

# Glycosylation at Asn<sup>91</sup> of H1N1 haemagglutinin affects binding to glycan receptors

Akila JAYARAMAN, Xiaoying KOH, Jing LI, Rahul RAMAN, Karthik VISWANATHAN, Zachary SHRIVER and Ram SASISEKHARAN<sup>1</sup>

Harvard–MIT Division of Health Sciences and Technology, Koch Institute for Integrative Cancer Research, Singapore–MIT Alliance for Research and Technology, Department of Biological Engineering, Massachusetts Institute of Technology (MIT), 77 Massachusetts Avenue, Cambridge, MA 02139, U.S.A.

The glycoprotein HA (haemagglutinin) on the surface of influenza A virus plays a central role in recognition and binding to specific host cell-surface glycan receptors and in fusion of viral membrane to the host nuclear membrane during viral replication. Given the abundance of HA on the viral surface, this protein is also the primary target for host innate and adaptive immune responses. Although addition of glycosylation sites on HA are a part of viral evolution to evade the host immune responses, there are specific

glycosylation sites that are conserved during most of the evolution of the virus. In the present study, it was demonstrated that one such conserved glycosylation site at Asn<sup>91</sup> in H1N1 HA critically governs the glycan receptor-binding specificity and hence would potentially impinge on the host adaptation of the virus.

**Key words:** glycan receptor, glycosylation, haemagglutinin, H1N1, influenza virus.

## INTRODUCTION

The specific binding of HA (haemagglutinin), a glycoprotein on the surface of influenza A virus, to glycan receptors on the host cell surface is one of the critical factors that govern host infection and adaptation of the virus [1–3]. HA is a homotrimeric transmembrane protein where each monomer has an ectodomain consisting of a globular head, which harbours the glycan RBS (receptor-binding site), and a stem region [4]. The host glycan receptors for HA are sialylated glycans [complex glycans terminated by sialic acid such as Neu5Ac (*N*-acetylneuraminic acid)]. Glycans terminated by Neu5Ac that is  $\alpha 2 \rightarrow 6$ -linked to the penultimate sugar are predominantly expressed in human upper respiratory epithelia (henceforth referred to as human receptors) and serve as receptors for human-adapted influenza A viruses [5,6]. On the other hand, glycans terminating in Neu5Ac that is  $\alpha 2 \rightarrow 3$ -linked to the penultimate sugar residue serve as receptors for the avian-adapted influenza viruses (henceforth referred to as avian receptors) [7]. Given the location of HA on the viral surface, this protein is also the primary antigen for the host immune response. Selection pressure from innate and adaptive host immunity drives HA to acquire mutations in order for the virus to escape host defences. These mutations are predominantly in the antigenic sites on HA that are proximal to the RBS (since interfering with receptor binding is the primary target for neutralizing antibodies). Therefore the mutations also impinge on the glycan receptor-binding properties of HA [8].

In addition to acquiring mutations that change antigenic sites on HA, influenza viruses acquire mutations that add or remove glycosylation sites on HA. Amino acid substitutions that lead to the sequence motifs (or sequons) required for glycosylation in the globular head domain of HA are believed to efficiently generate antigenic variants [9]. The complex glycans that are linked to the glycosylation sites are involved in masking the antigenic epitopes and have also been shown to play critical roles in maintaining the structural stability and modulating function of HA [10–12].

The glycans near the cleavage site and in the stem region of HA modulate proteolytic activation and maintain HA in the metastable form required for fusion activity [10].

Understanding the role of glycosylation on HA on its glycan receptor-binding properties has been an active area of research. Among the different types of glycosylation, N-linked glycosylation involving attachment of complex glycans to an asparagine residue in a consensus sequon (Asn-Xaa-Ser/Thr, where Xaa is any amino acid except proline) has been the focus of many studies.

Molecular dynamics simulation studies predicted that HA glycans may form interactions near the binding pocket to influence receptor binding [11]. Site-directed mutagenesis to knock out glycosylation sites on HA [10] or modifying structure of N-linked glycans on the virus by enzymatic treatment or transgenic cell lines [12] have shown distinct changes in glycan receptor-binding specificity. Site-specific loss-of-glycosylation mutations have also been shown to correlate with sensitivity to the innate immune response in mice [13].

A systematic sequence analysis of H1N1 HAs from 1918 pandemic to the 2009 pandemic strains (Supplementary Table S1 at <http://www.BiochemJ.org/bj/444/bj4440429add.htm>) has revealed that, although there are differences in the number and location of N-linked glycosylation sites in these HAs, there are specific conserved glycosylation sites [14]. In the present study, these conserved glycosylation sites were mapped on to the crystal structure of H1N1 HA (PDB code 1RUZ). Through structural analysis of the molecular contacts between glycosylation sites and the RBS, it was shown that glycosylation at the Asn<sup>91</sup> (numbering based on PDB code 1RUZ) sequon (conserved in most H1N1 HAs) was positioned to interact with the RBS of HA. On the basis of this analysis, it was hypothesized that loss of glycosylation at this sequon would affect receptor-binding properties of H1 HA. To test this hypothesis, three HAs derived from the 1918 pandemic H1N1 virus with distinct glycan receptor specificities were deglycosylated at Asn<sup>91</sup> through site-directed

Abbreviations used: Gal, galactose; GlcNAc, *N*-acetylglucosamine; HA, haemagglutinin; HRP, horseradish peroxidase; LN, type 2 lactosamine repeat unit; MALDI, matrix-assisted laser-desorption ionization; MS/MS, tandem MS; Neu5Ac, *N*-acetylneuraminic acid; RBS, receptor-binding site; S, sialyl group; SIN, Significant Interaction Network; TOF/TOF, tandem time-of-flight.

<sup>1</sup> To whom correspondence should be addressed (email rams@mit.edu).

mutagenesis. The binding properties of the mutant HAs were quantitatively characterized using a dose-dependent direct glycan-binding analysis using a glycan array platform. Through these analyses, it was demonstrated that loss of glycosylation affected human receptor binding of human HA, but not avian receptor binding of avian HA. Furthermore, using MALDI (matrix-assisted laser-desorption ionization)-MS and MS/MS (tandem MS) analyses, it was shown that this sequon predominantly comprises high-mannose-type glycans in the wild-type H1N1 HAs.

## MATERIALS AND METHODS

### HAs used in the present study

The HA from the prototypic 1918 H1N1 pandemic strain (A/South Carolina/1/18 or SC18), a single amino acid D225G mutant of SC18 HA (referred to as NY18) and a two amino acid D190E/D225G mutant of SC18 HA (referred to as AV18) were used.

