PRODUCTION AND SCREENING OF A RANDOM PEPTIDE LIBRARY FOR BLOCKING HIV GP120 ATTACHMENT TO CD4

A Major Qualifying Project Report

Submitted to the Faculty of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

in

Biology and Biotechnology

by

Andrew Bartley

March 9, 2010

APPROVED:

Paul Clapham, Ph.D.
Program in Molecular Medicine
UMASS Medical Center
Major Advisor

David Adams, Ph.D.
Dept. Biology and Biotechnology
WPI Project Advisor
ABSTRACT

Developing a vaccine against HIV has been very difficult mainly because HIV is highly variable. Many different avenues of overcoming this difficulty are being researched, including interfering with HIV attachment to host cells. In all stages of HIV infection, the viral envelope glycoprotein (gp120) is necessary for attachment to CD4 surface proteins on human cells. To identify potential reagents for blocking gp120/CD4 interaction, a random peptide library was constructed in search of a peptide capable of binding to gp120’s CD4 binding site. The library was screened using an E. coli two-hybrid system, and yielded 53 peptides that interact with a mimic of gp120’s CD4 binding site.
TABLE OF CONTENTS

ABSTRACT ........................................................................................................................................ 2

AKNOWLEDGEMENTS .................................................................................................................. 4

BACKGROUND .................................................................................................................................. 5

PROJECT OBJECTIVES ................................................................................................................ 14

MATERIALS AND METHODS ..................................................................................................... 15

RESULTS ......................................................................................................................................... 20

DISCUSSION .................................................................................................................................... 28

BIBLIOGRAPHY ........................................................................................................................... 30
ACKNOWLEDGEMENTS

I would first like to thank Paul Clapham and Alexander Repik for allowing me to work on this project in their laboratory at UMass Medical School. This was an amazing opportunity to work on a cutting edge project that few are lucky enough to see. I would like to express additional gratitude Alex Repik for teaching and helping me refine several lab techniques, including bacterial transformation, plasmid purification, and PCR. Thank you to Alex again for all of his insight on project discussions and for providing the necessary primers for creation of the protein library. Finally, I would like to give a tremendous thanks to Dr. Dave Adams for his advice throughout the entire project. His advice was critical from the project planning stages, right through to the end stages of finishing the paper.
BACKGROUND

HIV Background

Human Immunodeficiency Virus (HIV-1) is a sexually transmitted and blood-borne infection that targets the immune system of its hosts. Since its initial identification in 1981, HIV has been spreading in a high profile epidemic globally (CDC, 2008). New infections, such as HIV, have previously been difficult to monitor, but recently the advent of a serologic testing algorithm for recent HIV seroconversion (STARHS) has allowed a more accurate estimate of infection numbers (CDC, 2009). It is currently thought that over 1 million Americans, and 33 million people worldwide, are infected with HIV (UNAIDS, 2008), with about 40,000 new cases diagnosed annually in the US alone (Figure 1). This is an improvement from the peak U.S. infection rate in 1992 (CDC, 2007). The rates of both new infections and deaths dropped sharply around 1996, when a new cocktail of AIDS therapies became widely available.

![AIDS diagnoses and deaths](image.png)

**Figure 1**: The Annual Incidence of AIDS in the U.S. Shown are the annual numbers of AIDS diagnoses (orange) and deaths (brown) in the U.S. The peak diagnoses occurred in 1992, and the peak deaths in 1994. (CDC, 2007)
**HIV Cellular Entrance**

Since HIV cannot sustain itself very long outside of a human host, it requires a very direct means of transmission (CDC, 2008). HIV is transmitted from person-to-person in three ways: direct sexual contact, sharing hypodermic needles, or during the process of pregnancy, birth, or breastfeeding (CDC, 2008).

