SCIENTIFIC REPORTS

SUBJECT AREAS:
COMPUTATIONAL BIOLOGY
STRUCTURAL BIOLOGY
VIROLOGY
EVOLUTION

Received
26 July 2011

Accepted
9 November 2011

Published
19 December 2011

Correspondence and requests for materials should be addressed to R.S. (rams@mit.edu)

Networks link antigenic and receptor-binding sites of influenza hemagglutinin: Mechanistic insight into fitter strain propagation

Venkataramanan Soundararajan[1], Shu Zheng[2], Neel Patel[3], Ken Warnock[2], Rahul Raman[1], Ian A. Wilson[4], S. Raguram[1], V. Sasisekharan[1] & Ram Sasisekharan[1,5]

[1]Harvard-MIT Division of Health Sciences & Technology, the Koch Institute for Integrative Cancer Research, and the Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [3]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [4]Department of Molecular Biology and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA, [5]Singapore-MIT Alliance for Research and Technology, Centre for Life Sciences, Singapore 117456, Singapore.

Influenza viral passaging through pre-vaccinated mice shows that emergent antigenic site mutations on the viral hemagglutinin (HA) impact host receptor-binding affinity and, therefore, the evolution of fitter influenza strains. To understand this phenomenon, we computed the Significant Interactions Network (SIN) for each residue and mapped the networks of antigenic site residues on a representative H1N1 HA. Specific antigenic site residues are 'linked' to receptor-binding site (RBS) residues via their SIN and mutations within "RBS-linked" antigenic residues can significantly influence receptor-binding affinity by impacting the SIN of key RBS residues. In contrast, other antigenic site residues do not have such "RBS-links" and do not impact receptor-binding affinity upon mutation. Thus, a potential mechanism emerges for how immunologic pressure on RBS-linked antigenic residues can contribute to evolution of fitter influenza strains by modulating the host receptor-binding affinity.

Influenza A virus infects host cells via interaction of the HA attachment protein with sialylated glycan receptors on host-cell membranes[1–3]. The primary host immune response to influenza involves antibodies with high neutralizing activity that recognize epitopes on the antigenic sites of HA, designated Ca, Cb, Sa, and Sb for H1 subtype HA[4,5]. The ability to circumvent these host antibodies via accumulation of amino acid mutations within the antigenic sites of HA results in "antigenic drift" of influenza viruses. This capacity is a global burden to track, which challenge vaccine development efforts[6,7].

While antigenic and receptor binding sites were historically perceived as distinct regions on HA[8,9], recent studies have shown that mutations at antigenic sites — including those at sites distant from the RBS — can notably modulate glycan receptor binding properties[10–12]. Receptor-binding properties of HA are a critical determinant of influenza evolution, and there is a need to understand how host antigenic pressure shapes the receptor-binding site properties of HA[13–16]. Such an understanding is important to enhance pandemic preparedness, especially in light of still circulating virulent H5N1 strains, evolution and spread of Tamiflu-resistant H1N1 strains, and widespread cross-host reassortment at a global-scale (as evidenced by the 2009 swine-origin H1N1 pandemic strain)[17].

It has long been known that amino acid interactions are important determinants of protein fold-function-evolution relationships[18–21]. Towards understanding the structural underpinnings of how antigenic site mutations modulate RBS properties and thus influence influenza virus evolution, we considered the networks of amino acid residue interactions for each residue on HA – termed the *Significant Interactions Network (SIN)* (**Figure 1**). Inter-residue atomic interactions — including hydrogen bonds, disulfide bonds, pi-bonds, polar interactions, salt bridges, and van der Waals interactions — were computed between all pairs of amino acid residues within the

