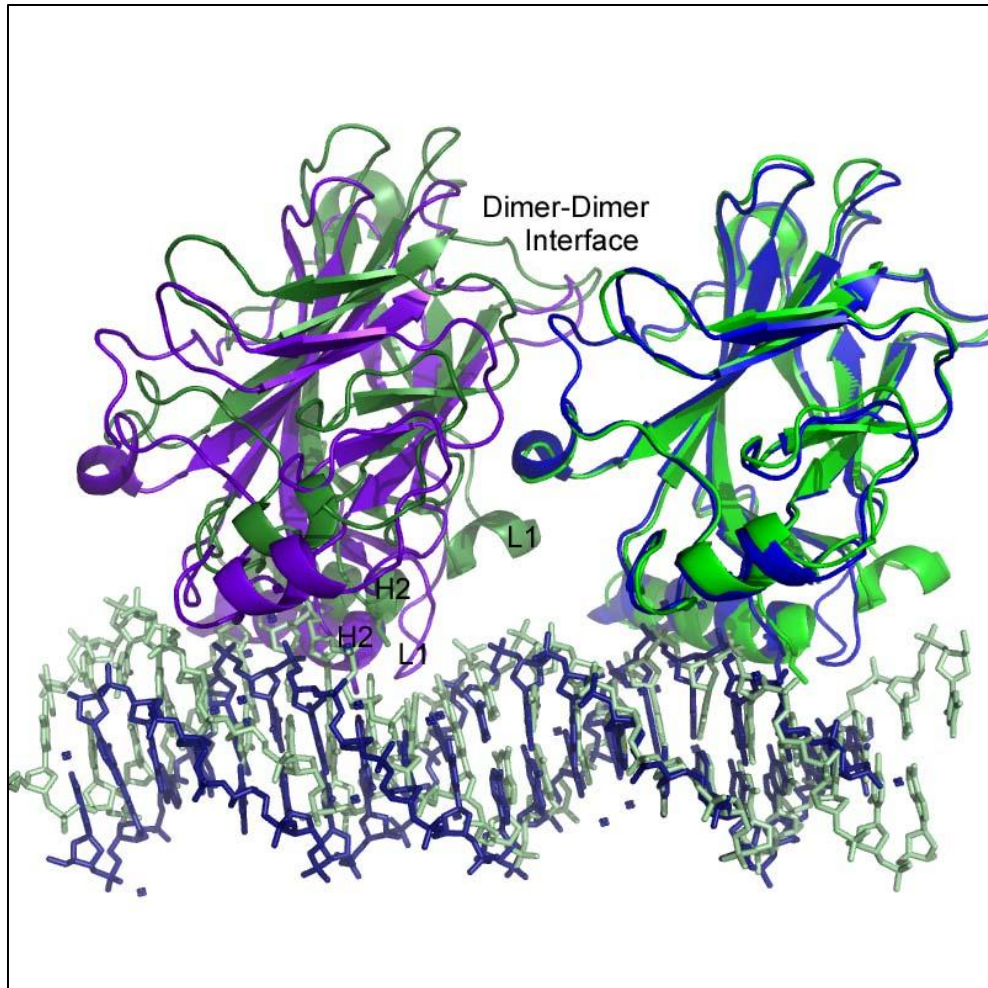
Similar to 3EXJ and 3EXL, each subunit of this new tetramer forms the previously reported DNA and dimerization contacts in addition to having the same overall structure ( $C_{\alpha}$  RMSD  $\leq$  0.601 Å). Each dimer-dimer interface (A-D, B-C) is also symmetrical and the B-form DNA has no discernable bend(Chen et al, 2010).

While the many of the previously highlighted residues are involved in dimer-dimer contacts, they mediate different interactions. Also well-ordered at the dimer-dimer interface are more p53DBD N terminal residues. In 3EXJ, the first modeled residue is 98 while in 3KMD, it is 92. These 6 residues are present at the dimer-dimer interface and contribute to the increase in contacting surface area (665 Å<sup>2</sup> vs 897 Å<sup>2</sup>)(Chen et al, 2010). There is also a shift in the way the subunits are interacting at the dimer-dimer interface, which also leads to more extensive contact between the two subunits.

The authors assert that the crosslinked tetramer could not form this conformation because of the tether(Chen et al, 2010). However, this seems unlikely for several reasons. Granted, the tether is longer than the distance between the sulfur of Cys274 and the N<sub>4</sub> of the crosslinked cytosine so the extra atoms must be somewhere. To the left, the H2 helix is packs close to the

DNA and to the right, the area has the flexible L1 loop followed by mostly open space. It is extremely likely that the linker loops towards the right. Because the new placement of the L1 loop is an area where nothing exists in 3KMD (Figure 2.42), it also seems unlikely that its new placement is hindering any sort of tetramerization.

It is possible that alternatively crosslinked tetramers will look much like 3KMD, 3EXJ or they could assemble with yet another unreported dimer-dimer interface. These differences simply underscore the inherent plasticity in the DNA double helix and the p53DBD protein under conditions of tetramerization.



**Figure 2.42: 3KMD vs. 3EXJ tetramerization.** Superposition of both tetramers via subunit A (right). Subunits D (left), A (right), and DNA of 3KMD (purple, blue, dark blue) and 3EXJ (dark green, bright green, light green) are shown. The shift of subunit D in 3KMD is easily seen compared to subunit D of 3EXJ. The L1 loop and H2 helix of subunit D is labeled for both tetramers. Cys274 is shown as a stick model to highlight the area of crosslinking for 3EXJ. The dimer-dimer interface for both tetramer is also labeled.

## 2.5. p53 Tetramerization and Conclusions

Considering the proposed variability of the dimer-dimer contacts within the tetrameric p53DBD bound to DNA, one might ask why p53 binds DNA as a tetramer. This could simply be to increase the affinity of p53 for DNA. It is also possible that other regions of p53 interact with the DNA binding domain to support tetrameric binding to DNA. Correlating with this possibility, the N- and C- terminal ends of the p53DBD are located at the dimer-dimer interface (Figure 2.31, Figure 2.32, Figure 2.33). Given that p53 has been reported to interact with many proteins, it is also possible that the p53DBD tetramer bound to DNA provides a unique protein contact surface. Taken together, these correlations suggest that the p53DBD domain may play a more active role in p53 function than simply providing a scaffold for DNA binding.



## 2.6. Future Directions

Tetramerization of p53 in the context of the full length protein is still a large research interest. Where is the tetramerization domain once the core domains are bound properly to DNA? Some small angle X-ray scattering(Tidow et al, 2007) as well as electron microscopy studies(Okorokov et al, 2006) have postulated where these domains reside, but it is still under investigation. The L1 loop and the importance (or not) of Lys117/Lys120 is still under debate, as well.

To that end, it would be useful to improve the crystals with mixed crosslinking and potentially collect a dataset. Chen et al. report that their Lys120 makes a base specific contact when no tethers are used and Kitayner et al. show that the L1 loop is almost always in the same conformation regardless of DNA sequence used, but the sequence of 5' – GAG –3' has not been tested in their systems. Given the low cancerous sensitivity of Lys120(Cho et al, 1994), it is still up in the air as to whether Lys117 needs to be in that conformation or it could move to the tetramerization interface when the DNA sequence is right.

All of the structural work thus far has only focused on icosameric sequences. El Deiry et al proved that p53 can bind to half sites that are separated by up to 13 base pairs(El-Diery et al, 1992). What do those tetramers bound to DNA look like? Do the DBDs have any contact at all? Do the DBDs bend the DNA in such a way that protein-protein interactions can still occur? It would be interesting to pursue some non-icosameric sequences for crystallization study and answer more of the questions that this fascinating protein poses.

## **CHAPTER 3**

### **Identification and characterization of small molecule human papillomavirus E6 inhibitors**

### 3.1. Introduction

#### 3.1.1. *Human papillomaviruses*

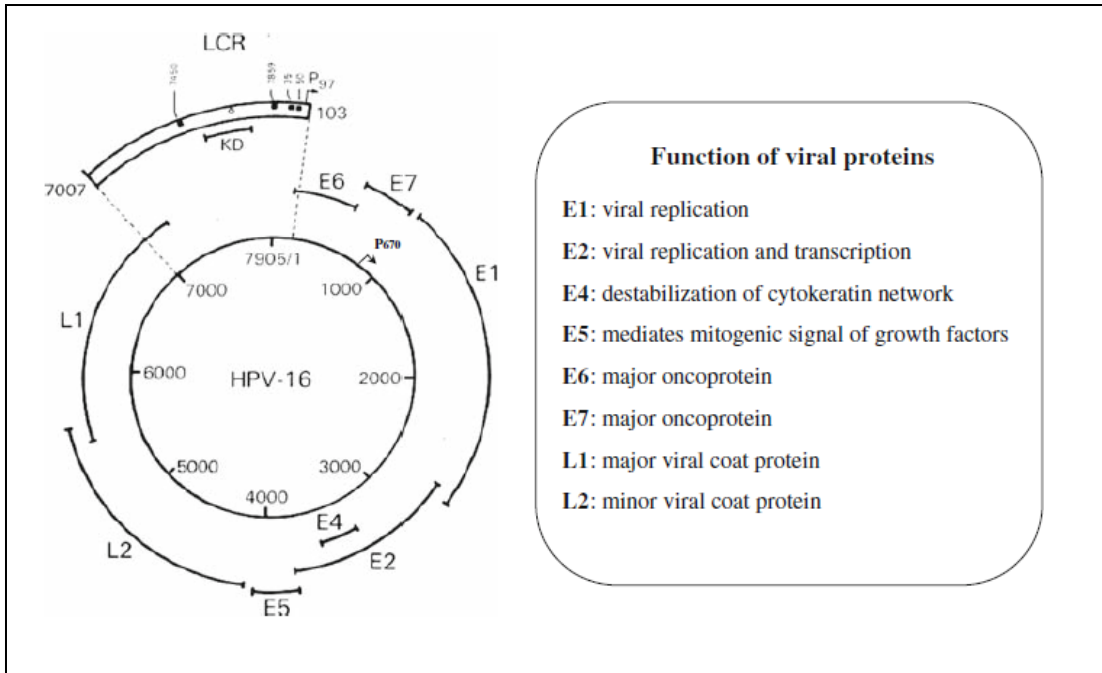
More than 130 different types of mucosal and cutaneous human papillomaviruses (HPVs) have been identified. A phylogenetic tree of HPVs, which has been developed based on the nucleotide sequence of the major capsid protein L1, breaks the family into five genera (de Villiers et al, 2004).

The beta genus is comprised of several HPV types that preferentially infect the skin and have circumstantially been linked to non-melanoma skin cancers. The gamma, mu, and nu all contain HPV types that lead to the development of cutaneous papillomas and warts (de Villiers et al, 2004).

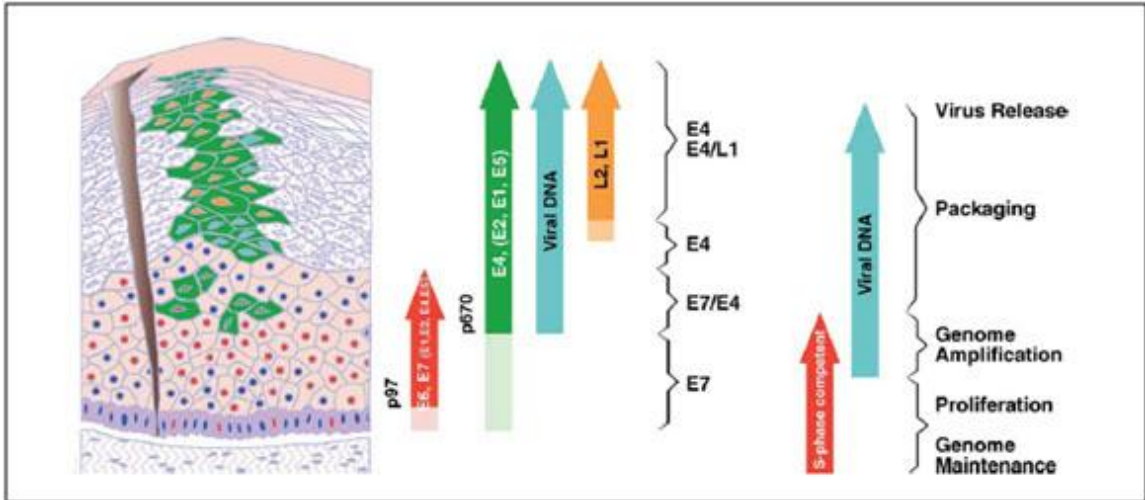
Approximately 30 mucosal HPV types fall within the alpha genus (de Villiers et al, 2004; Hausen, 2002). This group can be further subdivided into low risk HPVs (e.g. types 6 and 11) and high risk HPVs (e.g. types 16 and 18) with the majority infecting the mucosa of the genital tract. Low risk HPVs cause benign cellular proliferations leading to cutaneous or genital warts and papillomas. High risk HPVs, on the other hand, are associated predominantly with carcinomas of the anogenital tract (Hausen, 1996; Hausen, 2002). Cervical cancer is one of the leading causes of women's death in developing countries and is the sixth most common cancer in women worldwide (Baseman & Koutsky, 2005).

Regardless of genus, all HPVs have a circular, double stranded DNA genome of approximately 8000 base pairs and containing 8 – 9 open reading frames (Ghittoni et al, 2010)

(Figure 3.1). Following infection at the basal layers of stratified epithelia, the early genes (E1, E2, E4, E5, E6, and E7) are turned on to promote viral replication. As the virus moves up the stratification, the late genes L1 (major capsid protein) and L2 (minor capsid protein) are expressed so new viruses can be packaged and shed upon cellular rupture (Figure 3.2) (Doorbar, 2005; Ghittoni et al, 2010). An area of noncoding DNA, termed the long control region (LCR), is also present between the L1 and E6 open reading frames. It can vary in length between viral types and is involved in regulation of viral DNA replication/transcription (Ghittoni et al, 2010).



**Figure 3.1: Genome of the human papillomavirus type 16.** The circular, double stranded DNA genome is denoted schematically. The open reading frame for each gene is listed along with a designation for each gene product's function. This figure is adapted from (Ghittoni et al, 2010).



**Figure 3.2: Infection of epithelia by HPVs.** The far left panel shows a cross section of stratified epithelia. The stratification is created by cells at undergoing differentiation from basal cells (actively dividing, purple cells at the bottom of the cross section) to keratinocytes (terminally differentiated cells, top of the cross section). The green cells represent those infected with HPV and are still actively dividing throughout the stratification. The cluster of green cells near the top of the cross-section represents a lesion, which would be visible under the microscope. The right hand side of the figure shows where in the stratification each gene of the HPV genome is being expressed. This figure was adapted from (Doorbar, 2005).

### 3.1.2. *E6 and E7 oncoproteins*

Most HPV infections are cleared with time due to the cellular immune response. However, a small proportion of infections become chronic due to the integration of the high risk viral genome into the hosts' chromosomes. Expression of the viral oncoproteins E6 and E7 is typically increased once the viral genome has been integrated, which leads to cell proliferation by disruption of the cell cycle and eventual immortalization of infected cells, chromosomal instability and malignant transformation(Munger & Howley, 2002).

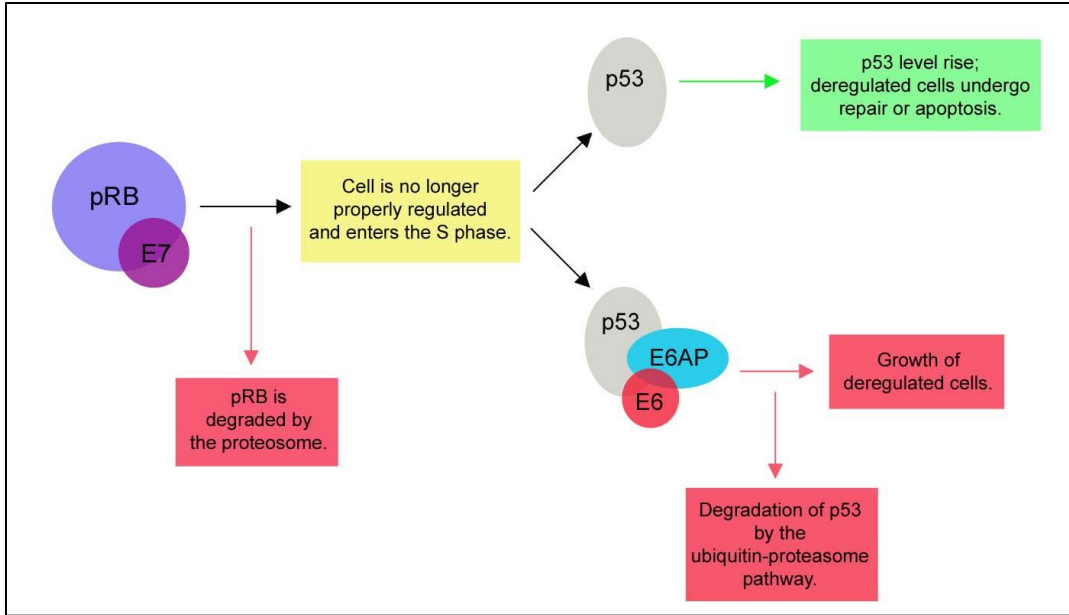
Integration of the HPV genome happens randomly within the host genome, but it is suspected that it may happen more frequently at fragile chromosomal spots. The frequency of integration can be correlated with the severity of cervical disease(Wentzensen et al, 2004). Following integration, most of the HPV genes are no longer expressed, except for the oncogenes E6 and E7. Continual expression of these two genes is necessary to maintain malignancy while suppression of both genes results in cellular death(Alvarez-Salas & diPaolo, 2007).

The major pathway of cell cycle disruption involves a concerted effort from both high risk E6 and E7. In healthy cells, the p110<sup>RB</sup> family of proteins (pRB) is bound by the family of transcription factors E2F/DP (E2F). Phosphorylation of pRB results in the disruption of the pRB/E2F interaction and subsequent entrance of the cell in to the S phase(Dyson, 1998). E7 is small, dimeric protein(Liu et al, 2006) that can bind to pRB and mimic the phosphorylation signal(Munger & al, 2001), which cause premature release of E2F and brings about the degradation of pRB (Figure 3.3). In response to this deregulated entrance into the S phase, the tumor suppressor protein p53 would be activated causing the cell to either undergo cell cycle

arrest or apoptosis. In an HPV positive cell, however, E6 binds to the cellular E3 ligase, E6-Associating Protein (E6AP), and this heterodimer then targets p53 for degradation via the ubiquitin-proteasome pathway(Huibregste et al, 1993; Scheffner, 1998; Scheffner et al, 1993) (Figure 3.3).