Another reason for HIV’s need for direct transmission, more directly related to this MQP, is the mechanism in which the virus attaches to human host cells. HIV virions typically attach to T-cells with CD4 (Cluster of Differentiation) glycoproteins expressed on their cell surface; these CD4$^+$ cells are referred to as T-Helper cells or CD4 T cells (Wyatt, et al., 1998). In order for the HIV virions to attach to T cells, the virus envelope proteins (or “spikes”) must attach to the CD4 receptors of human T cells (Wyatt, et al., 1998). After gp120 has bound to the T-cell’s CD4 receptor, an HIV co-receptor must also bind to the gp120 (Kasson, 1999). Depending on the stage of HIV infection, this co-receptor is usually the chemokine receptors CXCR4 or CCR5, but in very rare situations can also be different receptors for certain strains of HIV (Kasson, 1999). Once gp120 attaches to two receptors, the envelope trimer induces fusion of the viral and cellular membranes, so the viral contents are transferred into the cell and infection will begin.

The purpose of this MQP is to identify candidate proteins that will prevent the attachment of HIV particles to human CD4 T cells so infection does not occur. Our specific interest is finding candidates that will bind the CD4 binding site on gp120 to directly inhibit the interaction between gp120 and CD4 glycoproteins by direct competition.

The approach of blocking HIV infection by blocking viral entrance is strongly supported by previous studies showing that antibodies against CD4 or the co-receptors can block HIV
infection *in vitro* (Kasson, 1999). *In vivo*, if insufficient HIV particles infect a host, then the human immune system can naturally fight off the remaining particles.

**Glycoprotein 120**

gp120 is an ideal candidate for a target in HIV attachment interference. The HIV spike complex is composed of two proteins, surface glycoprotein 120 (gp120) and a transmembrane glycoprotein 41 (gp41), assembled as a trimer of heterodimers (Figure 2) (Stowell, 2006).

![Diagram of an HIV Envelope Spike](image)

**Figure 2: Diagram of an HIV Envelope Spike.** The diagram provides a basic visualization of the HIV envelope spikes that are responsible for HIV virion attachment to human T-cells. The spike is composed of gp120 on the surface and gp41 transmembrane protein. One spike consists of three subunits. (Stowell, 2006)

gp120 has five variable domains (V1-V5) and several conserved domains (Figure-3). The CD4 and chemokine receptor binding regions of gp120 are rather conserved, and should serve as stable and useful targets for binding peptides to block infection (Wyatt, et al., 1998). The two dimensional gp120 structure shown in Figure 3 indicates the conserved domains may not be available for binding competing peptides. The two-dimensional model leads the viewer to
believe that the V1, V2, and V3 loops (green and yellow in the figure) all partially block the conserved regions of the CD4 binding pocket (black arrows).

**Figure 3: Two Dimensional Structure of the gp120 Trimer.** The diagram shows a 2D model for gp120 trimerization. Black arrows point to the conserved CD4-binding sites. Yellow denotes the V3 loops, and green the V1 and V2 loops. The conserved chemokine-receptor-binding sites are shown in red (Wyatt, et al., 1998).

However, a three-dimensional view of the gp120 monomer (**Figure 4**) shows a different perspective in which there is space for a small competing protein to make its way into the conserved CD4-binding pocket (right panel, red site).
**Figure 4:** 3D Diagram of the HIV gp120 Trimer Core. **Left panel** shows the major areas of interest on the gp120 monomer, including V1-V3 variable regions, alpha-helices, and linker regions. **Right Panel** shows the residues within 4 Angstroms that bind to neutralizing antibodies. The CD4BS (CD4 binding site) is of specific interest to this project, and may be amenable for binding small protein competitor molecules (Wyatt et al., 1998).

