The Protein Journal (2020) 39:434–448
https://doi.org/10.1007/s10930-020-09930-z

Influenza Hemagglutinin Head Domain Mimicry by Rational Design

V Vamsee Aditya Mallajosyula[1] · Shiv Swaroop[1,2] · Raghavan Varadarajan[1]

Accepted: 9 October 2020 / Published online: 17 October 2020

Abstract
Despite diligent vaccination efforts, influenza virus infection remains a major cause for respiratory-related illness across the globe. The less-than-optimal immunity conferred by the currently prescribed seasonal vaccines and protracted production times warrant the development of novel vaccines. Induction of an epitope-focused antibody response targeting known neutralization epitopes is a viable strategy to enhance the breadth of protection against rapidly evolving infectious viruses. We report the development of a design framework to mimic the hemagglutinin (HA) head fragment of H1-subtype viruses by delineating the interaction network of invariant residues lining the receptor binding site (RBS); a site targeted by cross-reactive neutralizing antibodies. The incorporation of multiple sequence alignment information in our algorithm to fix the construct termini and engineer rational mutations facilitates the facile extension of the design to heterologous (subtype-specific) influenza strains. We evaluated our design protocol by generating head fragments from divergent influenza A H1N1 A/Puerto Rico/8/34 and pH1N1 A/California/07/2009 strains that share a sequence identity of only 74.4% within the HA1 subunit. The designed immunogens exhibited characteristics of a well-ordered protein, and bound conformation-specific RBS targeting antibodies with high affinity, a desirable feature for putative vaccine candidates. Additionally, the bacterial expression of these immunogens provides a low-cost, rapidly scalable alternative.

Keywords Protein engineering · Interaction network · In silico design · Energy minimization · Sequence conservation · Protein stability · Immune focusing · *Escherichia coli* · Post translation modifications · Subunit vaccine · Hemagglutinin (HA) · HA head · Glycosylation

1 Introduction

Influenza virus infections are a global health and economic burden despite concerted efforts towards surveillance and vaccination programs. Currently, a seasonal influenza vaccine is the preferred prophylactic modus operandi. Recommendations for the composition of the trivalent/quadrivalent seasonal influenza vaccine are made by the World Health Organization after elaborate analysis of virological and epidemiological influenza surveillance data across the globe. Development of the seasonal, inactivated influenza vaccine (IIV) and live attenuated influenza vaccine (LAIV) necessitate virus-culture in embryonated chicken eggs, a process that is time-consuming and expensive [1–3]. A mismatch between the vaccine formulations and circulating influenza virus strains abetted by a drawn-out vaccine production time have caused significant loss [4], observed as recently as the mismatched H3N2 component in the 2014–15 Northern Hemisphere influenza vaccines [5]. These lacunae in the existing influenza vaccination program warrant the development of improved algorithms for the prediction of circulating influenza strains, and novel influenza vaccines that are amenable for rapid scale-up and preferably confer an increased breadth of protection.

Influenza hemagglutinin (HA) is the principal antigen in stimulating a protective humoral response following virus infection/vaccination [6]. HA is the cognate binding partner

V Vamsee Aditya Mallajosyula and Shiv Swaroop have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10930-020-09930-z) contains supplementary material, which is available to authorized users.

✉ Raghavan Varadarajan
varadar@iisc.ac.in

[1] Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India

[2] Department of Biochemistry, Central University of Rajasthan, Kishangarh, Ajmer 305817, India

for the host cellular receptor, α2-6-linked or α2-3-linked sialic acid for human and avian influenza viruses respectively [7–9]. A high-avidity interaction between HA and the host receptor drives virus attachment to the cell surface initiating infection. Subsequently, HA undergoes an obligatory conformational rearrangement in the host endosome triggered by a low pH stimulus facilitating virus-host membrane fusion [10]. Influenza HA is synthesized as a precursor polypeptide (HA0) and is assembled into an oligomer along the cell secretory pathway [8, 11]. The fusion-competent, cleaved HA has a membrane-distal, globular head domain that encompasses the receptor binding site (RBS). A membrane-proximal, elongated stem domain anchors the HA glycoprotein to the virus envelope. At low pH, extensive rearrangement of the HA stem follows the disassembly of the receptor-binding head domain. The canonical 'antigenic sites' in the HA head domain have been extensively characterized [12, 13]. These sites acquire escape mutations driven by host immune pressure (antigenic drift) [14]. However, functionally critical residues in HA have been identified that are constrained by stringent structural requirements and therefore are recalcitrant to mutations. Evolutionarily silent 'cold spots' in HA span both the head and stem domains [15, 16]. Recent advances in high-throughput antibody repertoire analysis of vaccinated individuals have led to the isolation of cross-protective, neutralizing antibodies (nAbs) targeting these conserved epitopes [17–19]. The properties of these nAbs have been recently reviewed [20, 21]. Although the conserved, neutralization foot-print in the HA head is significantly limited in comparison to the HA stem, the former still affords an attractive, validated target to develop a vaccine candidate capable of eliciting potent nAbs that exhibit cross-reactivity. Furthermore, prophylactic and/or therapeutic monoclonal antibody treatment of influenza infection(s) in mice suggests better in vivo efficacy of HA head specific nAbs as opposed to those targeting the HA stem [17, 21, 22].

In attempts to focus the immune response towards the RBS and establish a general outline for designing stable head fragment vaccines, we evaluated structures/sequences of HA and developed an algorithm to ascertain the minimum interaction network for a set of identified invariant (within a subtype) 'foundation' residues that are essential in maintaining the structural integrity of the RBS, and critical for HA function. Mapping the interaction network enabled rational engineering of mutations in the designed fragments that aided head domain mimicry. This involved the evaluation of forces contributing to protein stabilization that have been extensively studied by Scheraga and others [23–25], including hydrogen bonds formed by ionizable side-chains interactions at surface exposed, non-polar sites that may lead to hydrophobic driven protein aggregation. Furthermore, incorporation of residue conservation score in fixing construct termini and engineering mutations facilitated facile extension of our design to heterologous (within subtype) influenza strains. We tested our design principle on unmatched influenza A virus strains (H1 subtype). The biophysical and biochemical properties of the designed immunogens produced using a prokaryotic expression system were consistent with characteristics of a well-ordered protein, a desirable trait in vaccine candidates.

2 Materials and Methods

2.1 Sequence Analysis

All non-identical, full-length H1N1 HA sequences (n = 5617) were obtained from the NCBI Influenza Virus Database (https://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html). These sequences were then clustered at 99% homology using CD-HIT (Cluster Database at High Identity with Tolerance) [26] to avoid over-representation in a multiple sequence alignment using ClustalX [27]. The aligned sequences were viewed using WebLogo 3 (https://weblogo.threeplusone.com/create.cgi). Alternatively, the quality score for each column in the alignment file was binned and mapped onto a crystal structure of H1N1 A/Puerto Rico/8/34 HA (PDB ID: 1RU7 [28]).

2.2 Cloning, Expression and Protein Purification

The codon-optimized genes corresponding to $H1HA9^{WT}$, H1HA9 and pH1HA9 were synthesized (Abexome, India or GenScript, USA) and cloned into the bacterial expression vector pET-15b (Novagen), between the NdeI and BamHI restriction sites.

All the constructs were over-expressed in *E.coli* BL21(DE3) cells and purified from the soluble fraction of the culture lysate. Briefly, a primary culture was grown overnight at 37 °C from a single colony of transformed *E.coli* BL21(DE3) cells. Next, 2lt (500 ml × 4) of LB media (HiMedia) was inoculated with 1% of the primary inoculum and grown at 37 °C until an OD_{600} of ~ 0.6–0.8 was attained. The culture was then induced with 1 mM isopropyl-β-thiogalactopyranoside (IPTG) and grown further for 12-16 h at 18 °C. The cells were harvested at 5000 × g and re-suspended in 100 ml of PBS (pH 7.4). The cell suspension was lysed by sonication (on ice) in presence of a phenylmethylsulfonyl fluoride (PMSF) and subsequently centrifuged at 14,000 × g. The supernatant was incubated with buffer-equilibrated Ni–NTA resin (GE HealthCare) for 2 h at 4 °C under mild-mixing conditions. The resin was then washed with 30 mM imidazole (in PBS, pH 7.4) and the bound protein was subsequently eluted using an increasing imidazole gradient (in PBS, pH 7.4) under gravity flow. Fractions containing the protein of interest were pooled and

dialysed against PBS (pH 7.4) containing 1 mM EDTA. The dialysed protein was concentrated in an Amicon (Millipore) stirred cell apparatus to a final concentration of ~ 1 mg/ml. The relative amount of protein in the soluble fraction or inclusion bodies was quantified using the Molecular Imager Gel Doc XR System (Bio Rad).

2.3 Circular Dichroism

The circular dichroism (CD) spectra of the proteins (concentration of ~ 10 μM in PBS, pH 7.4) was recorded using a Jasco J-715C spectropolarimeter flushed with nitrogen gas. Spectra were recorded at room-temperature (25 °C) between 200-250 nm in a 1 mm path-length quartz cuvette with a scan rate of 50 nm/min, a response time of 4 s, and a bandwidth of 2 nm. The represented data was averaged over five consecutive scans and corrected for buffer signals acquired under similar conditions. The mean residue ellipticity (MRE) was calculated as described previously [29].

The transition midpoint (T_m) of thermal denaturation for the proteins was measured using CD. The change in ellipticity as a function of temperature was monitored at a fixed wavelength (λ = 215 nm for H1HA9, and 213 nm for pH1HA9) between 25–95 °C. The temperature slope was fixed at 1 °C/min. The band width and response time were set to 2 nm and 4 s, respectively. The scan was repeated after cooling the samples back to 25 °C to estimate the reversibility of thermal unfolding.