Both E6 and E7 have been shown to have other functions outside of this pathway. E7 can associate with p21<sup>WAF1/CIP1</sup> and p27<sup>KIP1</sup> to disrupt the cell cycle, in addition to interacting with the cyclin A/CDK2 complex and altering kinase activity within the cell(Arroyo et al, 1993; He et al, 2003). On the other hand, E6 has been implicated in a wide range of other activities. The C terminus of E6 contains a PDZ binding motif that targets cytoplasmic membrane proteins, including hDLG, Scribble, MUPP1, and MAG1-3(Gardiol et al, 1999). E6 can bring about degradation of these proteins, which causes the loss of large protein scaffolds, cell-to-cell contacts, and cell polarity(Gardiol et al, 1999). Also, the C terminus of E6 can bind four-way DNA holliday junctions(Ristriani et al, 2000; Ristriani et al, 2001) suggesting other ways that it can lead to cellular transformation. Finally, E6 has been implicated in disrupting p53-dependant gene activation by inhibiting p300/CBP acetylation(Patel et al, 1999; Thomas & Chaing, 2005).





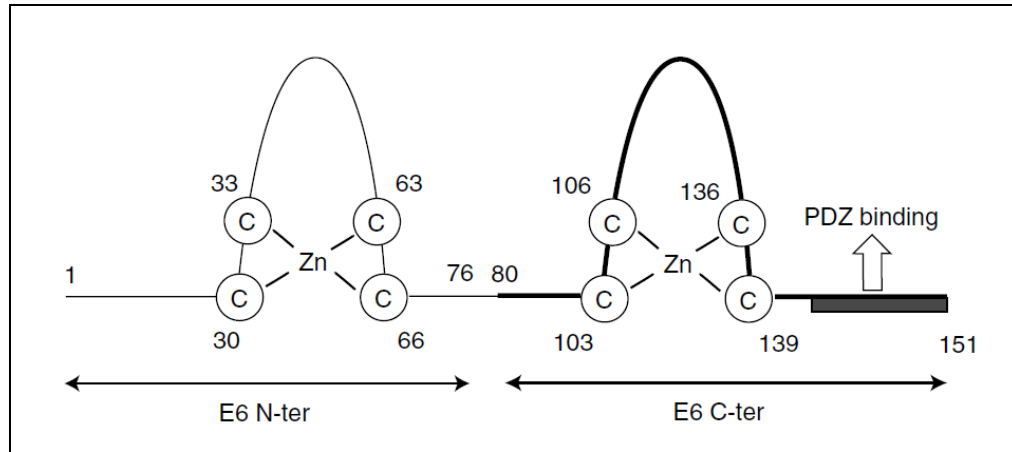
**Figure 3.3: The concerted effort of E6 and E7 to cause cell cycle disruption.** E7 binds pRB, forces the cell into the S phase and brings about pRB degradation. In a cell absent of E6, p53 levels would rise and transcribe genes involved in cell cycle arrest or apoptosis. In an HPV positive cell, E6 has bound E6AP and together they target p53 for degradation via the ubiquitin-proteasome pathway.

### 3.1.3. *Structural studies of E6*

The structure of E6 has been sought since its discovery in 1987(Seedorf et al, 1987). Unfortunately, this 158 amino acid protein has been difficult to study due to its propensity for aggregation and precipitation (Lipari et al, 2001; Nomine et al, 2003; Nomine et al, 2006).

E6 is known to be comprised of both an N and C terminal domain, one of which most likely arose from a gene duplication event. Sequence similarity between the two domains is low (< 20%), however quirks are seen in the sequence. For example, the amino acid motif of Cys-X-X-Cys is found four times within the entire protein with the sequence XXXLXXXE following both the first and third Cys-X-X-Cys motif(Cole & Danos, 1987). The chances of these quirks arising from a situation other than gene duplication are remote.

Schematically, the structure of full length E6 is shown in [Figure 3.4](#). Briefly, the N terminal domain is connected to the C terminal domain by a central, conserved linker. The extreme ends of the protein each have a loop region, with the C terminal loop containing a PDZ binding motif. Both domains each contain two Cys-X-X-Cys motifs, where all four cysteines ligate one Zn<sup>2+</sup> ion(Nomine et al, 2005).



**Figure 3.4: Schematic representation of the full length HPV16 E6 protein.** Residues numbers are given for each individual domain and the amino acids involved in  $Zn^{2+}$  ligation. This figure is adapted from (Nomine et al, 2005).

In 2001, Cordingley and colleagues studied the biochemical properties of the HPV16 E6 N terminal domain. Their construct contained residues 2 – 77 and was called E6-N. They assert that E6-N was highly soluble in low salt buffers. After an array of studies, they determined that one Zn<sup>2+</sup> ion was bound by the domain, the exact cysteines required for Zn<sup>2+</sup> ligation, and postulated that it was the C terminal domain of E6 that was responsible for the full length protein's purification difficulties(Lipari et al, 2001).

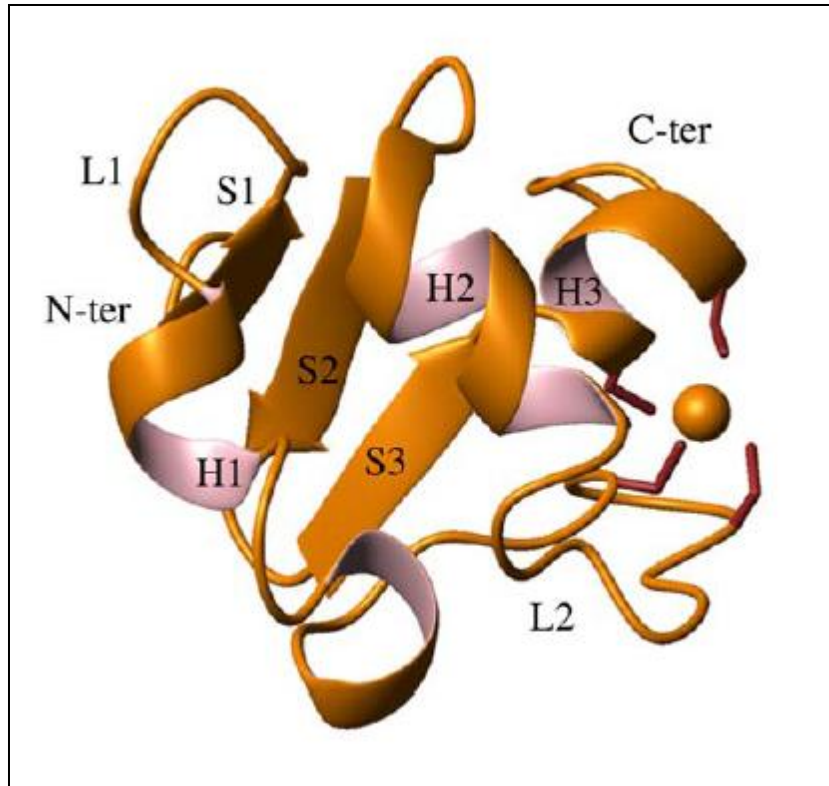
Two years later, work by the laboratory of Gilles Travé challenged Lipari et al's results by showing that it was the C terminal domain of HPV16 E6 that was well behaved. They expressed individual constructs of each domain (residues 7 – 83 and residues 87 – 158) and found that the N terminal domain was poorly behaved while the C terminal domain was easily expressed and concentrated up to 1 mM(Nomine et al, 2003). These results were also supported by work performed in our laboratory. The authors were then able to show that the domain was monomeric and had equal proportions of  $\alpha$  helix and  $\beta$  sheet. Interestingly, this domain also retained the full length protein's ability to bind four-way DNA Holliday junctions(Nomine et al, 2003).

The full length HPV16 E6 protein contains 14 cysteine residues. Only 8 are required for zinc ligation and the remaining 6 are not conserved. Nominé et al. used this information to their advantage and mutated the nonconserved cysteines to serines (known as E6 6C/6S) in hopes of minimizing the aggregation and precipitation problems of wild type E6(Nomine et al, 2003). They further showed that E6 6C/6S behaved indistinguishably from wild type E6 in both *in vitro* and *in vivo* p53 degradation assays(Nomine et al, 2003; Nomine et al, 2006).

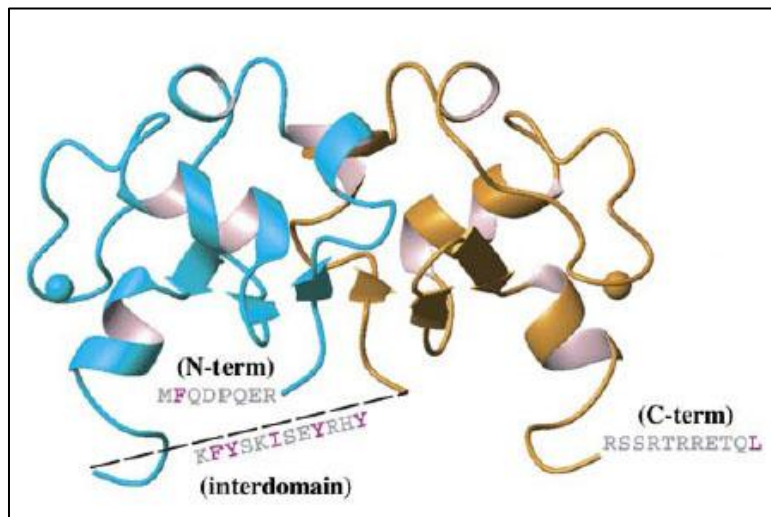
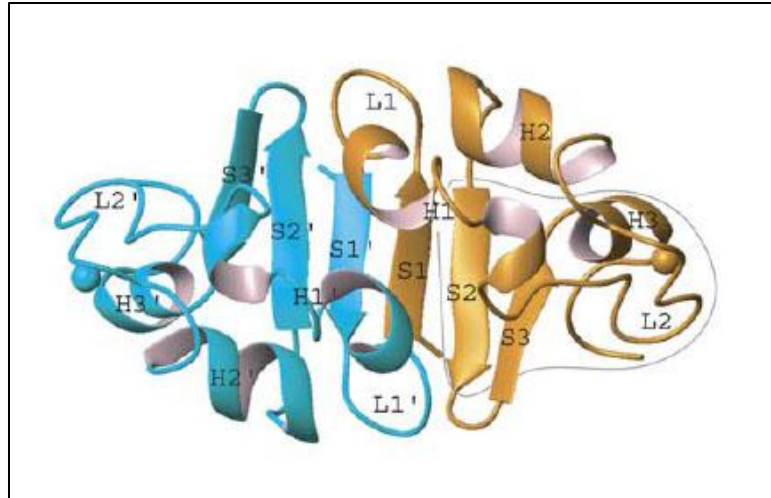
Travé and colleagues followed up their biochemical study with a nuclear magnetic resonance (NMR) structure of the C terminal domain(Nomine et al, 2006) (Figure 3.5). In support of the previous work, it was found to adopt a novel  $\alpha/\beta$  fold and ligate one  $Zn^{2+}$  ion. The structure, which is shown in Figure 3.5, has a three stranded anti-parallel  $\beta$  sheet with two of the three short  $\alpha$  helices packing on one side.  $Zn^{2+}$  ligation occurred via cysteines from H3 and the L2 loop. The domain has a main hydrophobic core that centers around Trp132, but also has a large hydrophobic patch including residues Tyr81, Tyr92, Leu99, and Asn127 that is solvent exposed(Nomine et al, 2006).

Given that these domains are probably related by a gene duplication event and were predicted to have similar secondary structure, Nominé et al modeled what the N terminal domain could look like by adding its residues to the  $C_{\alpha}$  backbone of the C terminal domain. Interestingly, several conserved hydrophobic residues were found to create a patch on one side of the N terminal model, leading the researchers to believe that these patches could be where the two domains contact each other(Nomine et al, 2006).

After exploring several possibilities for assembly, the final model minimized steric clashes, optimally fit secondary structure, and buried most of the hydrophobic patches. The full length model has a central six stranded anti-parallel  $\beta$  sheet with each  $Zn^{2+}$  ion at opposite ends of the protein(Nomine et al, 2006) (Figure 3.6).



**Figure 3.5: NMR structure of HPV16 C terminal domain.** The C terminal domain of HPV16 E6 is shown in cartoon representation. The three anti-parallel  $\beta$  sheets are labeled, as well as the three  $\alpha$  helices and two loops. The  $\text{Zn}^{2+}$  ion is shown as an orange sphere. This figure is adapted from (Nomine et al, 2006).



**Figure 3.6: Full length model of HPV16 E6.** The C terminal domain NMR structure is shown in orange and the N terminal domain model is shown in blue. All secondary structure in the modeled N terminal domain is labeled the same as the C terminal domain. The residues of missing N terminal, C terminal and linker regions are also listed with conserved residues colored in red. The top panel is a 90° rotation out of the page of the bottom panel. This figure was adapted from (Nomine et al, 2006).

Nominé et al's model accurately explains the reasons for several highly conserved residues within E6's sequence and even highlights some surface residues that could be necessary for E6AP recruitment(Nomine et al, 2006). Unfortunately, the model is an educated guess and lacks the loops of the protein, particularly the conserved linker region. The authors did attempt an NMR structure of the full length protein, but once the N terminal domain was added to the construct, it failed to behave well enough for NMR studies. Since a model cannot be probed in the same ways as a full length NMR or crystal structure could be, the hunt continues.

#### *3.1.4. Previous studies to inhibit E6 function*

E6AP is an E3 cellular ligase that is expressed ubiquitously in human cells. The full length protein is 875 amino acids, with the final 357 amino acids comprising the homologous to E6AP C terminus or HECT domain, which catalyzes the transfer of ubiquitin to target proteins. When E6 is not present, E6AP has no interaction or shows any propensity for targeting p53 for degradation(Be et al, 2001; Elston et al, 1998; Huang et al, 1999).

A short,  $\alpha$  helical peptide N terminal to the HECT domain (residues 403 – 421) is necessary and sufficient for E6 binding. A few other peptides have also been shown to bind E6 so a broad consensus sequence has been established: L h x  $\phi$  L s h, where L is leucine, h is an amino acid that can accept a hydrogen bond, x is any amino acid,  $\phi$  is a hydrophobic residue, and s is a small amino acid like glycine or alanine(Be et al, 2001).

When E6 is bound to E6AP, the complex binds p53's core domain for the transfer of ubiquitins(Cooper et al, 2003; Huibregste et al, 1993; Li & Coffino, 1996). Because the binding of



E6 to E6AP is a prerequisite to p53 degradation, many laboratories have focused on ways to block E6 binding to E6AP, which would subsequently block p53 degradation and bring about a return of p53 tumor suppressor function.

Early studies worked with RNAi(Niu et al, 2006) and anti-sense oligodeoxynucleotides (Marquez-Gutierrez et al, 2007b), which directly affected E6 levels within cell and animal models. Both showed an increase in apoptosis or a limit in tumor growth after introduction of the oligonucleotides(Marquez-Gutierrez et al, 2007a; Niu et al, 2006). Specific protein inhibitors were also developed, including intrabodies(Griffin et al, 2006) and alpha helical peptides(Butz et al, 2000; Liu et al, 2004), which were also able to show modest activity. Finally, ten small molecules inhibitors were identified by Baleja et al. after pharmacophore development and *in silico* screening(Baleja et al, 2006), but only one compound proved to be active in cells at high concentrations.

Historically, small molecules were developed to target enzymes, many of which already had small molecule substrates that could be mimicked(Arkin & Wells, 2004). Protein-protein interactions face many more challenges as drug targets due to the relative flatness and size of the interfaces, flexibility of the proteins upon binding, and stoichiometry of binding(Arkin & Wells, 2004). However, certain protein-protein interfaces seem to be amenable to small molecules, particularly those with small “hot spots” or involving binding to small peptides. The recent development of potent competitive antagonists towards the binding of BCL-X<sub>L</sub> to a 16 residue  $\alpha$  helix of BAK(Degeterev, 2001; Kutzki, 2002) shows that small molecules can be developed to disrupt protein-protein interactions. Similar to those proteins, E6 is a small, monomeric protein of 19 kDa(Lipari et al, 2001; Nomine et al, 2003) and the region of E6AP

necessary and sufficient for E6 binding is a 20 amino acid  $\alpha$  helix(Be et al, 2001; Cooper et al, 2003; Huibregste et al, 1993), making this protein-protein interaction a feasible target for inhibitor development.(Arkin & Wells, 2004)

With the recent release of the vaccines Gardasil (Merck) and Cervarix (GlaxoSmithKline), many feel that searching for more drug candidates to fight HPV infection is a moot point. However, the vaccine cost is still high and out of reach for some people in the United States, not to mention countries with less developed healthcare systems where rates of cervical cancer are still quite high. Also, vaccines offer little help to those already infected with HPV and showing signs of cervical intraepithelial neoplasia (CIN) or cancer. Considering that previous studies correlated cell cycle arrest and apoptosis with decreased of E6 activity, a small molecule inhibitor of E6 function could be a useful therapeutic tool and a crystal structure of E6 would a valuable asset to the field.

## 3.2. Materials and Methods

### 3.2.1. Growth and purification of E6M-GST, MBP-E6AP, and E6M

E6M-GST: Residues 1 – 158 of HPV16 E6 fused C terminally to glutathione-S-transferase were cloned into the pET28a vector (Invitrogen). Using the same idea as Nomine et al (Nomine et al, 2006), the nonconserved cysteine residues were mutated to the analogous residue in HPV 1A E6. DE3 *Escherichia coli* (*E. coli*) cells were transformed with the plasmid, grown in the presence of 50 µg/mL kanamycin, and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 75 µM Zn(OAc)<sub>2</sub> overnight at 18°C. Cell pellets were resuspended in 20 mM Hepes, pH 7.5, 500 mM NaCl, 10 mM β-mercaptoethanol (BME), 1 uM Zn(OAc)<sub>2</sub>, and 10 µg/mL phenylmethanesulfonyl fluoride (PMSF). Cells were lysed by sonication at 4°C and the insoluble fraction was separated by centrifugation at 18,000 rpm (SS-34, Sorvall) for 30 minutes. The soluble fraction was passed over pre-equilibrated glutathione resin (GE Healthcare), washed and protein was eluted with buffer containing 10 mM reduced glutathione (Sigma). Protein was aliquoted immediately and frozen on dry ice followed by long term storage at -80°C.

MBP-E6AP: Residues 363 to 440 of E6AP (isoform II) were cloned into the pMAL vector (NEB). DE3 *E. coli* cells were transformed with the plasmid, grown in the presence of 100 µg/mL ampicillin, and induced with 1 mM IPTG overnight at 18°C. Cell pellets were resuspended in 20 mM Hepes, pH 7.5, 200 mM NaCl, 10 mM BME, and 10 µg/mL PMSF. Cells were lysed by sonication at 4°C and the insoluble fraction was separated by centrifugation at 18,000 rpm (SS-34, Sorvall) for 30 minutes. The soluble fraction was passed over pre-equilibrated amylose resin

(NEB), washed, and protein was eluted with buffer containing 10 mM maltose (Fisher). Protein was concentrated to ~ 300  $\mu$ M, aliquoted and frozen on dry ice followed by long term storage at -80°C. A similar construct of residues 403 – 421 (isoform II) were cloned into the pMAL vector and purified in the exact same manner.

