Published in final edited form as:

Structure. 2020 October 06; 28(10): 1124–1130.e5. doi:10.1016/j.str.2020.07.012.

Identification of structurally related antibodies in antibody sequence databases using Rosetta-derived position-specific scoring

Jessica A. Finn^a, Jinhui Dong^b, Alexander M. Sevy^b, Erica Parrish^b, Iuliia Gilchuk^b, Rachel Nargi^b, Morgan Scarlett-Jones^b, Walter Reichard^b, Robin Bombardi^b, Thomas G. Voss^{b,c}, Jens Meiler^d, James E. Crowe Jr.^{a,b,c,*}

^aDepartment of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN 37232 USA

^bDepartment of Vanderbilt Vaccine Center, Vanderbilt University Medical Center, Nashville, TN 37232 USA

^cDepartment of Pediatrics and Vanderbilt University Medical Center, Nashville, TN 37232 USA

^dDepartment of Chemistry and Center for Structural Biology, Vanderbilt University, Nashville, TN 37235 USA.

SUMMARY

The amount of antibody (Ab) variable gene sequence information is expanding rapidly, but our ability to predict the function of Abs from sequence alone is limited. Here, we describe a sequence-to-function prediction method that couples structural data for a single Ab/Antigen (Ag) complex with repertoire data. We used a position-specific structure-scoring matrix (P3SM) incorporating structure-prediction scores from Rosetta to identify Ab variable loops that have predicted structural similarity to the influenza virus-specific human Ab CH65. The P3SM approach identified new members of this Ab class. Recombinant Ab expression, crystallography, and virus inhibition assays showed that the HCDR3 loops of the newly identified Abs possessed similar structure and antiviral activity as the comparator CH65. This approach enables discovery

Address correspondence to: James E. Crowe, Jr., M.D., Vanderbilt Vaccine Center, 11475 MRB IV, 2213 Garland Avenue, Vanderbilt University Medical Center, Nashville, TN 37232 USA. james.crowe@vumc.org, (615) 343-8064, or Jens Meiler, Ph.D., Department of Chemistry, Stevenson Center, Station B 351822, Room 7330, Nashville, TN 37235, USA jens.meiler@vanderbilt.edu.

***LEAD CONTACT:** James Crowe, james.crowe@vumc.org

AUTHOR CONTRIBUTIONS

J.A.F., J.M. and J.E.C. conceived and designed the experiments. J.A.F., E.P., I.G., A.M.S and T.G.V. performed the experiments. J.A.F., J.M. and J.E.C. performed data analysis. J.D. performed x-ray crystallography. R.N., M.S.-J., W.R. and R.B. prepared materials for the experiments. J.A.F. and J.E.C. wrote the manuscript with input from all other authors.

DECLARATION OF INTERESTS

J.E.C. has served as a consultant for GSK Vaccines, Lilly, Sanofi and Sanofi Pasteur, Pfizer, and Novavax, is on the Scientific Advisory Boards of CompuVax and Meissa Vaccines, and is Founder of IDBiologics, Inc. All other authors declare there are no conflicts.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

of new human Abs with desired structure and function using cDNA repertoires that are obtained readily with current amplicon sequencing techniques.

eTOC Blurb

Finn *et al.* describe a sequence-to-function prediction method that couples structural data for a single antibody-antigen complex with large-scale antibody variable gene repertoire data. Using a position-specific structure-scoring matrix incorporating structure-prediction scores from Rosetta, they identified Ab variable loops that have predicted structural similarity to an influenza virus-specific human antibody.

INTRODUCTION

Functional annotation of the emerging plethora of B and T cell immunome data will require development of new methods for predicting protein function from sequence. In some cases, when germline gene segment sequences encode a particular amino acid motif that binds an Ag in a canonical way, Ab specificity can be inferred because of germline gene usage. Examples include V_H1-69 -encoded influenza hemagglutinin stem Abs, V_H1-02 -encoded HIV-1 CD4 binding site Abs, or V_H4 -encoded autoreactive Abs that bind to poly(lactosamine) (Navis et al., 2014; Sui et al., 2009; Thompson et al., 1991). However, methods that predict the functional properties of encoded amino sequences without regard to the V_H gene segment origin or evolutionary gene history of Abs are lacking. Ab/Ag interactions are mediated by shape complementarity and biochemical properties of side chains in the interface (Kuroda and Gray, 2016). It is expected that diverse antibody genes and more than one Ab amino acid sequence can achieve a given structure needed to interact with an Ag, but currently there are no efficient methods for predicting members of such structural classes based on Ab gene sequence alone. We recently developed a structure-based Ab discovery method that uses a P3SM to predict structural homologs rapidly from Ab gene sequences (Willis et al., 2016).

Here we used Rosetta homology modeling and multiple linear regression to predict the stability and interaction energy of Ab heavy chain complementarity determining region 3 (HCDR3) loops with a target Ag. We interrogated a large repository of human Ab variable gene sequences from healthy individuals to identify Abs with similar specificity and function to the influenza hemagglutinin (HA)-specific human Ab CH65 (Whittle et al., 2011). We performed conventional sequence-based homology searches of the Ab gene repertoires and failed to identify Ab clones that bound to influenza HA. The central dogma of structural biology states that sequence determines structure determines function. Thus, we hypothesized that an Ab search strategy that predicts structural similarity of Ab compared to a pure sequence-based search. In effect, such a structure-based search weights each sequence position and the type of the mutation based on the predicted consequences on structure and binding. In contrast, in a sequence-based comparison all sequence positions are weighted equally, and the mutation is evaluated by evolutionary likelihood. To test our hypothesis, using only Ab variable region cDNA sequences, we developed and used a method to make predictions of Ab protein structures to identify a class of anti-influenza Abs with members that 1) shared structural features with the comparator Ab, 2) bound to the same epitope on

influenza HA, and 3) mediated potent virus inhibition. By integrating an efficient prediction method for the structure and function of Abs, we accelerated discovery of new members of this Ab structural class. This sophisticated yet efficient method of predicting structural and functional networks of Abs from sequences allows targeted discovery of new antiviral Abs, facilitates improved understanding of the diversification of function in Ab structural families, and reveals the potential for public structural solutions (i.e., inter-donor structural convergence) for target-specific Abs.