2.4 Fluorescence Spectroscopy

All fluorescence spectra were recorded on a Jasco FP-6300 spectrofluorimeter at 25 °C. The protein concentration used for intrinsic fluorescence measurements was ~ 2 μM. The sample was excited at a wavelength (λ_{ex}) of 280 nm, and fluorescence emission (λ_{em}) was monitored between 300-400 nm. The excitation and emission slit widths were set to 3 nm and 5 nm, respectively. The represented spectra were averaged over five consecutive scans and corrected for buffer signals acquired under similar conditions. The fluorescence measurements were recorded under native (PBS, pH 7.4) or denaturing condition (7 M GdmCl in PBS, pH 7.4) as indicated.

2.5 DTNB Assay

The DTNB assay was performed as described previously [30, 31]. Briefly, the proteins (in PBS, pH 7.4) were incubated with DTNB for 30mins under native, oxidizing (50 mM Tris–HCl (pH 8.0)) or reducing (6 M GdmCl, 1 mM DTT, 50 mM Tris–HCl (pH 8.0)) conditions, and the absorbance of TNB^{2-} ions at 412 nm was measured on a Varian Cary 100 Bio UV–Vis spectrophotometer. The protein reduced with DTT under denaturing conditions was desalted (to remove the reducing agent) prior to DTNB labelling.

2.6 Gel Filtration Chromatography

The oligomeric state of the proteins (~ 25 μg) under native condition (PBS, pH 7.4) was determined by gel filtration chromatography using a Superdex-200 analytical gel filtration column (GE HealthCare). The column was calibrated using a broad range of molecular weight standards (GE HealthCare).

2.7 Reverse Phase-HPLC (RP-HPLC)

12.5 μg of protein in PBS (pH 7.4) was injected into an analytical RP C18 column (150 × 4.6 mm, 5 μm particle size) (Vydac) and eluted using a gradient of water and acetonitrile (5–95%) containing 0.1% trifluoroacetic acid at a flow rate of 1 ml/min. For the reduced sample, 12.5 μg of protein was incubated with 7 M GdmCl, 5 mM DTT in PBS (pH 7.4) at 37 °C for 30mins before injection.

2.8 Enzyme-Linked Immunosorbent Assay (ELISA)

The binding of conformation-specific, RBS targeting antibodies and convalescent sera to the test antigens was determined by ELISA. Briefly, the antigens (H1HA9, pH1HA9, H1N1 A/Puerto Rico/8/34 HA, and pH1N1 A/California/04/2009 HA; 200 ng/well) were coated as indicated on 96-well Nunc plates (Thermo Fisher Scientific, Rochester, NY) in 50 μl PBS (pH 7.4) at 4 °C overnight. The full-length H1N1 A/Puerto Rico/8/34 HA, and pH1N1 A/California/04/2009 HA were purchased from Sino Biological Inc. (Beijing, China). The wells coated with ovalbumin (200 ng/well) served as a control for non-specific binding. The plates were washed with PBS containing 0.05% Tween-20 (PBST) and blocked with 1% BSA in PBST (PBSB) for 1 h. Next, the plates were washed with PBST before the addition of monoclonal antibodies or convalescent sera (as indicated) in 50 μl PBSB, followed by an incubation of 2 h at room temperature. Plates were washed with PBST. The binding of the monoclonal antibodies or convalescent sera was detected by alkaline phosphatase(ALP)-conjugated goat anti-mouse secondary antibody (Sigma) added to each well at a predetermined dilution (1:10,000) and incubated at room temperature for 2 h. The binding of the human nAb2D1 exclusively was detected by the appropriate ALP-conjugated goat anti-human secondary antibody (Sigma). The plates were washed and developed using the chromogenic substrate p-nitrophenyl phosphate (Sigma). The optical density was read at 405 nm (SPECTRAmax Plus 384, Molecular Devices, USA). The reported values are the mean ± SD of three independent experiments.

2.9 Binding Affinity Measurements Using Biolayer Interferometry (BLI)

The binding affinity of the head fragments, H1HA9 and H1HA9^{WT}, to antibody H36-4 was measured by BLI using an Octet RED96 instrument (Pall ForteBio, USA). Hydrated, Protein G (ProG) biosensors were used to capture the antibody (20 µg/ml in PBS with 0.05% Tween-20). The binding of H36-4 with a concentration series of the analytes (H1HA9 and H1HA9^{WT}) was determined. The non-specific binding of the analytes to the sensor surface was assessed by using unliganded ProG biosensors (reference sensors). After each binding experiment, the sensor surface was regenerated with 2 M $MgCl_2$. The data was processed using the ForteBio Data Analysis Software (v8.0). The kinetic parameters were obtained by fitting the data globally to a 1:1 Langmuir interaction model.

3 Results and Discussion

3.1 'Foundation' Residues Provide an Invariant Scaffold for Head Immunogen Design

We wanted to develop a HA head fragment immunogen with the following features; 1. Structural integrity of the RBS and the 'canonical' antigenic sites, 2. Residue contiguity to avoid any complexity in the 'folding' of individual fragments connected by synthetic linker sequences, 3. Incorporation of multiple sequence alignment information to build a robust platform that can be readily adopted to divergent strains.

Several structural studies have evaluated the binding modalities between HA and the host cell-surface receptor [32, 33]. A structural superposition of HA-receptor analogue complexes reveals alternate conformational modes that mediate effective binding. Previous studies have shown that torsion angles (Φ) of oligosaccharides containing α2-6-linked sialic acids in complex with proteins show considerable variation, ranging from 36° to 294° [34, 35]. Likewise, head-specific nAbs with shared specificities bind within the RBS with alternate conformations [36, 37] (Fig. S1). The central feature in all of the aforementioned diverse binding events is the RBS structural integrity. The rim of the RBS comprises three secondary structures; the 190-helix (residues 190–198), the 130-loop (residues 135–138) and the 220-loop (residues 221–228) [32]. In an attempt to build a robust design platform with wide applicability, we calculated the conservation score of residues in the rim of the RBS within a subtype(H1). All available influenza A H1N1 sequences (n = 5617) deposited at the National Center for Biotechnology Information (NCBI) Influenza Virus Database (https://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html) were analyzed to identify conserved residues (Fig. 1, Fig. S2). Next, the solvent accessible surface area (SASA) for all residues in the HA ectodomain of H1N1 A/Puerto Rico/8/34 were calculated. Residues in the rim of the RBS with extreme values of side-chain accessibility were identified (A137, Q191, Y195, P221, G228) (Fig. 1). Conserved residues with limited solvent accessibility (≤ 10%) are critical in maintaining the structural integrity, while residues with a high degree of exposure (solvent accessibility ≥ 90%) likely act as a scaffold for functional plasticity. Further, a set of conserved residues (Y98, W153, H183 and Y195) positioned at the base of the RBS are known to assist receptor binding [32, 33]. These invariant amino acids; Y98, A137, W153, H183, Q191, Y195, P221 and G228 constitute a set of 'foundation' residues on which we built the minimized head fragment from H1N1 A/Puerto Rico/8/34, viz. H1HA9.

We then determined the minimum interaction network of these foundation residues using the in-house software PREDBURASA as described previously [38]. Briefly, significant contacts (total buried accessible surface area ≥ 20.0% between residue pairs) for each of the foundation residues were determined to define a proximal layer of interaction network. The hydrogen bond and electrostatic interactions of the foundation residues were also evaluated. The contribution of side chain-side chain hydrogen bonds in protein stabilization has been elegantly described by Scheraga and can also be used to predict distance constraints in lieu of structural information as demonstrated by Scheraga in bovine pancreatic ribonuclease A [39]. Subsequently, the interaction network was extended outwards (distal to the RBS) by iterative calculations until stable breakpoints were identified strictly in non-secondary structural elements (loops) and spanning a contiguous residue stretch. The solvent accessible A52 was fixed as the N-terminus of the designed construct, H1HA9. A solvent accessible residue (N272) posterior to the edge strand of the beta-sheet (comprised of three strands; L54-G57, I83-E85, I268-T270) was set as the construct C-terminus. Protein minimization leads to loss of native contacts and exposure of previously buried hydrophobic residues that can potentially cause aggregation. We identified such exposed hydrophobic patches and resurfaced them by introducing polar mutations chosen by Rosetta Design (v3.0) [40] and visual inspection (Fig. 2a), an approach we have previously used to design soluble HA stem fragments [41–44]. Analysis of the interaction network facilitated engineering rational mutations (A52S, V84T, V205T, A218T, V223T, I244Q, I269R) that are expected to minimally perturb the conformation of the engineered head domain (Fig. 2a, Fig. S3). Our current model calculates solvation energy for each amino acid substitution with the use of implicit solvent and this significantly expedites energy calculations but results in loss

