

Potential dual purpose molecular probe for HPV16 E6: a computational and experimental approach

By

John Joseph Arthur Wigg

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science in Chemistry

Faculty of Science and Environmental Studies

Department of Chemistry

Lakehead University

December 2013

© John Wigg 2013

Abstract

The goal of this project is to find a molecular probe for HPV variant 16 protein E6. This goal is accomplished using a combination of Virtual Ligand Screening (VLS) and experimental ligand binding assays.

Protein E6 is a protein that is expressed in all individuals with the human papillomavirus (HPV) virus. A number of HPV 16 protein E6 genetic variants have been associated with increased risk for cervical cancer. Currently, there is no fast and reliable method for testing for the presence of these high-risk protein variants. We want to develop a method that is both fast and reliable for detecting the presence of this protein using molecular probes. These molecular probes could potentially be used for optical or PET diagnostic imaging, for risk-assessment and to guide early intervention.

Using virtual ligand screening, we have identified a number of molecular probe candidates with affinity for the binding site of prototype E6. This affinity is assessed using binding energy scores. Compounds possessing favorable calculated binding energy scores and other desirable molecular properties are identified as potential molecular probes. Selected candidates include: O-succinyl-L-homoserine (-46 kcal/mol), paclitaxel (-70 kcal/mol), and 3-Amino-5-fluorobenzo [E] [1,2,4] Triazine 1, 4 dioxide (-47kcal/mol), for 3-Chloro-2-(2-[(3-oxo-2-Benzfuran-1 (3H)-Yliden] methyl) Hydrazino)-5-(Trifluoromethyl) Pyridinium Acetate (-68.33kcal/mol), 3-Chloro-2-(3-[1-(Phenyl sulfonyl)-(H-pyrazol-3-YL] Phenoxy)-5-(Trifluoromethyl) Pyridine (-66.12kcal/mol), and N-(2,4-dinitrophenyl)-L-arginine) (-78.58kcal/mol).

We have performed experimental ligand-binding assays of the potential molecular probes against purified prototype protein E6. Both intrinsic tryptophan fluorescence and fluorescence

polarization assays have been performed. One of our compounds 3-Amino-5-fluorobenzo [E] [1,2,4] Triazine 1, 4 dioxide (AFTD) was shown to have affinity for our target protein (EC50 6 μ M). Due to its intrinsic fluorescence and the presence of a fluorine atom, this molecule could potentially be used for either PET or optical imaging modalities. Cell based assays have also been conducted to further characterize AFTD's potential as an *in vivo* imaging probe. Results indicate that AFTD is cytotoxic at 25 μ M, although it appears to be cell permeable. Further experiments need to be conducted to assess its potential use as *in-vivo* imaging probe. An alternative use for this molecule as a probe in biochemical assay tracking E6 protein-protein interactions is suggested. This is demonstrated using a monoclonal antibody (6F4) that is specific to HPV16 E6. These results set the stage for future work to further characterize this potential dual-purpose probe which may aid our understanding of the link between HPV infection and cancer.

Acknowledgements

I would like to acknowledge my supervisor Dr. Wely Floriano for her help and guidance in this project. I am truly grateful for your insightful guidance over the last two years. Also, my committee members Dr. Campbell, Dr. Zehbe, and Dr. Qin for their suggestions and guidance over the course of this master's project. Dr. Campbell for his help with LC-MS experiments and Dr. Zehbe for allowing me to use her laboratory for cellular based methods. I would also like to thank the Dr. Floriano laboratory group (Rhiannon, Saedeh, and Thomas). Rhiannon Kamstra for putting together the fluorine containing ligand compound database, and Tomas Sitter and Daryl Willick for the development of a script which was used to figure out which compounds required resubmission to SHARCNET. I would also like to thank Melissa Togtema for her help with cell culturing, and Morshed and Chris Phenix for their assistance with organic synthesis and conjugate identification.

I would like to thank SHARCNET for the use of computational resources over the course of this project. I would also like to thank the Thunder Bay Regional Research Institute (TBRRI) for the use of their laboratory equipment. The Lakehead University Chemistry Department for their help and support. Finally the Lakehead University Instrumentation Laboratory for all of their help and use of the facility at many stages of this project.

On a more personal note I would also like to thank my loving family for all their assistance over the course of these two years. For being there to listen to my concerns and always being a steady driving force to remind me that I can achieve great things with hard work. Without your influence and support I do not think this thesis would be possible.

Table of Contents

Abstract.....	I
Acknowledgements.....	III
Table of Contents.....	IV
List of Figures.....	VIII
List of Tables.....	XI
List of Abbreviations and Symbols.....	XII
Chapter 1. Introduction.....	1
1.1. Overview.....	1
1.2. Proteins expressed in the HPV genome.....	1
1.3. Protein E6 protein interactions leading to cell immortalization.....	2
1.4. Specific E6 protein variants lead to increased cancer susceptibility.....	3
1.5. The process of developing a molecular probe.....	4
1.6. Project Aims and Scope.....	6
1.7. References.....	7
Chapter 2. Background and Significance.....	11
2.1. Current clinical procedures.....	11
2.2. Molecular Genotyping for High Risk HPV Types.....	12
2.3. HPV16 E6 Detection.....	14
2.4. Antibodies against E6.....	15
2.5. Small Organic Molecules that bind specifically to E6.....	15
2.6. Biochemical Assay for HPV16 E6 Protein-Protein Interactions.....	17
2.7. HPV Protein E6 and Available Structures.....	17

2.8. Statement of Significance	22
2.9. References.....	22
Chapter 3. Methodology and Rational	25
3.1. COMPUTATIONAL APPROACHES	25
3.1.1. Molecular Docking	25
3.1.2. Virtual Ligand Screening (VLS).....	26
3.1.3. HierVLS and Scoring Functions.....	26
3.2. EXPERIMENTAL APPROACHES.....	28
3.2.1. Overview.....	28
3.2.2. Tryptophan Fluorescence.....	29
3.2.3. Fluorescence Polarization	30
3.2.4. Dose-Response Curves and Curve Fitting	31
3.3. Cell Based Methods	32
3.4. References:.....	33
Chapter 4. Materials and Procedures	37
4.1. Materials	37
4.2. Computational Screening.....	37
4.2.2. N-Terminal (PDB 2LJX) Docking.....	39
4.2.3. C-terminus Screening (2FK4).....	42
4.3. Tryptophan FluoresCense Experiment	45
4.4. Characterization of AFTD	46

4.4.1. Spectroscopy Characterization of AFTD	46
4.4.2. Standard Curve and AFTD Equilibrium	47
4.5. Fluorescence polarization concentration curve to determine EC50	48
4.6. Cell Based Assays	49
4.6.1. Cell culturing	49
4.6.2. Cell Viability.....	50
4.6.3. AFTD Cell Permeability	51
4.7. 6F4 Antibody Competition Experiment.....	52
4.8. O-succinyl-L-homoserine Conjugate.....	53
4.8.1. Synthesis of O-succinyl-L-homoserine-Bodipy	53
4.8.2. Spectroscopic Characterization of O-succinyl-L-homoserine-Bodipy	54
4.8.3. Fluorescence Polarization Assay	54
4.9. References.....	55
Chapter 5. Results and Discussion.....	57
5.1. Computational Screening.....	57
5.1.1. N-Terminus Computational Screening (2LJX).....	58
5.1.2. C-Terminus Computational Screening (2FJK).....	65
5.2. intrinsic Tryptophan Fluorescence Experiment.....	72
5.3. Characterization of AFTD	77
5.3.1. Spectroscopic Characterization of AFTD.....	77

5.3.2. Standard Curve for AFTD	78
5.3.3. Equilibrium Test for AFTD	79
5.4. Fluorescence Polarization and Ec50 determination.....	80
5.5. Cell Based Experiments.....	82
5.5.1. Cell viability assay.....	82
5.5.2. AFTD Cell Permeability Experiment	85
5.6. 6F4 Antibody competition assay	88
5.7. Characterization of O-succinyl-L-homoserine-BODIPY	92
5.8. O-succinyl-L-homoserine conjugate.....	94
5.9. References.....	96
Chapter 6. Conclusion and Future Work	97
Appendix A.....	100
Appendix B.....	122

List of Figures

Figure 1: Ubiquitination of p53 mediated through the presence of E6

Figure 2: Desired characteristics of molecular probe

Figure 3: Specific aims associated to the identification and characterization of potential molecular probes for HPV16 E6

Figure 4: Flow chart displaying the current clinical methods used in HPV detection

Figure 5: Amino acid sequence of E6 prototype highlighting the epitope (yellow) recognized by antibody 6F4

Figure 6: Amino acid sequence of E6 prototype highlighting tryptophan

Figure 7: Structure of C-terminus HPV16 E6 (2FK4)

Figure 8: Amino acid sequence of E6 prototype highlighting the sequence corresponding to the structure with PDB code 2FK4 (red)

Figure 9: Structure of N-terminus HPV16 E6 (2LJX)

Figure 10: Amino acid sequence of E6 prototype highlighting PDB 2LJX (red)

Figure 11: Full length HPV16 E6 structure (PDB 4GIZ)

Figure 12: Amino acid sequence of E6 prototype highlighting PDB 4GIZ (red)

Figure 13: Steps of HierVLS

Figure 14: Step by step process used for the selection of ligands for C-terminus domain.

Figure 15: Full surface structure of 2LJX N-terminal structure of E6

Figure 16: Scatter plot of results from the computational screening of NT-E6 (PDB code 2LJX) for binding region 1 (R1).

Figure 17: Scatter plot of results from computational screening of 2LJX for R2.

Figure 18: Ligand interaction diagram of 3-Chloro-2-(2-[(3-oxo-2-Benzfuran-1 (3H)-Yliden] methyl)Hydrazino)-5-(Trifluoromethyl) Pyridinium Acetate bound to N-terminal E6 (NT-E6). The ligand interaction diagram for 3-Chloro-2-(3-[1-(Phenyl sulfonyl)-(H-pyrazol-3-YL]

Figure 19: Ligand interaction diagram of 3-Chloro-2-(3-[1-(Phenyl sulfonyl)-(H-pyrazol-3-YL] Phenoxy)-5-(Trifluoromethyl) Pyridine bound to N-terminal E6 (NT-E6)

Figure 20: Ligand interaction diagram of N-(2,4-dinitrophenyl)-L-arginine to N-terminal E6 (NT-E6)

Figure 21: Surface representation of Protein E6 C-terminal domain (2FK4).

Figure 22: The Ramachandran plot of HPV16 E6 C-terminal domain (PDB 2FK4) The plot was generated using MOE 2010.10.

Figure 23: Chart representing statistical analysis of 3339 compounds. Used to determine the threshold (-35kcal/mol).

Figure 24: Ligand interaction diagram of 3-Amino-5-fluorobenzo [E] [1,2,4] Triazine 1, 4 dioxide (AFTD) bound to C-terminal E6 (CT-E6).

Figure 25: Ligand-protein interactions diagram of O-succinyl-L-homoserine bound to C-terminal (CT-E6).

Figure 26: Ligand-protein interactions diagram of Paclitaxel bound to C-terminal E6 (CT-E6)

Figure 27: Ligand-protein interactions diagram of Aminopterin bound to C-terminal E6 (CT-E6)

Figure 28: Change in intrinsic tryptophan fluorescence at 350nm induced by 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) after 30minutes incubation in the presence of E6. A) Fluorescence intensity B) Bar graph %change C) log scale % change plot of AFTD 300µM (black) and protein alone (red).

Figure 29: Change in intrinsic tryptophan fluorescence at 350nm induced by O-succinyl-L-homoserine after 30minutes incubation in the presence of E6. A) Fluorescence intensity B) Bar graph %change C) log scale % change D) plot of O-succinyl-L-homoserine 300µM (black) and protein alone (red).

Figure 30: Change in intrinsic tryptophan fluorescence at 350nm induced by Paclitaxel after 30minutes incubation in the presence of E6.

Figure 31: Excitation/Emission spectrum of 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (Fluoro).

Figure 32: Standard curve for 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide AFTD.

Figure 33: Time plot of 3-Amino-5-Fluorobenzo [1,4] Triazine dioxide and E6 complex.

Figure 34: FP curve of E6 (0.04mg/ml) incubated with varying the concentration of 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) (1 μ M-300 μ M).

Figure 35: CaSki cells after incubation with AFTD (25 μ M) for 24hrs.

Figure 36: CaSki cells incubated with ethanol (0.01%) for 24hrs.

Figure 37: C33A cells incubated with ethanol (0.01%). In this solvent control, the cells are alive (not floating), indicating that it is AFTD and not ethanol that is causing cell death.

Figure 38: C33A cell lines after incubation with AFTD (25 μ M) for 24hrs. Most of the cells are still alive as they are still attached to the plate.

Figure 39: Fluorescence intensity (FI) at 2 hour (A) and 3 hour (B) incubation.

Figure 40: Time curve showing the change in polarization of the protein and ligand complex and the antibody, protein ligand complex.

Figure 41: Fluorescence polarization bar graph varying the antibody concentration.

Figure 42: Plot of the fluorescence polarization versus log of the concentration (μ g/ml) of 6F4 antibody. A) Concentration scale and B) log scale.

Figure 43: Emission and excitation plot of O-succinyl-L-homoserine-BODIPY.

Figure 44: Linear fluorescence intensity plot of O-succinyl-L-homoserine-BODIPY (300 μ M, 100 μ M, 10 μ M, and 1 μ M) and protein E6 (0.04mg/ml).

Figure 45: Fluorescence polarization time plot of O-succinyl-L-homoserine-BODIPY (100 μ M, 10 μ M, and 1 μ M) and protein E6 (0.04mg/ml).

Figure 46: Bar graph of polarization (mP) at 60minutes for the O-succinyl-BODIPY binding experiment.

List of Tables

Table 1: Molecular methods used for HPV genotyping to date

Table 2: Zinc-finger ejecting molecules tested experimentally

Table 3: All conformation saved in the PDB file (2LJX). Shows that confirmation 18 has the lowest potential energy

Table 4: Summary table for 2LJX Screening

Table 5: Table of all compounds that based the R1 threshold of -53.12 kcal/mole.

Table 6: Compounds that passed the threshold of -58.20kcal/molHierVLS targeting regions R2 in NT-E6.

Table 7: Promising lead compounds from the NT-E6 virtual screening.

Table 8: Summary of results for CT-E6 Screening

Table 9: Top three compounds from the primary amine database in order of rank.

Table 10: Top 5 compounds from the fluorine atom containing molecule database in order of rank.

Table 11: Ligands selected for experimental testing. All four ligands were selected based on binding scores, ligand interactions to the protein, and commercial availability.

List of Abbreviations and Symbols

Human Papillomavirus (HPV)

Long Control Regions (LCR)

Retinoblastoma Protein (pRb),

E6 Associated Protein (E6AP)

Human Papillomavirus (HPV)

Cervical Intraepithelial Neoplasia (CIN)

High Intensity Focused Ultrasound (HIFU)

National Cancer Institute (NCI)

Glutathione-S-transferase (GST)

Glutathione (GSH)

Nuclear Magnetic Resonance (NMR).

Hierarchical Virtual Ligand Screening (HierVLS)

Analytical Volume Generalized Borne (AVGB)

3-amino-5-fluorobenzo [e] [1,2,4] triazine 1, 4 dioxide (AFTD)

High Performance Liquid Chromatography (HPLC)

Aminopterin (ALX-440-041-M010)

Dulbecco's Modification of Eagles Medium (DMEM)

Biograph Files (BGF)

Molecular Operating Environment (MOE)

Shared Hierarchical Research Computing Network (SHARCNET)

Phosphate Buffered Saline (PBS).

Virtual Ligand Screening (VLS)

Thunder Bay Regional Health Sciences Center (TBRHSC)

Position Emission Tomography (PET)

dimethylformamide (DMF)

Thin layer chromatography, (TLC)

Dimethyl sulfoxide, (DMSO)

Chapter 1.Introduction

1.1.OVERVIEW

Human Papillomavirus (HPV) is a double stranded DNA virus¹. To date this virus has more than 150 types known. Some types are considered high risk and some low risk with regards to likelihood of cancer development³. High risk HPVs can cause cancerous lesions, whereas low risk HPVs is associated with benign lesions⁴. This relationship was first described by Dr Harold zurHausen and collaborators during the 1980s. This was the first time that HPV DNA was isolated and discovered in a cervical tumor biopsy⁵. It is estimated that about 99.7% of cervical cancer cases are HPV positive⁶. The majority of these cases are shown to be HPV16 (50-70%)⁷. More recent numbers have indicated a more exact number of 54.4% of cervical cancer cases are linked to HPV16⁸. HPV not only has a link to cervical cancer but it is estimated that 90% of anogenital cancers are HPV positive⁸⁻⁹. HPV is also associated with head and neck cancers and non-melanoma skin cancer⁸⁻⁹.

1.2.PROTEINS EXPRESSED IN THE HPV GENOME

Long control regions (LCR) in the HPV genome encode six proteins (E1, E2, E4, E5, E6, E7, L1 and L2), which are expressed either early or late in the infection. L1 and L2 are expressed later in the infection and are found within the upper layers of the infected epithelium¹⁰. The early proteins (E1, E2, E4, E5, E6, and E7) are found in the lower to mid epithelial layers¹¹. E1 and E2 are responsible for the initiation of viral replication and control gene expression¹¹. E4 and E5 proteins aid in genome amplification¹¹. E6 and E7 are said to be oncoproteins; both play a role in cell immortalization¹².

E7 is known for its interaction with retinoblastoma protein (pRb), which leads to the release of transcription factor E2F which causes entry into S phase of the cell cycle¹³. This will lead to increased cell proliferation¹³. E6 plays a critical role in cell immortalization by degrading tumor suppressing protein p53¹⁴.

1.3.PROTEIN E6 PROTEIN INTERACTIONS LEADING TO CELL IMMORTALIZATION

Protein E6 is known as an oncoprotein due to its role in cancer development. E6 is shown to lead to the ubiquitination of p53¹⁵. The process of ubiquitination requires a protein ligase to transfer the high-energy thioester bond to the protein in order to transfer ubiquitin¹⁶. This process is facilitated through the binding of E6 and E6 associated protein (E6AP)¹⁷. This complex will work as an E3 ligase and will transfer the thioester bond to protein p53 leading to its ubiquitination (Figure 1)¹⁷. The presence of ubiquitin will signal the proteosomes to come and degrade the protein¹⁶. The presence of E6 can also suppress p53 function in an E6AP independent manner. This is through its interaction with p300/CBP complex¹⁸. Protein p53 is a tumor suppressor, important in the prevention of cancer. It functions as the guardian of the cell¹⁹. Protein p53 keeps the cell cycle regulated and it is involved in DNA repair¹⁹. The levels of p53 are regulated in the cell, and generally, p53 has a short half-life²⁰. Levels of p53 are also shown to increase upon any detected DNA damage, which can lead to mutations¹⁹. p53 will activate cell cycle arrest in these situation leading to potential apoptosis¹⁹. Low levels of p53 have been documented in almost all carcinomas¹⁹. This is not a surprise as destruction of p53 increases cancer susceptibility. In the absence of p53, cell division is no longer properly controlled and DNA repair mechanisms are depleted¹⁹.

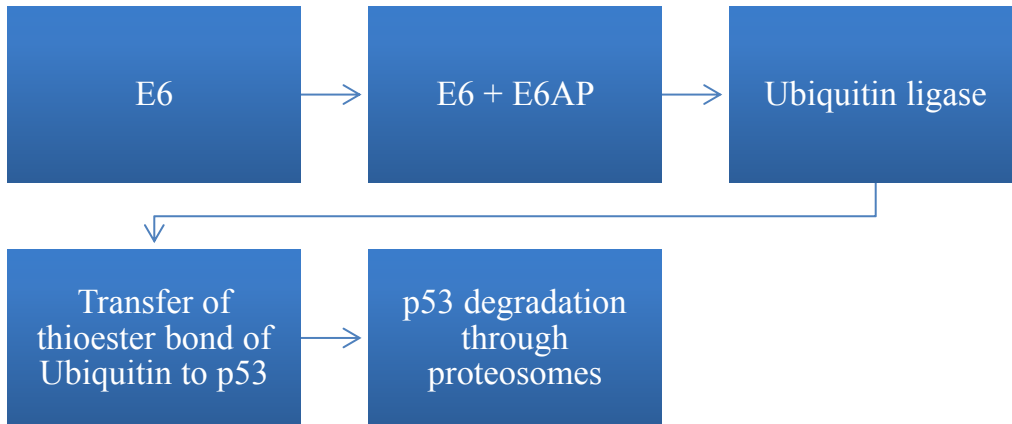


Figure 1: Ubiquitination of p53 mediated through the presence of E6. Flow chart shows the process leading the degradation of p53 in a series of steps. Certain variants of E6 bind to E6AP in such a way to allow this process to occur. The majority of these E6 variants are expressed in HPV type 16.

E6 is also shown to interact with other proteins such as MAGI-1²¹, MAGI-2, MAGI3, hScribble²², and hDIg²³. These PDZ contain domain-containing proteins are responsible for cell signalling and cell polarity, leading to cellular transformation to a malignant phenotype²². The presence of E6 also upregulates hTERT expression, which leads to E6 induced telomerase activity²⁴. This increased telomerase activity will protect chromosomes from damage during cell division²⁵.

1.4.SPECIFIC E6 PROTEIN VARIANTS LEAD TO INCREASED CANCER SUSCEPTIBILITY

Specific changes from the prototype E6 have been shown to increase the risk for the development of cancer⁷. Some high-risk E6 variants include: Q14H/H78Y/L83V and R10G/L83V⁷. Although all E6 proteins bind to E6AP it is believed that certain mutants bind in such a way to facilitate the interaction with p53²⁶. Much of these variants of E6 are associated

with HPV16⁷. In addition, recent studies have indicated that these high-risk variants of E6 alone can lead to cell immortalization and increased susceptibility to cancer development²⁷. This makes E6 an ideal biomarker for high-risk HPV. This creates an excellent opportunity to find a molecular probe that is specific for HPV16 E6, in particular those variants associated with cancer development. Prior to looking at high-risk variants, methods will have to be employed on the prototype in order to obtain proof of concept. An E6-specific molecular probe could serve two purposes: it could be used as a tool for diagnostic imaging for early intervention, or it could be used as a molecular probe for biochemical assays involving E6 protein interactions.

1.5.THE PROCESS OF DEVELOPING A MOLECULAR PROBE

Molecular probes aid in the visualization, characterization, and measurement of biological processes²⁸. A molecular probe can be defined as an agent that has the ability to characterize and quantify biological functions²⁹. Molecular probes can be target specific or non-specific³⁰. A specific probe is a probe that only binds to one particular biomarker³⁰. Non-specific probes do not bind to any specific target or have any biochemical activity in-vivo³⁰. Instead, they make use of high capacity systems such as blood and kidneys to provide image contrast.

Molecular imaging probes assist medical imaging techniques and may bind to a specific biomarker, usually a protein. Advantages associated with the use of molecular probes include improved prognosis and improved staging³¹. Improved staging is the ability to determine the stage of the specific disease. This leads to a more specific treatment regime. These advantages come from the probe's ability to specifically identify the biomarker at an early stage of the disease³. In order to obtain a probe with such advantages many steps need to be taken. The characteristics required to develop a molecular imaging probe include: high binding affinity for

target, high specificity to target, high sensitivity, high contrast ratio, high stability in vivo, low immunogenicity and toxicity, production and economic feasibility (Figure 2)³¹.



Figure 2: Desired characteristics of molecular probe. A molecular probe candidate must be evaluated for these characteristics. If the molecule does not pass one of these requirements, modifications to its structure may be introduced or an entirely new molecule may have to be considered.

In order to be marketed as a molecular probe, the compound must be shown to have high affinity for the target. This molecule must bind quickly to the target and be slow to unbind³¹. As it is to be used in diagnostic imaging it is critical that it accumulates in the target tissue. In order to acquire a well-defined image, non-specific binding must be avoided³¹. Additionally, off-target binding must be avoided in order to reduce false positive rates³¹. It is for these reasons that affinity and specificity are crucial in the process of molecular probe development. Once affinity and specificity have been assessed experimentally, the probe must be shown to have high sensitivity, meaning that a small amount of molecular probe is needed to produce an image³¹. Images obtained with the probe must have high contrast ratio, thus the signal to noise ratio must

be low³¹. Upon showing its potential to produce a high quality images, the probe must be shown to be stable *in vivo*. Due to the large amount of enzymes present in an organism, a probe could be easily degraded once placed *in vivo*. If the probe is easily degraded then it is not of much use³¹. Once all requirements have been met the compound can then be used as a molecular probe.

1.6.PROJECT AIMS AND SCOPE

The goal of this project is to identify and characterize molecular probe(s) for prototype HPV16 E6. In searching for a molecular probe, the requirements discussed earlier will be taken into consideration (Figure 2). For the scope of this thesis and the time frame allotted, we believe that we can identify probe candidates computationally, and evaluate both affinity and specificity to HPV16 E6 experimentally. Initial affinity to HPV16 E6 C-terminus, will be evaluated using computer software. Databases will be developed that contain compounds that are suitable for use as a molecular probe. Once docked these compounds will be further evaluated in order to determine, which compounds are worth testing experimentally.

Upon finding probe candidates, experimental ligand binding assays will take place in order to further validate the affinity to target HPV16 E6. Once affinity has been evaluated, the probe candidate can be placed in the cell environment in order to test specificity to HPV16 E6.

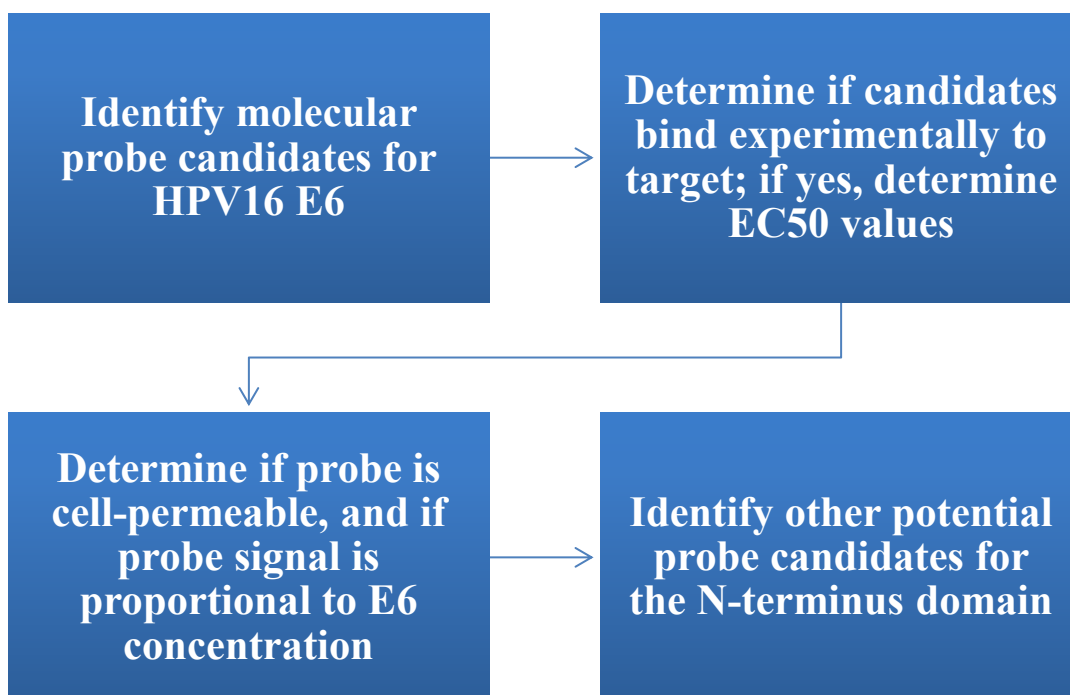


Figure 3: Specific aims associated to the identification and characterization of potential molecular probes for HPV16 E6.

Successfully achieving these aims will provide proof of concept for the development of a molecular imaging probe that is specific to those variants of HPV16 E6 shown to increase cancer susceptibility (Figure 3). Such probes will enable a specific test for high-risk HPV, thus facilitating diagnostics and early intervention in HPV-associated cancers.

1.7. REFERENCES

1. Fernandez, A. F.; Rosales, C.; Lopez-Nieva, P.; Graña, O.; Ballestar, E.; Ropero, S.; Espada, J.; Melo, S. A.; Lujambio, A.; Fraga, M. F.; Pino, I.; Javierre, B.; Carmona, F. J.; Acquadro, F.; Steenbergen, R. D. M.; Snijders, P. J. F.; Meijer, C. J.; Pineau, P.; Dejean, A.; Lloveras, B.; Capella, G.; Quer, J.; Buti, M.; Esteban, J.-I.; Allende, H.; Rodriguez-Frias, F.; Castellsague, X.; Minarovits, J.; Ponce, J.; Capello, D.; Gaidano, G.; Cigudosa, J. C.; Gomez-Lopez, G.; Pisano, D. G.; Valencia, A.; Piris, M. A.; Bosch, F. X.; Cahir-McFarland, E.; Kieff, E.; Esteller, M., The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer. *Genome Research* **2009**, *19* (3), 438-451.

2. <http://pave.niaid.nih.gov/#home>.
3. Zehbe, I.; Tachezy, R.; Mytilineos, J.; Voglino, G.; Mikyškova, I.; Delius, H.; Marongiu, A.; Gissmann, L.; Wilander, E.; Tommasino, M., Human papillomavirus 16 E6 polymorphisms in cervical lesions from different European populations and their correlation with human leukocyte antigen class II haplotypes. *International Journal of Cancer* **2001**, *94* (5), 711-716.
4. Quint, W.; Jenkins, D.; Molijn, A.; Struijk, L.; van de Sandt, M.; Doorbar, J.; Mols, J.; Van Hoof, C.; Hardt, K.; Struyf, F.; Colau, B., One virus, one lesion—individual components of CIN lesions contain a specific HPV type. *The Journal of Pathology* **2012**, *227* (1), 62-71.
5. Dürst, M.; Gissmann, L.; Ikenberg, H.; zur Hausen, H., A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. *Proceedings of the National Academy of Sciences* **1983**, *80* (12), 3812-3815.
6. Walboomers, J. M. M.; Jacobs, M. V.; Manos, M. M.; Bosch, F. X.; Kummer, J. A.; Shah, K. V.; Snijders, P. J. F.; Peto, J.; Meijer, C. J. L. M.; Muñoz, N., Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of Pathology* **1999**, *189* (1), 12-19.
7. Zehbe, I.; Richard, C.; DeCarlo, C. A.; Shai, A.; Lambert, P. F.; Lichtig, H.; Tommasino, M.; Sherman, L., Human papillomavirus 16 E6 variants differ in their dysregulation of human keratinocyte differentiation and apoptosis. *Virology* **2009**, *383* (1), 69-77.
8. Crow, J. M., HPV: The global burden. *Nature* **2012**, *488* (7413), S2-S3.
9. Forman, D.; de Martel, C.; Lacey, C. J.; Soerjomataram, I.; Lortet-Tieulent, J.; Bruni, L.; Vignat, J.; Ferlay, J.; Bray, F.; Plummer, M.; Franceschi, S., Global Burden of Human Papillomavirus and Related Diseases. *Vaccine* **2012**, *30*, F12-F23.
10. Doorbar, J., The papillomavirus life cycle. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **2005**, *32*, 7-15.
11. Doorbar, J.; Quint, W.; Banks, L.; Bravo, I. G.; Stoler, M.; Broker, T. R.; Stanley, M. A., The Biology and Life-Cycle of Human Papillomaviruses. *Vaccine* **2012**, *30*, F55-F70.
12. Münger, K.; Phelps, W. C.; Bubb, V.; Howley, P. M.; Schlegel, R., The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes. *Journal of Virology* **1989**, *63* (10), 4417-4421.

13. Chellappan, S.; Kraus, V. B.; Kroger, B.; Munger, K.; Howley, P. M.; Phelps, W. C.; Nevins, J. R., Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. *Proceedings of the National Academy of Sciences* **1992**, *89* (10), 4549-4553.
14. Kesis, T. D.; Slebos, R. J.; Nelson, W. G.; Kastan, M. B.; Plunkett, B. S.; Han, S. M.; Lorincz, A. T.; Hedrick, L.; Cho, K. R., Human papillomavirus 16 E6 expression disrupts the p53-mediated cellular response to DNA damage. *Proceedings of the National Academy of Sciences* **1993**, *90* (9), 3988-3992.
15. Zehbe, I.; Wilander, E.; Delius, H.; Tommasino, M., Human Papillomavirus 16 E6 Variants Are More Prevalent in Invasive Cervical Carcinoma than the Prototype. *Cancer Research* **1998**, *58* (4), 829-833.
16. Lowe, E. L.; Doherty, T. M.; Karahashi, H.; Ardit, M., Review: Ubiquitination and de-ubiquitination: role in regulation of signaling by Toll-like receptors. *Journal of Endotoxin Research* **2006**, *12* (6), 337-345.
17. Nakagawa, S.; Huibregtse, J. M., Human Scribble (Vartul) Is Targeted for Ubiquitin-Mediated Degradation by the High-Risk Papillomavirus E6 Proteins and the E6AP Ubiquitin-Protein Ligase. *Molecular and Cellular Biology* **2000**, *20* (21), 8244-8253.
18. Zimmermann, H.; Degenkolbe, R.; Bernard, H.-U.; O'Connor, M. J., The Human Papillomavirus Type 16 E6 Oncoprotein Can Down-Regulate p53 Activity by Targeting the Transcriptional Coactivator CBP/p300. *Journal of Virology* **1999**, *73* (8), 6209-6219.
19. Ryan, K. M.; Phillips, A. C.; Vousden, K. H., Regulation and function of the p53 tumor suppressor protein. *Current Opinion in Cell Biology* **2001**, *13* (3), 332-337.
20. Kastan, M. B.; Onyekwere, O.; Sidransky, D.; Vogelstein, B.; Craig, R. W., Participation of p53 Protein in the Cellular Response to DNA Damage. *Cancer Research* **1991**, *51* (23 Part 1), 6304-6311.
21. Kranjec, C.; Banks, L., A Systematic Analysis of Human Papillomavirus (HPV) E6 PDZ Substrates Identifies MAGI-1 as a Major Target of HPV Type 16 (HPV-16) and HPV-18 Whose Loss Accompanies Disruption of Tight Junctions. *Journal of Virology* **2011**, *85* (4), 1757-1764.

