Howard Hughes Medical Institute

Computational design of trimeric influenza neutralizing proteins targeting the hemagglutinin receptor binding site

EM Strauch[1], **SM Bernard**[2], **D La**[1], **AJ Bohn**[3], **PS Lee**[2], **CE Anderson**[4], **T Nieusma**[2], **CA Holstein**[4], **NK Garcia**[5], **KA Hooper**[6], **R Ravichandran**[1], **JW Nelson**[1], **W Sheffler**[1], **JD Bloom**[6], **KK Lee**[5], **AB Ward**[2], **P Yager**[4], **DH Fuller**[3,7], **IA Wilson**[2], and **D Baker**[*,1,8,9]

[1]Department of Biochemistry, University of Washington, Seattle, Washington, USA

[2]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California, USA

[3]Department of Microbiology, University of Washington, Seattle, Washington, USA

[4]Department of Bioengineering, University of Washington, Seattle, Washington, USA

[5]Department of Medicinal Chemistry, University of Washington, Seattle, Washington, USA

[6]Division of Basic Sciences and Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA

[7]Washington National Primate Research Center, University of Washington, Seattle, Washington, USA

[8]Institute for Protein Design, Department of Biochemistry, University of Washington, Seattle, Washington, USA

[9]Howard Hughes Medical Institute, University of Washington, Seattle, Washington, USA

Abstract

Many viral surface glycoproteins and cell surface receptors are homo-oligomers[1-4], and hence can potentially be targeted by geometrically matched homo-oligomers that engage all subunits

*corresponding author (dabaker@uw.edu).

simultaneously to attain high avidity and/or lock subunits together. The adaptive immune system cannot generally employ this strategy since the individual antibody binding sites are not arranged with appropriate geometry to simultaneously engage multiple sites in a single target homo-oligomer. We describe a general strategy for the computational design of homo-oligomeric protein assemblies with binding functionality precisely matched to homo-oligomeric target sites[5-8]. In the first step, a small protein is designed that binds a single site on the target. In the second step, the designed protein is assembled into a homo-oligomer such that the designed binding sites are aligned with the target sites. We used this approach to design high-avidity trimeric proteins that bind influenza A hemagglutinin (HA) at its conserved receptor binding site. The best characterized of these designed trimers can both capture and detect HA in a paper-based diagnostic format, neutralizes influenza in cell culture, and completely protects mice when given as a single dose 24 h before or after challenge with influenza.

Circulating influenza viruses undergo constant antigenic drift [9,10, 11], which renders acquired immunity against previous seasonal strains ineffective. The occasional adaptation of avian or swine viruses to human hosts[12] poses a serious threat to global health. Resistance has evolved to widely used anti-viral drugs[13, 14], and seasonal vaccines are unlikely to provide much if any protection against newly emerging viruses. Thus, general or quickly adaptable solutions are needed for both diagnostic and therapeutic applications. The surface glycoprotein HA enables viral entry into host cells by binding to cell-surface, sialic-acid-containing glycans (Fig. 1a) and mediating fusion between the viral and host membranes in endosomal compartments. As the regions on HA involved in binding and fusion are highly conserved, they are attractive sites for binder design. The affinity of an individual receptor-binding site (RBS) for a sialic-acid moiety is low ($K_d \sim 1$ mM)[15], and small-molecule analogs and oligomers of sialic acid have similarly low affinities, preventing their use in clinical settings to date[16]. Trimerization of terminal sialic-acid moieties using flexible alkyl chains resulted in only weak affinity, likely due to the entropic cost of fixing the long flexible alkyl chains[17]. Broadly neutralizing antibodies targeting the stem and head regions of HA have been identified[5-7, 18]; the stem region antibodies have broader binding specificity as they bind to a larger highly conserved patch, but the head region antibodies can neutralize more potently as the head region is more accessible[19, 20]. Neutralizing antibodies use two related strategies to bind the head region: direct mimicry of the receptor using the carboxylic acid of an aspartic acid and backbone atoms to mimic the carboxylate and acetamido groups of sialic acid respectively, or insertion of a large hydrophobic amino acid into the RBS pocket[5] (Fig.1b), with a backbone carbonyl group making interactions similar to those of the sialic-acid carboxylate[6, 21,7] (Supplementary Fig.1). The heavy-chain CDR3 (HCDR3) of the C05 broadly neutralizing antibody, which inserts into the RBS, contains internal, self-stabilizing hydrogen bond interactions, similar to a β-hairpin (Fig.1a), and provides extensive hydrogen bonding interactions with backbone atoms of the 130 loop of HA[5] (Fig.1b).

In a first step toward the design of a high-avidity trimeric flu inhibitor, we developed a computational approach to design small protein mimics of the C05 binding mode. We searched for close backbone alignments of the first and last residues in 7-15 residue segments of the C05 HCDR3 with pairs of positions in a set of 1718 native proteins

(Supplementary Table 1). At sites with close matches, the intervening segment in the scaffold protein was replaced with the antibody loop segment, the bonded geometry at the junctions between the introduced loop and the scaffold was idealized[22], and the interactions within the designed protein and with HA (positioned as in the C05 co-crystal structure, Fig. 1d) were optimized in sequence design calculations favoring key interface contacts and overall shape complementary. Designs with low computed monomer and HA interaction energies were selected for experimental characterization (see Methods). This approach goes beyond previous efforts to graft antibody loops in utilizing near perfect geometric alignment with the variable parts of the CDR, explicitly modeling and optimizing the junctions between the scaffold and CDR, and optimizing the scaffold to support the CDR and make additional interactions with the target [22, 23].

Genes encoding the 80 selected designs were screened as a pool against H3 A/Hong Kong/ 1/1968 HA (H3 HK68) using yeast surface display[24], while in parallel monitoring off-target binding (Fig. 1c, Strauch et al. *in preparation*). Five designs specifically bound H3 HK68 HA; all contained grafts of short 7-8 residue fragments from the tip of the C05 loop - longer loop fragments likely require more extensive supporting contacts to prevent alternative loop configurations. Screening with a panel of HAs from different flu strains showed that these initial designs only bound H3 HA, which is not surprising since there is considerable sequence and structure divergence in the periphery of the sialic-acid binding site in other subtypes (Fig. 2a). There are however no documented changes in vicinity of the binding site that would affect binding (Supplementary Fig. 2). Screening of simple random mutagenesis libraries resulted in only very limited extension of the binding spectrum (Supplementary Fig. 3).

To effectively target the diverse binding pockets presented by different HA molecules, we sought to match sequence variation in HA with complementary diversity in the designed binders starting from the design with the smallest buried surface area, which we refer to as HSB (Fig. 2a,b). Selections were carried out either alternating between strains representing Group 1 and Group 2 subtypes, or against individual strains (Supplementary Fig. 4). As described in Fig 2c-f, the varying specificities of the resulting clones can be understood based on interactions with features that vary in different HA molecules (Fig. 2c-f). As HSB. 1C had the broadest binding spectrum, we proceeded with this variant and further optimized its stability (Methods, Supplementary Fig. 6), resulting in variant HSB.2.

Other than a domain swap observed previously for cystatin folds, the crystal structure of HSB.2 agrees closely with the design model (0.44 Å RMSD, Supplementary Fig. 6a, Supplementary Table 2). To avoid the domain swap, we rigidified loop 2 (HSB.2A; Suppl. Methods), and succeeded in solving the structure of HSB.2A in complex with H3 HK68 HA at 2.6 Å resolution (Fig. 3a-c, Supplementary Table 2). The structure confirms the designed mode of binding, with three HSB.2A molecules bound to the HA (Fig.3a) and with all designed hydrogen bonds between the grafted loop and HA present (Fig. 3b,c). In the bound state, the contacting hairpin containing the grafted loop moves slightly away from the protein core (Supplementary Fig. 7b).

