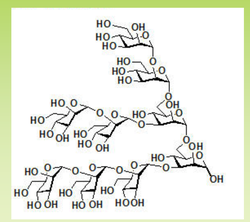
Discovery of Small-Molecule Inhibitors of HIV-1 gp120 Using Small-Molecule Microarrays

Brian, Olivia McPherson, and Angela Koehler, PhD

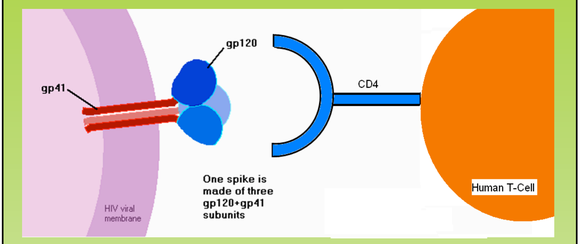
Broad Institute of MIT and Harvard, Cambridge, MA, USA



Introduction

More than 40 million people are currently infected with HIV/AIDS and 7.7% of them die annually of the disease.[1] Currently, in an attempt to combat HIV, there are twelve commonly prescribed HIV drugs, but these drugs do not eliminate viral infection; they are merely a means of preventing viral replication. None of these drugs target gp120, a important protein found on the surface of the HIV virus.

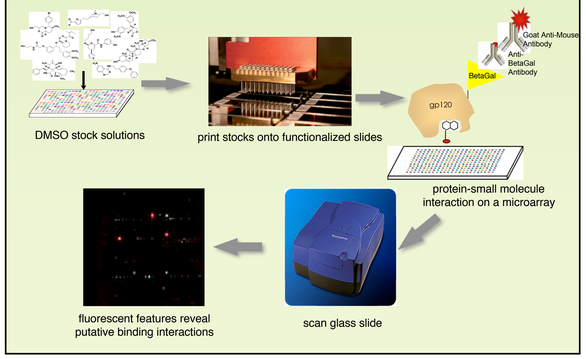
Overview of gp120



Adapted From: Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson

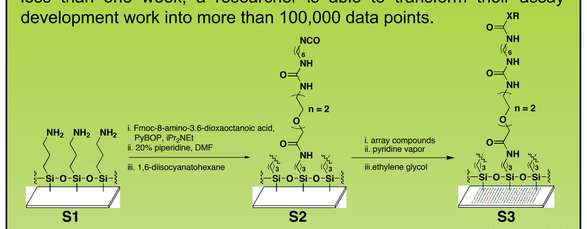
The role of gp120 in HIV is to bind to the CD4 receptor on T-cells, which are the cells in the human body that fight disease and infection. Structurally, one half of the molecular weight of gp120 is due to the carbohydrate side chains (the "glyco-" in "glycoprotein"). These are sugar residues, which form a sugar "dome" over the gp120 spikes, preventing gp120 from being recognized by the human immune response. As the HIV virus and the human $CD4^+$ T-cell come together, the gp120 binding site "snaps open" at the last minute.[2] If a drug could inhibit or block gp120's active site, it would prevent the protein from attaching to $CD4^+$ T-cells, preventing the entry of HIV. No drug has been created yet that will prevent the protein from binding to human T-cells. Although the gp120 protein has proven costly and difficult to inhibit, if inhibited, the benefits would be priceless.

Overview of Small-Molecule Microarrays



Materials and Methods

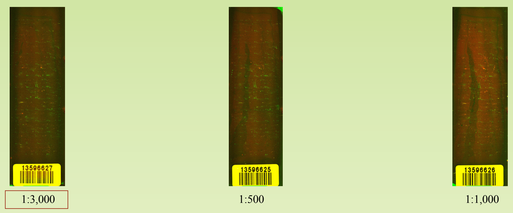
Small Molecule Microarrays, (SMMs), are a way to test over ten thousand different compounds for binding against a particular protein at one time. Using diversity-based (DIV) and natural products and commercials-based (NPC) libraries of compounds, 10,800 compounds are covalently bound to glass slides using isocyanate capture chemistry. This process is not only simple, used by everyone from high school interns to post-doctoral researchers, it is also cost effective. For around $15, a screener is able to test 10,800 compounds. Finally, in less than one week, a researcher is able to transform their assay development work into more than 100,000 data points.



Chem. Biol. **13**, 493-504, 2006
Nature Protocols **1**, 2344-2352, 2006

Assay Development

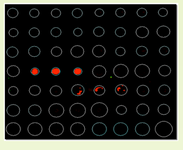
Many factors can affect the quality of a screen, and must be controlled. When developing a screen for a particular protein, one must consider concentrations of protein and antibody, buffer concentration, incubation time, and wash cycles. All of these factors affect to what is known as "signal to noise ratio." Taking all of these factors into consideration, we began at a protein/antibody concentration of 1ug/mL, or 1:1,000. From there, we increased or decreased the concentration based on how reactive the protein or antibody is and how it would affect the background of the slide.



