High-Throughput Epitope Determination of HIV-1 Polyclonal B Cell Repertoires

By

Rohit Venkat

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Chemical and Physical Biology

December 12, 2020

Nashville, Tennessee

Approved:

Ivelin Georgiev, Ph.D.
Cinque Soto, Ph.D.

CHAPTER 1

INTRODUCTION

Human immunodeficiency virus 1 (HIV-1) has posed a significant public health burden worldwide since its discovery in 1981. More than 75 million people have been infected with HIV-1 since the start of the epidemic (Deeks et al., 2015). As of 2019, around 38 million people were currently living with HIV-1 (Global HIV & AIDS statistics — 2020 fact sheet). Although effective prevention and treatment strategies have greatly reduced transmission and mortality rates, HIV-1 continues to infect close to two million people each year, and one-third of people living with HIV-1 lack basic access to antiretroviral therapy (Global HIV & AIDS statistics — 2020 fact sheet). To reverse the course of this global public health threat, the development of an HIV-1 vaccine is a top priority.

Due to the high mutability and tremendous levels of genetic diversity exhibited by HIV-1, traditional vaccine approaches have been unsuccessful to date. In recent years, HIV-1 vaccine development efforts have focused on understanding the interactions between B cell-encoded antibodies and the HIV-1 Env glycoprotein, the sole target of the HIV-1 antibody response, to rationally inform vaccine design (Burton et al., 2012; Kwong et al., 2013). This “reverse vaccinology” strategy involves isolating neutralizing antibodies from HIV-1 infected donors to be used as vaccine templates, studying the structural interactions and evolution between antibody and virus, and using this information to design immunogens that guide naïve B cell receptor (BCR) development toward a neutralizing antibody response (Rappuoli et al., 2016).

Broadly neutralizing antibodies (bnAbs) capable of neutralizing diverse strains of HIV-1 are ideally suited as templates for HIV-1 vaccine design because studies have

shown their efficacy in controlling viremic loads in infected individuals (Mendoza et al., 2018; Bar-On et al., 2018) and in preventing infection in animal models (Moldt et al., 2012). bnAbs can be classified based on the region of the HIV-1 Env trimer that they recognize, which generally fall into six conserved epitope categories: V1V2, V3-glycan, CD4-binding site (CD4bs), silent face center, fusion peptide, and subunit interface (Kwong and Mascola, 2018).

Since the conserved epitopes targeted by bnAbs represent distinct sites of vulnerability of the HIV-1 Env trimer's glycan shield, the characterization of antibodies targeting these epitopes is important for identifying potential new vaccine templates. Unfortunately, the BCR specificities in an individual are polyclonal and difficult to characterize using currently available methods. While BCR sequencing of the polyclonal B cell response in an individual allows the generation of thousands to millions of antibody sequences (DeKosky et al., 2013), epitope mapping of these antibodies requires considerable effort. Standard techniques to resolve the epitope specificities present in polyclonal B cell responses typically involve the synthesis of monoclonal antibodies for individual characterization by biolayer interferometry or electron microscopy (Gershoni et al., 2007; Guo et al., 2018; Bianchi et al., 2018). This bottleneck for antibody epitope determination makes it difficult to examine more than a small cross-section of the total B cell repertoire in a given individual.

Recently, our research group developed a technology termed LIBRA-seq (linking B cell receptor to antigen specificity through sequencing) that enables high-throughput mapping of B cell receptor sequence to antigen specificity for thousands of individual B cells in a polyclonal sample (Setliff et al., 2019). In this assay, B cells are mixed with a

panel of DNA-barcoded antigens so that both the BCR sequence and antigen barcodes are recovered via single-cell next-generation sequencing. For a given B cell, this allows the simultaneous recovery of paired heavy- and light-chain BCR sequences and antigen specificity for all bound antigens. Here, by including antigen variants with epitope-knockout mutations in the LIBRA-seq screening library, we look to extend the LIBRA-seq technology for the application of high-throughput epitope mapping.