With a well-characterized small binding domain in hand, we proceeded to the second step of our avid-binder design strategy and designed trimeric versions of HSB such that the three units of HSB were aligned with the three sialic-acid binding sites on trimeric HA. We sampled translations and rotations of homo-trimers from the PDB (Supplementary Table 3) along the symmetry axis of the HA trimers (Fig. 3d) searching for proximity between the termini of naturally occurring trimers and the termini of HSB when bound to the HA. Charges and potentially clashing bulky residues at the newly formed interface with HA were substituted with neutral and smaller amino acids, respectively, and the gap between the two domains (trimerization unit and binder) was closed with a Gly-Ser linker. Nine designs were experimentally tested; we chose the one with the most monodisperse size exclusion chromatography (SEC) trace (Tri-HSB.1C), which also expressed at high levels (more than 100 mg/L culture). Trimerization of HSB.1C and HSB.2 reduced the off-rate for HA by more than 1,000 fold (Fig. 3f, Supplementary Fig.8), resulting in sub- and single-digit nanomolar binding to a variety of H1, H2, H3, H9 and H12 strains (Supplementary Table 4; the ~3-fold decrease in the on-rate is likely due the fact each unit has to bind independently). Tri-HSB.2 binds H3 A/Victoria/361/2011 (Vic11), a close relative of current circulating strains, with low nanomolar affinity (the monomeric design does not give a binding signal; Fig. 2c). The trimer has a similar overall breadth in binding as the C05 IgG, with higher affinity than the IgG for several H2 and H12 strains and weaker affinity for several H3 strains (Supplementary Table 4). Because of its binding geometry, the IgG C05 can only engage one of three sialic-acid binding sites on a single HA molecule, but can achieve avidity by binding sites in two different HA trimers[5].

Asymmetric reconstruction of negative-stain, single-particle EM images of the Tri-HSB.1C – HA complex yielded clear density for the three trimerization domains and three binding domains (one of the connecting loops is not visible); the computational model of the trimer fits the density well (Fig. 3d, Supplementary Fig. 9). SAXS and HDX-MS data are also consistent with the design model (Supplementary Figs. 10 and 11); HA loop 130, which extensively hydrogen bonds with the edge β-strand of Tri-HSB.2, is protected from deuterium exchange in the bound state, as are residues near the RBS (Supplementary Fig. 11).

We tested the utility of Tri-HSB for detecting HA in a paper-based assay similar to what might be used for diagnosis of infectious diseases at point-of-care sites or in home-testing kits[25]. As shown in Figure 4a, Tri-HSB is effective both for capturing and detecting HA in a lateral flow assay, with a very clear band on the nitrocellulose strip and little or no background using either gold nanoparticles or diaminobenzidine (DAB) precipitation by conjugated HRP.

We next investigated whether Tri-HSB.2 could neutralize influenza virus *in vitro* and protect animals against infection. *In vitro*, Tri-HSB.2 potently neutralized human H3 HK68 (Fig. 4b) at sub-nanomolar concentrations with an IC_{50} below 150 pM (Fig. 4b), in the range of the apparent K_d. The weaker binder Tri-HSB.1C with an apparent K_d of 12 nM has an IC_{50} of 19 nM. To determine if Tri-HSB.2 can protect against influenza *in vivo*, as a therapeutic or prophylactic, we administered a single intranasal dose of Tri-HSB.2 (3 mg/kg) to BALB/c mice either 24 h prior to or post challenge with a high dose (1000 fold $TCID_{50}$) of H3 HK68

virus. Both treatments fully protected mice from disease (Fig. 4c), while untreated control mice and mice treated with a non-binding double point mutant, Tri-HSB.2.ko, exhibited significant disease and weight loss (>25%).

The adaptive immune system is capable of generating very high-affinity interactions with target proteins, but the fixed geometry of antibodies does not allow the positioning of the binding sites to maximally engage homomeric antigens. Our approach goes beyond elegant recent work using DNA linkers to optimize the spacing between Fab fragments for HIV binding[26] by (1) optimizing both the distance and the relative orientation of the three displayed binding moieties using structure-based design, (2) bringing together three binding domains rather than two, and (3) achieving the desired orientation with protein scaffolding rather than DNA, which has advantages for downstream applications. Our two-step protocol for optimally positioning mimics of antibody combining sites for maximally avid engagement of target homo-oligomers has considerable potential for both diagnostic and therapeutic applications. The power and applicability of our approach should continue to increase as additional crystal structures of antibodies in complex with both pathogen and homo-oligomeric surface receptors are elucidated.

Methods

Computational Methods

Design of the RBS Binding Proteins—For a scaffold database, we curated 1718 high-resolution, monomeric crystal structures from the PDB and relaxed them into the Rosetta energy function. Using Epigraft[22], segments of varying lengths extracted from the heavy chain CDR3 of the C05 antibody (PDB ID 4FP8) ranging from Ser97 to Leu100K (or Ala100I) were aligned with each scaffold. If the rms distance between the joining backbone endpoints was less than 0.25 Å and diverged less than 1.0 Å when aligning the replaced scaffold fragment with the CDR fragment motif, grafting was attempted. All predicted interface residues of suitable scaffolds were substituted with alanine and, using small fragment insertion and backbone minimization, we attempted to close broken bonds. For the top 10 closed solutions, the remainder was briefly redesigned by Epigraft[22]. Both intra- and intermolecular interfaces were then extensively redesigned using 3 rounds of the most current RosettaDesign[28] while also allowing rigid-body minimization, but maintaining key contacts (100A, 100C-F) of the grafted piece. The best scoring design for each successful graft solution was kept if hydrogen bonds to the HA 130 loop and Tyr98 were maintained and had a computed binding energy below -12 units. All residues that were changed and did not contribute to the computed ΔG of binding were reverted to their native identities. Source code is freely available to academic users through the Rosetta Commons agreement (http://www.rosettacommons.org/).

Computation-based Design of Disulfide Bridges—To identify possible locations for the incorporation of two cysteine residues to form a disulfide bridge in HSB variants without disturbing the backbone, we utilized RosettaRemodel[29]. Briefly, Cβ distances between all residue pairs were calculated and disulfide bridges were introduced and the model was relaxed and minimized when the appropriate distances were identified.

Loop 2 Re-design—To prevent domain swapping and enable crystallization with HA, loop 2 of HSB was redesigned. Ideal phi-psi angles and loop truncations were sampled using blueprint files[30] and sequence design was biased towards amino acid identities commonly seen within beta-sheet turns. Out of 10 tested designs, a 3-residue truncated version, HSB. 2A, expressed well and ran as a mono-disperse peak on SEC.

Homology Modeling—The query sequence was threaded onto the HA structure with the closest sequence similarity as identified by BLAST (ncbi.nlm.nih.gov). The model was then subjected to several rounds of relaxation while constraining the backbone atoms to their original coordinates. We used following PDBs as templates for homology models: chain A of 3sm5 for H1 SI and H1 NewCal, and chain A of 4hfu for H2 Ada.

HA Sequence Alignments and Shannon Entropy—Non-redundant human influenza A H3 sequences for HA were downloaded from the Influenza Research Database (fludb.org) and only sequences from complete genomes were extracted and aligned using Muscle[31]. Shannon Entropy was calculated using method calls from the Python-based ProDy package[32].

Design of Trimeric Binding Units—Translations and rotations of 300 homo-trimeric structures were sampled along the symmetry axis of the HA trimers and scored based on their proximity to either termini while avoiding clashes. Out of the successful solutions, we focused on small trimers of thermophilic organisms with compatible geometry for loop closure between the identified trimer and HSB. Residues of the trimer that could interfere with binding to HA were "shaved off", e.g. charges were designed to become neutral residues, and potentially clashing bulky residues were changed to smaller amino acids in FoldIt[33]. The original trimeric scaffolds were selected to have a minimal distance towards either terminus of HSB, which allows loop closure with a simple glycine-serine linker. Loops were closed via RosettaRemodel[29]. Blueprint files were generated so that 2, 3, 5 or 8 residue insertion between the trimerizing unit and HSB were modeled. Trimierization is based on the 1nza PDB coordinates. For initial experimental characterization, we fused the trimer adapter to the HSB.1C variant. Different loop lengths for most successful design were cloned, and expressed.

Library Constructions

Design library—DNA encoding the designed proteins (Gen9 Inc.) flanked by homologues region of the pETCON plasmid was co-transformed with linearized plasmid (NheI/XhoI/ BglII digested) into EBY100 cells[24] and expressed as Aga2 fusions.

PCR-Mutagenesis—The genes for HSB were subjected to an error-prone PCR amplification using the Metamorph Mutagenesis kit (Agilent). Briefly, using primers upGS, downMyc and 50 ng of total plasmid DNA as template, 31 cycles of 94°C, 54°C for annealing and 72°C for primer extension were performed. Genes were purified through Gel extraction (QIAgen). As the first mutagenesis PCR resulted in less than 20% of the genes with at least 1 mutation, we repeated the amplification using the mutagenized DNA as a

template, resulting in an error rate of on average 1-2 mutation per gene as determined by 10 randomly chosen clones.