22. Massimi, P.; Gammoh, N.; Thomas, M.; Banks, L., HPV E6 specifically targets different cellular pools of its PDZ domain-containing tumour suppressor substrates for proteasome-mediated degradation. *Oncogene* **2004**, *23* (49), 8033-8039.
23. Grm, H. S.; Banks, L., Degradation of hDlg and MAGIs by human papillomavirus E6 is E6-AP-independent. *Journal of General Virology* **2004**, *85* (10), 2815-2819.
24. Veldman, T.; Horikawa, I.; Barrett, J. C.; Schlegel, R., Transcriptional Activation of the Telomerase hTERT Gene by Human Papillomavirus Type 16 E6 Oncoprotein. *Journal of Virology* **2001**, *75* (9), 4467-4472.
25. Veldman, T.; Liu, X. F.; Yuan, H.; Schlegel, R., Human papillomavirus E6 and Myc proteins associate in vivo and bind to and cooperatively activate the telomerase reverse transcriptase promoter. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100* (14), 8211-8216.
26. Thomas, M.; Tomaić, V.; Pim, D.; Myers, M. P.; Tommasino, M.; Banks, L., Interactions between E6AP and E6 proteins from alpha and beta HPV types. *Virology* **2013**, *435* (2), 357-362.
27. Niccoli, S.; Abraham, S.; Richard, C.; Zehbe, I., The Asian-American E6 Variant Protein of Human Papillomavirus 16 Alone Is Sufficient To Promote Immortalization, Transformation, and Migration of Primary Human Foreskin Keratinocytes. *Journal of Virology* **2012**, *86* (22), 12384-12396.
28. Weissleder, R.; Mahmood, U., Molecular Imaging1. *Radiology* **2001**, *219* (2), 316-333.
29. Massoud, T. F.; Gambhir, S. S., Molecular imaging in living subjects: seeing fundamental biological processes in a new light. *Genes & Development* **2003**, *17* (5), 545-580.
30. Chopra, A.; Shan, L.; Eckelman, W. C.; Leung, K.; Latterner, M.; Bryant, S.; Menkens, A., Molecular Imaging and Contrast Agent Database (MICAD): Evolution and Progress. *Mol Imaging Biol* **2012**, *14* (1), 4-13.
31. Chen, K. C. a. X., Design and Development of Molecular Imaging Probes *Current Topics in Medicinal Chemistry* **2010** *10* (12), 1227-1236.

Chapter 2. Background and Significance

2.1. CURRENT CLINICAL PROCEDURES

Human Papillomavirus (HPV) and its association with various cancers has become a large concern within health care. HPV is now one of the most common sexually transmitted disease¹. HPV is divided into several types, with HPV 16 and 18 being the highest risk for cancer development². It is estimated that 68% of cervical cancer cases have HPV16 or HPV18 genotype³. Due to this association specific screening methods for these high-risk HPV types need to be employed. Current screening methods, such as pap smears and cytological screening, have resulted in decreased incidence of invasive cervical cancer⁴. However, HPV remains a major burden on health care systems around the world⁴.

Current burdens are over or under treatment of women with HPV infection⁴. This is due in large part to the lack of specificity of current screening methods⁴. The non-specific nature of screening leads to over/under treatment⁴. This issue could be quickly fixed with the use of a specific biomarker to properly evaluate the high-risk nature of the virus in a clinical setting. This biomarker could potentially reduce much of the current burden to the health care community.

Current methods used in HPV detection include: 1) Women with cytological abnormalities are brought in for additional screening 2) follow-up appointment for women with abnormal screening results 3) Cervical Intraepithelial Neoplasia (CIN) is treated following a risk assessment; 4) HPV DNA testing is conducted, usually in combination with a Pap smear, to look for cervical-cancer precursors; 5) Information on the persistence of specific HPV type is acquired (Figure 4)⁵.

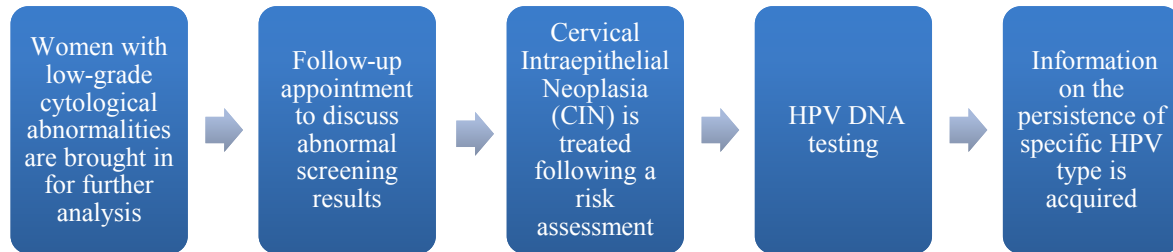


Figure 4: Flow chart displaying the current clinical methods used in HPV detection. This is the rational used to detect HPV. Although these tests have alleviated some of the burdens that HPV creates, more specific tests are needed that consider the proteome as opposed to the genome.

In addition to these methods vaccines are available for HPV infection. There is a bivalent vaccine (type 16 and 18) and a quadrivalent (type 6, 11, 16, and 18)⁶. The vaccination against type 16 and 18 is thought to be able to lower the incidence of cervical cancer by more than two thirds⁵. Broad spectrum HPV vaccines are currently in development⁵. However, such vaccines are only available in certain parts of the world. This creates a need for highly automated, inexpensive screening method that can be used in all parts of the world.

2.2.MOLECULAR GENOTYPING FOR HIGH RISK HPV TYPES

The current methods for HPV detection rely heavily on molecular testing in order to identify the type of HPV present in the sample⁵. Current molecular methods for genotyping are presented in Table 1⁵.

Table 1: Molecular methods used for HPV genotyping to date

Nucleic Acid Hybridization Assays	Signal Amplification Assays	Nucleic Acids Amplification Assays
Southern blot	Cervista HPV	Microarray
In situ hybridization	Hybrid capture 2 (he2)	PapilloCheck
Dot blot hybridization		PCR, PRC RFLP
		Real-time PCR
		Abboott Real-time
		Cobas 4800 HPV
		Genome sequencing
		CLART HPV-2
		Inno-LiPA
		The linear array
		Clinical Arrays HPV
		MCHA
		PreTeck proofer
		APTIMA HPV assay

All of these assays have advantages and disadvantages associated with them. The biggest disadvantage of these methods is the lack of automation and the price⁵. In an attempt to overcome these disadvantage a study was conducted to genotype HPV infection using PCR from urine samples⁷. This paper was able to conclude by comparing results collected from cervical brush technique that HPV DNA can be accurately detected in urine samples⁷. These positive samples were then genotypes using RFLP typing⁷. This leads to another disadvantage of such molecular techniques. Genotyping a HPV infection, however important, is not always going to

be indicative of a cancer-causing HPV infection⁵. This is the advantage of using specific disease biomarkers such as HPV proteins E6 and E7⁸. These proteins are important in specific detection of HPV because certain variants of these proteins are associated with higher cancer development risk⁸. Even if molecular tests are done and the HPV infection is declared type 16 it does not mean that these oncoprotein variants are present. Research has shown that certain variants of E6 can cause cell immortalization independent of E7⁹. This is due to E6s many interactions with other proteins including E6AP¹⁰. Although all E6 proteins bind protein E6AP, certain variants are shown to interact in such a way to make this possible¹⁰. It is due to this complexity that specific testing for E6 needs to be implemented in current clinical practise.

2.3.HPV16 E6 DETECTION

Since E6 has been well defined as a biomarker for high-risk HPV⁹, a lot of work is being done in order to detect this protein. This is because of its major role in cell immortalization⁹. One way by which this protein can be detected is through the use of antibodies. However, since these antibodies are not cell permeable techniques such as High Intensity Focused Ultrasound (HIFU) need to be employed for intercellular screening¹¹. Studies conducted showed proof of concept using a combination of HIFU and microbubbles for the delivery of antibodies¹¹. The antibodies used showed some therapeutic effects¹¹. This suggests the potential of using antibodies for intercellular delivery using such methods. Antibodies can also be delivered inside the cell, using transfection reagents; however these reagents are toxic¹¹. For this reason these reagents are only useful for *in-vitro* work. Being able to deliver an antibody to E6 can potentially be difficult as E6 mainly resides in the nucleus¹². Using a small organic molecule would be very beneficial for specific HPV16 E6 detection, even though small molecule interaction with proteins can be complex¹³. However, small organic molecules are more likely to be cell permeable than larger

molecules such as antibodies. These interactions are complex due to the high number of conformations that the molecule and protein can take¹³. As a result when selecting a small molecule as a potential ligand, conformational changes of the ligand need to be taken into account in order to make sure the binding event is properly scored.

An additional monoclonal antibody method involving test strips is also available and recently described. This is an *in vitro* method for specific E6 detection and has been shown to identify high risk HPV with 95% confidence¹⁴. Two monoclonal antibodies are used one specific for type 16 and the other for type 18 and 45¹⁴. This test gives great qualitative measure as it is shown to be able to detect high risk E6 variants.

2.4.ANTIBODIES AGAINST E6

Monoclonal antibodies are available for the N-terminus of E6. Monoclonal Antibody 6F4 was first developed by Giovane and colleagues¹⁵. It is a mouse monoclonal antibody specific to HPV16 E6. The epitope region was described by Mason and colleagues in 2003 is as following: F/Y X X P/L X X R (Figure 5)¹⁶.

MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV
YRDGNPYAVCDKCLKFYISKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE
EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL

Figure 5: Amino acid sequence of E6 prototype highlighting the epitope (yellow) recognized by antibody 6F4. This sequence was taken from GenBank (GenBank #AAA46939.1).

2.5.SMALL ORGANIC MOLECULES THAT BIND SPECIFICALLY TO E6

Little success has been made in the attempt to find a small organic molecule that binds specifically to E6. Some attempts have been made using zinc-finger ejecting molecules¹⁷. A

pharmacophore model was developed in order to target the zinc-finger domain to inhibit the binding of E6AP¹⁷. The reason for targeting E6AP binding is that E6 will not interact with p53 unless bound to E6AP¹⁷. Through screening two databases, one from the National Cancer Institute (NCI) and that other from Sigma Aldrich¹⁷, 10 molecules were selected for testing and IC50 values were experimentally determined (Table 2)¹⁷.

Table 2: Zinc-finger ejecting molecules tested experimentally. Two data bases were screened: one from the national cancer institute and the other was rare compounds available through Sigma Aldrich. A pharmacophore model was used in order to find compounds that would bind to a particular region in order to inhibit binding to E6AP. It was determined that none of the compounds generated an IC50 that was low enough to be used as a therapeutic drug.

Compound name	E6AP binding IC50 (μM)
National Cancer Institute (NCI) 83143	29
NCI 117907	29
NCI 135098	22
NCI 216029	12
Sigma Aldrich (SA) s32701	52
SA 207721	21
SA r218634	27
SA r225975	12
SA r278319	17
SA s204102	11

Also when using controls for non-specific inhibition it was determined that the inhibition may have been non-specific¹⁷. Cell permeability tests were also conducted with NCI 117907, which was intrinsically fluorescent¹⁷. This compound showed little to no uptake at 100μM¹⁷. This result along with the non-specific nature of the binding interactions led to the conclusion that the

pharmacophore model failed. It is suggested that a better model could be made once a series of known compounds that bind to E6 are discovered¹⁷.

To date there is no small organic molecule that is known to bind to HPV16 E6 with enough affinity to be marketed as a molecular probe. This makes our work extremely novel in the field of HPV detection.

2.6.BIOCHEMICAL ASSAY FOR HPV16 E6 PROTEIN-PROTEIN INTERACTIONS

Currently, there is not a biochemical assay available for looking at HPV16 E6 protein-protein interactions outside of the cell environment. Available assays tracking E6 protein-protein interactions take place in the cell environment. The most common is the Glutathione-S-transferase (GST) pull-down assay. In this assay, a protein is fused to GST and expressed in E-coli (bait protein) after which it is bound to glutathione (GSH)- coupled particles¹⁸. This will cause affinity purification of any proteins (prey) it comes into contact with¹⁸. This assay is deemed accurate¹⁸; however testing protein-protein interactions outside of the cell environment still serves an important purpose. Due to the massive amount of interactions inside the cell environment, testing protein-protein interactions outside of the cell environment can provide valuable information without confounding factors that would be present in the cell environment. Also, as the list of E6 protein-protein interactions continues to grow, having a biochemical assay to probe E6 interactions with other proteins will enable better understanding of HPV infection's role in cancer.

2.7.HPV PROTEIN E6 AND AVAILABLE STRUCTURES

Protein E6 is a small protein that is only composed of 151 amino acids (Figure 7)¹⁹. E6 contains two distinct domain an N-terminus and a C-terminus domain, which are both

approximately 75 amino acids in length¹⁹. Both domains contain a zinc binding domain¹⁷. Finding a stable folded form of this protein has been a real challenge and has limited full understanding of the full-length E6 structure¹⁹. This small protein interacts with many proteins in the cell, with the most well-known partner being E6AP. The E6-E6AP complex is responsible for the ubiquitination of tumor suppressor p53¹⁹. The structure of E6 only contains one tryptophan making suitable for intrinsic tryptophan fluorescence experiments (Figure 6).

```
MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREYDFAFRDLCIV  
YRDGNPYAVCDKCLKFYISKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE  
EKQRHLDKKQRFHNIRGRW TGRCMSCCRSSRTRRETQL
```

Figure 6: Amino acid sequence of E6 prototype highlighting tryptophan. This sequence was taken for GenBank (GenBank #AAA46939.1). E6 only contains one tryptophan making it suitable for intrinsic tryptophan fluorescence ligand binding experiments.

Various experimentally determined structures for E6 are currently available through the PDB, which creates an excellent starting point for molecular docking.

At the beginning of this project, structural information for HPV16 E6 was limited as there was only one human HPV16 E6 structure available. This structure represented the C-terminus of HPV16 E6 (PDB 2FK4)¹⁹(Figure 7).

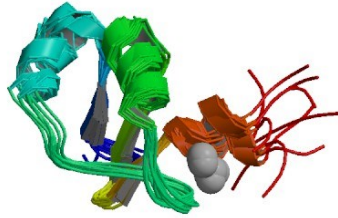


Figure 7: Structure of C-terminus HPV16 E6 (2FK4)¹⁹.

This structure was solved using Nuclear Magnetic Resonance (NMR). This structure had 75 amino acids (Figure 8)¹⁹ and was highly useful for our search for small organic molecules that bind to E6. In these structures some cysteine's were mutated to serine in order to obtain better protein folding for experimental determination¹⁹.

MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIIILECVYCKQQLLRREVYDFAFRDLCIV
 YRDGNPYAVCDKCLKFYISKISEYRH^YYC^SYSLYGT^STLEQQYNKPL^SCDLLIRCINCQKPL^SC^SPE
 EKQRHLDKKQRFHNIRGRWTGRCMSC^SCRSSRTRRETQL

Figure 8: Amino acid sequence of E6 prototype highlighting the sequence corresponding to the structure with PDB code 2FK4 (red). This sequence was taken from GenBank (GenBank #AAA46939.1). The yellow represents mutations from the original sequence primarily cysteine is mutated to serine for protein folding for experimental determination.

In 2012, the same group published the 3D structure of the N-terminal domain of protein HPV16 E6 (PDB 2LJX)²⁰(Figure 9).

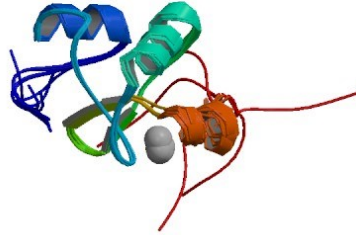


Figure 9: Structure of N-terminus HPV16 E6 (2LJX)²⁰.

This structure contains 84 amino acid residues and has some overlap with that of the C terminus domain (PDB 2FK4) published years prior (Figure 10)²⁰.

```
MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV  
YRDGNPYAVCDKCLKFYISKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE  
EKQRHLDKKQRFHNIRGRWTGRCMSSCRSSRTRRETQL
```

Figure 10: Amino acid sequence of E6 prototype highlighting PDB 2LJX (red). This sequence was taken for GenBank (GenBank #AAA46939.1). The yellow represents mutations from the original sequence primarily cysteine is mutated to serine for protein folding for experimental determination.

This structure was also resolved by NMR. This structure combined with the previous structure (2FK4) represents a complete structure for E6. However, we proceeded to computationally screen each structure individually due to insufficient information to assemble both structural domains together. Putting the full structure together computationally would be difficult and time-consuming, and the resulting structure may have lacked accuracy.

Just this year (2013) for the first time the full structure of HPV16 E6 was published in Science by the same group once again (PDB 4GIZ)²¹ (Figure 11).

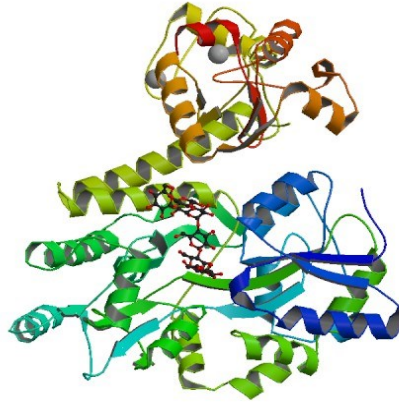


Figure 11: Full length HPV16 E6 structure (PDB 4GIZ)²¹.

This structure was experimentally determined using X-ray crystallography²¹. This structure was solved with bound peptides (LXXLL) isolated for Paxcillin and E6AP²¹. This structure contains 135 amino acids of the total 151 (Figure 12).

```
MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIIIECVYCKQQLLRREVYDFAFRDLCIV  
YRDGNPYAVCDKCLKFYISKISEYRHYCYSLYGTTLQYQNKPLCDLLIRCINCQKPLCPE  
EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
```

Figure 12: Amino acid sequence of E6 prototype highlighting PDB 4GIZ (red). This sequence was taken for GenBank (GenBank #AAA46939.1). The yellow represents mutations from the original sequence primarily cysteine is mutated to serine for protein folding for experimental determination.

A hydrophobic pocket is formed in the structure between the two zinc domains that are connected by a linker helix²¹. This pocket aids in the binding of proteins that contain the

(LXXLL) binding motif²¹. This was a big step forward in understanding E6 structurally and its interactions with other proteins.

2.8.STATEMENT OF SIGNIFICANCE

Finding a molecular probe of HPV16 E6 that can be used for either diagnostic imaging or in a biochemical assay will be very useful in the quest for a further understanding its role in HPV infection and HPV-related cancers. An E6-specific molecular probe suitable for medical imaging may also lay the foundation for new preventive screening and diagnostic tools, which could one day save the lives of many and potentially alleviate some of the burden on our health care system.

2.9.REFERENCES

1. Stanley, M., Pathology and epidemiology of HPV infection in females. *Gynecologic Oncology* **2010**, *117* (2, Supplement), S5-S10.
2. Khan, S.; Jaffer, N. N.; Khan, M. N.; Rai, M. A.; Shafiq, M.; Ali, A.; Pervez, S.; Khan, N.; Aziz, A.; Ali, S. H., Human papillomavirus subtype 16 is common in Pakistani women with cervical carcinoma. *International Journal of Infectious Diseases* **2007**, *11* (4), 313-317.
3. Muñoz, N., Human papillomavirus and cancer: the epidemiological evidence. *Journal of Clinical Virology* **2000**, *19* (1-2), 1-5.
4. Chow, L. T.; Broker, T. R.; Steinberg, B. M., The natural history of human papillomavirus infections of the mucosal epithelia. *APMIS* **2010**, *118* (6-7), 422-449.
5. Abreu, A. L.; Souza, R.; Gimenes, F.; Consolaro, M. E., A review of methods for detect human Papillomavirus infection. *Virology Journal* **2012**, *9* (1), 262.
6. Garland, S. M.; Hernandez-Avila, M.; Wheeler, C. M.; Perez, G.; Harper, D. M.; Leodolter, S.; Tang, G. W. K.; Ferris, D. G.; Steben, M.; Bryan, J.; Taddeo, F. J.; Railkar, R.; Esser, M. T.; Sings, H. L.; Nelson, M.; Boslego, J.; Sattler, C.; Barr, E.; Koutsky, L. A., Quadrivalent Vaccine against Human Papillomavirus to Prevent Anogenital Diseases. *New England Journal of Medicine* **2007**, *356* (19), 1928-1943.

7. Tanzi, E.; Bianchi, S.; Fasolo, M. M.; Frati, E. R.; Mazza, F.; Martinelli, M.; Colzani, D.; Beretta, R.; Zappa, A.; Orlando, G., High performance of a new PCR-based urine assay for HPV-DNA detection and genotyping. *Journal of Medical Virology* **2013**, *85* (1), 91-98.
8. Walboomers, J. M. M.; Jacobs, M. V.; Manos, M. M.; Bosch, F. X.; Kummer, J. A.; Shah, K. V.; Snijders, P. J. F.; Peto, J.; Meijer, C. J. L. M.; Muñoz, N., Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of Pathology* **1999**, *189* (1), 12-19.
9. Niccoli, S.; Abraham, S.; Richard, C.; Zehbe, I., The Asian-American E6 Variant Protein of Human Papillomavirus 16 Alone Is Sufficient To Promote Immortalization, Transformation, and Migration of Primary Human Foreskin Keratinocytes. *Journal of Virology* **2012**, *86* (22), 12384-12396.
10. Thomas, M.; Tomaić, V.; Pim, D.; Myers, M. P.; Tommasino, M.; Banks, L., Interactions between E6AP and E6 proteins from alpha and beta HPV types. *Virology* **2013**, *435* (2), 357-362.
11. Togtema, M.; Pichardo, S.; Jackson, R.; Lambert, P. F.; Curiel, L.; Zehbe, I., Sonoporation Delivery of Monoclonal Antibodies against Human Papillomavirus 16 E6 Restores p53 Expression in Transformed Cervical Keratinocytes. *PLoS ONE* **2012**, *7* (11), e50730.
12. Jackson, R.; Togtema, M.; Zehbe, I., Subcellular localization and quantitation of the human papillomavirus type 16 E6 oncoprotein through immunocytochemistry detection. *Virology* **2013**, *435* (2), 425-432.
13. Froloff, N.; Windemuth, A.; Honig, B., On the calculation of binding free energies using continuum methods: Application to MHC class I protein-peptide interactions. *Protein Science* **1997**, *6* (6), 1293-1301.
14. Schweizer, J.; Lu, P. S.; Mahoney, C. W.; Berard-Bergery, M.; Ho, M.; Ramasamy, V.; Silver, J. E.; Bisht, A.; Labiad, Y.; Peck, R. B.; Lim, J.; Jeronimo, J.; Howard, R.; Gravitt, P. E.; Castle, P. E., Feasibility Study of a Human Papillomavirus E6 Oncoprotein Test for Diagnosis of Cervical Precancer and Cancer. *Journal of Clinical Microbiology* **2010**, *48* (12), 4646-4648.
15. Giovane, C.; Trave, G.; Briones, A.; Lutz, Y.; Wasylyk, B.; Weiss, E., Targetting of the N-terminal domain of the human papillomavirus type 16 E6 oncoprotein with monomeric ScFvs blocks the E6-mediated degradation of cellular p53. *Journal of Molecular Recognition* **1999**, *12* (2), 141-152.

16. Masson, M.; Hindelang, C.; Sibler, A. P.; Schwalbach, G.; Trave, G.; Weiss, E., Preferential nuclear localization of the human papillomavirus type 16 E6 oncoprotein in cervical carcinoma cells. *Journal of General Virology* **2003**, *84*, 2099-2104.
17. Baleja, J. D.; Cherry, J. J.; Liu, Z.; Gao, H.; Nicklaus, M. C.; Voigt, J. H.; Chen, J. J.; Androphy, E. J., Identification of inhibitors to papillomavirus type 16 E6 protein based on three-dimensional structures of interacting proteins. *Antiviral Research* **2006**, *72* (1), 49-59.
18. Detection of protein-protein interactions using the GST fusion protein pull-down technique. *Nat Meth* **2004**, *1* (3), 275-276.
19. Nominé, Y.; Masson, M.; Charbonnier, S.; Zanier, K.; Ristriani, T.; Deryckère, F.; Sibler, A.-P.; Desplancq, D.; Atkinson, R. A.; Weiss, E.; Orfanoudakis, G.; Kieffer, B.; Travé, G., Structural and Functional Analysis of E6 Oncoprotein: Insights in the Molecular Pathways of Human Papillomavirus-Mediated Pathogenesis. *Molecular Cell* **2006**, *21* (5), 665-678.
20. Zanier, K.;ould M'hamedould Sidi, A.; Boulade-Ladame, C.; Rybin, V.; Chappelle, A.; Atkinson, A.; Kieffer, B.; Travé, G., Solution Structure Analysis of the HPV16 E6 Oncoprotein Reveals a Self-Association Mechanism Required for E6-Mediated Degradation of p53. *Structure* **2012**, *20* (4), 604-617.
21. Zanier, K.; Charbonnier, S.; Sidi, A. O. M. h. O.; McEwen, A. G.; Ferrario, M. G.; Poussin-Courmontagne, P.; Cura, V.; Brimer, N.; Babah, K. O.; Ansari, T.; Muller, I.; Stote, R. H.; Cavarelli, J.; Vande Pol, S.; Travé, G., Structural Basis for Hijacking of Cellular LxxLL Motifs by Papillomavirus E6 Oncoproteins. *Science* **2013**, *339* (6120), 694-698.

Chapter 3. Methodology and Rational

3.1. COMPUTATIONAL APPROACHES

3.1.1. Molecular Docking

Molecular docking involves virtually placing a molecule within a protein binding site, using computer software¹. The complexity of the conformational degrees of freedom of both the protein and ligand are taken into account¹. It is critical that these conformations are taken into consideration so that the conformation used for scoring is biologically significant¹. It is important that the conformation selected is similar to the conformation *in-vivo*. The main algorithms available for molecular docking can be classified into three approaches: systematic, stochastic, and simulation-based methods².

Systematic approach provides a complete coverage of all reasonable conformations². Such an approach is suitable for molecular fragments. Even small molecules have too high a number of conformation making this method impractical².

Stochastic approach is subdivided into two techniques Monte Carlo and Genetic Search algorithms. Monte Carlo will generate a random conformation, which is scored, then a series of additional random conformations are generated and scored². These conformations are compared to the original conformation. Genetic Search algorithms attempt to simulate the random events that take place during gene replication in order to generate various conformations².

Simulation-based methods are molecular dynamics simulations in which a molecular system is observed while its energy varies over time in order to observe all theoretical conformational states². Such methods can take a long time in order to produce a good sampling of conformational states. Such methods are only useful when necessary computational resources are available².

3.1.2.Virtual Ligand Screening (VLS)

Virtual ligand screening consists of screening large libraries of chemical compounds using computational methods in order to identify candidates with probable biological activity or desired properties. Two major subtypes of VLS include descriptor based methods and structure based methods. In descriptor based methods the screening of compounds is based on structural features that are shared amongst known ligands. Examples of this approach include similarity and substructure matching³, pharmacophore matching⁴, and 3D shape matching⁵. Structural based methods are those that use a known 3D structure of the target protein, and the affinity of a ligand for its target is evaluated based on how well the ligand is bound to the binding site of the protein target⁶. Our particular area of interest is in structural based methods as E6 has no known ligands at this point and attempts using pharmacophore modeling or descriptor based approaches have been unsuccessful and have demonstrated a need for a known ligand database for E6⁷.

3.1.3.HierVLS and Scoring Functions

Hierarchical Virtual Ligand Screening (HierVLS) is a method that uses molecular docking to screen large libraries of chemical compounds against proteins. HierVLS will dock a large number of conformations of a ligand to a protein which is kept at a fixed conformation until the last step in procedure, when the protein is allowed to relax to accommodate the bound ligand⁸. This method saves the most expensive computational procedures for the most promising conformations⁸. The docked ligands are scored using force fields⁸. Force field based scoring functions are commonly used for the estimation of potential energy of protein-ligand interactions⁹. These functions are based on principles of classical mechanics, and provide accurate potential energy estimates¹⁰. Force fields describe intramolecular and intermolecular

forces as a summation of bonded atoms and non-bonded atoms ($E_{\text{potential}} = E_{\text{Bonded}} + E_{\text{Non-Bonded}}$)¹⁰.

Drieding is the force field used in HierVLS¹⁰.

3.1.3. a) Steps in HierVLS

The first level is level 0: Coarse Grain Conformation Search using Dock 4.0 software package⁸. A scoring function along with Gasteiger charges¹¹ is used to obtain a predefined number of protein-bound conformations per ligand in a ligand library⁸. In this level multiple conformations of the ligand are docked to the protein in different orientations⁸. To move to level 1, a docked conformation must pass a user-defined buried surface cut off. This buried surface filter is used to eliminate those conformers that do not have the minimum specified amount of surface buried⁸. In order to pass through this filter the ligand must be buried at least 70% in the field of the protein or otherwise specified. Conformers that meet this requirement will be carried on to the next stage. In Level 1 energy minimization takes place on all conformations generated using Drieding¹⁰ force field⁸. Out of all the conformations energy-minimized in level 1, the top 3 conformations with the lowest energy will be selected to move to Level 2, which is an all-atoms energy minimization⁸. The best (lowest) docked conformer by energy for each ligand is selected and used for ligand ranking⁸. The binding energies used for ligand ranking take into account the solvation energies associated to the free protein, free ligand, and bound complex. Solvation energies are calculated using an implicit solvation model, which allows us to account for solvent effects without the explicit inclusion of water molecules⁸. The implicit solvation model used in this step is Analytical Volume Generalized Born (AVGB)⁸. The binding energies are calculated as:

$$\text{BindE} = E(\text{solvated complex}) - E(\text{solvated free protein}) - E(\text{solvated free ligand})^8.$$

Through taking the differences, as shown in the equation above, binding energy (BindE) between ligand and protein is estimated⁸ (Figure 13). VLS is an effective tool and useful for the discovery of a molecular probe, as it is both fast and cost effective. Computational screening is expected to provide a ranked list of potential molecular probes for prototype E6. The HierVLS method has been shown to be an effective virtual ligand screening method in several studies¹².

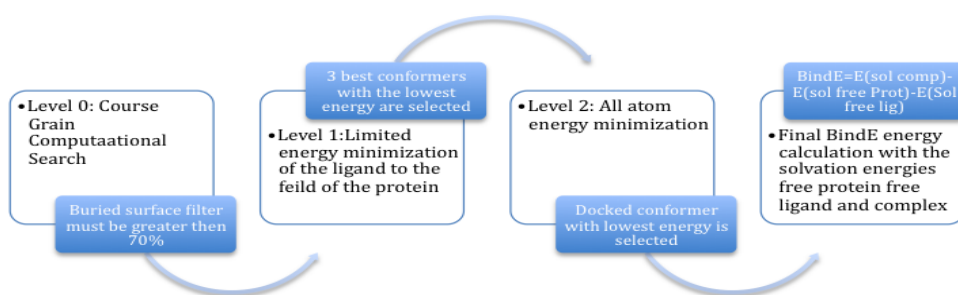


Figure 13: Steps of HierVLS. The best bound conformations are submitted to increasingly demanding computational steps, with the most demanding steps reserved for most promising conformations.

3.2.EXPERIMENTAL APPROACHES

3.2.1.Overview

Experimental ligand binding assays were selected carefully in order to properly characterize binding affinity of a molecule keeping in mind the hope of using the molecule as either a diagnostic imaging agent or tool for biochemical assays. Experimental approaches have been selected in order to not only demonstrate binding but also show that the binding is specific as opposed to non-specific. This was done by demonstrating dose-dependence. Binding assay

will either use the protein (tryptophan fluorescence) or the ligand (fluorescence polarization) as source of signal.

3.2.2. Tryptophan Fluorescence

Since molecules do not have to be fluorescent to be tested in tryptophan fluorescence, this method allows us to test affinity without conjugating the molecules to a fluorescent tag, which can negatively impact binding affinity. Tryptophan fluorescence is a biochemical assay that takes advantage of the indolechromophore naturally present in tryptophan¹³. Tryptophan has an observable intrinsic fluorescence that is sensitive to changes in its environment caused by conformational changes that occur upon ligand binding¹³. This intrinsic fluorescence is generally observed at an excitation wavelength of 288nm and an emission wavelength of 350nm¹³. Tryptophan is a rare amino acid in proteins, and is generally only seen once in every polypeptide chain¹³. Due to the rareness of tryptophan and its sensitivity to environment, Trp fluorescence is an ideal model for tracking conformational change. Changes shown in fluorescence can indicate a change in environment of the tryptophan due to ligand binding¹³. This method has been shown useful in showing affinity of potential drugs¹⁴. E6 only contains one tryptophan making this method adequate for probing ligand binding (Figure 6).

This assay will be used to calculate the change in intrinsic tryptophan fluorescence of top hit compounds obtained through HierVLS. In all experiments the E6 protein concentration will be held constant. Wells and controls are made to represent each ligand on its own at each concentration used. This is done in order to make sure that the fluorescence from the background is omitted when calculating the percentage change. Solvent controls were also present in later experiments, as different solvents will impact the environment of tryptophan differently. Curves were taken using an excitation wavelength of 288nm and an emission wavelength of 350nm.

Fluorescence intensity values were then manipulated to determine percentage change. A change in intrinsic tryptophan fluorescence that is greater than 30% is indicative of a binding event. This binding event will be further tested using fluorescence polarization.

3.2.3. Fluorescence Polarization

In fluorescence polarization, the fluorescent ligand is excited with polarized light¹⁵. An emission filter is then used to measure the fluorescence intensity emitted by the ligand in both the parallel and perpendicular directions¹⁵. When the ligand is stationary, it is not able to rotate or tumble, and the fluorescence emitted is polarized. Thus, the parallel intensity will be greater than the perpendicular which results in a higher polarization value¹⁵. That is the case when the ligand is bound to the protein target. When the fluorescent molecule is unbound, the fluorescence emitted is depolarized because the ligand is free to rotate. In this case, the parallel and perpendicular intensities should be close to equal, resulting in a low polarization value¹⁵. Thus the value of polarization of a system is proportional to the number of occupied binding sites.