### Site-directed mutagenesis of HA

Site-directed mutagenesis was carried out with the QuikChange® multi site-directed mutagenesis kit (Stratagene). The primer used for mutagenesis was designed by using the Internet-based program Primer3 (<http://frodo.wi.mit.edu/primer3/>) and synthesized by IDT (Integrated DNA Technologies). The primer sequence used for generating T93A mutants was 5'-GAA-ACATCGAACTCAGAGAATGGAGCATGTTACCCAGGA-GATTCATCG-3'. The sequences of the mutants were confirmed by DNA sequencing at Genewiz (Cambridge, MA, U.S.A.).

### Cloning, baculovirus synthesis, expression and purification of HA

Briefly, recombinant baculoviruses with the wild-type or mutant HA gene respectively were used to infect [MOI (multiplicity of infection) of 1] suspension cultures of Sf9 insect cells (Invitrogen) cultured in BD Baculogold Max-XP SFM (BD Biosciences). The infection was monitored and the conditioned medium was harvested 3–4 days after infection. The soluble HA from the harvested conditioned medium was purified using nickel-affinity chromatography (HisTrap HP columns; GE Healthcare). Eluting fractions containing HA were pooled, concentrated and buffer-exchanged into PBS (pH 8.0) (Gibco) using 100 kDa MWCO (molecular-mass cut-off) spin columns (Millipore). The purified protein was quantified using the BCA (bicinchoninic acid) method (Pierce).

### CD analysis of mutant HAs

CD analysis was performed using the Aviv Model 202 Circular Dichroism Spectrometer. The spectra were generated between 190 and 280 nm (Supplemental Figure S1 at <http://www.BiochemJ.org/bj/444/bj4440429add.htm>). Identical concentrations of the mutant and wild-type proteins were used to avoid any differences in the spectra arising from differences in their concentrations.

### Dose-dependent direct binding of wild-type and mutant HA

To investigate the multivalent HA–glycan interactions, a streptavidin plate array comprising representative biotinylated  $\alpha 2 \rightarrow 3$  and  $\alpha 2 \rightarrow 6$  sialylated glycans was used as described previously [15]. 3'SLN (where S indicates the sialyl group

and LN is a type 2 lactosamine repeat unit), 3'SLN-LN and 3'SLN-LN-LN are representative avian receptors. 6'SLN and 6'SLN-LN are representative human receptors. The biotinylated glycans were obtained from the Consortium for Functional Glycomics through their resource request program. Streptavidin-coated High Binding Capacity 384-well plates (Pierce) were loaded to the full capacity of each well by incubating the well with 50  $\mu$ l of 2.4  $\mu$ M biotinylated glycans overnight at 4 °C. Excess glycans were removed through extensive washing with PBS. The trimeric HA unit comprises three HA monomers (and hence three RBSs, one for each monomer). The spatial arrangement of the biotinylated glycans in the wells of the streptavidin plate array favours binding to only one of the three HA monomers in the trimeric HA unit. Therefore, in order to specifically enhance the multivalency in the HA–glycan interactions, the recombinant HA proteins were pre-complexed with the primary and secondary antibodies at a molar proportion of 4:2:1 (HA/primary antibody/secondary antibody). The identical arrangement of four trimeric HA units in the pre-complex for all of the HAs permits comparison between their glycan-binding affinities. A stock solution containing appropriate amounts of histidine-tagged HA protein, primary antibody (mouse anti-His<sub>6</sub>-IgG from Abcam) and secondary antibody [HRP (horseradish peroxidase)-conjugated goat anti-(mouse IgG) from Santa Cruz Biotechnology] at a proportion of 4:2:1 were incubated on ice for 20 min. Appropriate amounts of pre-complexed stock HA were diluted to 250  $\mu$ l with 1% (w/v) BSA in PBS. Then, 50  $\mu$ l of this pre-complexed HA was added to each of the glycan-coated wells and incubated at room temperature (25 °C) for 3 h followed by the wash steps with PBS and PBST (PBS containing 0.05% Tween 20). The binding signal was determined based on HRP activity using the Amplex Red Peroxidase Assay kit (Invitrogen) according to the manufacturer's instructions.

### Calculation of SIN (Significant Interaction Network) scores of RBS residues in SC18, NY18 and AV18 HA

The SIN of an amino acid in a protein is a metric developed previously by our group [16] to capture the extent of networking of that amino acid in terms of its inter-residue contacts with other amino acids in that protein. Starting from the co-ordinate PDB file of a protein instances of putative hydrogen bonds (including water-bridged ones), disulfide bonds,  $\pi$ -bonds, polar interactions, salt bridges and Van der Waals interactions (non-hydrogen) occurring between pairs of residues using appropriate distance thresholds were computed using protocols implemented in MATLAB (requiring Bioinformatics Toolbox). These data were assembled into an array of eight atomic interaction matrices. A weighted sum of the eight atomic interaction matrices were then computed to produce a single matrix that accounts for the strength of atomic interaction between residue pairs, using weights derived from relative atomic interaction energies and including weights for interchain interactions and long-range over short-range interactions. The resulting inter-residue energetic interaction matrix describes all first-order interactions for the molecular structure analysed. All interaction pathways regardless of length were then calculated to obtain the paths. Using the collection of paths identified (and their corresponding scores), the complete SIN matrix was created, wherein each element  $i, j$  is the sum of the path scores of all paths. The degree of networking score for each residue was computed by summing across the rows of the matrix, which was meant to correspond to the extent of 'networking' for each residue. The degree of networking score was normalized (SIN score) with the maximum score for each

protein so that the scores varied from 0 (absence of any network) to 1 (most networked). The homology models of the trimeric form of SC18, NY18 and AV18 HA were constructed using the Modeller program (<http://salilab.org/modeller/>) by adapting the python script from online documentation to model multiple chains related by symmetry (<http://salilab.org/modeller/manual/node28.html>). The template structure used for the modelling is that of SC18 HA (PDB code 1RUZ). These models were used to compute SIN scores of each amino acid in the respective protein.

### Digestion of H1N1 by Glu-C

In 50 mM ammonium bicarbonate solution, 10  $\mu\text{g}$  of H1N1 was denatured with 6 M urea, and reduced by 25 mM DTT (dithiothreitol) at 60°C for 45 min. The mixture was then cooled to room temperature. Iodoacetamide was added to a final concentration of 55 mM, and the mixture was incubated in the dark at room temperature for 30 min to alkylate the cysteine residue. The mixture was then diluted with 50 mM ammonium bicarbonate solution to reduce the urea concentration to 1 M. Then, 0.5  $\mu\text{g}$  of Glu-C (Sequencing Grade from Promega) was added, and the mixture was incubated at room temperature overnight to digest the H1N1 HA. The digestion reaction was quenched by adding formic acid until the pH fell below 4.