Generation of the Combinatorial Libraries—The two combinatorial gene libraries (one to address the HA diversity and the second to introduce stabilizing mutations identified via selections at increasing temperatures) were generated through recursive assembly of 4 ultramers (IDT) containing degenerate codons. Two fragment pieces annealing (HSBlib.gII.frag1 or HSB.1Cdis.frag1 and HSBlib.gII.frag2a or HSB.1Cdis.frag2) were amplified through outer primers (HSBlib.gII.For or HSB.1CdisYFor and HSBlib.gII.Rev or HSB.1CcysYRev). To limit the library size for the HA diversity-addressing libraries which increased due to additional diversity gained by the use of degenerate codons, two gene libraries were designed for fragment 1 (a and b). After assembly, the amplification reactions were combined. DNA was gel purified before transformation.

Generation of Site-Saturated Mutagenesis (SSM) library—All possible point mutations for each position were generated through overlap PCR. To simplify the preparation step and to generate standardized primers, we automated the primer design step with a custom Python script. We generated one shorter reverse primer and a longer forward primer that contained the mismatch encoded by NNK. Two fragments for each position were generated, here 2×87 positions. Outer PCR used 25 ng of the plasmid DNA encoding HSB. 1C with a disulfide bridge between position 8 and 40 as a template with primers colonyF and colonyR; The overlap fusing reaction for each position was performed by combining 1 μl of the two fragments followed by 30 cycles of amplification using the primer colonyF and colonyR primers and the same PCR conditions as the first reaction. All 87 overlap PCRs were checked for their success through agarose gel electrophoresis. Reactions (5 μl of each) were combined and gel extracted as a pool.

Yeast transformation

For library transformation, EBY100 [24] cells were transformed with 2 μg of the gene library together with 1 μg linearized pETCON (SalI, XhoI, NheI digested)[27] using previously reported protocol[34]. Transformation resulted in 2-5 $\times$ 10e7 cells.

Library Generation and Yeast Surface Selections and Titrations

Yeast cells were induced for about 16 – 18 h at 22°C, then washed once with PBS containing 0.1% BSA (PBSF). For the very first round of selections, unless noted, cells were typically incubated for 4 h at 4°C with the indicated HA, before adding a fourth of the used concentration streptavidin conjugated to phycoerythrin (SAPE, Invitrogen) and 2 ng/ml anti-c-Myc antibody conjugated to FITC and incubating for another hour. For non-avid conditions in later sorts, cells were incubated for 2 h at 4°C while rotating, unless noted otherwise, and then washed once with ice-cold PBSF before re-suspending in 100 mL with 72 nM SAPE and 2 ng/ml anti-c-Myc antibody conjugated to FITC for 30 min on ice. Cells were washed once with 1 mL ice-cold PBSF and stored as pellets on ice before their final re-suspension in 1 mL ice-cold PBSF followed by fluorescence-activated cell sorting (FACS). The combinatorial library was screened by sorting under less stringent, avid condition at

1000 nM with a 4:1 ratio of SAPE, followed by non-avid, more stringent selection at 500 nM, 125 nM and 50 nM H1 SI HA.

Library Preparation and Next-Generation Sequencing

Plasmids were extracted from 5×10^7 yeast cells for each selection and starting pool as previously described[35]. After ExoI (NEB) and Lambda exonuclease (NEB) treatment, the Illumina sequencing primers sequence, flow cell adapters and selection specific-barcodes were added *via* 2 nested PCRs using NEBnext. They additionally add 12 entirely degenerate bases at the beginning of the forward and reverse read, ensuring adequate diversity for the Illumina base-calling algorithms. A 250-PE kit was sequenced with 15% phiX at 10 pM concentration using a MiSeq (Illumina). After quality control filtering[35], we calculated enrichment for each sort by dividing the frequencies of selected by ones seen in the starting pool.

Protein Expression and Purification of Design Variants

For expression of the design variants in *E coli*, genes were cloned into pET29b between the restriction sites NdeI and XhoI and transformed into BL21Star (Invitrogen). For expression, 10 ml of overnight culture grown in Terrific Broth (BD Difco) was used to start a 500 ml culture using Studier's autoinduction medium but replacing the N-Z-amine fraction with TB. Cells were grown for 8 h at 37°C, before reducing the temperature to 18°C for another 14-16 h. Cells were re-suspended in 35 ml phosphate buffered saline (PBS, 150 mM NaCl and 25 mM phosphate buffer at pH 7.4) and lysed using a M110P Microfluidizer (Microfluidics). Insoluble cell debris was removed by centrifugation for 20 min at $40,000 \times g$. Supernatant was applied to gravity-flow columns containing 2 mL of Ni-NTA for each 500 ml of culture, washed with 50 ml PBS and 50 ml PBS containing 30 mM imidazole. Proteins were eluted with 20 ml of 250 mM imidazole in PBS. If necessary, proteins concentrated to 1.5 mg/ml using a Vivaspin 10kD MWCO centrifugal concentrator (Sartorius Stedim) at $4000 \times g$. For animal studies, proteins were purified using column and an imidazole gradient. The 400 mM elution fraction was subjected to further cleaning *via* SEC and further processed using bacterial endotoxin removal beads (Miltenyi Biotec).

Protein Expression of Designs

Design variants were cloned into pET29 and expressed in BL21* (Invitrogen), followed by purification using Ni-NTA (QIAgen). Recombinant HA was expressed and biotinylated as previously described[5]. Detailed descriptions can be found in the Supplementary Info under the Methods section.

Expression, Purification and Biotinylation of HA

Recombinant hemagglutinin was expressed, purified and biotinylated as previously described[18]. In brief, HAs were expressed as a fusion with a C-terminal Avi-tag, Foldon trimerization domain and hexahistidine tag in Hi5 cells. After 3 days of expression at 28°C, cells were removed by centrifugation and the culture supernatant was concentrated and buffer exchanged into 50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole prior to overnight incubation with Ni-NTA resin at 4°C. The resin was washed with 50 mM Tris pH

8.0, 300 mM NaCl, 10 mM imidazole and bound protein was eluted in the same buffer containing 250 mM imidazole. Proteins used in crystallographic or EM studies were digested using trypsin at a 1:1000 mass ratio overnight at 4°C and further purified by size exclusion chromatography on a Superdex S200 16/90 column (GE) in 25 mM Tris pH 8.0, 150 mM NaCl. Hemagglutinin used in library selections and bio-layer interferometry were site specifically biotinylated by incubation with recombinantly expressed BirA in 100 mM Tris pH 8.0, 10 mM ATP, 10 mM magnesium acetate, 50 μM biotin and purified by size exclusion chromatography as above.

Optimization of HSB.1C

After the introduction of a disulfide bridge to HSB.1C (see Methods), we generated a site-saturation mutagenesis (SSM) library of all positions and sorted for binding after incubation with HA at 22°C, 30°C, 37°C and 42°C (Supplementary Fig. 5). SSM libraries were incubated with 5 nM H1 SI in 50 μl at 22°C, 30°C, 37 °C and 42°C for 2 h while rotating, washed once with ice-cold PBSF and incubated on ice for suspending in 100 mL with 72 nM SAPE and 2 ng/ml anti-c-Myc antibody. After deep-sequencing of all pools, we used linear regression of the enrichment values for each sort at increasing temperatures to identify substitutions that enable better binding at higher temperatures and plotted these as a heat map using a custom Python script. Substitutions that were increasingly enriched over the temperature range were included in the design of a combinatorial library if they were not part of the interface with HA. The variant with the highest expression level in *E. coli* was selected, which also exhibited the most monodisperse SEC trace. This variant was used for further analysis, HSB.2, and it differs from HSB.1C at 8 positions. The combinatorial library based on the identified mutations was screened by sorting at 5 nM H1 SI at 37°C.

BioLayer Interferometry (BLI)

Titrations were performed on an OctetRED96 BLI system (ForteBio, Menlo Park, CA) using streptavidin-coated biosensors. Sensors were equilibrated for 20 min in PBSTB buffer (PBS, 0.002% Tween 20, 0.01% BSA). Titrations were executed at 30° C while rotating at 1000 rpm; baseline were set for 60 sec. Each sensor was loaded with 7 – 20 nM biotinylated HA for 100 sec. Data was fit to a 1:1 interaction. For measurements resulting in apparent K_d values below 1 nM, we report only as sub-nanomolar affinity, as the off-rates were too slow to measure accurately.

