Rapid synthesis and screening of natively paired antibodies against influenza hemagglutinin stem via oPool+ display

Wenhao O. Ouyang[1], Huibin Lv[1,2], Wenkan Liu[1], Zongjun Mou[3], Ruipeng Lei[1], Tossapol Pholcharee[1], Yiquan Wang[1], Katrine E. Dailey[1], Akshita B. Gopal[1], Danbi Choi[1], Madison R. Ardagh[1], Logan Talmage[1], Lucia A. Rodriguez[1], Xinghong Dai[3], Nicholas C. Wu[1,2,4,5,§]

[1] Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[2] Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[3] Department of Physiology and Biophysics, Case Western Reserve University, School of Medicine, Cleveland, OH 44106, USA

[4] Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[5] Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[§] To whom correspondence should be addressed. Email: nicwu@illinois.edu (N.C.W.)

ABSTRACT

Antibody discovery is crucial for developing therapeutics and vaccines as well as understanding adaptive immunity. However, the lack of approaches to synthesize antibodies with defined sequences in a high-throughput manner represents a major bottleneck in antibody discovery. Here, we presented oPool$^+$ display, which combines oligo pool synthesis and mRNA display to construct and characterize many natively paired antibodies in parallel. As a proof-of-concept, we applied oPool$^+$ display to rapidly screen the binding activity of >300 natively paired influenza hemagglutinin (HA) antibodies against the conserved HA stem domain. Structural analysis of 16.ND.92, one of the identified HA stem antibodies, revealed a unique binding mode distinct from other known broadly neutralizing HA stem antibodies with convergent sequence features. Yet, despite such differences, 16.ND.92 remained broadly reactive and conferred *in vivo* protection. Overall, this study not only established an experimental platform that can be applied in both research and therapeutics to accelerate antibody discovery, but also provides molecular insights into antibody responses to the influenza HA stem, which is a major target for universal influenza vaccine development.

Keywords: High-throughput screening, influenza, hemagglutinin, antibody, broadly neutralizing, oligo pools, mRNA display, cryo-EM

INTRODUCTION

Antibodies are central to the immune system for protection against pathogen infection. Therefore, identification of antibodies that target pathogens of interest is key to the understanding of adaptive immunity as well as the development of effective therapeutics and vaccines. In recent years, advances in single-cell B-cell receptor sequencing (scBCR-seq) have greatly improved the capacity to discover novel antibodies[1]. Thousands of natively paired antibody sequences can be obtained from a single scBCR-seq experiment. By contrast, downstream characterization of these antibody sequences remains costly, labor intensive, and time consuming, involving cloning, expression, purification, and testing the binding activities of different antibodies individually. At the same time, protein display technologies offer a high-throughput solution for characterizing antibody binding activity[2], with antibody library construction being an essential first step. Methods for constructing antibody libraries with random heavy-light chain pairing from B cell repertoires are well-established[3,4]. However, there is a lack of approaches to synthesize custom-made antibody libraries with precise heavy-light chain pairing from a defined list of antibody sequences. This technical barrier has restricted the application of protein display technologies in antibody research.

Hemagglutinin (HA) is the major antigen of Influenza A virus. HA is a homotrimeric glycoprotein consisting of a hypervariable globular head domain atop a conserved stem domain[5,6]. The functions of HA are critical for viral entry. The globular head domain engages the sialylated receptor, whereas the stem domain possesses the membrane fusion machinery. In the past two decades, many human antibodies to the HA stem have been discovered and characterized[7–11]. In contrast to HA head antibodies, which are usually strain-specific, HA stem antibodies often cross-neutralize multiple influenza subtypes. Several recurring sequence features have been observed among HA stem antibodies isolated from different individuals, such as IGHV1-69, IGHV1-18, IGHD6-1, and IGHD3-9[7,11–14]. However, many HA stem antibodies do not contain any known recurring sequence features, suggesting there are additional ones have yet to be discovered. Comprehending the recurring sequence features

67 of HA stem antibodies is critical for the molecular understanding of broadly neutralizing
68 antibody responses at the population scale, which will in turn benefit the development of
69 universal influenza vaccines.
70
71 In this study, we developed oPool^{+} display, a rapid and cost-effective *in vitro* platform that
72 combined oligo pool synthesis with mRNA display to assemble and screen natively paired
73 antibodies in a highly parallel manner. As a proof-of-concept, we first synthesized a library of
74 325 natively paired HA antibodies, then screened them for specificity to the HA stem. Our
75 screen identified 12 HA stem antibody candidates that were not previously known to target the
76 HA stem. Cryo-electron microscopy (cryo-EM) analysis revealed that the binding mode of one
77 of these HA stem antibodies, 16.ND.92, was dominated by immunoglobulin heavy diversity 3-
78 3 (IGHD3-3) gene. 16.ND.92 also exhibited a unique CDR H3 conformation compared to other
79 known HA stem antibodies that relied on IGHD3-3 for binding. Yet, 16.ND.92 remained broadly
80 reactive and protected against lethal influenza challenge. This observation substantiated that
81 IGHD3-3 is a recurring sequence feature of HA stem antibodies and demonstrated its
82 versatility in binding to HA stem.
83
84 **RESULTS**
85 **High throughput assembly of natively paired antibody library via oligo pools**
86 We previously curated a dataset containing 5,561 human monoclonal HA antibodies, where
87 1,082 antibodies (19.5%) are known to bind to either the head domain or the stem domain[15].
88 Of the remaining 4,479 (80.5%) HA antibodies which lack epitope information, 292 have
89 complete heavy chain variable (V_H) and light chain variable (V_L) sequences available and are
90 not encoded by well-characterized sequence features of HA stem antibodies, namely IGHV1-
91 69, IGHV1-18, IGHV6-1, and IGHD3-9[7,12–14] (**Figure S1A and Table S1**). These 292 HA
92 antibodies were included in the synthesis of our natively paired antibody library. In addition,
93 three known stem antibodies, namely 31.a.55, AG11-2F01, and 042-100809-2F04[7,16], as well

as 30 known HA head antibodies[6,17–25] were included as positive and negative controls, respectively, bringing the total library size to 325 antibodies **(Figure 1A)**.

To synthesize the natively paired antibody library in a high-throughput manner, we aimed to leverage recent advances in oligo pool synthesis. The maximum length of each oligo in commercial oligo pool synthesis is around 300 to 350 nucleotides. By contrast, the length of a single-chain variable fragment (scFv), which is the smallest format of a human antibody, is around 800 to 900 nucleotides. As a result, each given scFv sequence was split into 4 oligos with overlaps at the diverse complementary-determining regions (CDRs), namely the CDRs H1, H3, and L3 (**Figure 1B-C and Table S2**). We then performed codon randomization to ensure that the overlaps among oligos for the same scFv are unique at the nucleic acid level. This would help prevent mis-annealing between oligos from different scFvs, especially if they shared similar amino acid sequences (**Figure 1B, Material and Methods**). Through an overlap PCR, full-length scFv sequences could then be generated with the intended native V_H and V_L pairing. Subsequently, this strategy was applied to synthesize our library of 325 antibodies using 13 oligo pools with 25 antibodies per pool (**Figure 1C-F, Figure S1B-C, and Table S2**).

The length of the assembled product was consistent with that of full-length scFvs (**Figure 1D and Figure S2**). PacBio sequencing revealed that 319 of 325 (98%) natively paired antibodies were successfully assembled (**Table S4**). Over 88% of the assemblies represented natively paired scFvs, in which half had no mutations (**Figure 1E**). Only around 10% of the assemblies were non-natively paired scFvs with V_H and V_L from different antibodies. Mutations in natively paired scFvs were likely introduced during oligo pool synthesis, which is known to be error-prone[26,27]. Moreover, a Pearson correlation coefficient of 0.9 was observed between replicates (**Figure 1F**), demonstrating the reproducibility of our oligo pool-based strategy for antibody library synthesis. While we opted to assemble this antibody library by 25 antibodies per pool, we have also tested the assembly accuracy of 50 antibodies per pool. Among the 50

122 antibodies, 48 (96%) were successfully assembled with a Pearson correlation coefficient of
123 0.98 between replicates **(Figure S2A-B)**, demonstrating the potential to further increase the
124 throughput of our antibody library synthesis strategy.
125
126 **Rapid screening of the natively paired antibody library by mRNA display**
127 To identify HA stem antibodies from our natively paired antibody library, we utilized mRNA
128 display[28–30], a well-established *in vitro* screening approach that allows rapid screening for
129 protein binders (**Figure 2A-D**). Briefly, each scFv was covalently linked to the RNA molecule
130 that encoded it, thus providing a phenotype-genotype linkage. The mRNA-displayed antibody
131 library was then selected against H1 and H3 stem constructs that were designed based on
132 H1N1 A/Brisbane/59/2007 HA and H3N2 A/Finland/486/2004 HA, respectively[31,32]. Both pre-
133 and post-selection libraries were analyzed by PacBio sequencing to quantify the enrichment
134 of each scFv during selection. Pearson correlation coefficients of 0.95 and 0.93 were observed
135 between biological replicates of selections against H1 stem and H3 stem, respectively (**Figure**
136 **2B-C**), demonstrating the reproducibility of the selections. The positive controls were highly
137 enriched, whereas the negative controls were not (**Figure 2B-C and Table S3**), further
138 validating that selection had taken place. Besides, we identified one antibody candidate
139 targeting H1 stem and 11 antibody candidates targeting H3 stem (**Figure 2D**). Of note, one of
140 the H3 stem antibody candidates, AG2-G02, was concurrently identified as an H3 stem
141 antibody using a machine learning approach and experimentally confirmed in another study
142 of ours[15].
143
144 To validate our screening results, we used biolayer interferometry (BLI) to measure the binding
145 activities of selected antibody candidates (**Figure 3A**), including one targeting H1 stem and
146 five targeting H3 stem. We also selected two antibodies that had low enrichment in both H1
147 stem and H3 stem screens as negative controls. The binding activities of these antibody
148 candidates were successfully confirmed using both fragment antigen-binding (Fab) and scFv

149 formats (**Figure 3B-C, Figure S3 and S4)**. Together, our results showed that oPool$^+$ display
150 enabled rapid discovery of natively paired HA stem antibodies.
151
152 **AG11-2F01 and 16.ND.92 have similar sequence features but distinct binding modes**
153 The H1 stem antibody identified from our screen was 16.ND.92, which was originally isolated
154 from a young individual in an H5N1 influenza vaccine trial[7]. 16.ND.92 was encoded by IGHV3-
155 74/IGHD3-3/IGKV1-5 (**Table S3**). Coincidentally, both IGHD3-3 and IGKV1-5 were utilized by
156 AG11-2F01, which was one of the two positive controls against H1 stem in our screen (**Figure**
157 **2B and 3B**). Moreover, 16.ND.92 and AG11-2F01 shared a similar FG[V/L] motif encoded by
158 the reading frame +3 of IGHD3-3 (**Figure S5**). This observation led us to hypothesize that
159 16.ND.92 and AG11-2F01 engaged the HA stem via similar binding modes. Consequently, we
160 determined the cryo-EM structures of H1N1 A/Solomon Islands/03/2006 (SI/06) HA in complex
161 with AG11-2F01 and 16.ND.92 to resolutions of 2.89 Å and 2.82 Å, respectively (**Figure 4A-**
162 **E and Table S4**). Contrary to our hypothesis, the cryo-EM structures revealed very different
163 binding modes between the two antibodies. While AG11-2F01 bound to HA stem horizontally,
164 16.ND.92 had a downward approaching angle towards HA (**Figure 4A**). Relatedly, the epitope
165 of 16.ND.92 shifted slight upward compared to that of AG11-2F01 (**Figure 4B**).
166
167 Although both 16.ND.92 and AG11-2F01 were encoded by IGKV1-5, their light chains
168 interacted with the HA stem differently. For example, V_L S30 of AG11-2F01 H-bonded with
169 HA2 Q38, whereas that of 16.ND.92 H-bonded with HA1 K32 (**Figure 4C**). Similarly, despite
170 sharing an FG[V/L] motif in their IGHD3-3-encoded regions, 16.ND.92 and AG11-2F01 used
171 this motif to interact with different parts of the HA stem (**Figure 4D**). For the FGL motif in AG11-
172 2F01, V_H F100 inserted into a hydrophobic pocket in the HA stem centering at HA2 I48,
173 whereas V_H L100b inserted into a lower pocket centering at HA2 W21. By contrast, this lower
174 pocket was occupied by the V_H F100a of the FGV motif in 16.ND.92. The paratope of 16.ND.92
175 also involved IGHD3-3-encoded V_H V100, I100e, and I100f, allowing its CDR H3 to bind to
176 upper pockets in the HA stem that were not engaged by AG11-2F01 (**Figure 4D**). Together,

177 our structural analyses showed that AG11-2F01 and 16.ND.92 formed distinct molecular
178 interactions with HA stem.
179
180 **16.ND.92 utilizes IGHD3-3 in a unique manner for binding to HA stem**
181 Previous studies have determined the structures of several HA stem antibodies with an
182 FG[V/L/I] motif in the CDR H3 that was encoded by reading frame +3 of IGHD3-3, including
183 MEDI8852, 56.a.09, 54-1G05, 39.29, PN-SIA28, and 429 B01[7,9,33–36]. These six HA stem
184 antibodies used either IGHV6-1 or IGHV3-30, unlike AG11-2F01 and 16.ND.92, which used
185 IGHV4-38-2 and IGHV3-74, respectively. Nevertheless, the CDR H3 conformation of AG11-
186 2F01 resembled that of MEDI8852, 56.a.09, 54-1G05, 39.29, PN-SIA28, and 429 B01 (**Figure**
187 **4E**). Moreover, the IGHD3-3-encoded FG[V/L/I] motifs of these seven antibodies bound to the
188 HA stem in a similar fashion (**Figure 4E**). In comparison, the CDR H3 conformation of
189 16.ND.92 was different due to more extensive involvement of IGHD3-3 in binding (**Figure 4D-**
190 **E**).
191
192 The unique usage of IGHD3-3 for binding enables 16.ND.92 V_H to interact with the HA stem
193 exclusively through CDR H3, whereas the V_H paratopes of other IGHD3-3 HA stem antibodies
194 involved non-CDR H3 regions (**Figure 4F and Table S5**). Similarly, IGHD3-3 accounted for
195 98.6% of the buried surface area of the 16.ND.92 V_H paratope, but 38% to 63% of the V_H
196 paratopes of other IGHD3-3 HA stem antibodies (**Figure 4G and Table S5**). These
197 observations not only substantiated that reading frame +3 of IGHD3-3 was a recurring
198 sequence feature of HA stem antibodies, but also demonstrated that it could pair with diverse
199 IGHV genes and interact with HA stem via different binding modes.
200
201 **AG11-2F01 and 16.ND.92 are neutralizing antibodies with *in vivo* protection activity**
202 Given the different binding modes of AG11-2F01 and 16.ND.92, we further aimed to compare
203 their binding breath, *in vitro* neutralization, and *in vivo* protection activity. ELISA showed that
204 both AG11-2F01 and 16.ND.92 bound to all H1 and H5 HAs tested (**Figure 5A and Figure**