RESULTS

We sought to identify a structural class of Abs, starting with a representative Ab having an available co-crystal structure with its target Ag and known virus neutralizing function as a target for our structure-based search. The influenza HA-specific human Ab CH65 binds to many influenza virus type A subtype H1 HA proteins, and this Ab was crystallized in complex with the H1 HA (PDB ID: 5UGY) from A/Solomon Islands/3/2006 (designated here as SI06) (Whittle et al., 2011). A genetically related Ab with a differing junctional sequence, designated CH67 (PDB ID: 4HKX), also was described (Schmidt et al., 2013). These antibodies possess interesting structural features that determine their function, namely a dipeptide motif on the tip of HCDR3 that interacts with the HA protein in a way that directly mimics the atomic features of the HA/sialic acid interaction. In addition, the CH65 and CH67 HCDR3 hypervariable loops exhibit a high level of structural stability, and preconfigure the paratope to bind HA, reducing entropic cost for an optimal interaction (Xu et al., 2015).

We used an existing database of Ab variable domain sequences from over 100 subjects in the Vanderbilt Vaccine Center Biorepository that were obtained by next generation amplicon sequencing of peripheral blood B cells from human subjects using isotype-independent primers that target the heavy chain Fv domain. The database is dynamic, with nearly daily sequence deposits. At the time of the search, the database contained hundreds of millions of antibody gene sequences, from over 100 adult individuals. Human subjects selected from the database were expected to have prior exposure to influenza HA Ag either through vaccination or natural infection, however subjects were not chosen specifically for recent vaccination or infection status. The duration since influenza HA Ag exposure was unknown. The Abs CH65 and CH67 are encoded by the Ab germline V_H and J_H genes *IGHV1-2* and *IGHJ6* and they possess a HCDR3 length of 19 amino acids. Our Ab variable gene database contained ~67,000 unique junctional sequences using *IGHV1-2/IGHJ6* and HCDR3 length of 19. We did not identify any HCDR3 sequences in our database matching the CH65 or CH67 sequences. The nearest two sequences identified had three mutations each by substitution (only 84% identity). We expanded our search to additional HCDR3 lengths, and we identified an additional three sequences with a total of three mutations each (including gaps). Four of the five sequences notably lacked the Val106-Asp107 dipeptide motif previously described to be critical for CH65-HA binding. We chimerized the junctional sequences with the highest ranked sequence homology on CH65 and displayed these Abs as yeast surface-display single chain variable fragments (scFv), however, we did not detect binding for any of those Abs to the SI06 HA head domain protein in a flow cytometric yeast binding assay (Table S1).

We next screened our sequence databases for predicted Ab structural homology and Ag binding affinity using the P3SM method, a Rosetta-based heuristic model. Since our initial manuscript describing this protocol, we found that calculating the P3SM using linear ridge regression was a robust, reproducible way to map Rosetta Energy Units (REU) to a position-specific scoring matrix and outperformed our previous means-based approach in several use cases (data not shown). Using CH65 as the template antibody, we observed a modest improvement in correlation compared to our previous means-based calculation (Fig S1). To build the P3SM, we first randomly selected 400 HCDR3 sequences from the Ab gene dataset and threaded each sequence over each of the three CH65 Ag-binding Ab fragment (Fab) structures in the asymmetric unit of the co-crystal structure (PDB ID 5UGY). The best five out of ten models by score were selected for each sequence/Fab pair. We calculated the P3SM using ridge regression to fit a coefficient for each of the 20 amino acids at each position in the HCDR3, optimizing for prediction of the total Rosetta score of the loop (Fig 1a). We provide a script that utilizes the Python package scikit-learn to enable this calculation (F et al., 2011). Next, we scored each HCDR3 sequence in our dataset with this P3SM and selected the top 600 rank-ordered sequences for further analysis. These 600 HCDR3 sequences were homology modeled onto the CH65/SI06 HA complex crystal structure (PDB ID 5UGY) using Rosetta, and we filtered these hits by predicted stability and binding energy to identify 15 sequences for experimental validation (Fig 1b). These 15 HCDR3 sequences failed to score better than wild-type CH65. Additionally, none of the selected sequences retained the Val106-Asp107 dipeptide motif important for CH65-HA binding (Fig 1c). These 15 HCDR3 sequences were expressed in the framework of CH65 on the surface of yeast as scFv, but as before, none bound to the SI06 HA protein when screened using a flow cytometric binding assay.

Upon reviewing the 15 sequences, we became interested in one cluster with highly similar HCDR3 sequences (Fig 1c, **denoted by asterisk**). This cluster of sequences belongs to a population of related Abs observed in the sequence repertoires of one of the donors with repertoire sequence data obtained at four separate time points between 2004 and 2005. The cluster was of particular interest because of the frequency of the clonotype in the subject samples and also because timeline of the collected samples (2004–2005) predated the identification of CH65 in 2008 (Schmidt et al., 2015a) and major changes occurring to circulating influenza H1 viruses in 2009. In early 2009, significant changes were introduced to circulating H1 viruses due to genetic re-assortment between human and swine influenza viruses, and CH65 does not bind to the 2009 pandemic H1N1 virus HA. The seven sequences belonging to this cluster were very similar, and we selected two as representative of the group based on the localization of mutations to the antibody-antigen interface versus mutations in the HCDR3 torso domain

We expressed two representative members of this family as Fabs using the full heavy chain variable domain sequence from our database to see if the lack of binding affinity we observed in yeast surface display was caused by the chimerization of HCDR3 sequences onto the CH65 framework in the scFv format. Our database of Ab sequences does not contain linked heavy and light chain sequences, and the native light chain pairing for these Abs is unknown, therefore we paired these heavy chains with either the CH65 or CH67 light

chain. The resulting Fabs for those sequences identified using our P3SM bound SI06 HA when paired with the CH65 but not the CH67 light chain, albeit at significantly lower affinity than wild-type (Fig 2).

To determine if minimal mutations to the HCDR3 of the chimeric Abs rescued Ag binding, we performed *in silico* affinity maturation using Rosetta Design (Fig S2). A majority of the sequences suggested by Rosetta Design failed to improve the P3SM and Rosetta HCDR3 scores significantly (data not shown). However, a combination of three residue change altered the critical binding dipeptide (W101L, V102H and S107E). Rosetta Design converged on a similar solution for sequence 7969 (W101L, V102H and Y107D). These sequences were expressed as chimerized CH65 or CH67 scFv, and we observed native-like binding to SI06 HA for many of the designed HCDR3 sequences in the CH67 background. This bias toward better performance of the HCDR3s in the CH67 background was unexpected and may be caused by interactions between the HCDR3 loops and the CH67 light chain variable domain, or an overall improved fit between the CH67 variable domains in the scFv format.

To characterize the binding affinity and antiviral function of these designed HCDR3 loops further, we chimerized the HCDR3 loops on mammalian cell-expressed forms of CH65 or CH67 Fab proteins. We measured virus inhibition function using hemagglutinin inhibition (HAI) and virus neutralization assays using SI06 (Table 1). Many of the designed HCDR3 sequences had comparable binding to HA as that of wild-type CH65 and CH67 and mediated similar levels of antiviral function in HAI and neutralization assays. The exception was HCDR3 1203d4, which showed a bias toward superior performance in the CH67 Fab background.

