OXFORD

*Antibody Therapeutics*, 2024, tbae025
https://doi.org/10.1093/abt/tbae025
**Research Article**

# AI-based antibody discovery platform identifies novel, diverse, and pharmacologically active therapeutic antibodies against multiple SARS-CoV-2 strains

Cristina Moldovan Loomis [1,*], Thomas Lahlali[2], Danielle Van Citters[1], Megan Sprague[1], Gregory Neveu[2], Laurence Somody[2], Christine C. Siska[1], Derrick Deming[1], Andrew J. Asakawa[1], Tileli Amimeur[1], Jeremy M. Shaver[1], Caroline Carbonelle[2], Randal R. Ketchem [1], Antoine Alam[2], Rutilio H. Clark[1]

[1]Department of Discovery & Molecular Design, Just-Evotec Biologics Inc., 401 Terry Avenue N., Seattle, WA 98109, USA
[2]Department of Virology, Evotec ID, 40, Avenue Tony Garnier, 69007 Lyon, France

*Corresponding author. Just-Evotec Biologics Inc. 401 Terry Ave N., Seattle, Washington 98109, USA. Tel: +1 415-613-1960. E-mail: cristina.loomis@evotec.com

## Abstract

**Background:** We are entering a new era of antibody discovery and optimization where machine learning (ML) processes will become indispensable for the design and development of therapeutics.
**Methods:** We have constructed a Humanoid Antibody Library for the discovery of therapeutics that is an initial step towards leveraging the utility of artificial intelligence and ML. We describe how we began our validation of the library for antibody discovery by isolating antibodies against a target of pandemic concern, SARS-CoV-2. The two main antibody quality aspects that we focused on were functional and biophysical characterization.
**Results:** The applicability of our platform for effective therapeutic antibody discovery is demonstrated here with the identification of a panel of human monoclonal antibodies that are novel, diverse, and pharmacologically active.
**Conclusions:** These first-generation antibodies, without the need for affinity maturation, exhibited neutralization of SARS-CoV-2 viral infectivity across multiple strains and indicated high developability potential.

**Statement of Significance**: We describe a novel, cost-effective and accelerated approach to therapeutic antibody discovery, that couples *de novo* human antibodies derived *in silico* with high-throughput screening technologies. Our platform's applicability is demonstrated here by the identification of a panel of humanoid antibodies that are effective and with high developability potential.

**Keywords:** therapeutic antibody; discovery platform; artificial intelligence; developability; in silico

## Introduction

Antibodies represent an important class of biologics-based therapeutics with key benefits such as high specificity and affinity, longer-acting pharmacokinetics, and superior safety profiles compared to small molecules [1, 2]. Thus, antibody therapeutics are the fastest growing class of drugs on the market used for treatment of a wide range of human diseases, such as cancer, autoimmune, inflammatory, neural, metabolic, and infectious diseases.

The high potential of antibody therapeutics is, however, hampered by lengthy and expensive discovery and development processes. Indeed, candidate antibody therapeutics must undergo a complex multiobjective process and satisfy multiple criteria. This includes activity and specificity against a target, good pharmacokinetic, and safety profiles, as well as suitable biophysical and manufacturing properties. Traditionally, this is achieved in a sequential manner that lasts 10–15 years and costs approximately 2 billion USD [2]. Continual improvement of the antibody discovery and development process will benefit patients. This is especially true regarding situations such as pandemic response since the entire pipeline will be engineered end-to-end with all the salient criteria aligned, accelerating therapeutic delivery at low cost.

To achieve a more efficient and aligned antibody discovery-manufacturing end-to-end process we have begun to leverage AI and ML for design, execution, and re-learning. Progress in computational methods, technology and automation, and their increasing integration in multiple aspects of the biopharmaceutical pipeline have the potential to revolutionize therapeutic antibody discovery and development. We have employed ML to generate novel, humanoid antibody sequences that both represent natural repertoires and are biased towards desirable developability features. To enable properties such as broad target and epitope engagement, focused efficacy, and suitable developability, we have developed an antibody Generative Adversarial Network (GAN), a new *in silico* engineering approach for designing a novel class of diverse, hyper-realistic antibodies, termed "humanoid" antibodies [3]. The algorithm uses a modified Wasserstein-GAN for both single-chain (light or heavy chain) and paired-chain (light and heavy chain) antibody sequence generation [4]. These GANs

**Received:** March 9, 2024. **Revised:** July 5, 2024

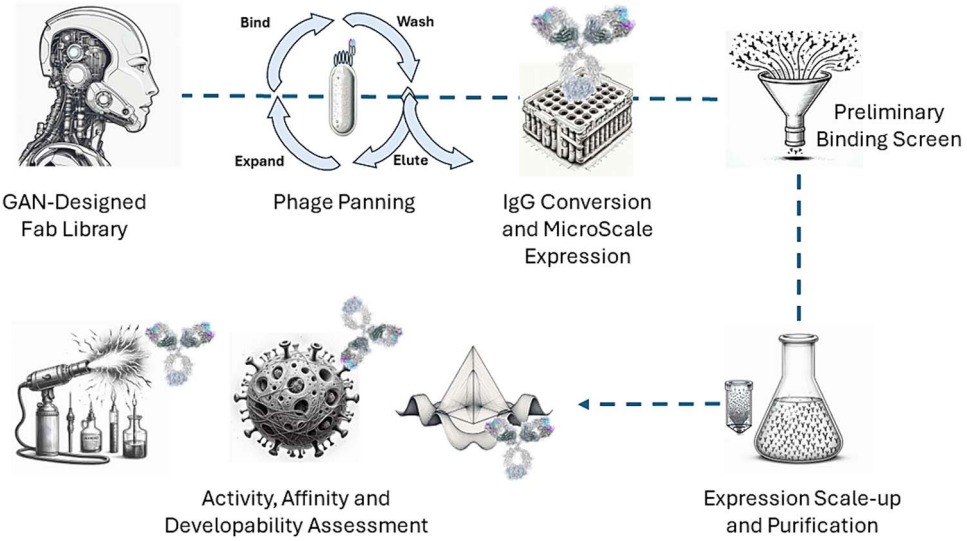


Figure 1. J.HAL® antibody discovery workflow. Depending on the number and complexity of the assays, the entire depicted process can be achieved in less than 4 months. Typical duration for phage panning: 2 weeks; IgG conversion and MicroScale expression: 3 weeks; preliminary binding screen: 2 weeks; expression scale-up and purification: 2 weeks; activity, affinity, and developability assessment: 4 weeks.

allow us to encode for antibodies with key properties of interest to create feature-biased libraries as the central benefit of our display libraries and to inform antibody engineering research. Our antibody GAN architecture (i) captures the complexity of the entire variable region of the standard human antibody sequence space, (ii) provides a basis for rationally generating novel antibodies that span a larger sequence diversity than is explored by standard *in silico* generative approaches, and (iii) provides, through transfer learning, an inherent method to bias the physical properties of the generated antibodies toward desired features, which could lead to improved efficacy as well as chemical and biophysical properties, which are critical for developability.

Here, we present the construction and utilization of a GAN-generated phage display library of approximately 1 billion Fab antibodies. This generalized Fab library was not designed for any particular target antigen nor to exhibit any particular type of paratope structure. As the first step in our validation of library utility, the Just Humanoid Antibody Library (J.HAL®) was successfully screened to isolate a panel of novel, diverse, and pharmacologically active human monoclonal antibodies against multiple SARS-CoV-2 strains. These first-generation antibodies, without the need of affinity maturation, bind to the SARS-CoV-2 spike protein with therapeutically relevant specificity and affinity, block the spike:human ACE2 receptor interaction, and neutralize SARS-CoV-2 viral infectivity across several strains. The antibodies also exhibit high developability potential as indicated by multiple conformational and colloidal stability assessments.

To maximize the efficiency of therapeutic discovery, we have paired the GAN-generated humanoid library with high-throughput and sensitive technologies, such as next-generation sequencing (NGS), high-throughput molecular biology, full-length IgG expression in 96 well blocks, and a high-throughput AlphaLISA-based binding screen (Fig. 1).

The integration of automation and streamlined experimental and analysis workflows allows for the rapid screening, isolation, and characterization of suitable IgG therapeutics that are also fit to our manufacturing platform. Our antibody discovery platform for addressing pathogenic disease demonstrates the promise of providing affordable, high-quality treatments to patients around the globe.

## Materials and methods

### Library *in silico* design

Details of how a modified Wasserstein-GAN was used for the design of the J.HAL® variable light and heavy v-regions can be found here [3]. Data for the training sets were derived from the Observed Antibody Space (OAS) repository (https://opig.stats.ox.ac.uk/webapps/oas/). Raw nucleotide sequences were automatically translated, classified, and structurally aligned using in-house software, Abacus™. The AHo structure numbering system (https://plueckthun.bioc.uzh.ch/antibody/Numbering/index.html) was used for structural alignments of the variable regions.

To create the training sets, variable regions were first filtered to remove any sequences, which were not classified as human variable regions by our in-house software, Abacus™. Sequences that contained stop codons, truncations, or insertions were removed. Any sequence that had less than 85% agreement to its closest germline was also removed. The initial training set of the single-chain model contained all represented germlines. If transfer learning was necessary, it was done on an extracted set of sequences whose closest germline belonged to the specific germline of interest.

The antibody-GAN code was developed in Python. The Keras and Tensorflow deep-learning libraries were primarily used to build and train the antibody-GAN. The Pandas and Numpy libraries were used to handle any data and training set construction. Other public libraries that were used in the development and analysis of the antibody-GAN include: Sklearn, Scipy, and Seaborn. The Wasserstein-GAN consists of a generator and a discriminator, which in the WGAN architecture is commonly referred to as a critic. The single-chain network generator, which was used to design J.HAL®, takes as input a noise vector of size 296. This vector is fed into a dense layer, followed by three up-sampling and 2D convolutional transpose layers, and a final SoftMax layer to produce a 2D array of size 148 × 22. This 2D array corresponds to a one-hot-encoded representation of 148 residues and 22 possible amino acids (including deletions and Xs) of a light or heavy chain in an antibody sequence. Antibody sequences aligned by AHo numbering have 149 residues in either chain; to make the network structure simpler, we chose to remove one residue, which appears

relatively constant in human repertoire, from each chain during encoding. When decoding, we add this constant residue back in. The discriminator, or critic, takes as input the same 148 × 22 encoding of an antibody chain and passes it through two 2D convolution layers, followed by a flattening, dense layer, and a single-node linear output. Training ends when the generated sequences begin showing sufficient quality as assessed by % germline alignment plateauing.

Once the antibody-GAN indicated sufficient learning, we selected five heavy, four kappa, and one lambda germline for v-region generation. Certain developability concerns were minimized such as free cysteines, N-linked glycosylation, and covariance violations. About, 6500 sequences of each germline were synthesized by Integrated DNA Technologies (www.idtdna.com) and cloned in every VH-VL combination into phagemids for phage display.

## Phage library construction

Phagemids pADL-10b and pADL-20c (Antibody Design Labs, San Diego, CA) were used for construction of the GAN discovery library and were modified for expression of Fab antibody fragments as N-terminal pIII fusion proteins in *Escherichia coli*. Pools of synthetic gene fragments encoding the variable heavy (VH) or variable light (VL) chains for each germline family were obtained from Integrated DNA Technologies (Coralville, IA). These fragments were flanked by >20 base pairs of sequence that included restriction sites used for cloning into the phagemid backbones.

A two-step cloning process was employed for combinatorial assembly of the V-regions into each phagemid. First, 50 ng of each of the five VH chain pools was digested with NcoI/NheI and ligated with 300 ng of each phagemid backbone. The ligation reactions were cleaned up and concentrated using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. Two microliters of this reaction were transformed into *E. coli* One Shot™ TOP10 electrocompetent cells (Thermo Fisher Scientific, Waltham, MA) in a 0.1-cm cuvette using a BTX ECM 630 electroporator (BTX, Holliston, MA) with the following settings: 1.8 kV, 600 Ohms, 25 $\mu$F. The transformants were recovered in 0.5 ml SOC media for 1 h. Cells from the transformation reaction were then used to inoculate 150 ml TB supplemented with 50 $\mu$g/ml carb and 2% glucose and incubated at 30°C and 250 rpm overnight. Cultures were harvested the following day, and plasmid DNA was isolated using GenElute™ HP Maxiprep Kit (Millipore Sigma) following the manufacturer's protocol.

Second, 5 $\mu$g of phagemid DNA harboring each VH chain pool was digested using BsiWI/RsrII (for kappa chains) or BsiWI/KasI (for lambda chains) and ligated with 0.5 $\mu$g of each of the 5 VL chain pools that had been digested with the appropriate enzymes. The ligation reactions were cleaned up and concentrated using QIAquick PCR Purification Kit following the manufacturer's protocol. Approximately, 2 $\mu$l of this reaction were transformed into MegaX DH10B $T1^R$ Electrocomp™ cells (Thermo Fisher Scientific) in a 0.1-cm cuvette with the following settings: 2.0 kV, 200 Ohms, 25 $\mu$F. As many as 3–8 replicate transformations were performed for each germline pair to ensure the number of transformants exceeded the theoretical diversity by at least 10-fold. The transformants were recovered in 1 ml recovery media for 1 h before pooling replicates and removing 10 $\mu$l for dilution plating on 2xYT plates supplemented with 100 $\mu$g/ml carb and 2% glucose, followed by overnight incubation at 30°C. The resulting colonies were used for estimating library size and for sequencing the VH and VL chains via colony PCR. The remainder of the transformation was used to inoculate 150 ml TB supplemented with 50 $\mu$g/ml carb and 2% glucose followed by overnight incubation at 30°C and 250 rpm. Cultures were harvested the following day, and plasmid DNA was isolated using GenElute™ HP Maxiprep Kit, following the manufacturer's protocol. A total of 50 phagemid DNA preps were synthesized in this manner to complete the library (25 sublibraries per each of 2 phagemid backbones). Colony PCR and Sanger sequencing were performed on a minimum of 24 clones per germline pair to assess correct insertion rate of the VH and VL domains.