GST-E6AP: Residues 403 – 421 (isoform II) were cloned into the pGEX vector (GE Healthcare). DE3 *E. coli* cells were transformed with the plasmid, grown in the presence of 100  $\mu$ g/mL ampicillin, and induced with 1 mM IPTG overnight at 18°C. Cell pellets were resuspended in 20 mM Hepes, pH 7.5, 100 mM NaCl, 10 mM BME and 10  $\mu$ g/mL PMSF. Cells were lysed by sonication at 4°C and the insoluble fraction was separated by centrifugation at 18,000 rpm (SS-34, Sorvall) for 30 minutes. The soluble fraction was passed over pre-equilibrated glutathione resin (GE Healthcare), washed and protein was eluted with buffer containing 10 mM reduced glutathione (Sigma). Protein was concentrated to ~ 15 mg/mL, aliquoted and frozen on dry ice followed by long term storage at - 80°C.

E6M: Residues 1 to 158 of HPV16 E6 was cloned into the pET-DUET vector multiple cloning site (MCS) 1. Using the same idea as Nomine et al, the nonconserved cysteine residues were mutated to the analogous residue in HPV 1A E6. The full length UBE3, the yeast small ubiquitin-like modifier (SUMO) protein and Smt3 cleavage site had already been inserted into this vector after the 6x His tag but before the multiple cloning site. DE3 or 520 *E. coli* cells were transformed with the plasmid, grown in the presence of 100  $\mu$ g/mL ampicillin, and induced with 1 mM IPTG and 75  $\mu$ M Zn(OAc)<sub>2</sub> overnight at 18°C. Cell pellets were resuspended in in 20 mM

Hepes, pH 7.5, 500 mM NaCl, 10 mM BME, 1  $\mu$ M Zn(OAc)<sub>2</sub>, 5 mM imidazole, and 10  $\mu$ g/mL PMSF. Cells were lysed by sonication at 4°C and the insoluble fraction was separated by centrifugation at 18,000 rpm (SS-34, Sorvall) for 30 minutes. The soluble fraction was passed over pre-equilibrated Ni-NTA resin (Qiagen), washed with buffer plus 30 mM imidazole and the protein was eluted on an imidazole gradient of 30 mM to 500 mM. Cleanest fractions were pooled and dialyzed overnight at 4°C into 20 mM Hepes, pH 7.5, 500 mM NaCl, and 10 mM BME. Protein was recovered and 500  $\mu$ g of previously purified Smt3 was added per 50 mL of purified SUMO-E6M. The cleavage reaction was incubated for 1 hour on ice, then passed over pre-equilibrated Ni-NTA resin (Qiagen). Purified E6M was washed off the Ni-NTA resin with 20 mM Hepes, pH 7.5, 500 mM NaCl, 10 mM BME, and 20 mM imidazole. Protein was either used within 12 hours or aliquoted and frozen on dry ice followed by long term storage at -80°C.

### 3.2.2. *Pulldown of E6M and MBP-E6AP*

50  $\mu$ L of resin was pre-equilibrated with 20 mM Hepes, pH 7.5, 500 mM NaCl, 10 mM BME. 30  $\mu$ g of matching tagged protein was added with an equimolar amount of the partner protein, excess buffer, and incubated at 4°C for one hour on the nutator. Resin was washed with buffer and then ran out on an SDS-PAGE gel to determine extent of binding between the proteins. Negative controls were performed with tag alone plus the partner protein.

### 3.2.3. *E6M complex purification and crystallization*

E6M was purified as described above in large volume (~ 15 mL per 3 L *E. coli* growth) at ~ 0.2 mg/mL. An equimolar concentration of previously purified MBP-E6AP or DSHS00884 was added to the dilute E6M protein. The mixture was allowed to sit on ice for one hour then concentrated with a 10,000 MWCO filter (Amicon). For E6M + MBP-E6AP, the complex was then further purified on Superdex 200a column (GE Healthcare) and concentrated to 15 mg/mL as determined by UV-absorbance at 280 nm. Crystals were grown in 1.5 M NaCl and 10% v/v ethanol. E6M + DSHS 00884 was concentrated to 1.8 mg/mL, as determined by a Bradford Assay. Crystals were grown in several conditions: 1. 200 mM potassium tartrate tetrahydrate, 20% PEG 3350, 2. 200 mM sodium malonate, pH 7.0, 20% PEG 3350, 3. 200 mM ammonium citrate tribasic, pH 7.0, 20% PEG 3350, and 4. 200 mM citrate tribasi dehydrate, 20% PEG 3350.

### 3.2.4. *Enzyme-linked immunosorbance assay (ELISA)*

50  $\mu$ L of 30 nM E6M-GST (20 mM Hepes, pH 7.5, 500 mM NaCl, 10 mM BME) was added to each well in a 384-well, black, glutathione coated plate (NUNC). After one hour of incubation with shaking, all wells were washed with 20 mM Hepes, pH 7.5, 500 mM NaCl, 10 mM BME, and 0.1% TWEEN-20. 50  $\mu$ L of 6  $\mu$ M MBP-E6AP (1x PBS, 10 mM DTT, 0.1% TWEEN-20, buffer will be used for the rest of the assay) was then added to each well followed by 0.5  $\mu$ L of DMSO/compound. After one hour of incubation with shaking, all wells were washed with buffer. 50  $\mu$ L of 1:15,000 anti-MBP-HRP (NEB) was then added to each well and plates were shaken for one hour. Wells were washed one final time with buffer and 50  $\mu$ L of ELISA Pico

chemiluminescence substrate (Pierce) was added and wells were read with a plate reader (Wallac) after thirty minutes. Either 30 nM GST or 6  $\mu$ M MBP were used for negative controls.

### 3.2.5. *High throughput assay screening / data processing*

Three different diverse chemical libraries were screened: Spectrum (Microsource; 2000 compounds), HitFinder (Maybridge; 14,000 compounds), and Orthogonally Pooled Screening libraries (OPS, Lankenau Institute for Medical Research; 60,000 compounds), resulting in nearly 80,000 total compounds screened. The Spectrum library was screened with a single compound per well in duplicate, while the Maybridge and OPS libraries were screened in a compressed format of 5 compounds per well based on the Devlin et al. published method for screening and deconvolution of data (Devlin et al, 1996). Spectrum and HitFinder compounds were tested at final concentrations 10  $\mu$ M and the OPS library was screened at final concentrations of 6 – 10  $\mu$ M.

Each 384 well plate receiving compounds during screening was set up to have positive controls (E6-GST + MBP-E6AP + DMSO) in columns 2 and 23 and negative controls (GST + MBP-E6AP + DMSO) in columns 1 and 24 and compounds in columns 3 – 22. Uniformity plates (192 positive controls, 192 negative controls) were placed throughout the screening plates to ensure both assay and results reliability.

Percent E6-E6AP binding was calculated from raw luminescent values for each test compound (% inhibition = [average positive signal – test compound signal] / [average positive signal – average minimum signal] \* 100%). For the all screens, compounds that displayed three

standard deviation units (~ 50% inhibition) from the average of E6-E6AP binding in the presence of DMSO was classified as active. Software applications developed by CeuticalSoft (OpenHTS) were used to deconvolute the orthogonally compressed data for both HitFinder and OPS. The data was grouped into four categories: actives (compounds that displayed > 50% inhibition of E6-E6AP binding and cleaning mapped to a unique well in both the horizontal and vertical direction), ambiguous (compounds that mapped to two or more wells in either dimension), orphan (compound displayed inhibition in one direction but not the other) and inactive compounds.

#### 3.2.6. *Fluorescent thermal stability assay*

Individual 20 uL reactions were set up with 4 uM E6M, 1:1000 SYPRO orange (Invitrogen), 10 μM compound (final DMSO concentration of 1%) and buffer with a final concentration of 20 mM HEPES, pH 7.5, 500 mM NaCl, and 10 mM BME. Thermal melt curves were obtained by heating the protein from 20°C to 80°C and monitoring fluorescence at 580 nM using a 7900HT Fast Real Time PCR System (Applied Biosystems). Because some compounds were autofluorescent, negative control reactions of each compound alone were also performed to show that their fluorescent properties did not change over the temperature range of the experiment and ensure that fluorescence was directly representing E6M melting only ([data not shown](#)).



### 3.2.7. *In vitro p53 degradation assay*

In vitro translation: Full length, wild-type human p53 was cloned into the pRSETa vector (Invitrogen). Full length, wild type human papillomavirus type 16 E6 was cloned into the pET28a vector (EMD Biosciences). Using the TNT T7 coupled rabbit reticulocyte lysate systems (Promega), each protein was translated in separate reactions.

Degradation: Individual 10  $\mu$ L degradation reactions were set up with 2  $\mu$ L p53 translation reaction, 5  $\mu$ L E6 translation reaction, 0.1  $\mu$ L compound, and 2.9  $\mu$ L buffer in a final concentration of 20 mM Tris, pH 7.5, 100 mM NaCl, and 2 mM DTT. Positive control reactions included 0.1  $\mu$ L DMSO; negative control reactions included 0.1  $\mu$ L DMSO and 5  $\mu$ L of a blank translation reaction instead of the E6 translation reaction. All reactions were incubated at 30°C for 2.5 hours. Reaction were then mixed with SDS loading dye and heated at 90°C for five minutes.

Western Blotting: 3  $\mu$ L of each reaction was run out on a 15% SDS-PAGE gel, transferred to PVDF membrane and then probed with anti-p53 (1:5000, Bp53-12, Santa Cruz Biotechnology) followed by anti-mouse conjugated to HRP (1:5000, Bio-RAD). Bands were visualized by chemiluminescence (Pierce) and exposure to film (Kodak).

### 3.2.8. *Cell maintenance*

C33A, SiHa, and HeLa cells were grown in 1x minimal eagle's media (MEM, Cellgro) supplemented with 10% fetal bovine serum (Hyclone), 10 µg/mL penicillin-streptomycin (Cellgro), 2 mM L-glutamine (cellgro), 1 mM sodium pyruvate (Cellgro), and 100 µM non-essential amino acids (Gibco). All experiments were carried out with cells passaged 30 times or less. C33A and SiHa cell lines were ordered from ATCC, while the HeLa cell line was a gift from laboratory of Susan Janicki, Ph.D.

### 3.2.9. *Cell assays: toxicity, p53 levels, APO-one*

Toxicity Assay: 40 µL of 500 cells/µL C33A cells or 50 cells/µL SiHa/HeLa cells (harvested in media containing 20% FBS) were plated on Day 1 in each well of a 384 well, clear, tissue culture plate (NUNC). On Day 2, 40 µL of DMSO/compound in media containing no FBS was added to each well and cells were incubated for 48 hours. On Day 4, 8 µL of MTS reagent (Promega) was added to each well and the absorbance at 490 nm was read (Wallac) after 5 hours.

p53 Levels: On Day 1, in a 6 well, clear, tissue culture plate, 400,000 SiHa or HeLa cells were plated with 2 mL of media. On Day 2, 5 µL of DMSO/compound was added to 500 µL media, mixed, and all 500 µL were added to each well. Cells were incubated for 48 hours and on Day 4, cells were washed with 2 mL of cold, sterile 1x PBS, then harvested with 600 µL of RIPA buffer with complete protease inhibitor cocktail (Roche). Protein sample concentrations were normalized with the bicinchoninic acid (BCA) protein assay (Pierce). 28 µL of each sample was

run on a 12% SDS-page gel and transferred to a PVDF membrane. All membranes were probed with three primary antibodies: p53 (1:1000 anti-p53 (Santa Cruz Biotechnologies)), cleaved PARP (1:1000 anti-cleaved PARP (Abcam)), and beta actin (1:5000 anti-beta actin (Abcam)). All bands were visualized by chemiluminescence (Pierce) and exposure to film (Kodak).

APO-One: 25  $\mu$ L of 80 cells/ $\mu$ L SiHa/HeLa cells were plated on Day 1 in each well of a 384 well, clear, tissue culture plate (NUNC). On Day 2, 0.5  $\mu$ L of DMSO/compound was added to each well and cells were incubated for 48 hours. On Day 4, 25  $\mu$ L of APO-one reagent (Promega) was added to each well and, following excitation at 485 nm, the fluorescence at 521 nm was read (Wallac) after incubation at room temperature for 8 hours.

### 3.3. High Throughput Screening Results and Discussion

#### 3.3.1. *Development of High Throughput Screen and Inhibitor Identification*

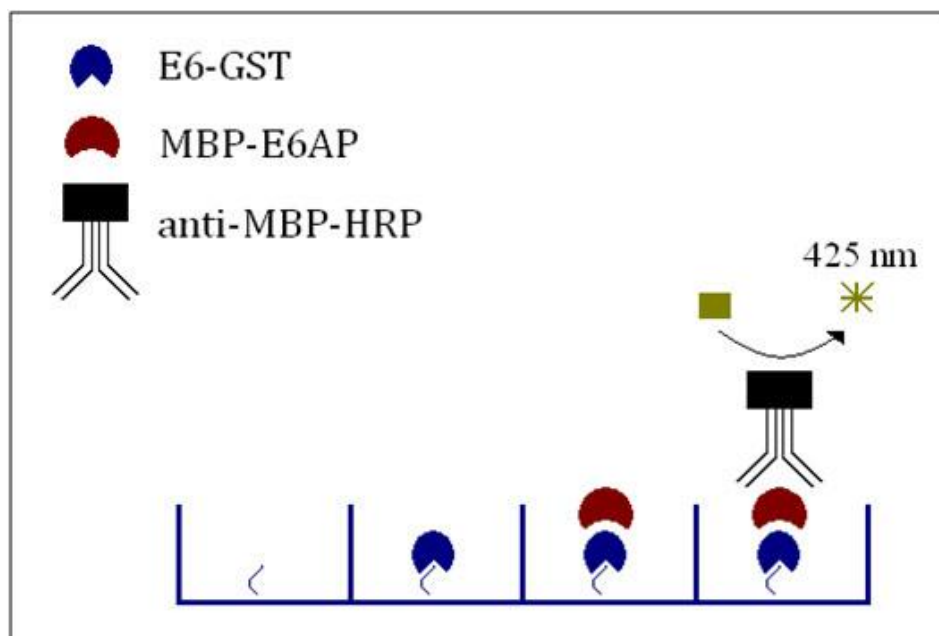
High throughput screening (HTS) for small molecule inhibitors of the E6 / E6AP interaction was designed around a modified sandwich enzyme-linked immunosorbance assay (ELISA). Full length HPV16 E6 was fused C terminally to glutathione-S-transferase (GST). MBP-E6AP (residues 363 – 440, isoform II) was added to the wells at a concentration of 6  $\mu\text{M}$  (the reported  $K_d$  of the protein-protein interaction(Zanier et al, 2005)), in the presence of a small molecule dissolved in DMSO or a DMSO control. The extent of binding between the two proteins was assessed by incubation with the anti-MBP-HRP antibody (New England BioLabs) and chemiluminescence (Pierce) (Figure 3.7).

The robustness of the assay was assessed using the Z factor parameter(Zhang et al, 1999). Using a 96 well format, we optimized conditions to obtain a Z factor of 0.784 (E6M-GST/GST + MBP-E6AP) and 0.624 (E6M-GST + MBP-E6AP/MBP). Miniaturization to a 384 well format and automation optimization resulted in Z factors consistently above 0.60 (Figure 3.8).

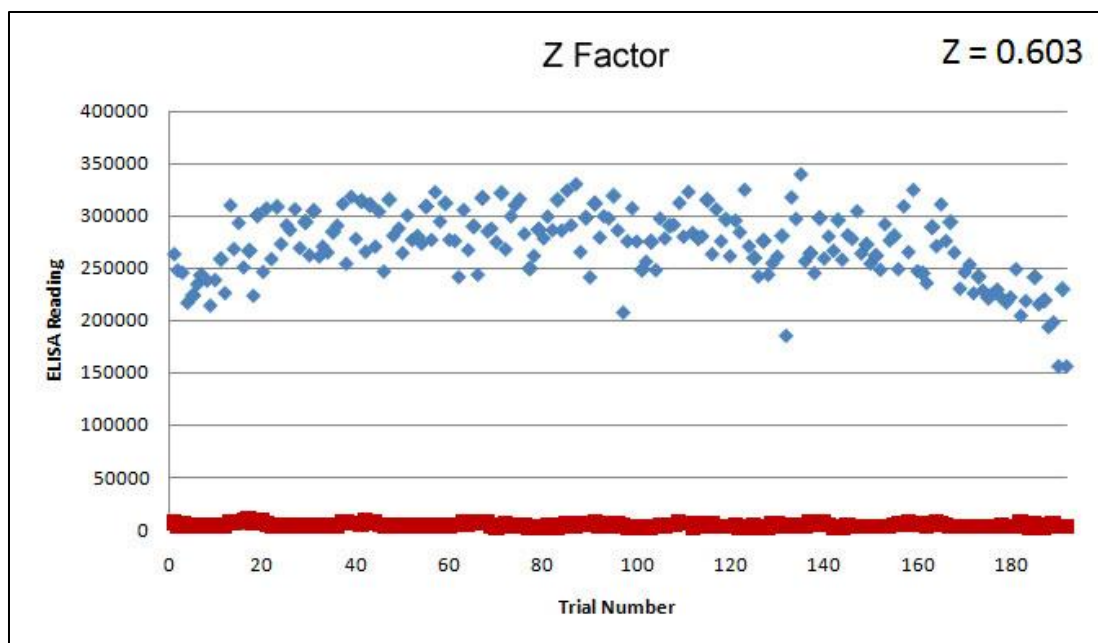
Three different diverse chemical libraries were screened: Spectrum (Microsource; 2000 compounds), HitFinder (Maybridge; 14,000 compounds), and Orthogonally Pooled Screening libraries (OPS, Lankenau Institute for Medical Research; 60,000 compounds), resulting in nearly 80,000 total compounds. The Spectrum library was screened with a single compound per well in duplicate, while the Maybridge and OPS libraries were screened in a compressed format of 5 compounds per well based on the Davlin et al. published method for screening and deconvolution of data(Devlin et al, 1996). The initial screening resulted in 201 potential

inhibitors of the E6 / E6AP interaction that were further cherry picked for follow-up studies. This included orphan and ambiguous hits that came out of the use of compressed libraries.

All potential hits, including orphans and ambiguous compounds, were cherry picked and retested at the screening concentration in the ELISA assay. Following retesting, the number of potential inhibitors dropped to 54, producing a hit rate from the screen of 0.070% (Table 3.1).



**Figure 3.7: Scheme for the high throughput screening assay.** The ELISA assay is represented here schematically. DMSO or compound dissolved in DMSO was added with the MBP-E6AP.



**Figure 3.8: Z factor graph.** A representative Z factor graph is shown for a 384 well plate: 192 positive control (E6M-GST + MBP-E6AP6) wells, 192 negative control (GST + MBP-E6AP6) wells. The Z factor was calculated according to Zhang et al, 1999.