Once we confirmed that these chimeric Fabs had functional homology to CH65 and CH67, we sought to validate that our method correctly predicted their structural homology. We crystallized and solved structures for three of our four constructs as apo proteins (*i.e.*, uncomplexed with HA; Fig 3 and Table S2). Our structural analysis aimed to confirm that, like CH65 and CH67, these chimeras undergo structural pre-configuration of the HCDR3 loop before binding the HA interface, taking on the conformation observed in the native crystal structures. Indeed, the three new structures possess HCDR3 conformations highly consistent with the apo crystal structures of CH65 and CH67. When superimposed, chimeric HCDR3s 1203d4 and 7969d2 in the CH65 background had RMSD to the native CH65 HCDR3 of 0.198 Å and 0.219 Å, respectively. For chimeric HCDR3 7969d2 in the CH67 background, the RMSD to native CH67 HCDR3 was 0.685 Å; the larger deviation is caused mostly by a backbone movement at residue T104, which differs from the position of the native CH67 residue R104.

DISCUSSION

In summary, we screened a large database of Ab variable gene sequences and filtered those whose HCDR3 structure was predicted to be similar to that of the previously characterized influenza virus hemagglutinin protein specific monoclonal Ab CH65, regardless of their amino acid sequence. We found that when these identified sequences were chimerized in the

original Ab framework and expressed as recombinant proteins, the newly identified HCDR3 sequences encoded antibodies that functioned with similar specificity as CH65, albeit lower affinity, even though they were unrelated in sequence and could not be identified by sequence similarity. Minimal mutations engineered into the HCDR3 sequences via computational design recovered wildtype-like affinity for the antigen and the resulting antibodies showed potent viral neutralization. Crystal structures of these antibodies validated that the HCDR3 structures possessed a high degree of structural homology to the original CH65 Ab on which the structural prediction was based, despite dissimilarity of sequence. This work develops a framework for large-scale structure homology prediction and functional assignment of human Abs using next generation sequencing of the immunome.

We used this framework to identify new members of the CH65 Ab structural class, which binds to influenza HA in a manner similar to the native host ligand, sialic acid. It was unclear at the beginning of this work how widely prevalent the CH65 structure was in human antibody repertoires. Influenza HA is a common antigen to which all humans have been exposed, either through natural infection or vaccination, but the response to this antigen is structurally and functionally diverse. Therefore, we made no assumptions about how many or few antibodies our method would identify. Prior to this study, Schmidt et al. had used a sequence motif to identify CH65-like sequences in three of the five subjects from the original vaccinated cohort out of which CH65 and CH67 were identified (Schmidt et al., 2015b). The motif used to search for these sequences was minimal, focusing mostly on the VD dipeptide in the center of the HCDR3 loop but leaving considerable room for sequence and structural variation. Schmidt *et al.* showed that one antibody identified using the motif, referred to as H2526, took on a significantly different overall structure that interacts with HA in a flipped orientation (~180 degrees) compared to CH67 (Schmidt et al., 2015b). It is unclear whether the sequences identified via sequence motif from that study recapitulate the CH65 structure, or if each finds its own structural solution to enable antigen binding. Conversely, our method predicts functional similarity via similarity of the overall Ab structure, even though the loops diverge in sequence and genetic origin. We succeeded in identifying such antibodies despite the fact that the subjects in our sequencing database were not associated with the original cohort. Such structure-based prediction of specific binding of Abs represents an attractive approach to human therapeutic monoclonal Ab discovery based on immunome sequencing data, especially when neutralizing determinants are already known as in the case we studied.

The principal major advance we achieved was successful searches of large DNA sequence databases for structural homologs based on structural predictions associated with an antiviral function, which is an advance adding value to the many emerging massive immune repertoire sequence databases. The method as presented here does, however, have limitations. The approach is designed to find antibody sequences inferred to encode an antibody with structure similar to that of a known structure. Based on the central dogma of structural biology, sequence determines structure determines function, the hypothesis is that many of these antibodies also will function in a similar way when compared with the template. Thus, the method can discover new sequences encoding antibodies within a structural class of antibody but cannot discover entirely new classes of antibodies or functions. It is also likely that the method as configured is most likely to succeed for mAbs

for which the HCDR3 is the principal determinant of binding. As structure is better conserved than sequence, structural siblings can be identified that might be missed by a sequence search. Thus, one also expects a higher 'hit rate' when searching by structure rather than sequence. It seems logical to think that the method could identify antibodies in existing repertoires with distinct sequences but high levels of structural homology that exhibit superior activity to the original antibody on which the structural search is based, but this was not the focus of our approach. Here, we did not find such antibodies, since the identified candidates by the P3SM bound with lower EC50 values than did the original CH65. Better binding was achieved after limited in silico affinity maturation.

Rosetta design can be a computationally demanding method, however in this study only minimal amounts of design were needed to converge on the functional sequences selected. It is possible to predict these mutations with other computational methods, or even by knowledge-based approaches wherein a trained individual may be able to compare the models of identified sequences with the co-crystal structure of the wild-type antibody interacting with antigen and identify mutations that would improve binding. In this case, mutagenesis studies were needed to optimize the function of the antibodies, and only a subset of the variants were improved. Three residues of the HCDR3 loop were designed. The first two mutations were conserved in both 1203d4 and 7969d2 (W101L and V102H) and are not associated with the distal tip of HCDR3 nor are involved directly in antigen binding. In CH65 and CH67, the critical dipeptide V106, D107 in the HCDR3 loop is involved in antigen binding. Our HCDR3 sequences were optimized by Rosetta design at position 107; in 1203d4 via the mutation S107E, and in 7969d2 via the mutation Y107D. While co-crystal structures of our designed antibodies with SI06 HA were not determined in this study, analysis of our Rosetta models of CH65:1203d4 and CH65:7969d2 in comparison to the co-crystal structure of CH65 suggest that these mutations to position 107 may recover the native CH65-like interaction with SI06 HA.

Another limitation of this present study is the lack of natively paired light chain sequences in our starting sequence databases. Since the inception of this study, techniques to sequence native pairs of antibody heavy and light chains have been developed, and public databases of paired sequences have become available (Setliff et al., 2019; Wang et al., 2018). It is possible to apply our P3SM method recursively to search for antibodies that are structurally homologous across many, even all, CDR loops. However, we have not yet expanded the method to model and score multiple CDR loops simultaneously, which we believe is a precursor to fully leveraging the information in paired antibody sequences. Both aspects are greatly aided by rapidly increasing antibody sequence databases, increasing the likelihood of finding good matches for multiple CDRs in heavy and light chain simultaneously. This goal is the focus of ongoing research.