### MALDI-MS and MALDI-TOF/TOF (tandem time-of-flight) MS analysis of glycosylation sites on H1N1

The protein digest was desalted by ZipTip 0.6  $\mu\text{l}$  C<sub>18</sub> resin (Millipore) before MALDI analysis. MS and MS/MS spectra were acquired on an Applied Biosystems 4800 Plus MALDI TOF/TOF Analyzer, equipped with a 355 nm Nd:YAG (neodymium-doped yttrium aluminium garnet) laser. A 5 mg/ml solution of  $\alpha$ -cyano-4-hydroxycinnamic acid in 1:1 (v/v) acetonitrile/water was used as the matrix. On a standard 384-well stainless steel MALDI sample plate, 0.5  $\mu\text{l}$  of the desalted protein digest and 0.5  $\mu\text{l}$  of the matrix solution was spotted in one well. Five duplicates of the above spot were made. One spot was used for MS analysis, and the rest were used for MS/MS analysis. MS and MS/MS spectra were both acquired under positive-ion reflector mode. For fragmentation, 2 kV acceleration voltage and CID (collision-induced dissociation) mode off was applied. External peptide standards were used for calibration.

## RESULTS

### Rationale for choice of HAs used in the present study

SC18 represents a prototypical human-adapted pandemic H1N1 virus. NY18 is a natural variant of SC18 that differs from its parental virus by a single amino acid mutation (D225G) in the RBS of HA. AV18 is a laboratory-generated recombinant virus from SC18 that differs from SC18 by two amino acid mutations (D190E/D225G). The glycan receptor specificity and affinity of the HA from these viruses have been well characterized [15,17]. SC18 HA shows strict specificity and high affinity for human receptors. On the other hand, AV18 HA shows strict specificity and high affinity for avian receptors. NY18 is an intermediate between SC18 and AV18 since it shows a mixed human/avian receptor binding, albeit at substantially lower affinities than SC18 (for the human receptor) and AV18 (for the avian receptor). Given that these viruses differ only in the RBS of HA, they have served as good model strains to link glycan receptor specificity and affinity with other biological properties such as viral transmission [17]. Whereas SC18 with specific high-affinity binding to human

receptors transmitted readily via respiratory droplets in a ferret animal model, AV18 did not show any such transmission. NY18, on the other hand, was able to transmit via respiratory droplets, albeit to a much lesser extent than SC18. Given the diversity of glycan-binding preferences of SC18, NY18 and AV18 HAs and differences in transmissibility of the viruses, these HAs served as ideal model systems to study the effect of HA glycosylation on HA-glycan interactions. These HAs were expressed in their soluble form (lacking transmembrane region) with a His<sub>6</sub> tag as described previously [18].

### Structural framework for identifying conserved glycosylation sites in H1N1 HA that potentially affects glycan receptor binding

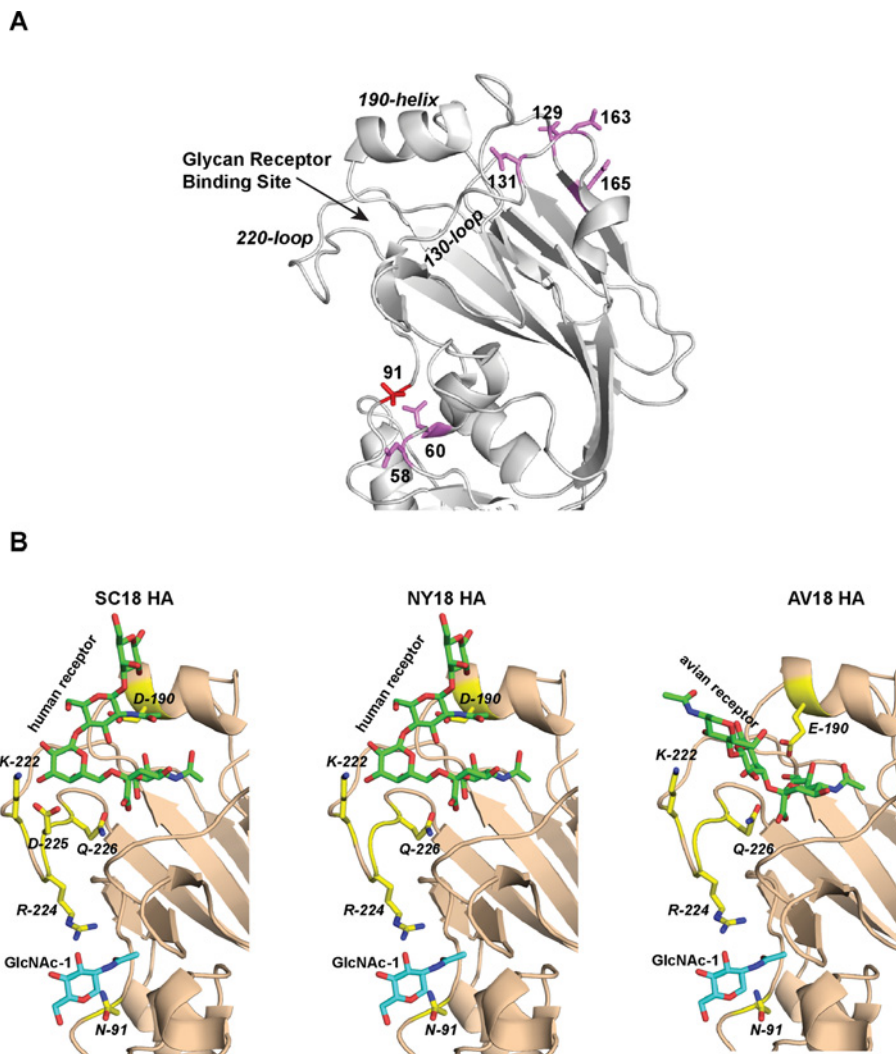
Building on a previous systematic analysis of sequons in H1N1 HAs from 1918 to the present [14] (Supplementary Table S1), the sequon positions in the side and top of the head region of HA were mapped on to a representative H1N1 crystal structure (PDB code 1RUZ). Analysis of these sequon positions on the crystal structure revealed that glycosylation at two specific residues, Asn<sup>91</sup> and Asn<sup>131</sup>, were positioned to interact with the 220-loop and 130-loop regions of the RBS respectively (Figure 1A). Among these two specific sequons, Asn<sup>91</sup> was observed in most of the H1N1 HAs (except those isolated between 1933 and 1940), including the recent strains after the 2009 pandemic, whereas Asn<sup>131</sup> was observed only in H1N1 HAs in 1934–1985 isolates. Therefore the effect of glycosylation at Asn<sup>91</sup> on glycan receptor binding of SC18, NY18 and AV18 HAs was investigated further.