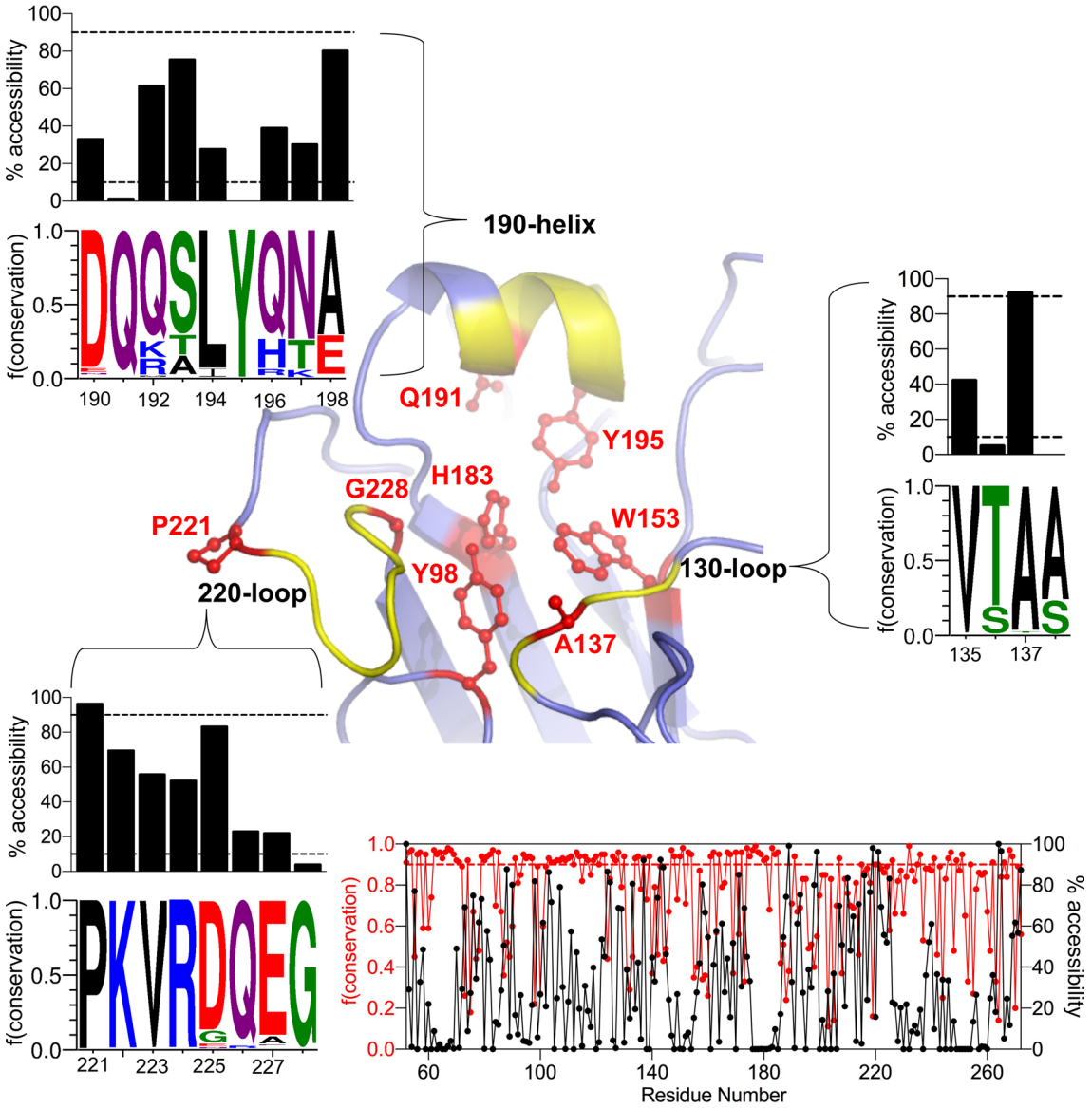

Fig. 1 Foundation residues in the HA RBS provide an invariant platform for head immunogen design. A proximal view of the H1 HA (PDB ID: 1RU7 [28]) RBS (blue, cartoon) is shown. The structural elements that form the RBS; the 190-helix, the 130-loop and the 220-loop are colored in yellow. The conservation score (across all available influenza A H1N1 sequences, n = 5617) and side-chain accessibility of residues (in H1N1 A/Puerto Rico/8/34) lining the RBS are shown. Conserved residues in the rim of the RBS with extreme values of side-chain accessibility ($x \leq 10\%$ or $x \geq 90\%$; A137, Q191, Y195, P221, G228, indicated by dashed lines) along with a set of residues (Y98, W153, H183, Y195) positioned at the base of the RBS known to assist receptor binding constitute the 'foundation' residues. The invariant foundation residues are represented as ball-and-stick (red). The fluctuations in the conservation score across all influenza A H1N1 sequences and the corresponding side-chain accessibility of residues (in H1N1 A/Puerto Rico/8/34) in the engineered head domain (residues 52–272) of HA are also shown. The figure was rendered using PyMOL (Color figure online)

of accuracy. In order to address the challenging problem of protein engineering without disturbing native contacts, the inclusion of explicit waters during energy minimization can be beneficial. However, this comes with high computational cost and improved methods to combine the efficiency of implicit solvation with the accuracy arising from methods that use explicit water molecules are desirable.

3.2 Resurfacing Exposed Hydrophobic Patches Reduces Aggregation

The production of conventional influenza vaccines (IIV, LAIV) is time-consuming and alternate approaches that are economical and expedite the process are being pursued [1]. Recombinant protein expression in *Escherichia coli* provides an ideal platform for manufacturing low-cost

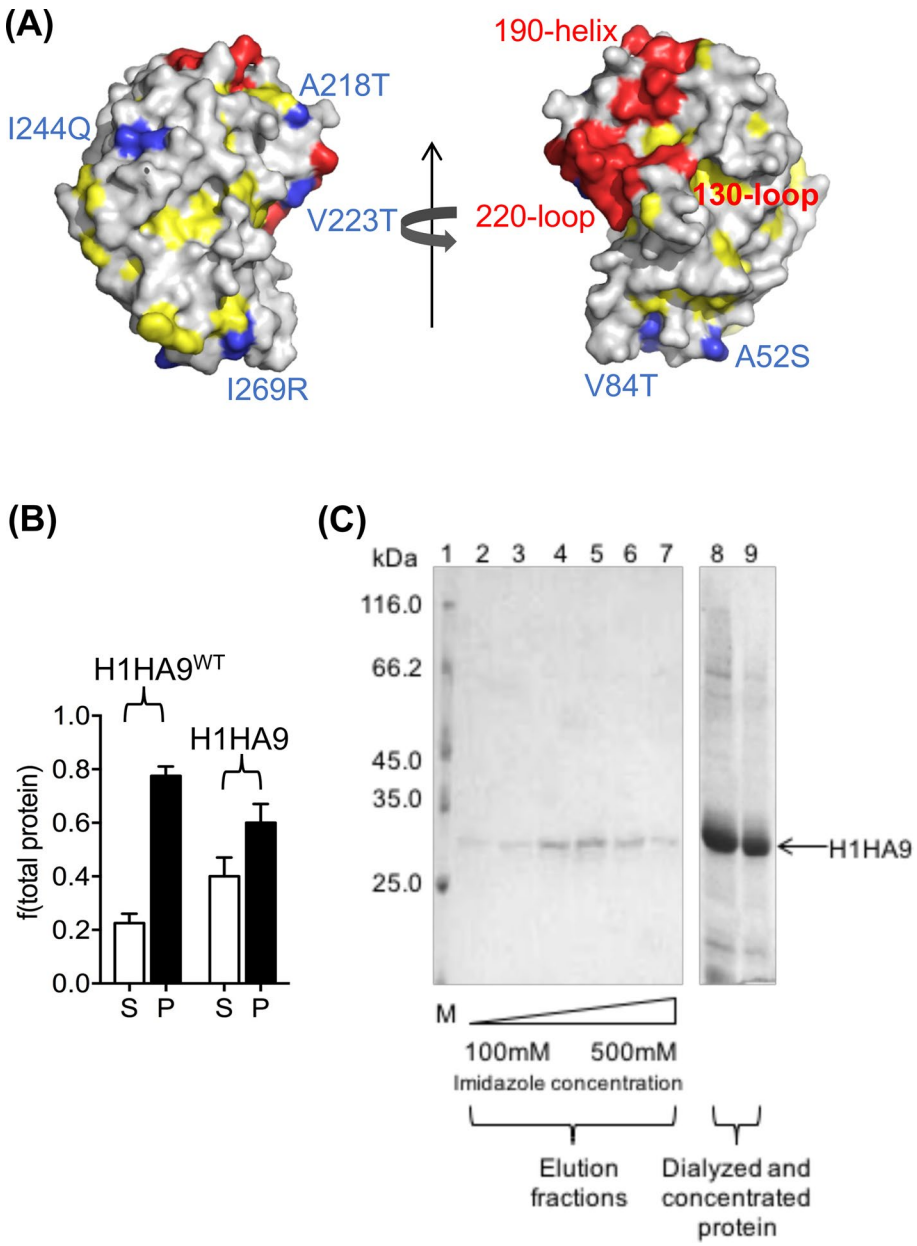

Fig. 2 Rational engineering of exposed hydrophobic patches decreases protein aggregation. **a** Surface representation (light gray) of a model of the head fragment, H1HA9, designed from H1N1 A/Puerto Rico/8/34 HA (PDB ID: 1RU7 [28]). The secondary structural elements lining the RBS (red), exposed hydrophobic residues (yellow), and mutations introduced to reduce aggregation (blue) are shown. The mutation A205S is not discernible in the represented orientations. The figure was generated using PyMOL. **b** The mutations engineered in H1HA9 reduced the fraction of heterologously expressed protein in the insoluble inclusion bodies with concomitant increase in the soluble fraction of the bacterial culture lysate as compared to $H1HA9^{WT}$. **c** The designed head fragment, H1HA9, could be purified to homogeneity from the supernatant of the culture lysate using a single-step of affinity (Ni–NTA) chromatography. SDS-PAGE profile of H1HA9 purification. Lane 1: Protein marker (ThermoFisher Scientific) (M), Lanes 2–7: Elution fractions. The fractions were pooled and dialyzed against PBS (pH 7.4), followed-by concentration: 5 µl (Lane 8) and 2 µl (Lane 9). None of the samples were treated with a reducing agent. The SDS-PAGE gel was stained with Coomassie (Color figure online)

vaccines [45]. However, soluble expression of proteins in a heterologous host remains challenging. Overexpression of eukaryotic proteins in *E.coli* often results in the formation of inclusion bodies and development of an efficient, reproducible refolding strategy by multi-parameter optimization (e.g., pH, osmolytes) can be a bottleneck [46, 47]. Stability of vaccine candidates in phosphate-buffered saline (PBS), a buffer isotonic to human blood, is another desirable trait. The theoretical isoelectric point (pI) corresponding to the wild-type sequence ($H1HA9^{WT}$) of the designed head fragment (52–272) is 7.38, resulting in an effective net-charge of − 0.1 in PBS (pH 7.4). Negligible electrostatic repulsion and exposure of previously buried hydrophobic surfaces cause aggregation of $H1HA9^{WT}$ when over-expressed in *E.coli*. In contrast, H1HA9 had a lower propensity to aggregate (Fig. 2b). We circumvented these problems by optimizing codons for bacterial expression. Importantly, exposed hydrophobic patches were rationally resurfaced to prevent aggregation. Additionally, the engineered mutations, specifically the polar, charged amino acid substitution (I269R) was selected to alter the theoretical pI of H1HA9 to 7.82, resulting in an effective net-charge of + 1.0 in PBS (pH 7.4). Previous efforts at over-expressing and purifying HA head fragment(s) independently or as fusion proteins in *E.coli* required recovery by solubilization and refolding from inclusion bodies [46–48]. In contrast, H1HA9 could be purified to homogeneity using a single-step affinity (Ni–NTA) chromatography (Fig. 2c).