Library Name	Library Size	Hits	Confirmed Cherry Picks	Confirmed Compounds
Spectrum	2000 compounds	7	4	4
Maybridge HitFinder	14,400 compounds	56*	13	12
OPS	60,000 compounds	138*	37	14
TOTALS	76,400 compounds	201	54	30

**Table 3.1: Screening summary.** Names of the libraries screened, their size, and the number of hits from each screen are listed. Hits: Number of potential inhibitor compounds from initial screening. Asterisks show which libraries included ambiguous and orphan hits to the original hit number. Confirmed cherry picks: Number of potential inhibitors that again showed inhibition in the ELISA assay at the screening concentration. Confirmed compounds: Compounds that, following purchase of powders, passed all secondary assays for inhibition.



### 3.3.2. *Characterization of HPV-E6 Inhibitors in vitro*

We ranked the potency of the inhibitors by determining the inhibitor concentration that reduced the extent of binding between E6M-GST and MBP-E6AP by 50% (IC<sub>50</sub> value). For each of the inhibitors tested, fresh powders were ordered directly from commercial suppliers and 42 of the 54 compounds had IC<sub>50</sub> values of 30 μM or less, with more than half of the inhibitors having IC<sub>50</sub> values in the low to mid-nanomolar. [Table 3.2](#) reviews the IC<sub>50</sub>s and structures for the fully confirmed 30 inhibitors.

To assay if compounds were nonspecifically inhibiting binding between E6M-GST and MBP-E6AP by precipitating E6M-GST, a fluorescent thermal stability assay was employed. This assay monitors protein unfolding in the presence of a given inhibitor where inhibitor binding is generally correlated with a change in the protein melting temperature curve. Full length, untagged E6M was incubated at increasing temperature (20°C - 80°C) with either DMSO control or with compound. The reporter dye, SYPRO orange (Invitrogen), which binds to hydrophobic regions of proteins that are generally exposed upon protein unfolding, was also added.

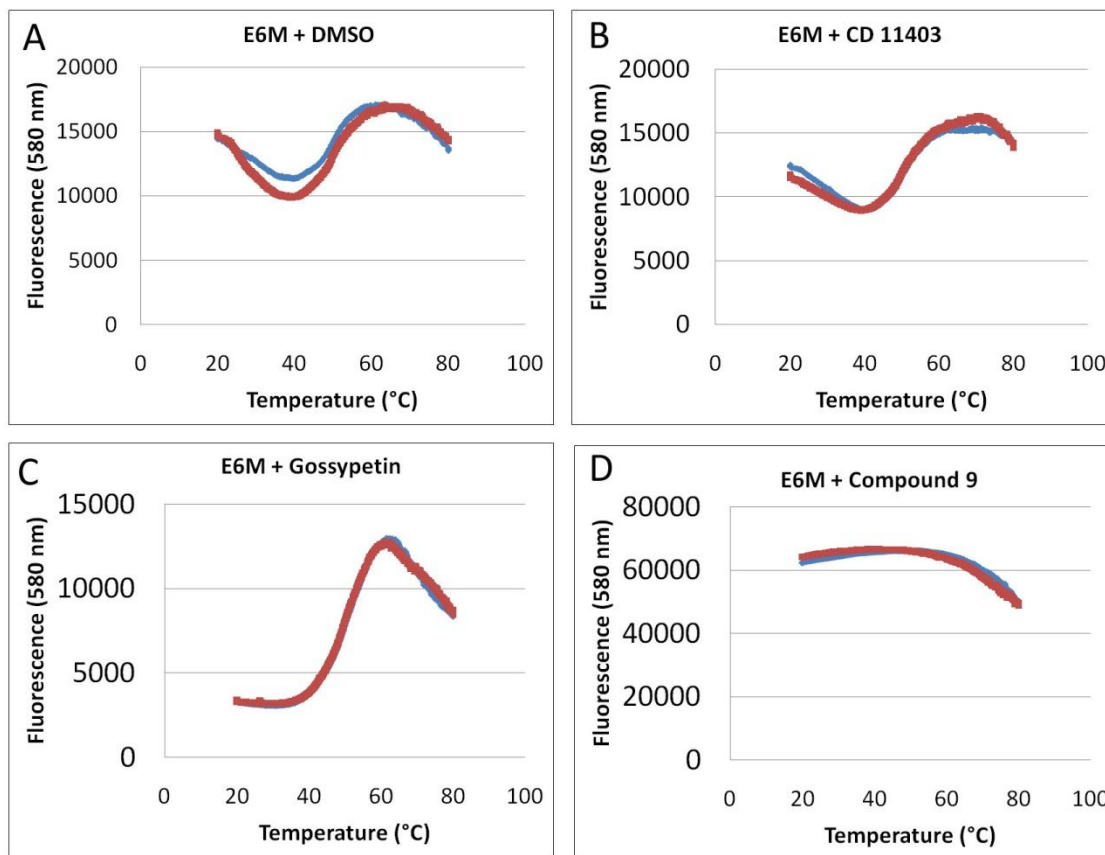
The control reaction with DMSO showed that E6M was unstable from 20°C to approximately 37°C, after which the protein melted in a standard sigmoidal fashion ([Figure 3.9, Table 3.2](#)). Overall, three different types of curves resulted from this assay with added compound. Fourteen compounds resulted in E6M melting curves that were indistinguishable from the protein melting curves without added compound. ([Figure 3.9, Table 3.2](#)). Unchanged melting curves could indicate that compounds were binding E6M but not significantly stabilizing its structure or that the compound mediated its function in other ways, most likely by binding E6AP. Twenty three compounds showed significant stabilization of E6M in the region of 20°C to

37°C, which suggested that these compounds were binding E6M (Figure 3.9, Table 3.2). Finally, E6M in the presence of five compounds had an essentially constant fluorescence reading over the temperature range. The flattened fluorescence curve was concluded to mean that the compound caused E6M to precipitate out of solution (Figure 3.9, Table 3.2). This conclusion was supported by a simple *in vitro* test of mixing purified E6M with an equimolar concentration of all five compounds individually. A white precipitate formed each time, which was pelleted by centrifugation and shown to be E6M by SDS-PAGE (data not shown). These five compounds were eliminated from the pool of potential inhibitors.

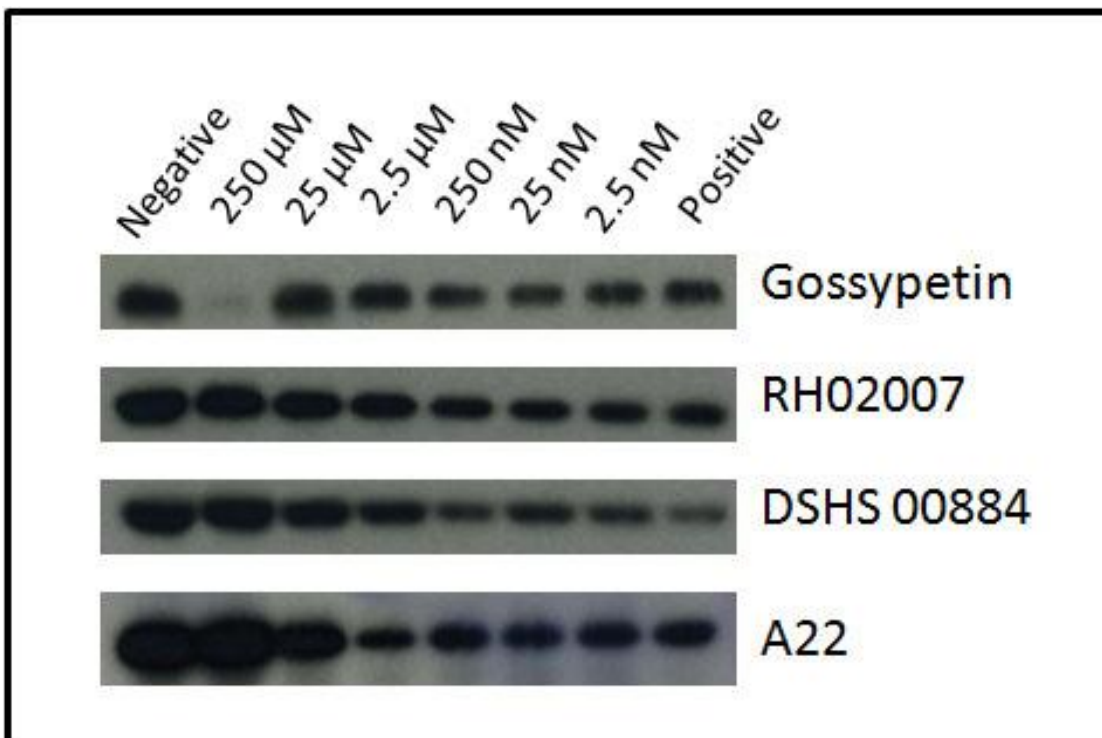
To establish if the E6 inhibitors could function to inhibit the ability of E6 to mediate p53 degradation, we assayed the remaining 37 compounds that did not appear to cause E6M precipitation in a p53 degradation assay. We first used a cell free system to factor out the ability of compounds to enter cells. We anticipated that this assay would also help eliminate compounds that were false positives in the ELISA assay, such as those compounds that inhibited GST/glutathione binding or HRP function, and possibilities that could not be easily tested the fluorescence assay, such as MBP or E6AP precipitation. To carry out these studies, full length, wild type HPV16 E6 and full length, wild type human p53 were translated in separate reactions using the TNT T7 Coupling Reticulocyte System (Promega). Equal amounts of each reaction were mixed together with compound and excess lysate, which contains E6AP and the machinery for ubiquitin-mediated degradation. The amount of p53 present in each sample was determined by a Western blot against the N terminus of human p53 (Santa Cruz Biotechnology).

These studies revealed that 7 compounds showed no effect on p53 degradation in the cell free system. Of the 30 compounds that did modulate p53 degradation, approximate IC<sub>50</sub>s

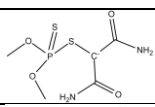
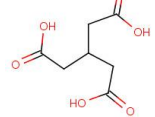
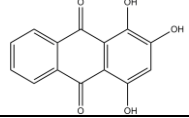
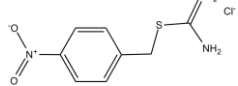
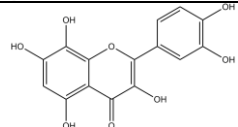
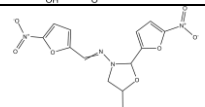
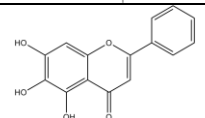
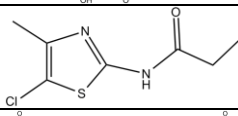
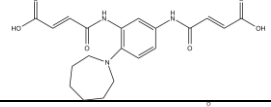
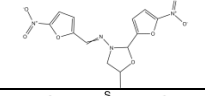
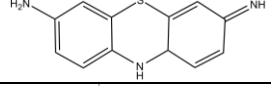
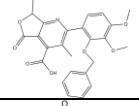
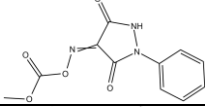
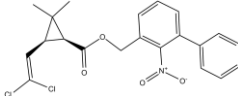
ranged from nanomolar to high micromolar (Table 3.2, Figure 3.10) with the  $IC_{50}$  value representing the concentration of compound needed to bring degradation of p53 down 50%. Together, these studies show that we have identified 30 HPV-E6 inhibitors that are able to block the ability of E6 to mediate its oncogenic property of p53 degradation in a cell free system.

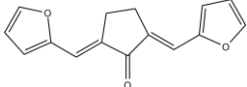
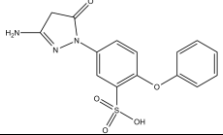
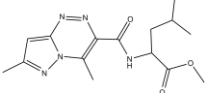
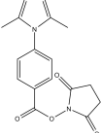
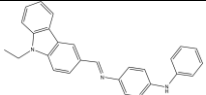
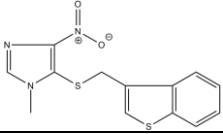
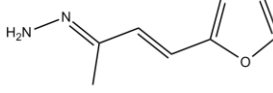
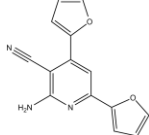
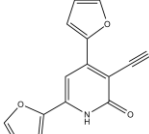
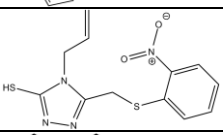
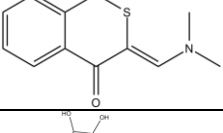
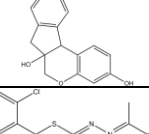
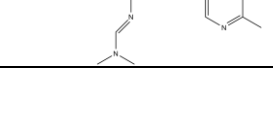


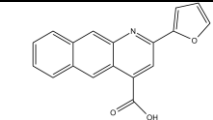
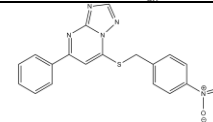
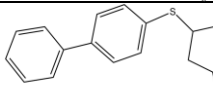
**Figure 3.9: Effect of inhibitors on E6 thermal stability.** Replicate data (blue and red) for fluorescence at 580 nm vs temperature for each compound is graphed. (A) E6M in 20 mM HEPES, pH 7.5, 500 mM NaCl, and 10 mM BME and 1% DMSO. Inherent protein instability is seen in the region of 20°C to 37°C. (B) E6M with CD11403, representing a compound that does not significantly stabilize E6M and results in a curve with a similar shape as E6M with DMSO alone. (C) E6M with gossypetin, representing the best compound to stabilize E6M. The results show a curve that no longer has the region of instability from 20°C to 37°C. (D) E6M with inhibitor Compound 9, a previously identified inhibitor from Baleja et al (Baleja et al, 2006), which results in E6M precipitation and the flattened curve.



**Figure 3.10: p53 degradation assay results.** Westerns of p53 degradation reactions treated with increasing concentrations of compound. Positive controls were E6 + p53 + DMSO while negative controls were p53 + blank reaction + DMSO.

Name	ELISA IC <sub>50</sub>	p53 Degradation ~IC <sub>50</sub>	Thermal Stability	Structure
C14	30.6 ± 4.7 nM	10 μM	Stabilized	
C17	86.1 ± 37.3 nM	10 – 100 μM	Unchanged	
Purpurin	143 ± 78 nM	250 nM	Stabilized*	
E17	148 ± 42 nM	10 – 100 nM	Stabilized	
Gossypetin	170 ± 21 nM	250 nM	Stabilized	
B15	178 ± 122 nM	10 – 100 μM	Unchanged	
Baicalein	235 ± 74 nM	250 nM	Stabilized*	
A19	295 ± 38 nM	10 – 100 μM	Stabilized	
S05659	311 ± 1.6 nM	250 nM – 2.5 μM	Unchanged	
D5	347 ± 218 nM	10 μM	Stabilized	
C6	363 ± 24 nM	1 – 10 μM	Unchanged	
A14	457 ± 14 nM	100 μM	Unchanged	
A22	477 ± 166 nM	1 – 10 μM	Stabilized*	
C9	496 ± 49 nM	10 – 100 μM	Stabilized	

DP00966	573 ± 323 nM	250 μM	Unchanged	
JFD00458	703 ± 54 nM	250 nM	Stabilized*	
B11	749 ± 310 nM	2.5 μM	Unchanged	
RH02007	1.41 ± 0.05 μM	2.5 – 25 μM	Stabilized*	
C22	1.79 ± 0.24 μM	100 μM	Stabilized	
HTS03324	2.27 ± 1.04 μM	2.5 nM	Stabilized*	
C16	3.98 ± 3.88 μM	2.5 – 25 μM	Unchanged	
HTS13545	4.11 ± 0.21 μM	2.5 – 25 μM	Stabilized*	
HTS10308	5.24 ± 0.03 μM	25 – 60 μM	Unchanged	
DSHS00884	~ 10 μM	250 nM – 2.5 μM	Stabilized	
CD11403	10.9 ± 0.6 μM	25 – 250 μM	Unchanged	
Brazilin	11.2 ± 4.2 μM	25 – 250 nM	Stabilized	
BTB 07267	15.1 ± 1.8 μM	250 nM – 2.5 μM	Stabilized*	

S05896	15.7 ± 1.2 μM	25 – 250 μM	Unchanged	
HTS11967	30.3 ± 10.9 μM	250 nM – 2.5 μM	Stabilized*	
D22	36.9 ± 10.3 μM	250 nM	Unchanged	

**Table 3.2: Secondary assay results summary for inhibitor compounds.** The names, structures and secondary assay results for all thirty identified inhibitor compounds are listed in rank order of ELISA IC<sub>50</sub> value. ELISA IC<sub>50</sub> values were determined by performing the assay in duplicate for each compound, normalization of data and subsequent data fitting in GraphPad with log(inhibitor) vs. response (variable slope) non-linear regression fit. p53 degradation approximate IC<sub>50</sub> is listed as a range. The thermal stability result is as described in text. Asterisks denote compounds that showed E6 stability, but to a lesser degree.



### 3.3.3. *Characterization of HPV-E6 Inhibitors in cells*

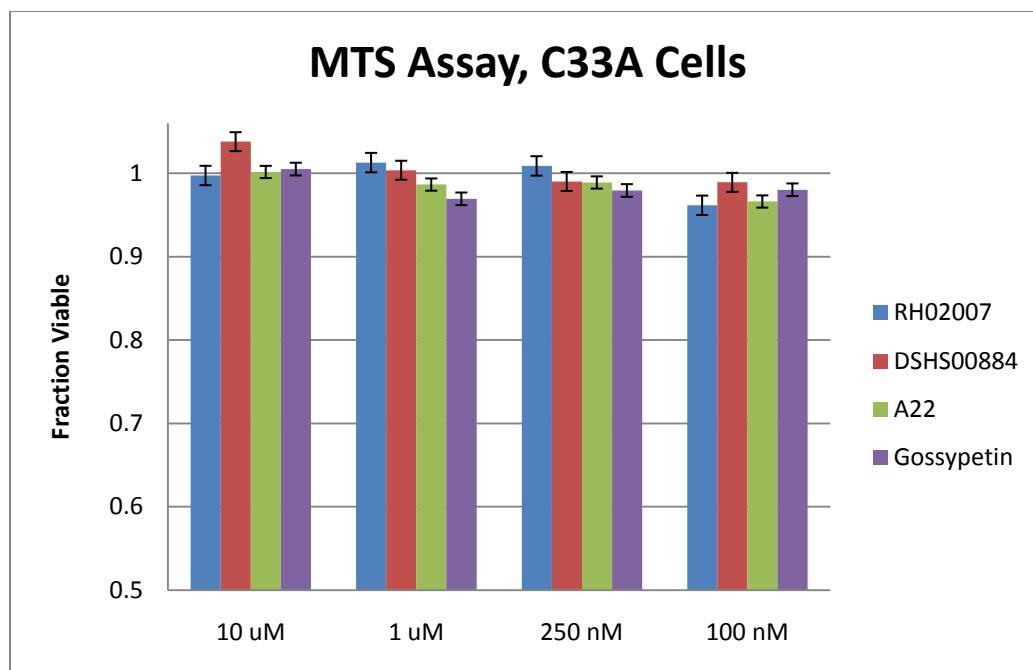
Through our high throughput screening assay and in vitro follow up assays, we identified 30 compounds that inhibit E6-E6AP binding in an isolated system (ELISA) and a complex cell lysate (p53 degradation assay) so we next addressed whether these compounds could work in a cellular environment. To do this, each compound was first tested in an MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) assay (Promega) to assess compound toxicity at a concentration range of 25  $\mu$ M to 100 nM. The control cell line used for this assay was C33A (ATCC), which is a tumor-derived, cervical cancer cell line that is not infected with HPV. 48 hours after the addition of compound, MTS assay reagent was added and the absorbance reading at 490 nm was directly related to the number of viable cells in the culture. Some compounds were toxic in the 10 – 25  $\mu$ M range, but none showed any sensitivity at concentrations below 10  $\mu$ M (Figure 3.11, remaining data not shown).