The co-ordinates of the core GlcNAc (*N*-acetylglucosamine) sugar of the N-linked glycosylation at Asn<sup>91</sup> have been resolved in the crystal structure of SC18 HA (PDB code 1RUZ). This sugar interacts with Arg<sup>224</sup>, which is a part of the 220-loop in the RBS (Figure 1B). Lys<sup>222</sup> and Asp<sup>225</sup> in SC18 HA make critical contacts with the Gal (galactose) sugar of the Neu5Ac $\alpha$ 2 $\rightarrow$ 6Gal-motif in the human receptor [15,19]. On the other hand, in AV18 HA, Glu<sup>190</sup> and Gln<sup>226</sup> make critical contacts with the Neu5Ac $\alpha$ 2 $\rightarrow$ 3Gal-motif in the avian receptor. The relatively lower-affinity binding of NY18 to the human and avian receptors has been attributed to the absence of the critical Asp<sup>225</sup> residue (for human receptor contacts) and Glu<sup>190</sup> (for avian receptor contacts) [15].

On the basis of these structural features of RBS and the glycosylation site at Asn<sup>91</sup>, it was hypothesized that any perturbation in the 220-loop due to loss of glycosylation at Asn<sup>91</sup> is likely to affect human receptor binding of SC18 HA, but have a relatively lesser effect on avian receptor binding of AV18 (given additional stabilization by Glu<sup>190</sup> that is not a part of the 220-loop). The 220-loop of NY18 HA makes fewer contacts with the human receptor and NY18 HA also lacks the longer side chain of Glu<sup>190</sup> for optimal contacts with the avian receptor. It was therefore hypothesized that perturbation of the 220-loop due to loss of glycosylation at Asn<sup>91</sup> would affect the binding of NY18 HA to both avian and human receptors. To experimentally test this hypothesis, a single T93A amino acid change that abrogates the sequon at Asn<sup>91</sup> (and hence removes glycosylation at this site) was introduced into SC18, NY18 and AV18 HAs. The glycan receptor binding of the mutant HAs were analysed using dose-dependent direct glycan binding on a glycan array platform.

### Dose-dependent direct binding analysis of the T93A mutants on the glycan array

The T93A mutant form of SC18 HA showed a characteristic binding to the human receptor (6'SLN-LN) (Figure 2) similar



**Figure 1** Structural framework for glycosylation at Asn<sup>91</sup> affecting glycan receptor binding

(A) Composite H1N1 HA monomer structure (derived from PDB code 1RUZ) with the key regions in the RBS and sequons shown in Supplementary Table S1 (at <http://www.BiochemJ.org/bj/444/bj4440429add.htm>) mapped on to the structure. (B) Structural complex of SC18 and NY18 with human receptor and AV18 with avian receptor where key residues involved in glycan receptor binding are shown. Homology models of these structures were obtained on the basis of the SC18 crystal structure (PDB code 1RUZ). The first GlcNAc sugar (indicated as GlcNAc-1) attached to Asn<sup>91</sup> critically interacts with Arg<sup>224</sup> in all of the structures.

to the wild-type SC18 HA. However, the binding affinity of the mutant was substantially lower than that of the wild-type protein. In the case of AV18, the T93A mutation did not alter the specificity or binding affinity to avian receptors (3'SLN, 3'SLN-LN and 3'SLN-LN-LN) (Figure 2) in comparison with the wild-type HA. Interestingly, the T93A mutation in NY18 completely abolished binding to both human and avian receptors when compared with wild-type HA. Therefore these observations validated our hypothesis on the basis of the structural framework.

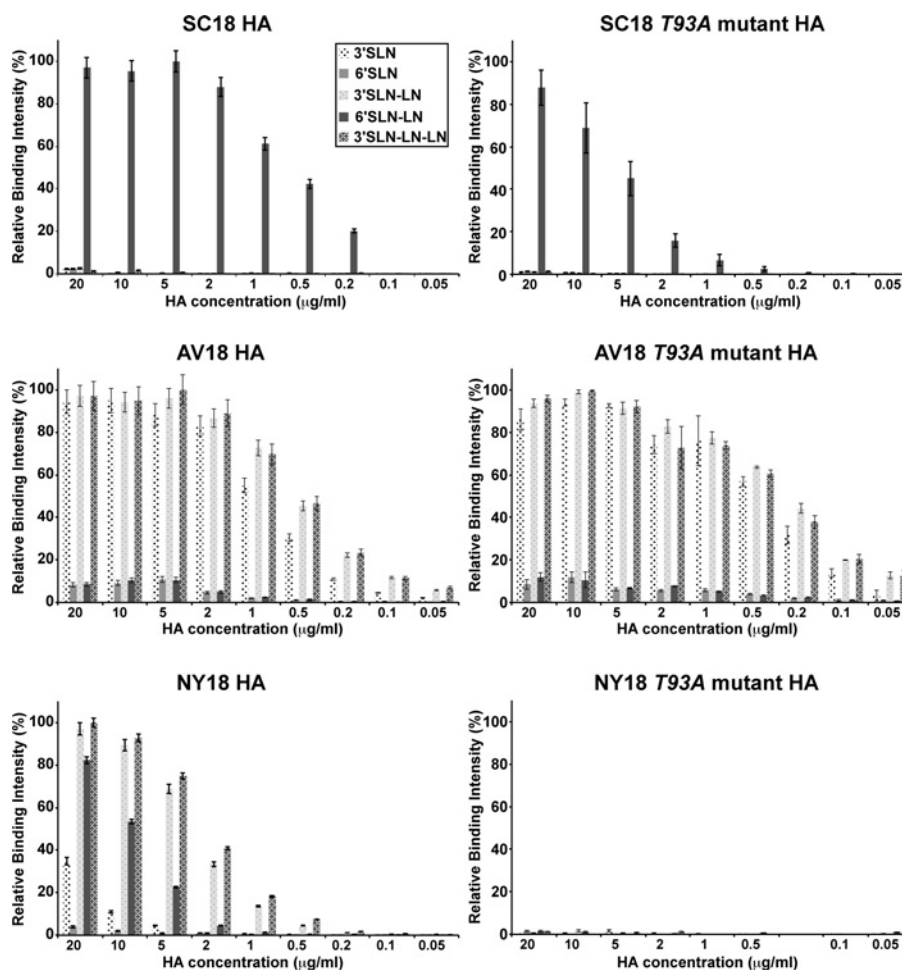
#### Network-based description of RBS: additional insights into observed differential effects of glycosylation

The SIN score calculated as described in the Materials and methods section captures the extent of networking of inter-amino acid interactions for each amino acid in the three-dimensional structure of HA. The SIN score has been correlated previously with constraints imposed on a residue in HA (including RBS) to undergo mutations [16]. A high SIN score for a residue implied

that it is the most constrained from undergoing mutations since it is highly networked in terms of interactions with surrounding residues. A low SIN score, on the other hand, implied that the residue is less constrained from undergoing mutations since it is poorly networked. Comparison of the SIN scores of key RBS residues showed that AV18 HA has the most networked residues in RBS, followed by SC18 HA and then by NY18 HA (Figure 3). Furthermore, the degree of networking of residues in RBS of these HAs appears to correlate with their ability to retain glycan receptor binding in the context of the loss of glycosylation at Asn<sup>91</sup>. Therefore, in addition to mutational constraints, the SIN scores of the residues appear to be linked to properties governed by the glycosylation of HA.