Crystallization of the Unbound Monomer and Complex

Initial crystal screening was carried out at the JCSG robotic crystallization facility using the Rigaku Crystalmation system. Crystals of HSB.2 were grown by sitting drop vapor diffusion in drops containing a protein and mother liquor at a 1:1 ratio with a well solution containing 15% PEG 8000, 0.1 M HEPES pH 7.9. Crystals were cryo-protected in mother liquor supplemented with 20% ethylene glycol prior to flash freezing in liquid N_2. Diffraction data were collected at SSRL beam line 12-2. For the HA-HSB complex, H3 HK68 and HSB.2A were incubated at a 1:1.5 molar ratio for 1 hour at 4°C before the sample was concentrated to 10 mg/ml in 25 mM Tris pH 8.0, 150 mM NaCl. Crystals were grown with a well solution of 10% PEG 6000, 100 mM MES pH 6.0 using the sitting drop vapor diffusion method and cryo-protected in well solution supplemented with 20% ethylene glycol before being flash

frozen in liquid N_2. Data were collected at APS beamline 23 ID-D (GM/CA @ APS) and processed with HKL2000[36]. Phaser[37] was used for molecular replacement within Phenix using two search models; a single protomer of H3 HK68 HA from PDB ID 4FNK and a monomeric model of HSB.2. The resulting solution contained one HA trimer and two copies of HSB. Density for the third HSB copy was clearly observable at the third RBS and a copy was manually placed by superimposing an HA-HSB complex from within the model. The position of the manually placed copy was further optimized by rigid body fitting within the F_O-F_C map in COOT[38] and rigid body refinement in Phenix[39]. The model was refined through iterative rounds of model building in COOT and refinement in Phenix. Several regions of the third copy of HSB are poorly ordered. The final model contains all regions of the protein with $2F_O$-F_C density for backbone atoms at 1.0 σ. Glycans and waters were manually added and edited in COOT. The final models were assed with quality metrics within the Phenix. refine interface and the JCSG quality control server, both of which utilize MolProbity[40].

Electron Microscopy

Samples were prepared by mixing purified HK68 HA with trimeric-HSB (Tri-HSB.1C) at a 1:2 molar ratio. The complex was then purified by size exclusion chromatography on Superdex S200 16/60 column equilibrated with TBS (25 mM Tris pH 8.0, 150 mM NaCl) to separate excess HSB. A thin layer of nitrocellulose was floated onto 400-mesh copper grids followed by a thin layer of evaporated carbon. The complex was diluted to ~60μg/ml and placed on a freshly glow discharged grid using a Gatan Solarus. Grids were stained with a 3 μL droplet of Nano-W for 7.5 seconds, blotted and allowed to dry. All samples were treated using this standardized method. Images were collected at 52kX magnification on a Tecnai T12 spirit at 120kV equipped with a Tietz F416 4kx4k CMOS camera, resulting in a final pixel size of 2.05 Å/pixel. The Leginon[41] software package was used for all automated image acquisition. Areas of the grids with thin crystalline Nano-W resulted in the best-resolved particles. Particles were picked using DoGpicker through the Appion web-based software package and stacked into 128 pixel boxes. Particles were aligned using either Xmipp CL2D[42] or IMAGIC[43] to remove suboptimal particles. Final alignment was performed with ISAC[44] and 16 classes were chosen to build an initial 3D asymmetric model using EMAN2[45]. Further model refinement with 14,956 raw particles was performed using EMAN.

Circular dichroism spectroscopy

HSB.2 was tested for folding and its melting temperature using an Aviv 420 circular dichroism spectrometer (Aviv Biomedical, Lakewood, NJ). Protein was diluted to 1 mg/mL in PBS and ellipticity between 205 and 260 nm was measured. To obtain its melting temperature by measuring ellipticity at 220 nm was measured in 2°C increments between 15° and 95°C.

Protease Resistance Assay

Before evaluating the neutralization potential of design variants, proteins were tested for its susceptibility to proteases as the cell-based neutralization assay contains high amounts of trypsin. Cleavage can occur either if a site is accessible in a loop area of if the protein is

partially unfolded and exposes a site. Proteins to be tested were diluted to 1mg/mL in PBS to which trypsin was added with a final concentration of 0.005% (w/v) (Gibco 0.25% (w/v) trypsin/phenol red, Life Technologies, Carlsbad, CA) and incubated at 37°C. Time points were taken by removing 6 μL aliquots of solution. Digestion was instantly quenched by mixing with 6 μL of 2× Laemmli Sample Buffer (Bio-Rad) and incubating it for 2 min at 95°C. Samples were loaded onto a 12% NuPage Bis-Tris gel (Life Technologies, Carlsbad, CA) and run 1× MES buffer (50 mM MES, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.3). Trypsin did not degrade monomeric or trimeric constructs over the monitored period of 24 h at 37°C.

Influenza and BHA preparation

X-31 (H3N2) influenza A virus (Charles River Laboratories) was used for production and purification of soluble bromelain-cleaved (BHA) as previously described[46].

Small Angle X-Ray Scattering (SAXS) and All-Atom Modeling

Using a 1:3 molar excess of the inhibitor, Tri-HSB.2 and BHA were incubated at 4°C overnight and complex was purified *via* size exclusion chromatography (SEC) directly followed by SAXS. SAXS measurements were conducted on beam line 4-2 at the Stanford Synchrotron Radiation Lightsource. The buffer for size exclusion was 150 mM NaCl, 0.02% NaN_3 10 mM sodium phosphate, 1 mM EDTA at pH 7.5. The eluate from the column was passed through a capillary cell and 1 sec X-ray exposures were collected every 5 sec. Scattering data were scaled for the transmitted beam intensity at each exposure, and azimuthally averaged. The radius of gyration (R_G) and I(0) for each frame were batch analyzed using autoRg, and frames with stable R_G[47] values were merged in primus[48]. The particle distance distribution function [P(r)] plots were generated from GNOM[49] and used to calculate a real-space R_G and maximum linear dimension (D_{Max}) of each protein.

All-Atom Ensemble Modeling of Glycosylated HA-Stem Binder

All-atom glycosylated models of unliganded H3 BHA (3HMG) and BHA-bound Tri-HSB.2 were generated using MODELER as previously described[46]. Glycans were added and set as flexible residues during simulated annealing. For the ensembles of unbound Tri-HSB.2, the glycine linkers connecting the binding domains to the trimerization domain were left unrestrained to allow them to sample conformational space while the trimerization fold was restrained as a rigid-body in the template. As for the BHA bound to Tri-HSB.2, three different ensembles were created in which one, two or three binding domains of Tri-HSB.2 were bound one, two or three HA1 subunits, respectively, within the BHA trimer while constraining Tri-HSB.2 trimerization domain to the rigid-body template. Consistency of the models with the experimental SAXS data was analyzed with the FoXS webserver[50, 51]. The goodness of fit (χ value) for each model relative to the experimental SAXS data of the corresponding protein complex was used to identify the models most consistent with the scattering data. FoXS was also used for multiple ensemble searches (MES)[52] of conformations within a collection of submitted ensembles with distinct conformations (at least 30 all-atom ensembles) to determine whether a single state or a combination of states was sufficient for minimizing the χ value.

Hydrogen Deuterium Exchange (HDX)

To establish an internal standard for intrinsic exchange rates, 1 μM of the tetra peptides PPPI and PPPF were added to unliganded H3 BHA and complexes (prepared as above). HDX reactions of each protein stock were initiated by diluting 3 μg of the protein or complex into deuterated buffer (10 mM Hepes, 150 mM NaCl, 0.02% NaN3) for a final pH of 7.5 and incubated at room temperature for 3 seconds, 20 seconds, 3.5 minutes, or 30 minutes. All reactions were quenched with ice cold quench buffer (100 mM TCEP [tris(2-carboxyethyl) phosphine], 0.2% formic acid) for a final pH of 2.5, rapidly frozen in liquid nitrogen and stored at -80°C until analysis. Undeuterated controls were performed as above with optima pure (Thermo-Fisher) water in place of D_2O. To account for "in-exchange" during online-pepsin digestion (described in the mass spectrometry section below), a "zero" time point was made by diluting 3 μg of stock protein or complex into the quenched deuteration buffer and freezing as mentioned above. Fully deuterated "protein" stocks were prepared by denaturing BHA (H3) in 10 mM DTT, 3.5 M Gnd-HCl, 10 mM Hepes, 150 mM NaCl at 80°C for 30 minutes, and 3 μg denatured of BHA was added to deuteration buffer containing PPPI and PPPF for 30 minutes at 40°C. Fully deuterated "peptide" stocks were made by collecting the BHA pepsin peptides that were digested by the online-pepsin column and eluted from the liquid chromatography gradient used for mass spectrometry (explained in further detail below). Following elution, the BHA peptides were dried and resuspended in the fully deuterated denaturant buffer, and deuteration of the peptides was carried out identically.