The yield using shake-flask cultures was ~ 3-5 mg/liter culture and further optimization is warranted. Here, we have adopted an approach to resurface exposed hydrophobic patches that can cause aggregation to achieve solubility in a prokaryotic system. A potentially complementary process to obtain well-folded, soluble designed HA subunit vaccine(s) is to test eukaryotic expression systems, since viral proteins are glycosylated and *E.coli* is not adapted to generate such post translational modifications (PTMs). However, our choice of *E.coli* as the expression system was to provide a platform that can be readily adapted in low resource settings.

3.3 Biophysical and Biochemical Characterization of the Designed Head Fragment

Solubility of the protein is a coarse indicator of correct folding, and further characterization of H1HA9 was carried out employing multiple biophysical and biochemical techniques. The circular dichroism (CD) spectrum of H1HA9 was characteristic of a well-folded protein rich in β-sheets (Fig. 3a). Thermal stability of H1HA9 was also monitored using CD. The thermal melt profile of H1HA9 indicates cooperative unfolding, with a transition midpoint (T_m) of 322.8 K (49.7 °C) (Fig. 3b). The overlay of consecutive

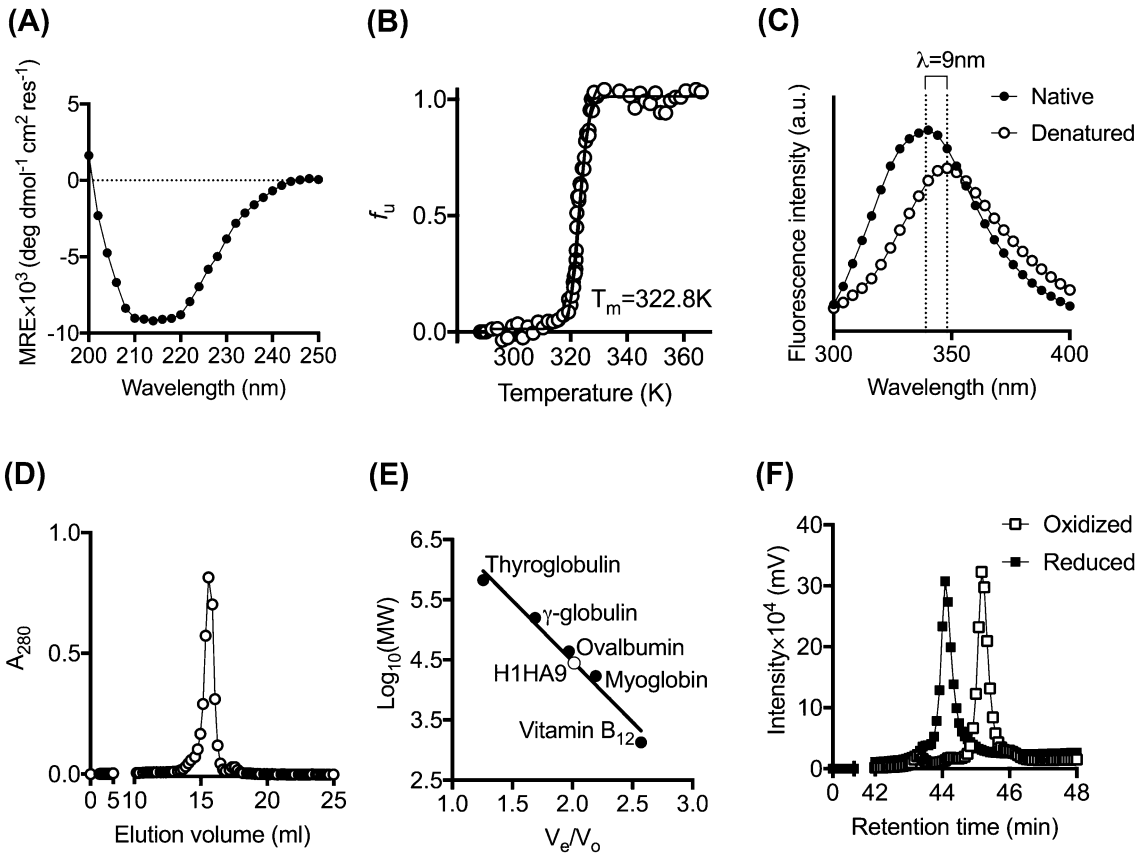

Fig. 3 Biophysical and biochemical characterization of H1HA9. **a** The CD spectra of H1HA9 in the far-UV region (200-250 nm) was characteristic of a well-folded, β-rich protein. The spectra averaged over five consecutive scans is shown. The protein was in PBS (pH 7.4). **b** Thermal stability of H1HA9 was monitored using CD (λ = 215 nm). H1HA9 showed co-operative unfolding with a transition mid-point (T_m) of 322.8 K (49.7 °C). **c** The intrinsic tryptophan fluorescence emission maximum of H1HA9 under native (PBS, pH 7.4) conditions was at 339 nm. The red-shift in the emission maximum of H1HA9 upon denaturation with GdmCl indicates a well-packed conformation with buried hydrophobic residues under native conditions. Signal averaged protein spectra recorded under native (PBS, pH 7.4) or denaturing conditions (7 M GdmCl in PBS, pH 7.4) after correcting for buffer signals are shown. An excitation wavelength (λ_{ex}) of 280 nm was used, and fluorescence emission (λ_{em}) between 300-400 nm was recorded. **d** The oligomeric state of H1HA9 in solution was determined by analytical gel-filtration chromatography under non-denaturing conditions (PBS, pH 7.4) using a Superdex-200 column. H1HA9 (27.5 kDa) eluted as a homogeneous monomer. **e** The column was calibrated using standards spanning a broad range of molecular weights (•) under similar conditions (PBS, pH 7.4). The data-point corresponding to the elution volume of H1HA9 is shown by an open circle (o). **f** RP-HPLC was used to further probe the conformational homogeneity of H1HA9. 12.5 µg of the oxidized (PBS, pH 7.4) and reduced (denatured with 6 M GdmCl prior to incubation with 5 mM DTT in PBS, pH 7.4) H1HA9 was analyzed at 25 °C. The protein was eluted using an increasing gradient of acetonitrile ranging from 5 to 95%at a rate of 1% per minute. A single elution peak corresponding to the oxidized protein, which is shifted upon reduction with DTT, indicates the presence of a homogenous, disulfide-bonded species in solution

scans recorded after cooling the sample back to 15 °C indicates reversibility of thermal denaturation. This cooperative phase transition of the designed immunogen arises from both short- and long-range interactions that stabilize the tertiary structure as described by Hao and Scheraga [49]. A red-shift in the intrinsic tryptophan fluorescence emission maximum of H1HA9 upon denaturation with guanidine hydrochloride (GdmCl) indicates a well-packed, folded conformation of the designed head fragment, with buried hydrophobic residues under native conditions (Fig. 3c).

Disulfide bonds in the head domain provide stability and assist in maintaining the structural integrity of HA. The formation of native disulfide bonds is critical for activity [50, 51]. H1HA9 contains four Cys residues. We assayed for free thiols in H1HA9 using Ellman's reagent (5, 5′-dithiobis-(2-nitrobenzoic acid), DTNB) [30, 31]. The absence of detectable TNB^{2-}(5-mercapto-2-nitrobenzoate acid) ions under native conditions indicate the absence of free thiols and hence the presence of two disulfides (Table S1). The retarded migration of H1HA9 in SDS-PAGE owing to the loss of a compact conformation upon reduction also indicates the formation of disulfide bonds (Fig. S4A). Non-native, inter-molecular disulfide bonds can lead to the formation of high molecular weight aggregates. Therefore, we probed the oligomeric state of H1HA9 by analytical gel-filtration chromatography under native conditions. H1HA9 eluted exclusively as a monomer in solution and no high molecular weight aggregates were detected. The oligomerization of full-length HA is facilitated predominantly by interaction of the long central alpha-helices in the HA stem. The total interface area per monomer in a full-length HA oligomer is 1992.7Å^2. The designed head fragment contributes to only one-fifth (447.2 Å^2) of the total interface area per monomer. H1HA9 is a monomer in solution probably because of the relatively weak interaction interface in the head domain (Fig. 3d-e). The conformational state of H1HA9 was further assessed using RP-HPLC. A single elution peak under oxidizing conditions indicates the presence of a homogenous population in solution. Furthermore, a shift in the elution peak upon reduction suggests the occurrence of disulfide bonds in H1HA9 under native conditions (Fig. 3f).

3.4 Rational Protein Minimization Leads to Minimal Perturbation of the Antigenic Sites in the HA Head

The canonical 'antigenic sites' in the head domain of HA have been comprehensively mapped. The primary antibody response post virus infection/vaccination is directed against distinct conformation-specific epitopes in the head domain of H1 HA; Sa, Sb, Ca (Ca_1 and Ca_2) and Cb. The Sa and Sb sites are located proximal to the RBS, while Ca and Cb sites are located away from the RBS [12]. The Sb site in part comprises residues lining the pocket of the RBS. We evaluated the binding of H1HA9 to monoclonal antibodies targeting distinct antigenic sites in the head domain to assess the structural integrity of the designed head fragment and to confirm that the engineered mutations do not abolish nAb binding. A panel of previously characterized antibodies; H2-6A1, H36-4 and H2-4C2 targeting the Sa, Sb and Cb antigenic sites respectively of H1N1 A/Puerto Rico/8/34 HA were used to profile the antigenic features of H1HA9 [52–54] using ELISA. The head fragment bound antibodies H2-6A1 and H36-4 targeting the Sa and Sb sites respectively with apparent affinities comparable to full-length H1N1 A/Puerto Rico/8/34 HA, validating our design protocol in achieving RBS mimicry (Fig. 4a, b). H1HA9 also bound H2-4C2, the Cb site targeting monoclonal antibody, although with lower affinity than full-length H1N1 A/Puerto Rico/8/34 HA (Fig. 4c). We further evaluated the antigenicity of H1HA9 by measuring reactivity to the anti-H1N1 A/Puerto Rico/8/34 virus convalescent sera. A sub-lethal dose of the H1N1 A/Puerto Rico/8/34 virus elicits a protective polyclonal antibody response. The measurable binding of H1HA9 to the convalescent sera is a crude indicator of the possibility of the designed head fragment to elicit a functional antibody response (Fig. 4d).