**Bacterial Two-Hybrid System**

A key component of this project is a bacterial two-hybrid system which tests protein-protein interactions by transcriptional activation. The two proteins being tested for interaction are referred to as the bait and the prey (target) (Love & Hochschild, 2004). The bait and prey are expressed in bacteria by transforming them with two plasmids, pBT and pTRG, respectively (Figure 5).
Figure 5: Schematic of the BacterioMatch II Two-Hybrid System for Detecting Protein-Protein Interactions. The figure shows how the bacterial two-hybrid system works. Bait and prey (target) encoding plasmids pBT and pTRG, respectively, must both reside in the reporter strain of *E. coli*, each expressing their own peptide fused to λ.cl or bacterial RNA polymerase (RNAP), respectively. If the bait and target peptides interact (lower diagram), then their fusion peptides RNAP and λ.cl will bind to the DNA and initiate transcription of the HIS3 reporter genes. If the reporter gene phenotypes are expressed, then the bait and prey peptides have interacted (Stratagene).

The mechanism of detecting interaction between the two bait and prey proteins is the expression of a reporter gene (or genes) that will only be transcribed if the two proteins physically interact (Love & Hochschild, 2004). The bait protein is fused to λ.cl, whose function is to bind the operator of the reporter genes by a DNA sequence-specific fusion. The prey represents the protein that is being tested for interaction with the bait or target. The prey protein is fused with a subunit of the bacterial RNA polymerase (RNAP)(Stratagene). If the bait and prey proteins interact, then the RNAP will be brought directly to the promoter region of the reporter gene and begin transcription. Typically, transcription will not occur when the bait protein alone is proximal to the operator since the binding of cellular RNAP is blocked from the promoter region.
If the bait and prey do not bind together, then the RNAP subunit will not be proximal to the promoter, and little or no transcription occurs.

Depending on which reporter genes are chosen, the result of the protein-protein interaction will be different for each two-hybrid system. In the example shown in Figure 5, commonly used reporters are *HIS3* and *aadA*, both of which are expressed with positive protein-protein interaction (Love & Hochschild, 2004) (Stratagene). *HIS3* encodes a component in the histidine biosynthetic pathway. The transformed cells are plated with 3-amino-1, 2, 4-triazole (3-AT), a competitive inhibitor of the cellular *HIS3* pathway (Stratagene). Bacterial cell growth will not occur unless sufficient *HIS3* reporter product is successfully produced. *aadA* encodes for resistance to the antibiotic streptomycin and acts as a double check for the system; the transformed bacteria are plated on 3-AT/streptomycin plates to help screen large protein libraries for different prey proteins (Love & Hochschild, 2004) (Stratagene).

With respect to this MQP, the reporter genes used were not *aadA* and *HIS3*, but instead *bla* (β-lactamase) and *lacZ* (β-galactosidase) were used. The *bla* gene confers resistance to carbenicillin, and *lacZ* expression will result in transformed colonies having a blue color when plated onto X-gal plates (Love & Hochschild, 2004). *LacZ* expression also allows for a quantitative measurement of protein-protein interaction through use of a liquid β-galactosidase assay (Love & Hochschild, 2004).

*Bait Mimic Protein*

A mimic protein of the CD4-binding site on gp120 will be used as bait in this project. This mimic will be used instead of natural gp120 due to the problem of expressing a *functional* form of gp120 in bacteria. A properly designed mimic protein of a small size will circumvent many of these synthesis issues, and facilitate production of the bait. The CD4-binding site
(CD4bs) on gp120 is discontinuous, and usually requires the whole molecule to correctly bind to CD4. To create the mimic, two discontinuous segments involved with CD4 binding were "excised" (amplified by PCR) from gp120’s gene then linked together by A. Repik in our laboratory (Figure 6) (Repik A, 2009). Additionally, the CD4 binding loop comprises IIFKQSSGGDPEIVT. This is likely to be the first site on gp120 contacted by CD4 and possibly the only part of the CD4bs exposed on the native trimeric envelope.