Each compound was then tested for its effectiveness at inhibiting p53 degradation in cells. Two different HPV-infected, tumor derived cell lines were used: SiHa (HPV16 integrated, ATCC) and HeLa (HPV18 integrated, ATCC). 48 hours after the addition of compound, cells were lysed and the level of p53 in each sample was assessed by Western blot using an anti-p53 antibody (Santa Cruz Biotechnologies). 4 compounds (DSHS 00884, RH02007, A22 and gossypetin) showed an increase in p53 levels in both cell lines over this concentration range while E17 and C22 protected against p53 degradation specifically in SiHa cells and C14 protected against p53 degradation in HeLa cells (Figure 3.12, Figure 3.13).

We were curious to determine if an increase in p53 levels could be correlated with apoptosis in the HPV-infected cells. This question was approached with the APO-One assay

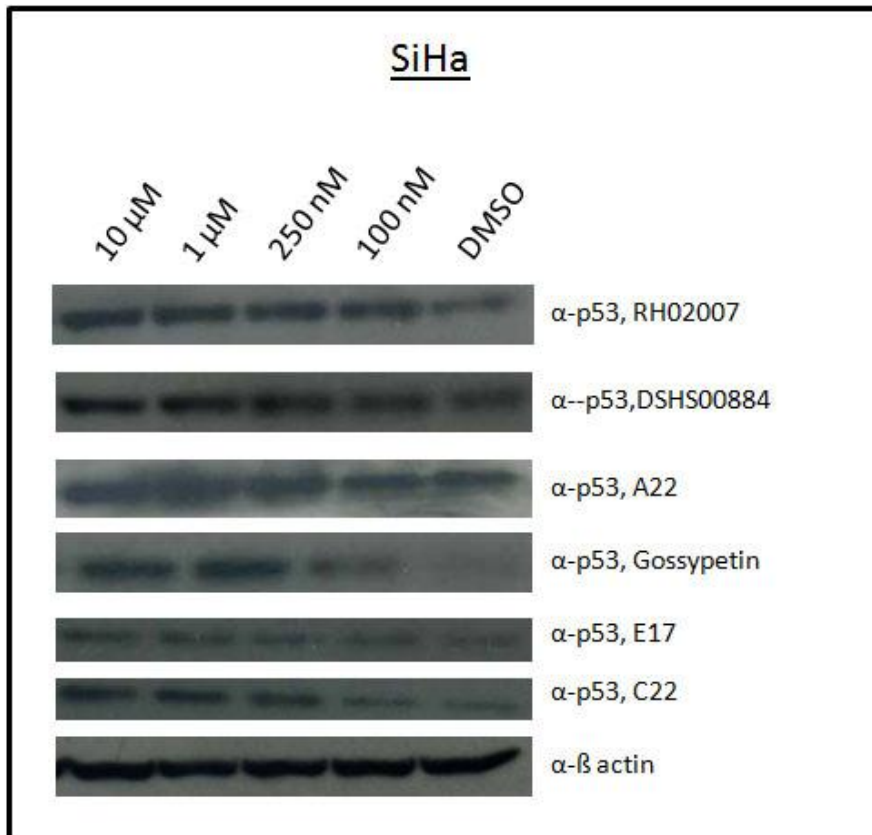
(Promega), which was carried out very similarly in HeLa cells to the MTS assay, except 48 hours after compound addition, a peptide bearing a rhodamine110 fluorescent tag was added to the wells. The peptide is a substrate for caspases 3 and 7 and, upon cleavage, releases the rhodamine110 that when excited at 499 nm results in an increase in fluorescence at 521 nm.

Based on this assays, two compounds showed an increase in apoptosis in both SiHa and HeLa cell lines while having minimal toxicity to C33A cells: DSHS 00884 and RH 02007 (Figure 3.11, Figure 3.14, Appendix). Both compounds fit the Lipinski guidelines for druglikeness(Lipinski et al, 1997). According to the APO-One assay, caspase activity was increased nearly three-fold in cells treated with 10  $\mu$ M or 250 nM DSHS 00884 over cells treated with DMSO alone (Figure 3.14). Similar to DSHS00884, the APO-one assay result for RH 02007 show that for cells treated with 10  $\mu$ M RH 02007, a 2 fold increase in caspase 3 and 7 activities were seen (Figure 3.14). In addition, cleaved Poly (ADP-ribose) polymerase (PARP), which is cleaved by caspase-3 during apoptosis, is also present in both cell lines treated with 10  $\mu$ M RH02007(Ellis et al, 1991; Nicolettie & Stella, 2003; Trucco & al., 1998). Given that the APO-one assay is very sensitive and PARP cleavage occurs late in the pathway, it is possible that the samples lacking cleaved PARP are either in the early stages of apoptosis or not enough cleaved PARP has accumulated that can be seen in a western blot. Taken together, these data support that both of these compounds bring about an increase in p53 levels, which is correlated with an increase in apoptosis in HPV-positive cells.



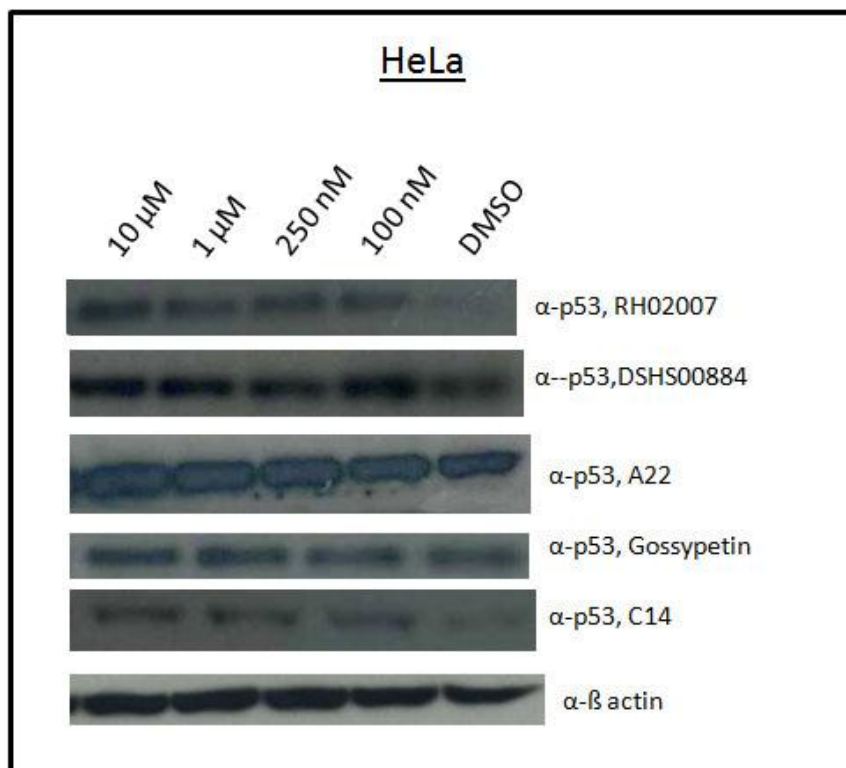
**Figure 3.11: Effect of E6 inhibitors on cell toxicity.** Normalized MTS assay data for C33A.

Absorbance values were normalized against wells containing C33A + DMSO.



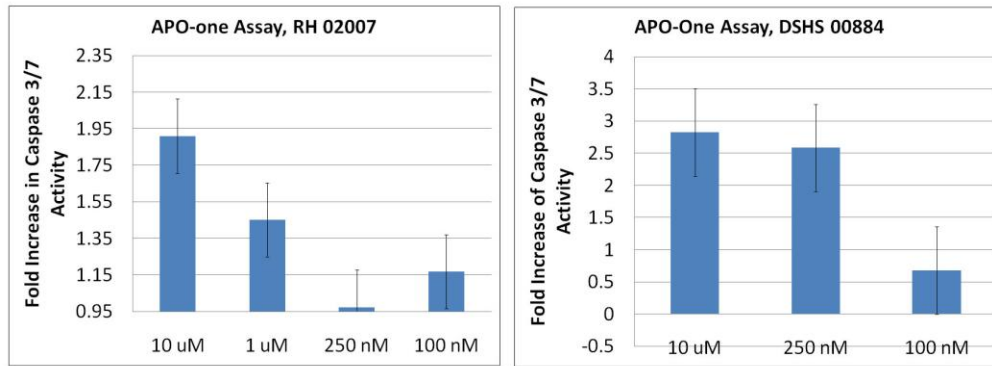
**Figure 3.12: Effect of E6 inhibitors on p53 degradation in SiHa (HPV16) infected cells.**

Westerns of cell samples treated with increasing concentrations of compound.



**Figure 3.13: Effect of E6 inhibitors on p53 degradation in HeLa (HPV18) infected cells.**

Westerns of cell samples treated with increasing concentrations of compound.



**Figure 3.14: Effect of E6 inhibitors on apoptosis in HeLa (HPV18) infected cells.** Fluorescence data was normalized against HeLa cells treated with DMSO.

#### 3.3.4. Discussion

The purpose of this high throughput screen was to identify inhibitors of the E6-E6AP interaction with the downstream effect being protection against p53 degradation in HPV-infected systems. We identified 30 compounds that are able to block p53 degradation in a cell-free system by a specific mechanism that does not involve E6 precipitation or degradation. The extension of these compounds into HeLa and SiHa cells relies on the compound's ability to get inside both cell types and to not be degraded over the time course of the experiment. Four compounds were able to show protection against p53 degradation in both cell lines while three compounds were effective in one cell line and not the other. It is conceivable that some of the remaining 23 compounds could be modified to more easily cross the cellular membrane and/or be made less resistant to degradation while still maintaining their E6/E6AP inhibitory properties.

Of the compounds that were able to penetrate cells and mediate E6 inhibition, two induced apoptosis in the experimental 48 hour time period. An increase in p53 levels can lead to any number of downstream effects so these compounds could be inducing cell cycle arrest or induction of DNA-damage repair genes instead. A22, gossypetin, E17, C22 and C14, which all display inhibition of p53 degradation in both cell lines, shows no signs of apoptosis after 48 hours. Alternatively, these compounds may still be able to induce apoptosis, but not in the time frame of the experiment. A recent study by DeFilippis et al looked at the HeLa cell response to the repression of the E6 gene (DeFilippis et al, 2003). Apoptosis was not detected until six days post repression of the E6 gene. It is conceivable that if these compounds were incubated to these time points, apoptosis would be detected.

The study reported here does provide some SAR information although it is quite limited at this point. For example, gossypetin, brazilin and baicalein all fall within the same scaffold, particularly gossypetin and baicalein, which have similar results in all secondary assays. Another related pair of compounds is HTS13545 and HTS10308, which differ only at one function group: HTS13545 bears a an amine group where HTS10308 has a double bonded oxygen. Both have very similar  $IC_{50}$ s, but HTS13545 stabilized E6M in the thermostability assay while HTS 10308 does not. Perhaps the amine group of HTS13545 donates two hydrogen bonds to its protein targeting leading to greater stabilization of the complex.

While E6 binds p53, the other oncogenic protein E7 targets pRB to prematurely release E2F and bring about pRB degradation. Both pathways are purposely inhibited so that viruses can be replicated. Previous studies show that continual expression of these two genes is necessary to maintain malignancy while suppression of both genes results in cellular death(Alvarez-Salas & diPaolo, 2007). While the studies reported here provide the first potent HPV-E6 inhibitors that might be further developed into therapeutic agents, other small molecules that can block pRB binding to E7 might take advantage of the synergistic effects of both compounds in a therapeutic setting to block HPV transforming effects and strongly bring about apoptosis in HPV-infected cells.

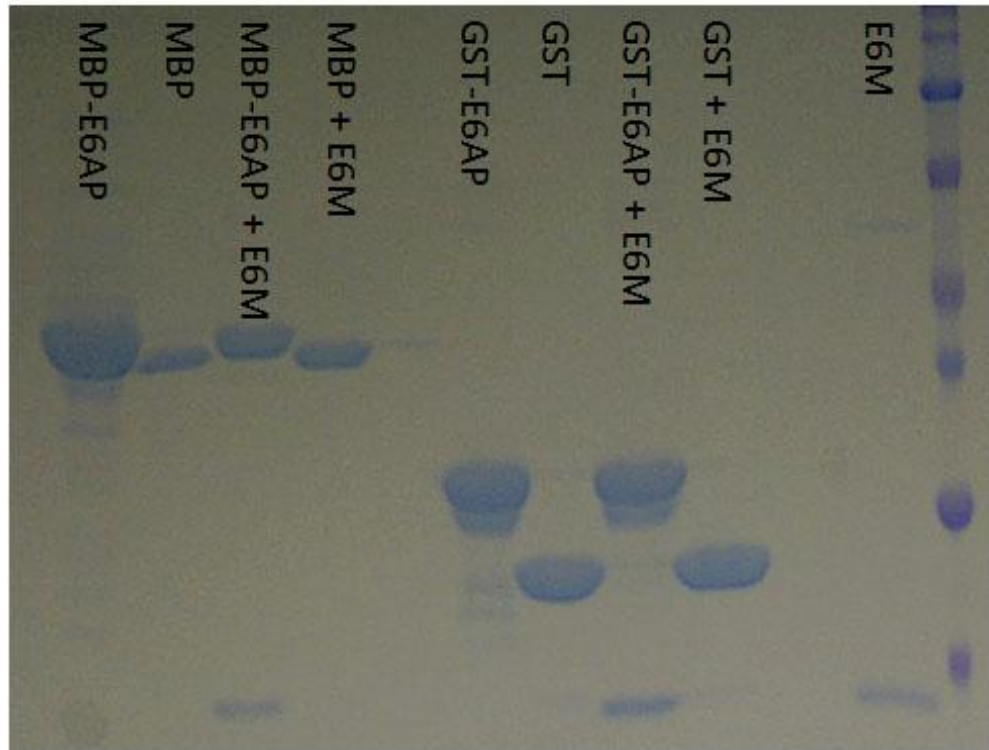


### 3.4. Crystallization Results and Discussion

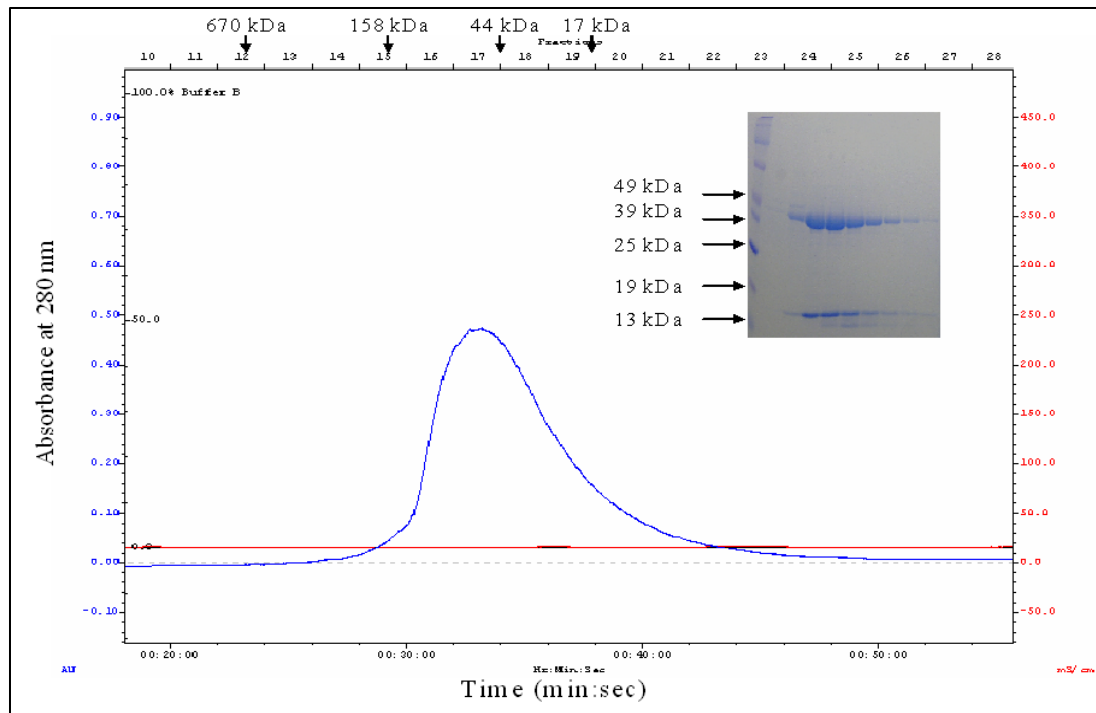
#### 3.4.1. E6M + MBP-E6AP

As already discussed, a crystal structure of E6 has been long sought and fraught with problems. Mainly, the protein has a strong propensity towards aggregation and precipitation. After several failed schemes, we were able to purify behaved E6M protein in large quantities but low concentrations after Ni-NTA purification. However, this purified protein faced two problems: 1. when run over the superdex s200a column (GE Healthcare) for size-exclusion chromatography, > 90% of the protein bound the matrix and failed to elute, 2. the protein would not concentrate above 0.2 mg/mL.

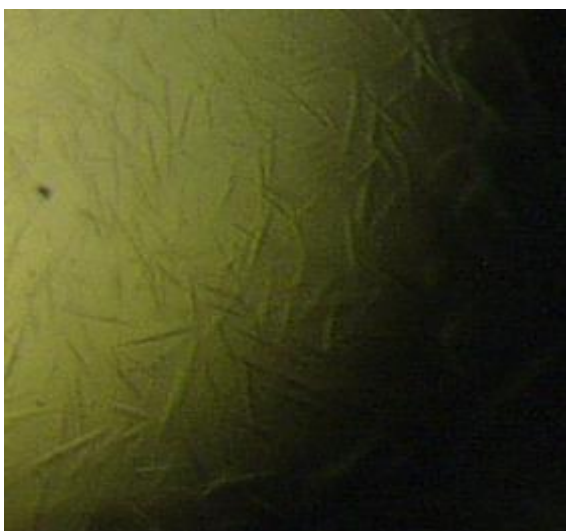
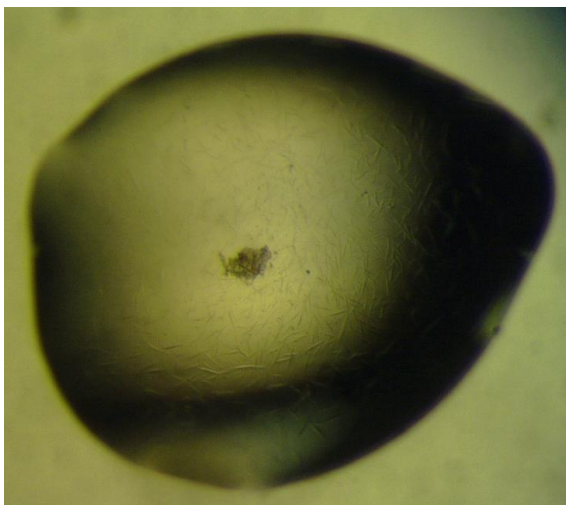
Luckily, the ~ 10% of protein that did elute from the gel filtration column was monomeric, which supports previous conclusions. Also, purified E6M specifically pulled down the peptide of E6AP necessary and sufficient for binding (Figure 3.15). Since E6M alone interacted with the gel filtration matrix and failed to concentrate, we considered the possibility that it may behave better when in the presence of a partner protein. To this end, we purified E6M alone, mixed an equimolar amount of MBP-E6AP, concentrated the complex and performed size-exclusion chromatography. One peak was seen on the gel filtration chromatogram at a size just larger than 44 kDa and an SDS-PAGE gel showed that both MBP-E6AP and E6M (complex 64.3 kDa) eluted together (Figure 3.16). These fractions were then pooled, concentrated and used for crystallization trials. One condition yielded crystals that were difficult to reproduce (Figure 3.17).