The binding of H1HA9 and $H1HA9^{WT}$ to H36-4 was also probed using bio-layer interferometry (BLI) to examine if the engineered mutations facilitated better binding in addition to improved solubility. Consistent with ELISA results, H1HA9 bound H36-4 with high affinity (33.8 nM) (Fig. 5a). In contrast, $H1HA9^{WT}$ showed significantly weaker binding to H36-4 (142.1 nM) owing to a ~ 15-fold lower on-rate (Fig. 5b), which is likely to be a consequence of incorrect RBS conformation and/or protein aggregation leading to steric occlusion of the antibody epitopes.

3.5 Multiple Sequence Alignment Information Guided Head Fragment Design

In order to evaluate the general applicability of our design framework, we attempted to generate a head-fragment from a heterologous influenza A H1N1 HA. Previous studies have attempted to address the growing disparity between the available HA sequences and experimentally determined high-resolution structures. The hemagglutinin structure prediction (HASP) server provides a convenient platform to initially screen HA models generated from sequence information, however, it adopts a conservative approach that omits insertions/deletions (indels) from the models [55]. The modelling of indels is non-trivial [56] and such unpredictable residue insertions/deletions within the head domain can present a challenge in fragment design. Therefore, in addition to examining HA models we incorporated multiple sequence alignment information from an exhaustive dataset

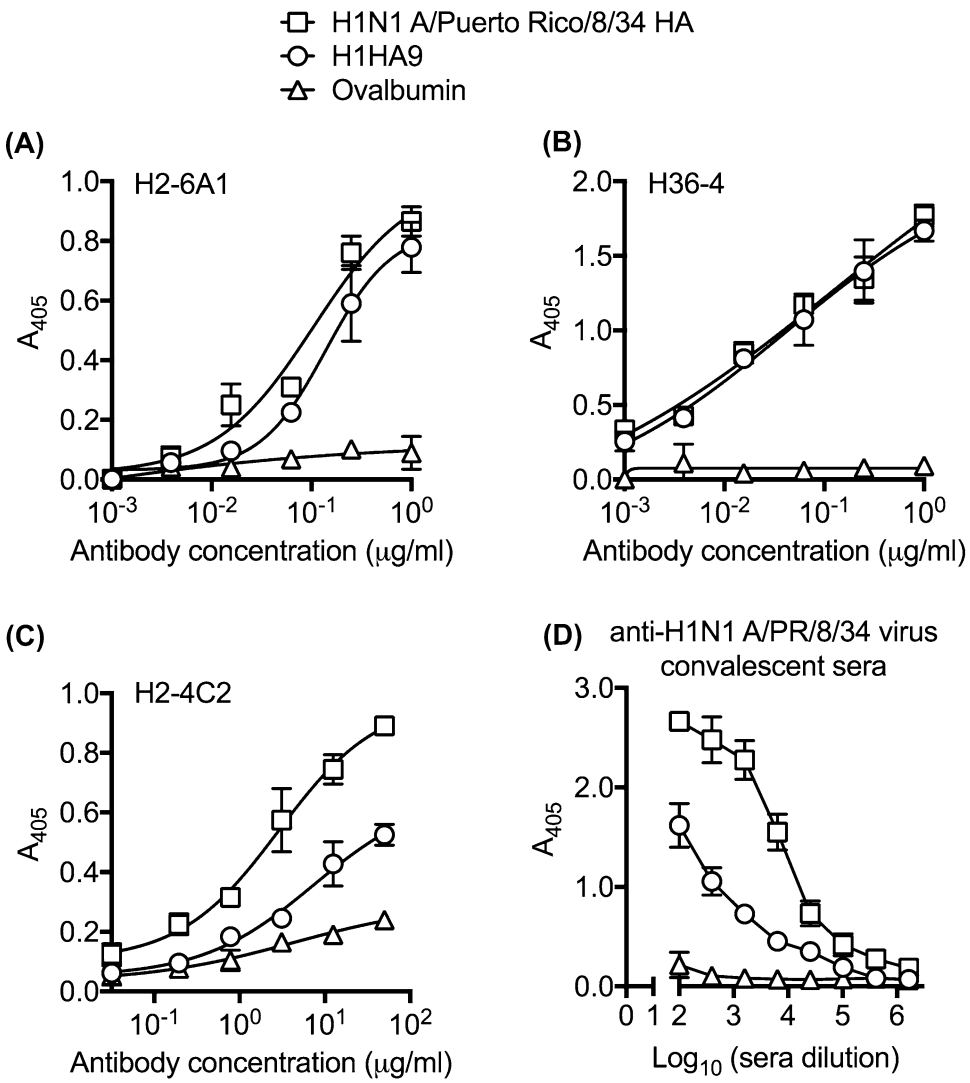

Fig. 4 Mutations engineered in H1HA9 do not perturb antibody binding. The binding of full-length H1N1 A/Puerto Rico/8/34 HA and H1HA9 to a panel of previously characterized antibodies; **a** H2-6A1 [54], **b** H36-4 [53], and **c** H2-4C2 [52] was evaluated by ELISA. The designed head fragment bound RBS targeting antibodies H2-6A1 (Sa) and H36-4 (Sb) with affinities comparable to full-length H1N1 A/Puerto Rico/8/34 HA. H1HA9 also bound the Cb site targeting antibody, H2-4C2, albeit with lower affinity than full-length HA. **d** H1HA9 also bound the anti-H1N1 A/Puerto Rico/8/34 virus convalescent sera. The measurable binding of H1HA9 to the sera suggests that the designed head fragment may conversely elicit a response with specificities that recognize the virus. The reported values are the mean ± SD of three independent experiments. Ovalbumin was used as a control for non-specific antibody binding in all the experiments

(n = 5617) to fix the construct termini in putative non-secondary structural elements (loops) based on the rationale that residues in loops that are not functionally critical are variable. We calculated the average conservation score over a sliding window with incremental lengths (n = 1–10) to identify a robust 'variance' signature (Fig. 6a). Robustness of the signature increases with the number of elements (residues), implying that a contiguous stretch with low average conservation score will probably be exposed. We analyzed sequence strings proximal to the putative termini and identified residue (n = 6) stretches; 47–52 and 271–276 with the lowest average conservation scores to flag the termini of HA head fragment (Fig. 6a). We suggest that any residue in this stretch can be fixed as the head fragment termini since they most likely will form unstructured loops, as validated by an evaluation of available H1 HA structures (Fig. 6b, c). Furthermore, the mutations (6/7) engineered in H1HA9 to prevent protein aggregation and enhance solubility were introduced at positions with a conservation score of > 85%. Hence, these mutations can be readily adopted in context of a heterologous influenza A H1N1 strain.

We generated a head fragment, pH1HA9 (Fig. S3), from the heterologous influenza A H1N1 HA (pH1N1 A/California/07/2009). The sequence identity between full-length H1N1 A/Puerto Rico/8/34 and pH1N1 A/California/07/2009 HAs is 81.3%, which further decreases to 74.4% within the HA1 subunits. The mutations introduced in pH1HA9 to resurface the exposed hydrophobic patches were adapted from H1HA9 following a sequence alignment. A mutation homologous to V205T in H1HA9 was not introduced in pH1HA9, because the wild-type residue at the equivalent position in pH1N1 A/California/07/2009 is glycine, an amino acid with intermediate hydrophobicity index and a high degree of conformational freedom. During protein minimization we typically preclude wild-type glycine residues from substitution. Another homologous mutation (I244Q) was not introduced, because the wild-type residue in pH1N1 A/California/07/2009 at the equivalent position

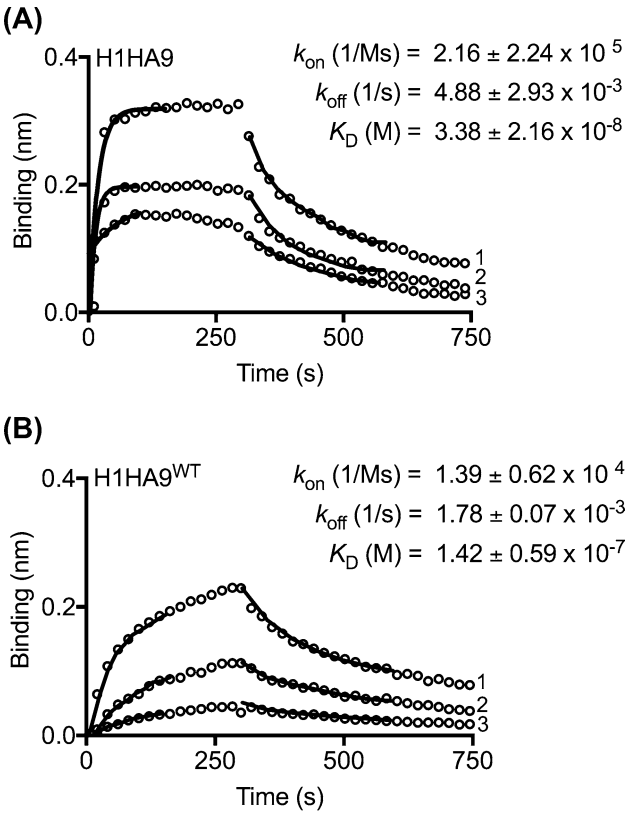

Fig. 5 The engineered mutations improve antibody binding. The kinetic parameters for binding of the head fragments to the Sb-site directed antibody (H36-4) were determined by BLI using an Octet RED96 instrument. The antibody was captured on Protein G (ProG) biosensors. A concentration series (trace 1–3: 500 nM, 250 nM and 125 nM) of the analytes; **a** H1HA9 and **b** $H1HA9^{WT}$ were used. The kinetic parameters were obtained by fitting the data to a 1:1 Langmuir model. The data points are represented by open circles, and the fits are shown by solid lines

is threonine, a polar amino acid. Therefore, given that mutations critical to resurface the exposed hydrophobic patches were introduced only at well conserved positions outside the RBS and canonical antigenic sites, the design protocol described here can be readily translated to divergent influenza A H1N1 strains.