![Figure 6: Primary Amino Acid Structure of the Bait Protein Used in this Project.](image)

The diagram shows the primary structure of the CD4-binding site mimic protein designed by A. Repik in our laboratory. The blue arrows represent the β-strand segments, the blue amino acids represent known CD4-binding residues. The red amino acids represent the linker regions connecting the previously discontinuous CD4 binding sites. The yellow highlighted region is the M5 mimic used as bait in the bacterial two hybrid system for this MQP (Repik A, 2009).

Peptide Prey Library

The last (and most important) component of the two-hybrid system is the library of peptides that will be screened for binding to the conserved CD4-binding epitopes to hopefully interfere with the gp120-CD4 viral/host interaction. This peptide prey library has been designed as a series of “gene snippets” that will be cloned as a part of this MQP into plasmid pTRG for expression in *E. coli* as part of the two-hybrid experiment. Each snippet is a 110-mer whose sequence is mostly random (Figure 7). The gene snippets are sense-oriented but carry specific
restriction enzymes at each end. NotI is at the forward end, and XhoI is at the reverse. A library of peptides was thus produced as described in the Materials and Methods section. This library will be screened in the two-hybrid system to identify candidates that bind the CD4-binding site of gp120. The NotI and XhoI primers used will be of great benefit to help sequence potential positives.

**Figure 7: Production of the Randomized Prey Peptides.** The diagram shows how the prey peptide library was prepared. Two primers, one forward and one reverse, were used to bookend (flank) 26 NNK base pairs, creating a mostly randomized gene library ready for direction-specific transformation. Once transformed, the genes will be expressed as our peptide library (Repik A., 2009).
PROJECT OBJECTIVE

The purpose of this MQP is to use a bacterial two-hybrid system to identify peptides that bind CD4-binding epitopes on HIV gp120, to create potential therapeutics that may prevent the binding of viral gp120 to host cellular CD4. The bait peptides represent various combinations of known CD4-binding epitopes present on gp120. The prey peptides represent randomized short peptides that should be short enough to penetrate to the site of CD4 binding. Any positives will signify candidates for further investigation as potential inhibitors of blocking gp120-CD4 binding.
MATERIALS AND METHODS

Preparation of Prey Peptide Library Genes

The random peptide library to serve as prey was created by PCR as a gene library using forward and reverse primers:

**CYSLIB** - Forward
5’- GCGGCCGGATCCGCGGCCGCAGGGGGGGGTCTTGC

**CYSLIB** - Reverse
5’-GGATCTCAGTTCAATTAATTTACTCGAGGCA-3’

The forward and reverse primers were used with an Invitrogen Platinum Taq Polymerase kit (available commercially) according to the manufacturer’s instructions to produce randomized sequences as follows:

5’GCCGCAGGGGGGCTCTTGCNNKNNKNNKNNNNKNNKNNKNNKNNKNKNNKNNN KNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNNKCCTCGAGTAATT AATTAAT -3’

The PCR was setup using three oligonucleotides. Two of which were the two primers shown above. The third was the randomized oligonucleotide that will serve as the template for PCR (shown previously in Figure 7). The process will yield randomized gene sequences that still have the conserved sequences on each end necessary for insertion into the plasmid.

Ligation of the Prey PCR Products into Plasmid pTRG

Equal nmols of plasmid pTRG and the PCR products were mixed together and purified using a Midi Kit (Qiagen) according to manufacturer’s specifications. 10 µL of NEB buffer 3 (New England Biolabs), 65 µL of H2O, 20 µL of the ligation mix, 2.5 µL of *NotI*, and 2.5 µL of
XhoI were added to a 1.5mL eppendorf tube. The reaction was incubated for 1 hour in a 37°C waterbath. The restriction digested ligation mix was then ligated using a Quick Ligation Kit (NEB) according to the manufacturer’s instructions, producing pTRG ligated with DNAs encoding random peptides.