CHAPTER 2

HIGH-THROUGHPUT EPITOPE DETERMINATION USING LIBRA-SEQ

2.1 Rationale

Current antibody epitope mapping technologies preclude high-throughput characterization of individual antibody specificities within a polyclonal B cell sample. Here, we sought to leverage LIBRA-seq to map the residue-level epitope specificities of individual B cells in a broadly neutralizing HIV-infected donor sample. The development of a LIBRA-seq based assay for antibody epitope determination would provide the ability to simultaneously obtain functional information about antibody epitope specificity together with paired heavy- and light-chain B cell receptor sequence for thousands of individual B cells in a polyclonal sample. No other experimental technique currently exists which allows high-throughput antibody epitope determination measurement for polyclonal B cell samples. Furthermore, applying high-throughput antibody epitope determination to HIV-1 infection samples has the potential to uncover novel epitopes, therapeutic antibodies, and vaccine templates.

2.2 Results

To test the feasibility of residue-level epitope mapping by LIBRA-seq, we performed a pilot experiment using a Ramos B cell line expressing on its surface VRC01, a CD4-binding site (CD4bs)-directed HIV-1 broadly neutralizing antibody. The antigen screening library consisted of seven pairs of Env SOSIP epitope-specific variants from two different backbones (BG505, CZA97) and two negative control antigens (Influenza HA NC99, hepatitis C virus JFH-1 E2). The seven Env SOSIP variants included one V3-glycan directed epitope-knockout mutant (N332T), two V1V2 directed epitope-knockout mutants (N160K, K169E), three CD4 binding site (CD4bs) directed epitope-knockout mutants (D279K, D368R, DKO), and the wild-type protein (N332).

Because VRC01 targets the conserved CD4bs epitope, we expected the VRC01 B cell population for this LIBRA-seq experiment to demonstrate significantly reduced sensitivity to the D279K, D368R, and DKO mutants compared to the wildtype and other epitope-knockout variants. Following single cell sequencing and quality control filtering (VRC01 CDRH3 sequence only), we recovered 3632 VRC01 cells with antigen specificity information. The negative control antigen HA NC99 appeared “sticky” due to uncharacteristically high read counts for VRC01 cells and was therefore discarded from further analysis. Examination of the LIBRA-seq score distributions for the seven Env SOSIP variants across two backbones revealed lower scores for the D279K, D368R, and DKO mutants when compared to the wildtype and other epitope-knockout variants, in agreement with the known reduced VRC01 affinity for D279K and D368R (**Figure 1**). When comparing LIBRA-seq score distributions across backbones, the CZA97 Env SOSIP variant set showed slightly elevated levels of noise as reflected by the number of

VRC01 cells with negative LIBRA-seq scores for the wildtype and non-CD4bs directed variants. Overall, these data suggested that LIBRA-seq could successfully be employed to extract residue-level epitope information for thousands of individual B cells in parallel.

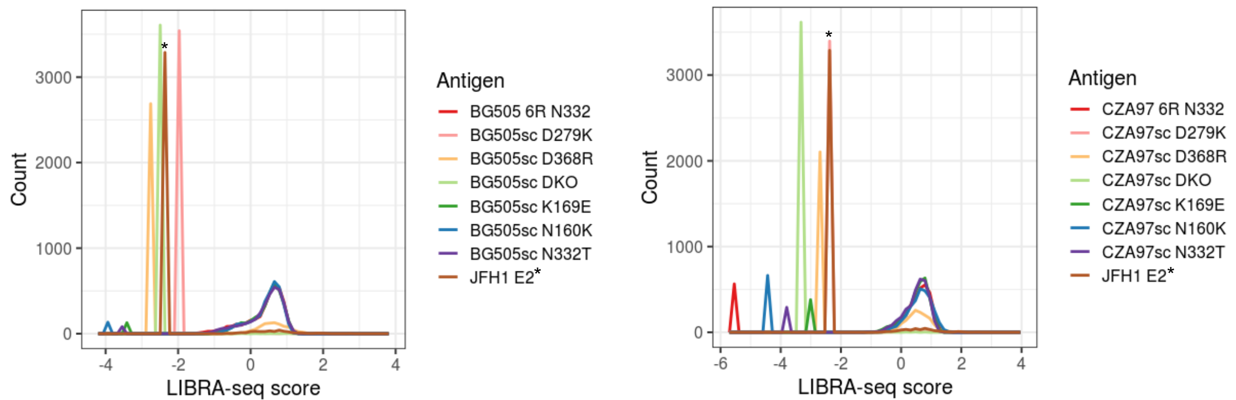

Figure 1. LIBRA-seq score distributions for BG505 (left) and CZA97 (right) Env SOSIP variants for VRC01 B cell population.