**Figure 3.15: Pulldown of E6AP by E6M.** Pulldowns were performed as described in the text. E6M specifically pulled down both MBP-E6AP (403 – 421, isoform II) and GST-E6AP (403 – 421, isoform II) but not MBP or GST alone.



**Figure 3.16: Purification of E6M + MBP-E6AP6.** Representative superdex 200a gel filtration chromatogram for E6M + MBP-E6AP purification is shown. The protein peaks just before the 44 kDa standard, which fits for a complex of 64.3 kDa. The 15% SDS-PAGE gel insert shows that the six fractions of the peak contain both proteins and that they elute together.



**Figure 3.17: Crystals of the E6M + MBP-E6AP complex.** These crystals were obtained with 15 mg/mL complex in 1.5 M NaCl and 10% v/v ethanol. The tiny needles appeared after approximately one week.

### 3.4.2. E6M + DSHS 00884

Encouraged by the success of binding E6 with MBP-E6AP for crystallization, we scoured the literature looking for other reported partner proteins. Included in our studies were the CH3 and CH1 domains of p300, an E6 intrabody, and both N terminal and C terminal fusions of E6M with the E6AP peptide (known as N3 and C3). We also tried three ternary complexes: E6M, MBP-E6AP + p53DBD, N3 + p53DBD and C3 +p53DBD. Unfortunately, none of these complexes resulted in crystals.

The high throughput assay was designed to search for small molecules that disrupted the E6 / E6AP interaction with the hopes that several of these compounds would mediate their function by binding to E6. Results of the thermostability assay show that some compounds are stabilizing E6 and are good choices for crystallization studies.

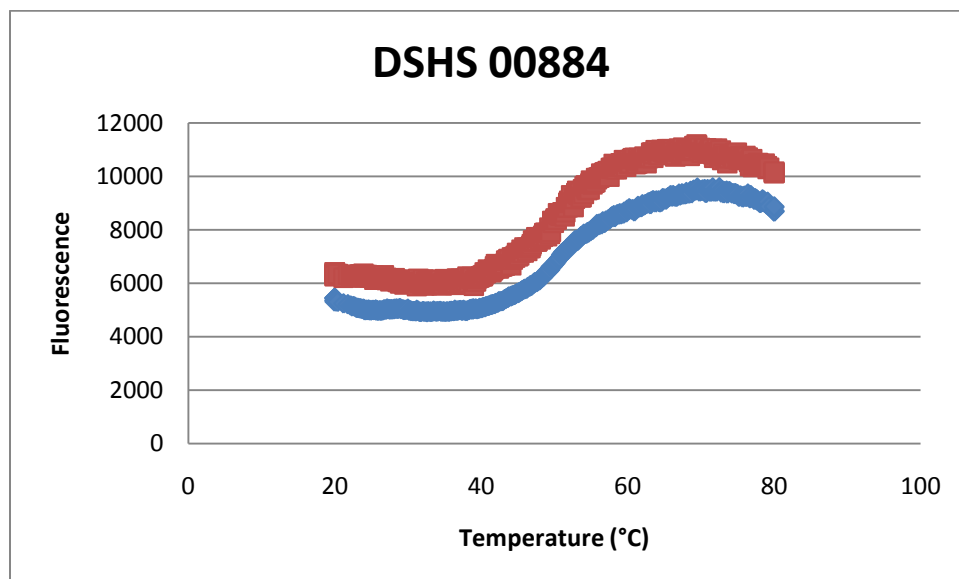
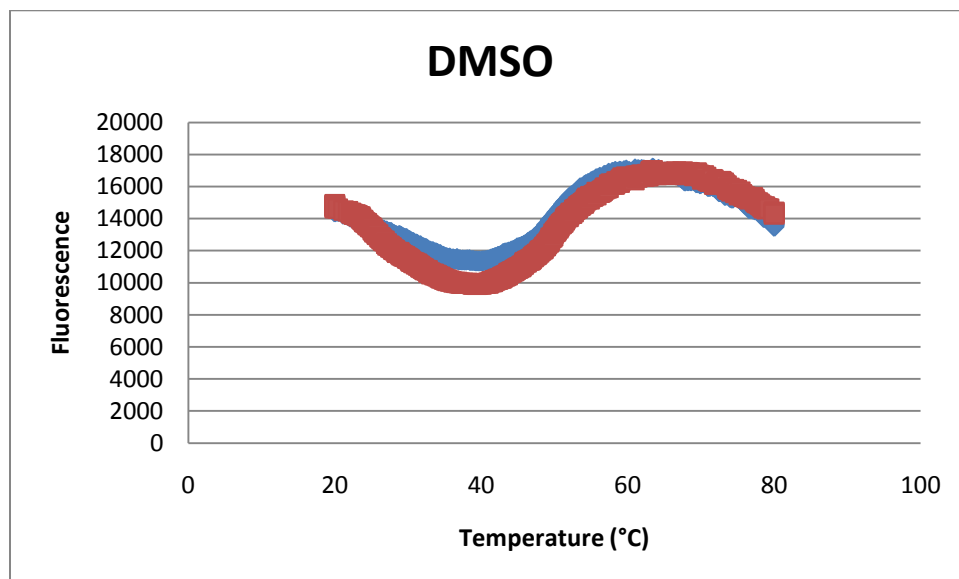
One compound, DSHS00884, has currently been pursued (Figure 3.18). Following purification of E6M, an equimolar amount of DSHS00884 was added and allowed to sit on ice for approximately one hour. Figure 3.19 shows that E6M concentrated well in the presence of DSHS00884. Crystallization trials were performed with complex at 1.8 mg/mL and several conditions yielded crystals (Figure 3.20), one of which had crystals that diffracted to 8 Å at the Advanced Photon Source (APS) microfocus beam.

Figure 3.21 shows the entire gel from the E6 + DSHS00884 preparation. In addition to E6M, a lower band (SUMO tag) is seen as well as two much larger protein impurities. It is unclear what protein actually crystallized. The purification of E6M with DSHS 00884 was repeated and concentrated to 1.4 mg/mL, but yielded no crystals. Because less impurities were present in the second preparation and the concentration was always determined by total protein amount, it's

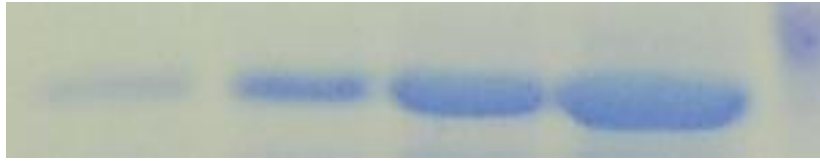
highly possible that E6M is a very different concentration in the second prep than the first. It's also a distinct possibility that we have crystallized an impurity, especially since the unit cell determined from the 8 Å data was very large.

Due to these problems, the procedure for purifying E6M would benefit from some improvements. Primarily, a way to remove impurities is necessary, especially since the impurities are not a big problem initially, but become so during concentration since E6M concentrates slower than the contaminants. Several possibilities have been attempted, but none to any great success because of the need for high salt concentrations and tendency of the protein to interact nonspecifically with resins. The best possibility may be to covalently link one of the identified compounds to resin and pass the impure E6M over it, then elute the protein with either excess of the same compound or another compound that can compete with the resin.

Regardless, 18 compounds have been identified from high throughput screening that have the potential to stabilize E6 for crystallization studies and need to be pursued as such. Perhaps a crystal structure of E6 is on the horizon.

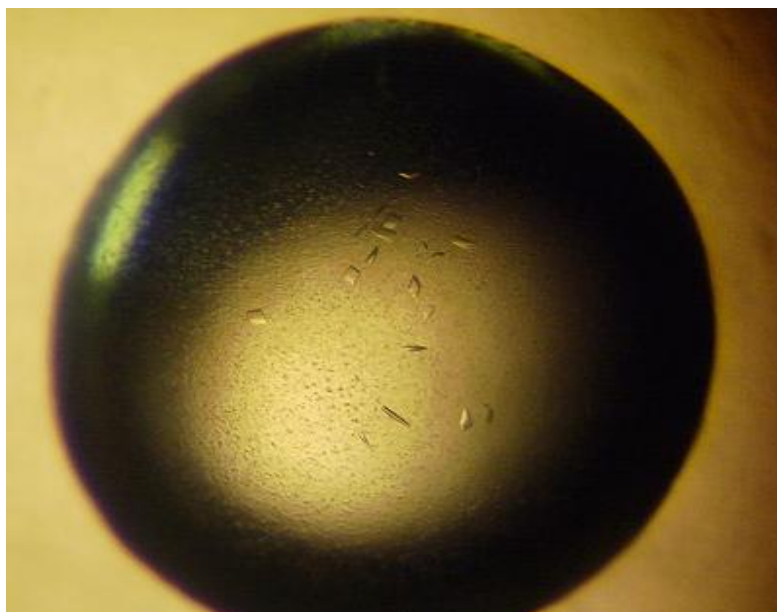


**Figure 3.18: Thermostability assay curves for E6M + DMSO, DSHS00884.** Replicate data (blue and red) for fluorescence at 580 nm vs temperature for DMSO (top) and DSHS 00884 (below) is shown. Inherent protein instability is seen in the region of 20°C and 37°C for E6M + DMSO. This region of instability is significantly flatter in E6M + DSHS 00884.

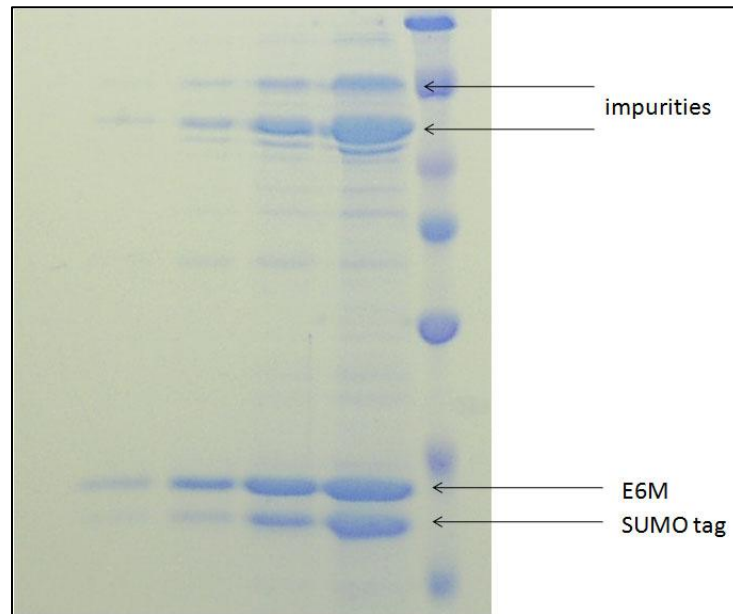


**Figure 3.19: Concentration of the E6M + DSHS00884 complex.** The first lane shows E6M after purification from the Ni-NTA resin. The next three samples were taken during various stages of concentration. The final lane is E6M + DSHS 00884 at a concentration of 1.8 mg/mL.





**Figure 3.20: Crystals of the E6M + DSHS00884.** These crystals were obtained with E6M + DSHS 00884 concentrated to 1.8 mg/mL in 200 mM ammonium citrate tribasic, pH 7.0 and 20% PEG 3350. They appeared after ~ 3 days.



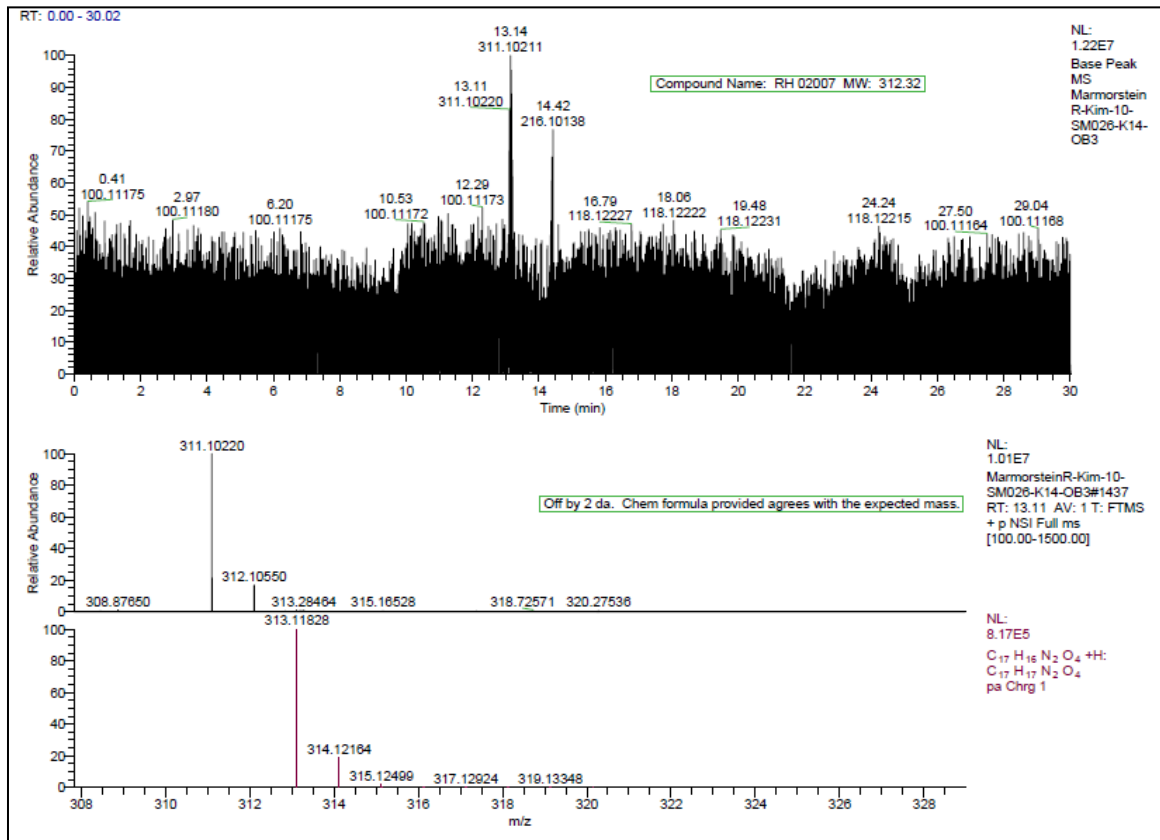
**Figure 3.21: Full gel of E6M + DSHS00884 concentration.** This is a larger view of the gel from [Figure 3.19](#). A lower contaminating band, which is the SUMO tag, is seen as well as two much larger contaminating bands, whose identity is unknown.

### 3.5. Future Work

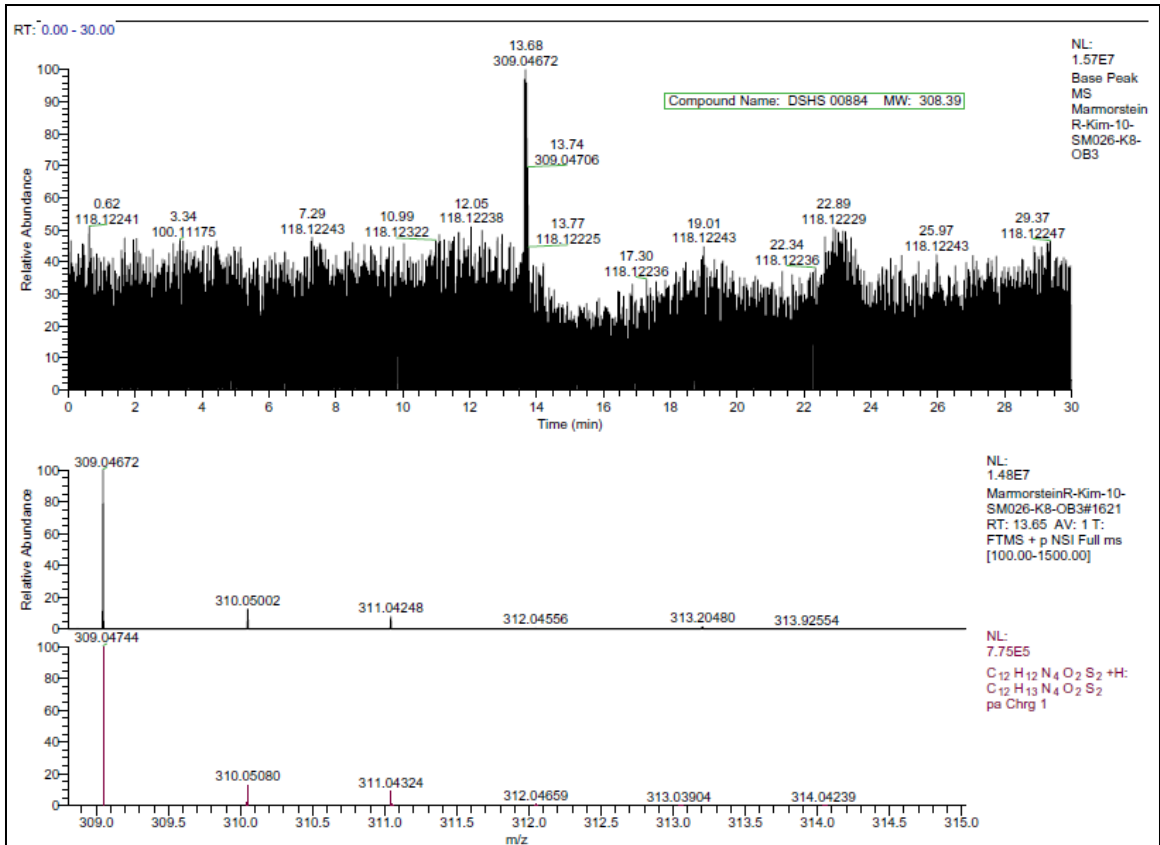
Aside from the crystallization studies that need to be pursued, follow up work on the compounds still needs to be completed. Several compounds inhibit p53 degradation in cells but do not appear to induce apoptosis. Since p53 function includes cell cycle arrest and induction of DNA-damage repair genes, it is possible that these compounds are mediating their function towards these ends. It would also be interesting to determine if some of the 23 compounds that did not work in cells could be modified to improve their behavior without loss of function. Finally, definitively showing the ability of these compounds to directly bind E6 would allow for structure-activity relationship (SAR) studies to find more compounds that can inhibit E6 function at greater potencies and lower toxicity.