**Bacterial Transformation to Prepare the Peptide Library**

One 1.5ml eppendorf tube containing 200 μL of BacterioMatch Reporter cells (relevant genotypes: *recA, F’ lacFl* (Stratagene Inc.), three additional sterile 1.5mL tubes, and tubes containing beta-mercaptoethanol, and ligated plasmid pTRG were all prechilled on ice for ten minutes. A 5mL tube of SOC media (Invitrogen) was left in a 37°C incubator to equilibrate. 50μL of competent cells were transferred into each sterile tube, leaving four tubes (including the original) with 50 μL of competent cells. 1 μL of pTRG, and 0.8 μL of beta-mercaptoethanol were added to each of the four tubes. After incubating in a wet ice bath for 30 minutes, the tubes were then subjected to a 45 second heat shock in a 42°C water bath. Immediately after the heat shock, the cells were returned to ice for 2 minutes.

1000 μL of SOC media was then added to each tube. The tubes were incubated at 37°C with gentle shaking for 2 hours. After the incubation, five LB-tetracycline plates (10 g bacto tryptone, 5 g yeast extract, 10 g NaCl, 10 μg/mL tetracycline, and 15 g bacto agar in 1 L distilled water) each had 200 μL of the transformed cells spread on them. The plates were incubated at 37°C overnight (about 16 hours).

Then number of tet’ colonies on each plate was counted (or estimated if too dense to count individually). All plates with colonies had 5mL of LB-tetracycline broth (10 g bacto tryptone, 5 g yeast extract, 10 μg/mL tetracycline, and 10 g NaCl in 1 L distilled water) added to
the plate. The plates containing broth were gently swirled around to spread the media and to dislodge some cells from each colony. The media from the plates was then transferred to a large 1 L beaker. This process was repeated until about 150,000 different colonies were collected.

Purification of Library Ligated pTRG

The plasmid DNAs from the cells in the beaker containing the colonies of the prey peptide library were collected and purified by use of a Midiprep Kit (Qiagen) according to the manufacturer’s specifications. These plasmids were sequenced by Genewiz to ensure that the peptide sequences were present in the plasmids, and that the sequences were different. The two primers (CYSLIB – Forward and CYSLIB – Reverse) were premixed with the samples, allowing the company to sequence our samples.

Preparation of Plasmid pBT

A 1.5mL eppendorf tube containing gp120 / CD4-binding sequence mimic m5 DNA (Repik, 2009) was mixed together in equal nmols with plasmid pBT (Stratagene). The ligation reaction was then digested by restriction enzymes NotI and XhoI. The restriction digested ligation mix was then ligated using an NEB Quick Ligation Kit according to the manufacturer’s instructions, producing pBT ligated with the mimic m5’s gene.

Transformation of Bacterial Reporter Strains with pBT and pTRG

One 1.5ml eppendorf tube containing 200 µL of BacterioMatch Reporter cells, three additional sterile 1.5mL tubes, and tubes containing beta-mercaptoethanol, ligated pBT, and ligated pTRG were all pre-chilled on ice for ten minutes. A 5mL tube of SOC media
(Invitrogen) was left in a 37°C incubator to equilibrate. 50 µL of competent cells were transferred into each sterile tube, leaving four tubes (including the original) with 50 µL of competent cells. 1 µL of pBT, 1 µL of pTRG, and 0.8 µL of beta-mercaptoethanol were added to each of the four tubes. After incubating in a wet ice bath for 30 minutes, the tubes were then subjected to a 45 second heat shock in a 42°C water bath. Immediately after the heat shock, the cells were returned to ice for 2 minutes.

500 µL of SOC media was then added to each tube. The tubes were incubated at 37°C with gentle shaking for 2 hours. After the incubation, each tube had 100 µL of the transformed cells spread onto each of four plates with antibiotics as follows: LB-agar plates containing 50 µg/ml kanamycin, 10 µg/ml tetracycline, 25 µg/ml chloramphenicol, 20 µM IPTG, 500-2000 µg/ml carbenicillin (each of the four plates contained a different concentration of carbenicillin: 500, 1000, 1500, or 2000 µg/ml). The plates were incubated at 37°C overnight (about 16 hours).