205 **S6)**, Microneutralization assay against six H1N1 strains further revealed their neutralizing
206 activity (**Figure 5B**). Both AG11-2F01 and 16.ND.92 also protected mice against a lethal
207 challenge of H1N1 A/Puerto Rico/8/1934 (PR8), based on the weight loss profiles (**Figure 5C-**
208 **D**), survival analyses (**Figure 5E-F**), and lung viral titers at day 3 post-infection (**Figure 5G-**
209 **H**). Nonetheless, our results indicated that the *in vivo* therapeutic protection activity of
210 16.ND.92 was stronger than AG11-2F01. While only 20% (1/5) of the mice therapeutically
211 treated with AG11-2F01 survived (**Figure 5E**), 80% (4/5) of the mice therapeutically treated
212 with 16.ND.92 survived (**Figure 5F**). Additionally, at day 3 post-infection, lung viral titers of
213 mice therapeutically treated with 16.ND.92 were ~15-fold lower than those treated with AG11-
214 2F01 (**Figure 5G-H**). Notably, 16.ND.92 had comparable, if not weaker, *in vitro* neutralizing
215 activity than AG11-2F01 against PR8 (**Figure 5B**). Consequently, the stronger *in vivo*
216 protection activity of 16.ND.92 against PR8 may at least be partly attributed to its more
217 downward approaching angle to the HA stem (**Figure 4A**), which could help position the Fc
218 region closer to effector cells.
219

220 **DISCUSSION**

221 Antibody discovery has led to significant advancements on many fronts, including antibody-
222 based therapeutics as well as vaccine designs[11,37–40]. Discovery of natively paired antibody
223 sequences has been hugely accelerated by scBCR-seq in the past few years[1]. However, going
224 from sequence information to specificity characterization remains a major bottleneck for
225 antibody discovery. In this study, we established oPool$^+$ display, an experimental platform that
226 allows specificity characterization of antibodies with defined sequences in a highly parallel
227 fashion. Importantly, oPool$^+$ display was more cost-efficient (~$30 per antibody) and faster
228 (~48-72 hours) than the conventional methods that require cloning and recombinant
229 expression of individual antibodies (~$175-275 per antibody, weeks to months) (**Table S6**). As
230 a proof-of-concept, we applied oPool$^+$ display to identify multiple HA stem antibodies from
231 hundreds of HA antibodies that were left uncharacterized in the literature. Structural analysis
232 of one of the identified HA stem antibodies, 16.ND.92, further revealed the importance of

233 IGHD3-3 in targeting the HA stem. Overall, these results provide not only a proof-of-concept
234 for oPool$^+$ display, but also molecular insights into our antibody responses against influenza
235 virus.
236
237 Previous studies have shown that IGHV6-1 and IGHV3-30 HA stem antibodies often utilize
238 IGHD3-3-encoded FG[V/L/I] motif for binding to HA stem[7,9,33–36]. As demonstrated by our work
239 here, IGHD3-3-encoded FG[V/L/I] motif can also pair with other IGHV genes to target HA stem,
240 substantiating that it is an IGHV-independent recurring sequence feature of HA stem
241 antibodies. Our results further revealed that IGHD3-3 can engage the HA stem via different
242 binding modes. These observations are comparable to those of IGHD3-9, which is utilized by
243 HA stem antibodies with various IGHV genes and can bind to HA stem in two different reading
244 frames[14]. Similarly, recent studies have identified IGHD3-22 as an IGHV-independent
245 recurring sequence feature of antibodies that target a conserved site on SARS-CoV-2
246 spike[41,42]. Although antibody sequence analysis typically focuses on IGHV genes, the
247 contribution of IGHD genes to antibody responses should not be overlooked since emerging
248 evidence suggests that it might be more important than previously thought.
249
250 A key feature of oPool$^+$ display is its relatively simple protocol. A previous study has shown
251 that the throughput for screening antibodies with defined sequences can be increased by
252 using liquid handlers to express individual antibodies one by one[43]. In comparison, oPool$^+$
253 display uses a near one-pot strategy for antibody library synthesis and screening, which has
254 higher potential for scaling. It also only requires standard benchtop equipment commonly
255 found in a regular molecular biology lab. Therefore, oPool$^+$ display can be easily integrated
256 into antibody discovery pipelines that rely on scBCR-seq[1]. After the natively paired antibody
257 sequences are obtained from scBCR-seq, oPool$^+$ display can be applied to validate and
258 characterize the specificity of a large panel of antibody candidates. Thus, this integration can
259 streamline the existing antibody discovery pipelines. Although this study focused on influenza
260 HA stem as a proof-of-concept, oPool$^+$ display can be generalized to any antigens of interest

261 as long as they can be recombinantly purified. Furthermore, the capability of constructing
262 custom-made antibody libraries means that oPool$^+$ display has the potential to benefit the
263 development of machine learning models for antibody engineering, specificity prediction and
264 *de novo* design in the future[15,44,45]. We envision that prediction results from these models can
265 be rapidly validated by oPool$^+$ display, which will in turn facilitate iterative refinement of the
266 models.
267
268 We acknowledge that oPool$^+$ display has some limitations. First, oPool$^+$ display requires
269 antibodies to be presented as scFv, which may lose functionality compared to its Fab
270 counterpart[46–48]. Second, antibodies with a fast off-rate may result in false negatives in mRNA
271 display selection, since it depends on monovalent binding. Nonetheless, this caveat can be
272 addressed by replacing mRNA display with other protein display technologies that support
273 multivalent binding, such as yeast display and phage display[49,50]. As the length of oligo pool
274 synthesis continues to improve, the cost and complexity of oPool$^+$ display will further decrease.
275 Besides, by including more antibodies in a given oligo pool, which is feasible without
276 compromising the assembly accuracy as demonstrated in this study, we can increase the
277 throughput of the antibody library construction. Moreover, the scale of antibody library
278 construction can potentially be increased by leveraging multiplexed gene synthesis
279 strategies[26,51]. Overall, we believe that oPool$^+$ display represents a starting point for the future
280 development of high-throughput approaches to screen antibodies with defined sequences.
281
282 **MATERIALS AND METHODS**
283 **Selection of HA antibodies for paired antibody library synthesis**
284 Members of the natively paired antibody library were selected from a previously curated
285 dataset containing 5,561 human monoclonal antibodies to influenza HA from 60 research
286 publications and three patents[15]. Filters were applied to exclude antibodies that 1) had
287 incomplete sequence information, 2) utilized germline genes that were regarded as recurring
288 sequence features of HA stem antibodies, namely IGHV1-69, IGHV6-1, IGHV1-18 and

289 IGHD3-9[7,12–14], and 3) were members of known HA stem antibody clonotypes. This resulted in
290 292 antibody sequences from 7 publications[7,16–18,23,52,53]. Three HA stem antibodies, namely
291 31.a.55, AG11-2F01, and 042-100809_2F04, as well as 30 HA head antibodies were randomly
292 selected as positive and negative controls[6,7,16–25]. Of note, both AG11-2F01 and AG2-G02
293 were not previously labeled as an HA stem antibody in the curated dataset[15]. Through
294 literature search, AG11-2F01 was found to compete with CR9114[10], which is an HA stem
295 antibody, for binding to H1[16], while 042-100809-2F04 was found to bind to group 2 HA stem
296 domain[54].
297
298 **Computational design of the oligo pool sequences for assembly**
299 A summary of the computational design pipeline is described below and summarized in **Figure**
300 **S1**.
301
302 *Sequence preparation*
303 Selected antibody sequences were first annotated using abYsis[55]. Any missing nucleotides at
304 the 5' and/or 3' ends were then filled in using the sequences from the IGHV and IGHJ genes
305 that had the highest identity with the given antibody.
306
307 *Codon randomization and pool assignment of the antibodies*
308 To decrease undesired assembly during antibody library synthesis, codon randomization of
309 each selected antibody sequence was first performed to reduce nucleotide sequence similarity
310 among different antibodies. For a given amino acid, codon usages <15% in *Escherichia coli*
311 were removed from consideration to prevent low translation efficiency during RNA display. For
312 325 antibodies, total of 2 million randomized sequences were generated to maximize
313 downstream sequence differentiation. All antibody sequences were then split into 8 segments,
314 followed by clustering using CD-HIT[56]. The clusters were generated using the criterion of 70%
315 sequence identity (at least 30% differences between each cluster). The antibody sequences
316 were then reconstructed by selecting necessary segments from different clusters, followed by

317 the deletion of the used clusters. Such reconstruction was repeated until a complete pool (total
318 of 25 antibody sequences) had been reassembled. The deleted clusters were then added
319 back with the reassembled antibody sequences removed. The process was then iterated to
320 generate the remaining pools.
321
322 *Selection of overlap regions and generation of the final oligo pools*
323 Upon the pool assignment of antibodies, the ideal overlap regions were searched over CDR
324 L3, H1, and H3 of each antibody sequence. Six 30-nucleotide long sequences were extracted
325 from each region through frame shifting, then aligned to the complete antibody sequences in
326 the corresponding pool using BLAST+[57]. For each antibody, the overlap sequences that are
327 least similar to other antibodies in the pool were selected. The antibody sequences were then
328 split at the overlap region, followed by the addition of universal adaptor regions at fragments
329 encoding the N-terminal and C-terminal of the antibodies, leading to the generation of the final
330 oligo pools.
331
332 **Overlap PCR assembly of the natively paired antibody library**
333 A total of 13 oligo pools were synthesized (Integrated DNA Technologies). The lengths of
334 oligos ranged from around 180 to 330 nucleotides. Each oligo pool contained 100 oligos
335 resuspended in 200 µL water. An assembly PCR was set up for each oligo pool using 1,600
336 ng of oligos as input. The assembly was performed using KAPA HiFi HotStart ReadyMix
337 (Roche) and a Mastercycler nexus GX2 (Eppendorf). PCR was set up in the absence of
338 primers. From cycles 1-40, PCR was performed with minimal ramp rate (0.1°C/s) in between
339 the denaturing (98°C, 20 s) and annealing steps (62°C, 15 s) to reduce erroneous annealing
340 events. After cycle 40, a universal forward primer 5'-TTC <u>TAA TAC GAC TCA CTA TAG</u> GGA
341 CAA TTA CTA AAG GAG TAT CC-3' and a universal reverse primer 5'-GGA GCC GCT ACC
342 <u>CTT ATC GTC GTC ATC CTT GTA ATC</u> GGA TCC T-3' were added to the PCR. The underlined
343 region in the forward primer sequence is the T7 promoter, whereas the underlined region in
344 the reverse primer sequence encodes a FLAG tag. Subsequently, another 15 cycles of PCR

345 were performed to amplify the assembled product. The final PCR product was purified using
346 a Monarch Gel Extraction Kit (New England Biolabs). Two replicates of the assembly were
347 performed.
348
349 **Preparation of the biotinylated H1 stem and H3 stem constructs**
350 The H1 stem (mini-HA #4900)[31] and H3 stem (H3ssF)[32] constructs were cloned into a
351 customized baculovirus transfer vector. Both constructs contained a N-terminal gp67 signal
352 peptide at the N-terminus as well as a BirA biotinylation site, a thrombin cleavage site, a
353 trimerization domain and a 6×His-tag at the C-terminus. Recombinant bacmid DNA that
354 carried the H1 stem construct or H3 stem construct was generated using the Bac-to-Bac
355 system (Thermo Fisher Scientific) per manufacturer's instructions. Baculovirus was generated
356 by transfecting the purified bacmid DNA into adherent Sf9 cells using Cellfectin reagent
357 (Thermo Fisher Scientific) per manufacturer's instructions. The baculovirus was further
358 amplified by passaging in adherent Sf9 cells at a multiplicity of infection (MOI) of 1.
359
360 Recombinant H1 stem and H3 stem constructs were then expressed using 1L of suspension
361 Sf9 cells at an MOI of 1. At day 3 post-infection, Sf9 cells were pelleted by centrifugation at
362 4,000 ×g for 25 min. Soluble recombinant H1 stem and H3 stem constructs were purified from
363 the supernatant by affinity chromatography using Ni Sepharose excel resin (Cytiva) and then
364 size exclusion chromatography using a HiLoad 16/100 Superdex 200 prep grade column
365 (Cytiva) in 20 mM Tris-HCl at pH 8.0, and 100 mM NaCl. The purified protein was concentrated
366 by an Amicon spin filter (Millipore Sigma) and filtered by a 0.22 mm centrifuge Tube Filter
367 (Costar). The purified H1 stem and H3 stem constructs were then biotinylated using a Biotin-
368 Protein Ligase-BIRA kit (Avidity) according to the manufacturer's instructions. The biotinylated
369 proteins were then purified again through size exclusion chromatography as described above.
370 The A280 absorbance values were measured using a Nanodrop One (Thermo Fisher
371 Scientific) to quantify the protein concentration.
372

373 **Antibody screening using mRNA display**

374 The mRNA display was performed based on the protocols provided by previous studies[29,30,58]
375 with slight modifications.

376

377 *Generation of the puromycin-conjugated mRNA templates*

378 The DNA library was first transcribed by a MEGAscript T7 Transcription Kit (Thermo Fisher
379 Scientific) and purified by a MEGAclear Transcription Clean-Up Kit (Thermo Fisher Scientific)
380 according to manufacturer's instructions. Ligation was performed using 1 nmol of the mRNA
381 product, 1.1 nmol of the splint oligo (5'-TTT TTT TTT TTT GGA GCC GCT ACC-3'), and 1.2
382 nmol of the puromycin linker (5'-/5Phos/-d(A)21-(C_9)3-d(ACC)-puromycin-3') by the T4 DNA
383 ligase (New England Biolabs) in a 100 μL reaction for 1 hour at room temperature, followed
384 by Lambda exonuclease (New England Biolabs) digestion for 30 mins at 37°C. The
385 puromycin-conjugated mRNA product was purified using a Dynabeads mRNA DIRECT
386 Purification Kit (Thermo Fisher Scientific), aliquoted, and stored at -20°C until used.