Fluorescence polarization is generally used as a ligand binding assay¹⁶. This is because of the positive correlation between it and the number of occupied binding sites¹⁵. Therefore a higher polarization values means more ligand is bound. This method has a few strong advantages, which include high sensitivity, highly reproducible, and its simple mix and read format¹⁵. Due to the dose relationship of fluorescence polarization and these advantages, this method is an ideal choice for obtaining a dose-response curve with a fluorescence-tagged lead or an intrinsically fluorescent compound. From this method we hope to obtain an accurate EC50 that is in the mid or lower micro-molar range to show necessary affinity for probe development. This was shown in the recent study where paclitaxel was identified as an inhibitor of botulism neurotoxin A, further validating our rationale^{12d}.

3.2.4.Dose-Response Curves and Curve Fitting

A dose response curve is used to track the response proportional to the concentration of a stressor¹⁷. In the case of probe/ drug development this stressor is a molecule binding to its target usually a protein. These curves are used for the visualization and calculation of binding constants such as EC50¹⁸. When building such a curve proper fitting methods need to be employed in order to accurately determine these constants.

When fitting a dose-response curve it is important that the properties of the hill slope are properly evaluated. The hill equation is traditionally a three parameter equation that shows a non-linear relationship between x (independent) and y (dependent) variables¹⁹. However, an additional parameter has been added that accounts for the baseline response giving the equation four parameters¹⁹. The hill equation is used for the fit of experimental data from physiochemical reactions¹⁹. It is particularly useful in pharmacology demonstrating dose effects, based on drug concentrations¹⁹.

The hill equation contains a coefficient known as the hill coefficient. The hill coefficient represents cooperative binding²⁰. Cooperative binding means that the binding affinity to the protein by the ligand is increased upon ligand binding²⁰. This binding could even take place at other binding sites. This is represented by a hill slope or hill coefficient greater than 1. A hill slope that is less than one indicates negative cooperative binding, meaning once the ligand is bound its affinity to the protein is decreased²⁰. If it is equal to one then the binding is completely independent of other binding events²⁰.

Prism Graphpad (<http://www.graphpad.com/scientific-software/prism/>) is a program that has several curve fitting functions, being widely used in dose-response curve fitting²¹. This program contains specific functions used for fitting a dose-response curve. Since we do not want

to assume that binding is not co-operative, the four parameter equation was used: $Y = \text{Bottom} + \frac{(\text{Top} - \text{Bottom})}{(1 + 10^{((\text{LogEC50} - X) * \text{HillSlope}))}}$ was used²¹. In this fitting, the hill slope is not assumed to be one. Since the binding of E6 is not understood at an experimental level, independent binding was not assumed.

3.3.CELL BASED METHODS

Cell permeability and viability was tested tested for promising compounds from ligand binding experiments. The main purpose of these assays is to demonstrate cell permeability and specificity within the cell environment. This was conducted using one of two methods. The first choice method is to use a microscope capable of inducing and detecting fluorescence. This will provide visualization of a binding event using photographs. This method and the other method will require the molecule to be labeled with a dye or be intrinsically fluorescent, however databases will be created that contain molecules that are suitable for fluorophore attachment. Prior to using the microscope cells, were incubated with a predefined concentration of compound for a fixed amount of time. These time points vary from one hour of incubation to 24 hours in order to get an idea of ligand uptake over time. After the various time points are complete, cells were placed under a microscope that induces fluorescence of the molecule and detect it. If the molecule has made it successfully through the cell membrane then it should be seen accumulating within the cell. Each well containing the drug has a corresponding control well containing all ingredients, except the compound, for proper comparison. If proper fluorescence filters are not available that match the excitation and emission wavelengths of the promising compounds, then an alternative method can be employed.

The second method uses fluorescence reads from the compound being tested using a microplate reader. Wells that contain potential probe were compared to control wells that contain

no potential probe. This was carried out by incubating cells with the compound being tested at two concentrations for various time lengths, after which cells will be lysed. The lysate was then filtered and fluorescence measurements were then taken from each sample and compared to a control that contains no compound.

In order to evaluate whether the compound targets E6 specifically, obtain proper cell lines need to be selected. CaSki, SiHa, and C33A cells were used in order to test cell viability, permeability and E6 dependency of promising compounds. CaSki contains 200-300 copies of the gene coding for E6²², whereas SiHa only contains 1-2 copies²³. C33A cells are cervical cancer cells, however they contain zero copies of E6, making them an ideal negative control²⁴. Due to the varying number of copies of E6 gene in each cell type, if binding is specific we expect higher accumulation of the compound being tested in CaSki cells than in the other cell lines, as more E6 is present. Either method will be sufficient to indicate cell permeability and will give insights on specificity.

3.4.REFERENCES:

1. Froloff, N.; Windemuth, A.; Honig, B., On the calculation of binding free energies using continuum methods: Application to MHC class I protein-peptide interactions. *Protein Science* **1997**, *6* (6), 1293-1301.
2. Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J., Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* **2004**, *3* (11), 935-949.
3. Mestres, J.; Knegt, R. A., Similarity versus docking in 3D virtual screening. *Perspectives in Drug Discovery and Design* **2000**, *20* (1), 191-207.
4. Mason, J. S.; Good, A. C.; Martin, E. J., 3-D Pharmacophores in Drug Discovery. *Current Pharmaceutical Design* **2001**, *7* (7), 567-597.

5. Srinivasan, J.; Castellino, A.; Bradley, E. K.; Eksterowicz, J. E.; Grootenhuys, P. D. J.; Putta, S.; Stanton, R. V., Evaluation of a Novel Shape-Based Computational Filter for Lead Evolution: Application to Thrombin Inhibitors. *Journal of Medicinal Chemistry* **2002**, *45* (12), 2494-2500.
6. Lyne, P. D., Structure-based virtual screening: an overview. *Drug Discovery Today* **2002**, *7* (20), 1047-1055.
7. Baleja, J. D.; Cherry, J. J.; Liu, Z.; Gao, H.; Nicklaus, M. C.; Voigt, J. H.; Chen, J. J.; Androphy, E. J., Identification of inhibitors to papillomavirus type 16 E6 protein based on three-dimensional structures of interacting proteins. *Antiviral Research* **2006**, *72* (1), 49-59.
8. Floriano, W. B.; Vaidehi, N.; Zamanakos, G.; Goddard, W. A., HierVLS Hierarchical Docking Protocol for Virtual Ligand Screening of Large-Molecule Databases. *Journal of Medicinal Chemistry* **2003**, *47* (1), 56-71.
9. Smith, G. R.; Sternberg, M. J. E., Prediction of protein-protein interactions by docking methods. *Current Opinion in Structural Biology* **2002**, *12* (1), 28-35.
10. Mayo, S. L.; Olafson, B. D.; Goddard, W. A., DREIDING: a generic force field for molecular simulations. *The Journal of Physical Chemistry* **1990**, *94* (26), 8897-8909.
11. Gasteiger, J.; Marsili, M., Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron* **1980**, *36* (22), 3219-3228.
12. (a) Li, X.; Bachmanov, A. A.; Maehashi, K.; Li, W.; Lim, R.; Brand, J. G.; Beauchamp, G. K.; Reed, D. R.; Thai, C.; Floriano, W. B., Sweet Taste Receptor Gene Variation and Aspartame Taste in Primates and Other Species. *Chemical Senses* **2011**, *36* (5), 453-475; (b) Floriano, W. B.; Vaidehi, N.; Goddard, W. A., Making Sense of Olfaction through Predictions of the 3-D Structure and Function of Olfactory Receptors. *Chemical Senses* **2004**, *29* (4), 269-290; (c) Vaidehi, N.; Schlyer, S.; Trabanino, R. J.; Floriano, W. B.; Abrol, R.; Sharma, S.; Kochanny, M.; Koovakat, S.; Dunning, L.; Liang, M.; Fox, J. M.; de Mendonça, F. L.; Pease, J. E.; Goddard, W. A.; Horuk, R., Predictions of CCR1 Chemokine Receptor Structure and BX 471 Antagonist Binding Followed by Experimental Validation. *Journal of Biological Chemistry* **2006**, *281* (37), 27613-27620; (d) Dadgar, S.; Ramjan, Z.; Floriano, W. B., Paclitaxel Is an Inhibitor and Its Boron Dipyrromethene Derivative Is a Fluorescent Recognition Agent for Botulinum Neurotoxin Subtype A. *Journal of Medicinal Chemistry* **2013**, *56* (7), 2791-2803.
13. Chen, Y.; Barkley, M. D., Toward Understanding Tryptophan Fluorescence in Proteins†. *Biochemistry* **1998**, *37* (28), 9976-9982.

14. Epps, D. E.; Raub, T. J.; Caiolfa, V.; Chiari, A.; Zamaï, M., Determination of the Affinity of Drugs toward Serum Albumin by Measurement of the Quenching of the Intrinsic Tryptophan Fluorescence of the Protein. *Journal of Pharmacy and Pharmacology* **1999**, *51* (1), 41-48.
15. Jameson, D.; Mocz, G., Fluorescence Polarization/Anisotropy Approaches to Study Protein-Ligand Interactions. In *Protein-Ligand Interactions*, Ulrich Nienhaus, G., Ed. Humana Press: 2005; Vol. 305, pp 301-322.
16. Parker, G. J.; Law, T. L.; Lenoçh, F. J.; Bolger, R. E., Development of High Throughput Screening Assays Using Fluorescence Polarization: Nuclear Receptor-Ligand-Binding and Kinase/Phosphatase Assays. *Journal of Biomolecular Screening* **2000**, *5* (2), 77-88.
17. Chrousos, G. P., Stressors, Stress, and Neuroendocrine Integration of the Adaptive Response: The 1997 Hans Selye Memorial Lecture. *Annals of the New York Academy of Sciences* **1998**, *851* (1), 311-335.
18. Vanewijk, P. H.; Hoekstra, J. A., Calculation of the EC50 and Its Confidence Interval When Subtoxic Stimulus Is Present. *Ecotoxicology and Environmental Safety* **1993**, *25* (1), 25-32.
19. Goutelle, S.; Maurin, M.; Rougier, F.; Barbaut, X.; Bourguignon, L.; Ducher, M.; Maire, P., The Hill equation: a review of its capabilities in pharmacological modelling. *Fundamental & Clinical Pharmacology* **2008**, *22* (6), 633-648.
20. Weiss, J. N., The Hill equation revisited: uses and misuses. *The FASEB Journal* **1997**, *11* (11), 835-41.
21. Harvey Motulsky, A. C., *Fitting Models to Biological Data Using Linear and Nonlinear Regression* Oxford University Press: New York 2004
22. O'Leary, J. J.; Browne, G.; Johnson, M. I.; Landers, R. J.; Crowley, M.; Healy, I.; Street, J. T.; Pollock, A. M.; Lewis, F. A.; Andrew, A., PCR in situ hybridisation detection of HPV 16 in fixed CaSki and fixed SiHa cell lines. *Journal of Clinical Pathology* **1994**, *47* (10), 933-938.
23. Heiles, H. B.; Genersch, E.; Kessler, C.; Neumann, R.; Eggers, H. J., In situ hybridization with digoxigenin-labeled DNA of human papillomaviruses (HPV 16/18) in HeLa and SiHa cells. *BioTechniques* **1988**, *6* (10), 978-981.

24. Jeon, S.; Lambert, P. F., Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis. *Proceedings of the National Academy of Sciences* **1995**, *92* (5), 1654-1658.

Chapter 4. Materials and Procedures

4.1. MATERIALS

O-succinyl –homoserine (MFCD00055782) was ordered from Sigma Aldrich Co in a 25mg quantity at 98% purity, determined by Sigma Aldrich using TLC. 125mg of 3-amino-5-fluorobenzo [e] [1,2,4] triazine 1, 4 dioxide (AFTD) was ordered from Sunbiochem Inc (MFCD14636665) at 98% purity, determined by Sunbiochem Inc using TLC. Paclitaxel was ordered from Sigma Aldrich Co (T7402) with a purity of 95% as confirmed by High Performance Liquid Chromatography (HPLC). Aminopterin (ALX-440-041-M010) was obtained from Enzo Life Sciences Inc, through Cedar Lane Inc, with a purity of 98%. These compounds were order and properly stored until usage in ligand binding experiments. Protein E6 was ordered from GenScript Inc in both unlabelled and biotinylated form. GenScript Inc determined the purity of E6 through the use of Coomassie blue-stained SDS-PAGE gel. Protein was stored in a buffer solution to keep the pH at 8.0 (50nM Tris-HCl, 10% Glycerol, pH 8.0). This protein was stored at -80°C until used in ligand binding assays. CaSki, SiHa, and C33A cell lines were obtained through American Type Culture Collection Inc (ATCC). 6F4 E6 monoclonal antibody was a gift from the Arbour Vita Co. Dulbecco’s Modification of Eagles Medium (DMEM) supplemented with 10% fetal bovine serum and anti/anti was obtained through Fisher Co.

4.2. COMPUTATIONAL SCREENING

Prior to conducting molecular docking, databases of drug like compounds needed to be created. The compounds in these databases all shared common drug like properties and had characteristics that were suitable for a potential molecular imaging probe for HPV16 E6. Three databases were created and screened.

4.2.1. a) Primary Amine Database

Compounds of this database were downloaded from PubChem¹. Lipinski's rule of 5 was used as part of the search criteria². A basic primary amine was drawn so that structural similarity was searched as well. 2,000 compounds were downloaded. Editing was then conducted to eliminate those molecules that contained counter ions and atoms that our docking software (Cassandra) cannot process, such as Boron. After editing, 1,153 molecules were left in the database. Gasteiger method³ was used to calculate partial charges and energy minimization was conducted using MMFF94x force field⁴. The database was then saved as a mol2 file in Molecular Operating Environment MOE 2010.10 (CC group)⁵ and converted to BGF file format using Babel⁶. The BGF file was then cleaned up using an already prepared script (clean_BGF.pl). BGF files were then converted into a BGF400FSM files using an additional in house script (BGF2BGF4fsm).

4.2.1. b) Imaging Compound Database

Imaging compound database was created using Discovery Gate software⁷. Lipinski's rule of 5 was used as part of the search criteria². In addition, the "known imaging agents compounds" tab was clicked for search criteria. Compounds were then downloaded from Discovery Gate in sdf format. Some compounds were edited to eliminate counter ions and atoms that Cassandra cannot process. A total of 4,965 were prepared for this database. These compounds were imported into MOE 2010.10 (CC group)⁴, compiled as a database, saved as a mol2 file and converted to BGF file us Babel⁶. Gasteiger method³ was used for calculating partial charges and energy minimization was conducted using MMFF94x force field². The BGF file was then cleaned up using an already prepared script (clean_BGF.pl). BGF files were then converted into a BGF400FSM files using an additional in house script (BGF2BGF4fsm).

4.2.1. c) Fluorine-Containing ligand Database

Fluorine containing compounds were downloaded from both Discovery Gate⁷ and PubChem¹ in sdf format. The major search criteria used was Lipinski's rule of 5⁵. In addition we searched for molecules that contained fluorine atoms. This generated a library of 4,146 compounds, which was used for molecular docking. These compounds were imported into MOE 2010.10 (CC group) and compiled as a database and saved as a mol2 file⁴. Gasteiger method³ was used for calculating partial charges and energy minimization was conducted using MMFF94x force field². Mol2 file was converted to BGF using Babel⁶. A BGF file was then cleaned up using an already prepared script (clean_BGF.pl). BGF files were then converted into a BGF400FSM files using an additional in house script (BGF2BGF4fsm).

4.2.1. d) Shelf Compounds Database

A series of compounds that were readily available within the laboratory were also prepared as a database. This database was also saved a mol2 string using MOE 2010.10 (CC group)⁵. Gasteiger method³ was used for partial charges were calculated and energy minimization was conducted using MMFF94x force field². Mol2 file was converted using Babel⁶ to a BGF file. A BGF file was then cleaned up using an already prepared script (clean_BGF.pl). BGF files were then converted into a BGF400FSM files using an additional in house script (BGF2BGF4fsm).

4.2.2.N-Terminal (PDB 2LJX) Docking

4.2.2. a) Editing Protein Structure.

The structure for the N-terminal domain of HPV16 E6 was downloaded from PDB.org (PDB 2LJX). This particular structure contained multiple conformations of HPV16 E6. To determine which confirmation to use, each conformation was energy-minimized using the

Dreiding force field⁸. The conformations and their binding energies were ranked according to total potential energy (Table 3). The best (lowest energy) conformer, conformation 18 with a potential energy of -1609.1491kcal/mol, was selected for further analysis.

Table 3: All conformation in the PDBfile (2LJX). Confirmation 18 was selected for screening.

Conformation Number	Potential Energy (kcal/mol)
1	-1494.0000
2	-1439.0000
3	-1571.5252
4	-1546.5809
5	-1544.9435
6	-1488.8126
7	-1569.3594
8	-1466.7676
9	-1539.3421
10	-1607.8191
11	-1537.3729
12	-1542.4821
13	-1539.1646
14	-1363.9693
15	-1587.9279
16	-1522.7819
17	-1416.2263
18	-1609.1491
19	-1462.3501
20	-1493.3796

Several file types needed to be created for the protein structure (2LJX) (mol2, PDB, BGF, and BGF 400 fsm). The structure of the protein is generally saved in MOE 2010.10 (CC group)⁵ as a mol2 and PDB files, then scripts were utilized to make the remaining files. Using the program Yassara⁹ we were able to save the PDB file with no hydrogen's. The scripts used included: BGF2BGF4fsm and PrepareBGF.

4.2.2. b) Submitting HierVLS jobs to SHARCNET

With the structure files created for the protein and the databases of ligands, jobs could be submitted to Shared Hierarchical Research Computing Network (SHARCNET)¹⁰, using the Cassandra interface¹¹ to the HierVLS method¹². SHARCNET is a high performance computer cluster, which allows thousands of jobs to run simultaneously. Default parameters (see appendix B) were used when submitting all databases except for shelf compounds database, where the buried surface area was lowered to 30% because of the large size of some of the molecule.

4.2.2. c) Analysis of HierVLS Results

When each molecule is docked, a series of files are made for each ligand. Files of particular interest to our analysis include a BindE table, which contains the raw (force field) binding energy score. A BGF file is also present, which contains the ligand bound to the protein structure. PDB files are made for the protein that contains the computationally determined binding sites as determined by PASS¹³. The N-terminal domain contained two binding sites, R1 and R2. HierVLS generated 2,911 entries for the N-terminal domain R1, and 6,009 for R2. Force field based binding energies for each ligand was taken from the results folders and compiled into an excel spreadsheet. This was done by going into each results folder and concatenating the BindE energy tables. Each of N-terminus domain two binding sites were individually analyzed. Once compiled the mean and standard deviation of the force field scores for all ligands surviving

the HierVLS process were calculated in Excel for each binding region. The binding threshold was determined by calculating two standard deviations below the mean for each site. These results were imported into Prism 6 (Graph Pad, inc)¹⁴, where graphs were made to depict these results. It was shown that 10 compounds passed the binding/nonbinding threshold for R2 and only 3 passed for R1. The 10 molecules that passed the threshold for R2 were examined, as well as the three molecules that passed the threshold for R1. Even though all three databases were docked, the advantageous properties of the primary amine database and the fluorine containing database made these databases first priority. Those compounds from the fluorine containing database and primary amine database were further examined to determine commercial availability and ligand interaction diagrams were also generated using MOE 2010.10 (CC group)

5.

4.2.3.C-terminus Screening (2FK4)

4.2.3. a)Obtaining and editing protein structure

We obtained our structure for the C-terminal domain of HPV 16 E6 protein using the PDB search engine (PDB code 2FK4). Once the structure was downloaded, energy minimization was performed MMFF94x force field⁴. The quality of this structure was checked using Procheck software¹⁵. For molecular docking, a few different file types of the protein are required (mol2, PDB, BGF, and BGF 400 fsm). The structure of the protein is generally saved in MOE 2010.10 (CC group)⁵ as mol2 and PDB file formats, then scripts are utilized to make the remaining files. Using the program Yassara⁹ we were able to save the PDB file with no hydrogen's. The scripts used included: BGF2BGF4fsm, and PrepareBGF.

4.2.3. b) Submitting jobs to SHARCNET HierVLS

With the structure files created for the protein and the databases of ligands, jobs could be submitted to Shared Hierarchical Research Computing Network (SHARCNET)¹⁰ using the Cassandra¹¹ interface to the HierVLS method¹². SHARCNET is a high performance computer cluster, which allows thousands of jobs to run simultaneously. Default parameters (see appendix B) were used when submitting all databases except for shelf compounds database, where the buried surface area was lowered to 30% because of the size of the molecule.

4.2.3. c) Analysis of Results

A script was used to sort through the HierVLS output files and identify molecules in the database that were missed in the initial screening (checkruns). The script created a mol2 file, which was then turned into BGF and BGF4fsm files using the same methods as described above. This was done until all our molecules had been docked, from each database. When each molecule is docked a series of files are made for each ligand. Files of particular interest to our analysis include a BindE table, which contains the raw binding energy score. A BGF file is also present, which contains the ligand found to the protein structure. In PDB files are made for the protein that contains the computationally determined binding sites from PASS⁵. For case of the C-terminal domain 3,339 entries acquired. All force field based binding energy tables were concatenated together and transferred into the Excel (Microsoft)¹⁶ where the binding threshold was calculated by as two standard deviations below the mean binding energy of the entire set. This was done as we are assuming normal distribution due to the large sample size. Two standard deviations will thus give us 95% confidence. Confidence intervals are commonly used in statistics to test the validity of an estimate. In our case, it was used to ensure that we select those ligands with the highest probability of binding, which are the 5% with the best scores.

This binding threshold was used as an ideal screening tool in order to eliminate those ligands that have little chance of binding. Using this threshold we went through the fluorine containing molecule database as well as the primary amine database to make a selection based a selection rational: Observe ligand interaction diagrams to make sure binding is specific, Check for commercial availability, obtain toxicology information (Figure 14).



Figure 14: Step by step process used for the selection of ligands for C-terminus domain. Molecules will selected for experimental ligand binding assays based on these steps

Five molecules were examined from the fluorine containing database and 3 molecules were examined for the primary amine database. Ideally we wanted to select one molecule from each database. In order to do so ligand interaction diagrams were created using MOE 2010.10 (CC group)⁵. This was easily done as each output file for each ligand contains a BGF file that contains the protein structure with the ligand bound. However, to further eliminate compounds commercial availability was looked at. Although all molecules were available some were more than others (AFTD). In addition some molecules had properties that made them a lot less toxic (O-succinyl-L-homoserine). The imaging agents database results can be further examined if those top candidates from other databases fail to show experimental affinity.

4.3. TRYPTOPHAN FLUORESCENCE EXPERIMENT

3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) (300 μ M, 100 μ M, 25 μ M, 10 μ M, 3 μ M, 1 μ M, 0.1 μ M, 0.01 μ M), O-succinyl-L-homoserine (300 μ M, 100 μ M, 25 μ M, 10 μ M, 3 μ M, 1 μ M, 0.1 μ M, 0.01 μ M,) and Paclitaxel (15 μ M, 10 μ M, 1 μ M) were tested in an intrinsic tryptophan fluorescence assay at several concentrations with a fixed concentration of protein E6 (0.03mg/ml). AFTD and O-succinyl-L-homoserine concentrations were accurately measured using a 1mM stock solution dissolved in deionized water, with dilutions prepared (100 μ M, 10 μ M, 1 μ M) in order to reach lower concentrations. E6 protein was accurately measured from the stock concentration (0.201mg/ml). The assay was buffered to a pH of 7.5 using 20mM HEPES, 0.01% tween (v/v) and incubated to a temperature of 37°C. Control wells were included for the protein alone, AFTD, O-succinyl-L-homoserine, and paclitaxel (at each concentration tested). Samples were plated as triplicate wells, and duplicate wells for all controls, in a 96-well black bottom microplate (Nunclon®). The experiment was run using a Synergy 4 Microplate Reader (Biotek). The monochromator was used to collect emission spectrums for each well from 310nm-700nm at an excitation of 288nm. Fluorescence intensity was read using a probe height of 4mm, a 150 μ sec delay between wells, and 20 reads per well. Data was collected every 30 minutes for 60 minutes using a sensitivity value of 100 with Gen5 software (Biotek)¹⁷. The read settings and filter/mirror combination were optimized during preliminary experiments. Changes in intrinsic tryptophan fluorescence were calculated using Prism 6 (Graph Pad, Inc)⁴ baseline reduction functions.

Water, buffer, and required solvent were added first to all the wells in the microplate. Once water buffer and solvent were added the ligands and antibody were measured out accurately using the stock solutions made at the desired concentrations. The reaction was

initiated with the addition of E6 protein (0.03mg/ml) after which the plate was incubated in the plate reader under the conditions specified above.

4.3.1. a) Data Analysis

Fluorescence intensity readings were imported into Excel 2010 (Microsoft, Inc)² then directly imputed into Prism 6 (Graph Pad, Inc)¹⁴. Means and standard errors were calculated using Prism 6 (Graph Pad, Inc)¹⁴ for all duplicate and triplicate wells using column statistic function. For each sample containing protein plus ligand, the background for each corresponding ligand at the appropriate concentration was deducted using baseline column math (Difference=Value-Baseline)¹⁴. Each individual ligand concentration on its own was assigned as the baseline for the corresponding sample, in order to deduct the background from the ligand. Baseline column math, was once again performed on each ligand and concentration, this time assigning the protein with corresponding solvent, as the baseline value to generate the percent change values for each wavelength (%difference:= 100*(value-baseline)/baseline)¹⁴. The baseline was assigned as the proper protein control depending on sample and concentration. The 350nm percent change values were then plotted in a bar graph with error bars representing the standard error that was propagated by Prism 6 (Graph Pad, Inc)¹⁴. In addition fluorescence intensity curves for each ligand concentration were also made from the background reduced values.

4.4.CHARACTERIZATION OF AFTD

4.4.1.Spectroscopy Characterization of AFTD

Characterization was conducted for 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) using a 50µM concentration. Experiment was run using a BioTek Synergy 4 Microplate reader. Using a microplate (Costar) 96 well black bottom for emission spectrum and

a clear microplate (Falcon) 96 well plate was used for absorbance and excitation spectrums. Sample was buffered to a pH of 7.5 (20mM 7.3 HEPES, 0.01% tween (v/v)) and ran at 37°C. A blank well that contained buffer (20mM 7.5 HEPES) and water was also used. All wells were topped up with distilled water to a final volume of 100 μ L. Initially an absorbance spectrum was run by collecting values every 5nm. This indicated a maximum absorbance at 270nm. This value was used as excitation wavelength for the emission spectrum, which was run from 300-700nm at a sensitivity of 100. Emission spectrum indicated a maximum emission at 535nm. This value was then used to run an excitation spectrum in the Falcon clear plate from 250-700nm. This indicated that the maximum excitation was at 260nm with a secondary peak at 435nm. In addition, the secondary peak max was found at 435nm. Emission spectrum was re-run using 260nm instead of 270nm from 300-700nm using a sensitivity of 100. Once again the maximum emission was shown to be 535nm. From these results, a suitable excitation filter (460/40nm), emission filter (528/20nm) and dichromatic mirror (510nm) were identified for use in the binding assays.

4.4.2. Standard Curve and AFTD Equilibrium

A standard curve for AFTD was made by testing 6 concentrations (300 μ M, 100 μ M, 50 μ M, 25 μ M, 10 μ M, 1 μ M). This was done using a sensitivity of 80 in a microplate (Nunclon®) at a pH of 7.5 (20mM HEPES, 0.01% tween (v/v)) at an incubation temperature of 37°C. Fluorescence intensity values were taken using a filter set combination (Ex 460/40nm, Em 528/20nm, and dichromatic mirror 510nm). The 25 μ M AFTD was also included in well with a set amount of E6 protein (0.04mg/ml) as a single point. Reads were taken every 15minutes under the same conditions described for the standard curve except fluorescence intensity values were taken in both the perpendicular and parallel directions to calculate polarization. This was done in

order to determine the time point our system reaches equilibrium. Data analysis was conducted as described in section 4.5.1.

4.5.FLUORESCENCE POLARIZATION CONCENTRATION CURVE TO DETERMINE EC50

AFTD was tested in a fluorescence polarization assay at seven concentrations (300 μ M, 100 μ M, 50 μ M, 25 μ M, 15 μ M, 10 μ M, 1 μ M) with a fixed concentration of protein E6 (0.04mg/ml). AFTD concentrations were accurately measured using 1mM and 100 μ M stock solutions dissolved in half water/half ethanol. E6 protein was accurately measured from the stock concentration (0.201mg/ml). The assay was buffered to a pH of 7.5 using 20mM HEPES, 0.01% tween (v/v) and incubated to a temperature of 37°C. Control wells were included for the protein alone, and AFTD without protein at each concentration tested. Samples were plated as duplicate wells in a 96-well black bottom microplate (Nunclon®). The experiment was run using a Synergy 4 Microplate Reader (Biotek). An excitation filter (460nm/40) and emission filter (528nm/20) were used with a dichromatic mirror (510nm). Fluorescence polarization was read using a probe height of 4mm, a 350 μ sec delay between wells, and 40 reads per well. Data was collected every 15 minutes for 75 minutes using a sensitivity value of 80 with Gen5 software (Biotek). The read settings and filter/mirror combination were optimized during preliminary experiments. Fluorescence polarization values were calculated with blank values discounted.

The reaction was initiated by first adding water, solvent, and buffer. Once these amounts were measured into the reaction wells AFTD was then accurately added from the 1mM or the 100 μ M stock solutions. Once all ingredients were in each well the protein was added to each protein-containing well from the stock solution to initiate the binding reaction. After which the

plate was incubated in the microplate reader under conditions described above. Readings were taken every 15 minutes starting at time zero to 120 minutes using Gen5 software (Biotek)¹⁷.

4.5.1. a) Data Analysis

Data was exported into Excel 2010 (Microsoft, Inc)¹⁶ from Gen5 (Biotek), where values for parallel and perpendicular intensities were manipulated for each time point. The blank was discounted by hand using excel in both perpendicular and parallel directions. The perpendicular values were manipulated using an experimentally determined G-Factor (0.96) ($G\text{-Factor} = \frac{\text{Parallel intensity} \cdot (1 - 0.02)}{\text{Perpendicular intensity} \cdot (1 + 0.02)}$)¹⁷. Calculating the G-factor is required to correct for instruments that have a variety of different optical designs. Using these values polarization was calculated in Microsoft excel according to equation ($\text{Polarization} = \frac{(\text{parallel} - G\text{-factor} \cdot \text{perpendicular})}{(\text{parallel} + G\text{-factor} \cdot \text{perpendicular})}$)¹⁷. Values were then imported into Prism 6 (Graph Pad, inc)¹⁴ where graphing and curve fitting took place. The second well value for polarization in the 300 μM was dropped from this calculation as it was inconsistent across the time points. The error bars represent the standard error of the duplicate polarization values. EC50 was estimated using Prism's non-linear dose-response fitting function (four parameters)¹⁴.

4.6. CELL BASED ASSAYS

4.6.1. Cell culturing

4.6.1. a) Cell lines and Routine Maintenance

CaSki, SiHa, and C33A cells were obtained for ATCC. CaSki cells contain 200-300¹⁸ copies of E6 protein, whereas SiHa only contain 1-2 copies¹⁹. C33A cells contain zero copies of E6²⁰. All cells were cultured in DMEM complete medium (10% fetal bovine serum and 1X anti/anti) obtained through Fisher Co. Cells were held in an incubator at constant conditions (37°C, 5% CO₂). These cells were feed every 2-3 days and split when they reached 60-80%

confluent. All cells were routinely screened for *Mycoplasma* contamination, which is a common contaminant in cell cultures and none was found.

4.6.1. b)Freezing and Storage

Freezebacks were made by adding 10% DMSO to complete DMEM medium. Cells were then frozen at -80°C then moved into liquid nitrogen.

4.6.1. c)Thawing Cells

Cells were taken out of liquid nitrogen then added to 10ml of complete medium (DMEM) and centrifuged for 5 min at 750rpm. Centrifugation is used in order to make sure that all DMSO is removed from the cells as this can be cytotoxic. Cells are then, re-suspend in clean complete medium in a T75 flask (Fisher, Inc).

4.6.2.Cell Viability

Cells were prepared as described in routine maintenance. Cells were then removed from flasks and seeded into 24 well plates (flat bottom, Fisher). Approximately 25,000 cells were in each well prior to application of AFTD. AFTD was added to the cells (CaSki, SiHa, and C33A) 24 hours after seeding. This was done by adding a calculated amount of AFTD stock (2mM) in complete medium (DMEM) accurately for each concentration (25µM and 10µM) used to make up treated medium which was added directly to each well (500µL). The effect of ethanol, used as solvent for AFTD, was controlled by adding the same volumes respectively for each concentration (25µM and 50µM) of a solution that contained half ethanol and half de-ionized water. The original medium (500µL) in each well was removed and the treated medium containing AFTD and control medium, which contained ethanol was then added (500µL). For each time point, there were three treated wells with AFTD at each concentration tested and 3

controls, which were treated with ethanol at the same percentage by volume as the corresponding concentration. Enough medium was made for all the plates created to serve each time point (3, 6, 12, and 24). Since no change was observed using a light microscope after 3 and 6 hours, the next time point observed was at 24 hours.