## Phage library production

*E. coli* SS320 host cells (Lucigen Corporation, Middleton, WI) were electroporated using 30–50 ng of each sublibrary DNA in a 0.1-cm cuvette with the following settings: 1.8 kV, 600 Ohms, 25 $\mu$F. Duplicate transformations were performed for each germline pair, recovered in 1 ml recovery media, and pooled before inoculating 2xYT broth containing 50 $\mu$g/ml carbenicillin and 2% glucose at an $OD_{600}$ of 0.07. The cultures were then incubated with shaking at 250 rpm and 37°C until an $OD_{600} \sim 0.5$, at which time the cells were infected with M13KO7 helper phage (Antibody Design Labs) at a multiplicity of infection (MOI) 25. The cells were continued incubating at 37°C without shaking for 30 min, followed by shaking at 200 rpm for 30 min. Cultures were centrifuged followed by medium replacement in 2xYT supplemented with 50 $\mu$g/ml carbenicillin and 25 $\mu$g/ml kanamycin. After overnight incubation at 30°C and 200 rpm, the phage particles were purified and concentrated by PEG/NaCl precipitation and resuspended in PBS containing 0.5% BSA and 0.05% Tween-20. Phage concentration was determined using a spectrophotometer, assuming 1 unit at $OD_{268}$ is equivalent to $5 \times 10^{12}$ phage/ml.

## Polyclonal phage ELISA (Fab display)

The display level of Fabs on phage was assessed for each sublibrary using polyclonal phage ELISA. Briefly, 96-well MaxiSorp® assay plates (Nunc) were coated overnight at 4°C with antihuman Fab (Millipore Sigma) diluted 1:500 in PBS, then blocked in PBS containing 1% BSA for 1 h at room temperature (RT). Phage preparations from each germline pair were serially diluted 1:25 in 2% nonfat dry milk in PBS, added to the plates, and allowed to incubate for 1 h at RT before captured virions were detected using a 1:5000 dilution of anti-M13-HRP (Santa Cruz Biotechnology, Dallas, TX) for 1 h at RT. All interval plate washes were performed three times in PBST (PBS supplemented with 0.1% v/v Tween-20). ELISAs were developed by addition of TMB solution (Thermo Fisher Scientific) and quenched using 10% phosphoric acid. Absorbance was read at A450 nm. Phage preparations derived from two commercial antibodies were used as high- and low-display Fab on phage controls (M-6240 and M-6239), respectively.

## Biopanning and clone sequencing

Biopanning was performed independently on two different antigens to isolate SARS-CoV-2 binders from the phage libraries. RBD (Cat. No. 40592-VNAH) and Spike S1 B.1.1.7 variant (Cat. no. 40591-V08H12) were acquired from Sino Biological (Beijing, China) and biotinylated using EZ-Link™ Sulfo-NHS-Biotin kit (Thermo Fisher Scientific) as recommended by the manufacturer. In the first round of selection, approximately $1 \times 10^{13}$ phage particles were incubated in 1 ml SuperBlock™ (Thermo Fisher Scientific) for 1 h. The blocked phages were next incubated with 100 nM biotinylated antigen for 1 h, followed by incubation with magnetic streptavidin beads (Dynabeads™ M-280 Streptavidin, Thermo Fisher Scientific) for 15 min. Panning was performed by washing the beads two times with PBS-T (PBS with 0.1% Tween-20) for 5 s each, followed

by two additional washes in PBS-LT (PBS with 0.01% Tween-20) for 5 s each. The bound phages were eluted from the beads using 0.2 M glycine (pH 2.5) for 10 min and neutralized with 1 M Tris–HCl (pH 8.0). All incubations were performed using the KingFisher™ Flex at RT. Eluted phages were used to infect *E. coli* ER2738 cells (Lucigen Corporation) for titer determination and overnight amplification. Prior to the next round, the amplified phage was precipitated in 20% PEG/2.5 M NaCl for at least 20 min on ice. Subsequent rounds of panning were performed similarly but had reduced input phage ($5 \times 10^{12}$ pfu), reduced antigen concentrations (second round: 25 nM and third round: 10 nM), and longer wash times (second round: 1.5 min and third round: 15 min) for increased stringency. Following the third round of panning, single clones were picked at random for colony PCR and Sanger sequencing of VH and VL chains. Unique sequences were identified for assessment of RBD or S1 B.1.1.7 binding using monoclonal phage ELISA.

## Next-generation sequencing and PCR recovery of enriched Fabs

The VH domain of the third-round RBD panning outputs was deep sequenced to enable targeted recovery of low-frequency Fabs using PCR. First, phage particles eluted off beads were used directly as template in PCR to amplify the VH region. These primers contained Read 1 and Read 2 adapters for paired-end sequencing on the Illumina platform. PCR products were column purified using the QIAquick PCR Purification Kit (Qiagen) and sent to Genewiz for sequencing using their Amplicon-EZ service. Paired-end sequences were merged using Geneious Prime software, followed by sequence alignment using internally developed proprietary software (Abacus). Sequences were ranked in descending order by count, and VH sequences with a frequency of at least 0.01% in the panning output were identified. In total, 164 VH sequences that had not previously been sampled in earlier screens were targeted for recovery by PCR. Next, reverse PCR primers were designed for each targeted VH sequence such that the 5′ end annealed in the CH1 domain and the 3′ end anchored in the unique HCDR3 region. Reverse primers were 72 bp in length and obtained from Integrated DNA Technologies in 11 distinct oligo pools ("oPools"). Each oPool was paired with a universal forward primer annealing in the signal peptide for light chain, enabling amplification of a DNA fragment containing VL, CL, and the targeted VH domains using phage panning eluates as a template. This fragment was gel purified, digested using appropriate restriction enzymes, and ligated back into the phagemid backbone. Sanger sequencing was performed on 384 picked colonies, and of these, 51 unique Fabs were selected for conversion to IgG using the methods described above.

## Monoclonal phage production

Single clones harboring a unique Fab fusion protein were inoculated into 150 $\mu$l 2xYT broth supplemented with 50 $\mu$g/ml carbenicillin, 15 $\mu$g/ml tetracycline, and 2% glucose and cultivated in 96 deep-well plates overnight at 37°C with rigorous shaking. Five $\mu$l of the overnight cultures was then transferred to new deep-well plates containing 100 $\mu$L 2xYT with 50 $\mu$g/ml carbenicillin, 15 $\mu$g/ml tetracycline, and 2% glucose and incubated at 37°C with rigorous shaking until an $OD_{600\,nm} \sim 0.5$. M13KO7 helper phage was added to each well at MOI 25, and plates were incubated without agitation at 37°C for 1 h before medium replacement to 2xYT containing 50 $\mu$g/ml carbenicillin and 25 $\mu$g/ml kanamycin and overnight incubation with rigorous shaking at 30°C. Phage supernatants were harvested after centrifugation and diluted 1:3 in 2% nonfat dry milk in PBS for use in the monoclonal phage ELISA.

## Monoclonal phage ELISA (antigen binding)

Binding of enriched Fabs displayed on phage to recombinant antigen was determined using phage ELISA. Briefly, 96-well MaxiSorp® assay plates (Nunc) were coated overnight at 4°C with 5 $\mu$g/ml NeutrAvidin (Thermo Fisher Scientific), then blocked in PBS containing 1% BSA for 1 h at RT. Biotinylated antigen (i.e. RBD or S1 B.1.1.7) was added at a concentration of 1 $\mu$g/ml for 30 min. Diluted phage preparations were added and allowed to incubate for 1 h at RT before captured virions were detected using a 1:5000 dilution of anti-M13-HRP (Santa Cruz Biotechnology) for 1 h at RT. All interval plate washes were performed three times in PBST (PBS supplemented with 0.1% v/v Tween-20). ELISAs were developed by the addition of TMB solution (Thermo Fisher Scientific) and quenched using 10% phosphoric acid. Absorbance was read at $A_{450\,nm}$. Plate wells coated with an irrelevant antigen (biotinylated CD40, Sino Biological, Cat. No. 10774-H08H) were processed in parallel to assess specificity of binders. Phage supernatants from two known RBD binders were included as positive controls.

## Batch conversion of fab to full-length IgG

After successful enrichment of SARS-CoV-2 binders was confirmed by phage ELISA, pooled panning outputs were subcloned into a mammalian expression vector in a manner that converts Fabs to full-length IgG while simultaneously maintaining VH-VL pairing for high-throughput screening. First, phagemid DNA was purified from phage panning outputs as follows. ER2738 cells infected with pooled phage output after the second and third rounds of panning were plated densely on 2xYT agar supplemented with 100 $\mu$g/ml carbenicillin and 2% glucose and incubated overnight at 30°C. Cells were then scraped in 1–2 ml 2xYT broth containing 50 $\mu$g/ml carbenicillin and 2% glucose, and plasmid DNA from 200 to 500 $\mu$l of the cell suspension was purified using QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's protocol.

The process of converting phage panning outputs from Fabs to full-length IgG was then performed in two steps. First, a linear fragment containing VL and VH was isolated from the phagemid backbone using restriction digest with BsiWI and NheI. This fragment was ligated into the mammalian vector such that the VL is cloned downstream of a CMV promoter and mammalian signal peptide, and VH is cloned upstream of IgG1 (i.e. CH1, hinge, CH2, and CH3 domains). This intermediate plasmid thus consists of antibody cassettes with complete light chain (LC) and heavy chain (HC) IgG1 regions but lacks the HC control elements. In the second cloning step, plasmid DNA was purified from the intermediate pool and digested with RsrII or KasI (for kappa or lambda chains, respectively) and NcoI to replace the fragment between VL and VH with a fragment containing the constant kappa or lambda domain, polyA signal, and a CMV promoter and signal peptide for HC expression. The final plasmid product contains dual CMV expression cassettes for full-length IgG expression while maintaining the VH-VL pairing of the original selected Fab fragment. A total of 487 single colonies from the batch cloning were randomly picked and cultured overnight in 0.9 ml TB supplemented with 30 $\mu$g/ml kanamycin. Colony PCR and Sanger sequencing of VH and VL domains was performed on each clone. In addition, plasmid DNA was isolated from each clone using QIAprep 96 Turbo Kit (Qiagen) following the manufacturer's protocol.

## Expression and quantitation of full-length IgG

The Expi293™ Expression System (Thermo Fisher Scientific) was utilized for transient production of full-length IgG from the 463 randomly picked sequences following batch reformatting. Expi293F cells were transfected using ExpiFectamine™ 293 reagent in 0.7 ml of culture volumes with approximately 0.7 $\mu$g plasmid DNA per well, following the manufacturer's protocol for 96-well microtiter plates. For a positive control, trastuzumab was cloned into the same mammalian expression vector and transfected in the same manner. Cells transfected with transfection reagents and no plasmid DNA served as negative controls. The transfected plates were covered with "System Duetz" sandwich covers (EnzyScreen BV, Netherlands), secured onto a 3-mm orbital platform shaker inside a humidified 37°C tissue culture incubator with 5% $CO_2$, and shaken at 1000 rpm for four days. Enhancers were added 18–24 h post-transfection. Culture supernatants were harvested by centrifuging the 96-well plates 3000 rpm for 5 min to pellet cell debris. The cleared supernatants were diluted in fresh expression medium for titer quantification on the Octet® RED96 System (ForteBio Inc, Fremont, California) using Protein A biosensors. The sample IgG titers were determined by comparison to a previously established standard curve using known IgG concentrations and reported as $\mu$g/ml.

## AlphaLISA binding screen

An AlphaLISA® (Perkin Elmer) bead-based luminescent amplification binding assay was performed to screen antibodies that specifically bound SARS-CoV-2 Spike protein and did not bind an irrelevant antigen. Briefly, 10 $\mu$l of unpurified transfection supernatant diluted 1:200 in 1x Immunoassay buffer (Perkin Elmer cat# AL000F) was preincubated with either 10 $\mu$l biotinylated SARS-CoV-2 spike protein (3 nM final concentration; R&D Systems cat# BT10549) or with 10 $\mu$l biotinylated irrelevant antigen (3 nM final concentration; R&D Systems cat# AVI10538–050) for 20 min at RT in 96 half-area white microplates (Corning cat# CLS3693). Subsequently, 10 $\mu$l of Protein A AlphaLISA® Acceptor beads (10 $\mu$g/ml final concentration; Perkin Elmer cat# AL101C) were added, and further incubated for 1 h at RT in the dark with shaking (250 rpm). Following this incubation, 20 $\mu$l of Streptavidin Donor beads (40 $\mu$g/ml final concentration; Perkin Elmer cat# 6760002S) were added, and the microplates were further incubated for 1 h at RT in the dark with shaking (250 rpm) and read on FLUOStar® Omega multimode microplate reader instrument (BMG Labtech) (excitation 680 nm, emission 615 nm). Data were graphed using GraphPad Prism software.