387

388 *Preparation of the mRNA-scFv fusion library*

389 The puromycin-conjugated mRNA templates were translated using a PURExpress In Vitro
390 Protein Synthesis Kit (New England Biolabs) with the addition of PURExpress Disulfide Bond
391 Enhancer (New England Biolabs) for 1 hour at 37°C. The reaction was then incubated with
392 500 mM KCl and 60 mM $MgCl_2$ for at least 30 mins at room temperature to promote fusion
393 between the translated scFv and puromycin. EDTA was then added to dissociate ribosomes.
394 The full-length mRNA-scFv product was then purified using Anti-FLAG M2 Magnetic Beads
395 (Millipore Sigma) followed by elution using 3×FLAG peptides (GlpBio). Subsequently, the
396 purified mRNA-scFv product was reverse transcribed using SuperScript IV reverse
397 transcriptase (Thermo Fisher Scientific). The cDNA/mRNA-scFv product was referred as the
398 "pre-selection library".

399

400 *Preparation of the magnetic beads coated with H1 stem or H3 stem*

401 Biotinylated H1 stem construct or H3 stem construct was coated onto the Dynabeads M280-
402 straptavidin (Thermo Fisher Scientific) according to the manufacturer's instruction. Briefly, 150
403 pmol of biotinylated proteins were incubated with 50 μL of the beads for 30 mins to 1 hour at
404 room temperature with gentle rotation. The beads were then washed with TBST (20 mM Tris-
405 HCl at pH 7.5, 100mM NaCl, and 0.025% Tween-20) five times using the DynaMag-2 magnetic
406 holder (Thermo Fisher Scientific) and then resuspended to the original volume.
407
408 *Antibody selection against H1 stem or H3 stem*
409 Selection of antibodies against H1 stem construct and H3 stem construct were carried out in
410 parallel. Briefly, the pre-selection library was mixed with 25 μL of beads coated with H1 stem
411 construct or H3 stem construct and incubated for 1 hour at room temperature with gentle
412 rotation. After incubation, the beads were washed thrice with 400 μL TBST. The beads were
413 then resuspended in water. These samples were referred as the "post-selection libraries".
414
415 **Next-generation sequencing of the scFv library**
416 The pre-selection libraries, post-selection libraries, and the small pool assemblies selected for
417 quality control were amplified using PrimeSTAR Max DNA Polymerase (Takara Bio) per
418 manufacturer's instruction with the following primers (5'-GTA AAA CGA CGG CCA GTT TCA
419 GGG GAC AAT TAC TAA AGG AGT ATC C-3' and 5'- CAG GAA ACA GCT ATG ACC CAC
420 TCG TCA TCC TTG TAA TCG GAT CCT CCG GA-3'. The PCR product was purified using a
421 Monarch Gel Extraction Kit (New England Biolabs). A second round of PCR was carried out
422 to add the adapter sequence and index to the amplicons (**Table S7**). The final PCR products
423 were sequenced on one SMRTcell 8M on a PacBio Revio system using the CCS sequencing
424 mode and a 30-hour movie time.
425
426 **Analysis of next-generation sequencing data**
427 Circular consensus sequences (CCSs) were generated from the raw subreads using
428 SMRTLink v13.0, setting the parameters to require 99.9% accuracy and a minimum of 3

429 passes. CSSs in FASTQ format were parsed using the SeqIO module in BioPython[59] and
430 filtered based on the base calling quality score, where any read with more than five nucleotides
431 of phred quality score <40 were removed. The adapter sequences were then identified on
432 each read and trimmed from the scFv sequences. Reads that did not have the complete
433 adapter sequences were also removed. The filtered reads were then aligned to the reference
434 scFv sequences and classified into three categories: 1) natively paired scFvs with no mutation,
435 2) natively paired scFvs with mutation, 3) others (non-natively paired scFvs and incomplete
436 assemblies). Only the reads encoding natively paired scFvs with no mutation were used for
437 downstream analysis. Frequency (*F*) of a scFv *i* of a given sample *s* was computed for each
438 replicate as follows:

439 $$F_{i,s} = \frac{readcount_{i,s}+1}{\sum_s(readcount_{i,s}+1)} \quad (1)$$

440 A pseudocount of 1 was added to each mutant to avoid division by zero in subsequent steps.
441 We then calculate the enrichment (*E*) of a scFv *i* of a given replicate *k* after the mRNA display
442 selection as follows:

443 $$E_{i,k} = \frac{F_{post-selection,i,k}}{F_{pre-selection,i,k}} \quad (2)$$

444

445 **Expression and purification of CR9114-FLAG scFv**

446 CR9114 scFv nucleotide sequence with a pelB secretion peptide at the N-terminal and a FLAG
447 (DYKDDDK) tag followed by a stop codon at the C-terminal were synthesized (Integrated DNA
448 Technologies) and ligated into a pET-28a plasmid vector backbone using NEBuilder HiFi DNA
449 Assembly Master Mix (New England Biolabs). The ligated product was then transformed into
450 DH5α competent cells (Thermo Fisher Scientific), and the plasmids were extracted using a
451 QIAprep Spin Miniprep Kit (Qiagen). The plasmids were then transformed into BL21(DE3)
452 competent cells (Thermo Fisher Scientific), followed by overnight growth at 37°C, 225 RPM
453 shaking. 100mL LB culture (Fisher Scientific) were inoculated at 1:500 ratio using the
454 overnight culture and shaked at 37°C, 225 RPM. Once the OD_{600} reached approximately 0.6,

455 IPTG (1mM final concentration) was added to induce protein expression for 16hrs at 30°C,
456 200 RPM.
457
458 The overnight culture was centrifuged at 4,500 ×g, 4°C for 1 hour to remove the supernatant.
459 The pellet was resuspended using 2 mL of ice-cold 1× TES buffer (200 mM Tris-HCl at pH 8.0,
460 0.65 mM EDTA, and 0.5 M sucrose). The resuspended mixture was incubated at 4°C for 2
461 hours with gentle shaking. 5 mL of ice-cold 0.25× TES buffer was then added, followed by
462 incubation at 4°C overnight. The mixture was centrifuged at 4,500 ×g, 4°C for 1 hour to
463 remove the pellet. The supernatant was subsequently clarified using a polyethersulfone
464 membrane filter with a 0.22 mm pore size (Millipore) and purified using ANTI-FLAG M2 Affinity
465 Gel (Millipore-Sigma) per manufacturer's instructions. The A280 was measured using the
466 Nanodrop One (Thermo Fisher) to calculate the sample concentration.
467
468 **Expression, purification, and quantitation of selected scFvs for validation**
469 Nucleotide sequences of selected scFvs with a T7 promoter at the N-terminal and a FLAG
470 (DYKDDDK) tag followed by a stop codon at the C-terminal were synthesized (Integrated DNA
471 Technologies) and amplified by PCR using Prime STAR Max DNA polymerase (Takara Bio).
472 The PCR product was purified using a PureLink PCR purification kit (Thermo Fisher Scientific)
473 and used as the template for *in vitro* translation using a PURExpress In Vitro Protein Synthesis
474 Kit (New England Biolabs) with the addition of PURExpress Disulfide Bond Enhancer (New
475 England Biolabs). The translated scFvs were then reverse purified using Pierce High-Capacity
476 Ni-IMAC magnetic beads (Thermo Fisher Scientific) per manufacturer's instructions to remove
477 all His-tagged translation kit components.
478
479 To measure the concentration of the purified scFv, quantitation assays were performed by
480 biolayer interferometry (BLI) using an Octet Red instrument (ForteBio). Briefly, rat anti-FLAG
481 tag monoclonal antibody (L5) (Thermo Fisher Scientific) at 5 μg/mL in 1× kinetics buffer (1×
482 PBS at pH 7.4, and 0.002% v/v Tween 20) were loaded onto ProG biosensors (Sartorius),
483 then incubated with the 40× diluted scFv sample (5 μL of sample added to 195 μL of 1×

484 kinetics buffer). The standard curve was generated via 2-fold serial dilutions using CR9114-
485 FLAG scFv. The assay consisted of five steps: (1) baseline: 60 s with 1× kinetics buffer; (2)
486 antibody capture: 180 s with rat anti-FLAG antibody; (3) baseline: 60 s with 1× kinetics buffer;
487 (4) binding rate measurement: 120 s with standard and scFv samples; and (5) regeneration:
488 5s in regeneration buffer (0.1 M Glycine at pH 3.0) followed by 5 s in neutralization buffer (1M
489 Tris-HCl at pH 7.5), repeated 3 times. The data were analyzed using Octet analysis software
490 9.0, where the first 30 s of the binding rate measurement were used for final quantitation.
491
492 **Expression and purification of Fabs and IgGs**
493 Heavy and light chains of the antibodies were cloned into a phCMV3 vector with a mouse
494 immunoglobulin kappa signal peptide in human IgG1 Fc or Fab format. Plasmids encoding the
495 heavy and light chains of antibodies were transfected into Expi293F cells using an
496 ExpiFectamine 293 transfection kit (Gibco) in a 2:1 mass ratio for IgG or a 1:1 mass ratio for
497 Fab following the manufacturer's protocol. Supernatant was harvested at 6 days post-
498 transfection and centrifuged at 4,000 ×g for 30 mins at 4°C to remove cells and debris. The
499 supernatant was subsequently clarified using a polyethersulfone membrane filter with a 0.22
500 mm pore size (Millipore). Antibodies were first purified by CaptureSelect CH1-XL beads
501 (Thermo Fisher Scientific). Then, the antibodies were further purified by size exclusion
502 chromatography using a HiLoad 16/100 Superdex 200 prep grade column (Cytiva) in 1× PBS.
503 The A280 were measured using the Nanodrop One (Thermo Fisher) to calculate the sample
504 concentration. Antibodies were stored at 4°C until used.
505
506 **Measurement of binding kinetics using biolayer interferometry**
507 *Fab K_D measurement*
508 The binding assay was performed by biolayer interferometry (BLI) using an Octet Red
509 instrument (ForteBio). Briefly, biotinylated H1 or H3 stem construct at 0.5 μM in 1× kinetics
510 buffer (1× PBS at pH 7.4, and 0.002% v/v Tween 20) were loaded onto SA biosensors
511 (Sartorius) and incubated with 33 nM, 100 nM, and 300 nM of purified Fabs. The assay

512 consisted of five steps: (1) baseline: 60 s with 1× kinetics buffer; (2) loading: 120 s with
513 biotinylated HA stem domains; (3) baseline: 60 s with 1× kinetics buffer; (4) association: 120
514 s with Fab samples; and (5) dissociation: 120 s with 1× kinetics buffer. For estimating the
515 exact K_D, a 1:1 binding model was used.
516
517 *scFv K_D measurement*
518 The binding assay was performed by biolayer interferometry (BLI) using an Octet Red
519 instrument (ForteBio). Briefly, biotinylated H1 or H3 stem construct at 0.5 μM in 1× kinetics
520 buffer (1× PBS at pH 7.4, and 0.002% v/v Tween 20) were loaded onto SA biosensors
521 (Sartorius) and incubated with 9× dilution (20 μL of sample added to 160uL of 1× kinetics
522 buffer) and 18× dilution (10 μL of sample added to 170 μL of 1× kinetics buffer) of the purified
523 scFv sample. The assay consisted of five steps: (1) baseline: 60 s with 1× kinetics buffer; (2)
524 loading: 180 s with biotinylated HA stem domains; (3) baseline: 60 s with 1× kinetics buffer;
525 (4) association: 120 s with Fab samples; and (5) dissociation: 120 s with 1× kinetics buffer.
526 For estimating the exact K_D, a 1:1 binding model was used.
527
528 **Expression and purification of HA ectodomains**
529 The HA ectodomains of H1N1 A/Puerto Rico/8/1934, H1N1 A/Beijing/262/1995, H1N1
530 A/Solomon Island/3/2006 (SI/06), H1N1 A/Michigan/45/2015, H1N1 A/Brisbane/02/2018 were
531 cloned, expressed, and purified as mentioned above for the H1 and H3 stem constructs. The
532 HA ectodomains of H1N1 A/California/04/2009 (NR-15749), H5N1 A/bald eagle/Florida/W22-
533 134-OP/2022 (NR-59476), H5N2 A/snow goose/Missouri/CC15-84A/2015 (NR-50651), and
534 H5N8 A/northern pintail/WA/40964/2014 (NR-50174) were obtained from BEI Resources
535 (https://www.beiresources.org/). The HA ectodomains of H1N1 A/USSR/90/1977 and H1N1
536 A/Taiwan/01/1986 were purchased from SinoBiological.
537
538 **Cryo-EM sample preparation, data collection, and data processing**
539 *AG11-2F01 Fab in complex with SI/06 HA*

540 The AG11-2F01 Fab was incubated with SI/06 HA on ice overnight followed by size exclusion
541 chromatography. The peak fraction of the Fab-HA complex was concentrated to around 1
542 mg/mL for cryo-EM sample preparation. Cryo-EM grids were prepared using a Vitrobot Mark
543 IV (Thermo Fisher Scientific). 3.5 µL of the sample was applied to a 300-mesh Quantifoil
544 R1.2/1.3 Cu grid pretreated with glow-discharge. Excess liquid was blotted away using filter
545 paper with blotting force -5 and blotting time 3 s. The grid was then flash frozen in liquid ethane.
546 Movies were then collected on a Titan Krios microscope equipped with Gatan BioQuantum K3
547 imaging filter and camera (Thermo Fisher Scientific). Images were recorded at 130,000×
548 magnification, corresponding to a pixel size of 0.33 Å/pixel at super-resolution mode of the
549 camera. A defocus range of -0.8 µm to -1.5 µm was set. A total dose of 50 $e^-/Å^2$ of each
550 exposure was fractionated into 50 frames. Both untilted and 30-degree-tilted data were
551 collected and combined to alleviate the preferred orientation problem of the sample.
552
553 CryoSPRAC[60] was used to process the cryo-EM data. For model building, ABodyBuilder[61] was
554 used to generate an initial model for AG11-2F01 Fab. This model, together with the model of
555 SI/06 HA (PDB 6FYT)[62], was fitted into the cryo-EM density map using UCSF Chimera[63]. The
556 model was manually adjusted in Coot[64] and refined with Phenix real-space refinement
557 program[65]. This process was iterated for several cycles until no significant improvement of the
558 model was observed.
559
560 *16.ND.92 Fab in complex with SI/06 HA*
561 The 16.ND.92 Fab was incubated with SI/06 HA and FISW84 Fab, a known HA anchor
562 antibody[66], on ice overnight followed by size exclusion chromatography. The peak fraction of
563 the Fab-HA complex was concentrated to around 3 mg/mL for cryo-EM sample preparation.
564 0.1% (w/v) of n-octyl-ß-D-glucoside was added to reduce orientation bias. Cryo-EM grids were
565 prepared using a Vitrobot Mark IV (Thermo Fisher Scientific). 3 µL of the sample was applied
566 to a 400-mesh Quantifoil R1.2/1.3 Cu grid pretreated with glow-discharge. Excess liquid was
567 blotted away using filter paper with blotting force 0 and blotting time 3 s. The grid was then