4.6.3.AFTD Cell Permeability

AFTD was added to each plate as described in (Section 4.5.2). However, the final concentrations in medium were 25 μ M and 50 μ M. Different time points were also used (1hour, 2hours, and 3hours). Plates were removed from the incubator at corresponding time points. Cells were then rinsed in 1X Phosphate Saline Buffer (500 μ L) of and then lysed for 1hour in lysing buffer (55% water, 32%DMSO, and 13% triton). The lysate was taken and transferred to 2ml centrifuge tubes and centrifuged at 750rpm for 5 minutes. The supernatant was added to a Nunclon® 96 well microplate at a volume of 100 μ L per well. Fluorescence was read using a Biotek Synergy 4 Microplate reader using an optical system that consisted of an excitation filter (460/40nm), emission filter (528/20nm), and mirror 510nm. 350 μ s delay was used with a 4mm probe height, 40 reads per well, and a sensitivity of 100. In addition, absorbance was also taken at 260nm, 280nm, and 535nm to control for cell density.

4.6.3. a)Data Analysis

Absorbance and fluorescence intensity for each time point was collected using Gen5 (Biotek) software. These values were then imported into Excel 2010 (Microsoft, Inc)¹⁶. Due to a large amount of bubbles in the microplate, the 1hour time point was omitted from analysis. The 2 hour and 3 hour time points were analyzed by importing values into Prism 6 (Graph Pad, Inc)¹⁴, where means and standard deviations were calculated for each of the triplicate samples. The means and standard deviations were calculated using Prism 6 (Graph Pad, Inc) column statistics

analysis. These values were then graphed as bar graphs. Sidak's multiple comparisons ANOVA was conducted using Prism 6 (Graph Pad, Inc) software¹⁴. This will calculate if any difference between the sample and its corresponding control is significant within a 95% confidence interval. Turbidity (optical density) at 535nm was used to verify that the density of cells in each sample and control wells were comparable. If there were discrepancies, then the measured fluorescence values would have to be corrected for the difference in cell density. Optical density values at 535nm were about the same for each well and, for this reason, no correction was needed. The ratio of absorbance values at 260nm (nucleic acids) and 280nm (proteins) was used to further verify that the cell cultures in each well were consistent.

4.7.6F4 ANTIBODY COMPETITION EXPERIMENT

A competition experiment was conducted varying the concentration of antibody 6F4 against fixed concentrations of AFTD (6 μ M) and E6 (0.04mg/ml). The concentrations of protein and AFTD were chosen based on the EC50 value determination of 6 μ M at a protein concentration of 0.04mg/ml (Section 5.4). The experiment was conducted using a 96 well Nunclon® black bottom microplate at 37°C and buffered to a pH of 7.5 (20mM 7.5 HEPES 0.01% tween (v/v)) in a Synergy 4 BioTek microplate reader. The AFTD 6 μ M wells were measured from a 100 μ M stock solution. The 100 μ M AFTD stock solution was diluted using de-ionized water from a 1mM stock solution in 2ml centrifuge tube. The 1mM AFTD stock solution was dissolved half ethanol half de-ionized water. All wells containing AFTD were topped up with 15 μ L of ethanol to make sure it did not precipitate out of solution. Antibody 6F4 concentrations varied from 4ng/ml all the way to 4x10⁻⁶ng/ml using 10 fold dilutions from a 6F4 antibody stock solution (2.69mg/ml). A standard antibody control with the antibody on its own in buffer solution was present at the highest concentration in a single well. A single protein control

well was also present. Controls wells were also present for the antibody at each concentration used (4ng/ml - $4 \times 10^{-6}\text{ng/ml}$) with AFTD ($6\mu\text{M}$). The ethanol was controlled by making sure each well controls and samples contained a consistent amount of ethanol ($15\mu\text{L}$). All wells were topped up to $100\mu\text{L}$ using deionized water. Perpendicular and parallel fluorescence intensity values were collected every 15 minutes using an optical system that consisted of an excitation filter ($485/20\text{nm}$), emission filter ($528/20\text{nm}$), and 510nm dichromatic mirror with a sensitivity of 80, a probe height of 4mm and a time delay of $350\mu\text{sec}$ per well.

4.7.1. a) Data Analysis

Manipulation of data to calculate polarization values were done in the same fashion as described in 4.4.1a). One of the 4ng/ml of antibody 6F4 was omitted from the dataset as inconsistencies were observed. Polarization was plotted against concentration of antibody in a log scale to see if any concentration-dependent behaviour had taken place.

4.8.O-SUCCINYL-L-HOMOSERINE CONJUGATE

4.8.1.Synthesis of O-succinyl-L-homoserine-Bodipy

O-Succinyl-L-homoserine-Bodipy was synthesized using 1:1:1 equivalents of Bodipy-NHS, O-succinyl-L-homoserine, and triethylamine under argon gas. Reaction was initiated by dissolving 5mg of Bodipy in $500\mu\text{L}$ of dimethylformamide (DMF). The reaction vial was placed, with stirring, in a water/ice bath at 0°C . While in the water/ice bath, one equivalent (2.730mg) of O-succinyl-L-homoserine was added and mixed, followed by $1.39\mu\text{L}$ of triethylamine. The remainder of the reaction vial was filled with argon gas. The reaction was left to proceed for 30 minutes at 0°C after which the vial was removed from the water/ice bath. Once removed from the water/ice bath, the reaction vial was left to stand at room temperature overnight. Separation was performed using water and a small amount of ethyl acetate. Our product was found to be in the

aqueous phase. This was confirmed using TLC (5:1 dichloromethane:methanol). Purity was also verified at 95% using TLC (5:1 dichloromethane:methanol). The sample was then freeze dried, and stored at -20°C. 3mg of product was obtained from this reaction. Liquid chromatography mass spectrometry (LC-MS) was used to confirm the identity of O-succinyl-L-homoserine-Bodipy(supplementary information).

4.8.2.Spectroscopic Characterization of O-succinyl-L-homoserine-Bodipy

Characterization of O-succinyl-L-homoserine-Bodipy was conducted using a 50 μ M concentration. This was conducted at 22°C in a Synergy 4 microplate reader (BioTek) buffered to a pH of 7.5 (20mM HEPES, 0.01% tween (v/v)). Absorbance spectrum was conducted using a microplate (Falcon) clear 96 well and absorbance spectrum was acquired collecting data every 5nm. An absorbance spectrum indicated a maximum absorbance at 500nm. This value was used as the excitation value to conduct an emission spectrum using a sensitivity of 70 and a probe height of 4mm with the same concentration of O-succinyl-L-homoserine-Bodipy (50 μ M) and was conducted in a black bottom microplate (Nunclon®). This demonstrated a maximum emission of 510nm. This value was then used as the emission wavelength to conduct an excitation spectrum at the same concentration sensitivity and probe height. The maximum excitation was confirmed to be 500nm. These values allowed us to determine a proper filter set combination to be used in the ligand binding assays, which includes an excitation filter (485/20nm) emission filter (528/20nm) and a dichromatic mirror (510nm).

4.8.3.Fluorescence Polarization Assay

O-succinyl-L-homoserine-Bodipy was tested in a fluorescence polarization assay at three concentrations (1, 10, and 100 μ M) with a fixed concentration of protein E6 (0.04mg/ml). The assay was buffered to a pH of 7.5 using 20mM HEPES, 0.01% tween (v/v) and incubated to a

temperature of 37°C. Control wells were included for the protein alone, and O-Succinyl-L-homoserine-Bodipy alone (at each concentration tested). Samples were plated as single wells in a 96-well black bottom microplate (Nunc®). The experiment was run using a Synergy 4 Microplate Reader (Biotek). An excitation filter (485nm/20) and emission filter (528nm/20) were used with a dichromatic mirror (510nm). Fluorescence polarization was read using a probe height of 4mm, a 350µsec delay between wells, and 40 reads per well. Data was collected every 15 minutes for 75 minutes using a sensitivity value of 65 with Gen5 software (Biotek). The read settings and filter/mirror combination were optimized during preliminary experiments. Fluorescence polarization values were calculated with blank values discounted. A standard curve was also included for the concentrations tested and included in the characterization of this molecule.

4.8.3. a) Data Analysis

Manipulation of data to calculate polarization values was done in the same fashion as described in 4.4.1a). The second well value for polarization in the 4ng/ml well was dropped from this calculation as it was inconsistent across the time points. Each concentration tested was plotted against time. The error bars represent the standard error of the duplicate polarization values.

4.9. REFERENCES

1. PubChem, PubChem3D.
2. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **2001**, *46* (1–3), 3-26.

3. Gasteiger, J.; Marsili, M., Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron* **1980**, *36* (22), 3219-3228.
4. Berl, V.; Huc, I.; Khoury, R. G.; Lehn, J.-M., Helical Molecular Programming: Folding of Oligopyridine-dicarboxamides into Molecular Single Helices. *Chemistry – A European Journal* **2001**, *7* (13), 2798-2809.
5. Group, C. C. *MOE, version 2010.10* Montreal Canada 2010
6. O'Boyle, N.; Banck, M.; James, C.; Morley, C.; Vandermeersch, T.; Hutchison, G., Open Babel: An open chemical toolbox. *Journal of Cheminformatics* **2011**, *3* (1), 33.
7. Symyx, Discovery Gate. 2008 ed.; 2008.
8. Mayo, S. L.; Olafson, B. D.; Goddard, W. A., DREIDING: a generic force field for molecular simulations. *The Journal of Physical Chemistry* **1990**, *94* (26), 8897-8909.
9. *Yet Another Another Scientific Reality Application (YASARA)* 10.8.2; Yasara Biosciences Inc: 2010
10. Shared Hierarchical Academic Research Computing Network (SHARCNET)
11. Ramjan, Z. H.; Raheja, A.; Floriano, W. B.; Ieee, A Cluster-Aware Graphical User Interface for a Virtual Ligand Screening Tool. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vols 1-8*, Ieee: New York, 2008; pp 4102-4105.
12. Floriano, W. B.; Vaidehi, N.; Zamanakos, G.; Goddard, W. A., HierVLS Hierarchical Docking Protocol for Virtual Ligand Screening of Large-Molecule Databases. *Journal of Medicinal Chemistry* **2003**, *47* (1), 56-71.
13. Brady, G. P., Jr.; Stouten, P. W., Fast prediction and visualization of protein binding pockets with PASS. *J Comput Aided Mol Des* **2000**, *14* (4), 383-401.
14. *Prism, 6 UPDATE for Windows*, GraphPad Software: La Jolla, California USA, 2013.

15. R. A. Laskowski, M. W. M., D. S. Moss and J. M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **1993** 26 (2), 283-291.
16. *Excel*, 2010; Microsoft: 2010.
17. Gen5 Windows 7 ed.; BioTek.
18. O'Leary, J. J.; Browne, G.; Johnson, M. I.; Landers, R. J.; Crowley, M.; Healy, I.; Street, J. T.; Pollock, A. M.; Lewis, F. A.; Andrew, A., PCR in situ hybridisation detection of HPV 16 in fixed CaSki and fixed SiHa cell lines. *Journal of Clinical Pathology* **1994**, 47 (10), 933-938.
19. Heiles, H. B.; Gensch, E.; Kessler, C.; Neumann, R.; Eggers, H. J., In situ hybridization with digoxigenin-labeled DNA of human papillomaviruses (HPV 16/18) in HeLa and SiHa cells. *BioTechniques* **1988**, 6 (10), 978-981.
20. Jeon, S.; Lambert, P. F., Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis. *Proceedings of the National Academy of Sciences* **1995**, 92 (5), 1654-1658.

Chapter 5. Results and Discussion

5.1. COMPUTATIONAL SCREENING

Computational screening was conducted using two different structures of HPV16 E6, in order to make sure both C-terminus and N-terminus domains were analyzed. This will allow us to be able to properly evaluate potential lead compounds that can then be used for further experimental testing. Note that the C-terminus domain was published first so more extensive analysis was conducted and the N-terminus was used in order to make further suggestions on potential molecular probes. Three libraries of molecules were downloaded using both PubChem and Discovery Gate online services. These libraries comprised a total of 10,264 molecules. One library was of known imaging agents; however, due to the lack of availability of these molecules and advantageous functional groups of the other databases, these were not fully examined. One of the other databases contained molecules with one or more fluorine atoms, which is advantageous for Positron Emission Tomography (PET). The third database was composed of molecules that contain a primary amine functional group. The primary amine compounds are suitable for labelling with amine-reactive fluorescent dyes available commercially. This represents a challenge in the identification of hits from this library. Compounds with promising binding energies to the target structures (hits) were selected for experimental testing from the fluorinated and the primary amine databases.

5.1.1.N-Terminus Computational Screening (2LJX)

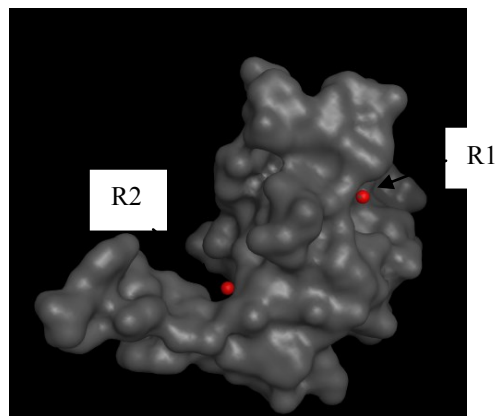


Figure 15: Full surface structure of 2LJX N-terminal structure of E6. The red dots show the two binding sites R1 and R2 as determined by PASS¹.

All three databases were computationally screened against the N-terminal domain of HPV16 E6 (2LJX) using Virtual Ligand Screening (VLS). Through examining the potential energy of each confirmation, it was concluded that the structure (2LJX) was in good shape for molecular docking. Two binding regions (R1 and R2) were identified within this domain (Figure 15) using PASS¹. Non-binder/binder thresholds were calculated for each region of the structure (Table 4).

Table 4: Summary of HierVLS results for N-terminal domain of HPV16 E6 (2LJX)

Binding Region	Number of Compounds Submitted	Number of Molecules Scored	Binding Threshold (kcal/mol)	Number of Molecules that Passed Threshold
R1	10,264	2,911	-53.12	3
R2	10,264	6,009	-58.20	10

Non-binder/binder threshold was calculated as the mean of the full data set for each binding site plus two standard deviations from the mean. Assuming normal distribution, the selected ligands will be at the 95% confidence interval. For R1 this threshold was calculated to be (-53.12kcal/mol) (Figure 16) and for R2 (-58.20kcal/mol) (Figure 17).

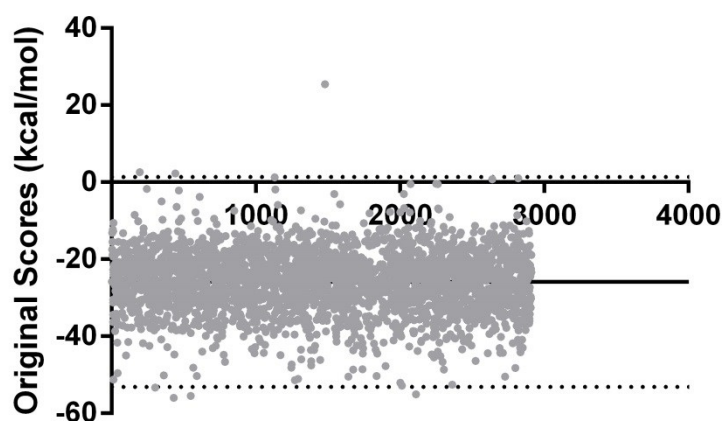


Figure 16: Scatter plot of results from the computational screening of NT-E6 (PDB code 2LJX) for binding region 1 (R1). The solid black line represents the mean score for the entire library docked to NT-E6. The dotted lines represent plus or minus two standard deviations from the mean, which was used to determine a binding threshold of -58.20kcal/mol. A total of three molecules passed the binding/nonbinding threshold. The complete list of hits passing the threshold for R1 is provided in Table 5.

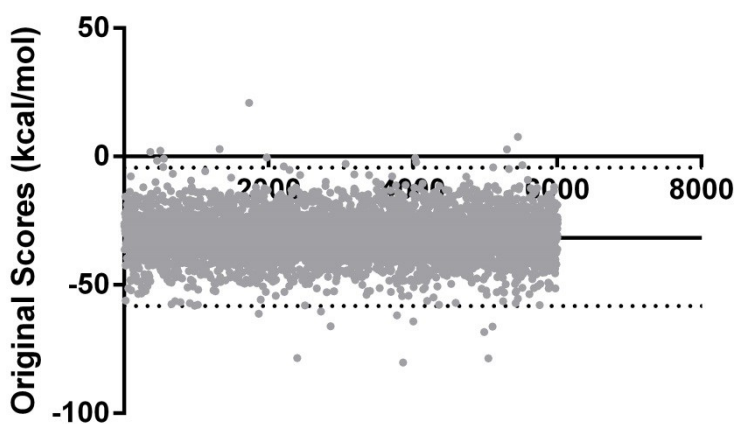


Figure 17: Scatter plot of results from computational screening of 2LJX for R2. The solid black line represents the mean and the dotted lines represent plus or minus two standard deviations. This was used to determine a binding threshold of -58.20kcal/mol. A total of 10 molecules based the binding/nonbinding threshold. The complete list of hits passing the threshold for R2 is displayed in table 6.

Table 6: Compounds that passed the threshold of -58.20kcal/mol HierVLS targeting regions R2 in NT-E6. Refer to Table B for molecular structures.

Rank	Identifier	Binding Score (kcal/mol)	IUPAC name	Smiles
1	AMINOPTERIN-3-5-7-9-3H-N	-80.13	Aminopterin-3-5-7-9-3H-N	<chem>O=C(O)C(NC(=O)C1C CC(CC1)NCC2NC3C(NC2)NC(NC3N)N)CC C(=O)O</chem>
2	PA24273	-78.58	N-(2,4-dinitrophenyl)-L-Arginine	<chem>OC(=O)[C@@H](Nc1 ccc([N+](=O)[O-])cc1[N+](=O)[O-])CCCNC(N)=N</chem>
3	PACLITAXEL-BENZOATE-RING-UL-14C-	-78.41	Paclitaxel-Benzoate-Rink-UL-14C	<chem>CC1=C2C(C(=O)C3(C (CC4C(C3C(C(C2(C)C) (CC1OC(=O)C(C(C5 =CC=CC=C5)NC(=O) C6=CC=CC=C6)O)O) OC(=O)C7=CC=CC=C 7)(CO4)OC(=O)C)O)C)OC(=O)C</chem>
4	MFCD01113139	-68.33	3-Chloro-2-(2-([3-oxo-2-Benzofuran-1(3H)-Ylinden]methyl)hydrazino)-5-(Trifluoromethyl)Pyridinium Acetate	<chem>C1c1cc(cnc1NNC=C\ 1/OC(=O)c2c/1cccc2) C(F)(F)F</chem>
5	MFCD00214696	-66.13	3-Chloro-2-(3-[1-(Phenylsulfonyl)-1H-Pyrazol-3-YL]Phenoxy)-5-(Trifluoromethyl)Pyridine	<chem>C1c1cc(cnc1Oe1cc(ccc 1)- c1nn(S(=O)(=O)c2cccc 2)cc1)C(F)(F)F</chem>
6	cmp450572	-66.06	2-amino-4-[[5-(6-aminopurin-9-yl)-3,4-dihydroxy-tetrahydrofuran-2-yl]methyl-sulfonio]butanoate	<chem>[S+](CC1OC(n2c3ncnc (N)c3nc2)C(O)C10)(C CC(N)C(=O)[O-])C</chem>
7	PA24775	-65.07	Beta-Nicotinamine Mononucleotide	<chem>P(OC[C@H]1O[C@@H]([n+]2cc(ccc2)C(=O)N)[C@H](O)[C@@H]1O)(O)(=O)[O-]</chem>
8	MFCD00172501	-61.85	4-(2-([3-Chloro-5-(Trifluoromethyl)-2-Pyridinyl] Amino) Ethyl) Benzenesulfonamide	<chem>C1c1cc(cnc1NCCc1ccc (S(=O) (=O)N)cc1) C(F) (F) F</chem>
9	MFCD15731617	-61.14	4-Anilino-2-([(4-Fluorobenzyl) (Methylsulfonyl) Amino] Methyl) Pyrimidine-5-Carboxylic Acid	<chem>S (=O) (=O) (N(Cc1ccc (F) cc1) Cc1nc (Nc2ccccc2)c(en1)C(O)=O)C</chem>
10	MONENSIN-3H-G	-60.28	Monensin-3H-G	<chem>01C(C(CC(C)C1(0)C0) C)C10C(C2(0C(CC2)C 2(0C3(0C(C(C(0C)C(C (0)=0)C)C)C(C)C(0)C 3)CC2)C)CC)C(C1)C</chem>

The compounds solanesyl-pyrophosphate, aminopterin, and paclitaxel were already available in the lab. However, information regarding vendors for the three other molecules needed to be obtained using information stored in the database (Table 7). Based on commercial availability, the compounds in Table 7 were recommended for experimental testing.

Table 7: Promising lead compounds from the NT-E6 virtual screening. Refer to Table C for molecular structures.

Computational Identifier	Rank/Region	Molecular name	Binding Energy Score (kcal/mol)	Commercial Availability	SMILES
MFCD01113139	4/R2	3-Chloro-2-(2-[(3-oxo-2-Benzfuran-1(3H)-Yliden]methyl)Hydrazino)-5-(Trifluoromethyl) Pyridinium Acetate	-68.33	Bionet	<chem>C1c1cc(cnc1NN\C=C\1/OC(=O)c2c1cccc2)C(F)(F)F</chem>
MFCD00214696	5/R2	3-Chloro-2-(3-[1-(Phenylsulfonyl)-(H-pyrazol-3-YL]Phenoxy)-5-(Trifluoromethyl) Pyridine	-66.13	Bionet	<chem>C1c1cc(cnc10c1cc(ccc1)c1nn(S(=O)(=O)c2cccc2)cc1)C(F)(F)F</chem>
PA24273	7/R2	N-(2,4-dinitrophenyl)-L-arginine	-78.58	Sigma Aldrich	<chem>OC(=O)[C@@H](Nc1ccc([N+](=O)[O-])cc1[N+](=O)[O-])CCCNC(N)=N</chem>

In addition to commercial availability, these ligands were also found to have specific interactions to the protein according to their HierVLS-predicted ligand-protein structures. Ligand interaction diagrams were generated using MOE 2010.10 (CC group)² and used to identify specific interactions between ligand and protein. The ligand interaction diagram for 3-Chloro-2-(2-[(3-oxo-2-Benzfuran-1 (3H)-Yliden] methyl) Hydrazino)-5-(Trifluoromethyl) Pyridinium Acetate bound to the NT-E6 (2LJX) shows the following non-bonded interactions between the ligand and protein E6: Hydrogen bonds Cys51 and Arg10 (Figure 18).

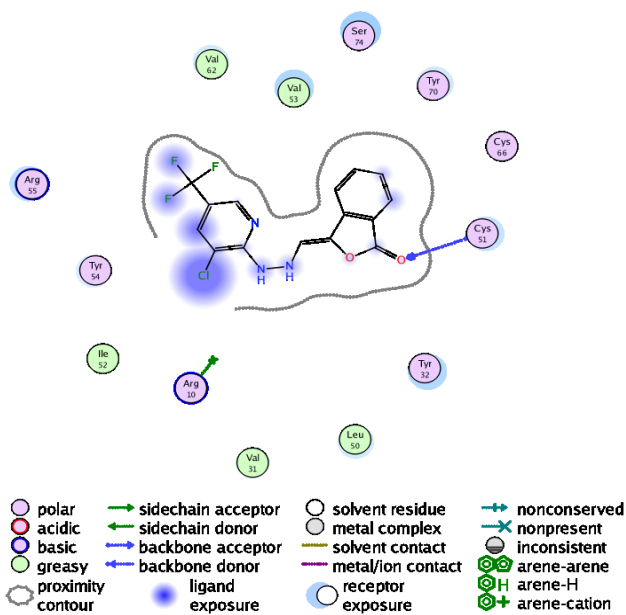


Figure 18: Ligand interaction diagram of 3-Chloro-2-(2-[(3-oxo-2-Benzfuran-1 (3H)-Yliden] methyl)Hydrazino)-5-(Trifluoromethyl) Pyridinium Acetate bound to N-terminal E6 (NT-E6).

The ligand interaction diagram for 3-Chloro-2-(3-[1-(Phenyl sulfonyl)-(H-pyrazol-3-YL] Phenoxy)-5-(Trifluoromethyl) Pyridine bound to the NT-E6 (2LJX) shows the following non-bonded interactions between the ligand and protein E6: Hydrogen bonds Cys51 as well as an arene hydrogen bond with Val53 (Figure 19).

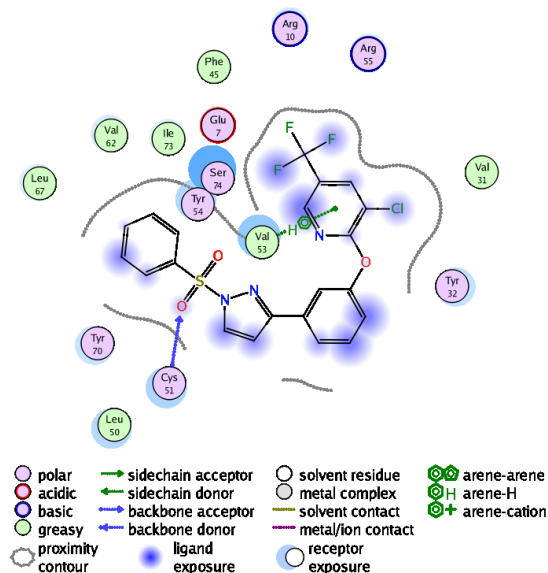


Figure 19: Ligand interaction diagram of 3-Chloro-2-(3-[1-(Phenyl sulfonyl)-(H-pyrazol-3-YL) Phenoxy]-5-(Trifluoromethyl) Pyridine bound to N-terminal E6 (NT-E6).

The ligand interaction diagram for N-(2,4-dinitrophenyl)-L-arginine) bound to the NT-E6 (2LJX) shows the following non-bonded interactions between the ligand and protein E6: Hydrogen bonds Cys51 and Arg10 (Figure 20).

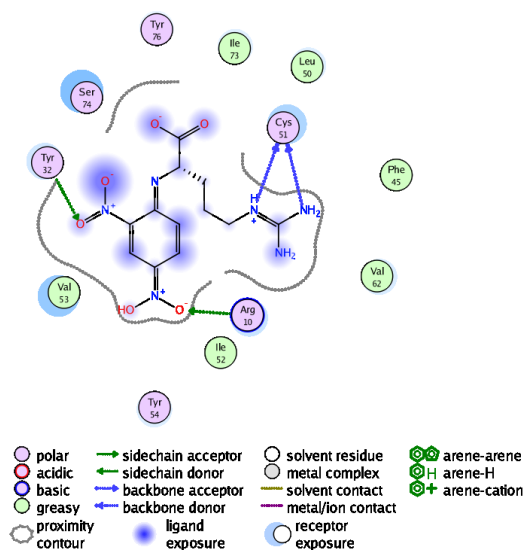


Figure 20: Ligand interaction diagram of N-(2,4-dinitrophenyl)-L-arginine to N-terminal E6 (NT-E6).

Based on these interactions we can conclude that these molecules interactions are specific to E6, making them highly favourable potential candidates for molecular probe development.

5.1.2.C-Terminus Computational Screening (2FKJ)

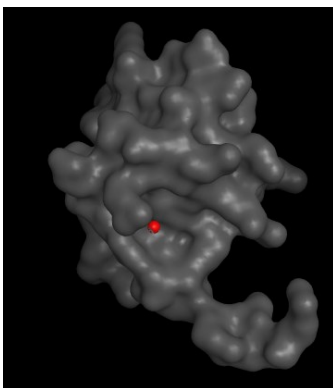


Figure 21: Surface representation of Protein E6 C-terminal domain (2FK4). The red dot represents the binding site as determined using PASS¹.

All three databases were also docked to the C-terminus HPV16 E6 structure (CT-E6), with PDB code 2FK4, which only had one binding region identified by Pass¹ (Figure 21). Prior to docking the protein structure (PDB 2FK4) was checked for quality of dihedral angles using a Ramachandran Plot (Figure 22). Most of these angles were found in the allowed region and for this reason this structure was deemed adequate for VLS.

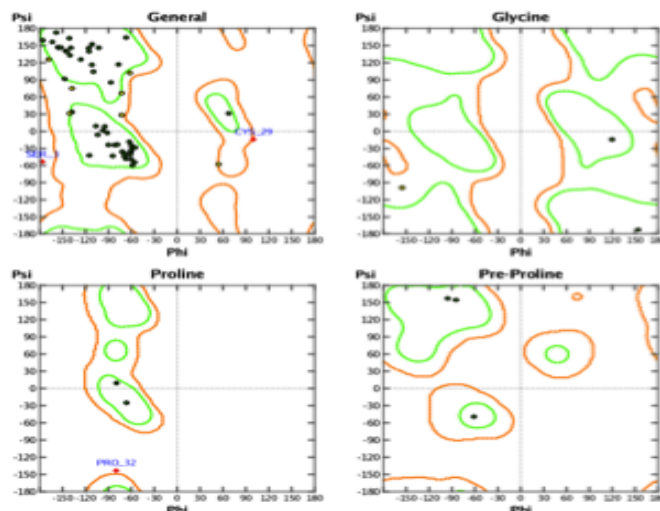


Figure 22: Ramachandran plot of HPV16 E6 C-terminal domain (PDB 2FK4) The plot was generated using MOE 2010.10 and used to check the dihedral angles of the experimentally determined structure prior to docking.

The screening of these databases against one binding site of E6 using a high performance computer cluster at the Shared Hierarchical Academic Research Computing Network (SHARCNET)³. A binding/nonbinding threshold (-35.47Kcal/mol) was established as the mean binding energy of the entire 3,339 compounds set plus two standard deviations from the mean (Figure 23). A summary of this screening is laid out in Table 8.

Table 8: Summary of HierVLS results for C-terminal domain of HPV16 E6 (2FK4)

Number of Compounds Submitted	Number of Molecules Scored	Binding Threshold (kcal/mol)	Number of Molecules that Passed Threshold
10,264	3,339	-35.47	295

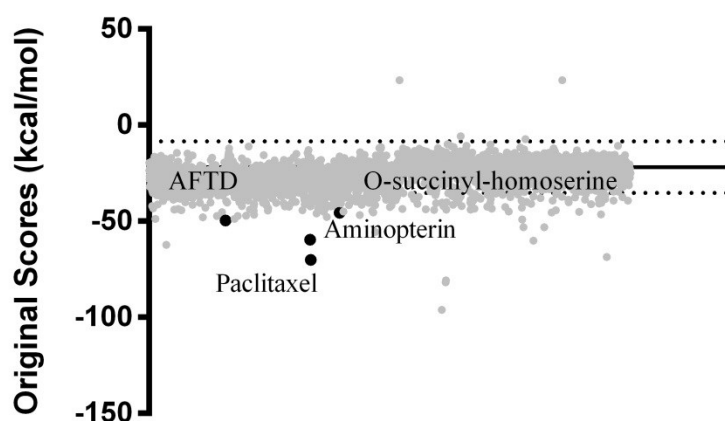


Figure 23: Force field scores for 3,339 compounds docked to CT-E6. binding/nonbinding threshold (-35.47Kcal/mol) was established as the mean binding energy of the entire 3,339 compounds set plus two standard deviations from the mean

. Establishing a statistical threshold is necessary because there are no available known small-molecule binders to HPV16 E6 that could be docked to the structure and used as reference.

Using this non-binder/binder threshold, three compounds from the primary amine database and five from the fluorine-containing database were selected for further analysis (Tables 9 and 10, respectively).

Table 9: Top three compounds from the primary amine database in order of rank. All molecules passed the binding/non-binding threshold of -35.47kcal/mol. Refer to Table D for molecular structures.

Compound ID	Compound Name	Binding Energy (Kcal/mol)	Commercial Availability	Smiles String
Cmp47400	L-mimosine	-46.81	Sigma Aldrich	<chem>OC1=CN(C=CC1=O)C[C@H](N)C(O)=O</chem>
Cmp144614	(S)-2-Amino-2-methyl-4-phosphonobutanoic acid	-45.83	Sigma Aldrich	<chem>P(O)(O)(=O)CC[C@](N)(C(O)=O)C</chem>
Cmp37088	O-succinyl-L-homoserine	-45.62	Sigma Aldrich	<chem>O(C(=O)CCC(O)(=O)CC[C@H](N</chem>

Table 10: Top 5 compounds from the fluorine-containing molecule database in order of rank. All these compounds passed the binding/nonbinding threshold of -35.47kcal/mol. Refer to Table E for molecular structures

Compound ID	Compound Name	Binding Energy (Kcal/mol)	Commercial Supplier	Smiles String
MFCD09040704	3-methyl-4(trifluoromethyl) Isoxazolo [5,4-B] Pyridin-6-YL-4-methyl benzene sulfonate	-62.22	Enamine	<chem>S(Oc1nc2onc(c2c(c1)C(F)(F)F)C(=O)(=O)c1ccc(cc1)C</chem>
MFCD14636665	3-Amino-5-fluorobenzo [E] [1,2,4] Triazine 1, 4 dioxide	-49.53	Sunbiochem	<chem>Fc1c2[n+]c(n+)[(O-)]c2ccc1N</chem>
MFCD00871869	Ofloxacin N-Oxide	-48.80	Sinochem	<chem>Fc1cc2c3N(C=C(C(O)=O)C2=O)[C@@](COc3c1N1CC[N+](O-)](CC1)C)C</chem>
MFCD00955225	(5-methylisoxazo-3-YL) ((2-(trifluoromethyl) phenyl) sulfonyl) amine	-48.70	ABCR and Ryan Scientific	<chem>S(=O)(=O)(Nc1noc(c1)C)c1ccccc1C(F)(F)F</chem>
MFCD11977074	Ethyl 6-chloro-2-(4-Fluorophenyl) imidazo [1,2-A] pyrind-3-carboxylate	-47.94	Golden Bridge Pharma	<chem>C1C=1C=Cc2n(C=1)c(C(OCC)=O)c(n2)-c1ccc(F)cc1</chem>

Both, the top three primary amines and top five fluorine-containing compounds, were further evaluated for commercial availability and safety, in addition to presence of specific ligand-protein binding interactions, as determined by ligand interaction diagrams created using MOE 2010.10 (CC group)⁴. These diagrams allow the visualization of the interactions that the ligand has with the residues of the protein. These interactions validate the binding energy score, since

favourable binding energies should correspond to specific interactions between ligand and protein.