For dose-dependent binding, positive unpurified supernatants were normalized to a starting concentration of 300 ng/ml (final concentration) and 9 point 1:3 serial dilutions were performed in 1x Immunoassay buffer prior to preincubation with biotinylated SARS-CoV-2 spike protein (3 nM final concentration) and an AlphaLISA® assay was performed as described above.

## Inhibition of SARS-CoV-2 spike binding to human angiotensin converting enzyme 2

Antibody supernatants that specifically bound SARS-CoV-2 spike protein were tested for their ability to block binding of SARS-CoV-2 spike protein to human angiotensin converting enzyme 2 (ACE-2). Briefly, MaxiSorp™ ELISA plates (Biolegend cat# 50–712-278) were coated with human ACE-2 Fc chimera protein (R&D Systems cat# 10544-ZN) in cold PBS (100 $\mu$l per well; 0.5 $\mu$g/ml) overnight at 4°C. The plates were then washed using PBS with 0.05% Tween 20 (wash buffer) and blocked with PBS, 1% (v/v) BSA for 1 h at 37°C in a 5% $CO_2$ incubator. Unpurified transient transfection supernatant samples (50 $\mu$l) were preincubated with biotinylated SARS-CoV-2 spike protein (R&D Systems cat# BT10549) (50 $\mu$l, 0.3 $\mu$g/ml final concentration in PBS with 1% BSA and 0.05% (v/v) Tween 20 (binding buffer)) for 20 min at RT, then transferred to the ELISA plates and incubated for 2 h at RT with shaking (250 rpm). After this incubation, the plates were washed three times with wash buffer, and 100 $\mu$l Streptavidin—Horseradish Peroxidase (Thermo Scientific cat# N100) (1:5000 dilution in binding buffer) was added onto the wells and incubated for 45 min at RT. After a 3× wash step with wash buffer, the plates were developed using 3,3′,5,5′-tetramethylbenzidine (TMB) substrate (100 $\mu$l; Pierce cat# 34022), and the reaction was quenched using phosphoric acid 10% (v/v) (100 $\mu$l; Ricca Chemical Company cat# 5850–16). The well absorbance was read at 450 nm on FLUOStar® Omega multimode microplate reader instrument (BMG Labtech). Antibody blockade, represented as percent inhibition relative to the binding of spike protein alone, was calculated and graphed using GraphPad Prism software.

For dose-dependent inhibition, unpurified supernatants were normalized to a starting concentration of 30 $\mu$g/ml (final concentration) and 10 point 1:3 serial dilutions were performed in binding buffer prior to preincubation with biotinylated SARS-CoV-2 spike protein (0.3 $\mu$g/ml final concentration) and ELISA assay performed as described above.

## rVSV production

The pseudotyped ΔG-luciferase (G*ΔG-luciferase) rVSV was purchased from Kerafast (catalog number: EH1020-PM) and was used to produce VSV-ΔG pseudotyped with the SARS-CoV-2 spike protein. Briefly, HEK-293 T cells were plated in T75 flask and incubated at 37°C, 5% CO2. Twenty-four hours postseeding, cells were transfected with 12 $\mu$g of plasmid expressing the spike from various variants using Fugene-6 (Promega; catalog number: E2691) according to the manufacturer's instructions and incubated at 37°C, 5% CO2. One hour later, cells were infected at a MOI of 5 with pseudotyped ΔG-luciferase (G*ΔG-luciferase) rVSV and incubated at 37°C, 5% $CO_2$ overnight. In the morning, cells were washed three times with PBS, fresh media was added. Cells were incubated for approximately 24 h at 37°C, 5% CO2, and the supernatants were then collected, clarified by low-speed centrifugation (1320 × g for 10 min), and were aliquoted and stored at −80°C. The rVSVΔG-SARS-CoV-2-spike proteins were titered by $TCID_{50}$ (50% tissue culture infectious dose). Briefly, pseudotyped virus titration is carried out by using 96-well plates. A549-hACE2/hTMPRSS2 cells were infected with 10-fold serial dilutions in four replicates. After infection with the pseudotyped virus at 37°C, and 5% CO2, for 24 h, the luciferase substrate was added for chemiluminescence detection (Promega; Luciferase Assay system; catalog number: E1500). The detected raw data were used to calculate the $TCID_{50}$ according to the Reed–Muench method. The positivity of the luciferase readout was set to four times the luminescence activity of the uninfected control.

## Spike variants used

The spike from the following variants of interest and concern were used to generate rVSVΔG-SARS-CoV-2-spike pseudotypes: SARS-CoV-2 USA-WA1/2020 (Lineage A) SARS-CoV-2 USA-WA1/2020, D614G (Lineage A), SARS-CoV-2 hCoV-19/England/204820464/2020, variant alpha (Lineage B.1.1.7), SARS-CoV-2 hCoV-19/South Africa/KRISP-K005325/2020, beta variant (Lineage B.1.351), SARS-CoV-2 hCoV-19/Japan/TY7–503/2021, gamma variant (Lineage P.1),

SARS-CoV-2 hCoV-19/USA/PHC658/2021, delta variant (lineage B.1.617.2), SARS-CoV 2 hCoV-19/USA/MD-HP20874/2021, omicron BA.1 variant (Lineage B.1.1.529.1), hCoV-19/South Africa/CERI-KRISP-K032307/2021, omicron BA.2 variant (Lineage B.1.1.529.2), hCoV-19/South Africa/NICD-N37417/2022, and omicron BA.5 variant (Lineage B1.1.529.5).

## Neutralization assay with rVSV∆G-SARS-CoV-2-spike

For the neutralization procedure, the virus corresponding to a MOI of 0.1 was incubated with increasing concentrations (10-fold serial dilutions) of test antibodies for 1 h at 37°C to allow the antibody to bind to the viral surface envelope glycoproteins. A549-hACE2/TMPRSS2 cells were then infected with the virus/antibody mixture and incubated for 20 h at 37°C, 5% CO2. At 20 h postinfection, the luciferase was detected using the Luciferase Assay system (Promega; catalog number: E1500) according to the manufacturer's instructions. Briefly, the cells were washed once with PBS, lysed using 1x Passive Lysis Buffer for 30 min. The luciferase substrate was then added, and luminescence was measured using a 96-well plate reader (Envision microplate reader; Perkin Elmer) with an integration time of 0.1 s per well.

The results were normalized to a percentage of neutralization (Neutralization %) where the cell control (noninfected condition) and the virus control (nontreated, infected condition) are used to set 100% neutralization and 0% neutralization, respectively.

## Neutralization assay with SARS-CoV-2 isolates

For the neutralization procedure, the virus at a MOI of 0.1 was incubated with increasing concentrations (10-fold serial dilutions) of S309 (11) or test antibodies for 1 h at 37°C to allow the antibody to bind to the viral surface envelope glycoproteins. A549-hACE2/TMPRSS2 cells were then infected with the virus/antibody mixture and incubated for 40 h at 37°C, 5% CO2. Remdesivir at 5 $\mu$M was also used as a control (Sigma-Aldrich, St. Louis, MO, USA). At 40 h, postinfection infectivity was assessed by quantifying intracellular viral genomes by RT-qPCR. Briefly, the supernatant was removed and cells were washed once with PBS 1X. Cells were then lysed with 100 $\mu$l of lysis buffer. Total RNA was extracted using the SV 96 Total RNA Isolation System and Vac-Man 96 (Promega; catalog number: Z3500), according to the manufacturer's instructions. Taqman Fast Virus 1-step kit (Thermo Fisher; catalog number 4444434) with oligos and probe specific to the N gene of SARS-CoV2 was used for SARS-CoV-2 RNA detection according to the manufacturer's instructions. Standard range was prepared by 10-fold serial dilution of qPCR Control RNA from inactivated SARS-CoV-2. The QuantStudio 7 Flex cycler (Thermo Fisher) was used for amplification. The results were expressed as copies of viral genome equivalent per $\mu$l of RNA.

## SARS-CoV-2 isolates

Isolates USA-WA1/2020 (A), cat. no.NR-52281, hCoV19/USA/PHC658/2021 (Delta B.1.617.2), cat. no. NR-55611 and hCoV-19/USA/MD-HP20874/2021 (Omicron B.1.1.529, BA.1), cat. no. NR-56461 were obtained through BEI Resources, NIAID, NIH. Isolate hCoV-19/France/ARA-HCL022074071401/2022 (Omicron B.1.1.529.5.3, BA5) was provided by the Hospices civils de Lyon.

## Surface Plasmon resonance kinetics characterization

Surface plasmon resonance (SPR) was used to determine the affinities of selected candidates for spike proteins from multiple SARS-CoV-2 strains (Wuhan (lineage A), Denmark—Mink Cluster-5 (lineage B.1.1.298), Japan (or Brazil) (lineage P.1), USA/CA San Diego (lineage B.1.1.7), Indian-Kappa variant (lineage B.1.617.1), UK variant (lineage B.1.1.7), and South African variant (lineage B.1.351)). Briefly, kinetics measurement was performed by SPR using the Carterra LSA instrument (carterra-bio.com). The instrument preparation and operation were conducted as recommended by the manufacturer. All kinetics measurements were performed at 25°C. Purified anti-SARS-CoV-2 antibodies of the invention were printed for 10 min using 20 $\mu$g/ml material onto sensor chips conjugated with anti-human Fc capture antibody. Appropriate controls were printed alongside novel antibodies to confirm capture integrity of sensor chip and associated system validation. Purified spike protein material was serially diluted to eight concentrations starting at 500 nM. Buffer blanks were passed over the sensor chip prior to spike protein injection for sensor array chip preparation and to assess baseline signal level. Each spike protein kinetic series was injected from low to high concentration to generate eight sensorgrams per antibody:spike variant interaction. Replicates for each 8-sensorgram set were performed. The data were processed and analyzed by Carterra's KIT software tool by interspot referencing and double referencing the data and then by fitting them to a simple Langmuir binding model using global $k_a$, $k_d$, and $R_{max}$ values per spot. Fit quality was determined by inspection of the residuals.

## Biophysical characterization sample preparation

Samples were buffer exchanged against 10 diavolumes of 20 mM sodium phosphate, 150 mM sodium chloride, pH 7.1 (PBS) using a centrifugal filter with a 30 kDa molecular weight cut off (Amicon). After buffer exchange, samples were normalized to 1 mg/ml using a Lunatic protein concentration plate format instrument (Unchained Labs).

## Differential scanning fluorimetry

Thermal transition temperature(s) and weighted shoulder scores (WSS) were determined by DSF according to the method previously described [5]. A single biological sample was divided, and the assay ran twice per molecule type. Additional information was also obtained from a parameter termed the WSS, which accounts for multiple pieces of information from the unfolding curve [6].

## Low pH stability

Stability during a low pH hold was determined according to the method previously described [5]. The increase in high molecular weight of the low pH exposed sample as compared with the control sample is reported.

## Chemical unfolding

The chemical unfolding assay was completed as previously described [7] with some modifications. After a three-day incubation in 32 guanidine hydrochloride (GND) concentrations, the samples were measured on a Fluorescence Innovations SUPR-UV plate reader (excitation: 275 nm, emission: 300–440 nm). The measured fluorescence intensity at 362 nm was corrected for scattering and stray light, the unfolding curve was generated by graphing each corrected intensity against the GND concentration and the inflection point is reported.

## Relative solubility

Solubility was assessed according to the method previously described [8]. Analysis was done in PBS buffer (20 mM sodium phosphate and 150 mM sodium chloride pH 7.1) and a final PEG

10 000 concentration ranging from 7.2% to 9.6%. Percent recovery relative to a 0% PEG control was determined and average recovery across the PEG concentration range is reported.

### Self-interaction nanoparticle spectroscopy

Self-interaction nanoparticle spectroscopy (SINS) measurements were performed according to the method previously described [9]. Briefly, gold nanoparticles (Ted Pella) were conjugated overnight with an 80:20 ratio of antihuman and antigoat antibodies (Jackson Immuno Research). Unreacted sites were blocked using an aqueous 0.01% (w/v) polysorbate 80 solution. Conjugated gold nanoparticles were then concentrated by centrifugation and removal of 95% of the supernatant. Analysis was carried out in PBS (20 mM phosphate, 150 mM NaCl, pH 7.1) at a protein concentration of 0.05 mg/ml reacted with 5 $\mu$l of concentrated conjugated gold nanoparticles. After a 2-h incubation, absorbance spectrum from 400–600 nm was collected using a Spectrostar Nano plate reader at 2 nm steps. The wavelength maximum of the spectrum peak is reported. Three assay replicates were made from each of the pooled expressed antibodies.

### Standup monolayer absorption chromatography

Standup monolayer absorption chromatography (SMAC) measurements were performed according to the method previously described [10]. Retention times were determined using a Dionex UPLC equipped with a Zenix column (Sepax Technologies) and a running buffer comprised of 150 mM sodium phosphate pH 7.0.