Mass Spectrometry

Analysis of Tri-HSB.2 Complex and Identification of HA peptides—Undeuterated BHA peptides were identified as described[46] on a Waters Synapt G1 HDMS. HDX samples were loaded onto a 250 μL steel line UPLC (ultra-pressure liquid chromatography) loop and injected onto 2.1×50 mm cartridge filled with POROS® AL/20 resin containing immobilized porcine pepsin (D. Smith 96 – pepsin immobilization for HDX paper) at 200 μL/min with solvent C (2% acetonitrile [ACN], 0.1% FA, 0.04% tri-fluoroacetic acid [TFA] pH 2.5) and peptides were trapped onto 1.7 μm 1×17 mm BEH reverse phase C_{18} (Waters) trap column in-line with the pepsin column for 5 minutes. The peptides were then resolved using a BEH 1.7 μm 1×100 mm C_{18} (Waters) reverse phase column running a gradient of 5% to 40% solvent B for 9 minutes (solvent A: 2% acetonitrile [ACN], 0.1% FA, 0.04% TFA pH 2.5; solvent B: 100% ACN, 0.1% FA). The injection loop, pepsin, trap and analytical columns were kept at a constant 1°C within the Waters HDX manager module. During the main chromatographic gradient, the pepsin column is switched offline and the LEAP Technologies dual arm PAL was programed to perform a series of pepsin column washes with two 250 μL injections of 2M guanidine hydrochloride (GndHCl), 0.1% TFA pH 2.2; one 250 μL 10% FA and one 250 μL injection of 0.1% FA with 200 μL/min of solvent C. After the main LC gradient, the pepsin column is back-flushed with 0.1% FA, 0.04% TFA at 100 μL/min for 15 minutes. During this time, the trap column was washed to minimize sample carryover by cycling through one 250 μL injection of 10% FA, one 250 μL injection of 50% trifluoroethanol, one 250 μL injection 80% methanol, one 250 μL injection of a 2:1 mixture of ACN to isopropanol and one 250 μL injection of 80% ACN (Fang et al 2011) and re-equilibrated with solvent C for subsequent injections. Mass shifts of each resolved

isotopic envelopes were calculated by fitting to a tailored binomial distribution within the semi-automated HX-Express v2 software[53] (HX-Express.com). Experimental error was calculated using the standard deviation from duplicate measurements. For glycoproteins, the exchange for the major glycoform is reported. Percent exchange was calculated for each time point relative to the "zero" and fully deuterated standards. In a few cases where the signal for the fully deuterated "protein" standard was insufficient (presumably due to aggregation prior to digestion or altered digestion efficiency) the fully deuterated "peptide" standard was used to calculate the percent exchange.

Cell Neutralization Assay—To evaluate whether the designed inhibitors could indeed neutralize *in vitro*, virus was grown as a green fluorescent protein (GFP)-based system, in which the *pb1* gene is replaced by *gfp*, previously described as PB1flank-eGFP viruses[54]. The trimeric inhibitor variants were titrated and incubated with indicated virus at 37°C for 1 h to allow protein binding before cells were added. After an 18-h incubation, the GFP fluorescence intensity of the cells was measured and the signal above background was normalized to the negative control. Values are reported as percent infectivity remaining averaged over triplicate measurements.

Diagnostic Application: Detection of Hemagglutinin on Paper—Protein solutions were filtered through a centrifugal filter (0.2 μm nylon membrane, VWR) before spotting. Capture proteins were immobilized to nitrocellulose membrane strips (GE FF80 HP, GE Healthcare) *via* a piezoelectric printer (sciFLEXARRAYER S3, Scienion AG) by depositing 12 spots (250 μm apart from each other) at 30 drops per spot and 450-500 pL per drop. Each test line received approximately 300 nL of total protein solution. The spotted membrane strips were assembled into a comb for ease of use and solutions were pre-loaded into 96-well plate (20 μL).. Tri-HSB.1C was prepared at 0.62 mg/mL, and recombinant streptavidin (Abcam) at 1.0 mg/mL. Membranes were moved from one well to another to apply the solutions. For the detection of HA using Tri-HSB.1C as a capture reagent, the following sequential steps were performed: (1) 100 nM recombinant HA, (2) 100 nM biotinylated HB36.6, and (3) Au-streptavidin. After each step, washing with PBST (PBS+0.1% (v/v) Tween 20 (PBST) and 1% (w/v) BSA) was performed. Recombinant HA was from one of two influenza strains, as indicated: A/California/04/2009 (Influenza Reagent Resource (IRR)) or A/Solomon Islands/03/2006 (IRR). For the inverted set-up, biotinylated HB36.6 was immobilized to nitrocellulose by mutant streptavidin[25] (Abcam) lined on the nitrocellulose membrane. Detection was through Tri-HSB.2 that was conjugated to HRP for signal amplification. Each step consisted of the following reagents delivered in order: 100 nM recombinant HA premixed with 10 nM bHB36.5 and 300 nM of Tri-HSB.2-HRP (2.) 1 mg/mL biotin-BSA (Biovision, Inc) (3.) Amplification buffer with 0.0125% H_2O_2 and 122 mg/mL of DAB as the substrate for HRP.

Protein Administration and Influenza Challenge—Animal studies were approved by the University of Washington Institutional Animal Care and Use Committee. For in vivo efficacy of Tri-HSB.2, female BALB/c mice (Envigo) of 6-8 weeks of age were randomly assigned groups of 5 mice per cage by the housing facility. N=10/group yields 80% power to see a significant difference in protection based on 100% infectivity in controls. Investigators

were not blinded to the treatment group identification. Mice were anesthetized with 2.5% isoflurane and intranasally (IN) administered protein binder (Tri-HSB.2) at a dose of 3 mg/kg body weight at a final volume of 30 μL per mouse. Mice were anesthetized with 2.5% isoflurane and challenged IN with 1000 TCID$_{50}$ (50% tissue culture infectious dose) of X-31 virus (A/Hong Kong/X31/1968) in 30 μL of PBS. The mice were monitored daily for weight loss and survival until 14 days post-infection. Animals that lost more than 25% of their initial body weight were euthanized by carbon dioxide and cervical dislocation in accordance with our animal protocols.

Statistical Analysis

Error bars, *n* values, and type of replicates for each experiment are defined in the respective section within Methods and figure legends. For comparison of the body weights of the animals for the Influenza challenge, an unpaired t test was performed to determine the P-value.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Wilson IA, Skehel JJ, Wiley DC. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. Nature. 1981; 289:366–373. [PubMed: 7464906]

2. Lyumkis D, et al. Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. Science. 2013; 342:1484–1490. [PubMed: 24179160]

3. Heldin CH, Miyazono K, ten Dijke P. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. Nature. 1997; 390:465–471. [PubMed: 9393997]

4. Heldin CH. Dimerization of cell surface receptors in signal transduction. Cell. 1995; 80:213–223. [PubMed: 7834741]

5. Ekiert DC, et al. Cross-neutralization of influenza A viruses mediated by a single antibody loop. Nature. 2012; 489:526–532. [PubMed: 22982990]

6. Schmidt AG, et al. Viral receptor-binding site antibodies with diverse germline origins. Cell. 2015; 161:1026–1034. [PubMed: 25959776]

7. Lee PS, et al. Heterosubtypic antibody recognition of the influenza virus hemagglutinin receptor binding site enhanced by avidity. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109:17040–17045. [PubMed: 23027945]

8. Xu R, et al. A recurring motif for antibody recognition of the receptor-binding site of influenza hemagglutinin. Nat Struct Mol Biol. 2013; 20:363–370. [PubMed: 23396351]