Protein or Antibody Tested	Concentration 1	Concentration 2	Concentration 3
Goat Anti-Mouse Antibody	1 To 500	1 to 1,000	1 To 3,000
Anti-BetaGal Mouse Antibody	1 To 1,000	1 To 3,000	1 To 5,000
Beta Galactosidase	1 To 500	1 To 1,000	1 To 3,000
HIV-1 gp120	1 To 500	1 to 1,000	1 To 2,000

Scanning and Data Analysis

Using an Axon 4000B scanner, slides were scanned and then analyzed through GenePix Pro 6.0 software. Slides were "spot-fitted" using GenePix Array Lists (GALs) and manually moved to fit individual spots of interest based upon their fluorescence intensities. After slides were analyzed manually, software created analyses of each individual spot. This information was then sent through the ChemBank Data Pipeline, where it was further analyzed through statistical methods.

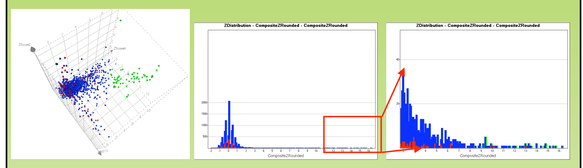


http://chembank.broad.harvard.edu

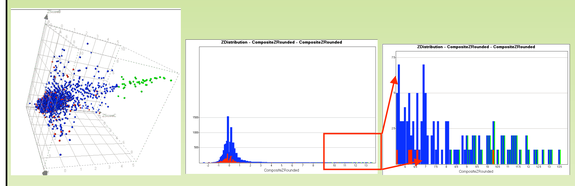
Results

Using a program called SpotFire, we analyzed screening data from Goat Anti-Mouse Antibody, Anti-BetaGal Antibody, and gp120. Using a template for the SMM format, three-dimensional graphs of z score intersection were generated. A z score is a number that takes into account the replicability of data points, and their power relative to negative controls. From those graphs, fifty compounds were selected due to their position as outliers, or data points farthest away from the general concentration of points. We also generated two-dimensional graphs, which express the same data, only as another representation. From that two-dimensional graph, the tail end, or area from which we selected our fifty data points, was enlarged to illustrate where data was selected from.

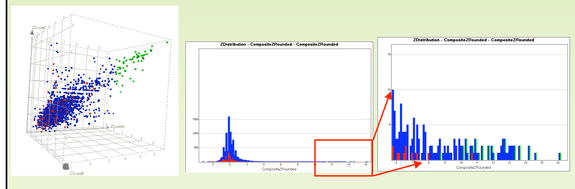
gp120



Anti-BetaGal Antibody



Goat Anti-Mouse Antibody

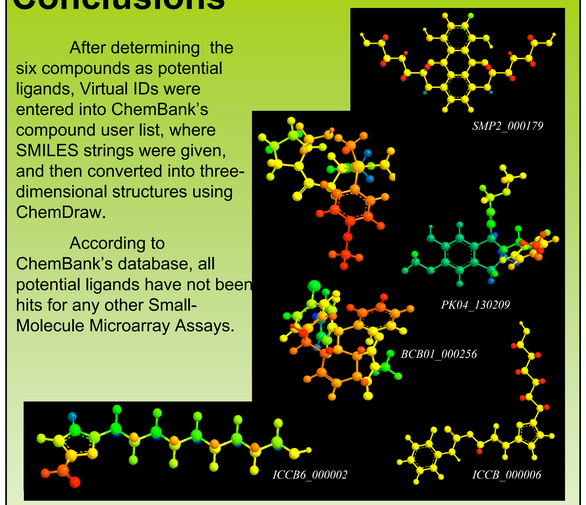


After the fifty compounds were mathematically selected from each assay, it was up to us to investigate the specificity of each molecule. Compounds that appeared as "hits" in more than one of the proteins' assays were eliminated because they were non-specific. Eleven compounds were left as "hits" for the gp120 assay; that list was then compared to a database of promiscuous binders across a large number of proteins. Finally, eight compounds remained as potential binders to gp120. Of those eight, plate and well data were unavailable for two, leaving the final potential ligand count at six.

Conclusions

After determining the six compounds as potential ligands, Virtual IDs were entered into ChemBank's compound user list, where SMILES strings were given, and then converted into three-dimensional structures using ChemDraw.

According to ChemBank's database, all potential ligands have not been hits for any other Small-Molecule Microarray Assays.



Six potential ligands were discovered due to their putative binding properties, which can later be studied more in depth to determine if they are in fact small-molecule inhibitors of gp120. From a cursory overview of the potential ligands, there does not appear to be skeletal similarities among the molecules, but there may be biological effects that they share, which have not yet been determined.

We were unable to analyze data from BetaGal because we discovered the pH of the PBS buffer was too acidic and the protein was compromised. In the future, I would like to run assays on BetaGal and compare the potential ligands of gp120 to any binders to BetaGal. I would also like to run tests using Biacore to determine Surface Plasmon Resonance in order to determine the affinity of the gp120 protein for each potential ligand.

Literature cited

1. http://globalhealth.org/view_top.php3?id=227
2. *Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody*, Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. **NATURE** 393 (6686): 648-659 JUN 18 1998
3. *Chem. Biol.* **13**, 493-504, 2006
4. *Nature Protocols* **1**, 2344-2352, 2006

Acknowledgments

I would like to thank Olivia McPherson and Angela Koehler for being my mentors over the past six weeks and helping me through any problems I encountered.

I would also like to thank Jason Fuller and Karen Rose for analyzing data.

I would also like to thank Megan Rokop, Julie Boehm, and Kate MacSwain for giving me this opportunity.

BROAD INSTITUTE