## Results

### J.HAL® library construction

Two different phagemids were used for construction of the GAN discovery library to allow for high-level display(pADL-20c) and low-level display (pADL-10b) of a gene of interest We hypothesized that binders would be easier to obtain from the pADL-20c phagemid due to higher display levels, but that the clonal distribution would be more normalized in the pADL-10b phagemid due to reduced effect of toxicity on growth rate. To increase the likelihood of isolating binders against the target of interest, the GAN antibody sequences were cloned into both phagemids independently, resulting in the production of two distinct libraries, each containing approximately $3.5 \times 10^{10}$ transformants representing a theoretical diversity of 1 billion VH and VL pairs. Sanger sequencing of VH and VL from randomly picked colonies indicated that approximately 70% of clones in the library had correctly assembled pairs. The remaining 30% of clones had nonproductive pairs, including stop codons, frameshifts, or truncated inserts. Fab-displaying phages were produced from each library. Phage ELISA was used to assess expression and display of the GAN-generated Fab fragments on phage (Fig. S1). Display levels were estimated by capturing serial dilutions of purified phage on ELISA plates coated with antihuman Fab and detecting with anti-M13 antibodies conjugated to HRP. There is considerable variability in display levels among the various germline pairings, but in general, all pairs tend to fall between the high- and low-display control Fabs (M-6240 and M-6239, respectively). Furthermore, germline pairs with IGHV5–51, IGKV4–1, and IGLV1–40 appear to have lower display levels than other pairs in this library. Whether or not these differences are significant or represent differential binding of the anti-human Fab capture antibody used in the ELISA is an area of further investigation. Germline pairs cloned into the higher display phagemid pADL-20c had similar trends as those in pADL-10b (data not shown).

### Biopanning and isolation of anti-SARS-CoV-2 Fabs

Biopanning was performed independently on two different SARS-CoV-2 antigens (Wuhan RBD and B.1.1.7 S1 (Sino Biological cat# 40592-VNAH and 40 591-V08H12, respectively)) to isolate SARS-CoV-2 binders from the phage libraries. Biopanning was performed with increasing stringency of selection at each round, ranging from 100 to 10 nM biotinylated antigen. Following the third round of panning, single clones were picked at random for colony PCR and Sanger sequencing of VH and VL chains. Unique sequences were identified for assessment of RBD or S1 B.1.1.7 binding using monoclonal phage ELISA.

In parallel, ELISA assays were performed on CD40, an irrelevant antigen, to verify the specificity of the binders for RBD. Duplicate clones harboring the same Fab sequence were included, if possible. A total of 22 unique Fab sequences were identified as binding specifically to RBD. This was defined as having $OD_{450}$ values greater than twice the $OD_{450}$ signal as compared to binding to CD40 (Fig. S2) thus confirming the efficiency of the selection strategy. The two positive control Fabs included in the assay were verified to be specific for RBD binding. Based on these results, panning was terminated after the third round, and selected phage outputs were converted in batch to full-length IgG for high-throughput screening and characterization.

### NGS and PCR recovery of enriched Fabs

To identify additional Fabs targeting the RBD of SARS-CoV-2, a high-throughput screen was implemented in parallel with random clone screening. The workflow used NGS coupled with PCR to recover less represented clones within the enriched third round outputs, which are not likely to be sampled by random colony picking. We first used PCR to amplify the VH region directly from the phage eluates after the third round of selection against RBD for analysis of clonal diversity by NGS. Based on these results, VH sequences with frequencies of 0.01% or higher that had not previously been tested by random screening were identified. We next designed pools of reverse PCR oligos that began in the CH1 domain and anchored in the unique HCDR3 region of each of those sequences. Using a universal forward primer that anneals in the signal peptide for light chain, we were able to recover linear fragments linking the targeted VH to its paired VL (VL-constant light chain-VH). The linear fragments were cloned back into the phagemid backbone, and four plates of single clones were picked for Sanger sequencing of the VH and VL domains. Of the 90 unique sequences identified in this panel, 68 were unique and had not been previously tested for binding to RBD. We selected 51 for IgG reformatting and transient mammalian expression.

### Fab to full-length IgG conversion and mammalian expression

We used a high throughput, two-step batch cloning method to convert enriched Fabs to full-length IgG, while simultaneously maintaining VH-VL linkage, for mammalian expression and further characterization. Following conversion of the Fabs to IgG, Sanger sequencing was performed on >400 picked colonies to obtain VH and VL sequences. Clones that contain the same sequence, as well as clones with nonproductive sequences containing frameshifts, stop codons, or truncations, are removed. Unique productive sequences were then transiently expressed in a mammalian system alongside positive and negative expression controls, and after four days, antibody titers were quantified using

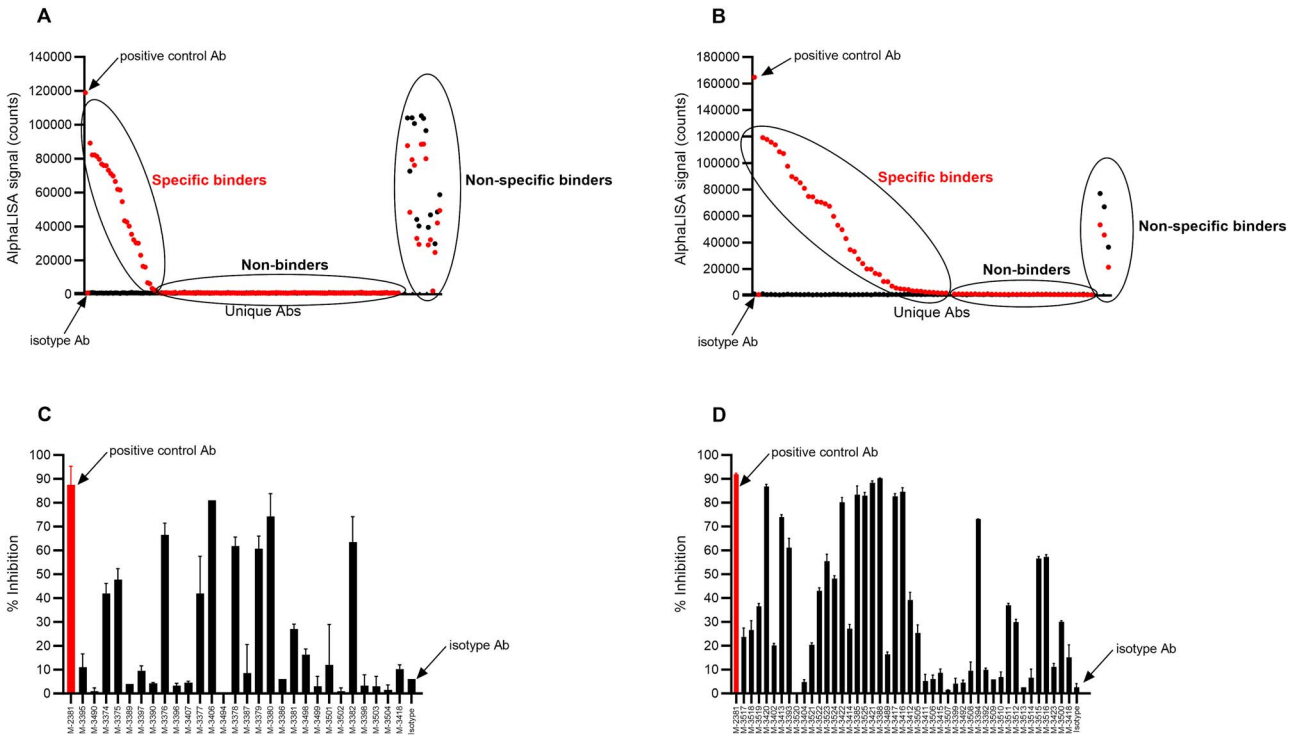


Figure 2. High-throughput binding and bioactivity screen allows identification of unique J.HAL® IgG candidates that bind SARS-CoV-2 spike protein and effectively block spike:Human ACE2 receptor interaction. (A and B) Unpurified transfection supernatants from two independent screens were tested by AlphaLISA for binding to biotinylated SARS-CoV-2 spike protein and in parallel to biotinylated irrelevant target protein. Positive control antibody M-3418 (S309) demonstrates binding specificity, whereas isotype control antibody demonstrates lack of binding. Data were graphed using GraphPad prism software. (C and D) Unpurified transfection supernatants from two independent screens were tested for their ability to block binding of SARS-CoV-2 spike protein to human ACE-2 using ELISA. Positive control antibody M-2381 (red symbol) demonstrates effective blockade, whereas isotype control antibody demonstrates lack thereof. Data were graphed as % inhibition +/− SD using GraphPad prism software.

bio-layer interferometry with Protein A biosensors. About, 88.1% of transfectants had detectable antibody titers, ranging from less than 1 up to 345 $\mu$g/ml (Fig. S3).

## AlphaLISA specificity binding screen

For accelerated discovery timelines, we developed a highly sensitive bead-based luminescent amplification binding assay using AlphaLISA technology that allows high-throughput screening of large antibody panels without the need for purification. Unpurified mammalian transfection IgG supernatants were assessed for binding to biotinylated SARS-CoV-2 spike protein, and in parallel for binding to biotinylated irrelevant antigen. Although phage hits were enriched via biopanning with either Wuhan RBD or B.1.1.7 S1 antigens, we utilized Spike protein for binding screen, as it contains RBD/S1 in a more physiological conformation. Positive and negative control antibodies, (S309 antibody [11] and isotype control antibody, respectively) expressed or spiked in transient transfection supernatants at concentrations matching the range of titers, were included to monitor assay performance. This allowed for fast screening and differentiation between nonbinders, nonspecific binder,s and specific binders to the spike protein (Fig. 2A and B). Since at this point, the supernatants are not concentration normalized, the AlphaLISA signal is a function of both affinity and titer, and thus binders cannot be ranked. Rather, this serves as an efficient process for down-selecting promising candidates that can then be further inquired.

A total of 73 unique antibody sequences specific for SARS-CoV-2 spike protein were identified: 66 unique antibody sequences identified from the biopanning with Wuhan RBD (20 of which were recovered via NGS mining), and 7 additional unique antibody sequences identified from the biopanning with B.1.1.7 S1 protein. Positive supernatants identified in the initial binding screen were concentration normalized and dose-dependent binding to the SARS-CoV-2 spike protein was assessed by AlphaLISA method described above (Figs 3 and S4). Thus, binding profiles of candidates were evaluated, and candidates were ranked relative to each other and to the positive control S309 antibody. S309 antibody, a high-affinity anti-SARS-CoV-2 spike monoclonal antibody [11] displayed dose-dependent binding with an $EC_{50}$ value of 4.2 ng/ml and an Emax value of 121 545 (counts) (Table S1). All J.HAL® antibodies tested displayed dose-dependent binding, with varying binding potencies ($EC_{50}$ values) and maximum binding (Emax values). Several J.HAL® antibodies displayed strong binding profile, comparable, albeit inferior to S309 control. For example, antibodies, M-3376, M-3406, and M-3382 displayed potent binding to the cognate antigen, with $EC_{50}$ values of 8.9, 7.2, and 4.4 ng/ml, respectively and Emax values of 75 566, 80 909, and 80 143 (counts), respectively. This demonstrates the potential of the J.HAL® library to produce specific binders with strong binding profile in absence of affinity maturation.

## Antibody inhibition screening of spike: Human ACE-2 receptor interaction

To quickly identify antibodies that block binding of SARS-CoV-2 spike protein to human ACE-2 receptor, unpurified antibody supernatants that specifically bound SARS-CoV-2 spike protein were tested for their ability to block binding of SARS-CoV-2 spike

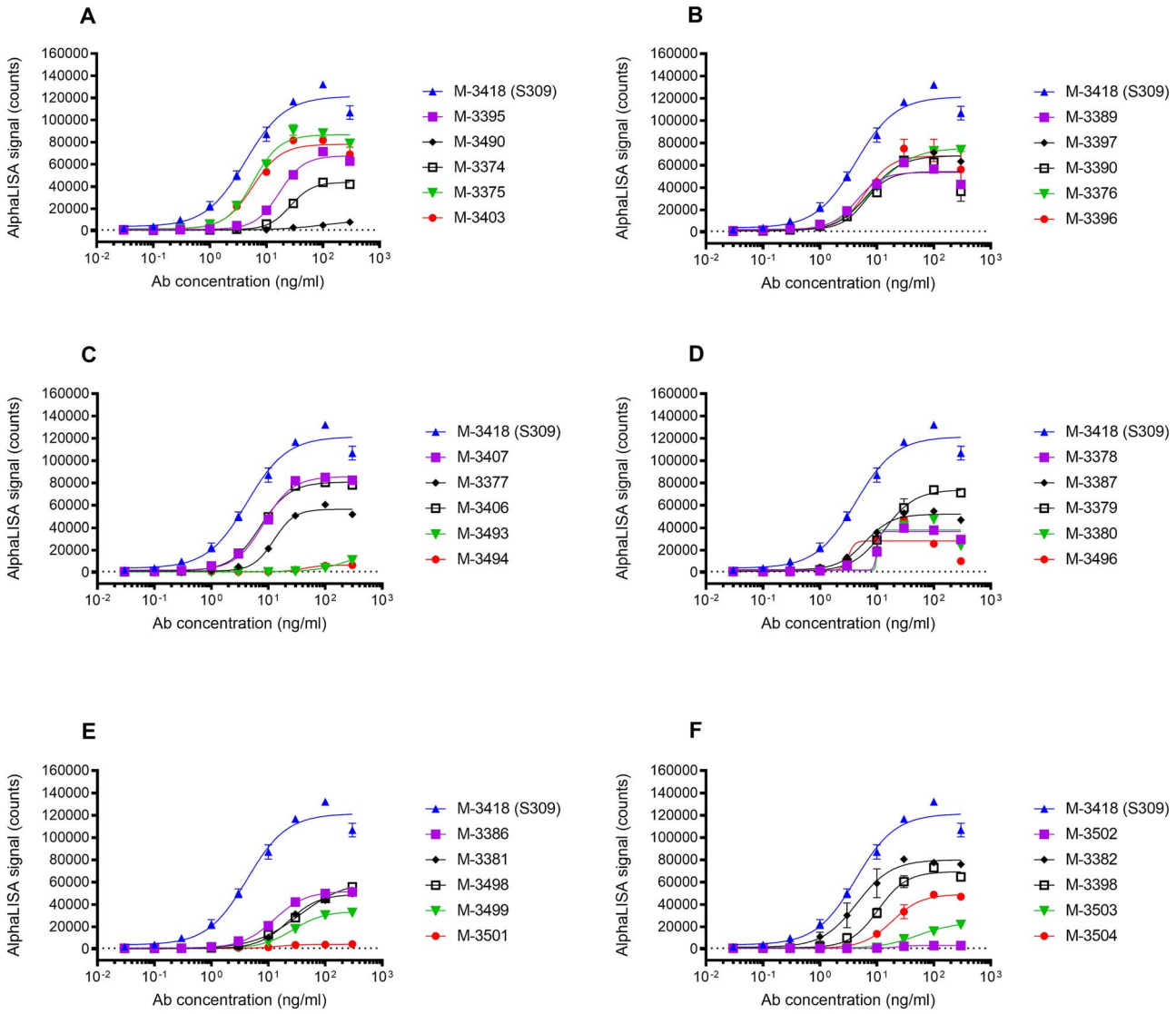


Figure 3. J.HAL® IgG antibodies exhibit dose-dependent binding to SARS-CoV-2 spike protein. Unpurified transfection supernatants were concentration normalized and tested for binding in a 9-point serial titration series to biotinylated SARS-CoV-2 spike protein by AlphaLISA. Representative binding profiles for 30 J.HAL® antibodies are shown in the subplots A-F above, along with the positive control antibody M-3418 (S309). Data were graphed as mean +/− SD using GraphPad prism software.