To overcome the lack of a native light chain sequence, the heavy chain sequences identified in our database were paired with the light chain sequences from the native CH65 and CH67 antibodies. We observed that the wild-type full length 1203 and 7969 heavy chain sequences (not chimerized onto CH65 frameworks or affinity matured through Rosetta design) were able to bind SI06 HA when paired with the CH65 light chain but not the CH67 light chain (Figure 2). It has recently been observed by Xiao *et al.* that light chain pairing is capable of

modulating HA binding by affecting the conformation of the heavy chain, and in particular the HCDR3 loop, even when the light chain does not interact directly with the HA antigen (Xiao et al., 2019). While the native antibodies CH65 and CH67 have very similar HCDR3 conformations, the loop is slightly displaced toward the light chain at its distal tip in the apo structure of CH67 (PDBID 4HKB), and the neighboring LCDR2 loop is not resolved. We observed a related change in the HCDR3 loop position within the context of the Fv domain in our crystallographic structures of CH65:1203d4 and CH67:1203d4, and we also were unable to resolve the structure of LCDR2 in CH67:1203d4. It is likely that the binding performance of our antibodies is impacted by the heavy-light chain pairing, but without resolution of the neighboring residues in the LCDR2 loop, we cannot completely assess this interface. In addition, while we hypothesize that the 1203 and 7969 sequences observed in our subject may have arisen in B cells responding to influenza infection or vaccination, we are unable to make this assertion based on the knowledge from Xiao *et al.* (Xiao et al., 2019) and the limitations of our available data.

Despite the limitations discussed here, these studies provide an advance enabling exploration of large adaptive immune repertoire datasets using structure-based predictions. This approach also could be applied to discovery of other medically important molecules such as chimeric Ag receptor (CAR) molecules used in T cell immunotherapy, new vaccine candidate Ags, and design of other protein molecules with structural diversity for which large collections of sequences have been generated.

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by James E. Crowe, Jr. (james.crowe@vumc.org).

Materials Availability—Plasmids for the antibodies to influenza generated in this study are available upon request for non-commercial research purposes but may require execution of a Materials Transfer Agreement.

Data and Code Availability—Atomic coordinates and structure factors for the crystal structures of apo Fabs CH65:7969d2, CH65:1203d4, CH67:1203d4 have been deposited in the Protein Data Bank with the accession codes 6DLA, 6DLB, and 6DL8, respectively. The Rosetta macromolecular modelling suite (<https://www.rosettacommons.org>) is freely available to academic and non-commercial users, commercial licenses are available via the University of Washington CoMotion Express License Program.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cells and viruses—The Madin Darby Canine Kidney Cell line was obtained from the American Type Culture Collection [MDCK (NBL-2) (ATCC® CCL-34™); the cell line was isolated originally from a normal adult female cocker spaniel]. The FreeStyle™ 293-F cell line was purchased from Thermo Fisher Scientific (#R79007); the parental cell line (human embryonic kidney 293 cells) were isolated originally from female human embryo tissues.

Cells in culture were tested monthly for the presence of mycoplasma and found to be negative. Cell line identity testing was performed with short tandem repeat (STR) DNA profiling for authentication of correct identity. MDCK cells were cultured at 37°C in Dulbecco's Modified Eagle's medium (DMEM) high glucose and Glutamax, supplemented with 10% heat inactivated fetal bovine serum (FBS) and 1% Gibco™ Penicillin-Streptomycin. The original seed stocks for the influenza virus SI06 was obtained through BEI Resources, NIAID, NIH: Influenza A Virus, A/Solomon Islands/3/2006 (H1N1), NR-41798. All of the working stocks were obtained by virus inoculation of MDCK cell culture monolayers in plain Dulbecco's Modified Eagle Medium (Gibco DMEM, Invitrogen, 11965) with 2 µg/mL of trypsin treated with N-tosyl-L-phenylalanine chloromethyl ketone (TPCK).

METHOD DETAILS

Next-generation DNA sequence analysis of expressed antibody variable genes.

—Total RNA was extracted from cryopreserved PBMCs, and a one-step RT-PCR was performed for using heavy-chain BIOMED-2 variable antibody gene-specific primers (van Dongen et al., 2003) and high-fidelity Taq polymerases. The Illumina-specific adapters were added using the Illumina TruSeq Library Preparation Kit (Illumina, FC-121–3001) according to the manufacturer's recommendations. The final amplicon libraries were sequenced on an Illumina MiSeq instrument using the MiSeq PE-300 v3 reagent kit (Illumina, MS-102–3001). Sequence analysis was performed using IgBLAST, and results were parsed to MongoDB for further study.

P3SM method—The command-line protocol for modeling antibody HCDR3 loops, calculating the P3SM, and scoring additional sequences is described in the Methods S1: **Rosetta Protocol Capture** and a suite of supporting scripts and example input files are provided in supplemental information (Methods S2: **P3SM_Protocol_Capture-directory-archive**). First, structure files for CH65 in complex with SI06 HA (PDB ID 5UGY) were downloaded from the PDB and cleaned such that each variable domain:Ag complex in the asymmetric unit were saved to their own files. These files were renumbered following instructions in the Rosetta Protocol Capture Section I. 400 randomly selected HCDR3 sequences from our database were homology modeled ($n=10$) over each of the template PDB files following instructions in the Rosetta Protocol Capture Section II. The 5 best-scoring models for each sequence:template pair were used to calculate the P3SM (6,000 total models) and instructions for identifying these models are in Rosetta Protocol Capture Section III. The P3SM was calculated and additional sequences were scored using support scripts described in Rosetta Protocol Capture Section III. Once sequences are scored with the P3SM, the top hits were homology modeled onto the template PDB file for further filtering, again following Rosetta Protocol Capture Section II. *In silico* affinity maturation was performed using Rosetta Design following existing protocols.