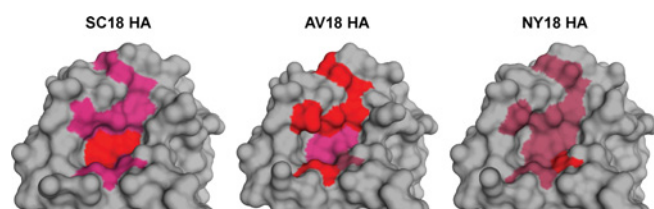
#### Characterization of glycosylation at Asn<sup>91</sup>

The effect of glycosylation at Asn<sup>91</sup> on receptor binding was primarily attributed to interactions between Arg<sup>224</sup> and the first core GlcNAc sugar of glycan N-linked to Asn<sup>91</sup>. The structure of



**Figure 2** Dose-dependent direct glycan binding of wild-type and loss of glycosylation mutant HAs

The dose-dependent direct binding of wild-type SC18, NY18 and AV18 HAs are adapted from our previous study [15]. Note that the human receptor (6'SLN-LN) binding drops from saturation at 20  $\mu\text{g/ml}$  for SC18 T93A mutant HA when compared with 2  $\mu\text{g/ml}$  for wild-type indicating substantial lowering of binding affinity as result of loss of glycosylation. On the other hand, the drop in binding from saturation for avian receptor (for example 3'SLN-LN) is the same between AV18 HA and its T93A mutant indicating that loss of glycosylation does not affect its avian receptor binding. CD analysis was performed on the wild-type and mutant HAs to show that loss of glycosylation does not affect the overall stability and fold of the protein (Supplementary Figure S1 at <http://www.BiochemJ.org/bj/444/bj4440429add.htm>).



**Figure 3** SIN in RBS of SC18, NY18 and AV18 HAs

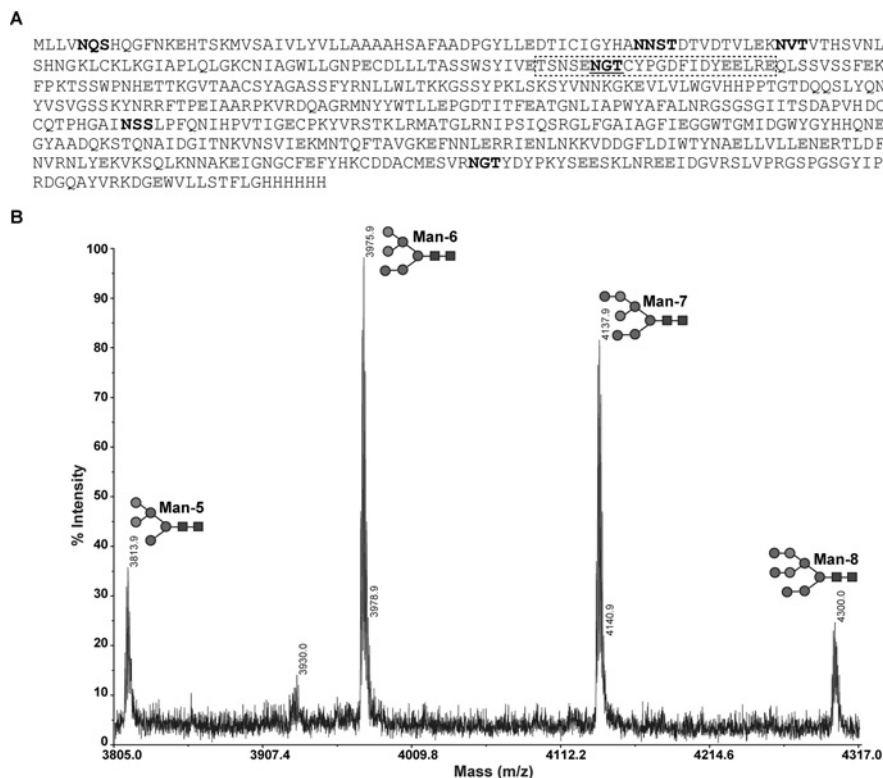
Surface rendering of the SC18, NY18 and AV18 RBS is shown in grey. The critical amino acids that interact with glycan receptor and have a threshold level of inter-residue networking (SIN score  $>0.1$ ) are shown and coloured as follows. For each critical amino acid among the three HAs, the one with highest SIN score is shown in red, the one with second highest value is shown in bright pink and one with lowest value is shown in dark pink.

the entire glycan at this site was characterized using MALDI-MS and TOF/TOF fragmentation.

The location of Asn<sup>91</sup> in the SC18 HA sequence was such that treatment with trypsin or Lys-C would result in a large peptide that, together with the glycan mass, would have a high molecular

mass. Such a molecular mass fragment is likely to cause detection issues due to poor ionization in MALDI-MS. To circumvent this issue, SC18 HA was digested using Glu-C (an enzyme that cleaves after the C-terminus of glutamic acid) as many glutamic acid residues flank the Asn<sup>91</sup> sequon-containing peptide (Figure 4A).

The glycopeptide peaks were identified on the basis of the following criteria. First, they did not match any predicted peptide mass peak upon simulating Glu-C digestion of SC18 HA. Secondly, the mass of these peaks differed by specific sugar units (such as hexose or *N*-acetylhexosamine). Thirdly, MS/MS fragmentation profiles of these peaks look similar given that the peptide fragment ions arise from the same parent peptide ion that is differentially glycosylated. An extensive search of the glycan structure database was performed to identify glycan compositions whose combined mass with the theoretical mass of peptide fragments containing Asn<sup>91</sup> sequon matched the experimentally observed MALDI-MS peaks. This analysis revealed that the compositions corresponding to the trimannosyl core, Man-5, Man-6, Man-7 and Man-8 glycans accurately matched the observed experimental mass peaks (Figure 4B). These results indicate



**Figure 4** Characterization of glycan structure at Asn<sup>91</sup>

(A) Sequence of SC18 with sequons indicated in bold. The Asn<sup>91</sup> sequon is underlined. The sites for cleavage of Glu-C protease are shown in grey. The glycopeptide fragment shown in a broken box containing the Asn<sup>91</sup> sequon when combined with glycan compositions having high-mannose structures matched experimental mass peaks. (B) Annotated experimental mass peaks of the glycopeptide fragment shown in the broken box in (A). Only that part of the MALDI-MS profile that matched this fragment is shown for clarity. The sequence of the glycopeptide fragment was confirmed further using MS/MS peptide fragmentation analysis. Circles are mannose sugar units and squares are GlcNAc sugar units.

that the Asn<sup>91</sup> site predominantly consists of high-mannose type N-linked glycans. The MALDI-MS peptide/glycopeptide profiling was also performed to confirm that the T93A mutation results in loss of glycosylation at the Asn<sup>91</sup> sequon (Supplementary Figure S2 at <http://www.BiochemJ.org/bj/444/bj44440429add.htm>).