protein to human ACE-2 by functional ELISA. Antibody blockade, represented as percent inhibition relative to the binding of spike protein alone, was calculated and graphed using GraphPad Prism software. Positive control included 2381 blocking antibodies isolated from a convalescent patient [12, 13] spiked at 30 $\mu$g/ml in transient transfection supernatant; isotype control antibody, expressed or spiked in transient transfection supernatants at concentrations matching the range of titers, were included to monitor assay performance. Replicates were included when possible and/or data were reproduced in independent experiments. Antibody blockade data are shown in Fig. 2C and D (Mean % Inhibition +/− SD). Since the supernatants are tested neat and are not concentration normalized, the % Inhibition value is a function of both potency and titer, and thus blocking antibodies cannot be ranked, rather selected for further analysis. Thirt-six IgG candidates demonstrated ≥30% blockade and were selected for further analysis. Their respective potencies were further characterized in dose-dependent studies relative to positive benchmark control antibody (Fig. S5). A range of potencies were observed, which is as expected. Several IgG candidates demonstrated effective blockade of spike protein binding to human ACE-2 receptor, notably M-3422, M-3388, and M-3406 displaying $IC_{50}$ values of 38, 70, and 618 ng/ml, respectively. (Table S2).

## Evaluation of the neutralization activity of IgG candidates using rVSV pseudotyped with the spike of SARS-CoV-2 UK variant (B1.1.7, Alpha)

To evaluate their potency to neutralize SARS-CoV-2 infection, IgG candidates were tested for their ability to neutralize the infection of an rVSV pseudotyped with the spike of the UK variant (B1.1.7, Alpha). For that, two successive screenings were performed in the A549-ACE2/TMPRSS2 cell line. The neutralization activity of IgG candidates was evaluated at 1 and 10 $\mu$g/ml. Infection data were compared to infected (rVSVΔG-SARS-CoV-2-Spike) and not infected controls. Among all IgG candidates evaluated in the first screen, M-3375, M-3380, M-3403, M-3376, M-3406, M-3412, and M-3420 demonstrated ≥70% of neutralizing activity at 10 $\mu$g/ml (Fig. 4A and B). Among these, M-3380 and M-3406 were also capable to inhibit infection by 70% at 1 $\mu$g/ml. These seven IgG

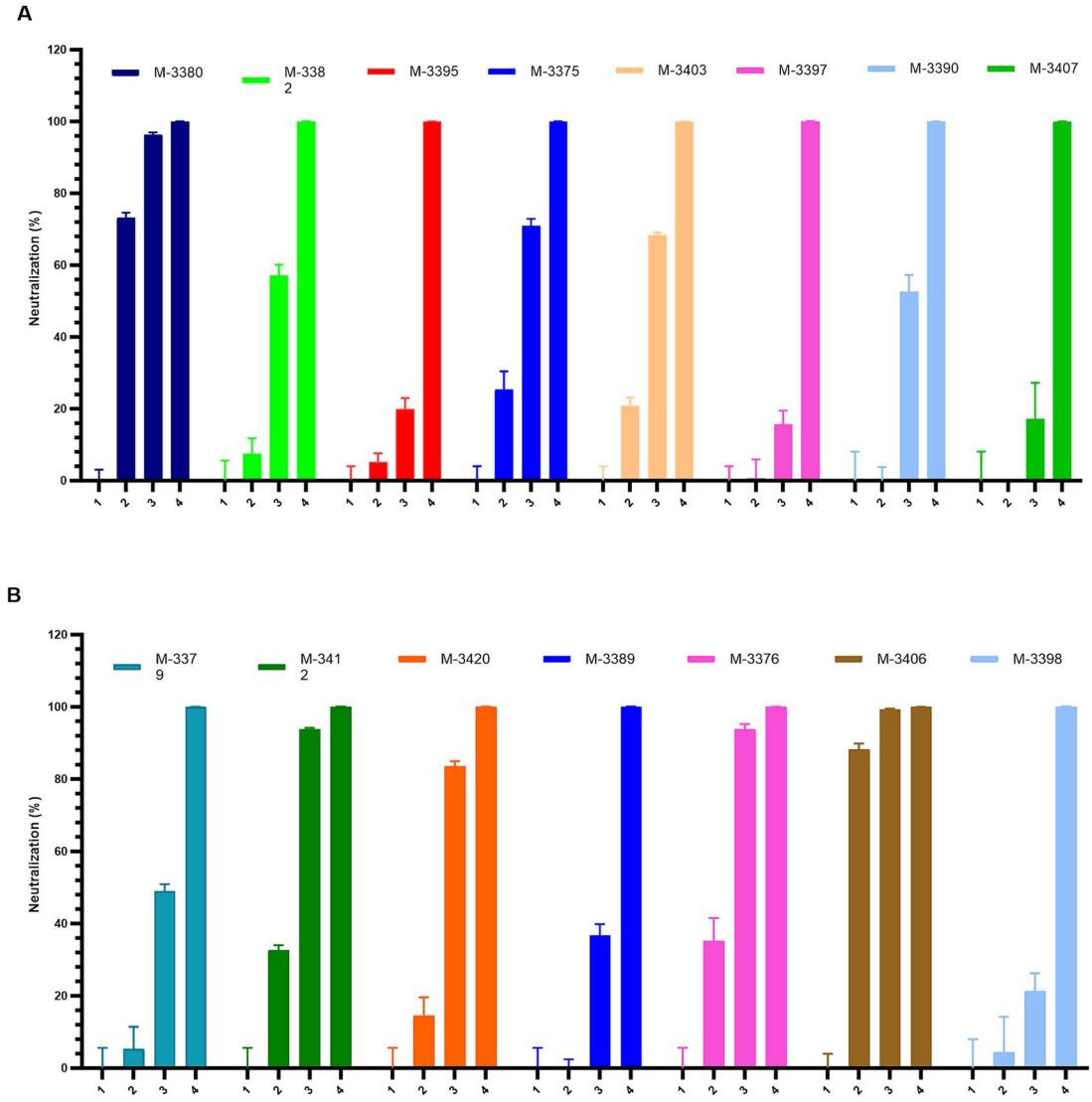


Figure 4. Neutralization activity of J.HAL® IgG candidates on VSV pseudotyped with SARS-CoV-2 UK variant spike (B1.1.7, Alpha) in A549-ACE2/TMPRSS2. (A and B) IgG candidates were tested for their ability to neutralize the entry of VSV pseudotyped with the spike of SARS-CoV-2 UK variant (B1.1.7, Alpha) in A549-ACE2/TMPRSS2 cells (A) M-3380 to M-3407 and (B) M-3379 to M-3398. For neutralization experiments, the pseudovirus was preincubated with two different concentrations of IgG candidates for 1 h at 37°C and the pseudovirus/antibody was subsequently added onto the cells and incubated for 20 h at 37°C. Infectivity was assessed at 20 h postinfection by luciferase assay. The results were normalized to a percentage of neutralization where the cell control (not-infected condition) and the virus control (nontreated, rVSVΔG-SARS-CoV-2-spike) are used to set 100% neutralization and 0% neutralization, respectively. Data were graphed as of % neutralization + SD using GraphPad prism software.

candidates were selected for further analysis. No IgG candidates were selected from the second screen (data not shown).

## Evaluation of the cross-neutralization activity of selected IgG candidates on a panel of rVSVΔG-SARS-CoV-2-spike pseudotypes

To determine whether the neutralizing activity of the eight IgG candidates is conserved across SARS-CoV-2 variants, neutralization assays were performed on a panel of rVSVΔG-SARS-CoV-2-Spike pseudotypes in A549-ACE2/TMPRSS2 cell line. This includes spike proteins from variants of interest and concern: Wuhan (lineage A), D614G mutant (lineage A), Alpha (lineage B.1.1.7), Beta (lineage B.1.351), Gamma (lineage P1), Delta (lineage B.1.617.2), Omicron BA.1 (lineage B.1.1.529), Omicron BA.2 (lineage B.1.1.529.2), and Omicron BA.5 (lineage B.1.1.529.5). To determine their 50% neutralization dose ($ND_{50}$), defined as the concentration of antibody that reduced the number of infected cells by 50%, increasing concentrations of IgG candidates ranging from 1 ng/ml to 10 $\mu$g/ml were used. Table 1 summarizes the $ND_{50}$s calculated for all the IgG candidates. Results indicated that IgG candidates can be classified into three groups. In the first group, M-3380 and M-3420 showed a potent neutralization across Wuhan, D614G, Alpha, and Delta variants with an $ND_{50}$ between 0.12 and 2.1 $\mu$g/ml, but they showed no effect on variants Beta, Gamma, and Omicron. In group 2, M-3375, M-3403, and M-3412 showed a decreased neutralization potency against Wuhan and D614G variants compared to Alpha but they were highly potent against Beta and Gamma variants with an $ND_{50}$ between 0.2 and 0.45 $\mu$g/ml. As for group 1, they showed no effect on Omicron variant. Finally, group three contains two complementary IgG candidates, M-3376 and M-3406. Both showed a potent neutralization across Wuhan, D614G, Alpha and Delta variants but M-3406 was the most potent with an $ND_{50}$ between 0.04 and 0.31 $\mu$g/ml. Conversely, only M-3376

Table 1. J.HAL® IgG candidates' neutralization activity on a panel of VSV pseudotyped with SARS-CoV-2 spike in A549-ACE2/TMPRSS2 cell line.

| | | | $ND_{50}$ ($\mu$g/ml) | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Lineage** | **Variant** | **Strains** | **M-3380** | **M-3420** | **M-3375** | **M-3403** | **M-3412** | **M-3376** | **M-3406** |
| A | / | SARS-CoV-2 USA-WA1/2020 | 1.1 | 2.1 | >10 | 9.25 | 7.4 | 1.31 | 0.22 |
| A | / | SARS-CoV-2 USA-WA1/2020, D614G | 1.4 | 1 | 3.7 | 6.4 | 6.35 | 1 | 0.15 |
| B1.1.7 | Alpha | SARS-CoV-2 hCoV-19/England/204820464/2020 | 0.5 | 1.6 | 1.5 | 5.3 | 1.3 | 1.42 | 0.31 |
| B1.351 | Beta | SARS-CoV-2 hCoV-19/South Africa/KRISP-K005325/2020 | >10 | >10 | 0.3 | 0.2 | 0.4 | 0.68 | >10 |
| P.1 | Gamma | SARS-CoV-2 hCoV-19/Japan/TY7–503/2021 | >10 | >10 | 0.4 | 0.25 | 0.45 | 0.13 | >10 |
| B1.617.2 | Delta | SARS-CoV-2 hCoV-19/USA/PHC658/2021 | 0.12 | 0.6 | >10 | >10 | >10 | 0.35 | 0.04 |
| B1.1.529.1 | Omicron BA.1 | SARS-CoV-2 hCoV-19/USA/MD-HP20874/2021 | >10 | >10 | >10 | >10 | >10 | 0.2 | >10 |
| B1.1.529.2 | Omicron BA.2 | SARS-CoV-2 hCoV-19/South Africa/CERI-KRISP-K032307/2021 | nd | >10 | nd | nd | nd | 0.1 | >10 |
| B1.1.529.5 | Omicron BA.5 | SARS-CoV-2 hCoV-19/South Africa/NICD-N37417/2022 | nd | nd | nd | nd | nd | >10 | 0.7 |
| | | | Group 1 | | Group 2 | | | Group 3 | |

The neutralizing effects of IgG candidates on entry of several rVSVΔG-SARS-CoV-2-Spike pseudotypes was evaluated in A549-ACE2/TMPRSS2 cells. VSV carrying the spike from Wuhan, Wuhan D614G mutant, Alpha, Beta, Delta, Gamma, Omicron BA.1, BA.2, and BA.5 were used. For neutralization experiments, the pseudovirus was preincubated with a dose range of IgG candidates for 1 h at 37°C and the pseudovirus/antibody was subsequently added onto the cells and incubated for 20 h at 37°C. Infectivity was assessed at 20 h postinfection by Luciferase assay. The 50% neutralization dose ($ND_{50}$), defined as the concentration of antibody that reduced the number of infected cells by 50% calculated using GraphPad Prism software are indicated in the graph. IgG candidates were classified in three groups. $ND_{50} \leq 1$ $\mu$g/ml, 1 $\mu$g/ml < $ND_{50}$ < 10 $\mu$g/ml, $ND_{50} > 10$ $\mu$g/ml. nd indicates not done.

was capable to neutralize Beta, Gamma and Omicron BA.1 and BA.2 variants with an $ND_{50}$ between 0.1 and 0.68 $\mu$g/ml. However, M-3376 was inactive against BA.5 variant contrary to M-3406. Detailed results obtained for M-3376 and M-3406 are depicted in Fig. 5. Since M-3376 and M-3406 were highly effective against the most recent variants of concern Delta and Omicron, they were selected for further analysis against SARS-CoV-2 isolates.