Next, we looked to map the residue-level epitope specificities of individual B cells in a broadly neutralizing HIV-infected donor sample. An HIV-infected donor sample from the NIAID cohort 3+ years post-infection (chronic) timepoint, donor N55, was selected for analysis based on its broad neutralization breath. The antigen screening library for this LIBRA-seq experiment was the same as the previous experiment. The negative control antigen HA NC99 was also discarded from this analysis due to quality control concerns. Following single cell sequencing and quality control filtering (duplicate barcode removal), we recovered 2018 B cells (or uniquely paired heavy-light chain sequences) with antigen specificity information, of which 1884 were IgG-specific. These IgG-specific B cells were used as the starting point for the remaining analysis.

Examination of the LIBRA-seq score distributions for the seven Env SOSIP variants across two backbones revealed some expected variability in LIBRA-seq scores for this polyclonal sample (**Figure 2**). However, general patterns with consistent LIBRA-seq score magnitudes or distributions across backbones or between variants targeting the same epitope were not observed. Complete linkage hierarchical clustering of the B cells by antigen LIBRA-seq scores (**Figure 3**) also demonstrated some inconsistent trends, such as (1) low LIBRA-seq scores for the CZA97 N332 variant but not other variants with the same mutation background and (2) lack of consensus in LIBRA-seq scores across backbones or between variants targeting the same epitope.

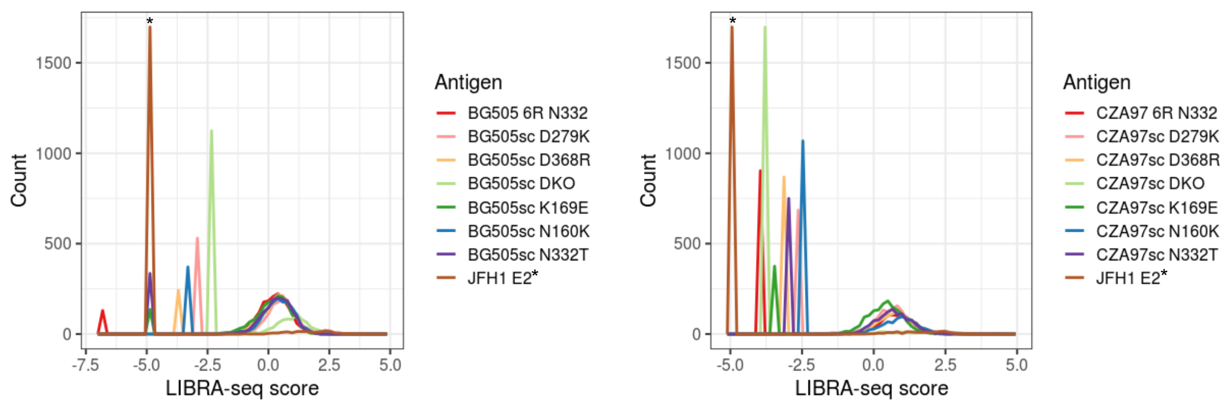

Figure 2. LIBRA-seq score distributions for BG505 (left) and CZA97 (right) Env SOSIP variants for donor N55 B cell population.