9. Cunha BA. Influenza: historical aspects of epidemics and pandemics. Infect Dis Clin North Am. 2004; 18:141–155. [PubMed: 15081510]

10. Both GW, Sleigh MJ, Cox NJ, Kendal AP. Antigenic drift in influenza virus H3 hemagglutinin from 1968 to 1980: multiple evolutionary pathways and sequential amino acid changes at key antigenic sites. Journal of virology. 1983; 48:52–60. [PubMed: 6193288]

11. Carrat F, Flahault A. Influenza vaccine: the challenge of antigenic drift. Vaccine. 2007; 25:6852–6862. [PubMed: 17719149]

12. Schafer JR, et al. Origin of the pandemic 1957 H2 influenza A virus and the persistence of its possible progenitors in the avian reservoir. Virology. 1993; 194:781–788. [PubMed: 7684877]

13. Bloom JD, Gong LI, Baltimore D. Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. Science. 2010; 328:1272–1275. [PubMed: 20522774]

14. Simonsen L, et al. The genesis and spread of reassortment human influenza A/H3N2 viruses conferring adamantane resistance. Molecular biology and evolution. 2007; 24:1811–1820. [PubMed: 17522084]

15. Sauter NK, et al. Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar dissociation constants: a 500-MHz proton nuclear magnetic resonance study. Biochemistry. 1989; 28:8388–8396. [PubMed: 2605190]

16. Sun XL. Recent anti-influenza strategies in multivalent sialyloligosaccharides and sialylmimetics approaches. Curr Med Chem. 2007; 14:2304–2313. [PubMed: 17896978]

17. Waldmann M, et al. A nanomolar multivalent ligand as entry inhibitor of the hemagglutinin of avian influenza. J Am Chem Soc. 2014; 136:783–788. [PubMed: 24377426]

18. Ekiert DC, et al. A highly conserved neutralizing epitope on group 2 influenza A viruses. Science. 2011; 333:843–850. [PubMed: 21737702]

19. Corti D, et al. A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. Science. 2011; 333:850–856. [PubMed: 21798894]

20. DiLillo DJ, Tan GS, Palese P, Ravetch JV. Broadly neutralizing hemagglutinin stalk-specific antibodies require FcgammaR interactions for protection against influenza virus in vivo. Nat Med. 2014; 20:143–151. [PubMed: 24412922]

21. Schmidt AG, et al. Preconfiguration of the antigen-binding site during affinity maturation of a broadly neutralizing influenza virus antibody. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110:264–269. [PubMed: 23175789]

22. Azoitei ML, et al. Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold. Science. 2011; 334:373–376. [PubMed: 22021856]

23. Azoitei ML, et al. Computational design of high-affinity epitope scaffolds by backbone grafting of a linear epitope. Journal of molecular biology. 2012; 415:175–192. [PubMed: 22061265]

24. Chao G, et al. Isolating and engineering human antibodies using yeast surface display. Nature protocols. 2006; 1:755–768. [PubMed: 17406305]

25. Holstein CA, et al. Immobilizing affinity proteins to nitrocellulose: a toolbox for paper-based assay developers. Anal Bioanal Chem. 2016; 408:1335–1346. [PubMed: 26427504]

26. Galimidi RP, et al. Intra-spike crosslinking overcomes antibody evasion by HIV-1. Cell. 2015; 160:433–446. [PubMed: 25635457]

27. Fleishman SJ, et al. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. Science. 2011; 332:816–821. [PubMed: 21566186]

28. Leaver-Fay A, et al. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. Methods in enzymology. 2011; 487:545–574. [PubMed: 21187238]

29. Huang PS, et al. RosettaRemodel: a generalized framework for flexible backbone protein design. PloS one. 2011; 6:e24109. [PubMed: 21909381]

30. Lin YR, et al. Control over overall shape and size in de novo designed proteins. Proceedings of the National Academy of Sciences of the United States of America. 2015; 112:E5478–5485. [PubMed: 26396255]

31. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC bioinformatics. 2004; 5:113. [PubMed: 15318951]

32. Bakan A, Meireles LM, Bahar I. ProDy: protein dynamics inferred from theory and experiments. Bioinformatics. 2011; 27:1575–1577. [PubMed: 21471012]

33. Cooper S, et al. Predicting protein structures with a multiplayer online game. Nature. 2010; 466:756–760. [PubMed: 20686574]

34. Benatuil L, Perez JM, Belk J, Hsieh CM. An improved yeast transformation method for the generation of very large human antibody libraries. Protein engineering, design & selection : PEDS. 2010; 23:155–159.

35. Strauch EM, Fleishman SJ, Baker D. Computational design of a pH-sensitive IgG binding protein. Proceedings of the National Academy of Sciences of the United States of America. 2014; 111:675–680. [PubMed: 24381156]

36. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Method Enzymol. 1997; 276:307–326.

37. McCoy AJ, et al. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]

38. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]

39. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–221. [PubMed: 20124702]

40. Chen VB, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 2010; 66:12–21. [PubMed: 20057044]

41. Suloway C, et al. Automated molecular microscopy: the new Leginon system. J Struct Biol. 2005; 151:41–60. [PubMed: 15890530]

42. Sorzano CO, et al. A clustering approach to multireference alignment of single-particle projections in electron microscopy. J Struct Biol. 2010; 171:197–206. [PubMed: 20362059]

43. van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M. A new generation of the IMAGIC image processing system. J Struct Biol. 1996; 116:17–24. [PubMed: 8742718]

44. Hohn M, et al. SPARX, a new environment for Cryo-EM image processing. J Struct Biol. 2007; 157:47–55. [PubMed: 16931051]

45. Tang G, et al. EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol. 2007; 157:38–46. [PubMed: 16859925]

46. Garcia NK, Guttman M, Ebner JL, Lee KK. Dynamic changes during acid-induced activation of influenza hemagglutinin. Structure. 2015; 23:665–676. [PubMed: 25773144]

47. Petoukhov MV, Svergun DI. Analysis of X-ray and neutron scattering from biomacromolecular solutions. Current opinion in structural biology. 2007; 17:562–571. [PubMed: 17714935]

48. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of Applied Crystallography. 2003; 36:1277–1282.

49. Svergun DI. Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. Journal of Applied Crystallography. 1992; 25:495–503.

50. Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. Accurate SAXS profile computation and its assessment by contrast variation experiments. Biophys J. 2013; 105:962–974. [PubMed: 23972848]

51. Schneidman-Duhovny D, Hammel M, Sali A. FoXS: a web server for rapid computation and fitting of SAXS profiles. Nucleic Acids Res. 2010; 38:W540–544. [PubMed: 20507903]

52. Pelikan M, Hura GL, Hammel M. Structure and flexibility within proteins as identified through small angle X-ray scattering. Gen Physiol Biophys. 2009; 28:174–189.

53. Guttman M, Weinkam P, Sali A, Lee KK. All-atom ensemble modeling to analyze small-angle x-ray scattering of glycosylated proteins. Structure. 2013; 21:321–331. [PubMed: 23473666]