3.6 Protein Purification and Characterization of Head Fragment Designed from pH1N1 A/California/07/2009

The HA head fragment, pH1HA9, designed from pH1N1 A/California/07/2009 was expressed in *E.coli* and purified from the soluble fraction of the culture lysate using a single-step affinity (Ni–NTA) chromatography. However, the yield for pH1HA9 was lower than H1HA9 under similar expression and purification conditions. The yield for pH1HA9 was about 0.5-2 mg/liter culture. The purified protein was characterized using multiple biophysical and biochemical techniques. The head fragment, pH1HA9, displayed properties of a well-folded protein, comparable to H1HA9.The CD spectrum of pH1HA9 was representative of a β-sheet rich protein and exhibited a thermal transition mid-point (T_m) at 318.6 K (45.4 °C) (Fig. 7a, b). The intrinsic tryptophan fluorescence emission maxima of pH1HA9 displayed a red-shift upon denaturation with GdmCl, indicating a well-packed conformation (Fig. 7c). The formation of disulfide bonds in pH1HA9 was confirmed using diverse assays. Labelling of pH1HA9 with a thiol modifying reagent (DTNB) could be achieved only under reducing conditions, indicating the formation of two disulfide bonds under native condition (Tables S1). Further, the migration of pH1HA9 reduced with DTT in SDS-PAGE is retarded suggesting a loss in compaction owing to breakage of disulfide bonds (Fig. S4B). The oligomeric state of pH1HA9 in solution was assessed by analytical gel-filtration chromatography under native conditions. The elution peak of pH1HA9 coincides with H1HA9, corresponding to the molecular weight of a monomer (Fig. 7d, e). The existence of a homogenous, monomeric population of pH1HA9 in solution was further confirmed by RP-HPLC (Fig. 7f).

3.7 Antibody Binding Substantiates Head Domain Mimicry by pH1HA9

The structural integrity of pH1HA9 was probed using two previously characterized conformation-specific nAbs, 2D1 and MA-2077, that bind proximal to the RBS and inhibit hemagglutination [57, 58]. Encouragingly, pH1HA9 bound both 2D1 and MA-2077 (Fig. 8a, b). However, binding of the head fragment to the RBS targeting nAbs was weaker than full-length pH1N1 A/California/04/2009 HA. We hypothesize that an increased functional affinity of pH1HA9 to the antibodies can be achieved by multimerization of the designed head fragment similar to oligomeric full-length HA which maybe essential to attain complete antigenicity [59, 60]. The head fragment, pH1HA9, also showed measurable binding to the anti-pH1N1 A/California/07/2009 virus convalescent sera (Fig. 8c). The binding of H1HA9 to the convalescent sera was slightly weaker than that of recombinant, soluble HA. This could be in part due to avidity effects with the latter protein. The binding studies in congruence with the biophysical and biochemical characterization of the head fragment, validate our design framework and demonstrates its utility.

4 Conclusions

We report a design methodology that integrates structural and sequence conservation features of influenza HA to rapidly generate head fragment mimics from unmatched

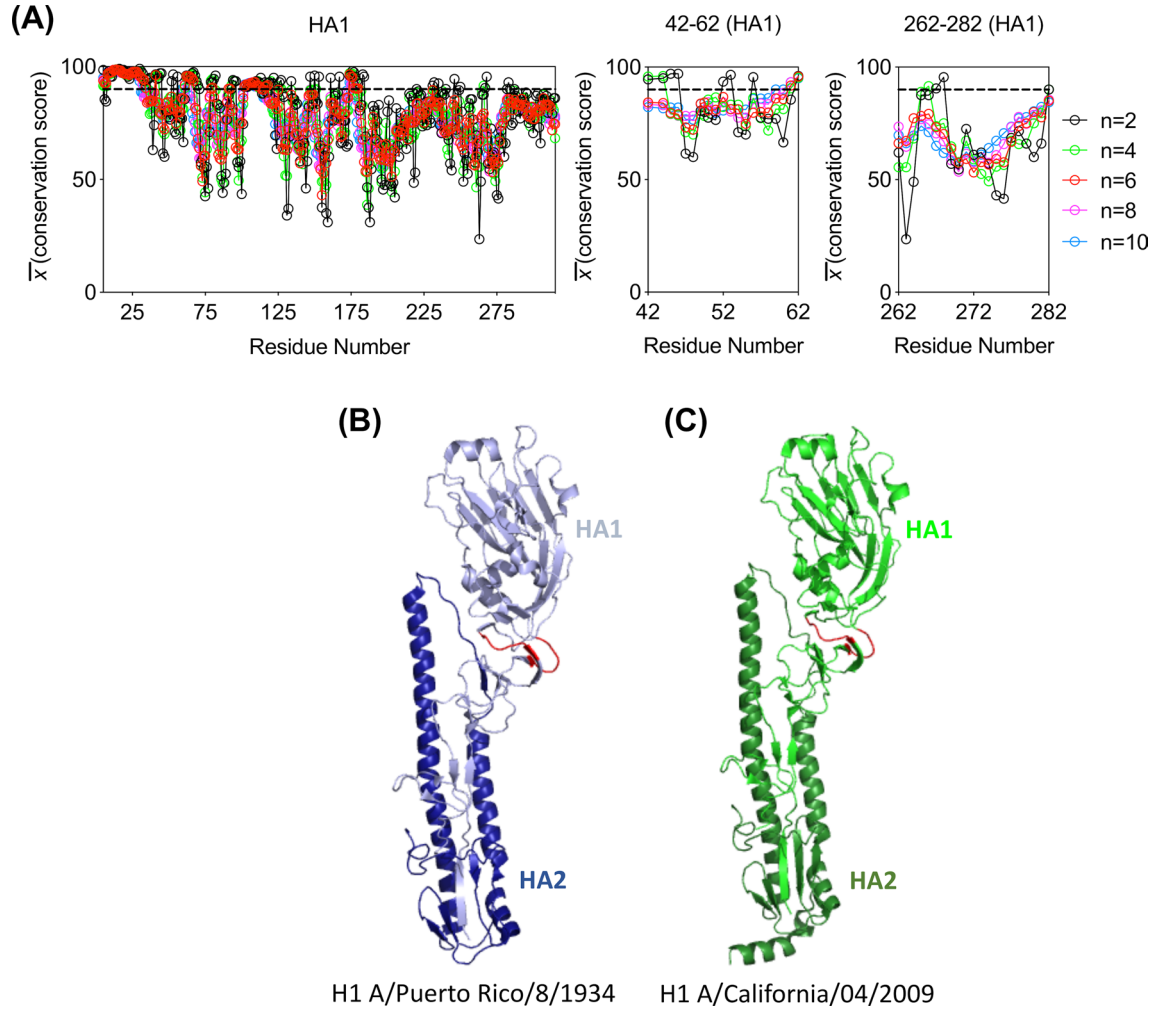

Fig. 6 Sequence features guided head fragment design from heterologous influenza strains. **a** A robust 'signature' to unambiguously fix the construct termini exclusively from sequence information was identified by calculating the average conservation score across all available influenza A H1N1 sequences (n = 5617) over a sliding window with incremental lengths (n = 1–10). The six-residue string (47–52 and 271–276) provides optimum sensitivity and specificity. The identified residue stretches with low average conservation score are suited for fragment termination since they will most likely form unstructured loops as validated by an evaluation of the equivalent positions in available HA structures. Representative structures are shown here; **b** H1 A/Puerto Rico/8/1934 (PDB ID: 1RU7) and **c** H1 A/California/04/2009 (PDB ID: 3LZG). The figure was rendered using PyMOL. The loop regions encompassing the chosen chain termini are shown in red in b and c (Color figure online)

influenza A H1N1 strains. In addition to structure-based interaction network determination, the identification of fundamental design characteristics, viz., the foundation residues, positions to introduce rational mutations and construct termini by incorporating multiple sequence alignment information facilitates the broad applicability of our protocol. X-ray crystal structures while invaluable in informing vaccine design, provide a static snapshot of protein conformation. The lack of protein dynamics, and "biological" water that influence protein conformation [61, 62] may limit a robust description of the minimal interaction network. In part, we have attempted to circumvent an overreliance on structural information to determine positional restraints of the designed HA-fragment by using sequence conservation information. We tested our algorithm by generating head fragments from heterologous influenza A strains, H1N1 A/Puerto Rico/8/34 and pH1N1 A/California/07/2009 HAs, which share a sequence identity of only 74.4% within the head domain.

Prior studies attempting to express HA head fragments in *E.coli* entailed protein recovery from inclusion bodies, often under controlled redox refolding conditions [46–48]. In some cases, the proteins could be produced only using engineered *E.coli* expression strains [46]. We countered these limitations by optimizing the nucleotide sequences encoding the head fragments for bacterial expression.