568 flash frozen in liquid ethane. Movies were then collected on a Titan Krios microscope equipped
569 with Gatan BioQuantum K3 imaging filter and camera (Thermo Fisher Scientific). Images were
570 recorded at 81,000× magnification, corresponding to a pixel size of 0.53 Å/pixel at super-
571 resolution mode of the camera. A defocus range of -0.5 μm to -5 μm was set. A total dose of
572 57.35 $e^-/Å^2$ of each exposure was fractionated into 40 frames.
573
574 CryoSPRAC[60] was used to process the cryo-EM data. DeepEMhancer[67] was used to generate
575 the sharpened density map for downstream model building. For model building, IgFold[68] was
576 used to generate an initial model for 16.ND.92 Fab. This model, together with the model of
577 SI/06 HA (PDB 6FYT)[62], was fitted into the cryo-EM density map using Phenix DockinMap
578 module[69]. The models were manually adjusted in Coot[64] and refined with Phenix real-space
579 refinement program[65]. This process was iterated for several cycles until no significant
580 improvement of the model was observed.
581
582 **Structural analysis of HA-antibody complexes**
583 Buried surface areas upon binding and paratope residues of AG11-2F01, 16.ND.92,
584 MEDI8852[9] (PDB 5JW4), 56.a.09[7] (PDB 5K9J), 54-1G05[33] (PDB 6WIZ), PN-SIA28[35] (PDB
585 8GV5), 39.29[36] (PDB 4KVN), and 429 B01[34] (PDB 6NZ7) were analyzed using PDBePISA[70].
586 The CDRH3 region and IGHD 3-3 usage of each antibody was annotated using IgBLAST[71].
587 The molecular interactions of AG11-2F01 and 16.ND.92 in complex with SI/06 HA were
588 analyzed and visualized using PyMOL (Schrödinger).
589
590 **Enzyme-linked immunosorbent assay (ELISA)**
591 Nunc MaxiSorp ELISA plates (Thermo Fisher Scientific) were coated overnight at 4°C with
592 100 μL of recombinant proteins at 1 μg/mL in 1× PBS. The next day, plates were washed
593 thrice with 1× PBS containing 0.05% Tween 20 and blocked with 200 μl of 5% non-fat milk in
594 1× PBS for 2 hours at room temperature. Monoclonal antibodies were serially diluted 10-fold
595 starting from 100 μg/mL, added to the plates, and incubated for 2 hours at 37°C. Plates were

596 then washed thrice and incubated with horseradish peroxidase (HRP)-conjugated goat anti-
597 human IgG antibody (Thermo Fisher Scientific) at 1:5,000 dilution for 1 hour at 37°C. After six
598 washes with 1× PBS containing 0.05% Tween 20, 100 μL of 1-Step TMB ELISA Substrate
599 Solution (Thermo Fisher Scientific) was added to each well. After incubation for 10 mins, the
600 reaction was stopped with 50 μL of 2 M H_2SO_4 solution, and absorbance values were
601 measured at 450 nm using a BioTek Synergy HTX Multimode Reader (Agilent).
602
603 **Recombinant virus construction and purification**
604 Recombinant viruses with HA and NA segments from the indicated H1N1 strains and six
605 internal segments from H1N1 A/Puerto Rico/8/1934 (PR8) were obtained from BEI Resources
606 (https://www.beiresources.org/). Recombinant viruses were rescued using the eight-plasmid
607 reverse genetics system[72]. Briefly, plasmids encoding the HA segments from H1N1
608 A/California/07/2009 and H1N1 A/Michigan/45/2015 along with seven plasmids encoding the
609 other seven segments from PR8 were transfected into a co-culture of HEK293T (human
610 embryonic kidney) cells and MDCK-SIAT1 (Madin-Darby Canine Kidney) cells at a 6:1 ratio.
611 Supernatants were injected into 8-10 days old embryonated chicken eggs and incubated at
612 37°C for 48 hours. Viruses in the allantoic fluid were plaque-purified on MDCK-SIAT1 cells
613 grown in Dulbecco's Modified Eagles Medium (Gibco) containing 10% fetal bovine serum
614 (Gibco) and a penicillin-streptomycin mix (100 U/mL penicillin and 100 μg/mL streptomycin,
615 Gibco). The HA sequence of each virus was confirmed by Sanger sequencing.
616
617 **Microneutralization assay**
618 For the microneutralization assay, MDCK-SIAT1 cells were seeded in 96-well plates. After
619 reaching 100% confluency, MDCK-SIAT1 cells were washed once with 1× PBS. Minimal
620 essential media (Gibco) containing 25 mM HEPES (Gibco) was then added to the cells.
621 Monoclonal antibodies were serially diluted 10-fold starting from 100 μg/ml and mixed with
622 100 $TCID_{50}$ (median tissue culture infectious dose) of viruses at equal volume and incubated
623 at 37°C for 1 hour. Subsequently, the mixture was inoculated into cells and incubated at 37°C
624 for another hour. Cell supernatants were discarded and replaced with minimal essential media

625 containing 25 mM HEPES, and 1 μg/mL TPCK-trypsin (Sigma). Plates were incubated at 37°C
626 for 72 hours, and virus presence was detected by hemagglutination assay to determine the
627 MN_{50} titers.
628
629 **Mice**
630 The animal experiments were performed in accordance with protocols approved by UIUC
631 Institutional Animal Care and Use Committee (IACUC). Six-week-old female BALB/c mice
632 (Jackson Laboratory) were used for all animal experiments.
633
634 **Prophylactic and therapeutic protection experiments**
635 Female BALB/c mice at 6 weeks old (n = 5 per group) were anesthetized with isoflurane and
636 intranasally infected with 5× lethal dose (LD_{50}) of recombinant PR8 virus. Mice were given the
637 indicated antibody at a dose of 5 mg/kg intraperitoneally at 4 hours before infection
638 (prophylaxis) or 4 hours after infection (therapeutics). Weight loss was monitored daily for 14
639 days. The humane endpoint was defined as a weight loss of 25% from initial weight at day 0.
640 Of note, while our BALB/c mice were not modified to facilitate interaction with human IgG1,
641 human IgG1 could interact with mouse Fc gamma receptors[73–75].To determine the lung viral
642 titers at day 3 post-infection, lungs of infected mice were harvested and homogenized in 1 mL
643 of minimal essential media with 10% bovine serum albumin using a gentleMACS C Tube
644 (Miltenyi Biotec). Subsequently, virus titers were measured by $TCID_{50}$ assay.
645
646 **ACKNOWLEDGEMENTS**
647 We thank Anders Olson, Chris Brooke, David Kranz, and Steve Sligar for helpful discussion
648 and the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign
649 for assistance with PacBio sequencing. We thank Kristen Flatt and the Materials Research
650 Laboratory Central Research Facilities at the University of Illinois at Urbana-Champaign for
651 access to cryo-EM instrumentation during the screening of the Fab-HA complexes. We are
652 grateful to the Cryo-EM Core and the Core Facility for Advanced Research Computing at Case

Western Reserve University for the access of cryo-EM instrumentation and data processing. We also thank Frank Vago and the Cryo-EM Facility at Purdue University for the access of cryo-EM instrumentation and data collection. This work is supported by National Institutes of Health R01 AI167910 (N.C.W.), DP2 AT011966 (N.C.W.), the Searle Scholars Program (N.C.W.), Howard Hughes Medical Institute Emerging Pathogens Initiative (N.C.W.), UIUC William T. and Lynn Jackson Graduate Student Fellowship (W.O.O.), and Carl R. Woese Institute for Genomic Biology Postdoctoral Fellowship (H.L.).

AUTHOR CONTRIBUTIONS

W.O.O. and N.C.W. conceived and designed the study. W.O.O., H.L., W.L., Y.W., K.E.D. and N.C.W. developed the methodology. W.O.O. and L.T. assembled the library and performed mRNA display experiment. W.O.O., W.L. and N.C.W. analyzed the PacBio sequencing data. W.O.O., H.L., R.L., A.B.G. and L.A.R expressed and purified recombinant proteins. W.O.O., Z.M., R.L., T.P., X.D., and N.C.W. performed structural analysis of the antibodies. W.O.O., H.L., D.C., and M.R.A performed functional characterization of the antibodies. X.D. and N.C.W. provided resources and support. W.O.O. H.L., W.L. and N.C.W. wrote the paper, and all authors reviewed and edited the paper.

DECLARATION OF INTERESTS

N.C.W. consults for HeliXon. The authors declare no other competing interests.

DATA AVAILABILITY

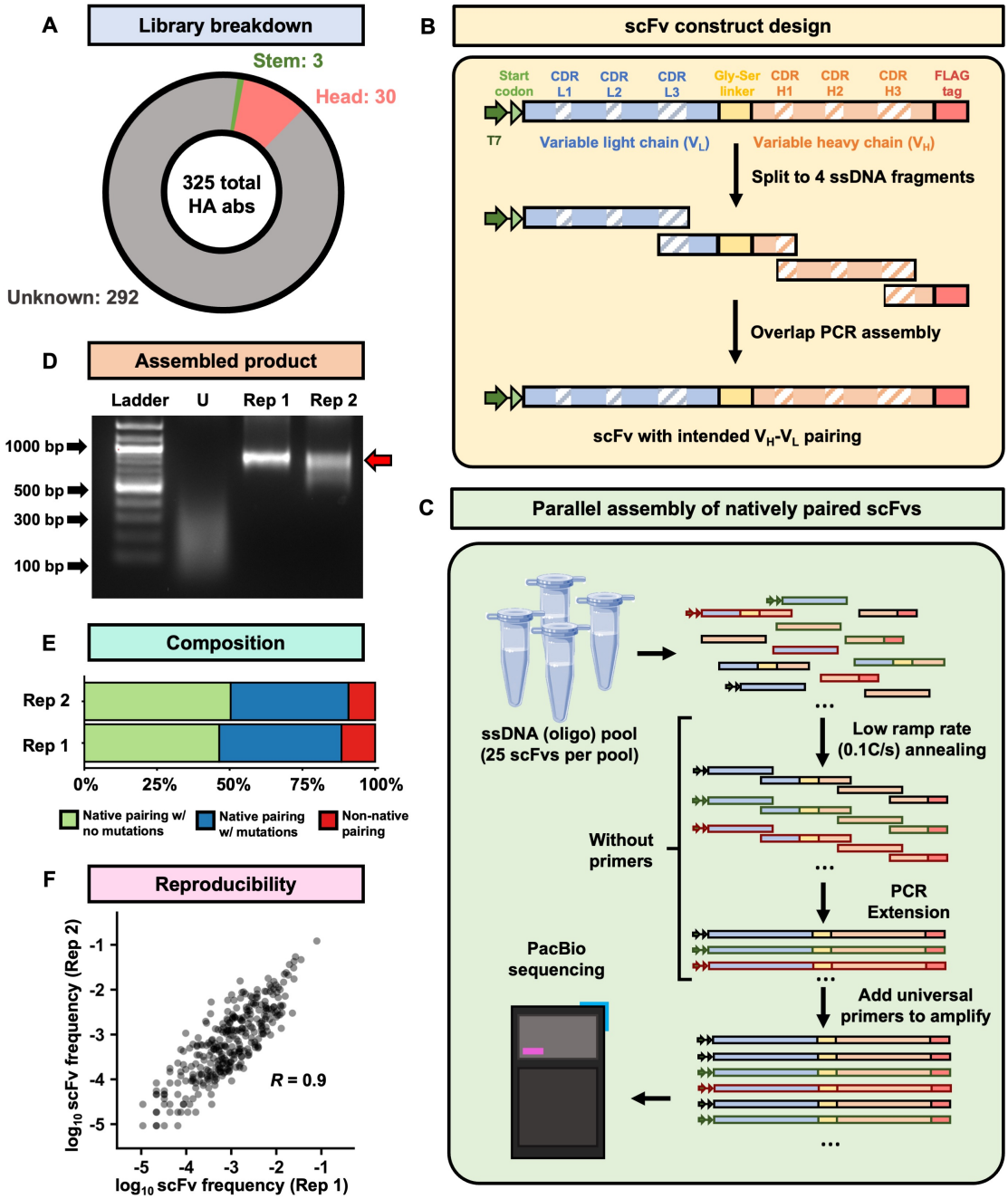
Raw sequencing data have been submitted to the NIH Short Read Archive under accession number: BioProject PRJNA1150188. Cryo-EM density maps and coordinates have been deposited to EMDB and PDB with accession numbers: EMD-46727 and EMD-45930; PDB 9DBX and 9CU7.

CODE AVAILABILITY

Custom scripts for all analyses have been deposited to:

https://github.com/nicwulab/oPool_display.

683 **MAIN FIGURES**



684

685 **Figure 1. Curation and synthesis of the natively paired HA antibody library. (A)** The
686 overall breakdown of the HA antibody library. **(B)** Design of oligos for scFv assembly. Each
687 given scFv construct contains a T7 promoter and a start codon at the N-terminal as well as a
688 FLAG tag at the C-terminal. The scFv sequences were then split into 4 fragments at the
689 selected CDR regions, with overlap between adjacent fragments. Through an overlap PCR,
690 oligos of the same construct would preferably anneal to each other, ensuring the assembly of

691 natively paired scFvs. **(C)** Synthesis of the natively paired HA antibody library. A total of 13
692 separate oligo pools (25 antibodies per pool) were synthesized, followed by the assembly,
693 which consisted of a two-stage PCR (**see Materials and Methods**). **(D)** Comparison between
694 the unassembled and assembled oligo pools were analyzed by agarose gel electrophoresis.
695 "U": unassembled oligo pool. "Rep1": replicate 1. "Rep2": replicate 2. The red arrow indicates
696 the target size (800-900 bp) for full length scFvs. **(E-F)** The composition **(E)** and reproducibility
697 **(F)** were assessed. The Pearson correlation coefficient (*R*) of the occurrence frequencies of
698 individual scFvs between the two replicates were indicated. Micro-tube icon by Servier
699 https://smart.servier.com/ is licensed under CC-BY 3.0 Unported, available via Bio Icons.