## Confirmation of the cross-neutralization activity of IgG candidates M-3376 and M-3406 on a panel of SARS-CoV-2 isolates

To validate their potency, the capacity of IgG candidates M-3376 and M-3406 to neutralize SARS-CoV-2 isolates, WA.1, Delta, Omicron BA.1, and Omicron BA.5 was evaluated in A549-ACE2/TMPRSS2 cell line. A negative control antibody was used in parallel as well as two positive controls, the S309 neutralizing antibody [11] and remdesivir, a nucleoside analogue inhibiting SARS-CoV-2 replication [14]. Neutralization was performed by incubating increasing concentrations of IgG candidates ranging from 0.01 to 100 $\mu$g/ml with the virus at MOI 0.1 for 1 h prior to infection of cells. Remdesivir was added 1 h postinfection. Viral titers (expressed as genome equivalent per $\mu$l of RNA) were quantified in cell lysates by RT-qPCR 40 h post-treatment and results of treated conditions were compared to the infected control. Results indicated that contrary to the negative control antibody, S309 was effective against the four isolates (Fig. 6). Indeed, the viral titers decreased in a dose dependent-manner to reach 1–1.5-log reduction at the highest concentration tested (100 $\mu$g/ml). Remdesivir also reduced viral replication of WA.1, Delta, and Omicron BA.1 and BA.5 isolates by 2-log, 2.5-log, and 3-log, respectively. A slightly higher reduction level (2.5-log) was observed against WA.1 isolate using M-3376 and M-3406 at the highest concentration (100 $\mu$g/ml) (Fig. 6A). However, no antiviral effect was observed at the other concentrations tested. These two IgG candidates were more potent than remdesivir and S309 against the Delta isolate with a 2.5–3-log reduction observed for M-3376 at 10 $\mu$g/ml and for M-3406 at 1 $\mu$g/ml (Fig. 6B). However, only M-3376 conserved an antiviral activity comparable to remdesivir and superior to S309 against Omicron BA.1 with a 1.5-log and 3-log reduction using 10 and 100 $\mu$g/ml, respectively (Fig. 6C). Conversely, only M-3406 showed an antiviral activity comparable to remdesivir and superior to S309 against Omicron BA.5 with a 1.5-log and 3-log reduction using 10 and 100 $\mu$g/ml, respectively (Fig. 6D) Altogether, these results indicate that M-3376 and M-3406 were able to neutralize SARS-CoV-2 isolates. They also confirm their complementarity since M-3376 neutralized Omicron BA.1 isolate and M-3406 neutralized Omicron BA.5 was superior to neutralize Delta isolate.

## Kinetic binding affinity assessment

To further assess the target-specific binding profile of the lead candidates, kinetic binding profiles of the lead candidates against eight clinically relevant strains were assessed by SPR using the Carterra LSA instrument (carterra-bio.com). All kinetics measurements were performed at 25°C using both purified and unpurified IgG candidates. Many of the J.HAL® lead candidates displayed respectable target binding affinity ($k_D$ 1–10 nM range) and cross-reactivity against spike target antigens from several viral strains of interest (Fig. S6). Additionally, we observed good correlation between the SPR binding profiles of purified candidates and unpurified supernatants (data not shown), validating the feasibility to use SPR technology to assess pan cross-reactivity of unpurified transfection supernatants to accelerate discovery timelines.

## Biophysical characterization

Figure 7 represents the biophysical performance of selected J.HAL® antibodies across six developability assays. Conformational stability was assessed by differential scanning fluorimetry (DSF), low pH hold, and chemical unfolding. Colloidal stability was assessed by polyethylene glycol (PEG) solubility, SINS, and SMAC.

DSF assesses the temperature at which certain regions of the antibody begin to unfold. More stable, and thus more developable

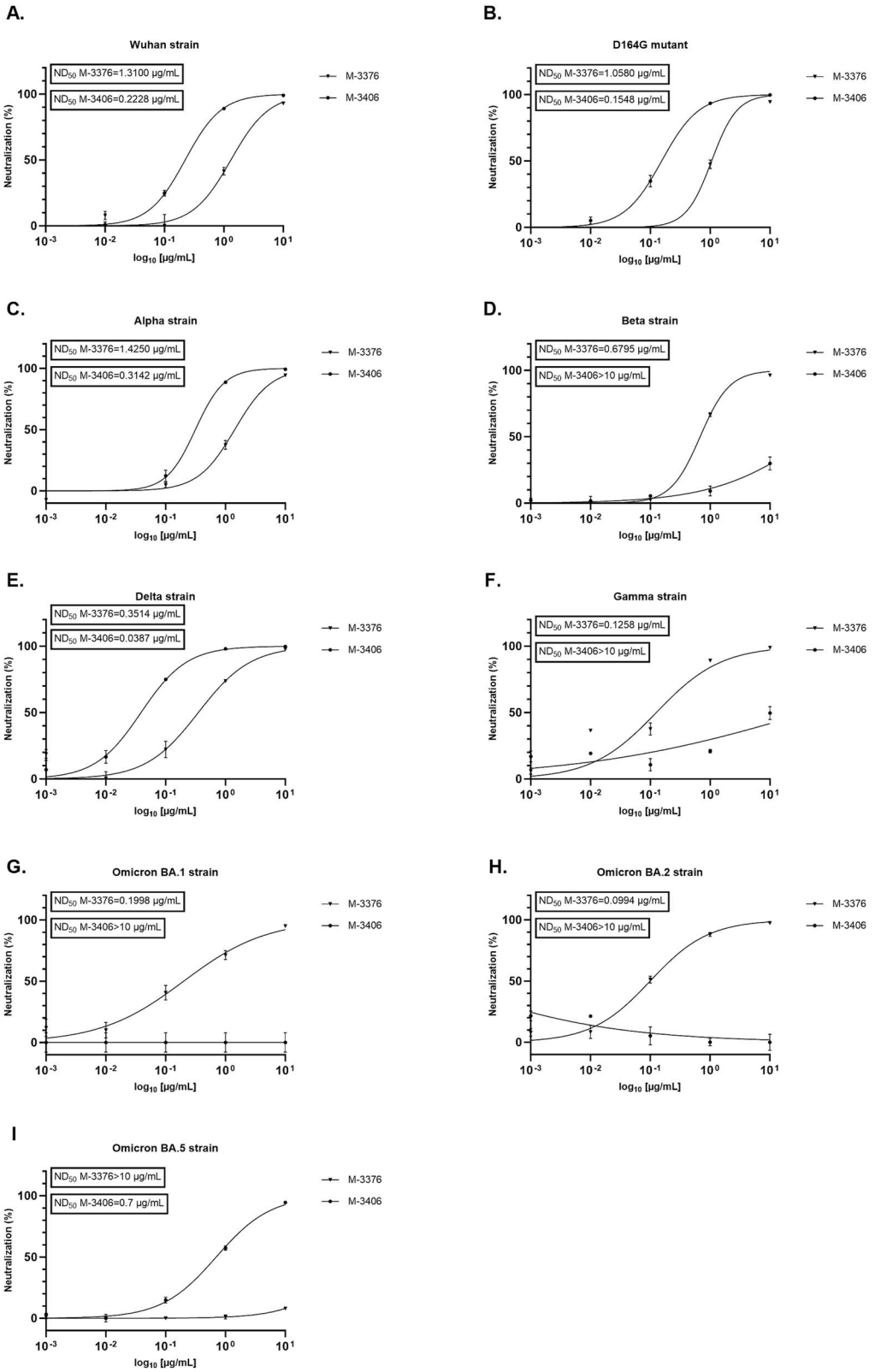


Figure 5. Evaluation of the cross-neutralization activity of J.HAL® IgG candidates M-3376 and M-3406 on a panel of VSV pseudotyped with SARS-CoV-2 spike in A549-ACE2/TMPRSS2 cell line. The neutralizing effects of IgG candidates M-3376 and M-3406 on entry of several rVSVΔG-SARS-CoV-2-Spike pseudotypes was evaluated in A549-ACE2/TMPRSS2 cells. VSV carrying the spike from Wuhan (A), Wuhan D614G mutant (B), Alpha (C), Beta (D), Delta (E), Gamma (F), Omicron BA.1 (G), Omicron BA.2 (H), and Omicron BA.5 (I) were used. For neutralization experiments, the pseudovirus was preincubated with a dose range of IgG candidates for 1 h at 37°C and the pseudovirus/antibody was subsequently added onto the cells and incubated for 20 h at 37°C. Infectivity was assessed at 20 h postinfection by Luciferase assay. The results were normalized to a percentage of neutralization where the cell control (not-infected condition) and the virus control (nontreated, rVSVΔG-SARS-CoV-2-Spike) are used to set 100% neutralization and 0% neutralization, respectively. Data were graphed as of % neutralization +/− SD and the 50% neutralization dose ($ND_{50}$), defined as the concentration of antibody that reduced the number of infected cells by 50% calculated using GraphPad Prism software.

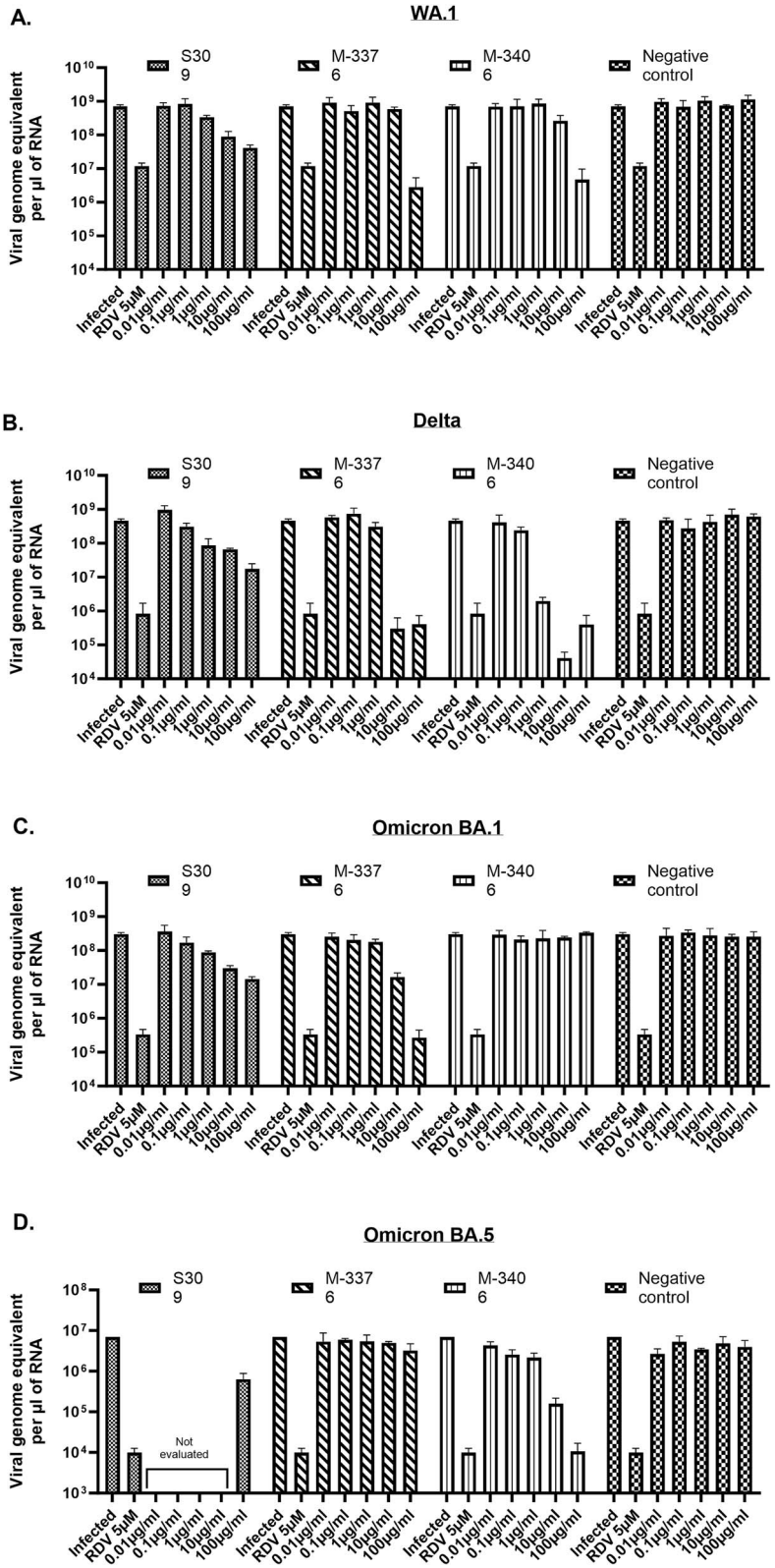


Figure 6. Confirmation of the cross-neutralization activity of IgG candidates M-3376 and M-3406 on a panel of SARS-CoV-2 isolates in A549-ACE2/TMPRSS2 cell line. The neutralizing effects of IgG candidates M-3376 and M-3406 on entry of SARS-CoV-2 isolates WA.1 (A), Delta (B), omicron BA.1 (C), and omicron BA.5 (D) was evaluated in A549-ACE2/TMPRSS2 cells. For neutralization experiments, the virus at MOI 0.1 was preincubated with a dose range of IgG candidates for 1 h at 37°C and the virus/antibody was subsequently added onto the cells. Positive (S309) and negative control antibodies were also subjected to the same protocol. Remdesivir, a nucleoside analogue, was added 1 h postinfection and used as a second positive control. Viral titers were quantified in cell lysates by RT-qPCR 40 h posttreatment using the standard curve method. Data were graphed as mean viral titers (expressed in genome equivalent per $\mu$l of RNA) +/− SD using GraphPad prism software.