**Quantifying Reporter Gene Expression**

Individual colonies from the transformation plates were inoculated into 3mL of LB containing appropriate antibiotics (composition as LB-agar plates described above, except uniform 500 µg/ml carbenicillin), and were incubated overnight at 37°C. 30 µl of each overnight culture was inoculated into 3 ml of LB with the same antibiotic supplements as the overnight culture. The cultures were incubated at 37°C with aeration until cultures reached an OD$_{600}$ of 0.3 to 0.7. The cultures were placed on ice for 20 minutes, after which 1ml of each culture was transferred to a cuvet. The OD$_{600}$ of each cuvet was recorded.

200 µl of each culture was transferred to a 14mL sterile tube containing 800 µl of Z-buffer (16.1 g Na$_2$HPO$_4$·7H2O, 5.5 g NaH$_2$PO$_4$H$_2$O·0.75 g KCl, 0.246 g MgSO$_4$·7H2O, 2.7 mL
β-mercaptoethanol, and distilled water to 1 L, pH adjusted to 7.0). 30 µl of 0.1% SDS and 60 µl of chloroform were then added to each assay tube. Each tube was then vortexed for 6 seconds and placed in a 28°C water bath for ten minutes. ONPG substrate solution was also placed in the water bath for ten minutes. The first assay tube had 200 µl of 28°C 4 mg/ml ONPG added to start the reaction. The time at which the assay started was recorded. The tube was gently vortexed until it reached an OD₆₀₀ of 0.6 to 0.9, at which point 500 µl of 1M Na₂CO₃ was added to stop the reaction. The time at which the reaction was stopped was recorded. The tube was vortexed briefly and left at room temperature until the remaining reactions were completed.

After all assays were complete, 1 mL of each assay tube was transferred to a cuvet to determine the optical density at 420nm and 550nm for each assay. The readings were zeroed against LB broth.
RESULTS

The main purpose of this MQP was to construct and screen a random protein library to identify peptides that interact with the m5 mimic of HIV’s gp120 CD4-binding site.

Construction of the Prey Random Peptide Library

A random gene library of about 150,000 plasmids encoding random peptides was constructed using the protocol described in Materials and Methods. The random gene sequences made by PCR were ligated into prey plasmid pTRG after each was digested by the restriction enzymes NotI and XhoI. Figure 8 shows an electrophoresis gel comparing the PCR product/pTRG ligation reaction mixture before and after ligation. The ligation (lane 2) clearly produced DNA bands of higher sizes than the unligated reaction, indicating the ligase was active.

Competent *E. coli* cells were transformed with the ligated pTRG mixture, and plated onto LB-tetracycline plates. The tet<sup>+</sup> colonies that grew on the tetracycline plates were counted; when growth was too dense to count, the number of colonies per plate was estimated. In order to check that the colonies that grew on the plates contained plasmid pTRG with the expected source of tetracycline resistance (rather than a mutation), several colonies were randomly selected for
testing. The selected colonies were “mini-preped” to isolate their plasmids. The plasmids obtained were restriction digested with NotI and XhoI, and analyzed by electrophoresis (Figure 9). Of the five randomly chosen tet’ colonies, each contained pTRG plasmid containing a PCR insert.

**Figure 9: Restriction Digestion Screening of pTRG Prey Library Colonies.**
The leftmost lane is a PCR marker. The next five lanes represent randomly chosen tet’ colonies from the pTRG/PCR ligation, double digested with XhoI and NotI. The upper strong bands in those lanes represent plasmid pTRG, and the smaller bands represent the PCR product inserts. Note that each tet’ colony contains pTRG plasmid containing an insert.