Yeast transformation—*Saccharomyces cerevisiae* Meyen ex E.C. Hansen EBY-100 yeast cells (EBY100 (a *GAL1-AGA1::URA3 ura3-52 trp1 leu2D1 his3D200 pep4::HIS2 prb1D1.6R can1 GAL*), Trp[−]Leu[−]) were transformed following the protocol by Gietz and Schiestl (Gietz and Schiestl, 2007). Briefly, yeast cells were washed and resuspended in

sterile water to 10^8 cells/mL. 10^7 cells were transferred to each well of a 96-well plate. The plate was centrifuged at $1,500 \times g$ for 5 minutes to pellet the cells and the supernatant was removed by inverting the plate. Transformation mix was prepared following the published protocol (Gietz and Schiestl, 2007). 50 μ L transformation mix added to each well, and cells were resuspended by pipetting. 100 μ L PEG3350 (50% w/v) was added to each well and contented mixed by pipetting. Cells were heat shocked by incubating at 42°C for one hour. The plate was centrifuged for ten minutes at $1,500 \times g$ and the supernatant again removed by inverting the plate. 50 μ L of sterile water was added to each well and the cells were resuspended by pipetting. The transformed yeast were plated on selective plates containing SD-CIT.CAA medium (defined synthetic dextrose medium with casamino acids and citrate buffer) and incubated at 30°C, for 72 hours.

Yeast growth and scFv expression—Transformed colonies were picked to 4 mL SD-CIT.CAA medium and grown at 30°C, shaking at 225 rpm for 24 hours. Next, yeast were pelleted by centrifugation, washed with SG-CIT.CAA expression medium, and transferred to SG-CIT.CAA. Yeast were incubated at room temperature shaking at 225 rpm for 24 to 48 hours to allow for scFv expression.

Recombinant antibody expression and purification—The heavy and light chain variable regions were cloned into Fab heavy chain or lambda chain vectors (McLean et al., 2000), respectively. The Fab fragment was expressed by transient co-transfection of the expression vector containing heavy chain and light chain into Expi-CHO cells. Recombinant Fab was purified from culture supernatant using an anti-CH1 CaptureSelect column (ThermoFisher Scientific). Purified Fab was measured by optical absorbance at 280 nm, and purity and integrity were analyzed by reducing and nonreducing sodium dodecyl sulfate polyacrylamide gel electrophoresis (**SDS-PAGE**). The purified Fab was concentrated to ~10 mg/mL for crystallization and K_D determination.

Production of recombinant soluble SI06 HA protein—We produced a recombinant form of the head domain of HA protein A/Solomon Islands/2006 (designated here as SI06) in a human-origin cell line similarly to methods used to express other H1 HAs previously reported (Yu et al., 2008). We chose this strain of influenza as the source of the HA sequence because the influenza HA-specific human Ab CH65 was crystallized in complex with the H1 HA of SI06 (PDB ID: 5UGY). cDNA sequences encoding the amino acids of the SI06 HA head domain (as shown in PDB: 5UGY) were synthesized to encode soluble trimeric protein constructs by replacing the transmembrane and cytoplasmic domain sequences with cDNAs encoding the *Saccharomyces cerevisiae* General Control Protein (GCN4) trimerization domain and a His-tag at the C-terminus. Synthesized genes were subcloned into the pcDNA3.1(+) mammalian expression vector (ThermoFisher Scientific), and then transfected and expressed in Expi-293F cells (ThermoFisher Scientific). SI06 HA protein was purified from culture supernatant using a HiTrap TALON crude column (GE Life Sciences).

Flow cytometric binding analysis of yeast surface display scFv—After inducing surface-display scFv expression in yeast cells, 5×10^5 cells were added to each well of a 96-well V-bottom plate. The cells were pelleted by centrifugation at $1,500 \times g$ for 5 minutes and

washed once in wash buffer made from phosphate buffered saline (PBS) containing 0.05% BSA (bovine serum albumin). The cells were resuspended in 50 μ L wash buffer containing 10 nM biotinylated SI06 HA. The cells were allowed to incubate in antigen for one hour at room temperature, after which they were washed three times. Next, the cells were resuspended in 50 μ L stain solution (1:250 fluorescein isothiocyanate (FITC)-conjugated V5 peptide and 0.1 μ g/well allophycocyanin (APC)-conjugated streptavidin prepared in wash buffer). The cells were again allowed to incubate for one hour at room temperature, after which they were pelleted and washed one time. Finally, the cells were resuspended in 250 μ L wash buffer and kept on ice until analyzed by flow cytometry.

Half maximal effective concentration (EC₅₀) analysis—ELISAs were performed to obtain EC₅₀ values for binding using 384-well plates coated with the HA of interest at a 2 μ g/mL concentration and incubated overnight at 4°C. The plates were blocked with 5% non-fat dry milk, 2% goat serum, and 0.1% Tween-20 in PBS for one hour. Three-fold dilutions of the mAb, starting from 50 μ g/mL, were added to the wells, incubated for one hour, followed by a one-hour incubation of 1:4,000 dilution of goat F(ab')₂ anti-human lambda light chain horseradish peroxidase conjugate (SouthernBiotech, 2072–05). The plates were washed three times between each step with PBS containing 0.1% Tween-20. 1-Step TMB (3,3',5,5'-tetramethylbenzidine) Ultra-ELISA Substrate solution (ThermoFisher) was added to the plates, incubated for ten minutes, and the optical density values were measured at 450 nm wavelength on a BioTek plate reader. Each dilution was performed in duplicate, and the EC₅₀ values were calculated in Prism software (GraphPad) using non-linear regression analysis.

Hemagglutination inhibition (HAI) assay—HAI assays were performed using standard procedures (Herfst et al., 2012). Briefly, an aliquot of 4 hemagglutination units of virus was pre-incubated with two-fold serial antibody dilutions for one hour. Virus-antibody mixtures were incubated with turkey red blood cells (Innovative Research Inc.) for one hour at room temperature. The HAI titer was defined as the highest dilution of antibody that inhibited hemagglutination of red blood cells.

Half maximal inhibitory concentration (IC₅₀) analysis—40 TCID₅₀ of A/Solomon Islands/3/2006 H1N1 (FR-331; IRR) virus was mixed with serial two-fold dilutions of Fabs starting from 40 μ L/mL and incubated for 1 hour at RT. This virus-antibody mixture was added to monolayers of MDCK cells and incubated for 24 hours at 37°C with 5% CO₂. A recombinant form of human IgG CH65 was used as a positive control, and an unrelated antibody, EEEV-16 Fab, was used as negative control. All experiments were performed in triplicate. After 24 hours cells were fixed, and virus was inactivated by 80% methanol in PBS. Cells were incubated with blocking buffer PBS with 0.05% Tween-20 for 1 hour at RT, primary anti-influenza nucleoprotein mouse antibodies (used 1:6,000; NR-4282; BEI Resources) for 1 hour at RT, and secondary anti-mouse IgG alkaline phosphatase (A conjugated antibodies (1:3,000, Novex, DKXMO AP AFFINITY, A16014) for 1 hour at RT. Wells with virus-infected cell culture monolayers were visualized by AP substrate. Low OD₄₀₅ values correspond to samples without influenza virus. High OD₄₀₅ values correspond

to the amount of influenza antigens in infected MDCK cells. EC₅₀ values were calculated using a non-linear regression analysis by Prism v. 5.0 (GraphPad).