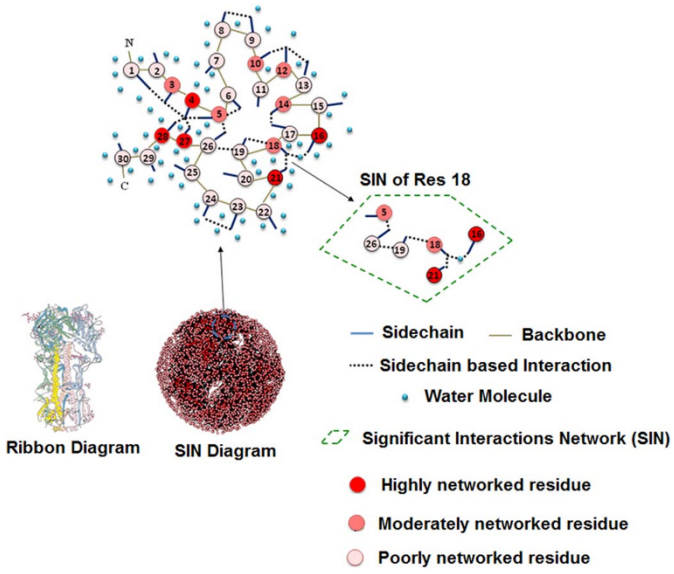

Figure 1 | Illustrating the significant interaction networks (SIN) for amino acid residues constituting the influenza virus HA structure. The significant interactions network (SIN) for each amino acid residue in a protein structure is its network of inter-residue interactions as computed from atomic interaction principles (see *Methods*). The degree of networking of each residue is assessed as a *SIN score* ranging from 0 to 1 (colored white to red) that may be considered a quantitative reflection of its network properties. Here, a illustrative depiction of these principles is provided for a randomly-selected region of 30 amino acids from the influenza H1N1 HA structure when immersed in an aqueous environment. Each amino acid residue is depicted as a circular node that is colored according to the SIN score of the residue — light pink for low SIN score (poorly networked; SIN score ranging from 0 to 0.25) residues, dark pink for medium SIN score (moderately networked) residues, and blood red for high SIN score (highly networked) residues. The N-terminus and C-terminus are highlighted for this region of the HA protein, and the peptide bonds constituting the backbone of this region are indicated (light brown lines). The side-chains of the 30 amino acid residues are also indicated (blue lines). Side-chain based atomic interactions are highlighted (broken black lines). Water molecules are shown as blue spheres. The SIN of residue 18 is highlighted as an example. The *ribbon diagram* perspective that depicts the HA protein structure in a biosynthetic manner (from N to C terminal along the series of peptide bonds) is shown in contrast with the *SIN diagram* perspective to the entire HA structure that comprehensively captures the network of inter-residue interactions.

trimeric HA structure. Integration of all such inter-residue interactions provided a quantitative measure for each HA residue, which we termed the *SIN score* (**Figure 1** - see *Methods* for details). The SIN scores of all HA residues were normalized based on the highest SIN score amino acid within HA, such that the scores varied from 0 (minimum) to 1 (maximum) for each residue.

The SIN perspective on HA structure provides a good correlation between SIN score of a residue and its conservation in sequence space across multiple HA subtypes (**Supplementary Figure S1**). Residues with higher SIN scores are highly conserved given that they are highly constrained to mutate from a network perspective. The residues with a high propensity to mutate all have low SIN scores due to lower constraints from a network perspective. Some residues with low SIN score are also seen to be highly conserved. These residues may have a higher propensity to mutate if there is any selection pressure (compared to high SIN score residues) due to lower constraints from a network perspective.

To classify the SIN scores of residues in HA, these scores were grouped based on the location of the residues in a representative trimeric H1N1 HA structure (**Supplementary Figure S2**). The solvent exposed residues (not involved in glycan receptor binding) predominantly had SIN scores in the range of 0–0.25. Given that these residues were outside the core or interface or RBS of the trimeric HA, they had a higher propensity to mutate which correlated with the lesser constraints on these residues from a network perspective. On the other hand a relatively higher fraction of residues (compared to solvent exposed residues) that were buried in the core or in the interface of trimeric HA structure or were involved in anchoring sialic acid of glycan receptor (as described below) had higher SIN scores (in the range of 0.25–0.5 or >0.5). Given the critical structural and functional role of these residues, they have a lower propensity to mutate which correlated with more constraints imposed on these residues from a network perspective (higher SIN scores). Based on the distribution of SIN scores in these contexts of residues in HA, each residue was classified as having a high SIN score [0.5–1], medium SIN score [0.25–0.5], or low SIN score [0–0.25]. In contrast to the classical *Ribbon diagram*, the resulting *SIN diagram* perspective to influenza HA structure captures all residues (*nodes*) and their integrated inter-residue atomic interactions (*edges*) (**Figure 1**).

The SIN perspective on HA structures permits intuitive contrast of the degree of "networking" of amino acid residues constituting HA structures, highlighted here for illustrative examples of "poorly networked" residues (**Figure 2A**) and "highly networked" residues (**Figure 2B**). In this study, we focus on the SIN of the residues constituting the antigenic sites of influenza H1N1 HA so as to evaluate the impact of antigenic site mutations on RBS residues. For this purpose, we use of the HA protein of the A/Puerto Rico/8/1934 (PR8) H1N1 influenza virus as a model system (see *Methods*). The PR8 HA protein was chosen as a model system due to the recently obtained *in vivo* experimental data on the antigenic site mutations escape mutants that emerged from PR8 virus passaging through pre-vaccinated mice or monoclonal antibody selection pressure[10].