## DISCUSSION

Several studies have pointed to the overall relationship between glycosylation states of HA in terms of number and structure of N-linked glycans and its glycan receptor-binding property. Glycosylation sites on HA are introduced or removed during evolution of viruses as a way to change its epitope surface in order to evade the host immune response. In the context of these changing sequons between various viral strains over time, even within a given subtype such as H1N1, there appear to be highly conserved sequons across these different strains. In addition to facilitating immune response evasion by the virus, the glycans at these conserved sequons may contribute to other key functions such as maintaining stability of different domains (RBS, membrane fusion, etc.).

The present study focused on understanding the functions of one such conserved glycosylation sequon (Asn<sup>91</sup>) in H1N1 HA using the SC18, NY18 and AV18 strains. These strains have distinct biological properties in terms of their glycan receptor-binding specificities that have been correlated with their ability to transmit via respiratory droplets in the ferret animal model. Analysis of HA crystal structures showed that glycosylation

at Asn<sup>91</sup> makes key contacts with residues that are a part of the 220-loop of the RBS. As predicted from this analysis, loss of glycosylation at Asn<sup>91</sup> affected human receptor binding of SC18 and NY18 HA, but not avian receptor binding of AV18 HA. Notably, the loss of glycosylation at Asn<sup>91</sup> had the most pronounced effect on NY18 HA by abrogating its binding to both avian and human receptors. NY18 was a natural variant of SC18 isolated from a fatal case during the 1918 pandemic. Therefore our results highlight the important role of this conserved glycosylation site in human receptor binding of H1N1 for efficient replication and sustainability in the human host.

In an earlier study, a new framework, termed SIN, was developed to capture and quantify the network of inter-residue interactions in HA. Using this framework, it was demonstrated that mutations in antigenic residues that were spatially distant from RBS residues on the three-dimensional HA structure, but connected via the inter-residue network, impinge on the glycan receptor-binding property of HA [16]. In the present study, we computed SIN scores of the RBS residues in SC18, NY18 and AV18 HA. It was interesting to note that the extent of networking of residues in the RBS captured by the SIN score correlated with glycan receptor-binding properties in the context of glycosylation at Asn<sup>91</sup>. This correlation may pave the way for future studies to understand the relationship between the extent of networking of a residue and other factors governing its stability, such as thermal motion, which is also influenced by glycosylation.

The present study also demonstrated that the predominantly observed glycosylation in this sequon is high-mannose-type glycans. Previous studies have shown that one of the mutations

involved in desensitizing H1N1 virus to the innate immune response in mice leading to increased virulence and rapid progression of disease is the loss of glycosylation at the Asn<sup>91</sup> sequon [13]. It has also been demonstrated that resistance to cyanovirin-N, a high-mannose-binding lectin, is acquired naturally by loss of glycosylation at Asn<sup>91</sup> during adaptation of a seasonal H1N1 virus in mice [20]. Given that a key part of the mouse innate immune response is binding of immune lectins to high-mannose-type glycans on pathogens facilitating their clearance [21], the presence of high-mannose-type glycans at Asn<sup>91</sup> is consistent with these observations.

In summary, by integrating structure-based network-based biochemical binding and glycan characterization analyses, the present study provides the structural basis for the effect of glycosylation at a conserved sequon in HA on the biological functions of the virus, such as glycan receptor-binding specificity and host adaptation.

#### AUTHOR CONTRIBUTION

Akila Jayaraman, Rahul Raman and Ram Sasisekharan were involved in the design of the study and writing the paper. Akila Jayaraman, Xiaoying Koh, Jing Li, Rahul Raman, Karthik Viswanathan and Zachary Shriver were involved in performing the various experiments and analyses including recombinant expression of HA and mutants, glycan array screening, MALDI-MS, and structural and network-based analyses.

#### FUNDING

This work was supported by the National Institutes of Health [grant number GM R37 GM057073-13] and in part by the Singapore-MIT Alliance for Research and Technology (SMART).