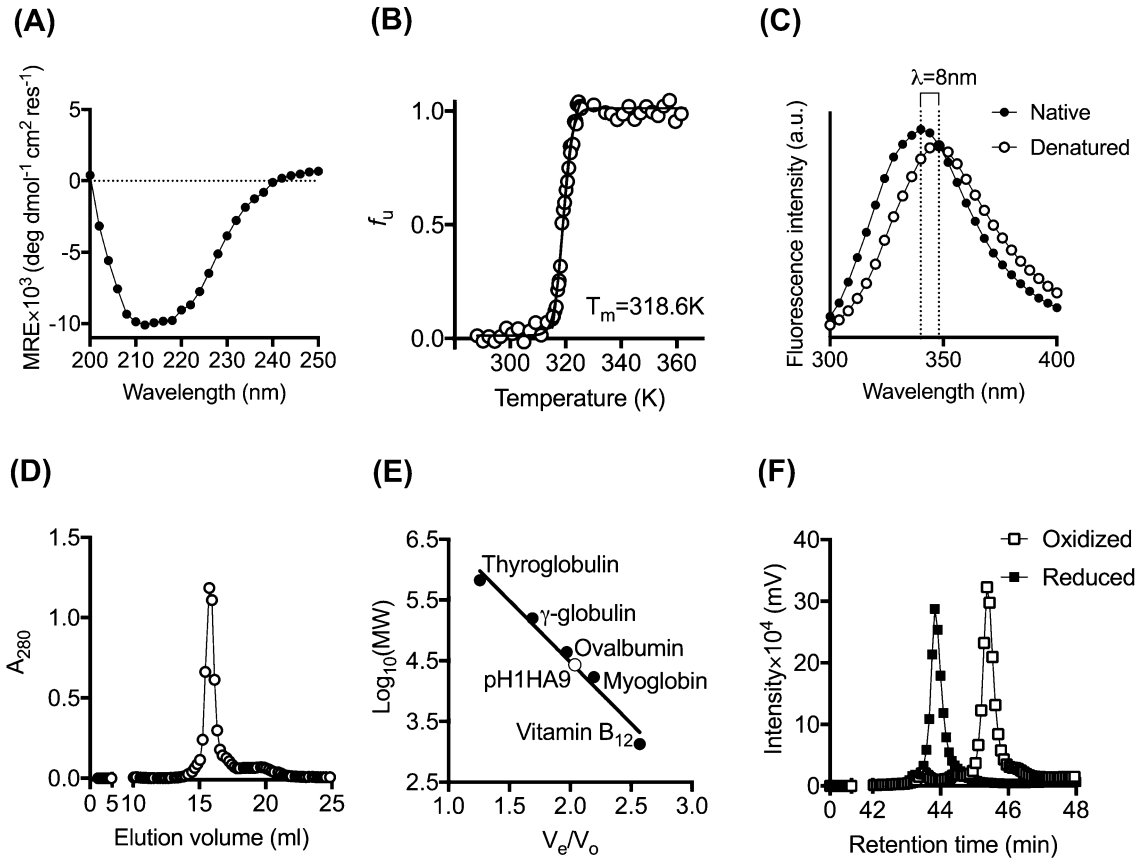

Fig.7 Characterization of pH1HA9; a head fragment designed from pH1N1 A/California/07/2009. **a** The CD spectra of pH1HA9 in the far-UV region (200-250 nm) was characteristic of a well-folded protein. The spectra were recorded at 25 °C and averaged over five consecutive scans. The protein was in PBS (pH 7.4). **b** Thermal stability of pH1HA9 was monitored using CD (λ = 213 nm). The thermal unfolding of pH1HA9 was co-operative with a transition mid-point (T_m) of 318.6 K (45.4 °C). **c** The intrinsic tryptophan fluorescence emission of pH1HA9 was recorded under native (PBS, pH 7.4) or denaturing conditions (7 M GdmCl in PBS, pH 7.4) as indicated. The red-shift in the emission maxima of pH1HA9 upon denaturation with GdmCl suggests a compact conformation with a buried hydrophobic core. Signal averaged protein spectra after correcting for buffer signals are shown. An excitation wavelength (λ_{ex}) of 280 nm was used and fluorescence emission (λ_{em}) was monitored between 300-400 nm. **d** The oligomeric state of pH1HA9 in solution was determined by analytical gel-filtration chromatography under non-denaturing conditions (PBS, pH 7.4) using a Superdex-200 column. pH1HA9 (27.2 kDa) eluted as a homogeneous monomer. **e** The column was calibrated using a broad range of molecular weight standards (•). The data-point corresponding to the elution volume of pH1HA9 is shown by an open circle (o). **f** RP-HPLC was also used to probe the conformational homogeneity of pH1HA9. The oxidized (PBS, pH 7.4) and reduced (denatured with 6 M GdmCl prior to incubation with 5 mM DTT in PBS, pH 7.4) forms of pH1HA9 were analyzed at 25 °C. The protein was eluted using a gradient of acetonitrile (5–95%) increasing at a rate of 1% per minute. A single elution peak corresponding to the oxidized protein which is shifted upon reduction with DTT, indicates the presence of a homogenous, disulfide-bonded species in solution

Lessened aggregation of the designed fragments was achieved by introducing rational mutations to resurface the newly exposed hydrophobic surfaces resulting from protein minimization. Mapping the interaction network of foundation residues facilitated incorporation of rational mutations while maintaining structural integrity of the head domain. The biophysical and biochemical properties of the designed head fragments were representative of a well-folded protein molecule. Monoclonal antibody binding further confirms the native-like conformation of the head fragments. A detailed comparison of the biophysical properties and antibody binding profiles of the engineered proteins reported in this study with previously described head fragments cannot be made because of unavailability of similar data for the latter fragments.

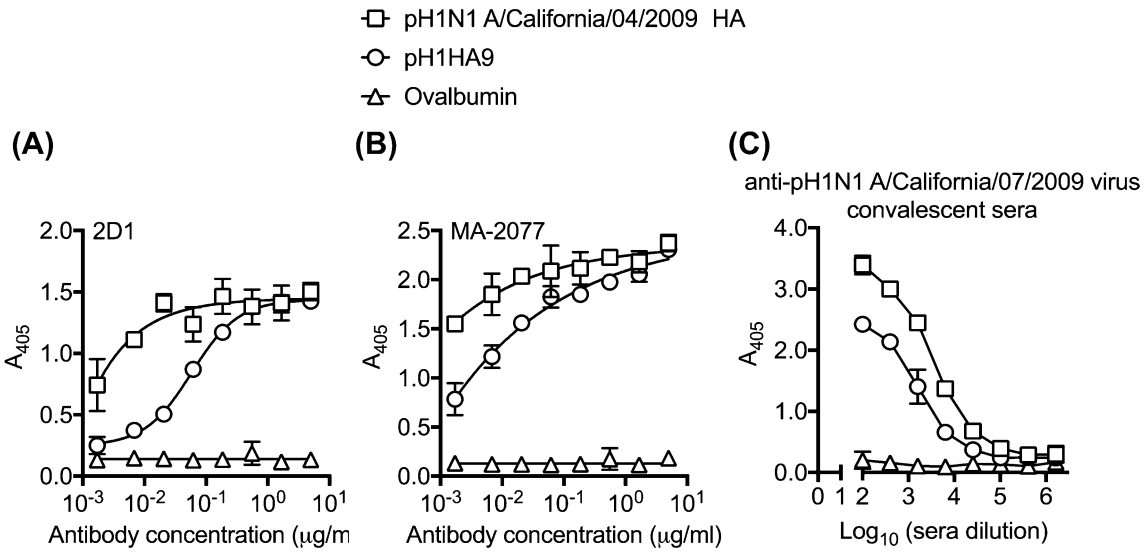

Fig. 8 Head domain mimicry by pH1HA9 facilitates antibody binding. The binding of full-length pH1N1 A/California/07/2009 HA and pH1HA9 to conformation specific RBS targeting nAbs; **a** 2D1 [58], **b** MA-2077 [57] and **c** anti-H1N1 A/California/07/2009 virus convalescent sera was evaluated by ELISA. The reported values are the mean ± SD of three independent experiments. Ovalbumin was used a control for non-specific antibody binding in all the experiments

Seasonal vaccination typically fails to robustly elicit/boost cross-reactive, neutralizing antibodies and alternative strategies are warranted to target the conserved epitopes of HA [16]. Alternatively, epitope-focused immunogens developed by protein minimization strategies have shown potential to elicit a humoral response with increased breadth against diverse viral pathogens [41, 63–68]. Reactivity with the convalescent sera as shown in this study indicates the possibility of the designed head fragments to elicit neutralizing antibodies similar to those elicited by vaccination or natural infection. In future studies, we intend to test the efficacy of the head fragments in vivo as a stand-alone immunogen and also in combination with previously tested platforms used to enhance the immunogenicity of head fragments, for instance, the potent TLR5 ligand (flagellin) as a fusion partner [47] or conjugation with bacteriophage-derived Q_β virus-like particles [69] or other nanoparticle display formats.

Our study describes a widely applicable structure-based design protocol aided by sequence conservation information to mimic the HA head domain and express it in bacteria. The study provides a framework that can be extended to heterologous influenza subtypes. Further, the diagnostic utility of HA head fragments in detection of influenza-specific antibodies post infection will also be explored.

Acknowledgements We thank Dr. James Crowe for providing the 2D1 antibody. The anti-H1N1 A/Puerto Rico/8/34 virus convalescent sera was provided by Dr. Jessica Flynn (Merck Research Laboratories, PA, USA). Dr. Satish Kumar Gupta provided the anti-pH1N1 A/California/07/2009 virus convalescent sera and MA-2077 antibody.

Author Contributions All authors contributed to the study conception and design. Experiments and analysis were performed by V VAM and SS under the supervision of RV. All authors read and approved the final manuscript.

Funding VVAM acknowledges the fellowship received from the Council of Scientific and Industrial Research, Government of India (GOI). SS is recipient of research associate fellowship from Department of Biotechnology (DBT), GOI. This work was supported in part by a grant from the DBT, GOI to RV.

Compliance with Ethical Standards

Conflict of interest The authors declare no conflict of interest.