The *150-loop* (W153, T155), *130-loop* (G134, T136), *180/190-loop* (H183, E190, L194), *90-loop* (Y98), and *220-loop* (Q226, G228) are involved in anchoring the Sialic Acid (SA) monosaccharide of the host glycan receptor to PR8 HA (**Figure 3A**). The composition, relative orientation of the side-chains, stability, and interactions for each of these receptor binding site (RBS) residues are critical determinants of host receptor-binding affinity for H1N1 HA[21–25]. The PR8 antigenic site residues are L79, L80, P81, V82, R83, S84 (*Cb antigenic site*); P128, N129, E156, K157, E158, G159, S160, P162, K163, L164, K165, N166, S167 (*Sa antigenic site)*; S140, H141, E142, G143, K144, S145, V169, N170, K171, K172, G173, T206, S207, N208, R224, D225, K238, P239, G240 (*Ca antigenic site);* and N187, S188, K189, E190, Q191, Q192, N193, L194, Y195, Q196, N197, E198 (*Sb antigenic site*) (**Figure 3B**). Thus, in this study, we focus on the SIN of each of these antigenic site residues and evaluate how antigenic mutations can impact HA affinity to the glycan receptor.

Results

SIN analysis of the PR8 trimeric HA crystal structure (obtained from PDB ID:1RVZ) shows that all of the experimentally observed mutations impinging on glycan receptor-binding affinity[10] are on antigenic residues with a SIN that includes SA-anchoring RBS residues (**Table 1**). An illustrative example is the SIN of K165 that includes H183 and E190 which are key SA-anchoring RBS residues, despite the fact that K165 has nearly 20 angstroms distance separation from H183/E190 on the PR8 HA structure (**Figure 4**). In addition, the SIN of K165 contains residues from one other neighboring HA monomer, thus "*connecting*" the HA glycoprotein across the HA1-HA1 protein-protein interface in the trimeric structure. Similarly, the SIN of I244 includes the SA-anchoring RBS residues H183 and L194 residues from the neighboring HA monomer, despite more than 20 angstroms distance between I244 and H183/L194 (**Figure 4**).