#### REFERENCES

- Taubenberger, J. K. and Kash, J. C. (2010) Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* **7**, 440–451
- Nicholls, J. M., Chan, R. W., Russell, R. J., Air, G. M. and Peiris, J. S. (2008) Evolving complexities of influenza virus and its receptors. *Trends Microbiol.* **16**, 149–157
- Russell, R. J., Stevens, D. J., Haire, L. F., Gamblin, S. J. and Skehel, J. J. (2006) Avian and human receptor binding by hemagglutinins of influenza A viruses. *Glycoconjugate J.* **23**, 85–92
- Skehel, J. J. and Wiley, D. C. (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu. Rev. Biochem.* **69**, 531–569
- Mansfield, K. G. (2007) Viral tropism and the pathogenesis of influenza in the mammalian host. *Am. J. Pathol.* **171**, 1089–1092
- Shinya, K., Ebina, M., Yamada, S., Ono, M., Kasai, N. and Kawaoka, Y. (2006) Avian flu: influenza virus receptors in the human airway. *Nature* **440**, 435–436
- van Riel, D., Munster, V. J., de Wit, E., Rimmelzwaan, G. F., Fouchier, R. A., Osterhaus, A. D. and Kuiken, T. (2006) H5N1 virus attachment to lower respiratory tract. *Science* **312**, 399
- Hensley, S. E., Das, S. R., Bailey, A. L., Schmidt, L. M., Hickman, H. D., Jayaraman, A., Viswanathan, K., Raman, R., Sasisekharan, R., Bennink, J. R. and Yewdell, J. W. (2009) Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. *Science* **326**, 734–736
- Igarashi, M., Ito, K., Kida, H. and Takada, A. (2008) Genetically destined potentials for N-linked glycosylation of influenza virus hemagglutinin. *Virology* **376**, 323–329
- Ohuchi, R., Ohuchi, M., Garten, W. and Klenk, H. D. (1997) Oligosaccharides in the stem region maintain the influenza virus hemagglutinin in the metastable form required for fusion activity. *J. Virol.* **71**, 3719–3725
- Kasson, P. M. and Pande, V. S. (2008) Structural basis for influence of viral glycans on ligand binding by influenza hemagglutinin. *Biophys. J.* **95**, L48–50
- Wang, C. C., Chen, J. R., Tseng, Y. C., Hsu, C. H., Hung, Y. F., Chen, S. W., Chen, C. M., Khoo, K. H., Cheng, T. J., Cheng, Y. S. et al. (2009) Glycans on influenza hemagglutinin affect receptor binding and immune response. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18137–18142
- Tate, M. D., Brooks, A. G. and Reading, P. C. (2011) Specific sites of N-linked glycosylation on the hemagglutinin of H1N1 subtype influenza A virus determine sensitivity to inhibitors of the innate immune system and virulence in mice. *J. Immunol.* **187**, 1884–1894
- Wei, C.-J., Boyington, J. C., Dai, K., Houser, K. V., Pearce, M. B., Kong, W. P., Yang, Z.-y., Tumpey, T. M. and Nabel, G. J. (2010) Cross-neutralization of 1918 and 2009 influenza viruses: role of glycans in viral evolution and vaccine design. *Sci. Transl. Med.* **2**, 24ra21
- Srinivasan, A., Viswanathan, K., Raman, R., Chandrasekaran, A., Raguram, S., Tumpey, T. M., Sasisekharan, V. and Sasisekharan, R. (2008) Quantitative biochemical rationale for differences in transmissibility of 1918 pandemic influenza A viruses. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 2800–2805
- Soundararajan, V., Zheng, S., Patel, N., Warnock, K., Raman, R., Wilson, I. A., Raguram, S., Sasisekharan, V. and Sasisekharan, R. (2011) Networks link antigenic and receptor-binding sites of influenza hemagglutinin: mechanistic insight into fitter strain propagation. *Sci. Rep.* **1**, 200
- Tumpey, T. M., Maines, T. R., Van Hoven, N., Glaser, L., Solorzano, A., Pappas, C., Cox, N. J., Swayne, D. E., Palese, P., Katz, J. M. and Garcia-Sastre, A. (2007) A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. *Science* **315**, 655–659
- Stevens, J., Corper, A. L., Basler, C. F., Taubenberger, J. K., Palese, P. and Wilson, I. A. (2004) Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. *Science* **303**, 1866–1870
- Liu, J., Stevens, D. J., Haire, L. F., Walker, P. A., Coombs, P. J., Russell, R. J., Gamblin, S. J. and Skehel, J. J. (2009) Structures of receptor complexes formed by hemagglutinins from the Asian Influenza pandemic of 1957. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 17175–17180
- Smee, D. F., Wandersee, M. K., Checketts, M. B., O'Keefe, B. R., Saucedo, C., Boyd, M. R., Mishin, V. P. and Gubareva, L. V. (2007) Influenza A (H1N1) virus resistance to cyanovirin-N arises naturally during adaptation to mice and by passage in cell culture in the presence of the inhibitor. *Antivir. Chem. Chemother.* **18**, 317–327
- Job, E. R., Deng, Y. M., Tate, M. D., Bottazzi, B., Crouch, E. C., Dean, M. M., Mantovani, A., Brooks, A. G. and Reading, P. C. (2010) Pandemic H1N1 influenza A viruses are resistant to the antiviral activities of innate immune proteins of the collectin and pentraxin superfamilies. *J. Immunol.* **185**, 4284–4291

Received 29 November 2011/23 February 2012; accepted 7 March 2012  
Published on the Internet 29 May 2012, doi:10.1042/BJ20112101



## SUPPLEMENTARY ONLINE DATA

# Glycosylation at Asn<sup>91</sup> of H1N1 haemagglutinin affects binding to glycan receptors

Akila JAYARAMAN, Xiaoying KOH, Jing LI, Rahul RAMAN, Karthik VISWANATHAN, Zachary SHRIVER and Ram SASISEKHARAN<sup>1</sup>

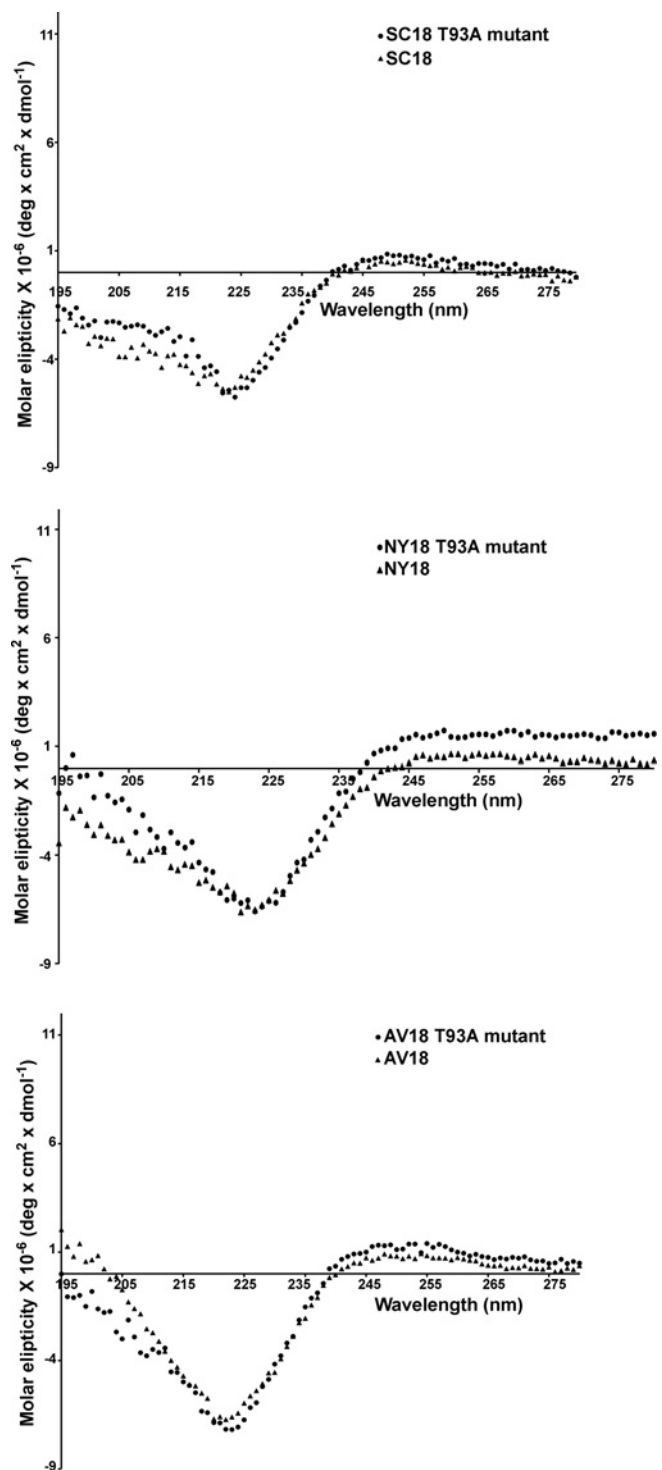
Harvard–MIT Division of Health Sciences and Technology, Koch Institute for Integrative Cancer Research, Singapore–MIT Alliance for Research and Technology, Department of Biological Engineering, Massachusetts Institute of Technology (MIT), 77 Massachusetts Avenue, Cambridge, MA 02139, U.S.A.