## APPENDIX

### Liquid Chromatography - Mass Spectrometry Results for RH 02007 and DSHS 00884



**Figure A.1: LC-MS results for RH 02007.** Predicted molecular weight for the compound is 312.32 g/mol. The data supports a molecule of 310.3 Da, which is a discrepancy of 2 Da (top panel). The middle panel shows the predicted break down of peaks around 312 Da for the compound based on its chemical structure and the bottom panel shows the actual break down of peaks.



**Figure A.2: LC-MS results for DSHS 00884.** Predicted molecular weight for compound is 308.39 g/mol, which is confirmed by the data (top panel). The middle panel shows the predicted break down of peaks around 308 Da for the compound based on its chemical structure and the bottom panel shows the actual break down of peaks. Both panels are the same.

## REFERENCES – Chapter 1

- Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2002) *Molecular Biology of the Cell*, 4th Edition edn. New York, New York: Garland Science.
- Ambrosini G, Sambol E, Carvajal D, Vassilev L, Singer S, Schwartz G (2007) Mouse double minute antagonist Nutline-3a enhances chemotherapy-induced apoptosis in cancer cells with mutant p53 by activating E2F1. *Oncogene* **26**: 3473 - 3481
- Berk AJ (2005) Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus. *Oncogene* **24**: 7673 - 7685
- Brooks C, Gu W (2003) Ubiquitination, phosphorylation, and acetylation: the molecular basis for p53 regulation. *Current Opinion in Cell Biology* **15**: 164 - 171
- Brosch R, Rotter V (2009) When mutants gain new powers: news from the mutant p53 field. *Nature Reviews Cancer* **9**: 701 - 713
- Brown CJ, Lain S, Verma CS, Fersht AR, Lane DP (2009) Awakening guardian angels: drugging the p53 pathway. *Nature Reviews Cancer* **9**: 862 - 873
- Campisi J, d'Adda di Fagagna F (2007) Cellular Senescence: when bad things happen to good cells. *Nature Reviews Molecular and Cell Biology* **8**: 729 - 740
- Chan T, Hermeking H, Lengauer C, Kinzler K, Vogelstein B (1999) 14-3-3 Sigma is required to prevent mitotic catastrophe after DNA damage. *Nature* **401**: 616-620
- Chen Q, Bartholomew J, Campisi J, Acosta M, Reagan J, Ames B (1998) Molecular analysis of H<sub>2</sub>O<sub>2</sub>-induced senescent-like growth arrest in normal human fibroblasts: p53 and Rb control G1 arrest but not cell replication. *The Biochemical Journal* **332**: 43 - 50
- Cho Y, Gorina S, Jeffrey PD, Pavletich N (1994) Crystal Structure of a p53 Tumor Suppressor-DNA Complex: Understanding Tumorigenic Mutations. *Science* **265**: 346 - 355
- de Rozières S, Maya R, Oren M, Lozano G (2000) The loss of mdm2 induces p53-mediated apoptosis. *Oncogene* **19** 1691 - 1697
- DeLeo A, Jay G, Apella E, Dubois G, Law L, Old L (1979) Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse. *Proceedings of the National Academy of Sciences* **76**: 2420 - 2424
- Di Micco R, Fumagalli M, d'Adda di Fagagna F (2007) Breaking news: high-speed race ends in arrest - how oncogenes induce senescence. *Trends in Cell Biology* **17**: 529 - 536

- Dimri G, Itahana K, Acosta M, Campisi J (2000) Regulation of senescence checkpoint response by the E2F1 transcription factor and p14(ARF) tumor suppressor. *Molecular and Cellular Biology* **20**: 273 - 285
- Eddy BE, Borman GS, Berkeley WH, Young RD (1961) Tumors induced in hamsters by injection of rhesus monkey kidney cell extracts. *Proceedings of the Society for Experimental Biology and Medicine* **107**: 191 - 197
- Eliyahu D, Michalovitz D, Eliyahu S, Pinhasi-Kimhi O, Oren M (1989) Wild-type p53 can inhibit oncogene-mediated focus formation. *Proceedings of the National Academy of Sciences* **86**: 8763 - 8767
- Eliyahu D, Raz A, Gruss P, Givol D, Oren M (1984) Participation of p53 cellular tumour antigen in transformation of normal embryonic cells. *Nature* **312**: 646 - 649
- Espinosa J (2008) Mechanisms of regulatory diversity within the p53 transcriptional network. *Oncogene* **27**: 4013-4023
- Feng H, Shuda M, Chang Y, Moore P (2008) Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* **319**: 1096 - 1100
- Ferbeyre G, de Stanchina E, AW L, Querido E, McCurrach M, Hannon G, Lowe S (2002) Oncogenic ras and p53 cooperate to induce cellular senescence. *Molecular and Cellular Biology* **22**: 3497 - 3508
- Ferbeyre G, de Stanchina E, Querido E, Baptiste N, Prives C, Lowe S (2000) PML is induced by oncogenic ras and promotes premature senescence. *Genes and Development* **15**: 2015 - 2027
- Finlay C, Hinds P, Levine A (1989) The p53 proto-oncogene can act as a suppressor of transformation. *Cell* **57**: 1083 - 1093
- Gil J, Peters G (2006) Regulation of the INK4b-ARF-INK4a tumor suppressor locus: all for one and one for all. *Nature Reviews Molecular and Cell Biology* **7**: 667 - 677
- Hardcastle I (2007) Inhibitors of the MDM2-p53 interaction as anticancer drugs. *Drugs of the Future* **32**: 883
- Haupt S, Berger M, Goldberg Z, Haupt Y (2003) Apoptosis - the p53 network. *Journal of Cell Science* **116**: 4077 - 4085
- Haupt Y (2003) Certainly no ARFterthought: oncogenic cooperation in ARF induction a key step in tumor suppression. *Cell Cycle* **2**: 113 - 115
- Hollstein M, Sidransky D, Vogelstein B, Harris C (1991) p53 mutations in human cancers. *Science* **253**: 49 - 53

- Huebner R, Rowe W, Lane W (1962) Oncogenic effects in hamsters of human adenovirus types 12 and 18. *Proceedings of the National Academy of Sciences* **15**: 2051 - 2058
- Huibregste J, Scheffner M, Howley P (1993) Localization of the E6-AP regions that direct human papillomavirus E6 binding, association with p53, and ubiquitination of associated proteins. *Molecular and Cellular Biology* **13**: 4918 - 4927
- Jenkins J, Rudge K, Chumaskov P, Currie G (1985) The cellular oncogene p53 can be activated by mutagenesis. *Nature* **317**: 816 - 818
- Jenkins J, Rudge K, Currie G (1984a) Cellular immortalization by a cDNA clone encoding the transformation-associated phosphoprotein p53. *Nature* **312**: 651 - 654
- Jenkins J, Rudge K, Redmond S, Wade-Evans A (1984b) Cloning and expression analysis of full length mouse cDNA sequences encoding the transformation associated protein p53. *Nucleic Acids Research* **12**: 5609 - 5626
- Jin S, Antinore N, Lung F, Dong X, Zhao H, Fan F, Colchagie A, Blanck P, Roller P, Fornace AJ, Zhan Q (2000) The GADD45 inhibition of Cdc2 kinase correlates with GADD45-mediated growth suppression. *Journal of Biological Chemistry* **275**: 16602 - 16608
- Kerr J (1965) A histochemical study of hypertrophy and ischaemic injury of rat liver with special reference to changes in lysosomes. *Journal of Pathology and Bacteriology* **90**: 419 - 435
- Kirchstein R, Gerber P (1962) Ependymomas produced after intracerebral inoculation of SV40 into newborn hamsters. *Nature* **195**: 299 - 300
- Kortlever R, Higgins P, Bernards R (2006) Plasminogen activator inhibitor-1 is a critical downstream target of p53 in the induction of replicative senescence. *Nature Cell Biology* **8**: 877 - 884
- Kuribayashi K, Krigsfeld G, Wang W, Xu J, Mayes P, Dicker D, Wu G, El-Diery W (2008) TNFSF10 (TRAIL), a p53 target gene that mediates p53-dependent cell death. *Cancer Biology & Therapy* **7**: 2034 - 2038
- Lane D (1992) Cancer: p53, guardian of the genome. *Nature* **358**: 15 - 16
- Lane D, Crawford L (1979) T antigen is bound to a host protein in SV40-transformed cells. *Nature* **278**: 261 - 263
- LaRusch G, Jackson M, Dunbar J, Warren R, Donner D, Mayo L (2007) Nutlin3 blocks vascular endothelial growth factor induction by preventing the interaction between hypoxia inducible factor 1alpha and Hdm2. *Cancer Research* **67**: 450 - 454
- Lau M, Nugent J, Zhao X, Irwin M (2008) HDM2 antagonists nutlin-3 disrupts p73-HDM2 binding and enhances p73 function. *Oncogene* **27**: 997 - 1003



- Lee J, Gu W (2010) The multiple levels of regulation by p53 ubiquitination. *Cell Death and Differentiation* **17**: 86 - 92
- Levine A, Hu W, Feng Z (2006) The p53 pathway: what questions remain to be explored? *Cell Death and Differentiation* **13**: 1027 - 1036
- Levine AJ (2009) The common mechanisms of transformation by the small DNA tumor viruses: The inactivation of tumor suppressor gene product: p53. *Virology* **384**: 285 - 293
- Linzer D, Levine A (1979) Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* **17**: 43 - 52
- Lodish H, Berk A, Matsudaira P, Kaiser CA, Scott MP, Ziprusky SL, Darnell J (2004) *Molecular Cell Biology*, 5th Edition edn. New York, New York: W. H. Freeman and Company.
- Lowe DB, Shearer MH, Jumper CA, Kennedy RC (2007) SV40 association with human malignancies and mechanisms of tumor immunity by large tumor antigen. *Cellular and Molecular Life Sciences* **64**: 803 - 814
- Luna RMdO, Wagner D, Lozano G (1995) Rescue of early embryonic lethality in mdm2-deficient mice by deletion of p53. *Nature* **378**: 203 - 206
- MacFarlane M, Williams A (2004) Apoptosis and disease: a life or death decision. *EMBO Reports* **5**: 674 - 678
- Maecker H, Koumenis C, Giaccia A (2000) p53 promotes selection for Fas-mediated apoptotic resistance. *Cancer Research* **60**: 4638 - 4644
- Malkin D, Li F, Strong L, Fraumeni JJ, Nelson C, Kim D, Kassel J, Gryka M, Bischoff F, Tainsky M (1990) Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* **250**: 1233 - 1238
- Maltzmann W, Czyzyk L (1984) UV irradiation stimulates levels of p53 cellular tumor antigen in nontransformed mouse cells. *Molecular and Cellular Biology* **4**: 1689 - 1694
- Marchenko N, Zaika A, Moll U (2000) Death signal-induced localization of p53 protein to mitochondria. A potential role in apoptotic signaling. *Journal of Biological Chemistry* **275**: 16202 - 16212
- Mihara M, Erster S, Zaika A, Petrenko O, Chittenden T, Pancoska P, Moll U (2003) p53 has a direct apoptogenic role at the mitochondria. *Molecular Cell* **11**: 577 - 590
- Milner J (1984) Different forms of p53 detected by monoclonal antibodies in non-dividing and dividing lymphocytes. *Nature* **310**: 143 - 145

- Momand J, Jung D, Wilczynski S, Niland J (1998) The MDM2 gene amplification database. *Nucleic Acids Research* **26**: 3453 - 3459
- Mowat M, Cheng A, Kimura N, Bernstein A, Benchimol S (1985) Rearrangements of the cellular p53 gene in erythroleukaemic cells transformed by Friend virus. *Nature* **314**: 633 - 636
- Mu X, Higgins P (1995) Differential growth state-dependent regulation of plasminogen activator inhibitor type-1 expression in senescent IMR-90 human diploid fibroblasts. *Journal of Cellular Physiology* **165**: 647 - 657
- Munro J, Barr N, Ireland H, Morrison V, EK P (2004) Histone deacetylase inhibitors induce a senescence-like state in human cells by a p16-dependent mechanism that is independent of a mitotic clock. *Experimental cell research* **295**: 525 - 538
- Oren M (1992) p53: the ultimate tumor suppressor gene? *FASEB J* **6**: 3169 - 3176
- Parrinello S, Samper E, Krtolica A, Goldstein J, Melov S, Campisi J (2003) Oxygen sensitivity severely limits the replicative lifespan of murine fibroblasts. *Nature Cell Biology* **5**: 741 - 747
- Pearson M, Carbone R, Sebastiani C, Cioce M, Fagioli M, Saito S, Higashimoto Y, Apella E, Minucci S, Pandolfi P, Pelicci P (2000) PML regulates p53 acetylation and premature senescence induced by oncogenic Ras. *Nature* **406**: 201 - 210
- Pearson S, Jia H, Kandachi K (2004) China approves first gene therapy. *Nature Biotechnology* **22**: 3 - 4
- Peng C, Graves P, Thoma R, Wu Z, Shaw A, Piwnicka-Worms H (1997) Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding to phosphorylation to Cdc25c on serine-216. *Science* **277**: 1501-1505
- Poulin D, DeCaprio J (2006) Is there a role for SV40 in human cancer? *Journal of Clinical Oncology* **24**: 4356 - 4365
- Rabson A, Kirchenstein R, Paul F (1964) Tumors produced by adenovirus 12 in mastomys and mice. *Journal of the National Cancer Institute* **32**: 77 - 87
- Roth J (2006) Adenovirus p53 gene therapy. *Expert Opinion on Biological Therapy* **6**: 55 - 61
- Rother K, Kirschner R, Sanger K, Bohlig L, Mossner J, Engeland K (2007) p53 downregulates expression of the G1/S cell cycle phosphatase Cdc25a. *Oncogene* **26**: 1949 - 1953
- Rotter V, Witte O, Coffman R, Baltimore D (1980) Abelson murine leukemia virus-induced tumors elicit antibodies against a host cell protein, P50. *Journal of Virology* **36**: 547 - 555
- Scheffner M (1998) Ubiquitin, E6-AP, and their role in p53 inactivation. *Pharmacological Therapy* **78**: 129 - 139

Scheffner M, Huibregste J, Viestra R, Howley P (1993) The HPV-16 E6 and E6-AP Complex Functions as a Ubiquitin-Protein Ligase in the Ubiquitination of p53. *Cell* **75**: 495 - 505

Serrano M, Lin A, McCurrach M, Beach D, Lowe S (1997) Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88**: 593 - 602

Sharpless N, Ramsey M, Balasubramanian P, Castrillon D, DePinho RA (2004) The differential impact of p16(INK4a) or p19(ARF) deficiency on cell growth and tumorigenesis. *Oncogene* **23**: 379 - 385

Sherman M, Gabai V, O'Callaghan C, Yaglom J (2007) Molecular chaperones regulate p53 and suppress senescence programs. *FEBS Letters* **581**: 3711 - 3715

Sherr CJ, DePinho RA (2000) Cellular Senescence: Mitotic Clock or Culture Shock? *Cell* **102**: 407 - 410

Shetty S, Shetty P, Idell S, Velusamy T, Bhandary Y, Shetty R (2008) Regulation of plasminogen activator inhibitor-1 expression by tumor suppressor protein p53. *Journal of Biological Chemistry* **283**: 19570 - 19580

Smogorzewska A, de Lange T (2002) Different telomere damage signaling pathways in human and mouse cells. *The EMBO Journal* **21**: 4338 - 4348

Society AC. (2010) Cancer Facts & Figures 2010. In Society AC (ed.), *American Cancer Society*. American Cancer Society, Atlanta, GA.

Spurgers K, Gold D, Coombes K, Boehstiehl N, Mullins B, Meyn R, Logothetis C, McDonnell T (2006) Identification of cell cycle regulatory genes as principal targets of p53-mediated transcriptional repression. *Journal of Biological Chemistry* **281**: 25134 - 25142

Srivastava S, Zou Z, Pirollo K, Blattner W, Chang E (1990) Germ-line transmission of a mutated p53 gene in cancer-prone family with Li-Fraumeni syndrome. *Nature* **348**: 747 - 749

St Clair S, Giono L, Varmeh Z, S, Resnick-Silverman L, Liu W, Padi A, Dastidar J, DaCosta A, Mattia M, Manfredi J (2004) DNA damage-induced downregulations of Cdc25c is mediated by p53 via two independent mechanisms: one involves direct binding to the cdc25c promoter. *Molecular Cell* **16**: 725 - 736

Suryadinata R, Sadowski M, Sarcevic B (2010) Control of cell cycle progression by phosphorylation of cyclin-dependent kinase (CDK) substrates. *Bioscience Reports* **30**: 243 - 255

Tegtmeyer P, Schwartz M, Collins JK, Rundell K (1975) Regulation of Tumor Antigen Synthesis by Simian Virus 40 Gene A. *Journal of Virology* **16**: 168 - 178

Vaseva A, Moll U (2009) The mitochondrial p53 pathway. *Biochimica et biophysica acta* **1787**: 414 - 420

Wang Z, Sun Y (2010) Targeting p53 for Novel Anticancer Therapy. *Translational Oncology* **3**: 1 - 12

Wolyniec K, Wotton S, Kilbey A, Jenkins A, Terry A, Peters G, Stocking C, Cameron E, Neil J (2009) RUNX1 and its fusion oncoprotein derivative, RUNX1-ETO, induce senescence-like growth arrest independently of replicative stress. *Oncogene* **28**: 2502 - 2512

Wotton S, Blyth K, Kilbey A, Jenkins A, Terry A, Bernardin-Fried F, Friedman A, Baxter E, Neil J, Cameron E (2004) RUNX1 transformation of primary embryonic fibroblasts is revealed in the absence of p53. *Oncogene* **23**: 5476 - 5486

Yao H, Li P, Venters B, Zheng S, Thompson P, Pugh B, Wang Y (2008) Histone Arg modifications and p53 regulate the expression of OKL38, a mediator of apoptosis. *Journal of Biological Chemistry* **283**: 20060 - 20068

Zamanian M, La Thangue N (1993) Transcriptional repression by the Rb-related protein p107. *Molecular Biology of the Cell* **4**: 389 - 396

Zhan Q, Antinore N, Wang X, Carrier F, Smith M, Harris C, Fornace AJ (1999) Association with Cdc2 and inhibition of Cdc2/CyclinB1 kinase activity by the p53-regulated protein Gadd45. *Oncogene* **18**: 2892 - 2900

Zuckerman V, Wolyniec K, Sionov RV, Haupt S, Haupt Y (2009) Tumor suppression by p53: the importance of apoptosis and cellular senescence. *Journal of Pathology* **219**: 3 - 15