The ligand interaction diagram for 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) shows hydrogen bonds (HBs) from the ligand to the Lys45, and Arg58 residues of the CT-E6 (2FK4) structure (Figure 24). These HBs indicate specific interactions between the partners, as opposed to non-specific atomic interactions. The ligand interaction diagram for O-succinyl-L-homoserine bound to the CT-E6 (2FK4) shows the following non-bonded interactions: HBs Arg40, Lys45, and Glu37 as well as a metal/ion contact with Arg40 (Figure 25). The presence of these non-bonded interactions justify the favourable binding energies calculated for this complex and, hence, increases our confidence that binding of this ligand to the target protein will be specific. This would be observed experimentally as a ligand concentration-dependent binding response.

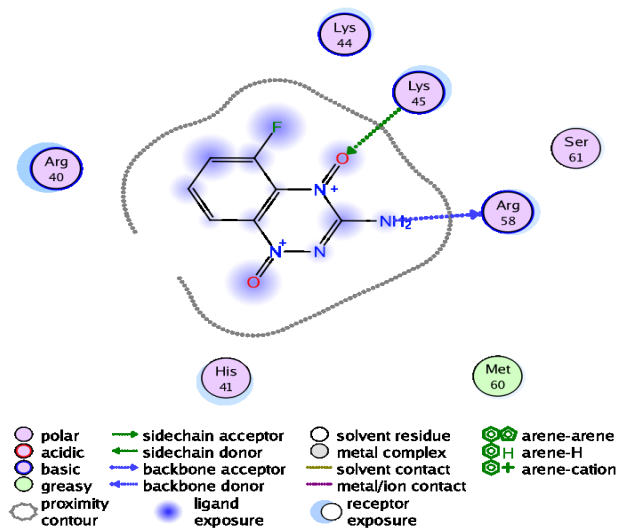


Figure 24: Ligand interaction diagram of 3-Amino-5-fluorobenzo [E] [1,2,4] Triazine 1, 4 dioxide (AFTD) bound to C-terminal E6 (CT-E6).

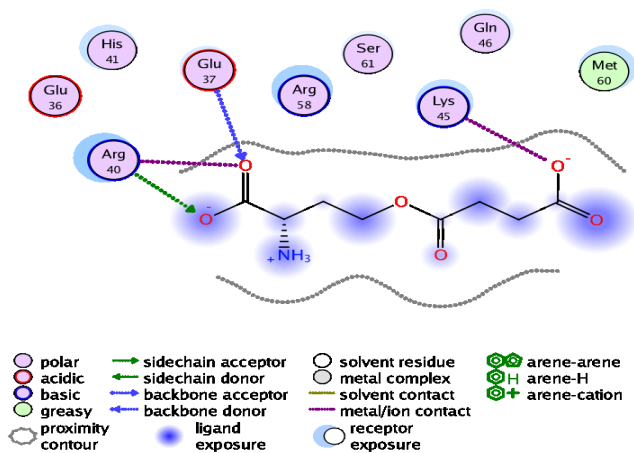


Figure 25: Ligand-protein interactions diagram of O-succinyl-L-homoserine bound to C-terminal (CT-E6).

With the objective of performing experimental assays, compounds that were not available commercially or were not safe to work with were deemed not useful for our purposes regardless of their binding energy score, and were eliminated from our list. Based on our selection criteria, O-succinyl-L-homoserine (O-succinyl) and 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) were selected for experimental testing (Table 11). Additionally, two of the shelf compounds, paclitaxel (-70.10 kcal/mol) and aminopterin (-59.61 kcal/mol), were found to have sufficiently high binding energy (above the non-binding/binding threshold), to justify experimental testing. The ligand interaction diagram for paclitaxel bound to the CT-E6 (2FK4) shows the following non-bonded interactions: H-bonds Arg 58, Lys45, and Arg40 (Figure 26). The ligand interaction diagram for Aminopterin bound to the CT-E6 (2FK4) shows the following non-bonded interactions: H-bonds Try7, Lys45, and Gln46 (Figure 27). Due to these specific interactions for both compounds, we expect binding to be specific.

The binding of all four compounds to the protein was assessed experimentally using intrinsic tryptophan fluorescence. Table 11 contains information about all ligands tested experimentally.

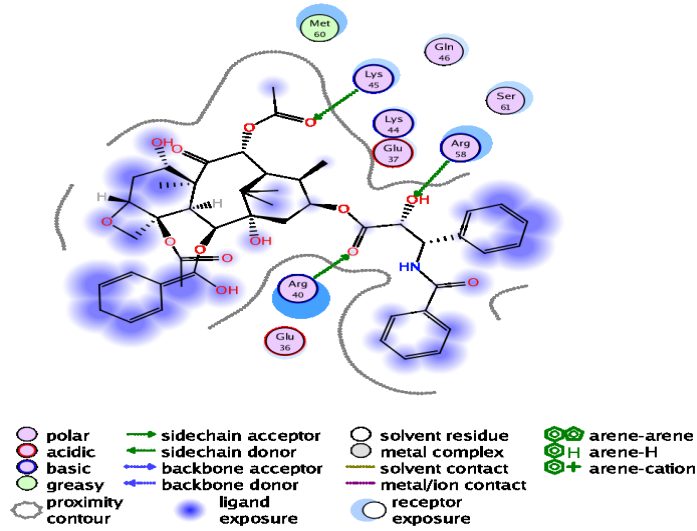


Figure 26: Ligand-protein interactions diagram of Paclitaxel bound to C-terminal E6 (CT-E6)

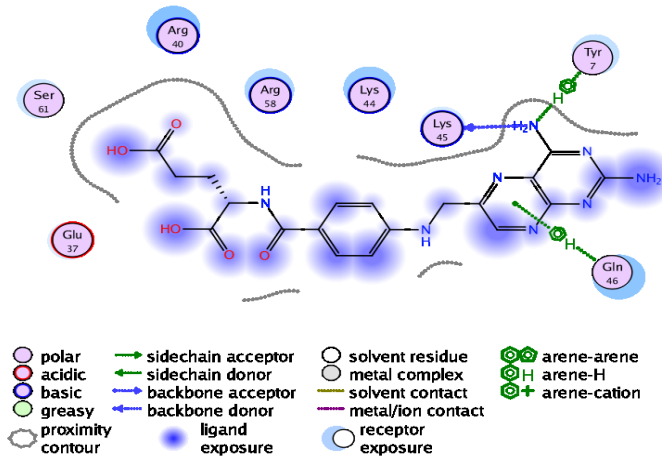


Figure 27: Ligand-protein interactions diagram of Aminopterin bound to C-terminal E6 (CT-E6)

Table 11: Ligands selected for experimental testing. All four ligands were selected based on binding scores, ligand interactions to the protein, and commercial availability. Refer to Table F for molecular structures.

Compound Name	Overall Rank	Binding Energy (kcal/mol)	Abbreviation	Similes
O-succinyl-L-homoserine	25	-45.62	O-succinyl	<chem>O(C(=O)CCC(O)(=O)CC[C@H](N)C(O)=O</chem>
3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide	12	-49.53	AFTD	<chem>Fc1c2[n+]c(n[n+](O-])c2ccc1)N</chem>
Aminopterin	8	-59.61	Amino	<chem>O=C(O)C(NC(=O)C1CCC(CC1)NCC2NC3C(NC2)NC(NC3N)N)CC(=O)O</chem>
Paclitaxel	4	-70.10	Pac	<chem>CC1=C2C(C(=O)C3(C(CC4C(C3C(C(C2(C)C)(CC1OC(=O)C(C(C5=CC=CC=C5)NC(=O)C6=CC=CC=C6)O)OC(=O)C7=CC=CC=C7)(CO4)OC(=O)C)O)C)OC(=O)C</chem>

5.2. INTRINSIC TRYPTOPHAN FLUORESCENCE EXPERIMENT

Intrinsic tryptophan fluorescence assays were employed in order to test against E6 ligands from computational screening that are not fluorescent labelled or intrinsically fluorescent. In these assays we hope to observe a change in the protein's intrinsic tryptophan fluorescence that is greater than 30% in the presence of the ligand. This will indicate a significant change to the environment of the tryptophan, suggesting a binding event. Principles of tryptophan fluorescence are outlined in the methodology section of this thesis (3.2.2).

The tryptophan (Trp) fluorescence assay was performed for paclitaxel, O-succinyl-L-homoserine and AFTD in an attempt to obtain binding constants which would facilitate the fluorescence polarization dose-response curve assay. The ligands were tested at 300 μ M, 100 μ M, 25 μ M, 10 μ M, 1 μ M, 0.1 μ M, and 0.01 μ M. Paclitaxel was tested at 15 μ M, 10 μ M, and 1 μ M in order to make sure it did not precipitate out at higher concentrations as it is not soluble unless the

concentration of DMSO is high. We wanted to keep the concentration of DMSO low as it could denature E6 protein. Preliminary experiments showed that a protein concentration of 0.02mg/ml was sufficient to obtain a good signal (See appendix A). However, to improve signal even further, the protein concentration was increased to 0.03mg/ml. Proper solvent controls were provided for DMSO, which is used to solvate paclitaxel. Unfortunately, our instrumentation setup was not adequate for obtaining emission spectra in the < 340 nm range with excitation at 288 nm. This prevented the use of intrinsic Trp fluorescence for determining binding constants. We instead used the percent change at 350 nm emission to probe binding. Percent change (% change) was considered indicative of binding if > 30%. An E6-specific antibody (6F4) was used as positive control and induced a change greater than 30%. The change in intrinsic fluorescence induced by 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine -1, 4 Dioxide (Figure 28) shows concentration-dependency, whereas the presence of O-succinyl-L-homoserine induces quenching albeit not concentration-dependent (Figure 29). In addition, concentration dependence was observed for Paclitaxel (Figure 30).

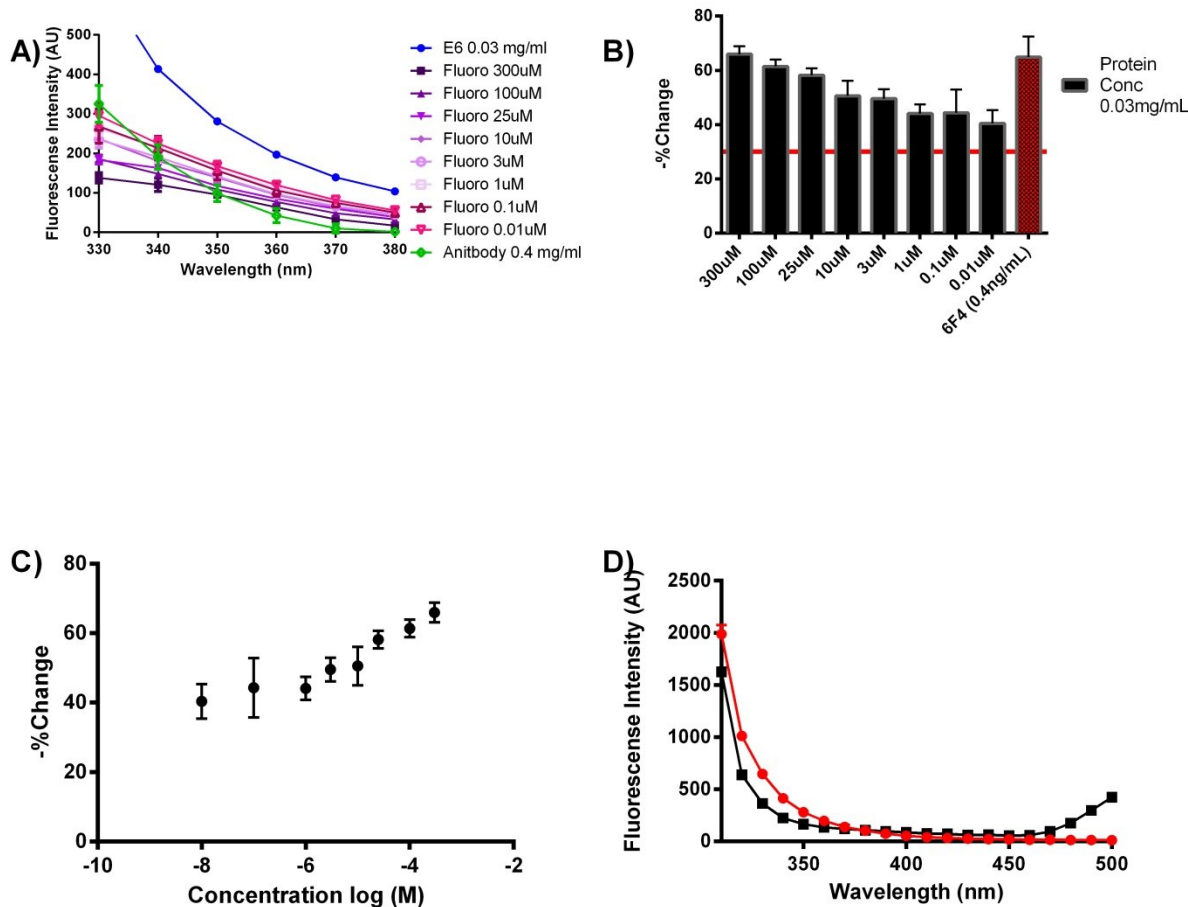


Figure 28: Change in intrinsic tryptophan fluorescence at 350nm induced by 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) after 30minutes incubation in the presence of E6. A) Fluorescence intensity. B) Bar graph %change. C) log scale % change plot. D) AFTD 300µM (black) and protein alone (red). Reaction was incubated at a temperature of 37°C and buffered to a pH of 7.3 (20mM HEPES 7.3, 0.01% tween (v/v)). Protein concentration was held constant at 0.03mg/ml. 6F4 antibody was used as a positive control at a concentration of 0.4ng/ml AFTD concentration was varied (300µM-0.01µM). The red line in graph B represents the 30% change baseline. Experiment was validated by our positive control, 6F4 antibody, producing a change above 30%. The effect of AFTD is concentration-dependent.

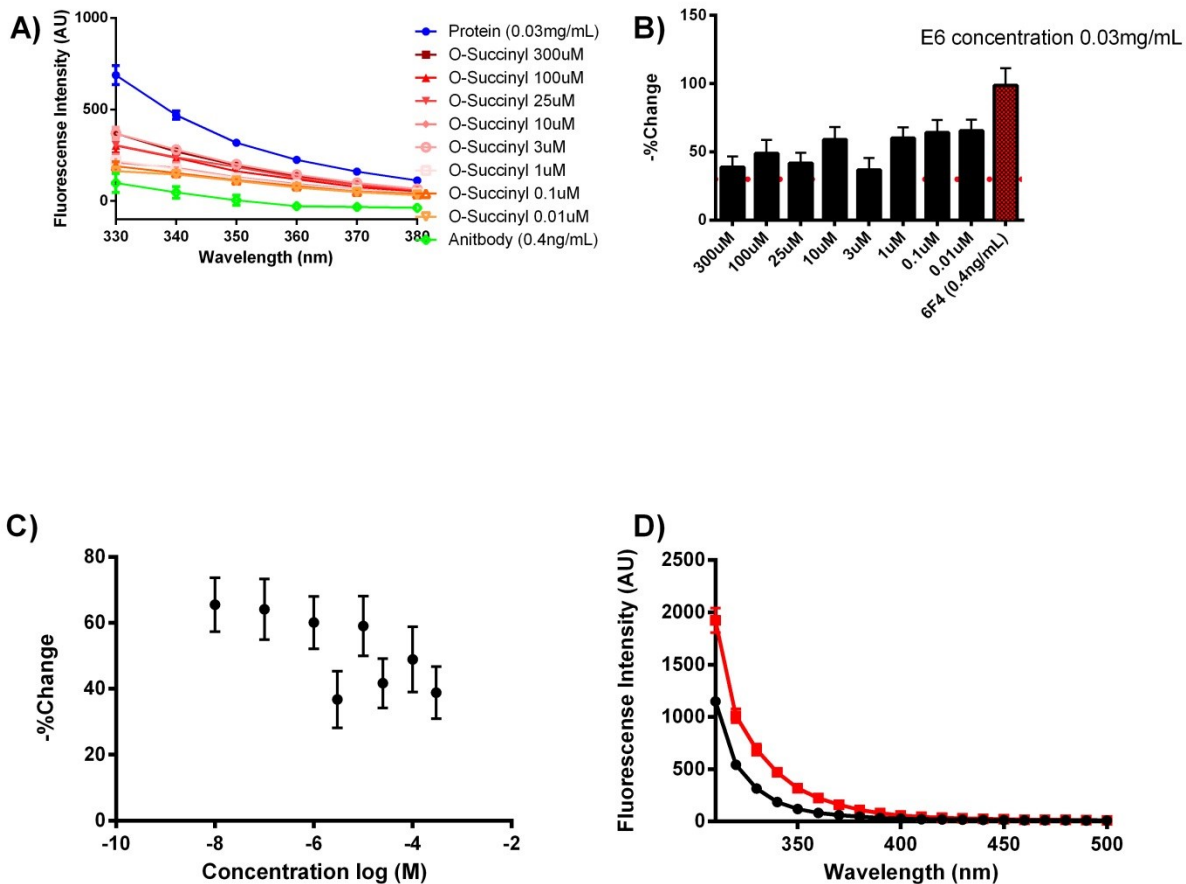


Figure 29: Change in intrinsic tryptophan fluorescence at 350nm induced by O-succinyl-L-homoserine after 30minutes incubation in the presence of E6. A) Fluorescence intensity. B) Bar graph %change. C) log scale % change. D) Plot of O-succinyl-L-homoserine 300μM (black) and protein alone (red). Reaction was incubated at a temperature of 37°C and buffered to a pH of 7.3 (20mM HEPES 7.3, 0.01% tween (v/v)). Protein concentration was held constant at 0.03mg/ml. 6F4 antibody was used as a positive control at a concentration of 0.4ng/ml. O-succinyl-L-homoserine concentration was varied (300μM-0.01μM). The red line in graph B) represents the 30% change baseline. Experiment was validated by our positive control, 6F4 antibody, producing a change above 30%. The effect of O-succinyl-L-homoserine did not follow a concentration-dependence, however showed multiple concentrations where quenching above 30% was observed.

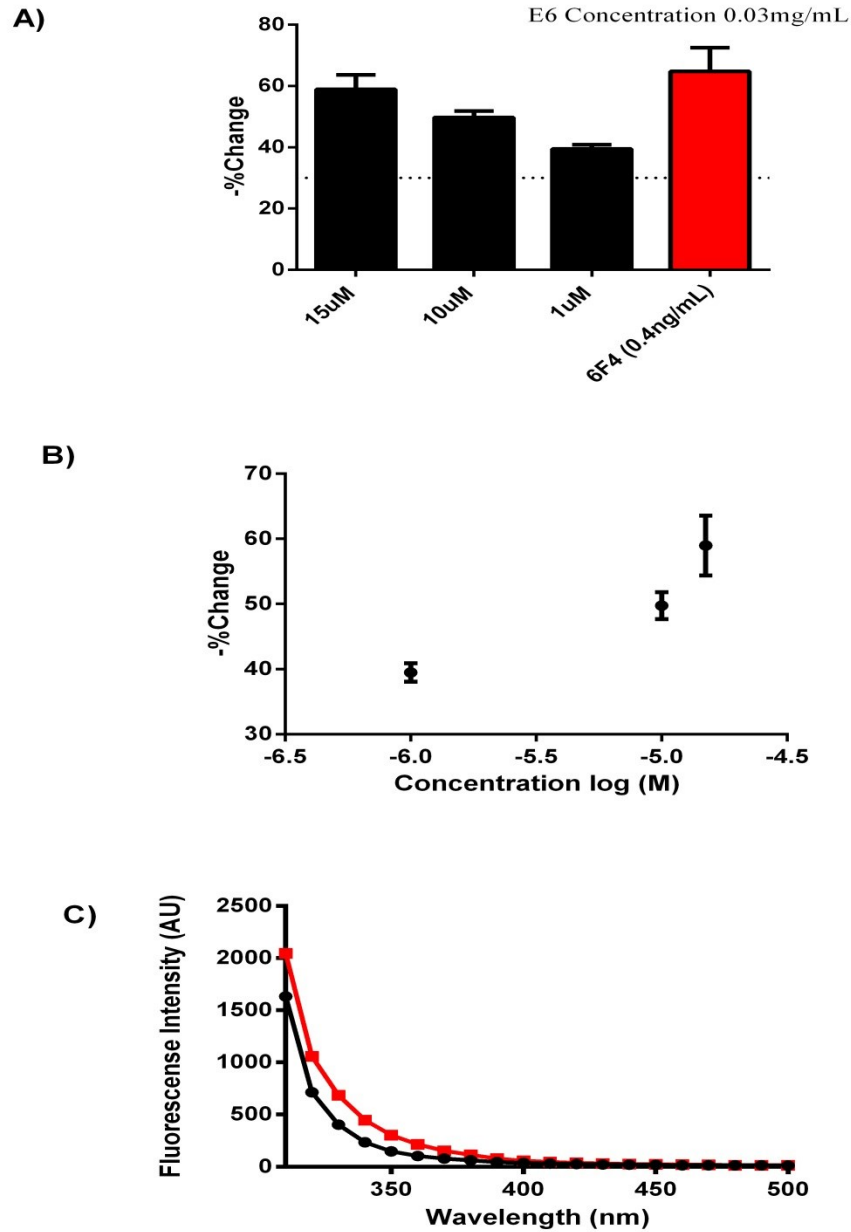


Figure 30: Change in intrinsic tryptophan fluorescence at 350nm induced by Paclitaxel after 30 minutes incubation in the presence of E6. Reaction was incubated at a temperature of 37°C and buffered to a pH of 7.3 (20mM HEPES 7.3, 0.01% tween (v/v)). A) Bar graph % change. B) %change log scale concentration. C) Protein alone 15µM control (red) and paclitaxel on its own 15µM (black). Protein concentration was held constant at 0.03mg/ml. 6F4 antibody was used as a positive control at a concentration of 0.4ng/ml. Paclitaxel was tested in three concentrations (15µM, 10µM, and 0.01µM). The dotted line in graph A) represents the 30% change baseline. Experiment was validated by our positive control, 6F4 antibody, producing a change above 30%. The effect of paclitaxel is concentration-dependent.

Although O-succinyl-L-homoserine did not show concentration-dependence it still induced a change in fluorescence greater than 30% at some concentrations. However, further ligand binding assays are needed before we can conclusively say binding has occurred. Paclitaxel will also need to be further tested for verification. The antibody used also appears to be giving inconsistent results. This is believed to be due to the antibodies sensitivity to the environment and the lack of sodium azide in the antibody's buffer. Sodium azide is usually in antibody buffers to prevent microbial contamination. Since the antibody had been in and out of the fridge between experiments, contamination might have occurred. In addition the antibody may have denatured or degraded when exposed to the external environment even for a short period of time.

5.3.CHARACTERIZATION OF AFTD

5.3.1.Spectroscopic Characterization of AFTD

In order to determine a proper filter set for AFTD, spectroscopy characterization needed to be conducted. Through this characterization a maximum excitation at 260nm with a second maximum at 435nm, and a maximum emission at 535nm was determined (Figure 31). This allows us to decide on a proper filter set suitable for AFTD, which contained an excitation filter (460/40nm), emission filter (528/20nm), and dichromatic mirror (510nm).

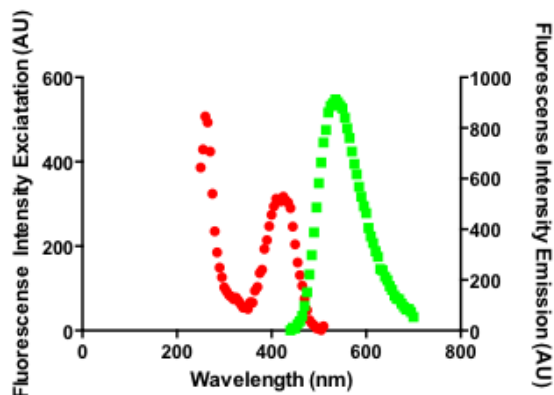


Figure 31: Excitation/Emission spectrum of 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (Fluoro). Experiment was run at 37°C at a pH of 7.5 (20mM 7.3 HEPES, 0.01% tween (v/v)). A maximum excitation was observed at 260nm with a secondary max at 435nm. Maximum emission was shown at 535nm.

5.3.2. Standard Curve for AFTD

A standard curve for AFTD, which was in the same microplate was used to show linearity in the concentration range used for the dose-response curve with a R^2 value of 0.99 (Figure 32). This experiment also demonstrated that a sensitivity of 80 is sufficient for the concentration range chosen of AFTD (300 μ M-1 μ M).

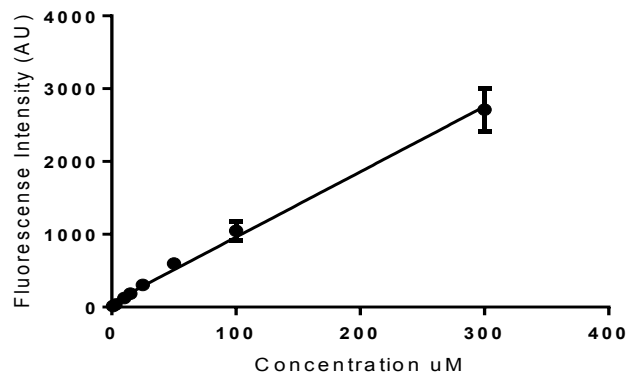


Figure 32: Standard curve for AFTD. This assay was conducted at an incubation temperature of 37°C using a BioTek Synergy 4 microplate reader. Data was taken only 2 minutes after incubation. Each concentration was buffered to a pH of 7.5 (20mM HEPES, 0.01% tween (v/v)) with a sensitivity of 80 and a probe height of 4mm. A correlation coefficient R^2 value of 0.99 was determined, indicating a high degree of linearity.

5.3.3. Equilibrium Test for AFTD bound to E6

In order to gain further information in regards to the equilibrium of our system, change in polarization was tested over time using a 25 μ M concentration of AFTD (Figure 33).

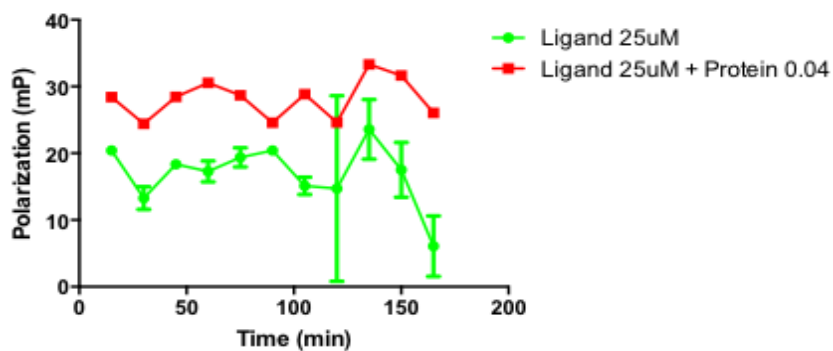


Figure 33: Time plot of AFTD and E6 complex. This was conducted at an incubation temperature of 37°C using a BioTek Synergy 4 microplate reader. Each concentration was buffered to a pH of 3.7 (20mM HEPES, 0.01% tween (v/v)). Excitation filter (460/40nm) and emission filter (528/20nm) was used and 510nm dichromatic mirror. A sensitivity of 80 was used as well as a probe height of 4mm. Our system does not follow any particular pattern however, a few time points show great separation. Through this experiment, it was evident that our system fluctuates over time. It also shows that the signal from AFTD appears to go down after the two hour time point. For this reason data was to be collected every 15 minutes for a two hour period.

5.4.FLUORESCENCE POLARIZATION AND EC50 DETERMINATION

Fluorescence polarization was used to determine the EC₅₀ value of AFTD, which was found to be a positive ligand in the intrinsic Trp fluorescence assay, against E6,. Through this assay we hoped to obtain an EC₅₀ for AFTD that was low enough to indicate sufficient affinity to E6 for molecular probe. The basic principles of this assay were outlined in section 3.2.3. In addition basic principles of dose-response curves, which will be used to estimate the EC₅₀, are discussed in section 3.2.4.

Prior to running the FP curve, calibration experiments (Section 5.3) and preliminary attempts (Appendix A) were ran in order to determine concentration ranges and parameters. With the information gathered in the calibration experiment, a concentration range for AFTD (1 μ M-300 μ M) was determined. Results indicated that 75 minutes was the best curve (Figure 34). This was not against what was shown in the equilibrium experiment (5.3.3) as the system appeared to oscillate over the various time periods. However in preliminary attempts time points around the hour mark seemed to create the best shaped curves (Appendix A).

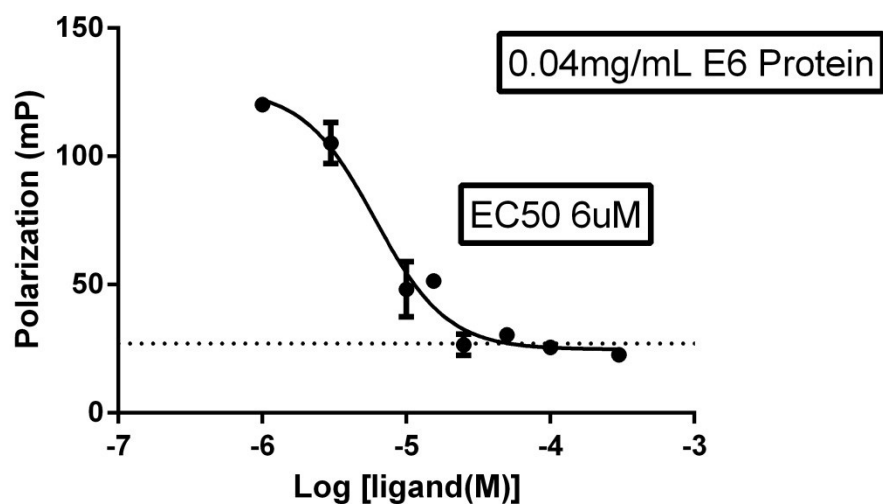


Figure 34: FP curve of E6 (0.04mg/ml) incubated with varying the concentration of 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) (1 μ M-300 μ M). This experiment was performed in a Nunclon® 96 well black bottom plate at an incubation time of 75 minutes, at 37°C. The reaction was buffered to a pH of 7.5 (20mM HEPES 7.5, 0.01% tween). The black dotted line represents the background polarization of AFTD without protein. EC50 value was estimated at 6 μ M, which indicates sufficient affinity for a molecular probe.

The estimated EC50 value from this curve is 6 μ M with a range of 3.6 μ M to 10 μ M, as determined using Prism 6 (Graph Pad, inc)⁵ non-linear regression dose-response four parameters function. This range is fairly small and not unreasonable based on previous results (Appendix A). The EC50 value is adequate for probe development as molecular probes should have binding constants that indicate high affinity. Ideally this value would fall in the nM range; however, μ M range is deemed sufficient affinity for a molecular probe⁶. The EC50 value is also aligned with the range of affinity generally associated with virtual ligand screening. This experiment validates the computational screening, confirming AFTD as a binder to E6 with an EC50 value in the micromolar range (6 μ M).

5.5.CELL BASED EXPERIMENTS

Cell permeability and toxicity are important factors in the assessment of AFTD as potential in-vivo molecular probe for diagnostic imaging. The cell assays were designed to give information regarding specificity of AFTD to E6. The cell types chosen for these assays were CaSki, SiHa, and C33A. CaSki and SiHa cell express E6, whereas C33A is a cervical cancer cell line that does not express E6.

5.5.1.Cell viability assay

The goal of this experiment is to determine if AFTD is toxic to cells at the concentrations proposed (25 μ M and 10 μ M), and whether toxicity is E6-dependent. The first cell-based experiment used an Axiovert microscope at the Thunder Bay Regional Health Sciences Center (TBRHSC) following exposure to two different concentrations of 3-Amino-5-Fluorobenzo 1,4-Triazine Dioxide (10 μ M and 25 μ M). Each cell type was incubated for 1, 3, 6, 12 and 24 hours then imaged using the microscope. Very little change was noticed after one hour and three hours. For this reason the cells were not viewed again until the 24 hour time point. It was discovered that all the CaSki and SiHa cells were dead in the 25 μ M concentration wells, and most of the cells were dead in the 10 μ M wells (Figure 35). To determine whether toxicity was caused by our compound and not the ethanol used in our stock solutions, the three cell types were exposed to the same concentration by volume of ethanol and incubated for 24hrs as a control in both cell lines (Figure 35). Upon observing the cells it was determined that it was not the ethanol killing the cells but our fluorescent compound AFTD (Figure 36).

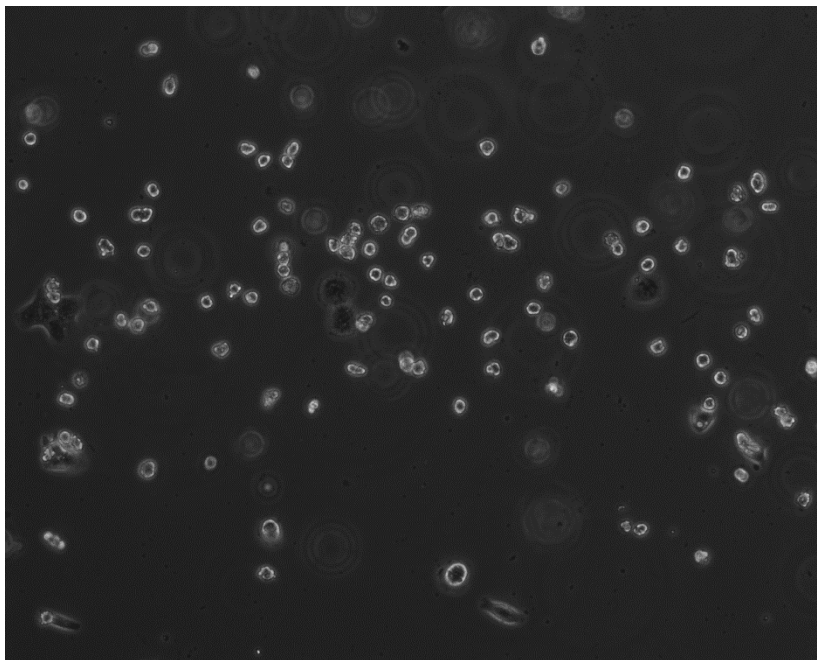


Figure 35: CaSki cells after incubation with AFTD (25uM) for 24hrs. Most of the cells have died at this time point: the cells are floating and appear to have a white ring.