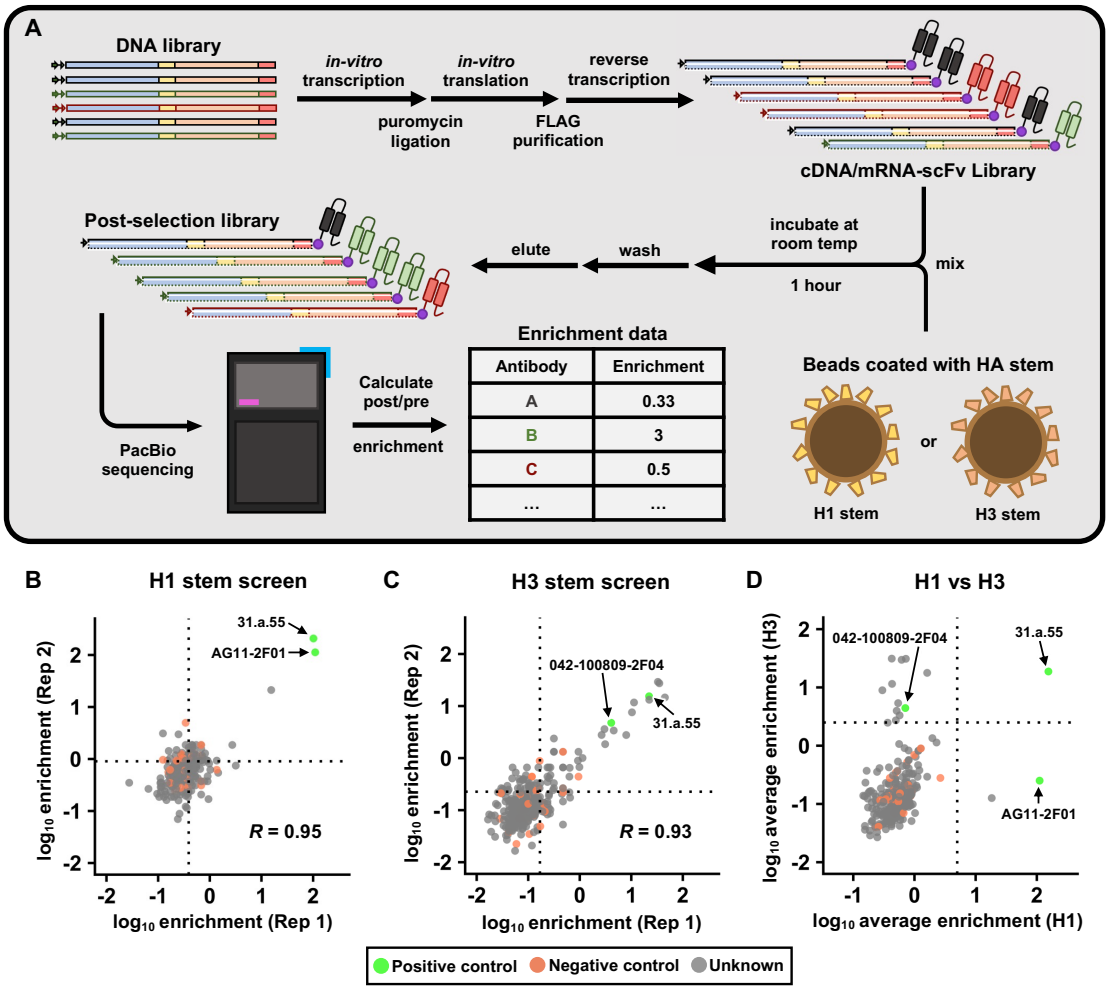


Figure 2. Rapid screening for HA stem antibodies via mRNA display. (A) Schematic overview of the mRNA display selection. Different colors represent different scFvs. The purple dot represents the puromycin. **(B-C)** Correlation between replicates of mRNA display selection against **(B)** H1 stem and **(C)** H3 stem. The green dots represent the positive controls (known HA stem antibodies), whereas the red dots represent negative controls (known HA head antibodies). The three positive controls in this study are 31.a.55 (both H1 and H3 stem-binding), AG11-2F01 (H1 stem-binding), and 042-100809-2F04 (H3 stem-binding)[7,16]. For each replicate, the average enrichment of negative controls is shown as a dotted line. Pearson correlation coefficient (*R*) between replicates is indicated. **(D)** The average enrichments against H1 stem and H3 stem are compared. The dotted line represents the cutoff for the identification of HA stem antibody candidates. H1 stem antibody candidates are in lower right quadrant, whereas H3 stem antibody candidates are in the upper left quadrant.

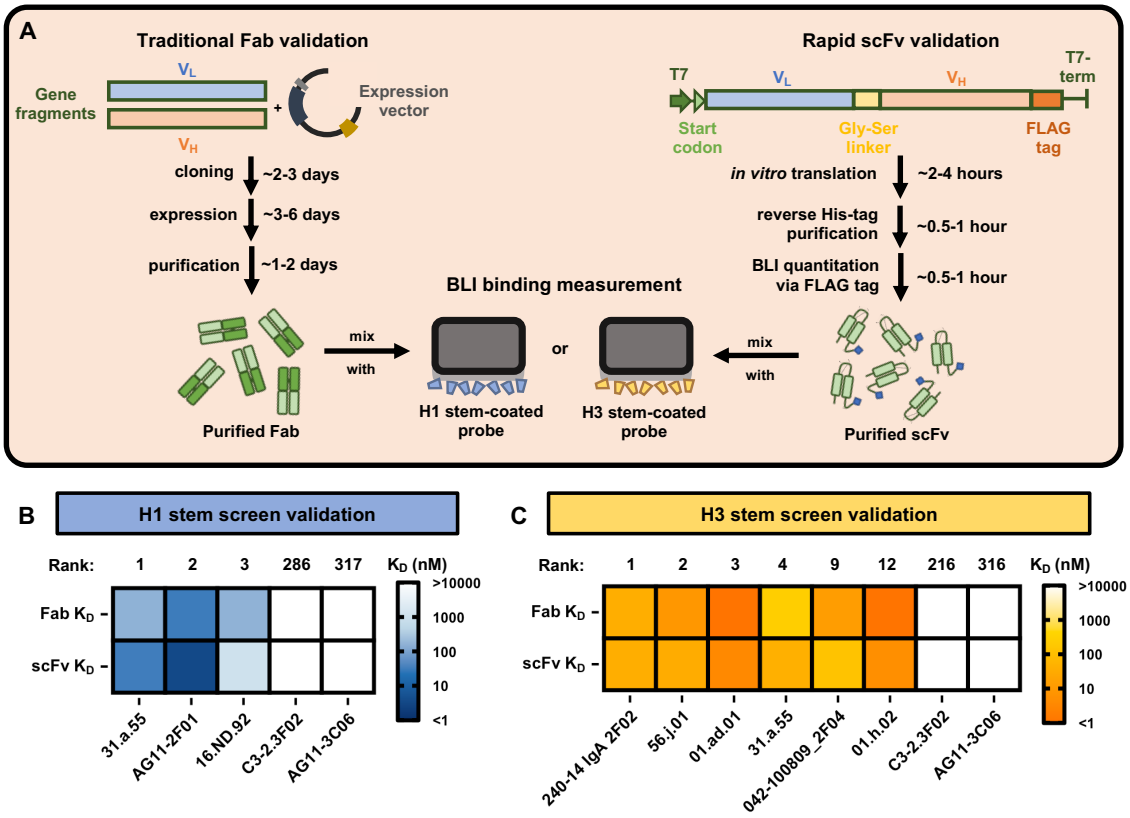
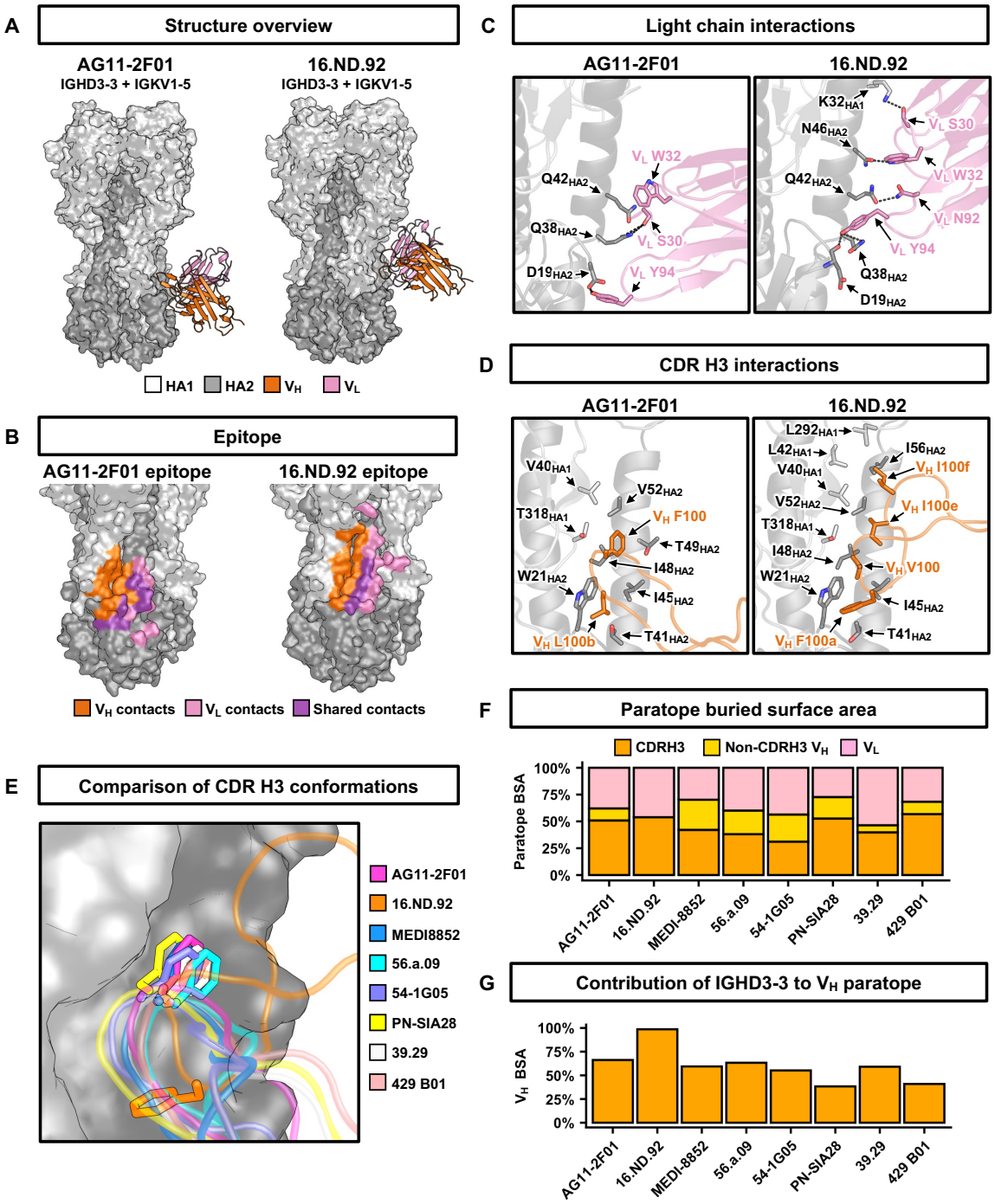


Figure 3. Validation of HA stem antibody candidates. (A) Schematic overview of the validation experiment. H1 and H3 stem antibody candidates in both scFv and Fab formats were validated using biolayer interferometry (BLI). **(B-C)** The dissociation constants (K_D) of antibodies against **(B)** H1 stem and **(C)** H3 stem are shown as heatmaps. The selected antibodies were listed from left to right based on their average enrichments. The enrichment rank of each antibody in the library were shown. Of note, 31.a.55 is a positive control for binding to both H1 stem and H3 stem[7], AG11-2F01 is a positive control for binding to H1 stem[16], and 042-100809_2F04 is a positive control for binding to H3 stem[54].



722

723 **Figure 4. Structural analysis of AG11-2F01 and 16.ND.92. (A)** Cryo-EM structures of AG11-
724 2F01 and 16.ND.92 in complex with SI/06 HA. HA1 is in light grey. HA2 is in dark grey. Heavy
725 chain variable domain (V_H) is in orange. Light chain variable domain (V_L) is in pink. **(B)**
726 Epitopes of AG11-2F01 and 16.ND.92. V_H contacts are in orange. V_L contacts are in pink.
727 Contacts shared by both V_H and V_L are in purple. **(C)** Interactions between light chain and HA
728 are shown. H-bonds are represented by black dashed lines. **(D)** Interactions between CDR

729 H3 and HA are shown. **(E)** Overlay of the CDR H3 loops from IGHD3-3 HA-stem antibodies.
730 HA is in surface representation. **(F)** Contributions of CDR H3 (orange), non-CDR H3 V_H
731 (yellow), and V_L (pink) to the paratope buried surface areas (BSA) of the indicated antibodies.
732 **(G)** Contributions of IGHD3-3-encoded residues to V_H paratope BSA of the indicated
733 antibodies.