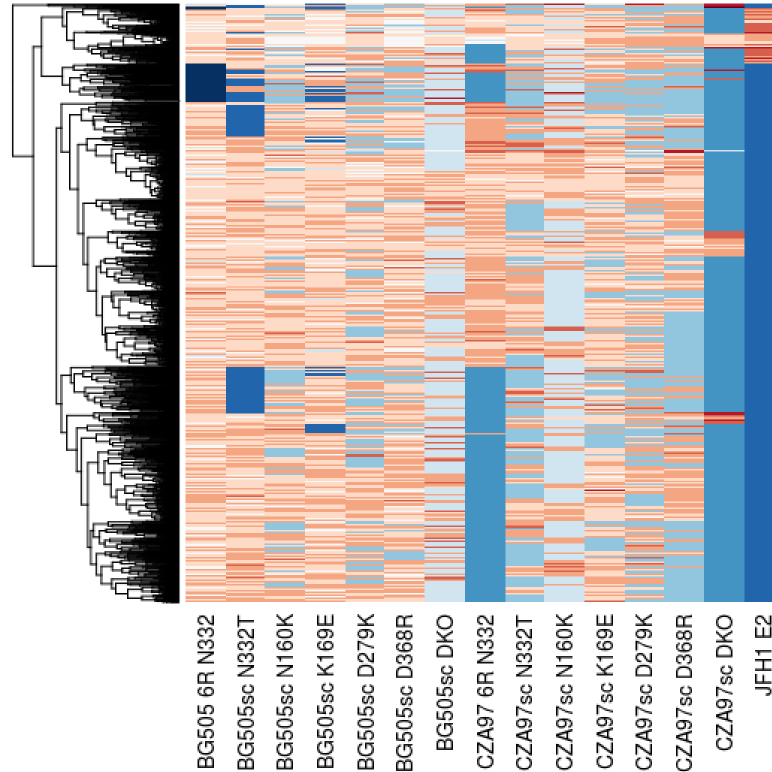

Figure 3. Complete linkage hierarchical clustering of LIBRA-seq scores for donor N55 B cell population. High and low LIBRA-seq scores are indicated by red and blue, respectively.

We next examined our ability to identify individual epitope-specific B cells based on the epitope-knockout variants included in the antigen panel. Due to difficulties in interpreting LIBRA-seq scores across backbones, we analyzed BG505 and CZA97 variant sets separately. We defined a B cell to be epitope-specific if one or more Env SOSIP variants targeting a particular epitope had a LIBRA-seq score less than −0.4 while all variants targeting other epitopes had LIBRA-seq scores above 0. Both of these LIBRA-seq score cutoffs represented arbitrary cutoffs. For the BG505 Env SOSIP variant set, we identified 28 V3-glycan directed B cells, 26 V1V2 directed B cells, 168 CD4bs directed B cells. For the CZA97 Env SOSIP variant set, we identified 4 V3-glycan directed B cells, 18 V1V2 directed B cells, and 136 CD4bs directed B cells. The number of B cells

overlapping between the two sets was 14. **Table 1** shows representative epitope-specific B cells that were identified with reasonable confidence.

Epitope Category	BG505 6R N332	BG505sc N332T	BG505sc N160K	BG505sc K169E	BG505sc D279K	BG505sc D368R	BG505sc DKO	CZA97 6R N332	CZA97sc N332T	CZA97sc N160K	CZA97sc K169E	CZA97sc D279K	CZA97sc D368R	CZA97sc DKO
V3-glycan	1.57	-5.63	1.70	0.54	1.31	0.56	1.42	1.70	-3.58	1.89	2.11	0.91	1.87	-0.39
CD4bs	0.99	1.21	1.28	-0.11	-2.00	-0.37	-0.73	2.06	2.58	1.73	1.64	-0.74	3.41	-1.96
CD4bs	1.05	1.12	0.98	-0.65	-1.23	-0.48	-0.49	1.75	2.28	1.66	1.11	-0.42	3.34	-1.25
V1V2	3.76	0.11	3.29	-1.26	2.55	2.49	4.60	-0.90	-1.79	0.27	-4.53	-1.99	-1.78	-0.44

Table 1. Representative epitope-specific B cells within donor N55 B cell population. High and low LIBRA-seq scores (relative to each antigen's distribution of LIBRA-seq scores) are indicated by red and blue, respectively.