54. Hooper KA, Bloom JD. A mutant influenza virus that uses an N1 neuraminidase as the receptor-binding protein. Journal of virology. 2013; 87:12531–12540. [PubMed: 24027333]
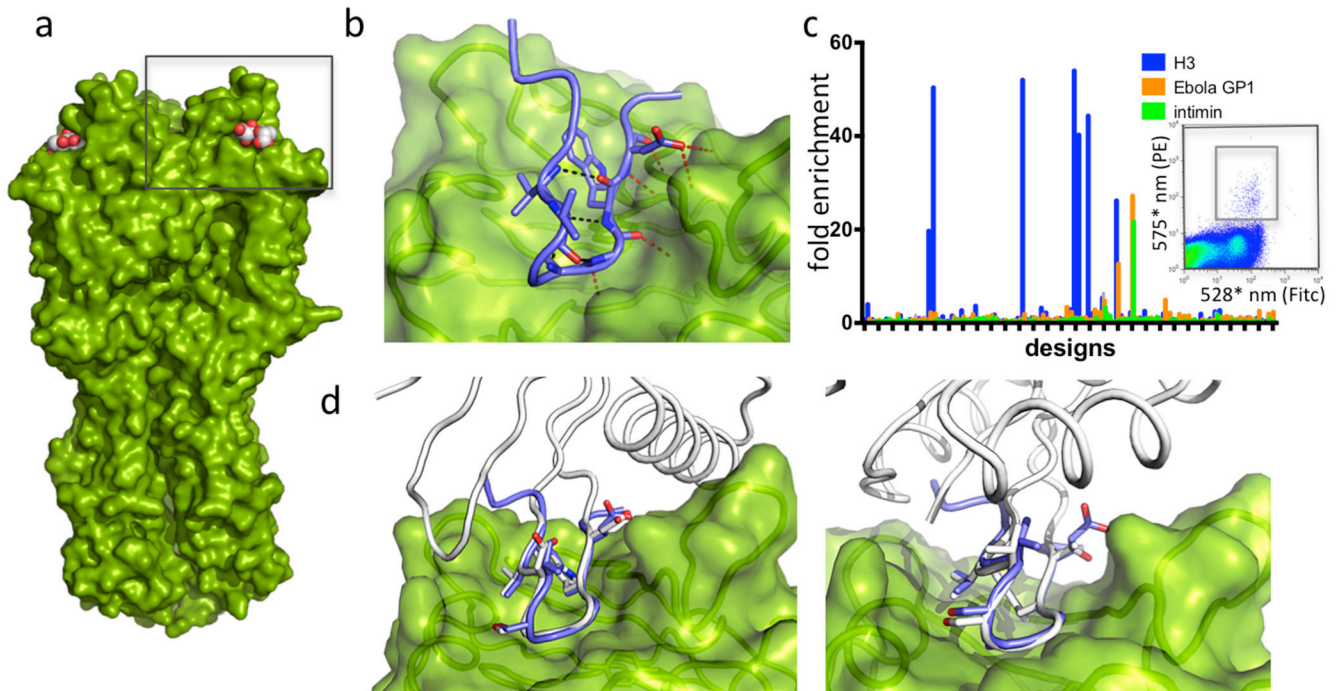
Figure 1. Design strategy

(a) Trimeric HA bound to receptor analog 6·SLN (PDB entry 4wea). (b) Broadly neutralizing antibody loops interact with the receptor binding site either by inserting an aromatic or aliphatic residue into a pocket above the conserved Trp153, as for HCDR3 of C05[5] and many other RBS-targeted antibodies[8] or by using an aspartate to mimic the carboxyl of sialic acid[6, 7] (Fig.S1). (c) Evaluation of 80 HA binder designs by yeast display followed by next-generation sequencing. The design pool was screened against H3 HK68 and two negative control targets (Ebola GP and *E. coli* intimin) to monitor off-target binding to HA. Enrichment values are the ratio of the frequency of a given design in FACS-selected populations for one of the three targets to the frequency in the original unselected population. A number of the designs specifically bind HA, but not Ebola GP or intimin. (d) Models of designs that specifically bind HA (#24 and HSB); the C05 derived loop is colored in purple.
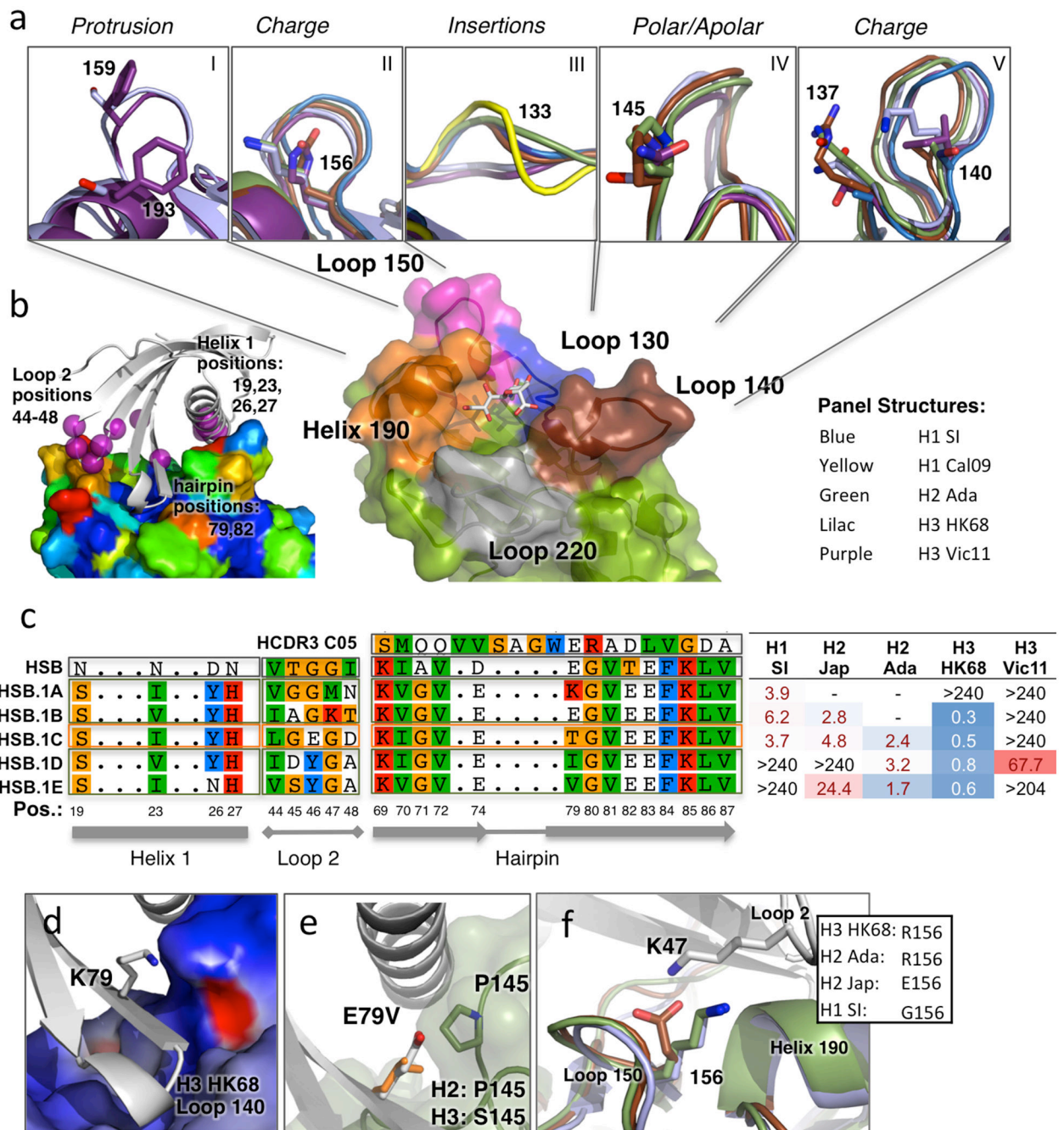
Figure 2. Complementing HA structural diversity with designed HSB variants

(a) Sequence and structural diversity around the receptor binding site of HA (4wea) presents a challenge for binding of molecules larger than sialic acid (PDB entries 4fqr, 3qqi, 4o5n and homology models, see Supp. Info). (b) Design of HSB library: residues of HSB (purple spheres) that come in contact with HA (4fp8) were allowed to mutate to match the diversity of nearby segments of HA. HA is colored by Shannon Entropy over a set of ~3700 H3 sequences (highly conserved in blue, less in red, Suppl. Info). (c) (left) Sequence alignment of HA-contacting residues for isolates identified in independent sorting trajectories against