References

1. Krammer F, Palese P (2015) Advances in the development of influenza virus vaccines. Nat Rev Drug Discov 14(3):167–182
2. Kumar A, Singh S (2015) Editorial: influenza virus vaccines and immunotherapies. Front Immunol 6:599
3. Schultz-Cherry S, Jones JC (2010) Influenza vaccines: the good, the bad, and the eggs. Adv Virus Res 77:63–84
4. Carrat F, Flahault A (2007) Influenza vaccine: the challenge of antigenic drift. Vaccine 25(39–40):6852–6862
5. Flannery B et al (2015) Early estimates of seasonal influenza vaccine effectiveness–United States, January 2015. MMWR Morb Mortal Wkly Rep 64(1):10–15
6. Gerhard W (2001) The role of the antibody response in influenza virus infection. Curr Top Microbiol Immunol 260:171–190
7. Shi Y et al (2014) Enabling the "host jump": structural determinants of receptor-binding specificity in influenza A viruses. Nat Rev Microbiol 12(12):822–831

8. Skehel JJ, Wiley DC (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem 69:531–569
9. Xiong X, McCauley JW, Steinhauer DA (2014) Receptor binding properties of the influenza virus hemagglutinin as a determinant of host range. Curr Top Microbiol Immunol 385:63–91
10. Carr CM, Kim PS (1993) A spring-loaded mechanism for the conformational change of influenza hemagglutinin. Cell 73(4):823–832
11. Braakman I et al (1991) Folding of influenza hemagglutinin in the endoplasmic reticulum. J Cell Biol 114(3):401–411
12. Caton AJ et al (1982) The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31(2 Pt 1):417–427
13. Wiley DC, Wilson IA, Skehel JJ (1981) Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature 289(5796):373–378
14. Neumann G, Kawaoka Y (2015) Transmission of influenza A viruses. Virology 479–480:234–246
15. Laursen NS, Wilson IA (2013) Broadly neutralizing antibodies against influenza viruses. Antiviral Res 98(3):476–483
16. Rathore U et al (2014) Immunogen design for HIV-1 and influenza. Biochim Biophys Acta 1844(11):1891–1906
17. Dreyfus C et al (2012) Highly conserved protective epitopes on influenza B viruses. Science 337(6100):1343–1348
18. Ekiert DC et al (2012) Cross-neutralization of influenza A viruses mediated by a single antibody loop. Nature 489(7417):526–532
19. Wine Y et al (2015) Serology in the 21st century: the molecular-level analysis of the serum antibody repertoire. Curr Opin Immunol 35:89–97
20. Corti D, Lanzavecchia A (2013) Broadly neutralizing antiviral antibodies. Annu Rev Immunol 31:705–742
21. Van, T.D., et al., *A Perspective on Rational Designs of a Hemagglutinin Based Universal Influenza Vaccine*. Curr Pharm Des, 2016.
22. Ellebedy AH, Ahmed R (2012) Re-engaging cross-reactive memory B cells: the influenza puzzle. Front Immunol 3:53
23. Andreatta RH, Liem RK, Scheraga HA (1971) Mechanism of action of thrombin on fibrinogen. I. Synthesis of fibrinogen-like peptides, and their proteolysis by thrombin and trypsin. Proc Natl Acad Sci U S A 68(2):253–256
24. Anfinsen CB, Scheraga HA (1975) Experimental and theoretical aspects of protein folding. Adv Protein Chem 29:205–300
25. Dyson HJ, Wright PE, Scheraga HA (2006) The role of hydrophobic interactions in initiation and propagation of protein folding. Proc Natl Acad Sci U S A 103(35):13057–13061
26. Huang Y et al (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics 26(5):680–682
27. Higgins DG, Sharp PM (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73(1):237–244
28. Gamblin SJ et al (2004) The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science 303(5665):1838–1842
29. Ganesh C et al (1997) Thermodynamic characterization of the reversible, two-state unfolding of maltose binding protein, a large two-domain protein. Biochemistry 36(16):5020–5028
30. Ellman GL (1959) Tissue sulfhydryl groups. Arch Biochem Biophys 82(1):70–77
31. Riddles PW, Blakeley RL, Zerner B (1979) Ellman's reagent: 5,5'-dithiobis(2-nitrobenzoic acid)–a reexamination. Anal Biochem 94(1):75–81
32. Russell RJ et al (2006) Avian and human receptor binding by hemagglutinins of influenza A viruses. Glycoconj J 23(1–2):85–92
33. Stencel-Baerenwald JE et al (2014) The sweet spot: defining virus-sialic acid interactions. Nat Rev Microbiol 12(11):739–749
34. Lutteke T, Frank M, von der Lieth CW (2005) Carbohydrate Structure Suite (CSS): analysis of carbohydrate 3D structures derived from the PDB. Nucleic Acids Res 33:D242–D246
35. Sabesan S, Bock K, Paulson JC (1991) Conformational analysis of sialyloligosaccharides. Carbohydr Res 218:27–54
36. Hong M et al (2013) Antibody recognition of the pandemic H1N1 Influenza virus hemagglutinin receptor binding site. J Virol 87(22):12471–12480
37. Whittle JR et al (2011) Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin. Proc Natl Acad Sci U S A 108(34):14216–14221
38. Sharma D et al (2005) Protein minimization of the gp120 binding region of human CD4. Biochemistry 44(49):16192–16202
39. Li LK, Riehm JP, Scheraga HA (1966) Structural studies of ribonuclease 23 Pairing of the tyrosyl and carboxyl groups. Biochemistry 5(6):2043–2048
40. Kuhlman B et al (2003) Design of a novel globular protein fold with atomic-level accuracy. Science 302(5649):1364–1368
41. Mallajosyula VV et al (2014) Influenza hemagglutinin stem-fragment immunogen elicits broadly neutralizing antibodies and confers heterologous protection. Proc Natl Acad Sci U S A 111(25):E2514–E2523
42. Mallajosyula VV et al (2015) Hemagglutinin Sequence Conservation Guided Stem Immunogen Design from Influenza A H3 Subtype. Front Immunol 6:329
43. Mallajosyula VV et al (2013) In vitro and in vivo characterization of designed immunogens derived from the CD-helix of the stem of influenza hemagglutinin. Proteins 81(10):1759–1775
44. Valkenburg SA et al (2016) Stalking influenza by vaccination with pre-fusion headless HA mini-stem. Sci Rep 6:22666
45. Jia B, Jeon CO (2016) High-throughput recombinant protein expression in Escherichia coli: current status and future perspectives. Open Biol 6(8):160196
46. Aguilar-Yanez JM et al (2010) An influenza A/H1N1/2009 hemagglutinin vaccine produced in Escherichia coli. PLoS ONE 5(7):e11694
47. Song L et al (2008) Efficacious recombinant influenza vaccines produced by high yield bacterial expression: a solution to global pandemic and seasonal needs. PLoS ONE 3(5):e2257
48. Xu R et al (2012) Structural characterization of the hemagglutinin receptor specificity from the 2009 H1N1 influenza pandemic. J Virol 86(2):982–990
49. Hao MH, Scheraga HA (1998) Molecular mechanisms for cooperative folding of proteins. J Mol Biol 277(4):973–983
50. Gahl RF, Scheraga HA (2009) Oxidative folding pathway of onconase, a ribonuclease homologue: insight into oxidative folding mechanisms from a study of two homologues. Biochemistry 48(12):2740–2751
51. Houry WA, Rothwarf DM, Scheraga HA (1994) A very fast phase in the refolding of disulfide-intact ribonuclease A: implications for the refolding and unfolding pathways. Biochemistry 33(9):2516–2530
52. Kavaler J et al (1990) A set of closely related antibodies dominates the primary antibody response to the antigenic site CB of the A/PR/8/34 influenza virus hemagglutinin. J Immunol 145(7):2312–2321
53. Staudt LM, Gerhard W (1983) Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin I Significant variation in repertoire expression between individual mice. J Exp Med 157(2):687–704
54. Yewdell JW, Caton AJ, Gerhard W (1986) Selection of influenza A virus adsorptive mutants by growth in the presence of a mixture of monoclonal antihemagglutinin antibodies. J Virol 57(2):623–628

55. Ambroggio XI et al (2013) HASP server: a database and structural visualization platform for comparative models of influenza A hemagglutinin proteins. BMC Bioinformatics 14:197
56. Kim R, Guo JT (2010) Systematic analysis of short internal indels and their impact on protein folding. BMC Struct Biol 10:24
57. Shembekar N et al (2013) Isolation of a high affinity neutralizing monoclonal antibody against 2009 pandemic H1N1 virus that binds at the "Sa" antigenic site. PLoS ONE 8(1):e55516
58. Xu R et al (2010) Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. Science 328(5976):357–360
59. Khurana S et al (2011) Recombinant HA1 produced in E coli forms functional oligomers and generates strain-specific SRID potency antibodies for pandemic influenza vaccines. Vaccine 29(34):5657–5665
60. Magadan JG et al (2013) Influenza A virus hemagglutinin trimerization completes monomer folding and antigenicity. J Virol 87(17):9742–9753
61. Bagchi B (2016) Untangling complex dynamics of biological water at protein-water interface. Proc Natl Acad Sci U S A 113(30):8355–8357
62. Wertz DH, Scheraga HA (1978) Influence of water on protein structur An analysis of the preferences of amino acid residues for the inside or outside and for specific conformations in a protein molecule. Macromolecules 11(1):9–15
63. Bhattacharyya S et al (2013) Design of an Escherichia coli expressed HIV-1 gp120 fragment immunogen that binds to b12 and induces broad and potent neutralizing antibodies. J Biol Chem 288(14):9815–9825
64. Impagliazzo A et al (2015) A stable trimeric influenza hemagglutinin stem as a broadly protective immunogen. Science 349(6254):1301–1306
65. Jardine J et al (2013) Rational HIV immunogen design to target specific germline B cell receptors. Science 340(6133):711–716
66. Joyce MG et al (2016) Iterative structure-based improvement of a fusion-glycoprotein vaccine against RSV. Nat Struct Mol Biol 23(9):811–820
67. Kanekiyo M et al (2015) Rational Design of an Epstein-Barr Virus Vaccine Targeting the Receptor-Binding Site. Cell 162(5):1090–1100
68. Yassine HM et al (2015) Hemagglutinin-stem nanoparticles generate heterosubtypic influenza protection. Nat Med 21(9):1065–1070
69. Jegerlehner A et al (2013) Bacterially produced recombinant influenza vaccines based on virus-like particles. PLoS ONE 8(11):e78947