2.3 Discussion

By utilizing epitope-specific variants in the aforementioned LIBRA-seq experiments, we were able to survey both the potential and limitations of high-throughput epitope mapping using LIBRA-seq. In the VRC01 Ramos B cell line experiment, the clear, unambiguous differences in LIBRA-seq scores across backbones and between epitope-knockout variants based on the known specificity of VRC01 demonstrated that LIBRA-seq can be used to query epitope specificity for thousands of individual B cells in parallel. By using epitope-specific variants from two different backbones in our antigen screening library, we were also able to demonstrate the ability to assess both neutralization breath and epitope specificity in parallel for an HIV-1 broadly neutralizing B cell line. While the CZA97 Env SOSIP variant set did show higher levels of noise compared to the BG505 Env SOSIP variant set, this may reflect differences in VRC01's affinity to BG505 and to CZA97. In order to reduce noise and more clearly distinguish epitope specificity based on

LIBRA-seq scores in future experiments, it may be useful to calculate LIBRA-seq scores separately for different backbones.

In the donor N55 LIBRA-seq experiment, the ability to assess epitope specificity based on LIBRA-seq scores was far more difficult. The observed inconsistencies in LIBRA-seq scores across backbones and between variants targeting the same epitope may have been due to one or more of the following reasons: (1) B cells in the sampled population were broadly reactive to varying degrees or not at all; (2) B cells in the sampled population did not target one of the epitopes queried by the antigen screening library; (3) epitope knockout mutations may abrogate binding to varying degrees based on an antibody's angle of approach (Zhou et al., 2016); (4) differences in antigen quality may have resulted in variations in antigenicity. Taken together, the difficulties we experienced in interpreting LIBRA-seq scores for this experiment illustrate the vast heterogeneity that exists within a polyclonal B cell sample and the complexities of performing high-throughput epitope determination for polyclonal samples.

2.4 Methods

2.4.1 Antigen production and DNA barcoding

Recombinant AviTag-labeled antigens were expressed in 293F mammalian cells. The proteins were purified over a lectin affinity chromatography column and then size-separated on an AKTA fast protein liquid chromatography (FPLC) system. Fractions corresponding to correctly folded antigen were analyzed by SDS-PAGE. Antigen binding was assessed by ELISA using a panel of monoclonal antibodies with known binding properties to ensure proper antigenic and structural properties. After purifying and testing

the antigens, each antigen was biotinylated using the BirA Biotin-Protein Ligase Kit. A unique DNA barcode was then directly conjugated to each biotinylated antigen using the Solulink Protein-Oligonucleotide Conjugation Kit. FPLC was used to remove excess oligonucleotide from the protein-oligo conjugates, and the concentrations of the antigen-oligo conjugates were determined by a BCA assay.

2.4.3 Enrichment of antigen-specific B cells

For a given sample, cells were stained with the DNA-barcoded biotinylated antigens, then with streptavidin labeled with the fluorophore phycoerythrin (PE), and then finally with a panel of antibodies (CD3-APCCy7, IgG-FITC, CD19-BV711, CD14-V500) for detecting IgG-positive memory B cells and the fluorescently labeled antigen-oligo conjugates. Cells were bulk sorted by fluorescence-activated cell sorting (FACS) at the Vanderbilt Flow Cytometry Shared Resource (FCSR) core. Antigen-positive cells were delivered to the Vanderbilt Technology for Advanced Genomics (VANTAGE) sequencing core at an appropriate target concentration for 10x Genomics library preparation and next-generation sequencing (NGS).

2.4.4 Computational analysis

BCR and antigen barcode sequence reads were processed using a custom computational pipeline developed in-house. Briefly, paired-end FASTQs files of the BCR and antigen barcode libraries were processed using the 10x Genomics software Cell Ranger (v2.2.0 for VRC01 Ramos B cell line experiment; v4.0.0 for donor N55 experiment). For BCR libraries, Cell Ranger aligned sequence reads against the GRCh38

VDJ reference to generate heavy and light chain BCR sequence contigs for each cell barcode. These BCR sequence contigs were then submitted to IMGT HighV-Quest (Alamyar et al., 2012) for germline sequence annotation. For antigen barcode libraries, Cell Ranger generated a cell barcode–UMI count per antigen matrix. Antigen UMI counts were then normalized to obtain LIBRA-seq scores (Setliff et al., 2019). After merging BCR sequence information and LIBRA-seq scores by cell barcode, a table was generated containing information on a cell-by-cell basis, including LIBRA-seq scores, heavy and light chain sequences, clonotype family, and annotated sequence features such as V gene, J gene, and CDR3 sequence.