All tet’ colonies that grew were pooled by collecting together into LB-tetracycline broth. After about 150,000 colonies had been collected into the broth, the mix was midi-preped to isolate plasmid DNA, yielding a pTRG- peptide-encoding library with 150,000 randomized prey peptides.

**Co-Transformation of Bait and Prey Plasmids**

After the purification of plasmids from the prey library, competent *E. coli* cells were double transformed with pBT bait and pTRG prey plasmids (the bait plasmid containing the CD4-binding site mimic was provided by A. Repik of our laboratory). The transformed colonies were plated onto petri dishes containing X-gal (blue color formation in the presence of the β-galactosidase reporter gene), 50 μg/ml kanamycin, 10 μg/ml tetracycline, 25 μg/ml
chloramphenicol, 20 μM IPTG, 500-2000 μg/ml carbenicillin. **Figure 10** shows randomly chosen colonies patch-plated onto the petri dishes. If both pTRG and pBT were successfully taken up, and the bait and prey proteins interact with each other, then the reporter gene β-galactosidase will be expressed. If the reporter β-gal is expressed, the X-gal on the plate will be metabolized, giving the colonies a blue color.

![Figure 10: Photographs of Colonies Containing Bait and Prey Plasmids pBT and pTRG. Shown are photographs of X-gal plates of randomly chosen colonies after dual transformation, and prior to β-galactosidase screening. While the photos are monochrome, the dark color of the colonies represents their blue color indicative of bait and prey interaction.](image)

**Quantitation of the β-Gal Reporter Activity**

The blue color of the colonies is only a *qualitative* expression of the interaction between the library peptide and the CD4bs mimic. However, the strength of the reporter activity reflects the strength of the bait and prey interaction. So to obtain a quantitative expression of the protein-protein interaction, a β-gal activity assay was conducted. The β-gal activity assay measures the metabolism of a fixed amount of ONPG β-gal substrate. The readings necessary for this assay, and the required calculations are shown in **Calculation 1**. Abs$_{420}$ represents the amount of orthonitrophenol, a product of β-gal metabolizing ONPG substrate. Abs$_{550}$ measures cellular
debris, which can interfere with the Abs$_{420}$ reading. Abs$_{600}$ measures the cell density, which can be used to standardize the readings for each specific level of bacterial growth. Time in seconds is represented by $t$, and the volume in mL is represented as $v$.

$$1000 \times \frac{(\text{Abs}_{420} - (1.75 \times \text{Abs}_{550}))}{(t \times v \times \text{Abs}_{600})}$$

**Calculation 1: Formula Used for Quantifying β-Gal Expression.** The formula above details how to compute the β-gal activity from various OD readings on a chosen *E. coli* sample. The activity is expressed as Miller units, a normalized, quantitative expression of β-gal activity (Love & Hochschild, 2004).

The bluest colonies (over 50% blue coloring) were selected for sequence analysis to remove any repeated plasmids. Then, all of their β-Gal activities were determined (Table 1). The activity data is presented graphically in Figure 11.
Table 1: Calculation of β-Galactosidase Activities. The β-gal activity assay results from the bluest, non-repeated colonies were calculated using the formula shown in calculation 1. Not shown are the time of each assay (45 seconds) and the volume (0.2 mL).
Figure 11: Summary of the $\beta$-Galactosidase Activity Assay Results. The $\beta$-gal activity assay results from Table 1 are shown graphically, with the individual sample numbers delineated on the Y-axis, and the activity assay results on the X-axis.
The frequency distribution of β-Gal reporter activities is shown in Figure 12. The six highest activity values shown on the right half of the figure (Miller values between 270-449) represent prey proteins that interact more strongly with the m5 mimic (bait).

Figure 12: Frequency Distribution of the β-Galactosidase Activity Assays from Table I. The β-gal activity assay results from Table 1 are shown as a frequency distribution. Each histobar represents the number of screened peptides having a particular Miller activity value range within a set of 20 integers.