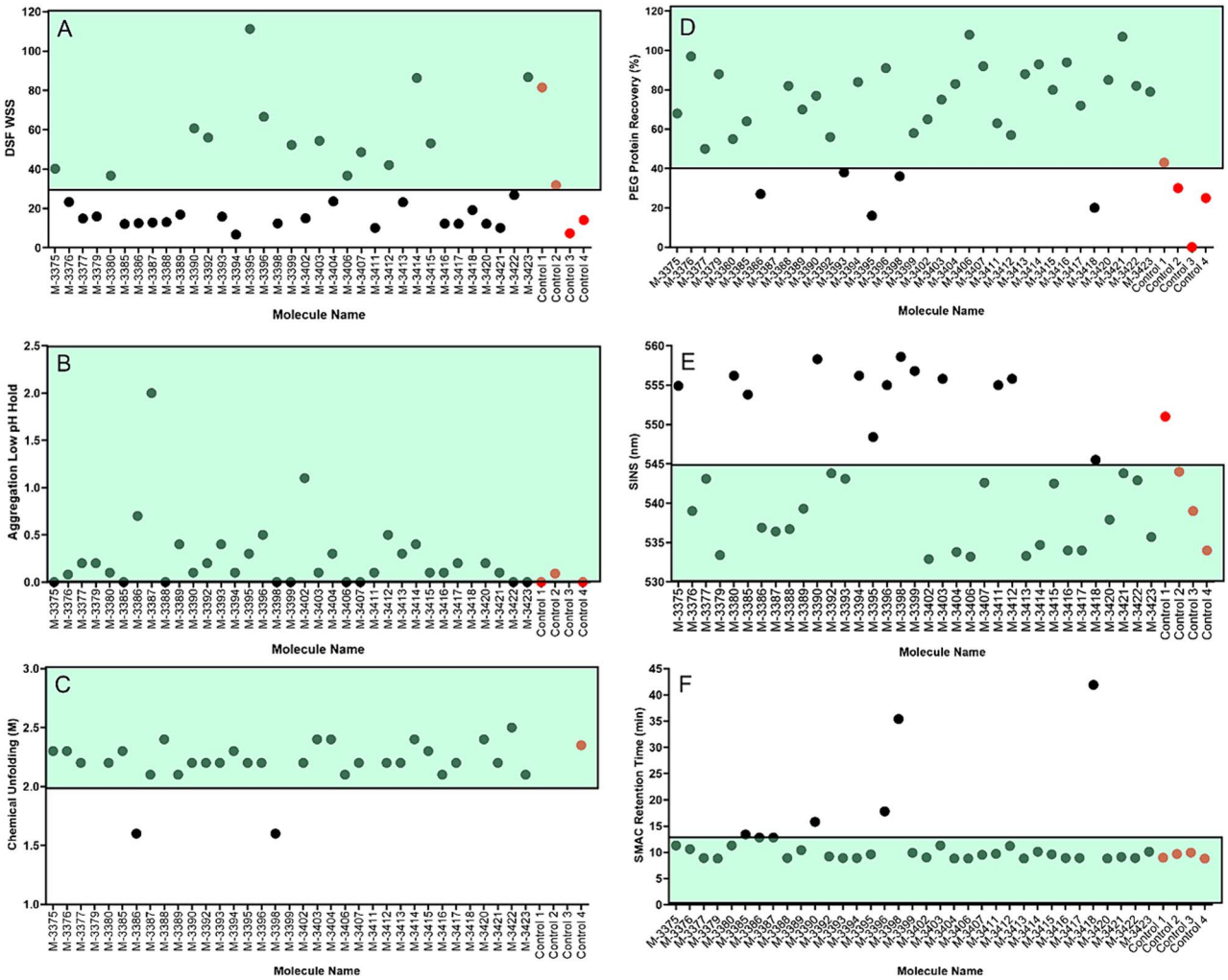


Figure 7. Antibody biophysical characterization indicating potential for manufacturability. For respective LC and HC amino acid sequences of these homodimeric antibodies, listed here by name and molecule ID number. Conformational stability results are shown in panels. (A) DSF WSS, (B) aggregation after low pH hold, and (C) chemical unfolding. Colloidal stability results are shown in panels, (D) relative PEG solubility, (E) SINS wavelength maximum, and (F) SMAC retention time. Acceptable value range for favorable developability characteristics are encompassed in the top areas in panels A, B, C and D, and in the bottom areas in panels E and F. M-3418 is the S309 control. S309 low pH-induced aggregation was observed at 9.6%. IgG1 controls 1–3 are (broadly neutralizing antibodies) bNAbs from convalescent patients and control 4 is a commercial product. Control 3 experiences 80% aggregation upon low pH hold, and therefore, is not shown on the above graph. Only control 4 was included in the chemical unfolding assay.

antibodies tend to have one or more regions unfold at higher temperatures and have a higher first-unfolding transition temperature. The DSF data in Fig. 7A are reported as WSS, which is a metric for thermal stability that can be derived from a DSF spectrum. The WSS algorithm consists of identifying the first temperate transition (T1) peak, normalizing the spectrum to the intensity of the T1 peak, and integrating the area under the curve between the T1 and the temperature where the intensity drops below the baseline. The integration is weighted by the squared temperature difference from T1. A higher WSS value indicates multiple transition temperatures and therefore improved stability. The WSS value increases as the second transition temperature increases. Many of the J.HAL® lead candidates, those in the green shaded region, showed high WSS indicating suitable thermal stability. Almost all mAbs with KV1–39 germlines were in this category (11 of 14). Similarly, 11 of 14 mAbs with HV3–30 HC germlines were also in this category of high thermal stability. All mAbs with HV1–18 (S309 control), HV1–46, and HV5–51 germlines were outside of this category. Only one lambda-containing mAb had high thermal stability as determined by DSF.

The inflection point of a chemical unfolding curve is thought to be related to conformational stability and to stability during long-term storage, with a greater inflection point relating to a more structurally or conformationally stable molecule. The only two mAbs that did not exhibit high chemical stability were M-3386 and M-3398, both of which have lambda variable regions (Fig. 7C). Exposure to low pH during viral inactivation could result in significant unfolding and aggregation in a conformationally unstable molecule. Aggregation that results from low pH instability such as during viral inactivation is cleared by subsequent purification steps but can negatively impact final yield. None of the J.HAL® candidates showed an increase in aggregate of more than 2% and were therefore all stable when exposed to pH 3.3 (Fig. 7B).

Protein–protein self-interaction has been correlated to low solubility, high viscosity, and increased manufacturability difficulties. Therefore, it is helpful to screen out mAbs early in development that self-interact. The SINS assay utilizes gold colloid surfaces to measure protein–protein interaction. Gold colloids have unique optical absorption properties dependent on their aggregation state, absorbing light at longer wavelengths as nanoparticle

aggregation occurs. Protein self-interaction is assessed by capturing test antibodies on the surface of a gold colloid and measuring the shift of the absorbance maxima. If the immobilized antibodies self-interact, the absorbance maximum of the spectrum red shifts to longer wavelengths (545–565 nm), or blue shifts to shorter wavelengths (530–545 nm), if less interaction occurs. The J.HAL® candidates displayed a wide variety of SINS results, with many of the candidates showing short wavelengths indicating no protein–protein self-interaction (Fig. 7E). The 11 mAbs above 550 nM may experience viscosity challenges above 100 mg/ml. Surface properties such as charge distribution leading to a strong dipole moment could drive this phenomenon. These Abs were not modelled to explore this possibility and the library was not designed to minimize this potential liability.

The inherent solubility of a molecule is thought to be related to aggregation levels throughout manufacturing process development and during long-term stability. Solubility was assessed by exposing the protein to PEG, which precipitates the protein out of solution. After incubation and filtration of insoluble precipitates, protein concentration is measured, and recovery calculated. Molecules with higher solubility will show more soluble protein recovered after exposure to PEG. Most of the J.HAL® candidates showed well over 50% protein recovery after exposure to PEG solutions, suggesting inherently high solubility (Fig. 7D). The degree of solubility is likely a property of surface hydrophobicity and charge distribution. The more that the former is not compensated by the latter, the greater the likelihood of poor solubility. Note that the control M-3418 is one of the poorest performers in this regard. This was also the case in SMAC method. The SMAC method is used to assess colloidal stability of a molecule by monitoring the retention of a molecule on a Zenix HPLC column. A commercially available Zenix® column (Millipore Sigma) with a hydrophobic standup monolayer can be used. We hypothesize that molecules with colloidal instability interact with the column resulting in a delayed retention time. Very few of the J.HAL® candidates interacted with the Zenix® column indicating that the lead candidates have favorable colloidal stability as measured by this method (Fig. 7F). Note that there are very few examples of a J.HAL® Ab performing poorly across more than two stability assays. M-3398 is one example, which consists of lambda. There are no obvious sequence drivers except that the 16-residue long H3 loop (slightly above average for this panel), consists of 5 Tyr, 2 Pro, 2 Gly, 1 Val, and 1 Phe with charged and polar residues sequestered potentially, suboptimally.

Altogether, these results demonstrate the potential of our AI-derived library and our high-throughput screening approach to identify panels of antibodies with good binding profile and activity without need of affinity maturation. Furthermore, these humanoid antibodies exhibit desirable developability features.

## Discussion

We describe here a novel, cost-effective and accelerated approach to therapeutic antibody discovery, that couples *de novo* human antibodies derived *in silico*, which we refer to as "humanoid antibodies", with high-throughput screening technologies. We have developed the antibody-GAN, a new *in silico* engineering approach to designing a novel class of diverse antibody therapeutics that mimic somatically hypermutated human repertoire response. The platform allows antibody biasing towards desirable developability attributes, such as suitability for common protein manufacturing processes, high stability during long-term storage conditions, low viscosity for injectability at high concentration and long elimination half-lives to reduce dosing frequency [3]. Here, we employed the antibody-GAN to design and construct a 1 billion theoretical diversity phage Fab display library. This library is in process of expanding to over 50 billion theoretical diversity by 2024. The first validation campaign described herein was conducted with the intent to discover antibodies to the SARS-CoV-2 spike protein and demonstrate library utility. Phage display methodology has proven to be a robust, and efficient platform to discover and develop human therapeutic antibodies [15]. Our antibody-GAN generated phage library is comprised of 25 sublibraries, five heavy chain germlines paired combinatorially with five variable chain (4 kappa and 1 lambda) germlines. Fv sequence diversity encompasses both framework and CDR regions, which drives affinity, efficacy, and developability [3]. Most synthetic libraries focus only on CDR diversity [16]. Due to the sensitive balance between affinity, efficacy, and developability, it is likely advantageous to consider the entire Fv sequence space when engineering the optimal therapeutic candidate. Since it is known that the framework residues can impact both developability and binding [19, 21]. We, therefore, chose to engineer this degree of diversity when designing J.HAL®. Both *in vitro* and *in vivo* antibody discovery methodologies have limitations whether it is library diversity or lack of an appropriate immune response. We do not know how J.HAL® will compare and perform against any particular methodology or antigen. However, as data are collected on the sequence-function/stability relationship, further library refinement (biasing via the transfer learning process) and diversity expansion will be conducted, improving discovery and developability success.

These first-generation antibodies, without the need for affinity maturation, bind to the SARS-CoV-2 spike protein with therapeutically relevant specificity and affinity and display efficient inhibition of spike:human ACE2 receptor binding. The IgG binding affinity covered two orders of magnitude with most of the candidates within the 10–100 nM $k_D$ range for all strains. The binding affinity exhibited for some candidates proved sufficient for effective functional activity, first demonstrated through the blockade of spike:ACE2 interaction, then confirmed by neutralization of SARS-CoV-2 viral infectivity across several strains. Affinity maturation is a common subsequent step after isolating initial antibody binders from *in vitro* antibody discovery systems and sometimes, *in vivo* systems, especially regarding cross-reactivity preference [17]. Being able to avoid the affinity maturation process while still addressing certain therapeutic applications with high developability potential is a significant benefit.