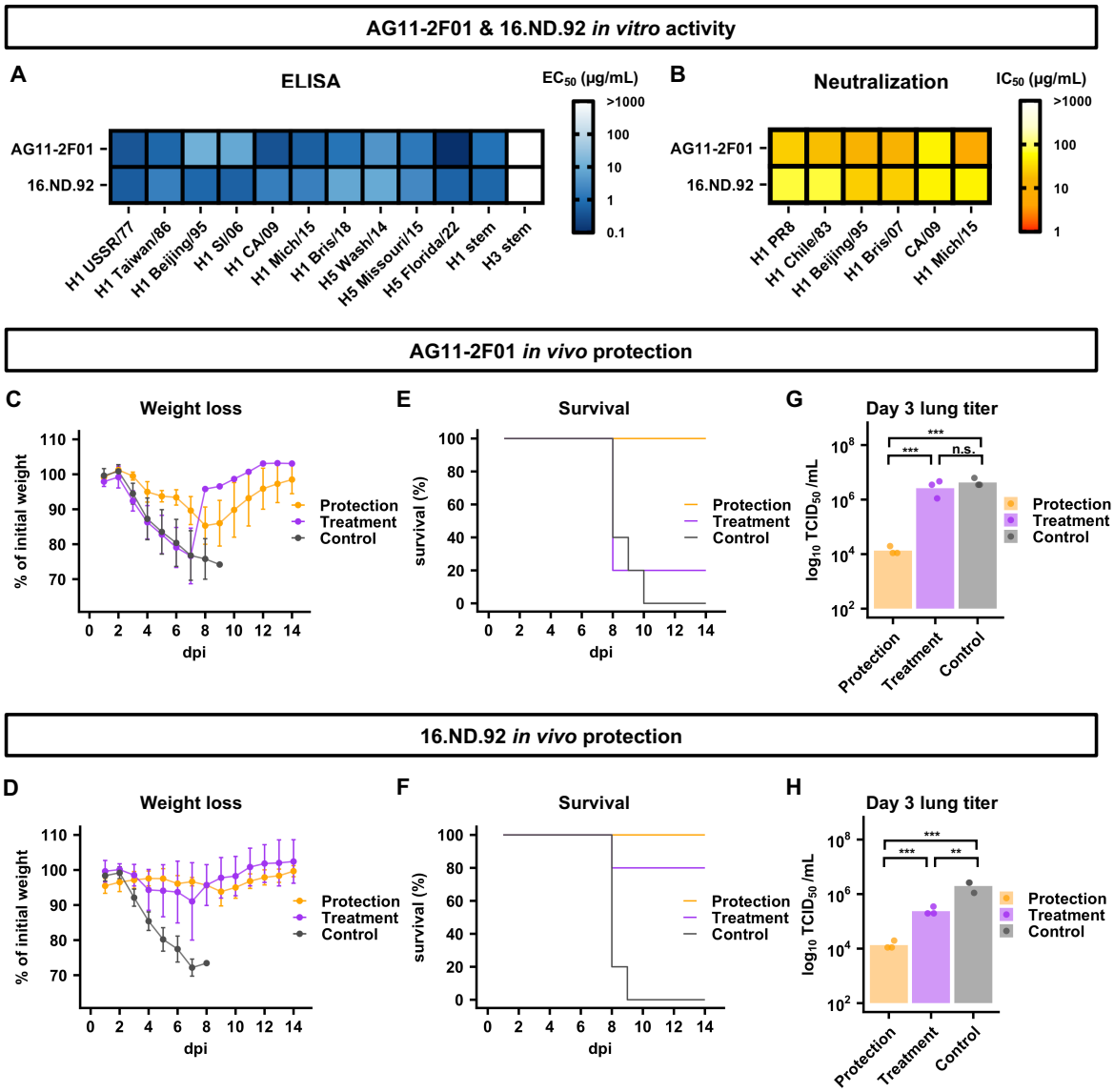


Figure 5. *in vitro* and *in vivo* protection of AG11-2F01 and 16.ND.92. (A) The binding activities of AG11-2F01 and 16.ND.92 against recombinant HA proteins from the indicated H1 and H5 strains were measured by ELISA. The EC_{50} values are shown as a heatmap. **(B)** The neutralization activity of AG11-2F01 and 16.ND.92 against different recombinant H1N1 viruses was measured by a microneutralization assay. The IC_{50} values are shown as a heatmap. **(A-B)** Strain names are abbreviated as follows: H1N1 A/Puerto Rico/8/1934 (H1 PR8), H1N1 A/USSR/90/1977 (H1 USSR/77), H1N1 A/Chile/1/1983 (H1 Chile/83), H1N1 A/Taiwan/01/1986 (H1 Taiwan/86), H1N1 A/Beijing/262/1995 (H1 Beijing/95), H1N1 A/Solomon Island/3/2006 (H1 SI/06), H1N1 A/Brisbane/59/2007 (H1 Bris/07), H1N1 A/California/04/2009 (H1 CA/09), H1N1 A/Michigan/45/2015 (H1 Mich/15), H1N1

745 A/Brisbane/02/2018 (H1 Bris/18), H5N8 A/northern pintail/WA/40964/2014 (H5 Wash/14),
746 H5N2 A/snow goose/Missouri/CC15-84A/2015 (H5 Missouri/15), and H5N1 A/bald
747 eagle/Florida/W22-134-OP/2022 (H5 Florida/22). H1 stem and H3 stem represents the HA
748 stem constructs designed based on H1N1 A/Brisbane/59/2007 HA and H3N2
749 A/Finland/486/2004 HA, respectively[31,32]. **(C-H)** The *in vivo* protection activity of AG11-2F01
750 and 16.ND.92 against lethal challenge of PR8 virus was assessed by **(C-D)** weight loss profiles,
751 **(E-F)** Kaplan-Meier survival curves, and **(G-H)** lung viral titers at day 3 post-infection. *P* values
752 were computed by two-tailed student's t-test. ***: $P < 0.001$; **: $P < 0.01$; n.s.: not significant.
753

754 **SUPPLEMENTAL FIGURES**

755 **Figure S1. Curation and design of the natively paired HA antibody library.**

756 **Figure S2. Quality assessment of oligo pool-based antibody library assembly.**

757 **Figure S3. BLI sensorgrams for antibody validation in Fab format.**

758 **Figure S4. BLI sensorgrams for antibody validation in scFv format.**

759 **Figure S5. CDR H3 sequence analysis of AG11-2F01 and 16.ND.92.**

760 **Figure S6. ELISA titration curves.**

761

762 **SUPPLEMENTAL TABLES**

763 **Table S1: Selected HA antibodies.**

764 **Table S2: Oligo pool sequences.**

765 **Table S3: Enrichment results of oPool$^+$ display.**

766 **Table S4: Cryo-EM data collection, refinement and validation statistics.**

767 **Table S5: Buried surface areas upon binding of IGHD3-3 antibodies.**

768 **Table S6: Cost and time comparison between traditional approaches for antibody**
769 **specificity characterization and oPool$^+$ display.**

770 **Table S7: Sequences of primers used in this study.**

771

772 **REFERENCES**

773 1. Pedrioli, A. & Oxenius, A. Single B cell technologies for monoclonal antibody discovery.
774 *Trends Immunol.* **42**, 1143–1158 (2021).
775 2. Valldorf, B. *et al.* Antibody display technologies: selecting the cream of the crop. *Biol.*
776 *Chem.* **403**, 455–477 (2022).
777 3. Persson, M. A., Caothien, R. H. & Burton, D. R. Generation of diverse high-affinity human
778 monoclonal antibodies by repertoire cloning. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 2432–2436
779 (1991).
780 4. Sheets, M. D. *et al.* Efficient construction of a large nonimmune phage antibody library:
781 the production of high-affinity human single-chain antibodies to protein antigens. *Proc.*
782 *Natl. Acad. Sci. U. S. A.* **95**, 6157–6162 (1998).
783 5. Kirkpatrick, E., Qiu, X., Wilson, P. C., Bahl, J. & Krammer, F. The influenza virus
784 hemagglutinin head evolves faster than the stalk domain. *Sci. Rep.* **8**, 10432 (2018).
785 6. Zost, S. J., Wu, N. C., Hensley, S. E. & Wilson, I. A. Immunodominance and antigenic
786 variation of Influenza virus hemagglutinin: implications for design of universal vaccine
787 immunogens. *J. Infect. Dis.* **219**, S38–S45 (2019).
788 7. Joyce, M. G. *et al.* Vaccine-induced antibodies that neutralize group 1 and group 2
789 Influenza A viruses. *Cell* **166**, 609–623 (2016).
790 8. Corti, D. *et al.* A neutralizing antibody selected from plasma cells that binds to group 1 and
791 group 2 Influenza A hemagglutinins. *Science* **333**, 850–856 (2011).
792 9. Kallewaard, N. L. *et al.* Structure and function analysis of an antibody aecognizing all
793 Influenza A subtypes. *Cell* **166**, 596–608 (2016).
794 10. Dreyfus, C. *et al.* Highly conserved protective epitopes on Influenza B viruses. *Science*
795 **337**, 1343–1348 (2012).
796 11. Wu, N. C. & Wilson, I. A. Influenza hemagglutinin structures and antibody recognition.
797 *Cold Spring Harb. Perspect. Med.* **10**, a038778 (2020).
798 12. Andrews, S. F., Graham, B. S., Mascola, J. R. & McDermott, A. B. Is it possible to develop
799 a “universal” influenza virus vaccine? Immunogenetic considerations underlying B-cell
800 biology in the development of a pan-subtype Influenza A vaccine targeting the

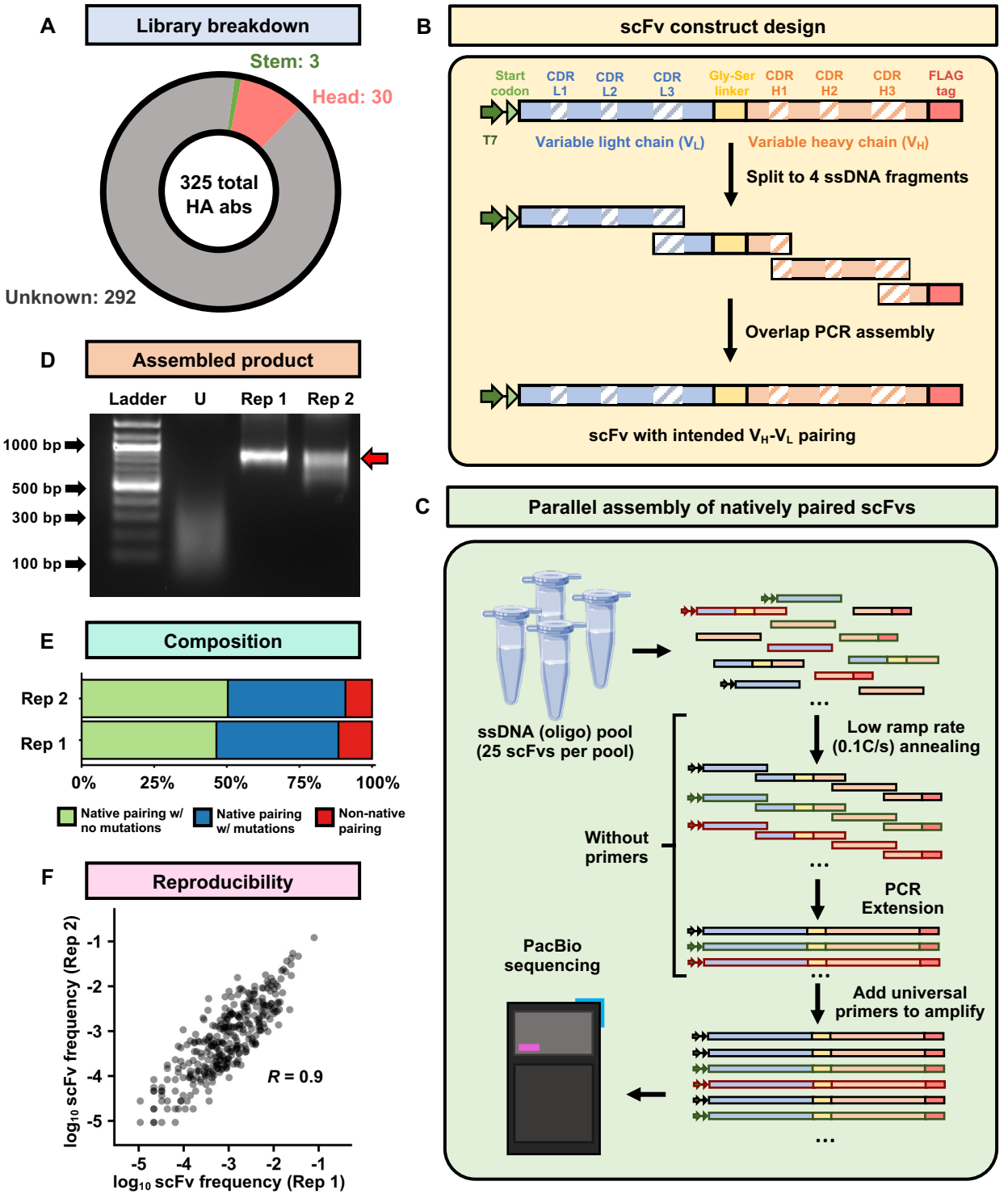
801 hemagglutinin stem. *Cold Spring Harb. Perspect. Biol.* **10**, a029413 (2018).
802 13. Pappas, L. *et al.* Rapid development of broadly influenza neutralizing antibodies through
803 redundant mutations. *Nature* **516**, 418–422 (2014).
804 14. Wu, N. C. *et al.* Recurring and adaptable binding motifs in broadly neutralizing antibodies
805 to Influenza virus are encoded on the D3-9 segment of the Ig gene. *Cell Host Microbe* **24**,
806 569-578.e4 (2018).
807 15. Wang, Y. *et al.* An explainable language model for antibody specificity prediction using
808 curated influenza hemagglutinin antibodies. *Immunity* (2024)
809 doi:10.1016/j.immuni.2024.07.022.
810 16. Henry, C. *et al.* Influenza virus vaccination elicits poorly adapted B cell responses in elderly
811 individuals. *Cell Host Microbe* **25**, 357-366.e6 (2019).
812 17. Andrews, S. F. *et al.* Immune history profoundly affects broadly protective B cell responses
813 to influenza. *Sci. Transl. Med.* **7**, 316ra192 (2015).
814 18. Dugan, H. L. *et al.* Preexisting immunity shapes distinct antibody landscapes after
815 influenza virus infection and vaccination in humans. *Sci. Transl. Med.* **12**, eabd3601 (2020).
816 19. Cheung, C. S.-F. *et al.* Identification and structure of a multidonor class of head-directed
817 Influenza-neutralizing antibodies reveal the mechanism for its recurrent elicitation. *Cell*
818 *Rep.* **32**, 108088 (2020).
819 20. Ekiert, D. C. *et al.* Cross-neutralization of influenza A viruses mediated by a single antibody
820 loop. *Nature* **489**, 526–532 (2012).
821 21. Henry, C. *et al.* Monoclonal antibody responses after recombinant hemagglutinin vaccine
822 versus subunit inactivated Influenza Virus vaccine: a comparative study. *J. Virol.* **93**,
823 e01150-19 (2019).
824 22. Hu, H. *et al.* A human antibody recognizing a conserved epitope of H5 hemagglutinin
825 broadly neutralizes highly pathogenic avian influenza H5N1 viruses. *J. Virol.* **86**, 2978–
826 2989 (2012).
827 23. Liu, Y. *et al.* Cross-lineage protection by human antibodies binding the influenza B
828 hemagglutinin. *Nat. Commun.* **10**, 324 (2019).