K_D determination—K_D values were determined by bio-layer interferometry using an Octet RED instrument (FortéBio, Inc.), as described previously (Thornburg et al., 2016). Biotinylated SI06 HA protein (10 µg/mL) was loaded onto streptavidin-coated biosensors in kinetics buffer (1× PBS, pH 7.4, 1% bovine serum albumin [BSA], and 0.05% Tween 20) for 300 sec. For measurement of *k*_{on}, association of Ab was measured for 120 sec by exposing the sensors to seven concentrations of Fab (2-fold dilutions starting at 200 nM) in kinetics buffer. For measurement of *k*_{off}, dissociation of Ab was measured for 120 sec in kinetics buffer. Experiments were performed at 30°C. K_D was calculated as the ratio of *k*_{off} to *k*_{on} determined from binding curves of at least four concentrations for each Fab.

Crystallization and x-ray structure determination—All recombinant Fabs were concentrated to ~ 10 mg/mL in 20 mM Tris-HCl, 50 mM NaCl for crystallization trials, and Fab crystals were grown using sitting-drop vapor diffusion method at 18 °C. Crystals of Fab CH65:1203d4 were obtained in 1.8 – 2.2 M (NH₄)₂SO₄, 100 mM Tris-HCl pH 8.3, crystals of Fab CH65:7969d2 in 1.8 – 2.2 M (NH₄)₂SO₄, 100 mM Bis-tris pH 6.5, and crystals of Fab CH67:1203d4 in 22% - 32% PEG 1000, 100 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.0 – 8.0. Crystals were cryo-protected in mother liquor supplemented with 20% (w/v) glycerol (CH65:1203d4 and CH65:7969d2) or 40% poly(ethylene glycol) (PEG) 1000 (CH67:1203d4), flash frozen, and stored in liquid nitrogen until data collection. X-ray diffraction data for the CH65:7969d2, CH65:1203d4, CH67:1203d4 apo Fabs were collected to 2.00 Å, 2.20 Å or 3.80 Å resolutions at the Cornell High Energy Synchrotron Source (CHESS) F1 beamline. The diffraction data sets were processed and scaled with the *XDS* (Kabsch, 2010) and *scala* (Winn et al., 2011) program software packages. The crystal structure of Fabs was determined by molecular replacement with Phaser (McCoy et al., 2007) using the variable and constant domains of Fabs in the PDB (4WUK, 4HKB) (Lee et al. 2015, Schmidt et al., 2013) as search models for CH65:1203d4/CH65:7969d2 and CH67:1203d4, respectively. The model was rebuilt iteratively using Coot (Emsley and Cowtan, 2004) and refined in Phenix (Adams et al., 2010). Final refinement statistics are summarized in Table S2.

QUANTIFICATION AND STATISTICAL ANALYSIS

Prism software version 5 (GraphPad) was used for all statistics. The HAI titer was defined as the highest dilution of antibody that inhibited hemagglutination of red blood cells; mean values are shown in Table 1. EC₅₀ values for binding were calculated using data from experiments in which each dilution was performed in duplicate, and the EC₅₀ values were calculated in Prism software version 5 (GraphPad) using non-linear regression analysis, in Table 1 and Figure 2; mean values are shown. K_D values were calculated as the ratio of *k*_{off} to *k*_{on} determined from binding curves of at least four concentrations for each Fab, in Table 1; mean values with standard deviations are shown.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

A grant from the National Institutes of Health (U19 AI117905 to J.E.C. and J.M.) and a contract from the NIH (HHSN272201400024C to J.E.C.) supported this work. J.A.F. was supported by NIH T32 AI 089554 and through the Vanderbilt Trans-Institutional Program (TIP) “Integrating Structural Biology with Big Data for next Generation Vaccines”.

REFERENCES

- Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66, 213–221. [PubMed: 20124702]
- Emsley P, and Cowtan K (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126–2132. [PubMed: 15572765]
- Gietz RD, and Schiestl RH (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* 2, 31–34. [PubMed: 17401334]
- Herfst S, Schrauwen EJ, Linster M, Chutinimitkul S, de Wit E, Munster VJ, Sorrell EM, Bestebroer TM, Burke DF, Smith DJ, et al. (2012). Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* 336, 1534–1541. [PubMed: 22723413]
- Kabsch W (2010). Xds. *Acta Crystallogr D Biol Crystallogr* 66, 125–132. [PubMed: 20124692]
- Kuroda D, and Gray JJ (2016). Shape complementarity and hydrogen bond preferences in protein-protein interfaces: implications for antibody modeling and protein-protein docking. *Bioinformatics* 32, 2451–2456. [PubMed: 27153634]
- Lee PS, Arnell AJ, Wilson IA (2015) Structure of the apo anti-influenza CH65 Fab. *Acta Crystallogr F Struct Biol Commun* 71, 145–148 [PubMed: 25664786]
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, and Read RJ (2007). Phaser crystallographic software. *J Appl Crystallogr* 40, 658–674. [PubMed: 19461840]
- McLean GR, Nakouzi A, Casadevall A, and Green NS (2000). Human and murine immunoglobulin expression vector cassettes. *Mol Immunol* 37, 837–845. [PubMed: 11257305]
- Navis M, Tran K, Bale S, Phad GE, Guenaga J, Wilson R, Soldemo M, McKee K, Sundling C, Mascola J, et al. (2014). HIV-1 receptor binding site-directed antibodies using a VH1–2 gene segment orthologue are activated by Env trimer immunization. *PLoS Pathog* 10, e1004337. [PubMed: 25166308]
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. (2011). Scikit-learn: machine learning in Python. *J Mach Learn Res* 12, 2825–2830.
- Schmidt AG, Do KT, McCarthy KR, Kepler TB, Liao HX, Moody MA, Haynes BF, and Harrison SC (2015a). Immunogenic stimulus for germline precursors of antibodies that engage the influenza hemagglutinin receptor-binding site. *Cell Rep* 13, 2842–2850. [PubMed: 26711348]
- Schmidt AG, Therkelsen MD, Stewart S, Kepler TB, Liao HX, Moody MA, Haynes BF, and Harrison SC (2015b). Viral receptor-binding site antibodies with diverse germline origins. *Cell* 161, 1026–1034. [PubMed: 25959776]
- Schmidt AG, Xu H, Khan AR, O'Donnell T, Khurana S, King LR, Manischewitz J, Golding H, Suphaphiphat P, Carfi A, et al. (2013). Preconfiguration of the antigen-binding site during affinity maturation of a broadly neutralizing influenza virus antibody. *Proc Natl Acad Sci U S A* 110, 264–269. [PubMed: 23175789]
- Setliff I, Shiakolas AR, Pilewski KA, Murji AA, Mapengo RE, Janowska K, Richardson S, Oosthuysen C, Raju N, Ronsard L, et al. (2019). High-throughput mapping of B cell receptor sequences to antigen specificity. *Cell* 179, 1636–1646 e1615. [PubMed: 31787378]