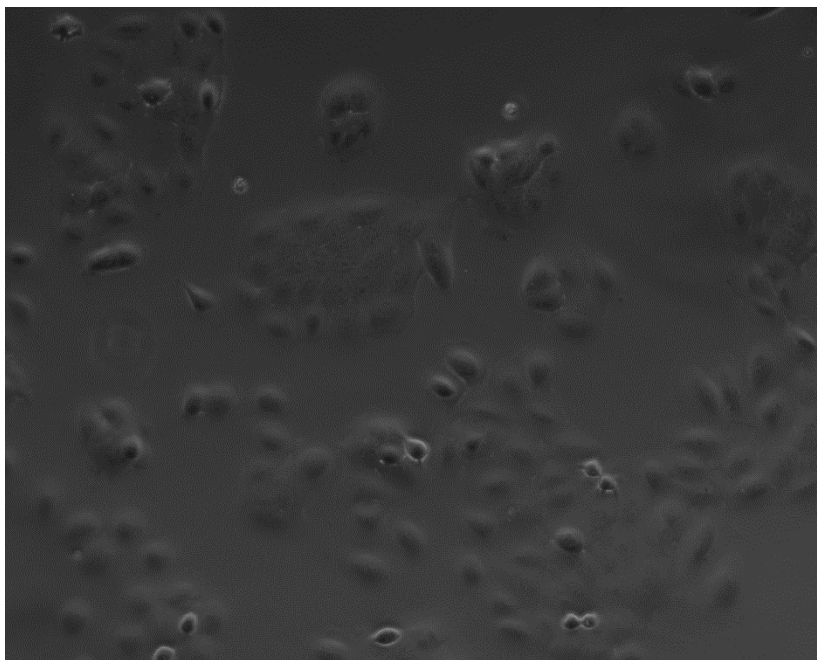


Figure 36: CaSki cells incubated with ethanol (0.01%) for 24hrs. In this solvent control, the cells are indeed healthy as they look normal, indicating that it is AFTD and not ethanol that is causing cell death.

Since both these cells types contained copies of E6, it was possible that AFTD is killing the cells in an E6 specific manner. To further test this hypothesis, an additional cell line C33A was tested. C33A cells were used in order to test if the cell death induced by AFTD is E6-dependent. C33A is a cervical cancer cell line, however contains zero copies of E6, making it an ideal choice for a negative control. Wells were incubated with AFTD (25 μ M and 10 μ M) for 24 hours. Ethanol control wells were included in this experiment. Upon examination under a light microscope at the 24 hour time point, it was determined that our ligand is generally cytotoxic at the concentrations tested (Figure 37). This was further concluded using the ethanol control (Figure 38)

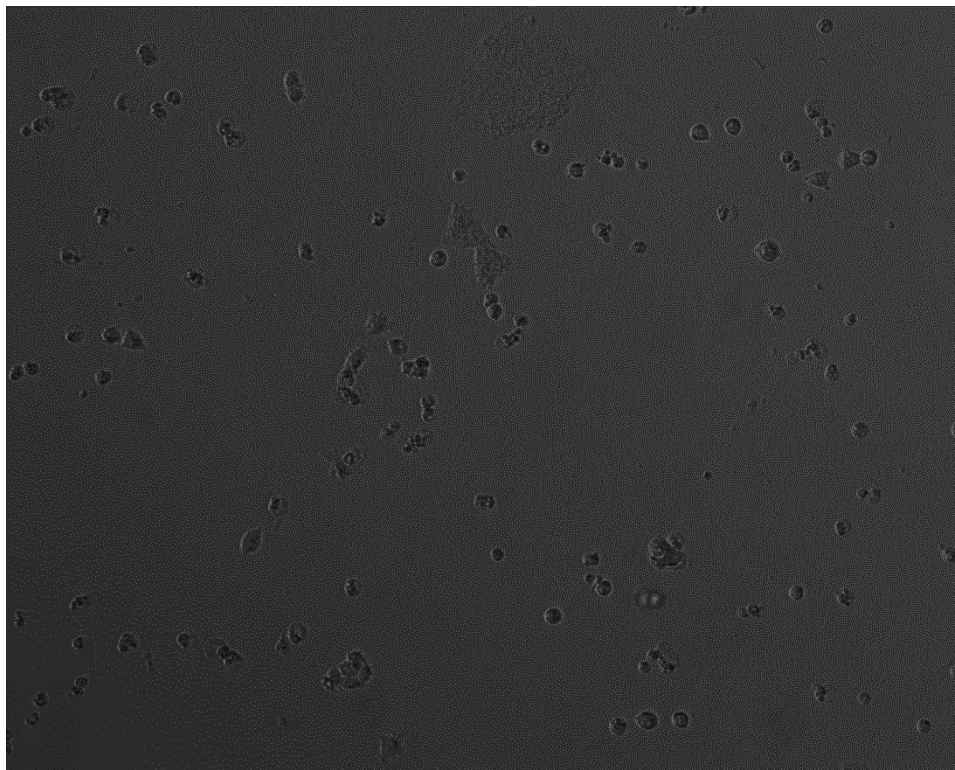


Figure 37: C33A cells incubated with ethanol (0.01%). In this solvent control, the cells are alive (not floating), indicating that it is AFTD and not ethanol that is causing cell death.

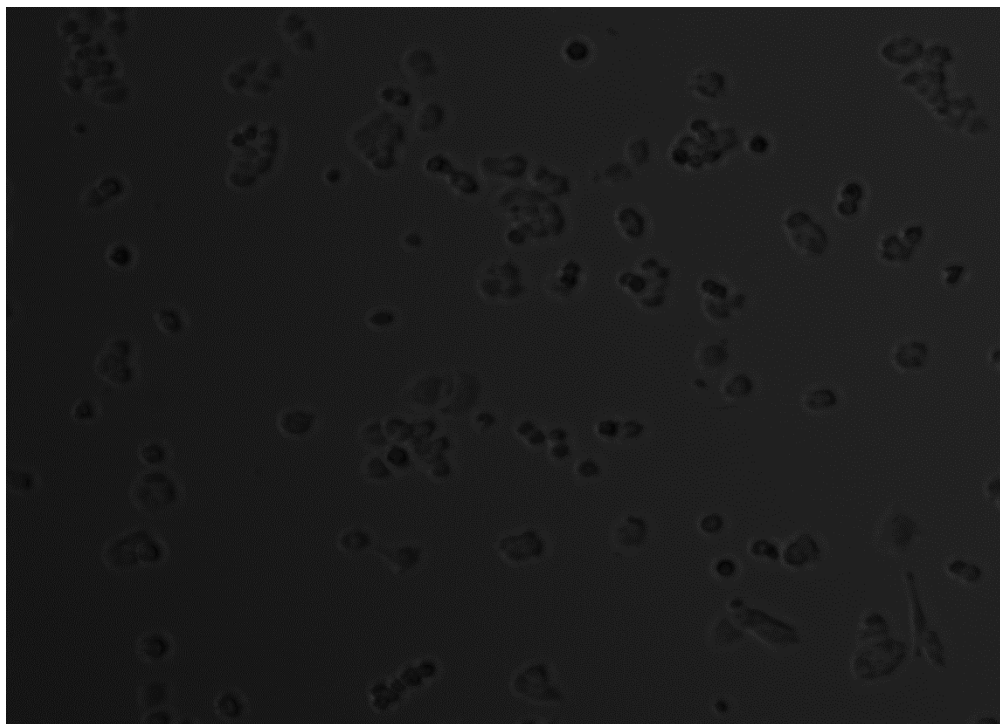


Figure 38: C33A cell lines after incubation with AFTD (25uM) for 24hrs. Most of the cells are dead as they are still attached to the plate.

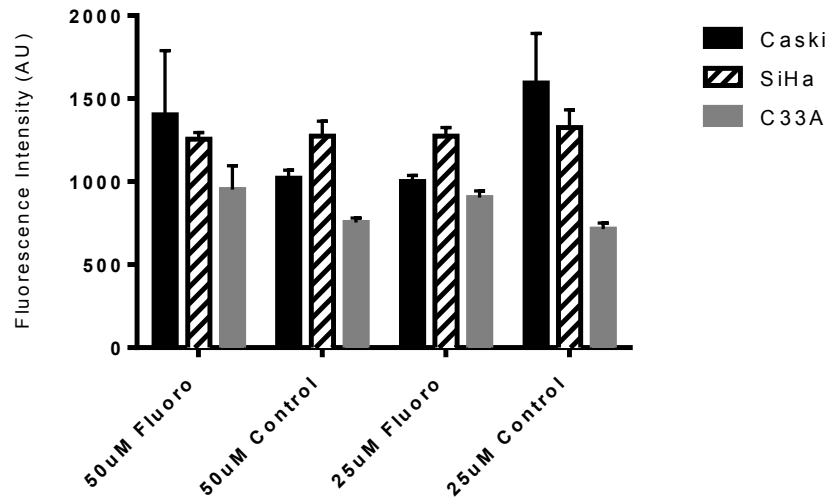
Although cytotoxic at 10 μ M, AFTD would still be useful as imaging probe in imaging modalities such as Position Emission Tomography (PET), where very small amounts, generally in the nM range, are used⁷. AFTD can be labelled with ¹⁸F and it is, hence, suitable as PET probe. However, further cytotoxicity studies need to be conducted to investigate if AFTD is safe *in vivo* at lower concentrations. Regardless of toxicology, AFTD may be useful as a biochemical probe for fluorescence-based assays involving E6.

5.5.2. AFTD Cell Permeability Experiment

In order to determine if AFTD is cell permeable, a fluorescence ligand uptake experiment was designed utilizing AFTD's intrinsic fluorescence. We tested AFTD against the three cell types (CaSki, SiHa and C33A) in order to determine if AFTD is cell permeable, and if uptake is E6-specific uptake should be seen most in CaSki cells as they contain the most copies of the E6

genome, with SiHa and C33A following respectively. All three cell types CaSki, SiHa and C33A cells were incubated with two concentrations of AFTD (25 μ M and 50 μ M). The time points selected were 1, 2, and 3 hours (Figure 39). We also controlled for cell density by reading absorbance at 260, 280 and 535nm using the microplate reader. The one hour time point was not analyzed as it was inconsistent with the other time points. This is because a lot of air bubbles were present in the plates and our technique had to be adjusted for the consecutive time points to make sure more accurate reads could be made.

A)



B)

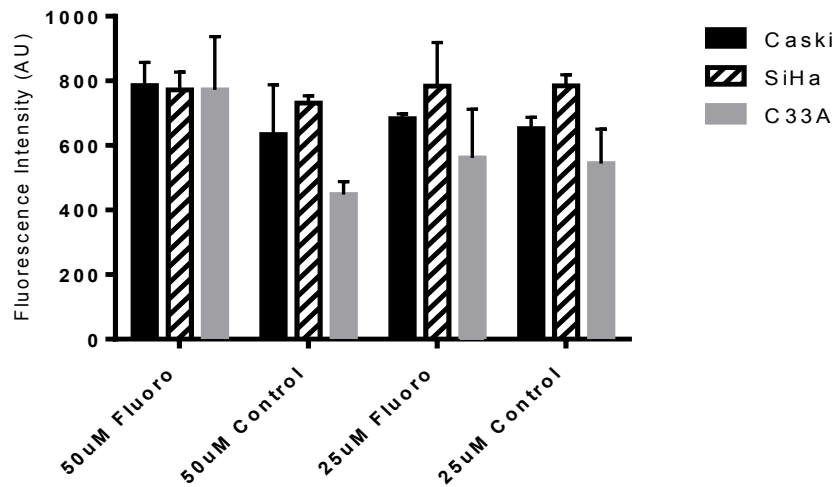


Figure 39: Fluorescence intensity (FI) at 2 hour (A) and 3 hour (B) incubation. Intensities collected from the second cell permeability test. Cells were incubated AFTD at 25µM and 50µM. Experiment was conducted in a BioTek Synergy 4 Microplate reader at an incubation temperature of 37°C. AFTD-associated FI showed significant difference ($P < 0.05$) relative to control at 50µM in C33A cells, indicating cell permeability. These controls are cells under the same environment as those wells that contain cells that have been exposed to AFTD just with no AFTD. Fluorescence was measured using a fluorescence filter system (Ex 460/40nm, Em 528/20nm, and 510nm dichromatic mirror).

Multiple comparison ANOVAs were ran for each cell type, time point and concentration. The C33A 50 μ M sample was shown to be statistically higher in fluorescence than its corresponding control (Figure 39). This allows us to conclude that our molecule did get inside this cell type. This cell type does not contain any copies of E6, however further testing can be done outside the scope of the project to look at specificity of binding to E6. The fact that AFTD was not present in the CaSki cells does not mean it does not accumulate in this cell type. CaSki cells grow in tight clusters making it difficult for compounds to get inside the cells⁸. Another potential reason for the negative result could be drug-resistant pumps with affinity for AFTD, such as MDR-1, which has been detected in SiHa cells⁹. Drug resistance has been seen in CaSki cells although by estrogenic hormones¹⁰. This hypothesis would need to be further investigated using an alternative method, such as a microscope capable of inducing fluorescence. Cells would be incubated with different concentration of AFTD (50 μ M and 25 μ M) and multiple time points would be viewed using the microscopes where pictures would be taken. Pictures would have to be taken at time point shorter than one hour to be able to observe drug resistance pumps at play. Further experiments are thus needed to determine if AFTD accumulates in the CaSki cell lines versus SiHa or C33A.

5.6.6F4 ANTIBODY COMPETITION ASSAY

Preliminary fluorescence polarization assays using AFTD and the 6F4 antibody indicated that these molecules bind to different regions within E6. This was evidenced as an increase in polarization observed when the antibody 6F4, AFTD, and E6 were all in the same reaction well, as shown in appendix A. In order to further demonstrate this cooperative binding event, we tested various concentrations of 6F4 antibody against fixed concentrations of both AFTD (6 μ M) and E6 (0.04mg/ml). These correspond to the EC50 concentration conditions as determined in a

separate FP assay (Section 5.4). Seven concentrations of 6F4 antibody were tested (4ng/ml to 0.000004ng/ml). These results are shown in Figure 40. The system appears to be equilibrated after 90 minutes. For this reason, the 105min time point was used for further analysis.

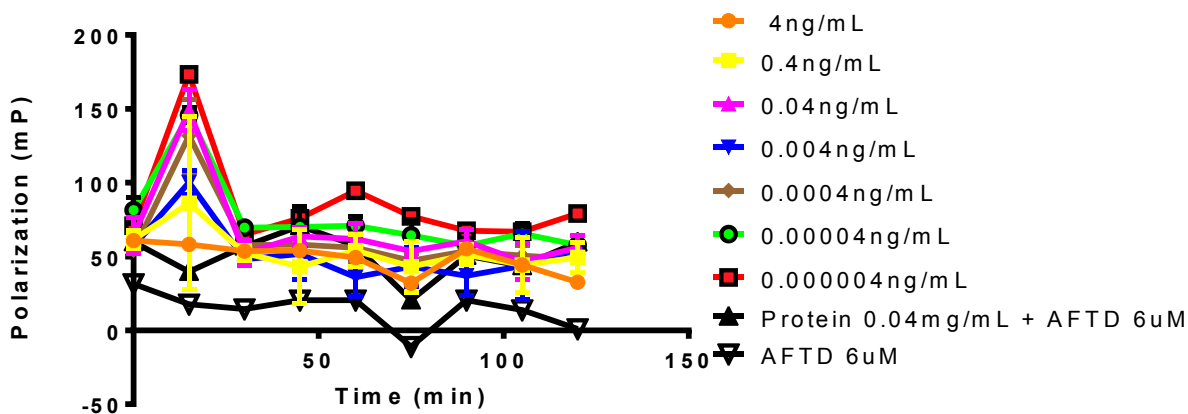


Figure 40: Time curve showing the change in polarization of the protein and ligand complex and the antibody, protein ligand complex. This experiment was run using a fixed concentration of ligand 6 μ M and 0.04mg/ml of protein. This experiment was performed in a Nunclon® 96 well black bottom plate at a temperature of 37°C and buffered at a pH of 7 (20mM HEPES 7.3, 0.01% tween). Our system seems to be equilibrated after 90 minutes. However no real clear pattern of change in polarization is observed within the concentrations used of 6F4.

A bar graph was created for the 105 minute time point to compare the FP values for all concentration of antibody tested (Figure 41). This bar graph shows the higher concentrations of antibody (6F4) do not produce a significant change in polarization when comparing them to the positive control of protein E6 and AFTD. The last two concentrations appear to show a significant result. One potential explanation of this couple could be that AFTD and antibody (6F4) are binding at the higher concentration, which was further investigated.

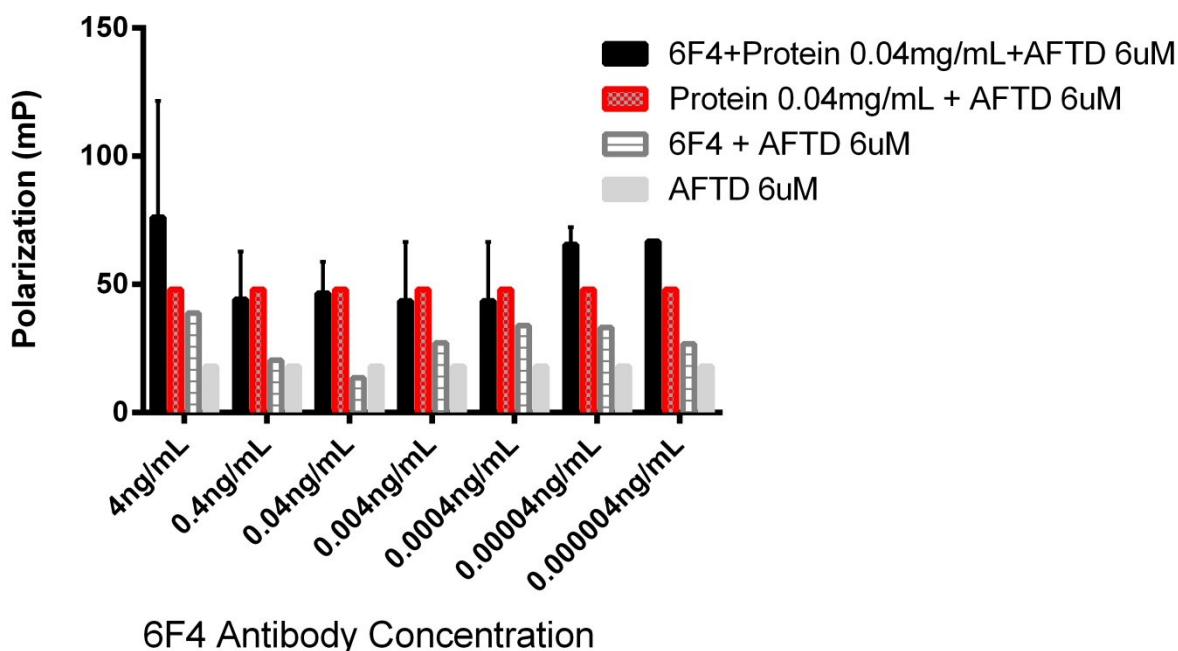
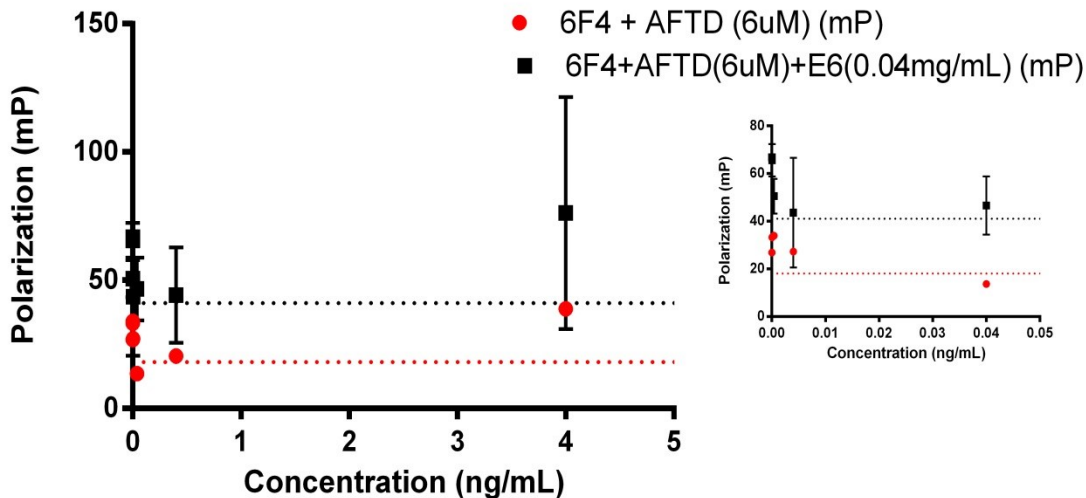


Figure 41: Fluorescence polarization varying the antibody concentration. Seven concentrations of 6F4 antibody (4ng/ml-0.000004ng/ml) were used with a constant concentration of AFTD (6 μ M) and E6 protein (0.04mg/ml). This reaction was conducted at an incubation temperature of 37°C and at a pH of 7.5 (20mM 7.5 HEPES 0.01% tween (v/v)). An increase in polarization was observed for samples containing 6F4 in presence of AFTD and E6 relative to samples containing only AFTD and E6, indicating that 6F4 binds to the AFTD-E6 complex. Some binding between the antibody 6F4 and AFTD is also observed.

The fluorescence polarization calculated for AFTD and 6F4 shows some binding interaction but with an oscillating behaviour relative to antibody concentration, within the range tested (Figure 42). Additional assays are necessary to better understand this behaviour, but antibody aggregation may play a role (Figure 42).

A)



B)

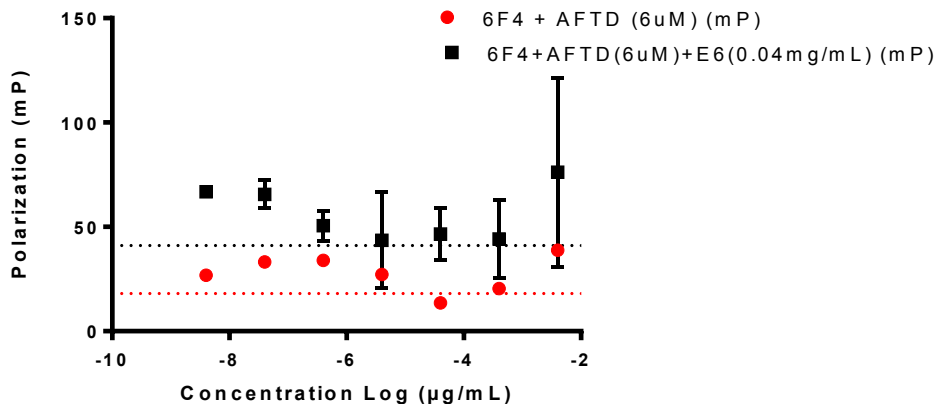


Figure 42: Fluorescence polarization versus log of the concentration ($\mu\text{g/ml}$) of 6F4 antibody. A) Concentration scale and B) log scale. Plate was incubated for 105 minutes at a temperature of 37°C , and pH was buffered to 7.5 (20mM 7.5 HEPES, 0.01% tween (v/v)). Concentrations of 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) at $6\mu\text{M}$ and E6 protein at 0.04mg/mL were held constant while varying the concentration of antibody 6F4. The black dotted line represents the polarization of AFTD and E6 and the red dotted line represents the background polarization from AFTD on its own. It can be concluded that 6F4 and AFTD are not interacting more at the higher concentrations than at the lower concentrations, within the range tested.

In order to overcome the inconsistencies shown with the antibody, an additional experiment should be performed, where E6 interaction with E6AP, a protein known to bind to E6, will be tracked using AFTD. This will give further validation of use of AFTD for the purpose of a tool for such biochemical assays.

Although additional assays need to be performed, the competition assay results indicate that AFTD may be useful as a probe for biochemical assays involving E6 interactions with other proteins. E6 is known to interact with many proteins within the cell. The best known interaction being with E6AP, which can lead to ubiquitination of p53 if high risk variants are present¹¹. E6 is also shown to interact with other protein such as MAGI-1¹², MAGI-2¹³, MAGI3¹³, hScribble¹⁴, hDIg¹³ and p300/CBP complex¹⁵. With these interactions in mind, developing a tool to further track protein-protein interaction of E6 would be very useful. Current tools used are GSP-pull down assays, which are accurate, however are time consuming and only show interaction inside the cell environment¹⁶. It is important to look at these interactions outside the cell as many factors within the cell can impact the interaction of these proteins. Testing E6 protein binding interactions outside of the cell would enable a better understanding of these interactions. A small-molecule molecular probe to allow for protein-protein interaction to be tracked outside of the cell environment would be the first of its kind for E6.

5.7.CHARACTERIZATION OF O-SUCCINY-L-HOMOSERINE-BODIPY

O-succinyl-L-homoserine was conjugated with Bodipy dye (3-Bodipy-proanoylaminocaproic Acid, N-Hydroxysuccinimide) in order to further validate O-succinyl-L-homoserine's potential use as a molecular imaging probe. Following successful conjugation of O-succinyl-L-homoserine, the conjugate was characterized in order to determine what filters and mirror were optimal for fluorescence experiments. It was determined, based on emission and

excitation analysis, that the conjugate compound has well-defined maximum excitation (500nm) and maximum emission (510nm) (Figure 43). Given these values, an excitation filter (485/20nm) and emission filter (528/20nm), and a dichromatic mirror (510nm) were selected.

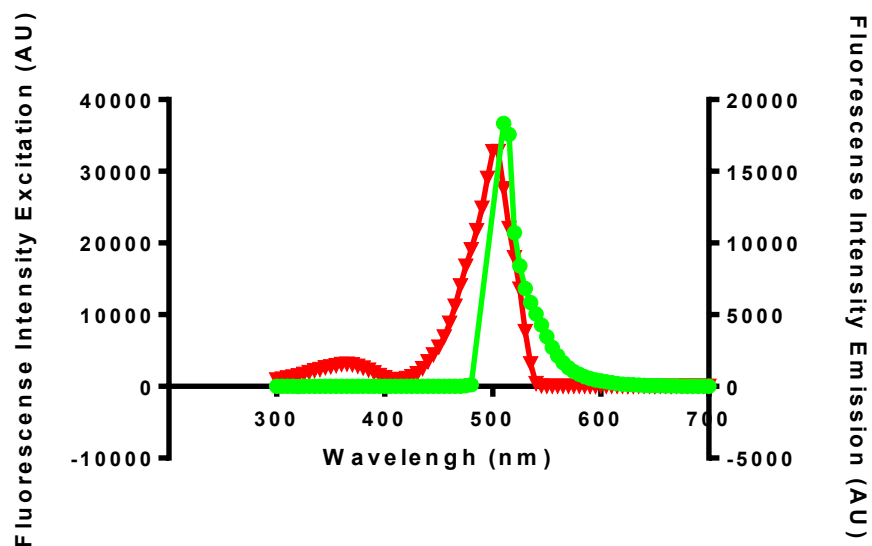


Figure 43: Emission (green) and excitation (red) plot of O-succinyl-L-homoserine-BODIPY. This plot shows a maximum excitation at 500nm and emission at 510nm. Using these values we were able to develop a proper optical system (Ex filter: 485/20nm, Em filter: 528/20nm, and dichromatic mirror 510nm)

Prior to conducting this experiment a standard curve was run using all the concentration proposed for the actual experiment to make sure that our concentrations demonstrated linearity (Figure 44). It was determined that this filter set up was adequate as an R^2 value of 0.90 was determined indicated linearity (Figure 44).

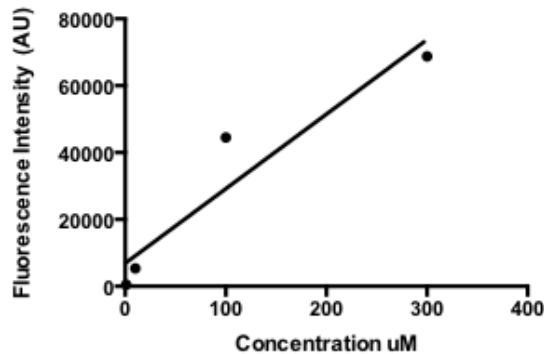


Figure 44: Linear fluorescence intensity plot of O-succinyl-L-homoserine-BODIPY (300µM, 100µM, 10µM, and 1µM) and protein E6 (0.04mg/ml). Plate was incubated at a temperature of 37°C and pH was buffered to 7.5 (20mM HEPES, 0.01% tween (v/v)). R^2 was determined using a linear function to be 0.90.

5.8. FP FOR O-SUCCINYL-L-HOMOSERINE CONJUGATE

Fluorescence polarization was conducted using three concentration of O-succinyl-L-homoserine-BODIPY (100µM, 10µM, and 1µM) and 0.04mg/ml of E6 (Figure 45).

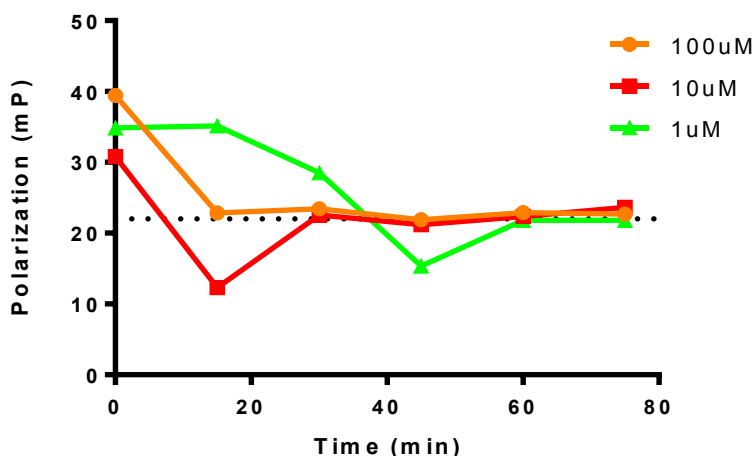


Figure 45: Fluorescence polarization time plot of O-succinyl-L-homoserine-BODIPY (100µM, 10µM, and 1µM) and protein E6 (0.04mg/ml). Plate was incubated at a temperature of 37°C and pH was buffered to 7.5 (20mM HEPES, 0.01% tween (v/v)). Protein concentration was held constant at 0.04mg/ml. The dotted line represents the baseline polarization of O-succinyl-L-homoserine-BODIPY. Equilibrium is reached around 1hour, but FP values are at baseline by then. Our results indicate that the O-succinyl-L-homoserine-BODIPY is not binding at any concentrations tested.

Based on the time plots (Figure 45), it is evident that the system equilibrated after 1 hour. Upon equilibrium there is no difference in polarization between free O-succinyl-L-homoserine-BODIPY and O-succinyl-L-homoserine-BODIPY added to E6 protein at any of the concentrations tested.

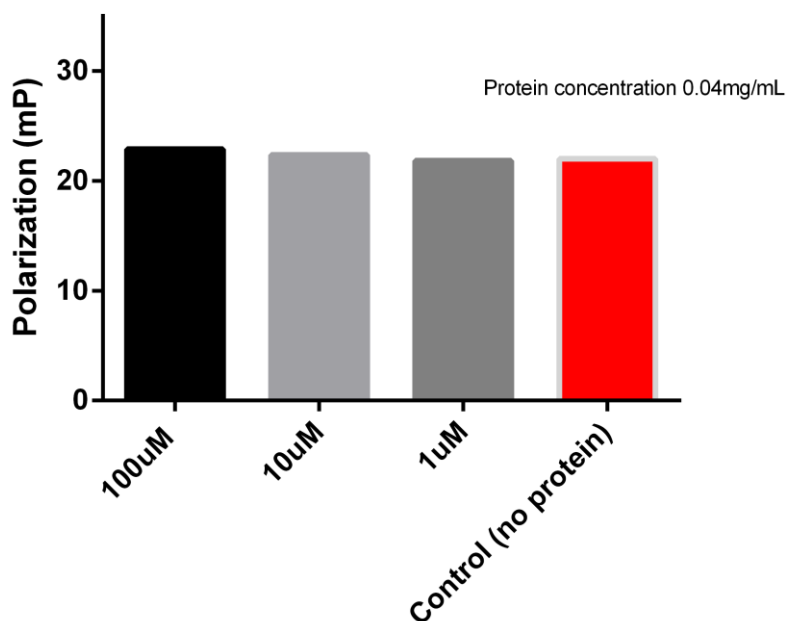


Figure 46: Bar graph of polarization (mP) at 60 minutes for the O-succinyl-BODIPY binding experiment. Plate was incubated at a temperature of 37°C and pH was buffered to 7.5 (20mM HEPES, 0.01% tween (v/v)). E6 protein concentration was held constant at 0.04mg/ml. All three concentration were plotted (100µM, 10µM and 1µM) with the control in red, which represents the polarization (mP) of O-succinyl-L-homoserine-BODIPY on its own. This bar graph demonstrates that O-succinyl-L-homoserine-BODIPY does not bind to E6 at the concentrations tested.

Based on these results it is not believed that O-succinyl-L-homoserine-BODIPY binds to E6. However, results for unlabeled O-succinyl-L-homoserine are still potentially promising, as quenching above 30% was observed in tryptophan fluorescence experiments.