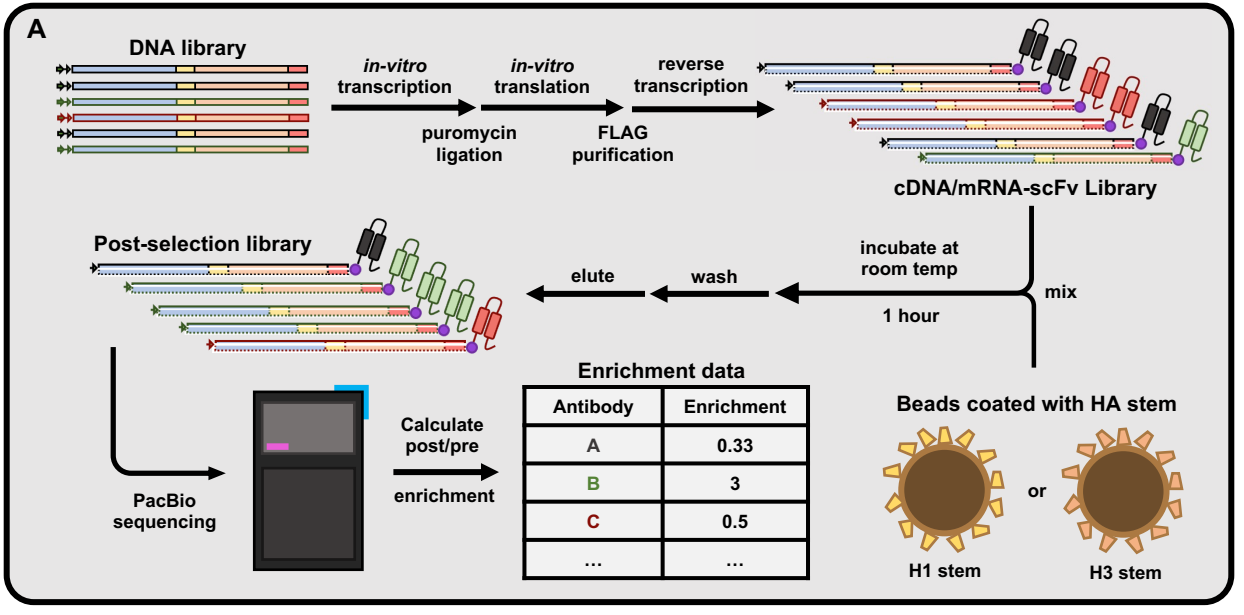
829 24. Watanabe, A. *et al.* Antibodies to a conserved Influenza head interface epitope protect by
830 an IgG subtype-dependent mechanism. *Cell* **177**, 1124-1135.e16 (2019).
831 25. Lee, J. *et al.* Molecular-level analysis of the serum antibody repertoire in young adults
832 before and after seasonal influenza vaccination. *Nat. Med.* **22**, 1456–1464 (2016).
833 26. Plesa, C., Sidore, A. M., Lubock, N. B., Zhang, D. & Kosuri, S. Multiplexed gene synthesis
834 in emulsions for exploring protein functional landscapes. *Science* **359**, 343–347 (2018).
835 27. Moravec, Z. *et al.* Discovery of tumor-reactive T cell receptors by massively parallel library
836 synthesis and screening. *Nat. Biotechnol.* 1–9 (2024) doi:10.1038/s41587-024-02210-6.
837 28. Roberts, R. W. & Szostak, J. W. RNA-peptide fusions for the in vitro selection of peptides
838 and proteins. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12297–12302 (1997).
839 29. Wu, N. C., Dai, L., Olson, C. A., Lloyd-Smith, J. O. & Sun, R. Adaptation in protein fitness
840 landscapes is facilitated by indirect paths. *eLife* **5**, e16965 (2016).
841 30. Tanaka, S. *et al.* Rapid identification of neutralizing antibodies against SARS-CoV-2
842 variants by mRNA display. *Cell Rep.* **38**, 110348 (2022).
843 31. Impagliazzo, A. *et al.* A stable trimeric influenza hemagglutinin stem as a broadly protective
844 immunogen. *Science* **349**, 1301–1306 (2015).
845 32. Corbett, K. S. *et al.* Design of nanoparticulate group 2 influenza virus hemagglutinin stem
846 antigens that activate unmutated ancestor B cell receptors of broadly neutralizing antibody
847 lineages. *mBio* **10**, e02810-18 (2019).
848 33. Wu, N. C. *et al.* Convergent evolution in breadth of two VH6-1-encoded Influenza antibody
849 clonotypes from a single donor. *Cell Host Microbe* **28**, 434-444.e4 (2020).
850 34. Matsuda, K. *et al.* Prolonged evolution of the memory B cell response induced by a
851 replicating adenovirus-influenza H5 vaccine. *Sci. Immunol.* **4**, eaau2710 (2019).
852 35. Chen, Y. *et al.* Structural basis for a human broadly neutralizing influenza A hemagglutinin
853 stem-specific antibody including H17/18 subtypes. *Nat. Commun.* **13**, 7603 (2022).
854 36. Nakamura, G. *et al.* An in vivo human-plasmablast enrichment technique allows rapid
855 identification of therapeutic Influenza A antibodies. *Cell Host Microbe* **14**, 93–103 (2013).
856 37. Lanzavecchia, A., Frühwirth, A., Perez, L. & Corti, D. Antibody-guided vaccine design:

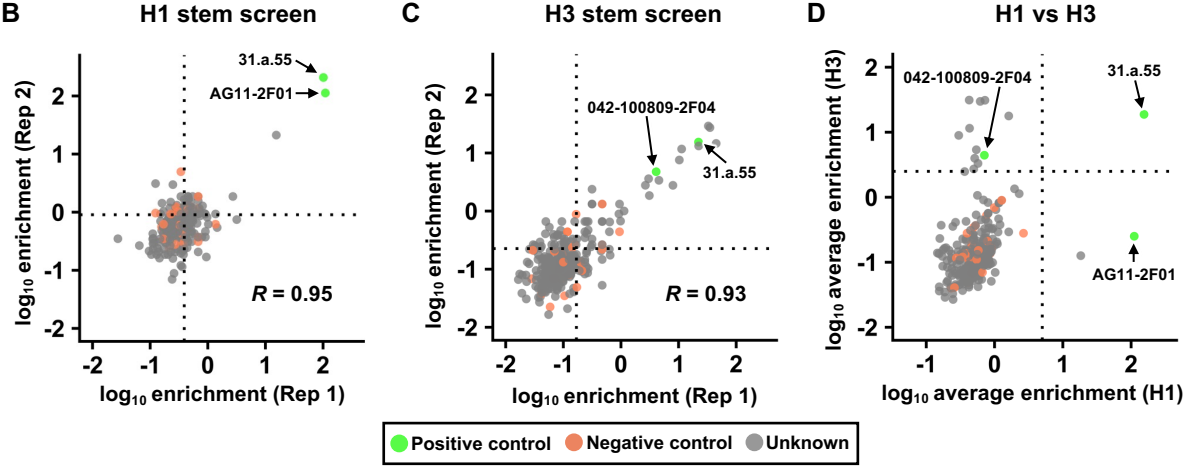
857 identification of protective epitopes. *Curr. Opin. Immunol.* **41**, 62–67 (2016).
858 38. Haynes, B. F. *et al.* Strategies for HIV-1 vaccines that induce broadly neutralizing
859 antibodies. *Nat. Rev. Immunol.* **23**, 142–158 (2023).
860 39. Toussi, S. S., Hammond, J. L., Gerstenberger, B. S. & Anderson, A. S. Therapeutics for
861 COVID-19. *Nat. Microbiol.* **8**, 771–786 (2023).
862 40. Schoeder, C. T. *et al.* Epitope-focused immunogen design based on the ebolavirus
863 glycoprotein HR2-MPER region. *PLOS Pathog.* **18**, e1010518 (2022).
864 41. Liu, L. *et al.* An antibody class with a common CDRH3 motif broadly neutralizes
865 sarbecoviruses. *Sci. Transl. Med.* **14**, eabn6859 (2022).
866 42. Liu, H. *et al.* Human antibodies to SARS-CoV-2 with a recurring YYDRxG motif retain
867 binding and neutralization to variants of concern including Omicron. *Commun. Biol.* **5**, 766
868 (2022).
869 43. Hunt, A. C. *et al.* A rapid cell-free expression and screening platform for antibody discovery.
870 *Nat. Commun.* **14**, 3897 (2023).
871 44. Joubbi, S. *et al.* Antibody design using deep learning: from sequence and structure design
872 to affinity maturation. *Brief. Bioinform.* **25**, bbae307 (2024).
873 45. Bennett, N. R. *et al.* Atomically accurate de novo design of single-domain antibodies.
874 *BioRxiv Prepr. Serv. Biol.* 2024.03.14.585103 (2024) doi:10.1101/2024.03.14.585103.
875 46. Wu, N. C. *et al.* In vitro evolution of an influenza broadly neutralizing antibody is modulated
876 by hemagglutinin receptor specificity. *Nat. Commun.* **8**, 15371 (2017).
877 47. Quintero-Hernández, V. *et al.* The change of the scFv into the Fab format improves the
878 stability and in vivo toxin neutralization capacity of recombinant antibodies. *Mol. Immunol.*
879 **44**, 1307–1315 (2007).
880 48. Kang, T. H. & Seong, B. L. Solubility, stability, and avidity of recombinant antibody
881 fragments expressed in microorganisms. *Front. Microbiol.* **11**, (2020).
882 49. Rondot, S., Koch, J., Breitling, F. & Dübel, S. A helper phage to improve single-chain
883 antibody presentation in phage display. *Nat. Biotechnol.* **19**, 75–78 (2001).
884 50. Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display.

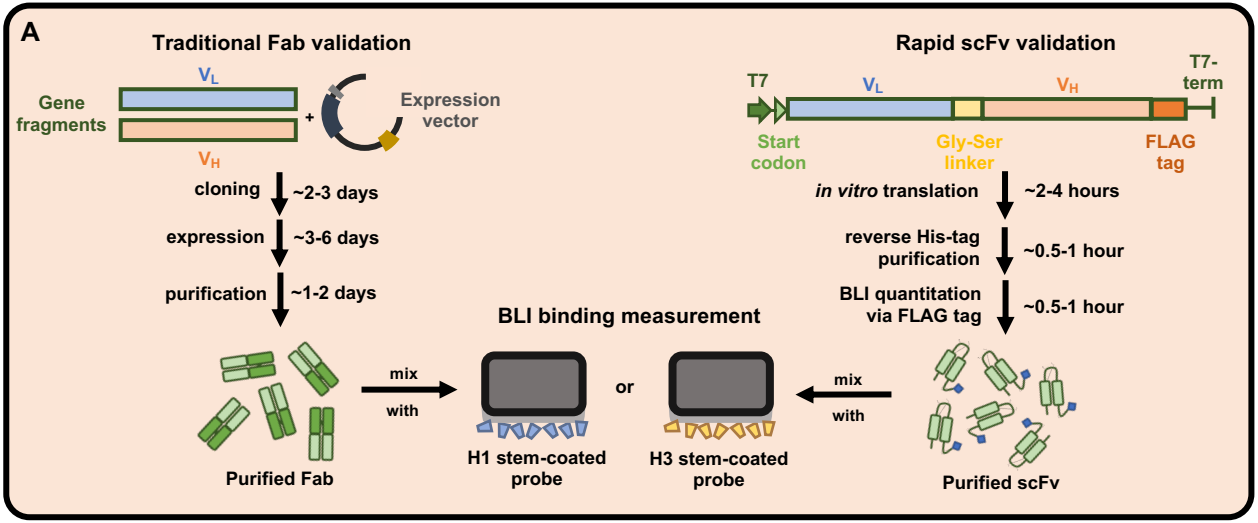
885 *Nat. Protoc.* **1**, 755–768 (2006).
886 51. Sidore, A. M., Plesa, C., Samson, J. A., Lubock, N. B. & Kosuri, S. DropSynth 2.0: high-
887 fidelity multiplexed gene synthesis in emulsions. *Nucleic Acids Res.* **48**, e95 (2020).
888 52. Huang, K.-Y. A. *et al.* Focused antibody response to influenza linked to antigenic drift. *J.*
889 *Clin. Invest.* **125**, 2631–2645 (2015).
890 53. Li, G.-M. *et al.* Pandemic H1N1 influenza vaccine induces a recall response in humans
891 that favors broadly cross-reactive memory B cells. *Proc. Natl. Acad. Sci. U. S. A.* **109**,
892 9047–9052 (2012).
893 54. Henry Dunand, C. J. *et al.* Preexisting human antibodies neutralize recently emerged
894 H7N9 influenza strains. *J. Clin. Invest.* **125**, 1255–1268 (2015).
895 55. Swindells, M. B. *et al.* abYsis: Integrated antibody sequence and structure-management,
896 analysis, and prediction. *J. Mol. Biol.* **429**, 356–364 (2017).
897 56. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-
898 generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
899 57. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421
900 (2009).
901 58. Olson, C. A., Wu, N. C. & Sun, R. A comprehensive biophysical description of pairwise
902 epistasis throughout an entire protein domain. *Curr. Biol.* **24**, 2643–2651 (2014).
903 59. Cock, P. J. A. *et al.* Biopython: freely available Python tools for computational molecular
904 biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009).
905 60. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for
906 rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
907 61. Leem, J., Dunbar, J., Georges, G., Shi, J. & Deane, C. M. ABodyBuilder: Automated
908 antibody structure prediction with data–driven accuracy estimation. *mAbs* **8**, 1259–1268
909 (2016).
910 62. Laursen, N. S. *et al.* Universal protection against influenza infection by a multidomain
911 antibody to influenza hemagglutinin. *Science* **362**, 598–602 (2018).
912 63. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and

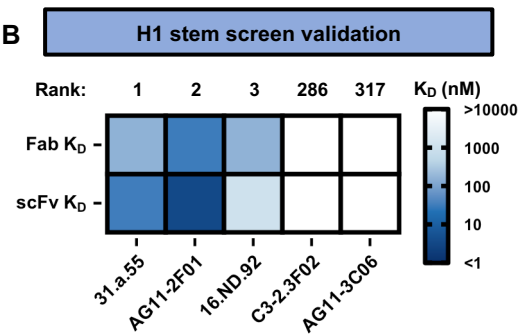
913 analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
914 64. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta*
915 *Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
916 65. Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography.
917 *Acta Crystallogr. Sect. Struct. Biol.* **74**, 531–544 (2018).
918 66. Benton, D. J. *et al.* Influenza hemagglutinin membrane anchor. *Proc. Natl. Acad. Sci. U.*
919 *S. A.* **115**, 10112–10117 (2018).
920 67. Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume
921 post-processing. *Commun. Biol.* **4**, 874 (2021).
922 68. Ruffolo, J. A., Chu, L.-S., Mahajan, S. P. & Gray, J. J. Fast, accurate antibody structure
923 prediction from deep learning on massive set of natural antibodies. *Nat. Commun.* **14**,
924 2389 (2023).
925 69. Millán, C., McCoy, A. J., Terwilliger, T. C. & Read, R. J. Likelihood-based docking of models
926 into cryo-EM maps. *Acta Crystallogr. Sect. Struct. Biol.* **79**, 281–289 (2023).
927 70. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state.
928 *J. Mol. Biol.* **372**, 774–797 (2007).
929 71. Ye, J., Ma, N., Madden, T. L. & Ostell, J. M. IgBLAST: an immunoglobulin variable domain
930 sequence analysis tool. *Nucleic Acids Res.* **41**, W34-40 (2013).
931 72. Neumann, G., Fujii, K., Kino, Y. & Kawaoka, Y. An improved reverse genetics system for
932 influenza A virus generation and its implications for vaccine production. *Proc. Natl. Acad.*
933 *Sci. U. S. A.* **102**, 16825–16829 (2005).
934 73. Lei, R. *et al.* Leveraging vaccination-induced protective antibodies to define conserved
935 epitopes on influenza N2 neuraminidase. *Immunity* **56**, 2621-2634.e6 (2023).
936 74. Overdijk, M. B. *et al.* Crosstalk between human IgG isotypes and murine effector cells. *J.*
937 *Immunol. Baltim. Md 1950* **189**, 3430–3438 (2012).
938 75. Dekkers, G. *et al.* Affinity of human IgG subclasses to mouse Fc gamma receptors. *mAbs*
939 **9**, 767–773 (2017).
940