- Sui J, Hwang WC, Perez S, Wei G, Aird D, Chen LM, Santelli E, Stec B, Cadwell G, Ali M, et al. (2009). Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. *Nat Struct Mol Biol* 16, 265–273. [PubMed: 19234466]
- Thompson KM, Sutherland J, Barden G, Melamed MD, Randen I, Natvig JB, Pascual V, Capra JD, and Stevenson FK (1991). Human monoclonal antibodies against blood group antigens preferentially express a VH4–21 variable region gene-associated epitope. *Scand J Immunol* 34, 509–518. [PubMed: 1718028]
- Thornburg NJ, Zhang H, Bangaru S, Sapparapu G, Kose N, Lampley RM, Bombardi RG, Yu Y, Graham S, Branchizio A, et al. (2016). H7N9 influenza virus neutralizing antibodies that possess few somatic mutations. *J Clin Invest* 126, 1482–1494. [PubMed: 26950424]
- van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL, Delabesse E, Davi F, Schuurin E, Garcia-Sanz R, et al. (2003). Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98–3936. *Leukemia* 17, 2257–2317. [PubMed: 14671650]
- Wang B, DeKosky BJ, Timm MR, Lee J, Normandin E, Misasi J, Kong R, McDaniel JR, Delidakis G, Leigh KE, et al. (2018). Functional interrogation and mining of natively paired human VH:VL antibody repertoires. *Nat Biotechnol* 36, 152–155. [PubMed: 29309060]
- Whittle JR, Zhang R, Khurana S, King LR, Manischewitz J, Golding H, Dormitzer PR, Haynes BF, Walter EB, Moody MA, et al. (2011). Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin. *Proc Natl Acad Sci U S A* 108, 14216–14221. [PubMed: 21825125]
- Willis JR, Finn JA, Briney B, Sapparapu G, Singh V, King H, LaBranche CC, Montefiori DC, Meiler J, and Crowe JE Jr. (2016). Long antibody HCDR3s from HIV-naïve donors presented on a PG9 neutralizing antibody background mediate HIV neutralization. *Proc Natl Acad Sci U S A* 113, 4446–4451. [PubMed: 27044078]
- Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, et al. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67, 235–242. [PubMed: 21460441]
- Xiao H, Guo T, Yang M, Qi J, Huang C, Hong Y, Gu J, Pang X, Liu WJ, Peng R, et al. (2019). Light chain modulates heavy chain conformation to change protection profile of monoclonal antibodies against influenza A viruses. *Cell Discov* 5, 21. [PubMed: 30993000]
- Xu H, Schmidt AG, O'Donnell T, Therkelsen MD, Kepler TB, Moody MA, Haynes BF, Liao HX, Harrison SC, and Shaw DE (2015). Key mutations stabilize antigen-binding conformation during affinity maturation of a broadly neutralizing influenza antibody lineage. *Proteins* 83, 771–780. [PubMed: 25524709]
- Yu X, Tsibane T, McGraw PA, House FS, Keefer CJ, Hicar MD, Tumpey TM, Pappas C, Perrone LA, Martinez O, et al. (2008). Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. *Nature* 455, 532–536. [PubMed: 18716625]

Highlights

- New position-specific structure-scoring matrix method using Rosetta
- Effective sequence-to-function prediction for identifying virus-specific antibodies
- Crystallography and virus inhibition assays validated the models

Figure 1. A position-specific structure scoring matrix (P3SM) identifies potential structural homologs to CH65.

A. Linear ridge regression analysis of per-residue Rosetta energy scores from homology modeling simulations of 400 HCDR3 sequences over the CH65 crystal structure 5UGY determined a weight for each amino acid at each HCDR3 position (PDB residues 312–330), resulting in a position-specific structure scoring matrix (P3SM).

B. The top 1,000 sequences rank-ordered by P3SM score each were homology modeled on the CH65/SI06 complex structure, and the average Rosetta score (in Rosetta energy units, REU) versus the root mean square deviation (RMSD) of the HCDR3 loop is shown. The native CH65 sequence is shown in red for comparison. Limited HCDR3 loop deviation was observed and sequences with good scores (low REU) retained the native CH65 structure during modeling.

C. The 15 best homology modeled sequences rank-ordered by HCDR3 score were selected for further analysis of the Ab/Ag interaction. Sequence IDs denoted by an asterisk (*)

belong to a cluster of clonally related sequences identified from a single donor. Apo HCDR3 scores as well as the complex DDG were calculated by separating the Ab from the Ag and rescoring the Ab while allowing limited side-chain minimization. These sequences were expressed as chimeric HCDR3 on yeast surface-displayed CH65 scFv, and the percentage of cells expressing scFv and the percentage of scFv(+) cells that bound to SI06 HA in a flow cytometric assay are listed.

A

Heavy chain	EC ₅₀ (μg/mL) in indicated background	
	CH65	CH67
wt	0.05	0.67
1203	13.3	>50
7969	28.4	>50

B

Figure 2. Two HCDR3 sequences identified using P3SM were chimerized into either CH65 or CH67 recombinant Fab backgrounds.

A. Binding to SI06 HA was assessed as half-maximal binding (EC₅₀ values) calculated from representative binding curves shown in panel B.

B. The sequences identified using our method showed preferential bias to the CH65 Fab background. ELISA was performed with two biological replicates, each time with technical triplicate replicates; one experiment is shown.

Figure 3. X-ray crystallography confirms structural homology of P3SM-selected antibody sequences.

A. The structure of Fab CH65:7969d2 is presented in rainbow, aligned to the variable domain of CH65 (PDB ID 4WUK) in gray.

B. CH65:1203d4 in rainbow, aligned to the variable domain of CH65 in gray, as in A.

C. The structure of Fab CH67:1203d4 is presented in rainbow, aligned to the variable domain of CH67 (PDB ID 4HKB) in gray. In each structure, the HCDR3 loop is colored cyan.