5.9. REFERENCES

1. Brady, G. P., Jr.; Stouten, P. W., Fast prediction and visualization of protein binding pockets with PASS. *J Comput Aided Mol Des* **2000**, *14* (4), 383-401.
2. Floriano, W. B.; Vaidehi, N.; Zamanakos, G.; Goddard, W. A., HierVLS Hierarchical Docking Protocol for Virtual Ligand Screening of Large-Molecule Databases. *Journal of Medicinal Chemistry* **2003**, *47* (1), 56-71.
3. PubChem, PubChem3D.
4. Group, C. C. *MOE, version 2010.10* Montreal Canada 2010
5. *Prism, 6 UPDATE for Windows*, GraphPad Software: La Jolla, California USA, 2013.
6. Kelly, K. A.; Fred Reynolds; Kristof, K. R., High-Throughput Screening for Probe Development. In *Molecular Imaging in Oncology*, pp 179-188.
7. Wells, J. A.; McClendon, C. L., Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007**, *450* (7172), 1001-1009.
8. Togtema, M.; Pichardo, S.; Jackson, R.; Lambert, P. F.; Curiel, L.; Zehbe, I., Sonoporation Delivery of Monoclonal Antibodies against Human Papillomavirus 16 E6 Restores p53 Expression in Transformed Cervical Keratinocytes. *PLoS ONE* **2012**, *7* (11), e50730.
9. Laochariyakul, P.; Ponglikitmongkol, M.; Mankhetkorn, S., Functional study of intracellular P-gp- and MRP1-mediated pumping of free cytosolic pirarubicin into acidic organelles in intrinsic resistant SiHa cells. *Canadian Journal of Physiology and Pharmacology* **2003**, *81* (8), 790-799.
10. Laochariyakul, P.; Ponglikitmongkol, M.; Mankhetkorn, S., Functional study of intracellular P-gp- and MRP1-mediated pumping of free cytosolic pirarubicin into acidic organelles in intrinsic resistant SiHa cells. *Can J Physiol Pharmacol* **2003**, *81* (8), 790-9.
11. Nakagawa, S.; Huibregtse, J. M., Human Scribble (Vartul) Is Targeted for Ubiquitin-Mediated Degradation by the High-Risk Papillomavirus E6 Proteins and the E6AP Ubiquitin-Protein Ligase. *Molecular and Cellular Biology* **2000**, *20* (21), 8244-8253.

12. Kranjec, C.; Banks, L., A Systematic Analysis of Human Papillomavirus (HPV) E6 PDZ Substrates Identifies MAGI-1 as a Major Target of HPV Type 16 (HPV-16) and HPV-18 Whose Loss Accompanies Disruption of Tight Junctions. *Journal of Virology* **2011**, *85* (4), 1757-1764.
13. Grm, H. S.; Banks, L., Degradation of hDlg and MAGIs by human papillomavirus E6 is E6-AP-independent. *Journal of General Virology* **2004**, *85* (10), 2815-2819.
14. Massimi, P.; Gammoh, N.; Thomas, M.; Banks, L., HPV E6 specifically targets different cellular pools of its PDZ domain-containing tumour suppressor substrates for proteasome-mediated degradation. *Oncogene* **2004**, *23* (49), 8033-8039.
15. Zimmermann, H.; Degenkolbe, R.; Bernard, H.-U.; O'Connor, M. J., The Human Papillomavirus Type 16 E6 Oncoprotein Can Down-Regulate p53 Activity by Targeting the Transcriptional Coactivator CBP/p300. *Journal of Virology* **1999**, *73* (8), 6209-6219.
16. Detection of protein-protein interactions using the GST fusion protein pull-down technique. *Nat Meth* **2004**, *1* (3), 275-276.

Chapter 6. Conclusion and Future Work

Much work has gone into further understanding HPV as it remains a large concern in today's health care system, due to its association with various cancers, including cervical. HPV remains a complex virus with many known genotypes. However, the complexity does not end there as even high risk types do not necessarily express variants of E6 responsible for the cell immortalization leading to increased susceptibility to cancer development. Finding a way to detect protein E6, specifically those variants associated with invasive cancer, has the potential to aid in early intervention, thus improving the life of those impacted. Finding a molecule that can be used as a molecular probe for biochemical assays for E6 protein-protein interactions (PPIs) would also be beneficial as E6 is known to bind to many proteins within the cell. To date no method for probing PPIs outside of the cell environment has been developed. This would allow for better understanding of these interactions, and may lead to new therapeutic and diagnostic strategies.

To date there is not a small organic molecule known to bind specifically to E6. This left a great starting point in the pursuit to find such a compound. Through using a combination of computational methods and experimental methods our hope was to find a small organic molecule that could bind to wild-type HPV variant 16 protein E6 specifically. Such small organic molecule could be a very useful tool in the further understanding of HPV infection.

3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD), showed great promise and for this reason was tested extensively, both computationally and experimentally. Due to its intrinsic fluorescence, and fluorine atom this molecule could be used for optical imaging or Positron Emission Tomography (PET), respectively, if shown to be specific and non-toxic.

Both tryptophan and fluorescence polarization assays for AFTD against E6 were conclusive of a concentration-dependent binding. An EC50 value of 6 μ M was determined for AFTD and this is deemed sufficient for molecular probe development. This not only demonstrates affinity but also indicates that binding is indeed specific. Cell based assays indicated that AFTD is cytotoxic at 25 μ M; however further cytotoxicity experiments would need to be conducted to test at what concentration AFTD could potentially be used safely *in-vivo*. Cell based experiments also demonstrated evidence of cell permeability; however this was shown significantly in a cell line that contains no copies of E6 (C33A). This event was believed to be the result of drug resistance pumps shown and documented in SiHa cells, and morphological characteristics documented and shown in CaSki cells. Microscopic images along with shorter time points will allow us to validate this hypothesis. High Intensity Focused Ultrasound (HIFU) methods could be employed in order to overcome the tightly packed clusters that limit permeability, shown and documented in CaSki cells.

AFTD also appears suitable for usage as a molecular probe for a biochemical protein-protein interaction assay. AFTD was able generate results that indicate protein-protein interactions between E6 and HPV16 E6 specific antibody (6F4). However, due to inconsistencies seen with this antibody other proteins known to bind to E6, will be tested in the future such as E6 Associated Protein (E6AP).

In addition to AFTD, O-succinyl-L-homoserine, and paclitaxel also demonstrated binding to E6 when using intrinsic tryptophan fluorescence and could be further investigated. O-succinyl-L-homoserine conjugated to a bodipy dye failed to bind to E6, which is disappointing. Conjugation with other dyes could be explored, both computationally and experimentally in an attempt to find a conjugate that retains binding affinity to E6. Paclitaxel showed an excellent

dose-dependence at the concentrations tested using intrinsic tryptophan fluorescence. This is a particularly interesting result as Paclitaxel is a commonly used chemotherapy drug. However, in order to fully characterize Paclitaxel, additional experimental testing would be required in order to gain an accurate EC50. Due to availability of radio-labelled Paclitaxel this could be easily done using a scintillation proximity assay. Paclitaxel would be an interesting probe for E6 due to its current use and availability within a clinical setting. Paclitaxel it is already a characterized drug approved for clinical use.

Computational screening conducted using the structure of the N-terminal domain of E6 suggests additional compounds that could be characterized in the future as potential molecular probes. Also with the publishing of a full structure of HPV16 E6 (PDB 3GIZ) further computational studies can be conducted with the already created databases. This will allow us to further validate our computational work with the two separate domains.

Through the utilization of a variety of computational and experimental techniques we were able to find and begin the characterization of AFTD. Affinity was demonstrated both computationally and experimentally. Cell permeability was also demonstrated in C33A, however due to potential drug resistance shown in CaSki and SiHa cells further work will need to be conducted in order to determine specificity for E6. AFTD is also believed to be a suitable probe for biochemical assays for E6 protein-protein interactions. Other molecules such as O-succinyl-L-homoserine and paclitaxel also showed great promise for use as a molecular probe. However, AFTD due to advantageous characteristics was the major focus of this study. We believe that AFTD as a dual propose molecular probe could be a very powerful tool in furthering our understanding of HPV's link to cancer. The same approach that identified AFTD may lead to the

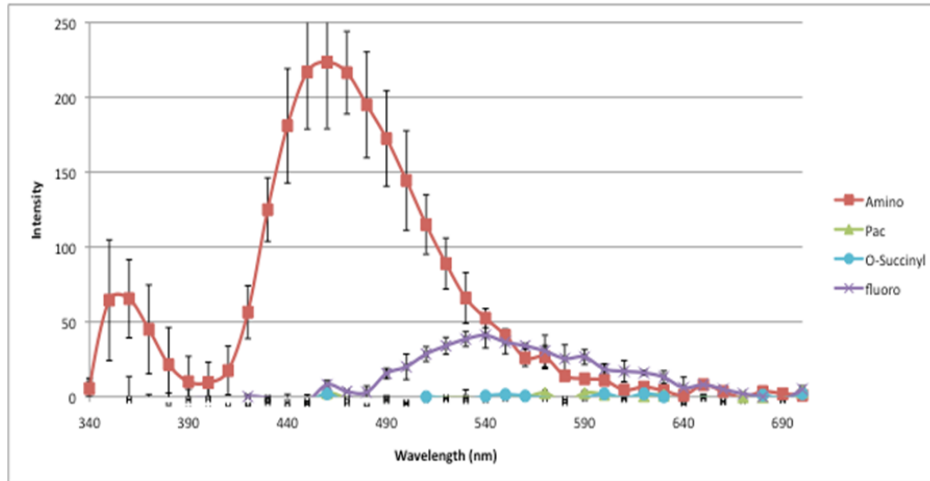
identification of molecular probes for high-risk variants of protein E6 associated with cancer development.

Chapter 7. Appendix A: Calibration and Preliminary Results

7.1. OBTAINING FLUORESCENCE AND ABSORBANCE SPECTRUMS FOR LIGANDS

Prior to running ligand binding assays absorbance and emission spectrums were run for each ligand. Emission spectrums were run at an excitation of 288nm in order to make sure ligands that could cause quenching in tryptophan fluorescence were identified. Absorbance spectra for all four ligands were obtained using a Take3 microplate. Fluorescence emission spectrum of all for ligands was obtained using a Nunclon® 96 well- microplate. Both plates were run on a Synergy 4 BioTek microplate reader at an incubation temperature of 37°C. All ligands were kept at a constant concentration (10uM) (Figure A).

A)



B)

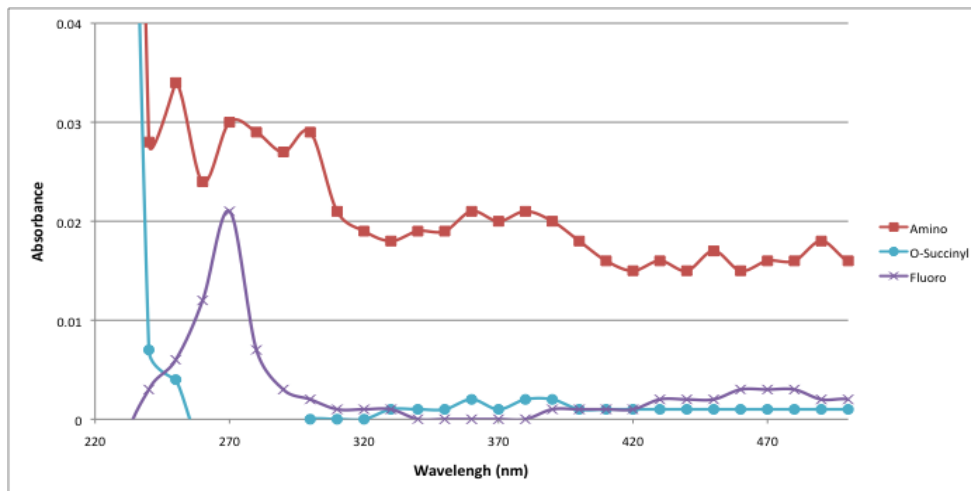


Figure A: Shows the emission (A) absorbance (B) and for the four ligands tested ($10\mu\text{M}$). These graphs were taken using a Nunclon® 96 well plate incubated at 37°C at a constant pH using a HEPES 7.0 buffer system (20mM Hepes pH 7.3). These ligands were excited at 288nm (B) and the blank containing the buffer and water. Each well was topped up to a total volume of $100\mu\text{L}$. Each ligand was at a concentration of $10\mu\text{M}$. From these graphs we can see that Aminopterin is the only ligand that may cause quenching in tryptophan fluorescence assays. In addition a maximum absorbance was identified at 270nm with a secondary max at 470nm for AFTD with a max emission at 535nm. The other ligands tested were deemed non-fluorescently active.

It was determined that the only ligands that were fluorescently active were Aminopterin and AFTD. AFTD was determined to have a max absorbance at 270nm and a max emission at 535nm. Since this compound was to be used in fluorescence polarization assays a proper filter set would need to be established. Due to availability of filters the secondary absorbance maximum was used (470nm) (Figure A). From this a filter set containing an excitation filter (485/20nm), emission filter (528/20nm), and dichromatic mirror (510nm) was implemented. Paclitaxel appeared to come out of solution and for this reason was not included in the absorbance spectrum .

7.2.PRELIMINARY RESULTS

7.2.1.6F4 antibody preliminary test to select concentration

Lots of work has been conducted with an antibody specific to E6 (6F4), as it has taken part in both tryptophan fluorescence, and fluorescence polarization experiments.

Preliminary experiments were conducted to determine the proper concentration of antibody to be used in tryptophan fluorescence assays. Four 100 fold dilutions of the antibody from a 4 μ g/mL stock solution were prepared. Change in intrinsic tryptophan fluorescence induced on E6 in presence of the antibody was measured. Results indicated that a change of at least 30% was possible with each of the concentrations used (Figure B). However the 0.04ng/mL concentration showed the highest percentage change (80%) (Figure B).

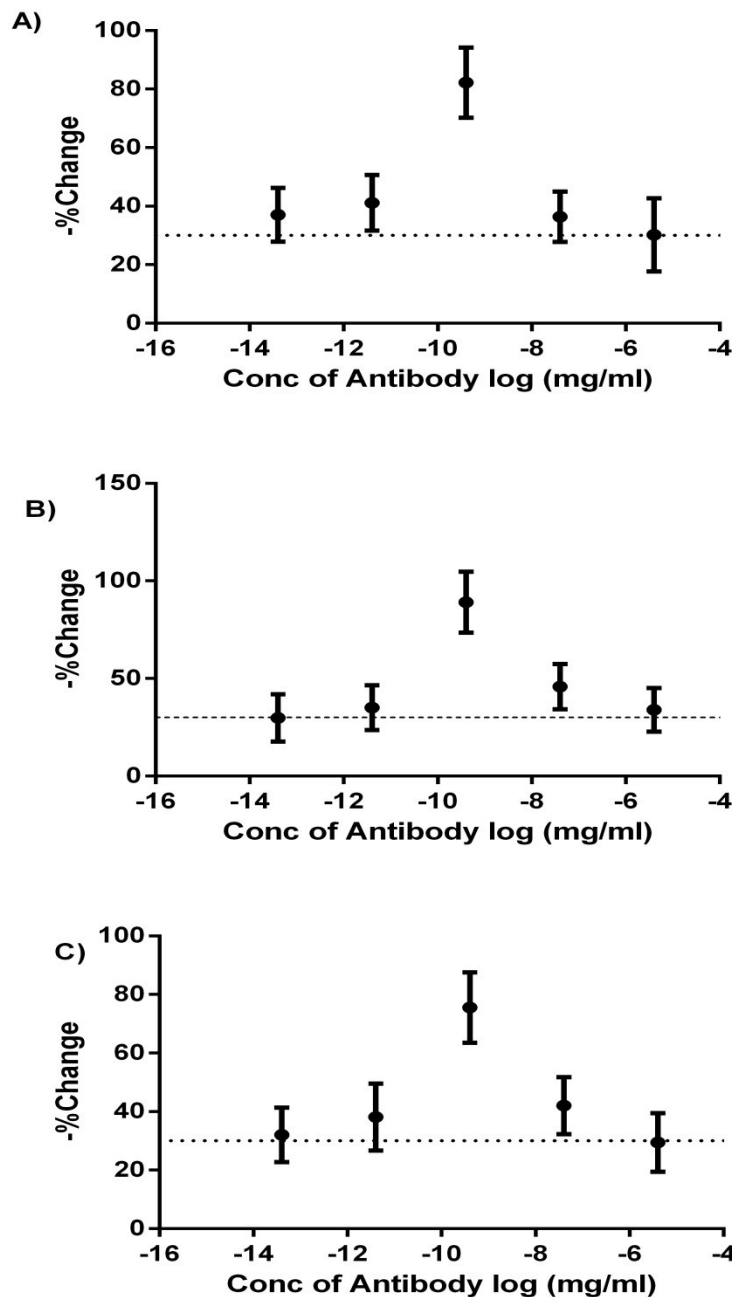


Figure B: %change curves showing the change in intrinsic tryptophan fluorescence of the protein and ligand complex and the antibody, protein ligand complex. A) 5 minutes B) 30 minutes and C) 1 hour. This experiment was run for 60 minutes using $4\mu\text{g/mL}$, 4ng/mL , 0.4ng/mL , 4pg/mL , 0.04pg/mL of antibody, $10\mu\text{M}$ of ligand, and $0.24\mu\text{M}$ of protein. This experiment was performed in a Nunclon® 96 well black bottom plate at a temperature of 37C and buffered at a pH of 7 (20mM HEPES 7.3, 0.01% tween). From this we can see that all the concentrations of antibody produce about the same change in tryptophan fluorescence. From this we selected a concentration of 0.4ng/mL to be the optimal concentration.

In order to assess whether or not the antibody contains tryptophan curves of each concentration were examined with an excitation of 288nm (Figure C). As shown in figure C, the 6F4 antibody contains tryptophan and proper controls for its contribution to the measurements need to be used in other tryptophan fluorescence assays for each concentration tested.

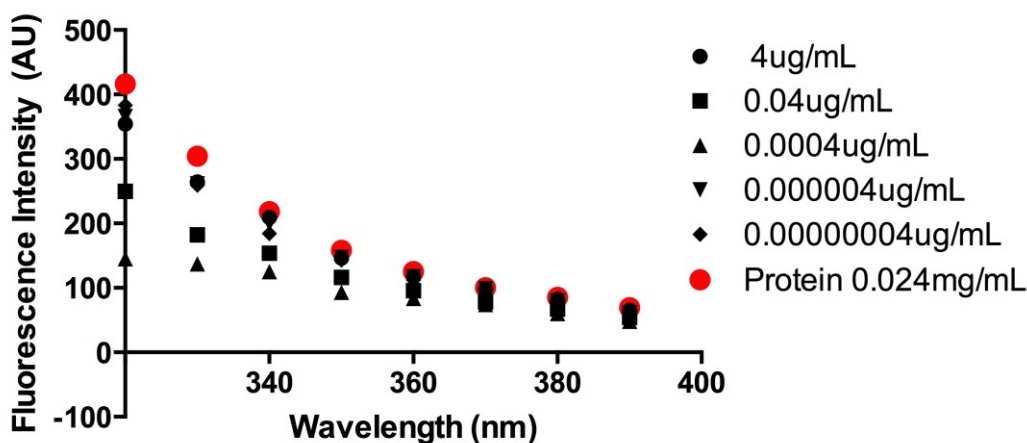


Figure C: Fluorescence intensity curve of the antibody (6F4) at various concentrations on its own. This experiment was run and data was taken at 60 minutes using 4 μ g/mL, 0.04 μ g/mL, 0.0004ng/mL, 0.000004 μ g/mL, 0.00000004 μ g/mL of antibody, 10 μ M of antibody, and 0.24 μ M of protein. This experiment was performed in a Nunclon® 96 well black bottom plate at a temperature of 37°C and buffered at a pH of 7 (20mM HEPES 7.3, 0.01% tween). It is shown based on the fluorescence intensity in the 300nm-400nm that tryptophan's are present in the antibody (6F4).

7.2.2. Fluorescence polarization using AFTD and antibody 6F4

In order to further investigate AFTD binding to E6, 6F4 antibody was tested in fluorescence polarization competition assays (Figure D). A competitor would decrease the polarization associated to ligand binding to E6, as it would displace the ligand from the binding site. It was observed in this assay that at 60min, and 15min that the polarization actually increased (Figure D). This was an interesting result as polarization is proportional to the molecular weight of the complex. It is expected that the polarization values for the complex of

the antibody, AFTD and E6 would be higher than the AFTD complex values. This increase in polarization indicates that both antibody (6F4) and AFTD are bound at the same time. This property could be used to track protein-protein interactions involving E6. The results obtained for the 6F4 antibody show that AFTD is a good candidate for the development of a fluorescence-based assay to study protein-protein interactions involving E6.

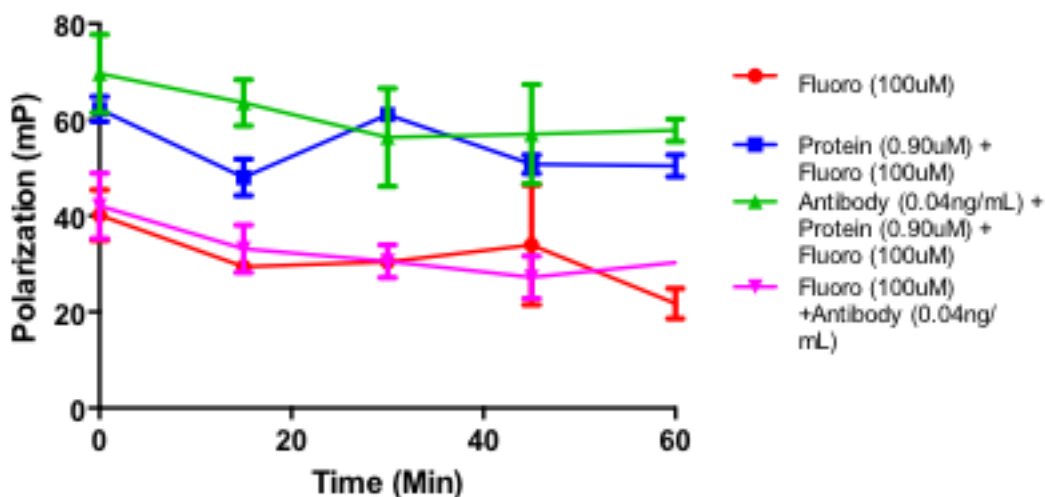


Figure D: Time curve showing the change in polarization of the protein and ligand complex and the antibody, protein ligand complex. This experiment was run for 60 minutes using 0.4ng/mL of antibody, 100 μ M of ligand, and 0.90 μ M of protein. This experiment was performed in a Nunclon® 96 well black bottom plate at a temperature of 37C and buffered at a pH of 7 (20mM HEPES 7.3, 0.01% tween). This graph demonstrates the increase in polarization at 15minutes and 60minutes when the antibody (6F4) is present. This increase in polarization indicates a complex between AFTD, 6F4, and E6.

This result is in agreement with the modeling results, as the E6 specific antibody is known to bind the N-terminus, whereas the docking results suggest that AFTD binds to the C-terminus domain of E6. Further concentrations of the antibody and of AFTD need to be tested in order to better characterize the complexes. From this assay we can conclude that there is some evidence to indicate that antibody (6F4) is binding and can be indicated by AFTD. This would be

further validated by attempting a dose response curve varying the concentration of 6F4 (section 5.5).

7.2.3. Tryptophan Fluorescence Results

Prior to obtaining these results four previous attempts were made. These results were all inconclusive due to lack of solvent controls and concentration of protein. In addition, the error between wells was extremely high in some cases. This was attempted prior to obtaining accurate results (section 5.2).

7.2.3. a) First attempt

Tryptophan fluorescence was initially conducted using a fixed concentration of both protein (0.01mg/mL) and ligand (10 μ M). Error was extremely high so error bars were not plotted, in order to make graphs easier to read. For this reason this plot was not included.

7.2.3. b) Second attempt,

The protein concentration was slightly increased from 0.01mg/mL to 0.02mg/mL. It was noted that the error between wells was significantly lower than in the first attempt. It was concluded that mechanical issues were partly to blame for the inconsistencies in the first attempt. Some parameters were adjusted in order to reduce error, such as the number of reads per well. The default number of scans is 10. We changed this value to 20 scans per well. The samples were run in triplicate instead of duplicate, which also aids in lowering the error. Only one ligand showed (O-succinyl-L-homoserine) a significant change (> 30%) in the protein's intrinsic tryptophan fluorescence (figure 5A). O-succinyl-L-homoserine also showed an increase in time starting at 39% at 30minutes then 42.9% at the one hour time point (Figure E).

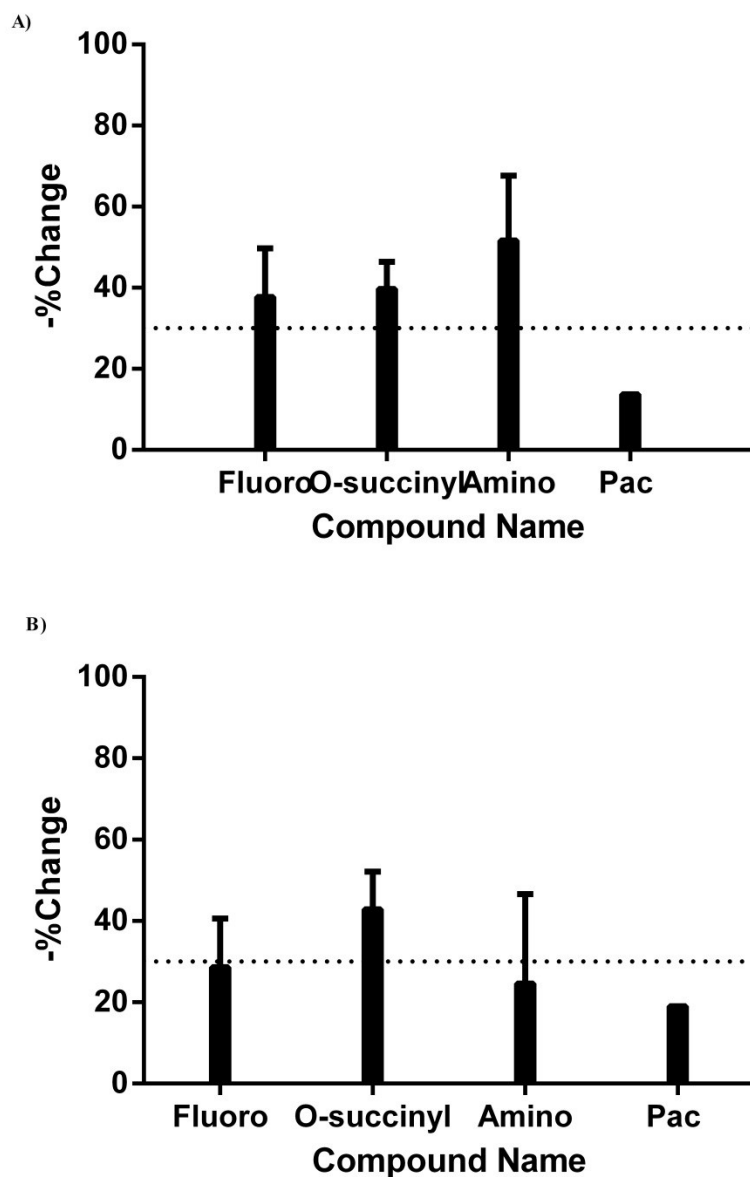


Figure E: Percent change in protein intrinsic tryptophan fluorescence upon ligand interaction. Experiment (A) 30minutes and (B) 1hour. The concentration of E6 was 0.02mg/mL for all samples. Each ligand was tested at 10 μ M. This experiment was performed in a Nunclon® 96 well black bottom plate at a temperature of 37°C and buffered at pH of 7.3 (20mM HEPES 7.3, 0.01% tween (v/v)). The change in tryptophan fluorescence was monitored over time and the percentage change relative to E6 without ligand was plotted using prism graph pad. At 1 hour incubation, O-succinyl-L-homoserine induces a 42.9%. Increase in Trp fluorescence. The dotted line represents the 30% threshold.

Some of the other ligands such as 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) and Aminopterin showed a decrease in change in intrinsic tryptophan fluorescence at 30 minutes versus 1 hour. Paclitaxel did not cause significant change in intrinsic Trp fluorescence, appearing not to bind, and its error bars were very high, which is why they were not included in the plot (Figure E).

7.2.3. c) *Third attempt*

The last experiment described appeared successful, however we did not control for solvent effects. O-succinyl-L-homoserine was dissolved in water so proper controls for this ligand were already in place. Other ligands such as 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine 1, 4- Dioxide (Fluoro), were dissolved in a half water/half ethanol solution. In addition, paclitaxel and aminopterin, were dissolved in 100% DMSO. Since solvent can have impact on fluorescence and potentially the environment of tryptophan, we repeated this experiment a third time to include solvent controls. A positive control was also added (Monoclonal antibody 6F4). This antibody was previously tested (Section 7.2.1), and shown to induce a significant change in intrinsic tryptophan fluorescence relative to the unbound protein. Since the previous experiments were positive for some of the ligands, we tested different concentrations of each ligand (25,10, and 1 μ M). In theory, the change in tryptophan fluorescence should be proportional to the concentration of the ligand. Unfortunately this experiment was unsuccessful for the concentrations tested, with no significant change in intrinsic fluorescence for any of the ligands tested, including our positive control (Figure F).

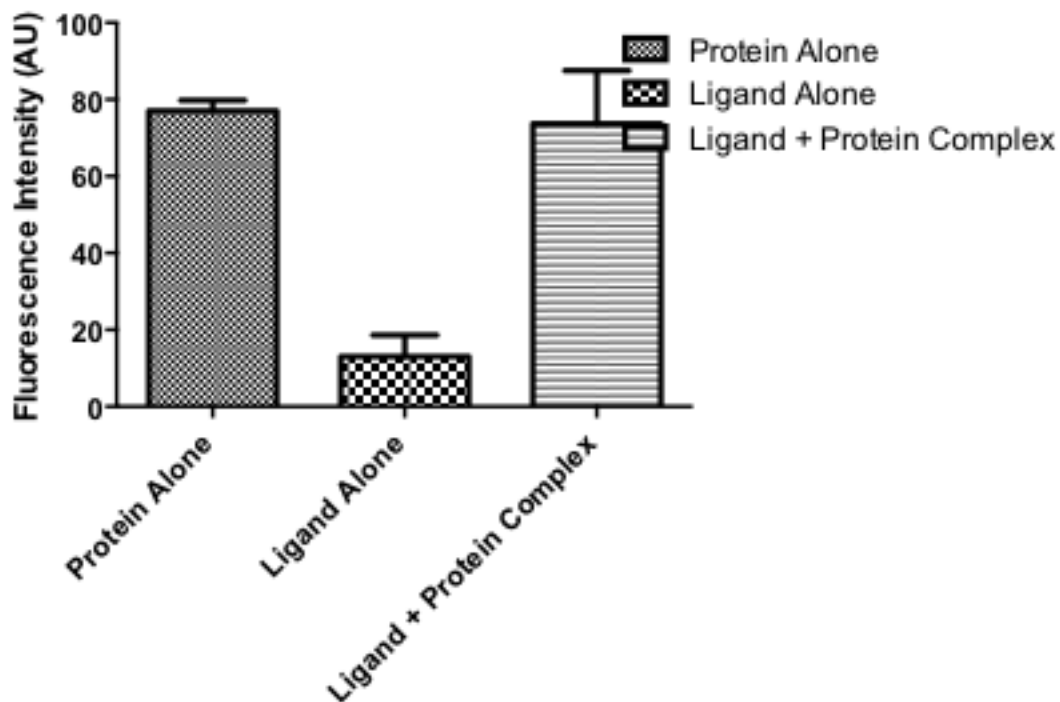


Figure F: Intrinsic Trp fluorescence. Protein, antibody (ligand) and antibody protein complex. Assay was run in a Nunclon® 96 well microplate in a BioTek Synergy 4 microplate reader. Samples were incubated at a temperature of 37°C and buffered to a pH of 7.3 (20mM HEPES 7.3, 0.01% tween (v/v)). Protein was at a concentration of 0.01 mg/mL and the antibody was at a concentration of 0.04µg/mL. Emission was collected at 350nm with an excitation wavelength of 288nm. No evidence of the antibody binding to the protein is observed, even though the antibody is E6-specific. This is believed to be in large part to the low concentration of the protein used.

It was concluded that the protein concentration used was too low to conduct tryptophan fluorescence. No significant change in intrinsic Trp fluorescence were seen associated to any of the ligands tested (Paclitaxel, Aminopterin, AFTD, and O-succinyl-L-homoserine). Comparing this result to our first tryptophan fluorescence attempt this experiment confirmed that a protein concentration of 0.01mg/mL is too low for the tryptophan fluorescence experiments.

7.2.3. d)Fourth attempt

The tryptophan fluorescence experiment was repeated using a higher concentration of protein (0.02mg/mL) under the same conditions as the previous experiment. Antibody 6F4 appeared to be below the 30% threshold in this experiment, which indicates experimental issues. O-succinyl-L-homoserine appears to be a positive result, however does not show dose dependence (Figure G).