different flu strains (HSB.1A and B were selected for H1 A/Solomon Islands/3/2006 (SI), HSB.1C for H1 A/New Caledonia/20/1999 (NewCal), HSB.1D and E alternately against H3 HK68/Vic11 and H2 A/Adachi/2/1957 (Ada)). (right) mean of the apparent K_d (nM) from yeast surface titrations of two independent yeast cultures measured on different days. (d) Model of HSB.1A (grey cartoon) with H3 HK68 (colored by electrostatic potential, see Fig. S3) indicates electrostatic repulsion between K79 and the positively charged RBS of H3 HK68; H1 SI is the only strain tested with a net negative charge within Loop 140 (Supplementary Fig. 5) explaining the high preference of HSB.1A for this strain (e) Substitution E79V of HSB.1E (grey cartoon) promotes binding to H2 strains (green) which commonly have a proline at position 145 (Fig.2a, panel IV). (f) Between loop 150 and helix 190 of HA (Fig. 2a, panel II), HSB.1B projects a positive charge (K47) which interacts with the negatively charged E156 of H2 A/Japan/305/1957 (H2 Jap). In contrast, HSB.1D has a negative charge within this loop (D45), which diminishes binding to H2 Jap.
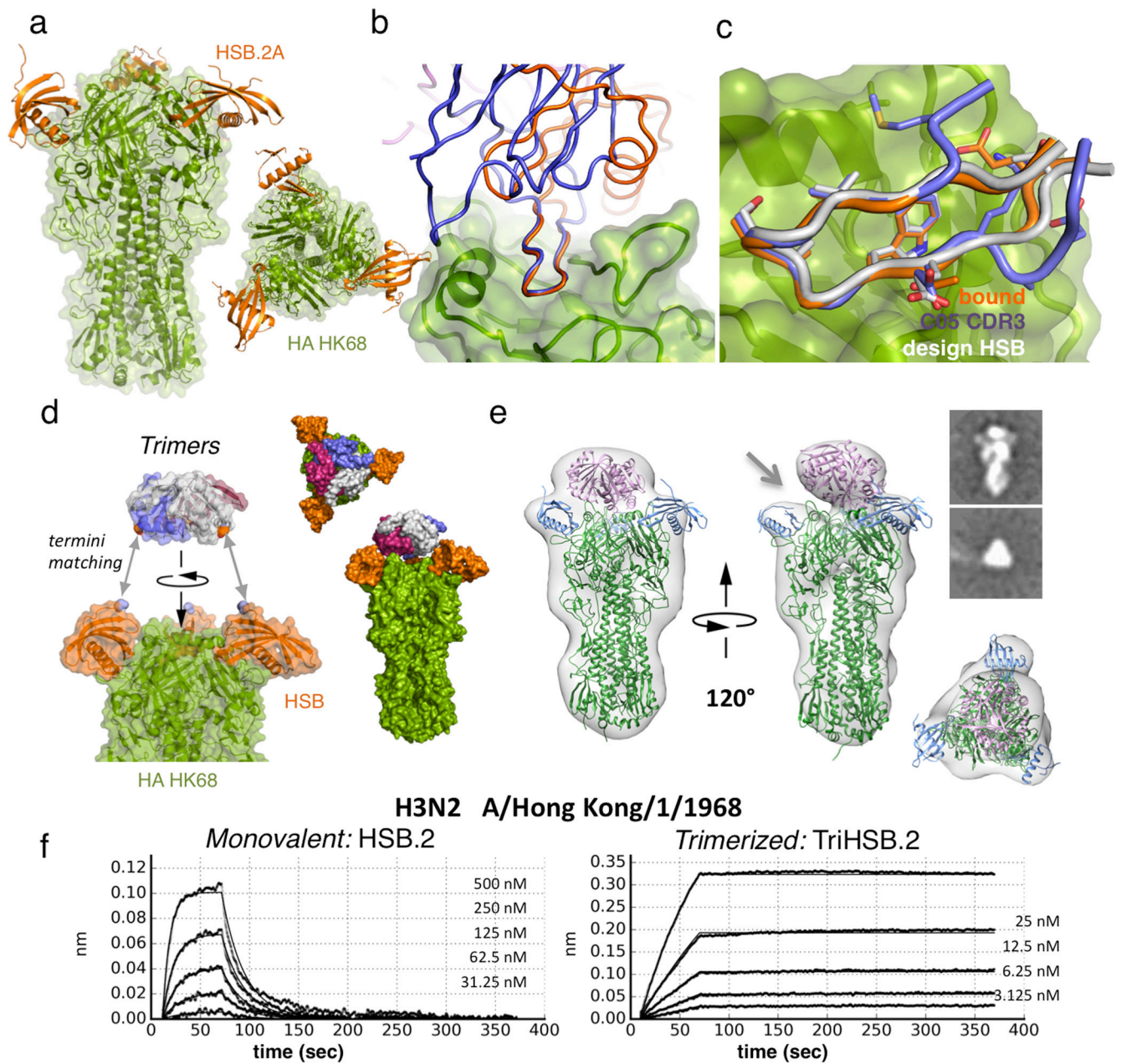
Figure 3. Trimerization of HSB to match the HA trimer
(a) A co-crystal structure of HSB.2A bound to HK68 HA shows that the design binds to the RBS as designed. (b) Superposition of antibody C05 and the Tri-HSB.2A-HA crystal structure. (c) Close agreement is found for the contacts with HA between the tip of HCDR3 of C05 (blue), the HSB designed model (grey), and the bound crystal structure HSB.2A (orange). (d) Translational and rotational sampling of trimeric protein scaffolds (grey, magenta, blue) to identify trimers that connect the termini of three HSB (orange) molecules bound to trimeric HA (green). Termini of HSB are colored purple and those of the trimerizing unit orange. (e) (top left and middle panels) EM reconstruction of Tri-HSB.1C bound to HK68 HA in side view and top view. (top right and bottom panels) Top view and

two side views of the asymmetric EM reconstruction fitted with X-ray structure of HK68 HA bound to HSB.2A monomers; a break in the EM density is indicated by the arrow. (f) BLI titrations of H3 HK68 HA binding to monomeric HSB.2 and trimer Tri-HSB.2.
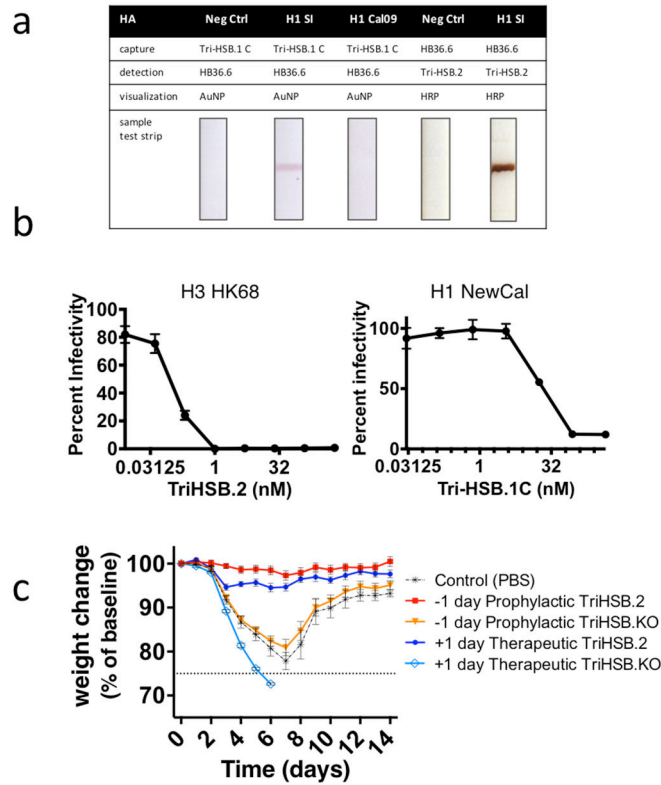
Figure 4. Diagnostic and therapeutic applications

(a) Detection of immobilized HA *via* lateral flow on nitrocellulose paper strips using either a Tri-HSB variant for capture and the stem-region binding design HB36.6[27] (coupled to streptavidin-gold via biotinylation) for detection, or *vice versa.* Tri-HSB.2 readily binds to nitrocellulose and can be directly used as an immobilization reagent for HA. HB36.6 was immobilized through biotinylation followed by nitrocellulose-binding streptavidin. For detection *via* the head-region binder, Tri-HSB.2 was conjugated to HRP for signal amplification. (b) Inhibition of infection of MDCK-SIAT1-CMV-PB1 cells by H3N2 HK68 virus at indicated concentrations. Values are reported as percent infectivity compared to untreated, infected cells averaged over triplicate measurements. TriHSB.2 has an IC_{50} of 0.15 nM for HK68 neutralization, and Tri-HSB.1C an IC_{50} of 19 nM for neutralization of H1 NewCal; both values are close to their apparent K_d, consistent with a simple receptor blocking mechanism. (c) Mice (n=10) treated with a single dose of 3 mg/kg TriHSB.2/body weight either before (prophylactically) or after (therapeutically) intranasal challenge of H3N2 (X-31) virus show substantial protection from weight loss (P < 0.00001) following influenza challenge compared to PBS or non-binding Tri-HSB.2ko-treated controls. Mean values are plotted for each group and with s.e.m.as error bars.