## REFERENCES – Chapter 2

- Balagurumoorthy P, Sakamoto H, Lewis MS, Zambrano N, Clore GM, Gronenborn AM, Apella E, Harrington RE (1995) Four p53 DNA-binding domain peptides bind natural p53-response elements and bend the DNA. *PNAS USA* **92**: 8591 - 8595
- Banerjee A, Santos WL, Verdine GL (2006) Structure of a DNA Glycosylase Searching for Lesions. *Science* **311**: 1153 - 1157
- Chen Y, Dey R, Chen L (2010) Crystal structure of the p53 core domain bound to a full consensus site as a self-assembled tetramer. *Structure* **18**: 246 - 256
- Cho Y, Gorina S, Jeffrey PD, Pavletich NP (1994) Crystal Structure of a p53 Tumor Suppressor-DNA Complex: Understanding Tumorigenic Mutations. *Science* **265**: 346 - 355
- DeLano WL (2002) *The PyMOL Molecular Graphics System*.
- El-Diery WS, Kern SE, Pietenpol JA, Kinzler KW, Vogelstein B (1992) Definition of a consensus binding site for p53. *Nature Genetics* **1**: 45 - 49
- Ellenberger T, Brandl C, Struhl K, Harrison S (1992) The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. *Cell* **71**: 1223 - 1237
- Emsley P, Cowtan K (2004) Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallography Section D - Biological Crystallography* **60**: 2126 - 2132
- Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExpASY: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* **31**: 3784 - 3788
- Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: multiple sequence alignments in PostScript. *Bioinformatics* **15**: 305 - 308
- Harms K, Chen X (2006) The functional domains in p53 family proteins exhibit both common and distinct properties. *Cell Death and Differentiation* **13**: 890 - 897
- Ho WC, Fitzgerald MX, Marmorstein R (2006) Structure of the mouse p53 core domain dimer bound to DNA. *Journal of Biological Chemistry* **281**: 20494 - 20502
- Huang H, Chopra R, Verdine GL, Harrison SC (1998) Structure of Covalently Trapped Catalytic Complex of HIV-1 Reverse Transcriptase: Implications for Drug Resistance. *Science* **282**: 1669 - 1675

- Jeffrey PD, Gorina S, Pavletich NP (1995) Crystal Structure of the Tetramerization Domain of the p53 Tumor Suppressor at 1.7 Angstroms. *Science* **267**: 1498 - 1502
- Kitayner M, Rosenberg H, Kessler N, Rabinovich D, Shaulov L, Haran TE, Shakked Z (2006) Structural Basis of DNA Recognition by p53 Tetramers. *Molecular Cell* **22**: 741 - 753
- Klein C, Georges G, Kunkele KP, Huber R, Engh RA, Hansen S (2001) High thermostability and lack of cooperative DNA binding distinguish the p63 core domain from the homologous tumor suppressor p53. *Journal of Biological Chemistry* **276**: 37390 - 37401
- Liu J, Fan Q, Sodeoka M, Lane W, Verdine G (1994) DNA binding by an amino acid residue in the C-terminal half of the Rel homology region. *Chemistry & Biology* **1**: 47 - 55
- Lu X-L, Olson WK (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Research* **31**: 5108 - 5121
- MacMillan AM, Verdine GL (1991) Engineering Tethered DNA Molecules by the Convertible Nucleoside Approach. *Tetrahedron* **47**: 2603 - 2616
- Malecka KA, Ho WC, Marmorstein R (2009) Crystal Structure of a p53 core tetramer bound to DNA. *Oncogene* **28**: 325 - 333
- McClure KG, Lee PWK (1998) How p53 binds DNA as a tetramer. *The EMBO Journal* **17**: 3342 - 3350
- McLure KG, Lee PWK (1998) How p53 binds DNA as a tetramer. *The EMBO Journal* **17**: 3342 - 3350
- McNamara PT, Bolshoy A, Trifonov EN, Harrington RE (1990) Sequence-dependent kinks induced in curved DNA. *Journal of Biomolecular Structural Dynamics* **8**: 529 - 538
- Nagaich AK, Apella E, Harrington RE (1997a) DNA Bending is Essential for the Site-specific Recognition of DNA Reponse Elements by the DNA Binding Domain of the Tumor Suppressor Protein p53. *The Journal of Biological Chemistry* **272**: 14842 - 14849
- Nagaich AK, Bhattacharyya D, Brahmachari SK, Bansal M (1994) CA/TG sequence at the 5' end of oligo(A)-tracts strongly modulates DNA curvature. *Journal of Biological Chemistry* **269**: 7824 - 7833
- Nagaich AK, Zhurkin VB, Durell SR, Jernigan RL, Apella E, Harrington RE (1999) p53-induced DNA bending and twisting: p53 tetramer binds on the outer side of a DNA loop and increases DNA twisting. *PNAS USA* **96**: 1875 - 1880
- Nagaich AK, Zhurkin VB, Sakamoto H, Gorin AA, Clore GM, Gronenborn AM, Apella E, Harrington RE (1997b) Architectural Accomodation in the Complex of Four p53 DNA Binding Domain

Peptides with the p21/waf1/cip1 DNA Response Element. *The Journal of Biological Chemistry* **272**: 14830 - 14841

Okorokov AL, Sherman MB, Plisson C, Grinkevich V, Sigmundsson K, Selivanova G, Milner J, Orlova EV (2006) The Structure of p53 tumour suppressor protein reveals the bases for its functional plasticity. *The EMBO Journal* **25**: 5191 - 5200

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**

Pan Y, Nussinov R (2007) Structural Basis for p53 Binding-induced DNA Bending. *The Journal of Biological Chemistry* **282**: 691 - 699

Pan Y, Nussinov R (2008) p53-Induced DNA Bending: The Interplay between p53-DNA and p53-p53 Interactions. *The Journal of Physical Chemistry B* **112**: 6716 - 6724

Pietenpol JA, Tokino T, Thiagaligam S, El-Diery WS, Kinzler KW, Vogelstein B (1994) Sequence-specific transcriptional activation is essential for growth suppression by p53. *PNAS USA* **91**: 1998 - 2000

Rhodes G (2000) *Crystallography made crystal clear*, 2nd edn. San Diego, CA: Academic Press.

Tidow H, Melero R, Mylonas E, Freund SMV, Grossman JG, Carazo JM, Svergun DI, Valle M, Fersht AR (2007) Quaternary structures of tumor suppressor p53 and a specific p53-DNA complex. *Proceedings of the National Academy of Sciences*

Verdine GL, Norman DPG (2003) Covalent Trapping of Protein-DNA Complexes. *Annual Reviews in Biochemistry* **72**: 337 - 366

Viadiu H (2008) Molecular Architecture of Tumor Suppressor p53. *Current Topics in Medicinal Chemistry* **8**: 1327 - 1334

Wang Y, Rosengarth A, Luecke H (2007) Structure of the human p53 core domain in the absence of DNA. *Acta Crystallography* **D63**: 276 - 281

Weinberg RL, Veprintsev DB, Fersht AR (2004) Cooperative Binding of Tetrameric p53 to DNA. *Journal of Molecular Biology* **341**: 1145 - 1159

Willis M, Hicke B, Uhlenbeck O, Cech T, Kock T (1993) Photocrosslinking of 5-iodouracil-substituted RNA and DNA to proteins. *Science* **262**: 1255 - 1257

Xu Y-Z, Zheng Q, Swann PF (1992) Synthesis of DNA Containing Modified Bases by Postsynthetic Substitution. *Journal of Organic Chemistry* **57**: 3839 - 3845

Zhao K, Chai X, Johnston K, Clements A, Marmorstein R (2001) Crystal Structure of the Mouse p53 Core DNA-binding Domain at 2.7 Å Resolution. *Journal of Biological Chemistry* **276**: 12120 - 12127

Zhurkin VB, B. UN, Gorin AA, Jernigan RL (1991) Static and statistical bending of DNA evaluated by Monte Carlo simulations. *Proceedings of the National Academy of Sciences, USA* **88**: 7046 - 7050

Zupnick A, Prives C (2006) Mutational Analysis of the p53 Core Domain L1 Loop. *The Journal of Biological Chemistry* **281**: 20464 - 20473



## REFERENCES – Chapter 3

- Alvarez-Salas L, diPaolo J (2007) Molecular approaches to cervical cancer therapy. *Current drug discovery technologies* **4**: 208 - 219
- Arkin MR, Wells JA (2004) Small Molecule Inhibitors of Protein-Protein Interactions: Progressing Towards the Dream. *Nature Reviews Drug Discovery* **3**: 301
- Arroyo M, Bagchi S, Raychaydhuri P (1993) Association of the human papillomavirus type 16 E7 protein with the S-phase-specific E2F-cyclin-A complex. *Molecular and Cellular Biology* **13**: 6537
- Baleja J, Cherry J, Liu Z, Gao H, Nicklaus M, Voigt J, Chen J, Androphy E (2006) Identification of inhibitors to papillomavirus type 16 E6 protein based on three-dimensional structures of interacting proteins. *Antiviral Research* **72**: 49 - 59
- Baseman J, Koutsky L (2005) The epidemiology of human papillomavirus infections. *Journal of Clinical Virology* **32(suppl1)**: S16 - S24
- Be X, Hong Y, Wei J, Androphy EJ, Chen JJ, Baleja JD (2001) Solution Structure Determination and Mutational Analysis of the Papillomavirus E6 Interacting Peptide of E6AP. *Biochemistry* **40**: 1293 - 1299
- Butz K, Denk C, Ullmann A, Scheffner M, Hoppe-Seyler F (2000) Induction of apoptosis in human papillomavirus-positive cancer cells by peptide aptamers targeting the viral E6 oncoprotein. *Proceedings of the National Academy of Sciences* **97**: 6693 - 6697
- Cole S, Danos O (1987) Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome. Phylogeny of papillomaviruses and repeated structure of the E6 and E7 gene products. *Journal of molecular biology* **193**: 599 - 608
- Cooper B, Schneider S, Bohl J, Jiang Y-h, Beaudet A, Pol SV (2003) Requirement of E6AP and the features of human papillomavirus E6 necessary to support degradation of p53. *Virology* **306**: 87 - 99
- de Villiers EM, Fauquet C, Broker TR, Bernard HU, Hausen Hz (2004) Classification of papillomaviruses. *Virology* **324**: 17
- DeFilippis RA, Goodwin EC, Wu L, DeMaio D (2003) Endogenous Human Papillomavirus E6 and E7 Proteins Differentially Regulate Proliferation, Senescence, and Apoptosis in HeLa Cervical Carcinoma Cells. *Journal of Virology* **77**: 1551 - 1563
- Degeterev A (2001) Identification of small molecule inhibitors of interaction between the BH3 domain and Bcl-XL. *Nature Cell Biology* **3**: 173 - 182
- Devlin JJ, Liang A, Trinh L, Polokoff MA, Senator D, Zhang W, Kondracki J, Kretschmer PJ, Morser J, Lipson SE, Spann R, Loughlin JA, Dunn KV, Morrissey MM (1996) High Capacity Screening of

Pooled Compounds: Identification of the Active Compound Without Re-Assay of Pool Members. *37Drug Development Research*

Doorbar J (2005) The papillomavirus life cycle. *Journal of Clinical Virology* **32**: S7 - S15

Dyson N (1998) The regulation of E2F by pRB-family proteins. *Genes and Development* **12**: 2245 - 2262

Ellis RE, Yuan JY, Horvitz HR (1991) Mechanisms and functions of cell death. *Annual Review of Cellular Biology* **7**: 663 - 698

Elston R, Napthine S, Doorbar J (1998) *Journal of General Virology* **79**: 371 - 374

Gardiol D, Kuhne C, Glaunsinger B, Lee SS, Javier R, Banks L (1999) Oncogenic human papillomavirus E6 proteins target the discs large tumour suppressor for proteasome-mediated degradation. *18*

Ghittoni R, Accardi R, Hasan U, Gheit T, Sylla B, Tommasino M (2010) The Biological Properties of E6 and 7 oncoproteins from Human Papillomaviruses. *Virus Genes* **40**: 1 - 13

Griffin H, Elston R, Jackson D, Ansell K, Coleman M, Winter G, Doorbar J (2006) Inhibition of Papillomavirus Protein Function in Cervical Cancer Cells by Intrabody Targeting. *Journal of Molecular Biology* **355**: 360 - 378

Hausen Hz (1996) Papillomavirus infections - a major cause of human cancers. *Biochim Biophys Acta* **1288**

Hausen Hz (2002) Papillomaviruses and cancer: from basic studies to clinical application. *Nature Reviews Cancer* **2**: 342

He W, Staples D, Smith C, Fisher C (2003) Direct activation of cyclin-dependent kinase 2 by human papillomavirus E7. *Journal of Virology* **77**: 10566

Huang L, Kinnucan E, Wang G, Beaudenon S, Howley PM, Huibregste JM, Pavletich NP (1999) Structure of an E6AP-UbcH7 Complex: Insights into Ubiquitination by the E2-E3 Enzyme Cascade. *Science* **286**: 1321 - 1326

Huibregste JM, Scheffner M, Howley PM (1993) Localization of the E6-AP Regions that Direct Human Papillomavirus E6 Binding, Association with p53, and Ubiquitination of Associated Proteins. *Molecular and Cellular Biology* **13**: 4918 - 4927

Kutzki O (2002) Development of a potent Bcl-XL antagonist based on alpha helix mimicry. *Journal of American Chemical Society* **124**: 11838 - 11839

Li X, Coffino P (1996) High-Risk Human Papillomavirus E6 Protein has Two Distinct Binding Sites within -53, of Which Only One Determines Degradation. *Journal of Virology* **70**: 4509 - 4516

- Lipari F, McGibbon GA, Wardrop E, Cordingley MG (2001) Purification and Biophysical Characterization of a Minimal Functional Domain and of an N-terminal Zn<sup>2+</sup>-Binding Fragment from the Human Papillomavirus Type 16 E6 Protein. *Biochemistry* **40**: 1196 - 1204
- Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **23**: 3 - 25
- Liu X, Clements A, Zhao K, Marmorstein R (2006) Structure of the Human Papillomavirus Oncoprotein and Its Mechanism for Inactivation of the Retinoblastoma Tumor Suppressor. *Journal of Biological Chemistry* **281**: 578 - 586
- Liu Y, Liu Z, Androphy E, Chen J, Baleja J (2004) Design and Characterization of Helical Peptides that Inhibit the E6 Protein of Papillomavirus. *Biochemistry* **43**: 7421 - 7431
- Marquez-Gutierrez M, Benitz-Hess M, DiPaolo J, Alvarez-Salas L (2007a) Effect of Combined Antisense Oligodeoxynucleotide Directed against the Human Papillomavirus Type 16 on Cervical Cancer Carcinoma Cells. *Archives of Medical Research* **38**: 730 - 738
- Marquez-Gutierrez MA, Benitez-Hess ML, DiPaolo JA, Alvarez-Salas LM (2007b) Effect of Combined Antisense Oligodeoxynucleotides Directed against the Human Papillomavirus Type 16 on Cervical Carcinoma Cells. *Archives of Medical Research* **38**: 730 - 738
- Munger K, et al (2001) Biological Activities and molecular targets of the human papillomavirus E7 oncoprotein. *Oncogene* **20**: 7887 - 7888
- Munger K, Howley PM (2002) Human papillomaviruses immortalization and transformation functions. *Virus Research* **89**: 213 - 228
- Nicoletti VG, Stella AM (2003) Role of PARP under stress conditions: Cell death or protection? *Neurochemistry Research* **26**: 187 - 194
- Niu X, Peng Z, Duan W, Wang H, Wang P (2006) Inhibition of HPV 16 E6 oncogene expression by RNA interference in vitro and in vivo. *International Journal of Gynecological Cancer* **16**: 743 - 751
- Nomine Y, Charbonnier S, Miguët L, Potier N, Dorselaer AV, Atkinson A, Trave G, Kieffer B (2005) 1H and 15N resonance assignment, secondary structure and dynamic behavior of the C-terminal domain of human papillomavirus oncoprotein E6. *Journal of Biomolecular NMR* **31**: 129 - 141
- Nomine Y, Charbonnier S, Ristriani T, Stier G, Masson M, Cavusglu N, Dorselaer AV, Weiss E, Kieffer B, Trave G (2003) Domain Substructure of HPV E6 Oncoprotein: Biophysical Characterization of the E6 C-Terminal DNA-Binding Domain. *Biochemistry* **42**: 4909 - 4917

Nomine Y, Masson M, Charbonnier S, Zanier K, Ristriani T, Deryckere F, Sibler A-P, Desplancq D, Atkinson RA, Weiss E, Orfanoudakis G, Kieffer B, Gilles T (2006) Structural and Functional Analysis of E6 Oncoprotein: Insights in the Molecular Pathways of Human Papillomavirus-Mediated Pathogenesis. *Molecular Cell* **21**: 665 - 678

Patel D, Huang S-M, Baglia LA, McCance DJ (1999) The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. *The EMBO Journal* **18**: 5061 - 5072

Ristriani T, Masson M, Nomine Y, Laurent C, Lefevre J-F, Weiss E, Trave G (2000) HPV Oncoprotein E6 is a Structure-dependant DNA-binding Protein that Recognizes Four-Way Junctions. *Journal of Molecular Biology* **296**: 1189 - 1203

Ristriani T, Nomine Y, Masson M, Weiss E, Trave G (2001) Specific Recognition of Four-way DNA Junctions by the C-terminal Zinc-binding Domain of HPV Oncoprotein E6. *Journal of Molecular Biology* **305**: 729 - 735

Scheffner M (1998) Ubiquitin, E6-AP, and their role in p53 inactivation. *Pharmacological Therapy* **78**: 129 - 139

Scheffner M, Huibregste JM, Viestra RD, Howley PM (1993) The HPV-16 E6 and E6-AP Complex Functions has a Ubiquitin-Protein Ligase in the Ubiquitination of p53. *Cell* **75**: 495 - 505

Seedorf K, Oltersdorf T, Krammer G, Rowekamp W (1987) Identification of early proteins of the human papillomavirus viruses type 16 (HPV 16) and type 18 (HPV 18) in cervical carcinomas. *The EMBO Journal* **6**: 139 - 144

Thomas MC, Chaing C-M (2005) E6 Oncoprotein Represses p53-Dependent Gene Activation via Inhibition of Protein Acetylation Independently of Inducing p53 Degradation. *Molecular Cell* **17**: 251 - 264

Trucco C, al. e (1998) DNA repair defect in poly (ADP-ribose) polymerase-deficient cell lines. *Nucleic Acids Research* **26**: 2644 - 2649

Wentzensen N, Vinokurova S, M. vKD (2004) Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. *Cancer Research* **64**: 3878 - 3884

Zanier K, Charbonnier S, Mireille Baltzinger, Nomine Y, Altschuh D, Trave G (2005) Kinetic Analysis of the Interactions of Human Papillomavirus E6 Oncoproteins with the Ubiquitin Ligase E6AP Using Surface Plasmon Resonance. *Journal of Molecular Biology* **349**: 401 - 412

Zhang J, Chung T, Oldenburg K (1999) A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of Biomolecular Screening* **4**: 67 - 73