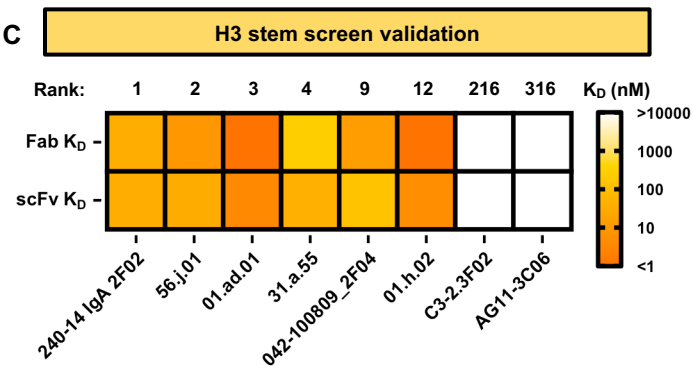

A
Library breakdown
Stem: 3
Head: 30
325 total HA abs
Unknown: 292
B
scFv construct design
Start codon
CDR L1
CDR L2
CDR L3
Gly-Ser linker
CDR H1
CDR H2
CDR H3
FLAG tag
T7
Variable light chain (VL)
Variable heavy chain (VH)
Split to 4 ssDNA fragments
Overlap PCR assembly
scFv with intended VH-VL pairing
C
Parallel assembly of natively paired scFvs
ssDNA (oligo) pool (25 scFvs per pool)
Low ramp rate (0.1C/s) annealing
Without primers
PCR Extension
Add universal primers to amplify
PacBio sequencing
D
Assembled product
Ladder
U
Rep 1
Rep 2
1000 bp
500 bp
300 bp
100 bp
E
Composition
Rep 2
Rep 1
0%
25%
50%
75%
100%
Native pairing w/ no mutations
Native pairing w/ mutations
Non-native pairing
F
Reproducibility
log10 scFv frequency (Rep 2)
log10 scFv frequency (Rep 1)
-1
-2
-3
-4
-5
R = 0.9

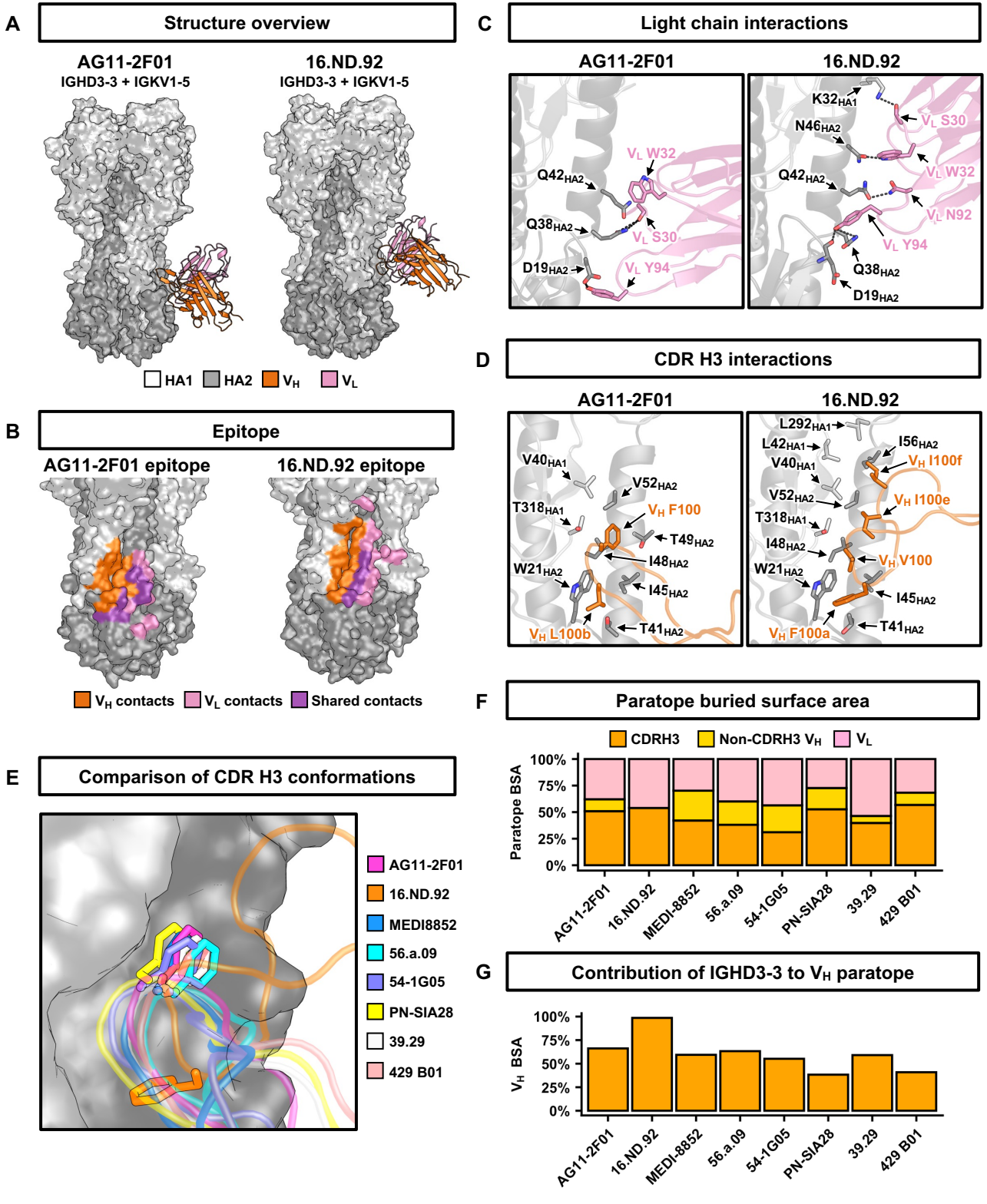

A
DNA library
in-vitro transcription
puromycin ligation
in-vitro translation
FLAG purification
reverse transcription
cDNA/mRNA-scFv Library
mix
incubate at room temp
1 hour
wash
elute
Post-selection library
PacBio sequencing
Calculate post/pre enrichment
Enrichment data
Antibody
Enrichment
A
0.33
B
3
C
0.5
...
...
Beads coated with HA stem
H1 stem
or
H3 stem


B
H1 stem screen
31.a.55
AG11-2F01
R = 0.95
log10 enrichment (Rep 2)
log10 enrichment (Rep 1)
C
H3 stem screen
042-100809-2F04
31.a.55
R = 0.93
log10 enrichment (Rep 2)
log10 enrichment (Rep 1)
D
H1 vs H3
042-100809-2F04
31.a.55
AG11-2F01
log10 average enrichment (H3)
log10 average enrichment (H1)
Positive control
Negative control
Unknown

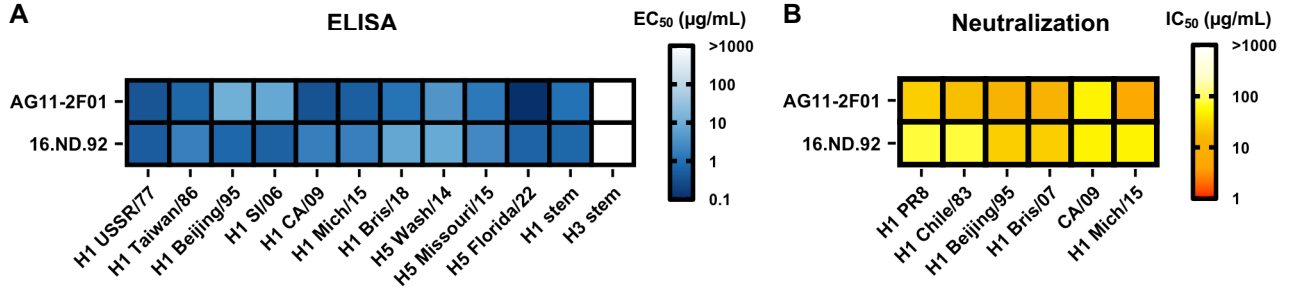

A
Traditional Fab validation
Gene fragments
V_L
V_H
Expression vector
cloning
~2-3 days
expression
~3-6 days
purification
~1-2 days
Purified Fab
Rapid scFv validation
T7
V_L
V_H
T7-term
Start codon
Gly-Ser linker
FLAG tag
in vitro translation
~2-4 hours
reverse His-tag purification
~0.5-1 hour
BLI quantitation via FLAG tag
~0.5-1 hour
Purified scFv
mix with
BLI binding measurement
H1 stem-coated probe
or
H3 stem-coated probe
mix with


B
H1 stem screen validation
Rank: 1 2 3 286 317
K_D (nM)
Fab K_D
scFv K_D
>10000
1000
100
10
<1
31.a.55
AG11-2F01
16.ND.92
C3-2.3F02
AG11-3C06

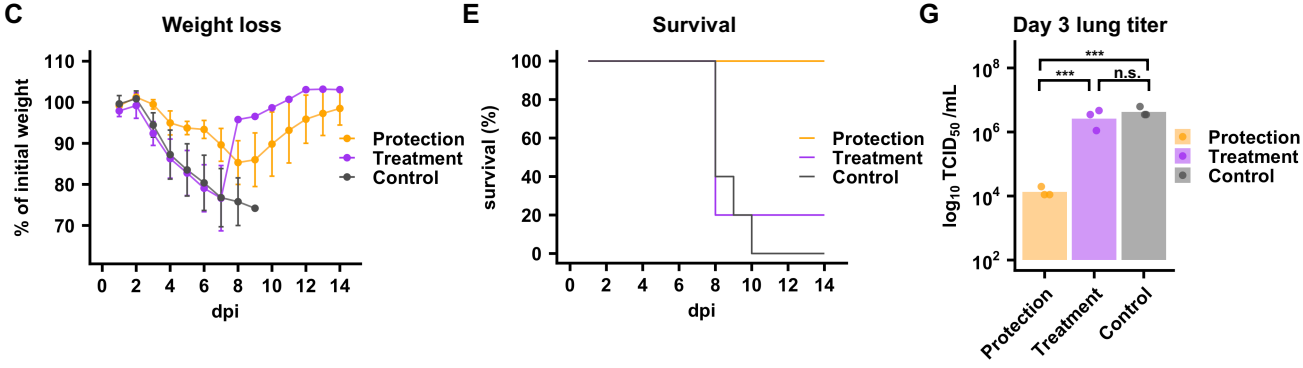
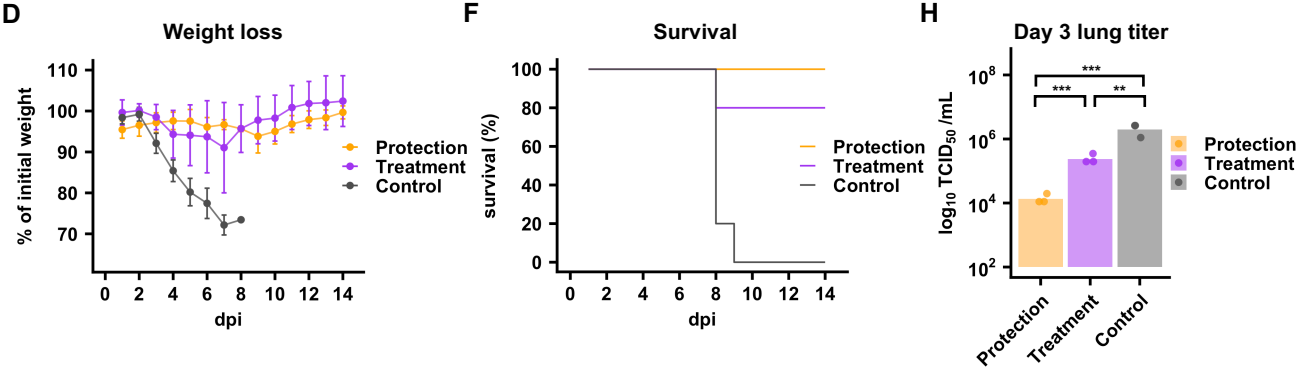

C
H3 stem screen validation
Rank: 1 2 3 4 9 12 216 316
K_D (nM)
Fab K_D
scFv K_D
>10000
1000
100
10
<1
240-14 IgA 2F02
56.j.01
01.ad.01
31.a.55
042-100809_2F04
01.h.02
C3-2.3F02
AG11-3C06


A
Structure overview
AG11-2F01
IGHD3-3 + IGKV1-5
16.ND.92
IGHD3-3 + IGKV1-5
HA1
HA2
VH
VL
B
Epitope
AG11-2F01 epitope
16.ND.92 epitope
VH contacts
VL contacts
Shared contacts
C
Light chain interactions
AG11-2F01
16.ND.92
Q42HA2
Q38HA2
D19HA2
VL W32
VL S30
VL Y94
K32HA1
N46HA2
VL N92
D
CDR H3 interactions
AG11-2F01
16.ND.92
V40HA1
T318HA1
W21HA2
V52HA2
VH F100
T49HA2
I48HA2
I45HA2
T41HA2
VH L100b
L292HA1
L42HA1
I56HA2
VH I100f
VH I100e
VH V100
VH F100a
E
Comparison of CDR H3 conformations
AG11-2F01
16.ND.92
MEDI8852
56.a.09
54-1G05
PN-SIA28
39.29
429 B01
F
Paratope buried surface area
CDRH3
Non-CDRH3 VH
VL
Paratope BSA
100%
75%
50%
25%
0%
AG11-2F01
16.ND.92
MEDI-8852
56.a.09
54-1G05
PN-SIA28
39.29
429 B01
G
Contribution of IGHD3-3 to VH paratope
VH BSA
100%
75%
50%
25%
0%
AG11-2F01
16.ND.92
MEDI-8852
56.a.09
54-1G05
PN-SIA28
39.29
429 B01

AG11-2F01 & 16.ND.92 *in vitro* activity



AG11-2F01 *in vivo* protection



16.ND.92 *in vivo* protection