Prey Plasmid Insert Sequencing

The plasmid DNA was isolated from each of the 53 screened colonies shown in Figure 10 (plus others representing 309 totally), and their DNA inserts were again sequenced by GENEWIZ, a commercial sequencing service. From the prey plasmid insert sequences, the primary protein structure, molecular weight and pI were calculated for the 6 prey peptides with the highest activity values (Table 2). For confidentiality reasons, the DNA sequence and primary protein structure cannot be shown.
<table>
<thead>
<tr>
<th>Peptide Number</th>
<th>Peptide Length</th>
<th>Peptide MW</th>
<th>Peptide pI</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>27</td>
<td>3004.34</td>
<td>12.6</td>
</tr>
<tr>
<td>155</td>
<td>25</td>
<td>2976.57</td>
<td>10.05</td>
</tr>
<tr>
<td>159</td>
<td>26</td>
<td>3029.44</td>
<td>6.77</td>
</tr>
<tr>
<td>185</td>
<td>26</td>
<td>3164.87</td>
<td>10.95</td>
</tr>
<tr>
<td>230</td>
<td>26</td>
<td>3030.49</td>
<td>9.5</td>
</tr>
<tr>
<td>260</td>
<td>26</td>
<td>2691.05</td>
<td>8.9</td>
</tr>
</tbody>
</table>

Table 2: Basic Peptide Information for the 6 Highest Activity Prey Peptides. Using ExPASy software tools (Swiss Institute of Bioinformatics, 2010) to analyze the six DNA sequences of the peptides in Table 1 allows the prediction of each peptide’s primary structure, length, molecular weight, and isoelectric point.
DISCUSSION

The HIV-1 glycoprotein 120 (gp120) is a protein necessary for all stages of HIV attachment to CD4 T-cells (Kasson, 1999). For this reason, gp120 was selected as the target for our attempt to identify short peptide sequences that might bind gp120’s CD4-binding site to prevent HIV attachment. If any peptide could effectively bind to gp120, this would occupy gp120 hopefully preventing it from attaching to CD4 receptors on T-cells. This could stop or slow the infection to a level that the host immune system could overcome the infection. The specific purpose of this MQP was to create a randomized library of prey proteins, and to screen it using the E. coli two-hybrid system for peptides that bind to the conserved CD4-binding sites (CD4-bs) on a gp120 mimic. This goal was accomplished, identifying 53 peptides that bind the CD4-bs (Figure 10). Six peptides interacted with the gp120 mimic at a higher level than the other peptides (Figure 11), and our lab is interested in testing these peptides for activity against blocking HIV entrance in vitro and in vivo.

It should be noted that the protein-protein interaction was measured against a gp120 mimic. A mimic is clearly not equivalent to gp120’s 3D structure in vivo, but the mimic was used as a starting point for finding a suitable peptide to block gp120’s CD4 binding site. Additional testing is underway to measure protein-protein interaction of the peptides against native gp120. Initial results show that a fraction of the peptides recognize native gp120 from heterologous HIV-1 strains in ELISAs (A. Repik, personal communication). The peptides will next be tested for their capacity to inhibit HIV-1 infection in vitro.

While the β-gal activity assay quantifies the strength of each prey peptide’s interaction with the gp120 mimic, it cannot specify how and where each peptide interacts. Weaker signals
found in this monomeric mimic assay may prove to bind much stronger to several different clades of HIV when tested in vivo. Based on the findings and reasoning presented in this report, testing protein-protein interaction of these peptides against HIV-1 envelope trimers, and testing their capacity to inhibit infection of diverse HIV-1 strains is required. If peptides that bind the conserved CD4bs on the native trimer and that also block infection are identified, then they will have potential for use as microbicides and other therapies.
BIBLIOGRAPHY


Repik, A. (2009, December 2). The application of a bacterial two hybrid system for identifying strong peptide binders against the artificial CD4 binding site mimic proteins of gp120 . . Worcester, MA, USA.