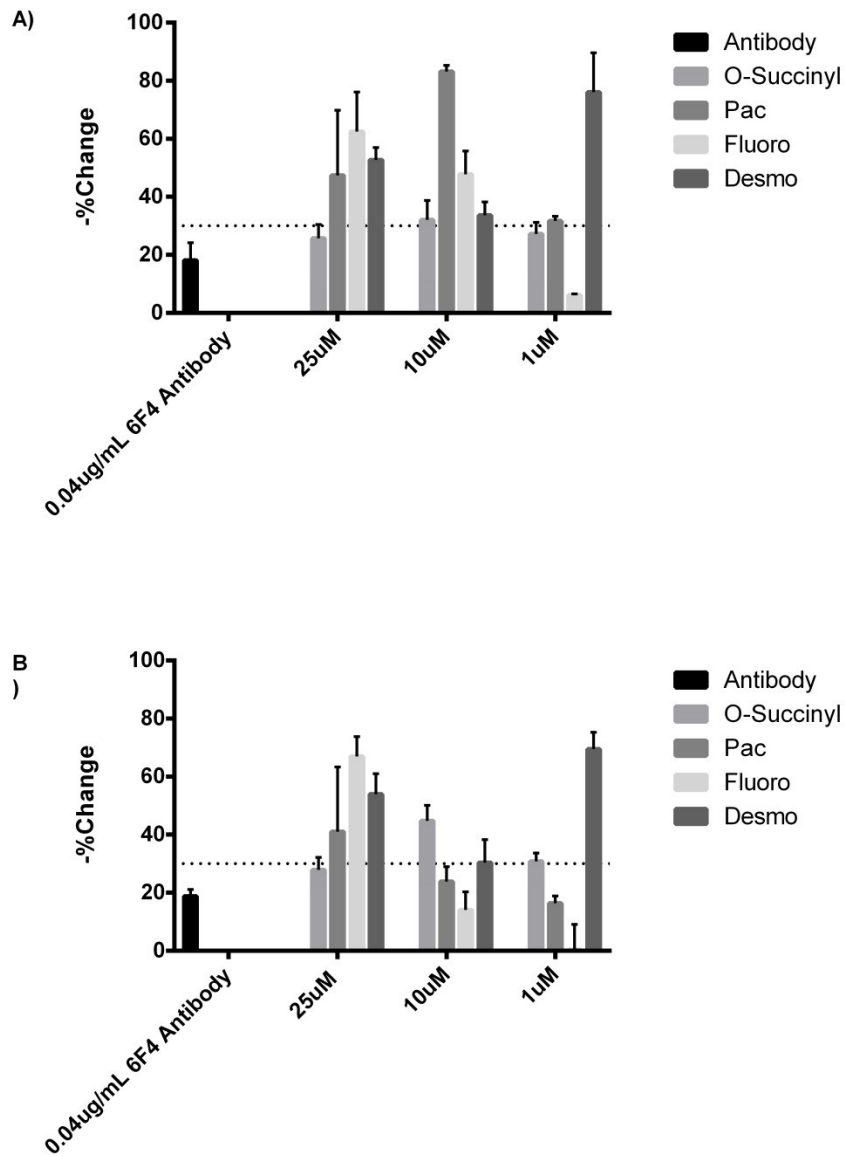


Figure G: Bar graph of change in E6's intrinsic tryptophan fluorescence experiment A) 30min B) 1hour. Change is shown at 30min (A) and 1hour (B) incubation times. Assay was run using Nunclon® 96 well microplate in a BioTek Synergy 4 microplate reader. Samples were incubated at a temperature of 37°C and buffered to a pH of 7.3 (20mM HEPES 7.3, 0.01% tween (v/v)). Protein was at a concentration of 0.02mg/mL and the antibody was at a concentration of 0.04µg/mL. Ligands were tested at 25µM, 10µM, and 1µM. Emission was collected at 350nm with an excitation wavelength of 288nm. The dotted line represents the 30% change in intrinsic fluorescence which is considered to be significant and associated to binding. The change in Trp fluorescence is dose-dependence for 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD). However our positive control (antibody) was negative.

In this case the best result was AFTD (figure G). Dose dependence was shown for this ligand at the concentrations used in this assay. The 10 μ M and 1 μ M appeared to show dose dependence for Paclitaxel, however the 25 μ M concentration due to high error was inconclusive (Figure G). This high error is most likely associated with paclitaxel precipitating out of solution at high concentration as it requires a high concentration of DMSO to stay in solution. One interesting result from this assay was that the antibody (6F4), which was previously shown to be positive, was negative. In this experiment the result was closer to 20% at best. One possible reason for this inconsistent result is due to improper storage of the antibody. This antibody had been stored in a minus 20°C freezer for a number of months and had been warmed to room temperature and refrozen a number of times. It is believed that this sample of our antibody was most likely denatured. Desmosine was added as a negative control, however appeared to be positive. When examining the computational results again this was not a surprise as Desmosine was still above our Binding Energy mean. O-succinyl-L-homoserine is not showing dose dependence. However the positive result at 10 μ M is consistent with previous experiments.

7.2.4. Fluorescence Polarization Results

7.2.4. a) Preliminary polarization assessment

Prior to attempting a dose response curve we need identify a concentration of fluorescent compound to use, and perhaps gain some hints as to potential concentration ranges we can use. The concentrations of fluorescent compound used were 10nM, 100nM, 10 μ M, and 100 μ M (Figure H). Each concentration was tested against various concentration of protein (Figure H). It was evident that the lower concentrations (10nM and 100nM) were not able to produce a substantial signal. The 10 μ M and 100 μ M concentrations were selected for further assays (Figure H).

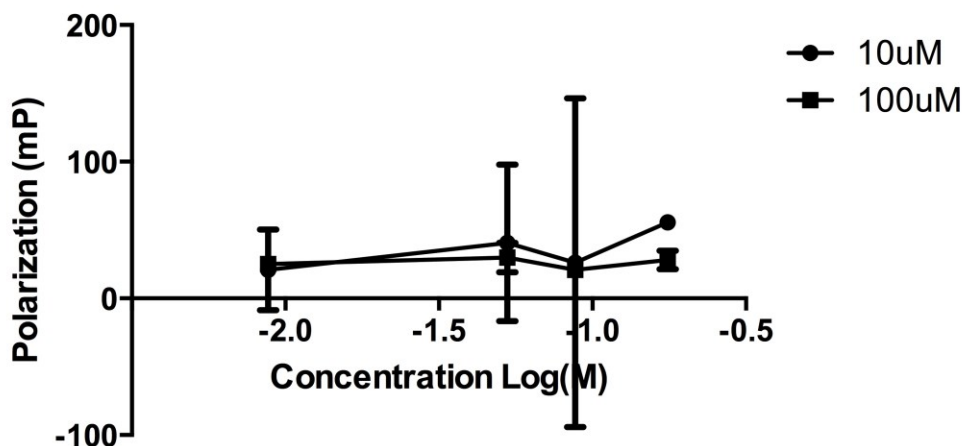


Figure H: Fluorescence polarization assay at 10 μ M and 100 μ M of 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) with varying the concentration of E6 (0.18 μ M, 0.09 μ M, 0.05 μ M, and 0.009 μ M) . Reaction was buffered at pH of 7 (20mM HEPES 7.3, 0.01% tween) and incubated at a temperature of 37C for one hour. Excitation filter (485/12) and emission filter (528/12) were used to measure fluorescence intensity in the parallel and perpendicular directions. The 100 μ M ligand concentration produces a cleaner signal with less standard error.

7.2.4. b) First dose-response curve attempt 100 μ M AFTD

A dose response curve for AFTD was first attempted using a fixed concentration of the compound (100 μ M). Initially results were poor and little to no convergence was found. This led to adjusting the probe height to 4mm, to better suit the total volume of 100 μ L. Using the adjusted probe height, a better curve was obtained at the three-hour time point (Figure I). From this graph we were able to use non-linear regression dose-response fitting four parameter was implemented in Prism (GraphPad, Inc.) to estimate EC50 (0.7 μ M) (Figure I), however this was just an estimate as although it converged the range was quite high.

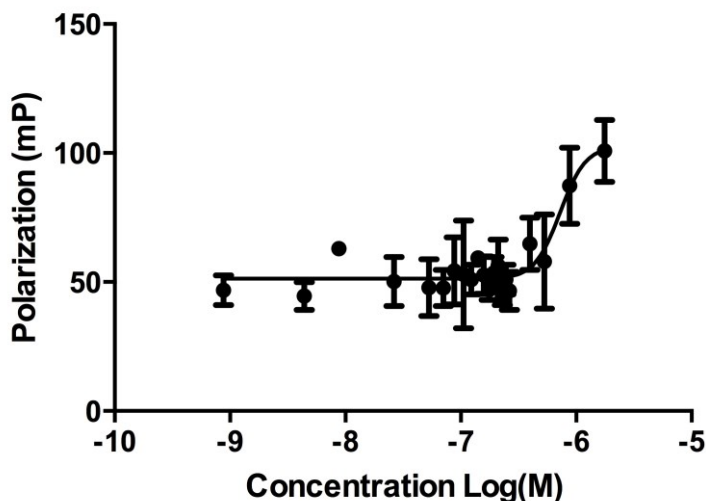


Figure I: Dose-response curve at 3 hour incubation using 100 μ M of 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) at varying protein (E6) concentrations (0.0010 μ M -1.8 μ M). Probe height was 4mm and a sensitivity of 50 was used. This experiment was performed in a Nunclon® 96 well black bottom plate at a temperature of 37°C and buffered at a pH of 7 (20mM HEPES 7.3, 0.01% tween). Excitation filter (485/12) and emission filter (528/12) were used to measure fluorescence intensity in the parallel and perpendicular directions. The top part of the curve is not well defined. Therefore greater concentrations of protein are required. An EC50 value of 0.7 μ M was estimated using Prism Graph non-linear regression four parameter equation.

Since the top part of the curve is not complete this EC50 value is merely an estimate.

Other time points to show that the system is equilibrated could not be analyzed as the probe height was not properly adjusted until the three hour time point. With this in mind, the experiment was repeated a second time with higher concentrations of protein. In addition, the concentrations were chosen to better space the points as the first curve had most points too close together (Figure I).

7.2.4. c) Dose-response curve evaluation 25 μ M and 50 μ M AFTD

A preliminary fluorescence polarization test was conducted at 25 μ M and 50 μ M to determine which of the two concentrations produces the best signal to noise ratio for a dose-

response curve. Values in this experiment were collected every 15 minutes and it was determined that our system equilibrates around 45min-60min incubation time. The 50uM curve at one hour (figure 10A) produced an adequate signal to noise ratio, however the shape of the dose-response curve was not ideal as the concentrations of AFTD used were not sufficient to produce a significant polarization signal (Figure J). The 25uM curve (Figure J) also produced good signal to noise ratio with a more satisfactory curve (Figure J).

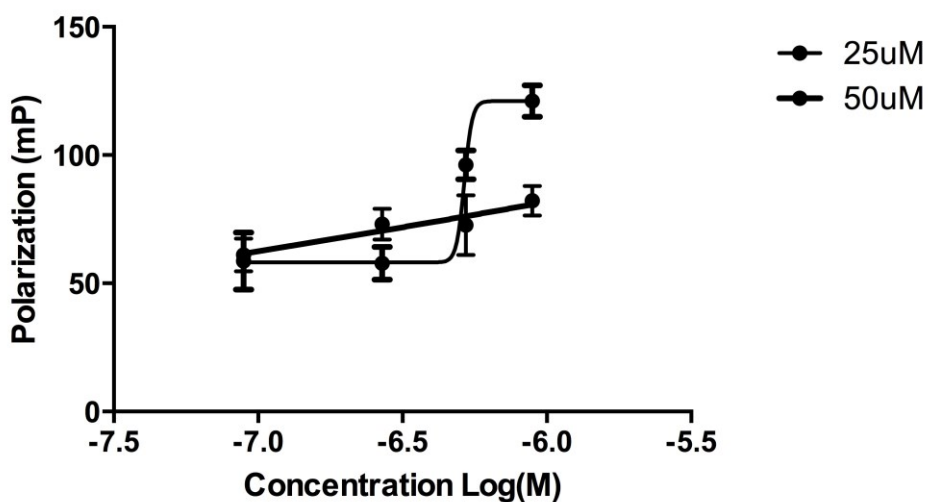


Figure J: 1 hour dose response attempt using 25uM and 50uM of AFTD. Various concentrations of E6 protein were used (0.007mg/mL -0.07mg/mL). This experiment was performed in a Nunclon® 96 well black bottom plate at a temperature of 37°C and buffered at a pH of 7.3(20mM HEPES 7.3, 0.01% tween). From this plot it is evident that the shape of the 25uM is much better than the 50uM curve. 25uM curve had an EC50 value of 0.04mg/mL where the 50uM curve had an EC50 values of 0.25mg/mL.

An EC50 value of 0.04mg/mL was estimated for the 25uM curve using the non-linear regression dose-response four parameter function. It was concluded that a ligand concentration of 25uM was better suited for future dose-response curves with variable protein concentrations.

7.2.4. d) Full 25 μ M Dose-response attempt

With the information gathered above, a new dose response curve was obtained at 25 μ M of 3-amino-5-fluorobenzo [e] [1,2,4] triazine 1, 4 dioxide. The protein concentrations varied from 0.09-1.07 μ M (Figure K).

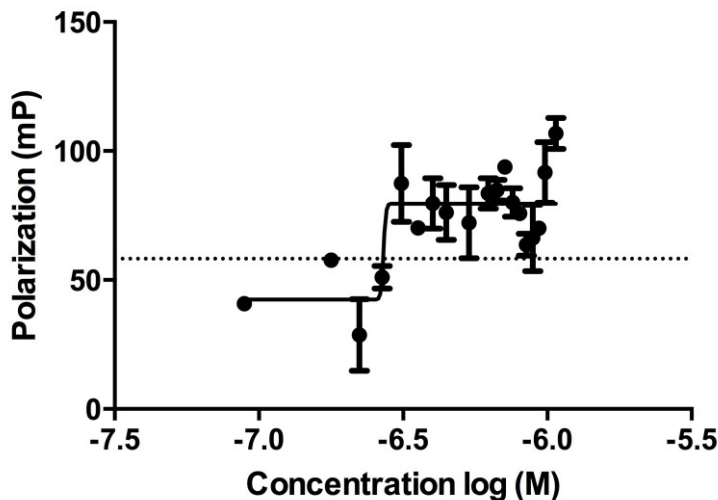


Figure K: One hour dose response attempt using 25 μ M AFTD. Various concentrations of E6 protein were used (0.007mg/mL -0.07mg/mL). This experiment was performed in a Nunclon® 96 well black bottom plate at a temperature of 37°C and buffered at a pH of 7.3 (20mM HEPES 7.3, 0.01% tween) at fixed concentration of AFTD 25 μ M. Unfortunately this curve did not converge and it was determined that a different approach would need to be taken in order to obtain a good quality dose-response curve.

The curve obtained was of poor quality and could not be used to estimate EC50 with confidence.

Our concentrations were noticeably too close in log units and the top part of the curve is not well-defined. The best time point curve was observed after 45minutes (figure K). Since a good-quality dose-response curve could not be obtained varying the concentration of the protein, a different approach was taken.

7.2.4. e) Dose-response curve varying AFTD concentration

Based on the equilibrium analysis, it was determined that data would be collected every 15 minutes after 30 minutes until the 75 minute mark. Another important test was to determine the proper sensitivity setting to use. This sensitivity needs to be high enough to give us a reproducible signal for our lowest concentration and not so high that it over flows our highest concentration. Through testing a few sensitivities, it was determined that a sensitivity of 100 was appropriate for the proposed concentration range of 3-Amino-5-Fluorobenzo-1,4-Triazine Dioxide ($300\mu\text{M}$ - $1\mu\text{M}$).

Using the preliminary work a dose-response curve was obtained with variable ligand concentration, using the parameters determined in the preliminary experiments. Numbers right from the first time point at 30minutes appeared to be off. One big concern was that the perpendicular values appeared to be higher than the parallel values. Data was collected up to the one hour mark. The remainder of the scans collected were done using a slightly different protocol, however this appeared to not fix the problem. Once experiment was completed, data was still analyzed in order to determine what the issue was. Upon further investigation, it was determined that our protein sample was contaminated. This was evident as the protein alone well had a high fluorescence intensity value. Upon analysis it was determined that it was contamination from the ethanol solution used to create the protein blank. Despite this, data analysis was conducted by discounting the fluorescence in the protein alone well in both directions. In addition, our lower concentration ($1\mu\text{M}$) did not produce a signal strong enough to calculate polarization and was therefore eliminated. The best plot was obtained at 30 minutes (Figure L)

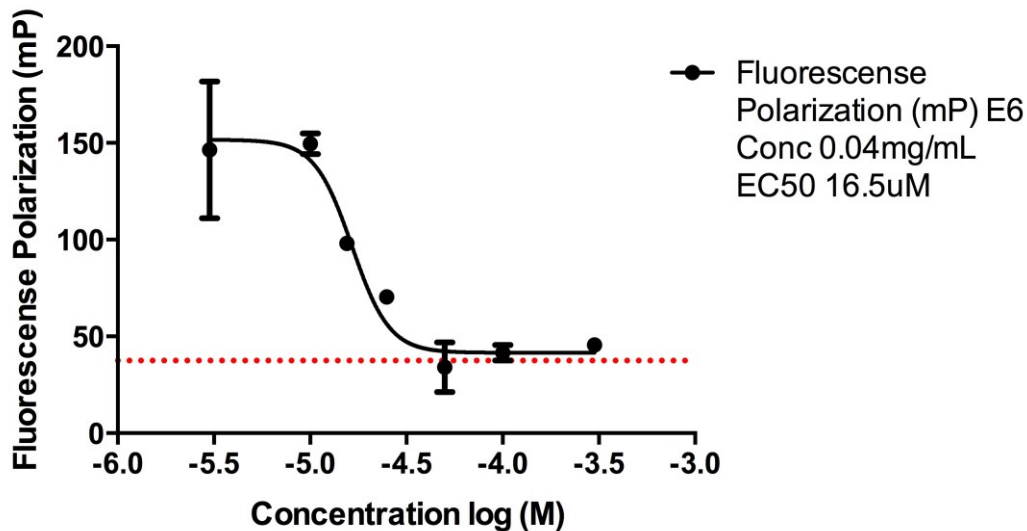


Figure L: 30minute does response attempt using 0.04mg/mL of E6 protein and varying the concentration of 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) (1 μ M-300 μ M). Probe height was 4mm and a sensitivity of 100 was used. This experiment was performed in a Nunclon® 96 well black bottom plate at a temperature of 37°C and buffered at a pH of 7.3 (20mM HEPES 7.3, 0.01% tween). Red line represents the background polarization of 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) on its own. EC50 value was estimated at 16.5 μ M.

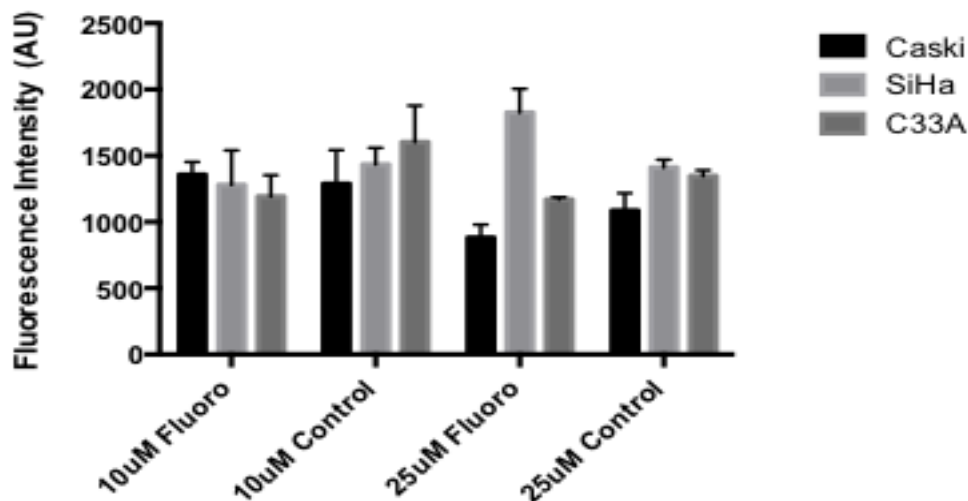
The 30minute time point converged and generated an EC50 value of 16.5 μ M. This value is close to our expected value. However due to contamination from the ethanol solution used this curve was not able to be trusted.

7.2.5.Preliminary Ligand Uptake Experiment

Cell permeability needs to be tested. Since our molecule is intrinsically fluorescent this can be used to trace if the molecule is getting inside the cell. CaSki, SiHa, and C33A cell types were used for cell permeability assays at 10 μ M and 25 μ M of AFTD. Ethanol control wells were present for each cell type and concentration. The cells were incubated at various time points (1, 3, and 6 hours), lysed and filtered with a 0.2 micron filter. Each well was kept as an individual sample for testing. Standards were also made so a calibration curve could be made to estimate

the concentration of our molecule within the cell. Each sample from each time point was then plated (100 μ M) and read in the Microplate reading using an excitation of 270nm and emission at 525nm. Data was then exported into Prism Graph Pad where bar graphs were made (Figure M).

A)



B)

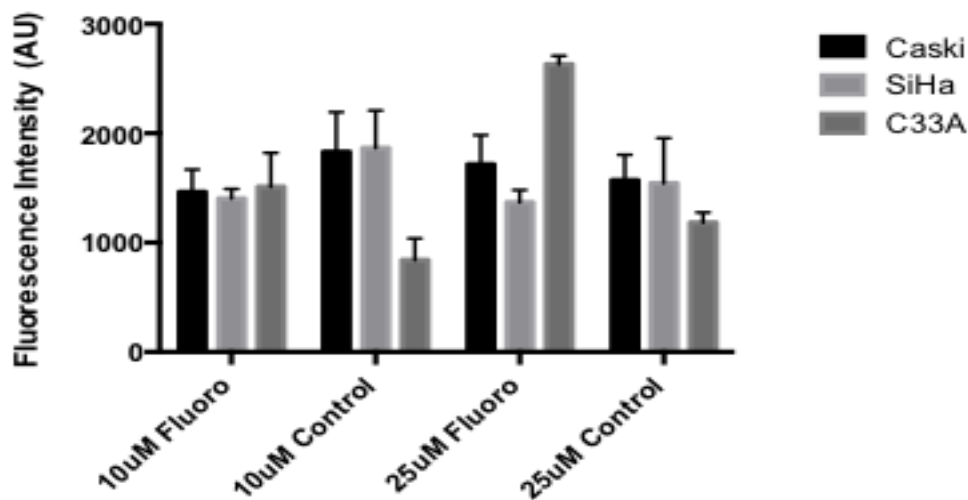


Figure M: Fluorescence intensity at 1 hour (A) and 3 hour (B) incubation of 3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide (AFTD) at 10µM and 25µM with CaSki, SiHa, and C33A cell lines. Intensities collected from the first cell permeability test. Cells were lysed after incubation and filtered with 0.2 micron filter. 100µM of each sample were analyzed with an excitation of 270 nm, and emission at 525 nm in a Biotek Synergy 4 microplate reader. Values were blank-discounted. Samples were compared to their corresponding controls which account for cellular background fluorescence. Some evidence of cell permeability was obtained in particular in the C33A cells. However, this experiment needs to be repeated to show reproducibility and validity.

The most significant differences between samples and controls were observed for the C33A cells. Some difference was also present in the SiHa cells; however the background control appeared to be higher than the sample after three hours. Some of these inconsistencies could be due the use of a monochromator instead of optical filters for the readings. Also, the readings were not normalized for cell density. It is assumed that the same number of cells in each well is the same, however this was not determined and many factors may alter cell growth and proliferation differently across samples. In addition by the time the fluorescence was read the cells had been sitting in DMSO for over 48hours. Since DMSO is a harsh solvent it could cause degradation of many components contributing to the fluorescence readings. The experiment was thus repeated to address these issues and show reproducibility.

Chapter 8. Appendix B: Supplementary Material

Parameters used in Cassandra for:

protocol HierDock

job_distribution_mode per_ligand

protein_type globular

diversity_option none

grid_margin 15

grid_spacing 0.25

flexible_ligand yes

maximum_orientations 10000

torsion_drive no

nconf_level0 1500

bump_maximum 10

anchor_search no

nconf_filter 150

bursurface_cutoff 70 *note changed to 30 for shelved compounds

energy_cutoff 100

mpsim_threads 1

min_method_mpsim CONJUGATE_BDO

cmm_level_level1 0

cmm_level_level2 3

nb_update_frequency_level1 5

nb_update_frequency_level2 10

force_field_file dreidii322-mpsim.par

nconf_level1 3

ptn_option_level1 fix

max_steps_level1_min 30

eps_ff_min 2.5
ptn_option_level2 move
max_steps_level2_min 20
solv_method_oneE AVGB
solv_method_min GAS
eps_ff_oneE 1.0
solv_in_eps 1.11
solv_out_eps 78.2
solv_probe_radius 1.4
solv_ionic_strength 0.1
solv_ptn_option move
nconf_diversity 50
diversity_threshold 0.6
rejection_threshold 2.2
sample_min 500
no_family 50
child_avg 6
enrichment_maximum 1000000

Parameters Used for MMFF94x Energy Minimization of E6 and ligand Databases

Gradient 0.05

Options: Calculate Forcefeild Partial Changes

The remaining boxes were unchecked.

Enable: bonded, van der Waals, Electrostatics, Restraints, and cutoff

Solvation: Distance

Scale Like: 1

On: 8

Off: 10

Dielectrin: 1

Exterior:80

Unlike: 0

Wild: 1

Treads: 0

Table A: Compounds that passed the R1 threshold with structures

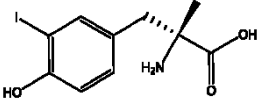
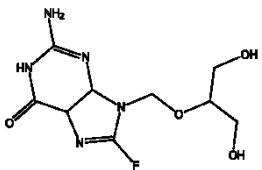
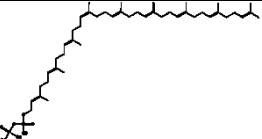
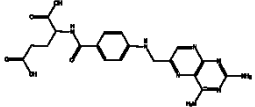
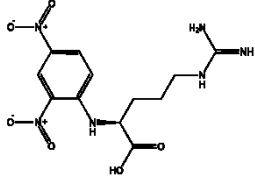
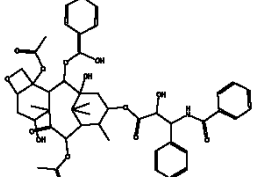
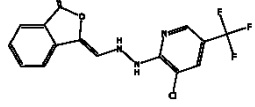
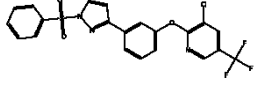
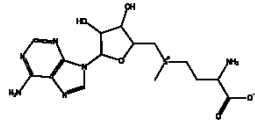
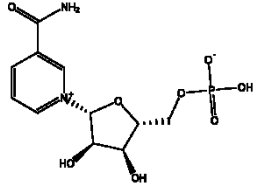
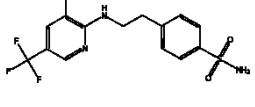
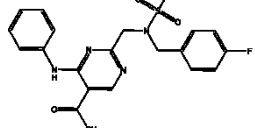
Rank	Identifier	Binding Score (kcal/mol)	IUPAC name	Molecular Structure
1	cmp450432	-55.93	(2R)-2-amino-3-iodanyl-phenyl)-2-methyl-propanoic acid	
2	cmp450769	-55.29	2-amino-8-fluoranyl-9-[[2-hydroxy-1-(hydroxymethyl)ethoxy] methyl]-4,5-dihydro-1H-purine-6-one	
3	SOLANESYL-PYROPHOSPHAT E-1-3H	-55.06	Solanesyl-PyrophosphatE-1-3H	

Table B: Compounds that passed R2 threshold with structures

Rank	Identifier	Binding Score (kcal/mol)	IUPAC name	Molecular Structure
1	AMINOPTERIN-3-5-7-9-3H-N	-80.13	Aminopterin-3-5-7-9-3H-N	
2	PA24273	-78.58	N-(2,4-dinitrophenyl)-L-Arginine	
3	PACLITAXEL-BENZOATE-RING-UL-14C-	-78.41	Paclitaxel-Benzoate-Rink-UL-14C	
4	MFCD01113139	-68.33	3-Chloro-2-(2-([3-oxo-2-Benzofuran-1(3H)-Ylinden]methyl)hydrazino)-5-(Trifluoromethyl)Pyridinium Acetate	
5	MFCD00214696	-66.13	3-Chloro-2-(3-[1-(Phenylsulfonyl)-1H-Pyrazol-3-YL]Phenoxy)-5-(Trifluoromethyl)Pyridine	
6	cmp450572	-66.06	2-amino-4-[[5-(6-aminopurin-9-yl)-3,4-dihydroxy-tetrahydrofuran-2-yl]methyl-sulfonio]butanoate	
7	PA24775	-65.07	Beta-Nicotinamine Mononucleotide	
8	MFCD00172501	-61.85	4-(2-([3-Chloro-5-(Trifluoromethyl)-2-Pyridinyl] Amino) Ethyl) Benzenesulfonamide	
9	MFCD15731617	-61.14	4-Anilino-2-((4-Fluorobenzyl) (Methylsulfonyl) Amino) Methyl) Pyrimidine-5-Carboxylic Acid	

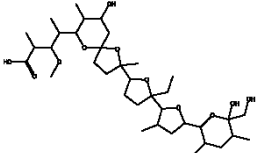
10	MONENSIN-3H-G	-60.28	Monensin-3H-G	
----	---------------	--------	---------------	---

Table C: Promising lead compounds from the NT-E6 virtual screening with molecular structure

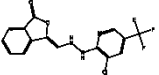
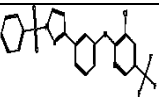
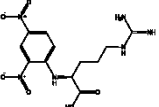
Computational Identifier	Rank/Region	Molecular name	Binding Energy Score (kcal/mol)	Commercial Availability	Molecular Structure
MFCD01113139	4/R2	3-Chloro-2-(2-[(3-oxo-2-Benzfuran-1(3H)-Yliden] methyl)Hydrazino)-5-(Trifluoromethyl) Pyridinium Acetate	-68.33	Bionet	
MFCD00214696	5/R2	3-Chloro-2-(3-[1-(Phenylsulfonyl)-(H-pyrazol-3-YL] Phenoxy)-5-(Trifluoromethyl) Pyridine	-66.13	Bionet	
PA24273	7/R2	N-(2,4-dinitrophenyl)-L-arginine	-78.58	Sigma Aldrich	

Table D: Top three compounds from the primary amine database in order of rank with molecular structure

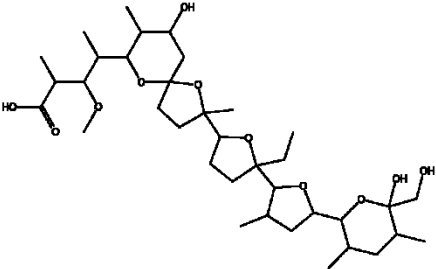
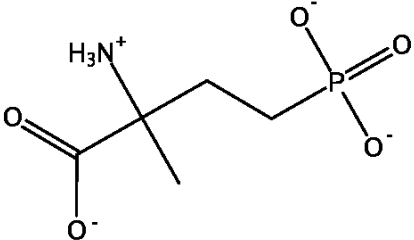
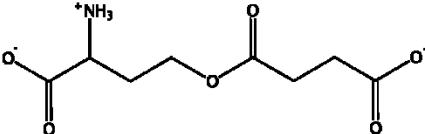
Compound ID	Compound Name	Binding Energy (Kcal/mol)	Commercial Availability	Molecular Structure
Cmp47400	L-mimosine	-46.81	Sigma Aldrich	
Cmp144614	(S)-2-Amino-2-methyl-4-phosphonobutanoic acid	-45.83	Sigma Aldrich	
Cmp37088	O-succinyl-L-homoserine	-45.62	Sigma Aldrich	

Table D: Top 5 compounds from the fluorine-containing molecule database in order of rank with molecular structure

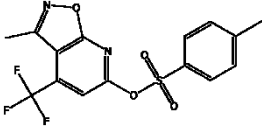
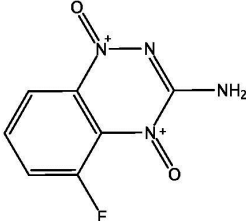
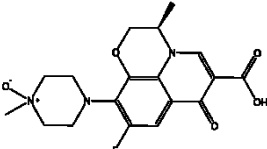
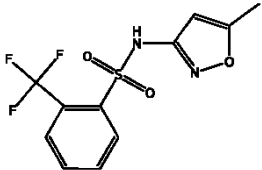
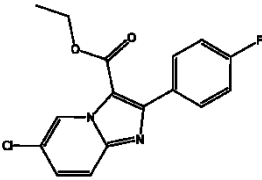
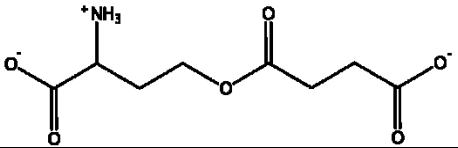
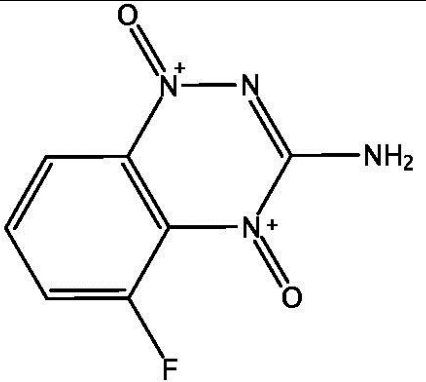
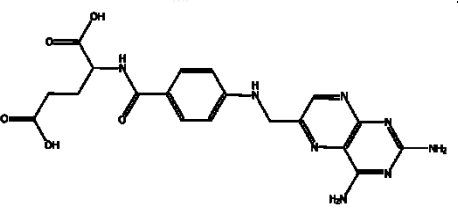
Compound ID	Compound Name	Binding Energy (Kcal/mol)	Commercial Supplier	Molecular Structure
MFCD09040704	3-methyl-4(trifluoromethyl) Isoxazolo [5,4-B] Pyridin-6-YL-4-methyl benzene sulfonate	-62.22	Enamine	
MFCD14636665	3-Amino-5-fluorobenzo [E] [1,2,4] Triazine 1, 4 dioxide	-49.53	Sunbiochem	
MFCD00871869	Ofloxacin N-Oxide	-48.80	Sinochem	
MFCD00955225	(5-methylisoxazo-3-YL) ((2-(trifluoromethyl) phenyl) sulfonyl) amine	-48.70	ABCR and Ryan Scientific	
MFCD11977074	Ethyl 6-chloro-2-(4-Fluorophenyl) imidazo [1,2-A] pyrindine-3-carboxylate	-47.94	Golden Bridge Pharma	

Table E: Ligands selected for experimental testing with molecular structures

Compound Name	Overall Rank	Binding Energy (kcal/mol)	Abbreviation	Molecular Structures
O-succinyl-L-homoserine	25	-45.62	O-succinyl	
3-Amino-5-Fluorobenzo [E] [1,2,4] Triazine-1,4 Dioxide	12	-49.53	AFTD	
Aminopterin	8	-59.61	Amino	
Paclitaxel	4	-70.10	Pac	