Multiple controls exhibiting both good and poor developability were included in all the biophysical stability assays. These controls have undergone extensive assessment through our process and product development workflow and associations have been made between the large-scale processes and the small-scale biophysical stability assays. There is high confidence in the relevance of the small-scale biophysical performance as it pertains to manufacturability. Antibodies from convalescing patients or animal immunizations often do not have suitable developability profiles [18–20]. Although such candidates might have superior efficacy, they often require developability optimization. The speed of the entire discovery through manufacturability process needs to be considered. The S309 control used in this study along with many other antibodies such as broadly neutralizing antibodies (bNAbs) from *in vivo* discovery processes are not optimized for manufacturability and require greater resources, including time, for development. Therapeutic development and manufacturing resources will be less taxed if affinity and developability

optimization is unnecessary while also providing sufficiently effective therapeutics to patients sooner.

The antibody panel identified encompasses diverse antibodies across multiple germlines with different selectivity and potency (Fig. S7). For example, a group of antibodies, such as M-3380 and M-3420 showed potent neutralization across Wuhan, D614G, Alpha, and Delta variants, but no effect on variants Beta, Gamma, and Omicron. The potency trended with strain affinity as measured by SPR for M-3380 where the $k_D$ were in the single-digit nM for Wuhan, D614G, Alpha, and Delta but for the Beta and Gamma strains (Omicron SPR not performed) were in the double-digit nM. For M-3420, a similar trend was present where the most potent activity was observed for Delta, for which it had the highest affinity. Whereas M-3420 exhibited the lowest potency was for Beta and Gamma both for which it had lowest affinity. Although it is possible these two antibodies bound to a similar epitope, it is probably unlikely since their sequences were so diverse. The heavy and light v-regions were HV3–30/KV3–15 and HV5–51/KV4–1 for M-3380 and M-3420, respectively. The CDRH3 pairwise sequence identity was 66% between the two with each having CDRH3 lengths greater than the average length of all isolated binders, 18 and 20 residues, respectively. The pairwise sequence identity of the CDRL3 loops were 45% but their lengths were shorter (both with nine residues) than the average isolated binder and doubtful were the driver for binding compared to the CDRH3. A second group of antibodies, M-3375, M-3403, and M-3412 showed a decreased neutralization potency against Wuhan and D614G variants compared to Alpha but highly potent neutralization against Beta and Gamma variants. These all had a similar $k_D$ profile with the notable exception of M-3412, which exhibited 30-fold higher affinity for Gamma. $k_D$ similarity was mostly unsurprising since the only sequence differences were in the frameworks. It is quite interesting that in one instance, the framework differences impacted the affinity approximately 30-fold. Additionally, the framework differences did confer stability variations such as M-3412 having the poorest chemical and pH stability of the three. However, the developability profile of all three mAbs was still better than the S309 control. The library was not designed with such a high identity in the CDRH3, and therefore these candidates with slight differences in the framework were likely due to variation caused by inherent molecular biology or process manipulation.

Of particular interest are two antibody candidates, M-3376 and M-3406, which displayed potent and complementary neutralization activity across all strains considered here. Both exhibited potent neutralization across Wuhan, D614G, Alpha, and Delta variants, although M-3406 was the most potent and the only one capable of inhibiting Omicron BA.5. Conversely, only M-3376 was able to neutralize Beta, Gamma, and Omicron BA.1 and BA.2 variants. This corresponds to the SPR affinity data. M-3406 exhibited single-digit nM $k_D$ across all strains but no binding to Beta or Gamma, whereas M-3376 exhibited low double-digit nM to all strains. The tested strains were chosen due to high seroprevalence during this study. Newer strains, such as XBB/BA.2.86, were not available at the time. Both antibody candidates consist of HV3–30, but M-3376 has a KV1–39 whereas M-3406 is KV4–1. The pairwise identity of the heavy v-regions was 78% whereas for the CDRH3 loops, it was 44%. Given the variation in strain binding, neutralization and sequence differences, the epitopes were likely different. The developability profile of M-3406 was quite good even though the chemical unfolding stability and DSF was nearer the bottom of the cohort. M-3376 exhibited good stability in all assays except for DSF where the weighted shoulder-score was slightly lower than our typical recommendation for developability.

Combined, these two antibody candidates represent an attractive therapeutic option, either co-administered in cocktail form or engineered as a bi-specific therapeutic modality, effective across multiple SARS-CoV-2 strains. Epitope binning would have been informative for further understanding the landscape of binding and associated activity of these antibodies. It is possible that aside from BLI or SPR epitope binning, an *in silico* method of SARS-CoV-2 spike protein structure superposition of each strain could have led to deductions of binding location. Due to resource limitations, we chose not to pursue these activities but instead, focused our efforts on the areas described.

There is the possibility that some developability optimization would be at odds with antibody efficacy. Improved algorithms via biasing will better address these challenges along with *de novo* antibody design for binding to specific epitopes. The former will allow for developability optimization while minimizing detriment to efficacy. The algorithm will be able to discern residues that are distal from the paratope or otherwise unlikely to drastically disrupt the paratope but potentially still have value in improving developability. However, if the liability is within the paratope then this will be a more challenging problem. In this case, surveying the target surface for a more suitable epitope while still maintaining a good developability profile is likely the better strategy. This latter methodology is under development and will greatly benefit speed and efficient utilization of resources. The new discovery process will constitute a different set of challenges but overall, it will gain the advantage since candidates will enter stable cell line development sooner.

This validation work demonstrates the effectiveness of the humanoid antibody library, J.HAL®, and the associated discovery process. Multiple broadly neutralizing antibodies against multiple SARS-CoV-S strains with desirable developability attributes were isolated. The viral neutralization assays described exhibited good correlation with the rVSVΔG-SARS-CoV-2-Spike pseudotypes and SARS-CoV-2 isolates. These observations highlight the robustness of the VSV pseudotyped model for the rapid identification of effective neutralizing antibody candidates. These diverse and effective antibodies are anticipated to be a great fit-to-platform for product and process development, resulting in efficient resource allocation and shortened timelines, which is critical for rapid response and a benefit to patients.

## Acknowledgements

We gratefully thank Alex Taylor for bioinformatical support, and Kathryn McLean and Lindsay Pautsch for the generation and purification of material.

## Author contributions

Cristina Loomis (Conceptualization [equal], Data curation [equal], Formal Analysis [equal], Investigation [equal], Methodology [equal], Supervision [equal], Writing—original draft [equal], Writing—review & editing [equal]), Thomas Lahlali (Conceptualization [equal], Data curation [equal], Formal Analysis [equal], Investigation [equal], Methodology [equal], Supervision [equal], Writing—original draft [supporting], Writing—review & editing [supporting]), Danielle Van Citters (Conceptualization [equal], Data curation [equal], Formal Analysis [equal], Investigation [equal], Methodology [equal], Supervision [equal], Writing—original draft [equal], Writing—review & editing [supporting]), Megan Sprague (Data curation [equal], Formal Analysis [supporting], Investigation [supporting], Methodology [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Gregory

Neveu (Conceptualization [equal], Data curation [equal], Formal Analysis [equal], Investigation [equal], Methodology [equal], Supervision [equal], Writing—review & editing [supporting]), Laurence Somody (Investigation [supporting], Methodology [supporting]), Christine Siska (Data curation [equal], Formal Analysis [equal], Investigation [equal], Methodology [equal], Writing—review & editing [supporting]), Derrick Deming (Formal Analysis [supporting], Writing—review & editing [supporting]), Andrew Asakawa (Investigation [supporting], Methodology [supporting], Writing—review & editing [supporting]), Tileli Amimeur (Methodology [equal]), Jeremy M. Shaver (Methodology [equal]), Caroline Carbonelle (Conceptualization [equal], Supervision [equal], Writing—review & editing [supporting]), Randal Ketchem (Conceptualization [equal], Supervision [equal], Writing—review & editing [supporting]), Antoine Alam (Conceptualization [equal], Resources [equal], Supervision [equal], Writing—review & editing [supporting]), and Rutilio Clark (Conceptualization [equal], Formal Analysis [equal], Resources [equal], Supervision [equal], Writing—original draft [equal], Writing—review & editing [equal]).

## Supplementary data

Supplementary data is available at *ABT* Online.

## Conflict of interest

All authors were employees of Just-Evotec Biologics/Evotec at the time all work was performed. All work was performed by employees of Just-Evotec Biologics/Evotec.

## Funding

This work was funded by Just-Evotec Biologics/Evotec Inc.

## Data availability

The data underlying this article are available in the article and in its online supplementary material.

## Ethics and consent statement

Consent was not required.

## Animal research statement

Not applicable.

## References

1. Lu R-M, Hwang Y-C, Liu I-J. *et al.* Development of therapeutic antibodies for the treatment of diseases *J Biomed Sci*. 2020;**27**:1. https://doi.org/10.1186/s12929-019-0592-z.
2. Naranayan H, Dingfelder F, Butté A. *et al.* Machine learning for biologics: opportunities for protein engineering *Developability and Formulation Trends Pharmacol Sci*. 2021;**42**:151–65. https://doi.org/10.1016/j.tips.2020.12.004.
3. Amimeur T, Shaver JM, Ketchem RR. *et al.* Designing feature-controlled humanoid antibody discovery libraries using generative adversarial networks. BioRxiv 2020. https://doi.org/10.1101/2020.04.12.024844.
4. Gulrajani I, Ahmed F, Arjovsky M. *et al.* Improved training of Wasserstein GANs. *arXIV* 2017; arXiv:1704.00028.
5. Kerwin BA, Bennett C, Brodsky Y. *et al.* Framework mutations of the 10-1074 bnAb increase conformational stability, manufacturability, and stability while preserving full neutralization activity *J Pharm Sci*. 2020;**109**:233–46. https://doi.org/10.1016/j.xphs.2019.07.009.
6. Gillespie A, Shaver JM. High throughput interrogation of physiochemical properties of a protein. United States patents US11408840B2; US2020003714A1 2020.
7. Floyd JA, Siska C, Clark RH. *et al.* Adapting the chemical unfolding assay for high-throughput protein screening using experimental and spectroscopic corrections *Anal Biochem*. 2018;**563**:1–8. https://doi.org/10.1016/j.ab.2018.08.027.
8. Toprani V, Joshi S, Kueltzo L. *et al.* A mico-polyethylene glycol precipitation assay as a relative solubility screening tool for monoclonal antibody design and formulation development *J Pharm Sci*. 2016;**105**:2319–27. https://doi.org/10.1016/j.xphs.2016.05.021.
9. Liu Y, Caffry I, Wu J. *et al.* High-throughput screening for developability during early-stage antibody discovery using self-interaction nanoparticle spectroscopy *MAbs*. 2014;**6**:483–92. https://doi.org/10.4161/mabs.27431.
10. Kohli N, Jain N, Geddie ML. *et al.* A novel screening method to assess developability of antibody-like molecules *MAbs*. 2015;**7**:752–8. https://doi.org/10.1080/19420862.2015.1048410.
11. Pinto D, Park Y-J, Beltramello M. *et al.* Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody *Nature*. 2020;**583**:290–5. https://doi.org/10.1038/s41586-020-2349-y.
12. Winkler ES, Gilchuk P, Yu J. *et al.* Human neutralizing antibodies against SARS-CoV-2 require intact fc effector functions for optimal therapeutic protection *Cell*. 2021;**184**:1804–1820.e16. https://doi.org/10.1016/j.cell.2021.02.026.
13. Zost SJ, Gilchuk P, Case JB. *et al.* Potently neutralizing and protective human antibodies against SARS-CoV-2 *Nature*. 2020;**584**:443–9. https://doi.org/10.1038/s41586-020-2548-6.
14. Pruijssers AJ, George AS, Schäfer A. *et al.* Remdesivir inhibits SARS-CoV-2 in human lung cells and chimeric SARS-CoV expressing the SARS-CoV-2 RNA polymerase in mice *Cell Rep*. 2020;**32**:107940. https://doi.org/10.1016/j.celrep.2020.107940.
15. Almagro JC, Pedraza-Escalona M, Arrieta HI. *et al.* Phage display libraries for antibody therapeutic discovery and development *Antibodies (Basel)*. 2019;**8**:44. https://doi.org/10.3390/antib8030044.
16. Shim H. Synthetic approach to the generation of antibody diversity *BMB Rep*. 2015;**48**:489–94. https://doi.org/10.5483/BMBRep.2015.48.9.120.
17. Doria-Rose NA, Joyce MG. Strategies to guide the antibody affinity maturation process *Curr Opin Virol*. 2015;**11**:137–47. https://doi.org/10.1016/j.coviro.2015.04.002.
18. Patel A, Gupta V, Hickey J. *et al.* Coformulation of broadly neutralizing antibodies 3BNC117 and PGT121: analytical challenges during preformulation characterization and storage stability studies *J Pharm Sci*. 2018;**107**:3032–46. https://doi.org/10.1016/j.xphs.2018.08.012.
19. Bailly M, Mieczkowski C, Juan V. *et al.* Predicting antibody Developability profiles through early stage discovery screening *MAbs*. 2020;**12**:1743053. https://doi.org/10.1080/19420862.2020.1743053.
20. Clark RH, Latypov RF, De Imus C. *et al.* Remediating agitation-induced antibody aggregation by eradicating exposed hydrophobic motifs *MAbs*. 2014;**6**:1540–50. https://doi.org/10.4161/mabs.36252.