CHAPTER 3

CONCLUSIONS AND FUTURE DIRECTIONS

Despite the aforementioned limitations of our work, the identification of a number of epitope-specific B cells in the donor N55 experiment gives confidence that adjustments in assay design and quality control procedures can yield dramatic improvements in high-throughput epitope determination for HIV-1 polyclonal samples in future experiments. Some adjustments that would lead to better interpretability of results include: (1) using diverse, non-epitope-specific backbones to assess neutralization breath; (2) using epitope-specific variants from a single backbone to assess epitope specificity; (3) calculating LIBRA-seq scores separately for antigens that query neutralization breath and epitope specificity; (4) incorporating at least two or more epitope knockout mutations for each epitope-specific variant; (5) ensuring broad and single representation of epitope-

specific variants in the antigen screening library; (6) using several negative control antigens. Overall, the incorporation of these changes has the potential to significantly improve the interpretability of LIBRA-seq scores for epitope determination and ensure the success of future experiments.

REFERENCES

Alamyar E, Duroux P, Lefranc M-P, Giudicelli V. IMGT(®) tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. *Methods Mol Biol*. 2012;882(D):569-604.

Bar-On Y, Gruell H, Schoofs T, et al. Safety and antiviral activity of combination HIV-1 broadly neutralizing antibodies in viremic individuals. *Nat Med*. 2018;24(11):1701-1707.

Bianchi M, Turner HL, Nogal B, et al. Electron-Microscopy-Based Epitope Mapping Defines Specificities of Polyclonal Antibodies Elicited during HIV-1 BG505 Envelope Trimer Immunization. *Immunity*. 2018;49(2):288-300.e8.

Burton DR, Ahmed R, Barouch DH, et al. A Blueprint for HIV Vaccine Discovery. *Cell Host Microbe*. 2012;12(4):396-407.

Chuang G-Y, Zhou J, Acharya P, et al. Structural Survey of Broadly Neutralizing Antibodies Targeting the HIV-1 Env Trimer Delineates Epitope Categories and Characteristics of Recognition. *Structure*. 2019;27(1):196-206.e6.

DeKosky BJ, Ippolito GC, Deschner RP, et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotechnol*. 2013;31(2):166-169.

Gershoni JM, Roitburd-Berman A, Siman-Tov DD, Tarnovitski Freund N, Weiss Y. Epitope mapping: the first step in developing epitope-based vaccines. *BioDrugs*. 2007;21(3):145-156.

Global HIV & AIDS statistics — 2020 fact sheet. Available at: http://www.unaids.org/en/resources/fact-sheet.

Guo Z, Wilson JR, York IA, Stevens J. Biosensor-based epitope mapping of antibodies targeting the hemagglutinin and neuraminidase of influenza A virus. *J Immunol Methods*. 2018;461(July):23-29.

Kwong PD, Mascola JR, Nabel GJ. Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nat Rev Immunol*. 2013;13(9):693-701.

Kwong PD, Mascola JR. HIV-1 Vaccines Based on Antibody Identification, B Cell Ontogeny, and Epitope Structure. *Immunity*. 2018;48(5):855-871.

Mendoza P, Gruell H, Nogueira L, et al. Combination therapy with anti-HIV-1 antibodies maintains viral suppression. *Nature*. 2018;561(7724):479-484.

Moldt B, Rakasz EG, Schultz N, et al. Highly potent HIV-specific antibody neutralization in vitro translates into effective protection against mucosal SHIV challenge in vivo. *Proc Natl Acad Sci U S A*. 2012;109(46):18921-18925.

Rappuoli R, Bottomley MJ, D'Oro U, Finco O, De Gregorio E. Reverse vaccinology 2.0: Human immunology instructs vaccine antigen design. *J Exp Med*. 2016;213(4):469-481.

Setliff I, Shiakolas AR, Pilewski KA, et al. High-Throughput Mapping of B Cell Receptor Sequences to Antigen Specificity. *Cell*. 2019;179(7):1636-1646.e15.

Zhou T, Lynch RM, Chen L, et al. Structural Repertoire of HIV-1-Neutralizing Antibodies Targeting the CD4 Supersite in 14 Donors. *Cell*. 2015;161(6):1280-1292.