Table 1.
***In silico* affinity maturation of P3SM-identified sequences rescues wild-type function.**

Half-maximal binding determined by ELISA (EC_{50}), binding affinity calculated using Octet BLI (K_D), hemagglutinin inhibition (HAI) and half-maximal viral neutralization (IC_{50}) of the designed HCDR3 sequences chimerized to CH65 or CH67 Fab are shown. Two negative controls were included in the study; CH65 V106D is a loss-of-function point mutant of the wild-type CH65 sequence, while EEEV-16 is a recombinant human Fab not specific to influenza. The binding, HAI and neutralization tests were performed two or three times with two technical replicates in each assay. The K_D experiment was performed once, with four technical replicates.

Antibody	HCDR3 Sequence	ELISA EC_{50} ($\mu\text{g/mL}$)	K_D (nM)	HAI endpoint ($\mu\text{g/mL}$)	Neut. IC_{50} ($\mu\text{g/mL}$)
CH65	ARGGLEPRSDYYYYYGMDV	0.04	17.8 ± 0.08	6.3	0.77
CH65:1203d4 Chimera	ARAGLHPTTTEYYYYGMDV	0.92	116 ± 2.55	25.0	3.03
CH65:7969d2 Chimera	ARAGLHPTTTDYYYYGMDV	0.12	36.5 ± 0.04	3.1	0.96
CH67	ARAGLEPRSDYFFYGLDV	0.58	45.5 ± 0.46	12.5	0.65
CH67:1203d4 Chimera	ARAGLHPTTTEYYYYGMDV	0.26	55.6 ± 1.34	12.5	0.83
CH67:7969d2 Chimera	ARAGLHPTTTDYYYYGMDV	0.10	26.2 ± 0.02	6.3	0.58
CH65 V106D	ARGGLEPRSDYYYYYGMDV	> 50	<i>n.d.</i>	<i>n.d.</i>	> 40
EEEV-16	ARADGYNFDY	> 50	<i>n.d.</i>	<i>n.d.</i>	> 40

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rosetta sequence 1203 designed antibody	This study	N/A
Rosetta sequence 16996 designed antibody	This study	N/A
Rosetta sequence 8200 designed antibody	This study	N/A
Rosetta sequence 7969 designed antibody	This study	N/A
Rosetta sequence 18444 designed antibody	This study	N/A
Rosetta sequence 4492 designed antibody	This study	N/A
Rosetta sequence 11225 designed antibody	This study	N/A
Rosetta sequence 3083 designed antibody	This study	N/A
Rosetta sequence 10961 designed antibody	This study	N/A
Rosetta sequence 17038 designed antibody	This study	N/A
Rosetta sequence 12374 designed antibody	This study	N/A
Rosetta sequence 13458 designed antibody	This study	N/A
Rosetta sequence 9987 designed antibody	This study	N/A
Rosetta sequence 14857 designed antibody	This study	N/A
Rosetta sequence 3342 designed antibody	This study	N/A
CH65 (recombinant CHO-produced IgG1)	Based on sequence in PDB ID: 5UGY	N/A
CH67 (recombinant CHO-produced IgG1)	Based on sequence in PDB ID: 4HKX	N/A
CH65:1203d4 Chimera; incorporating Rosetta designed HCDR3 1203d4 into a backbone of CH65	This study	N/A
CH65:7969d2 Chimera; incorporating Rosetta designed HCDR3 7969d2 into a backbone of CH65	This study	N/A
CH67:1203d4 Chimera; incorporating Rosetta designed HCDR3 1203d4 into a backbone of CH67	This study	N/A
CH67:7969d2 Chimera; incorporating Rosetta designed HCDR3 7969d2 into a backbone of CH67	This study	N/A
CH65 V106D mutant antibody	This study	N/A
Anti-influenza nucleoprotein mouse antibodies	BEI Resources	Cat# NR-4282
Anti-mouse IgG AP conjugated antibodies	Novex, DKXMO AP AFFINITY	Cat# A16014
Goat F(ab') ₂ anti-human lambda light chain horseradish peroxidase conjugate	Southern Biotech	Cat# 2072-05
Bacterial and Virus strains		
Influenza A Virus, A/Solomon Islands/3/2006 (H1N1)	Biodefense and Emerging Infections Research Resources Repository [BEI Resources]	Cat# NR-41798
Chemicals, Peptides, and Recombinant Proteins		
1-Step Ultra TMB-ELISA	ThermoFisher	Cat# 34029
Freestyle 293 expression medium	ThermoFisher	Cat# 12338002
ExpiCHO Expression Medium	ThermoFisher	Cat# A2910001
Fetal Bovine Serum, ultra-low IgG	ThermoFisher	Cat# 16250078
Dulbecco's Modified Eagle Medium	Invitrogen	Cat# 11965

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HiTrap TALON crude column	GE Life Sciences	Cat# 28953766
Deposited Data		
Crystal structure of apo Fab CH65:7969d2	This paper	PDB: 6DLA
Crystal structure of apo Fab CH65:1203d4	This paper	PDB: 6DLB
Crystal structure of apo Fab CH67:1203d4	This paper	PDB: 6DL8
Crystal structure of CH65 in complex with SI06 HA	Whittle et al., 2011	PDB ID: 5UGY
Crystal structure of CH67 in complex with SI06 HA	Schmidt et al., 2013	PDB ID: 4HKX
Crystal structure of apo CH65	Lee et al., 2015	PDB ID: 4WUK
Crystal structure of apo CH67	Schmidt et al., 2013	PDB ID: 4HKB
Experimental Models: Cell Lines		
EBY-100 yeast cells	ATCC	Cat# MYA-4941
Hamster: ExpiCHO-S	ThermoFisher	Cat# A29127
Human: FreeStyle 293F	ThermoFisher	Cat# R79007
Madin Darby Canine Kidney Cell line	ATCC	Cat# CCL-34
Recombinant DNA		
Heavy-chain BIOMED-2 variable antibody gene-specific primers	Genscript	N/A; Custom synthesis
TruSeq Library Preparation Kit	Illumina	Cat# RS-122-2001
MiSeq PE-300 v3 reagent kit	Illumina	Cat# MS-102-3001
Plasmids: Fab heavy chain or lambda chain expression vectors	Genscript	N/A; Custom synthesis
Software and Algorithms		
Prism 7	GraphPad	N/A
FlowJo 10	FlowJo 10	N/A
UCSF Chimera	University of California	N/A
PyMOL	Schrödinger	N/A; PyMOL version 2.3
Other		
iQue Screener Plus flow cytometer	iQue Screener Plus flow cytometer	N/A
Synergy H1 microplate reader	Synergy H1 microplate reader	N/A
EL406 washer dispenser	EL406 washer dispenser	N/A
Anti-CH1 CaptureSelect column	ThermoFisher	Cat# 19432001L
MiSeq instrument	MiSeq instrument	N/A
Octet RED instrument	Octet RED instrument	N/A