See the following pages for Supplementary Figures S1 and S2, and Supplementary Table S1.

---

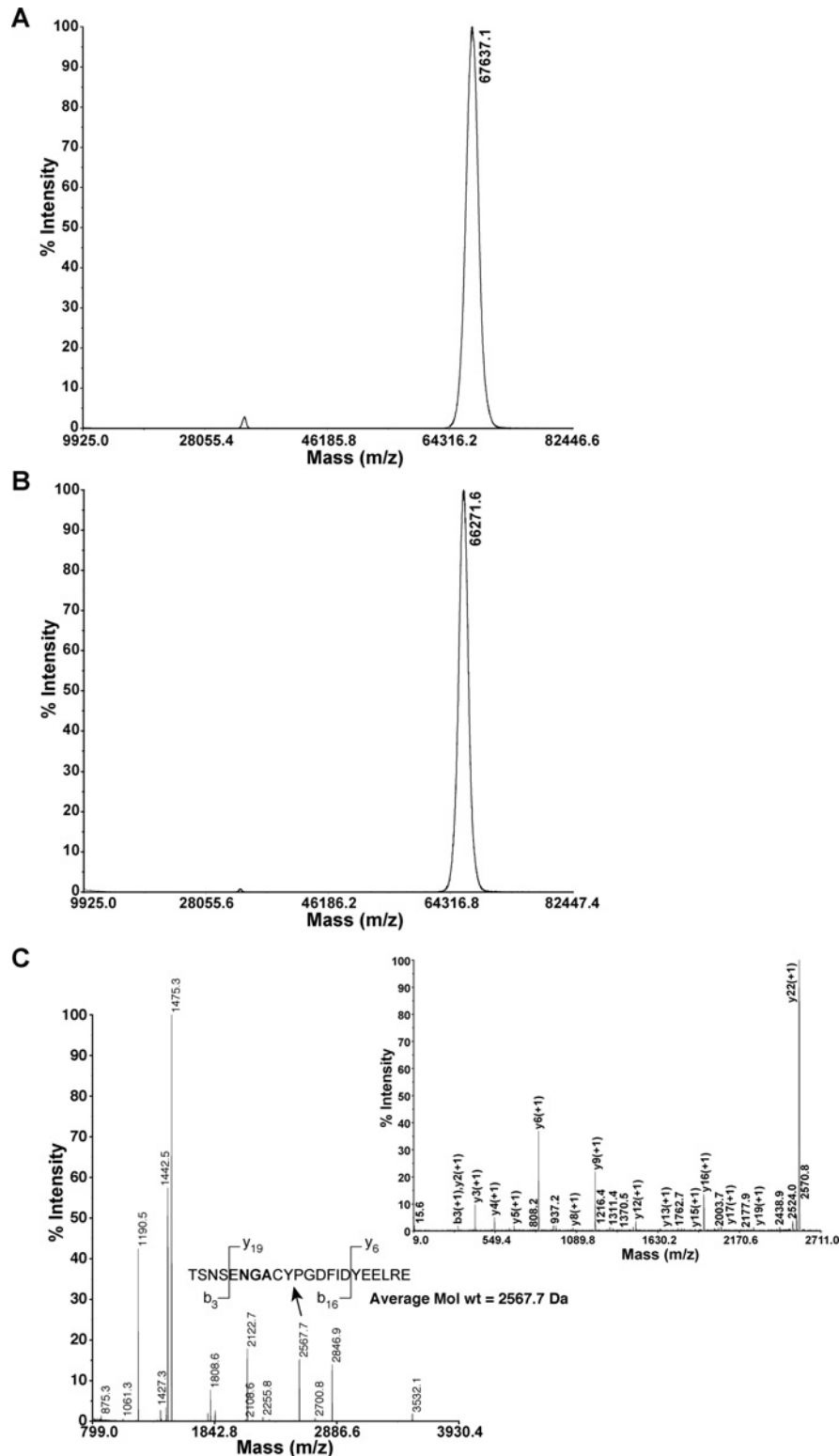
<sup>1</sup> To whom correspondence should be addressed (email rams@mit.edu).





**Figure S1 CD of wild-type and mutant HAs**

CD analysis of wild-type and Asn<sup>91</sup> loss-of-glycosylation mutant HAs. This analysis shows that the loss of glycosylation does not affect the overall protein fold and stability.



**Figure S2 MS analysis of loss of glycosylation resulting from the T93A mutation on HA**

(A) MALDI-MS peak of recombinantly expressed and purified SC18 HA (monomer). (B) MALDI-MS peak of the recombinantly expressed and purified T93A mutant form (monomer) of SC18 HA. The difference of  $\sim 1366$  in  $m/z$  values of the predominant peaks between SC18 wild-type and T93A mutant HA is consistent with a loss of N-linked glycans with a composition of Hex<sub>6</sub>HexNAc<sub>2</sub> (after accounting for the reducing end hydroxy group). (C) MALDI-MS analysis of peptides formed by digestion of SC18 T93A mutant HA with Glu-C and MS/MS fragmentation of the peptide (inset) containing Asn<sup>91</sup> (shown in red). Inspection of peptide mass and signature fragmentation ions confirms the absence of glycosylation at Asn<sup>91</sup> due to the T93A mutation that abrogates the sequon.

**Table S1 Glycosylation sequons in H1N1 HA**

The Table is adapted from [1]. The sequons are numbered based on the SC18 HA crystal structure (PDB code 1RUZ).

| Period       | Side of the head region of HA | Top of the head region of HA |
|--------------|-------------------------------|------------------------------|
| 1918         | 91                            |                              |
| 1933         | 60                            | 129                          |
| 1934–1939    |                               | 131                          |
| 1940–1948    | 91                            | 131                          |
| 1949–1957    | 91                            | 131 163                      |
| 1977–1985    | 91                            | 131 163                      |
| 1986–2008    | 58                            | 129 163                      |
| 2009–present | 91                            |                              |

## REFERENCE

- 1 Wei, C.-J., Boyington, J. C., Dai, K., Houser, K. V., Pearce, M. B., Kong, W. P., Yang, Z.-y., Tumpey, T. M. and Nabel, G. J. (2010) Cross-neutralization of 1918 and 2009 influenza viruses: role of glycans in viral evolution and vaccine design. *Sci. Transl. Med.* **2**, 24ra21

Received 29 November 2011/23 February 2012; accepted 7 March 2012

Published on the Internet 29 May 2012, doi:10.1042/BJ20112101