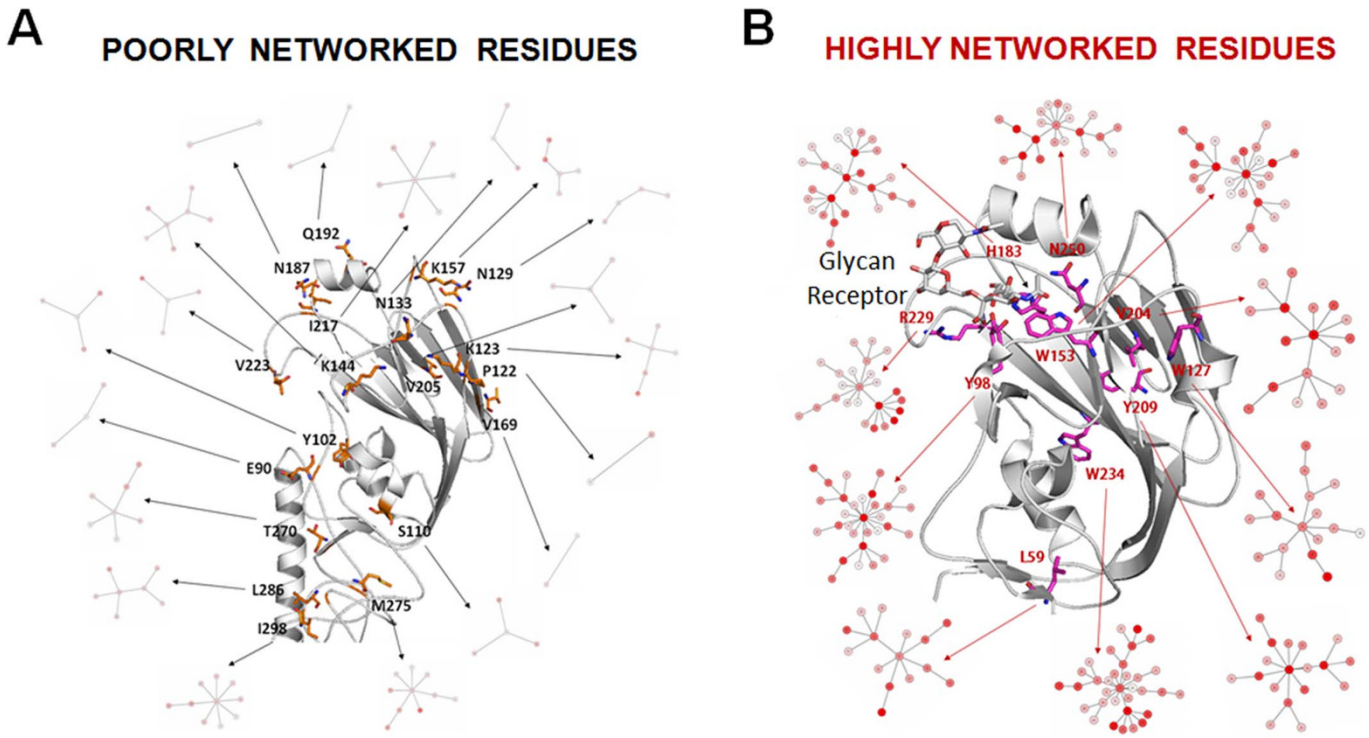

Figure 2 | Illustrative examples of poorly networked (low SIN score) residues and highly networked (high SIN score) residues of influenza H1N1 HA-1 domain. The SIN of a few (A) poorly networked (low SIN score) residues and (B) highly networked (high SIN score) residues for PR8 HA1 (receptor-binding) domain are provided to highlight the stark difference in the extent of "networking" for low SIN score and high SIN score residues. Each amino acid residue constituting the SINs represents a node in the network (shown as circles colored white-to-red based on the SIN score ranging from 0 to 1) with all significant inter-residue interactions representing an edge of the network (gray lines connecting a pair of nodes). The white-to-red coloring gradient implies that high SIN score residues are colored a darker shade of red for the node, whereas low SIN score residues are colored a lighter shade for the node. In addition to the stark differences in extent of "networking" between low and high SIN score residues, low SIN score residues (whose SINs are highlighted in *2A*) are noted to have very few high SIN score residues (red colored nodes) in their network, whereas high SIN score residues (whose SINs are highlighted in *2B*) are seen to have a large number of other high SIN score residues (red colored nodes) in their network. Another point of interest discussed in this study is that many of the low SIN score residues (highlighted in 2A) occur on the antigenic sites of PR8 HA1 whereas many of the sialic acid (SA) anchoring RBS residues happen to be high SIN score residues (highlighted in *2B*). The SA-containing host glycan receptor is indicated in *2B* for reference as sticks (with carbon atoms colored gray, oxygen atoms colored red and nitrogen atoms colored blue).

In addition to mutations on PR8 HA antigenic residues that have cross-monomer "RBS links", antigenic mutations affecting receptor-binding affinity are also seen to be on residues that have intra-monomer "RBS links". An illustrative example is the SIN of L164 which includes the SA-anchoring RBS residues W153, H183, Y98, and Q226 (**Figure 4**). An additional example is the SIN of N129 that includes the SA-anchoring RBS residue T155. Thus, escape mutations that impinge on glycan receptor binding affinity are observed to be on antigenic residues with intra-monomer or inter-monomer RBS-links (i.e. antigenic residues with a SIN that contains SA-anchoring RBS residues).

Further analysis of the specific antigenic site escape mutants with modified receptor-binding affinity shows that these HA mutants possess altered SIN for one or more SA-anchoring RBS residues (**Figure 5**). For instance, we observed a general trend of increase in the SIN score of H183 (a critical SA-anchoring RBS residue) as compared to its score in wild-type PR8 HA in each of the following antigenic mutants, N129K, E156G, E156K, L164Q, K165E, N166K, Q196R, E198G, R224I, and I244T. H183 is contained within the SIN of each of these antigenic residues that were mutated (**Table 1**). The PR8 HA-SA co-complex crystal structure shows that His-183 forms a hydrogen bond with the 9-hydroxyl group of SA, in addition to a hydrogen bond with the SA-anchoring Tyr-98 as well as with the RBS-proximal Y195 (**Supplementary Figure S3**). Site-directed mutageneses confirm that the extended hydrogen bond network associated with these residues is an important determinant of SA-binding affinity of HA[1]. The increase in SIN score of H183 correlates with increasing its stability in the RBS and offers an explanation for increase in glycan-receptor binding affinity by antigenic mutations that are far removed from the RBS in three-dimensional space. Conversely, the SIN score of H183 is lowered by I93T (another antigenic residue in its SIN) and this correlates with reducing the stability of this residue in the context of this mutation and hence offers an explanation of the observed reduction in glycan-binding affinity.

The above relationship between antigenic mutations that increase or decrease SIN score of key SA-anchoring RBS residues and the respective increase or decrease in glycan-binding affinity is consistently observed for all the antigenic escape mutants (**Figure 5; Table 1**). For instance, some of the mutations on other "RBS-linked" antigenic residues of PR8 are observed to modify the SIN of the critical SA-anchoring residues W153 (that stabilizes the RBS via extensive van der Waals interaction networks), Y98 (that hydrogen bonds with the 8-hydroxyl group of the SA moiety on the receptor), and L194 (that makes non-polar contacts with the N-acetyl methyl group of the SA moiety on the receptor)[1].

Amongst the few emergent PR8 escape mutants that showed no change in glycan receptor binding affinity, the mutating residues

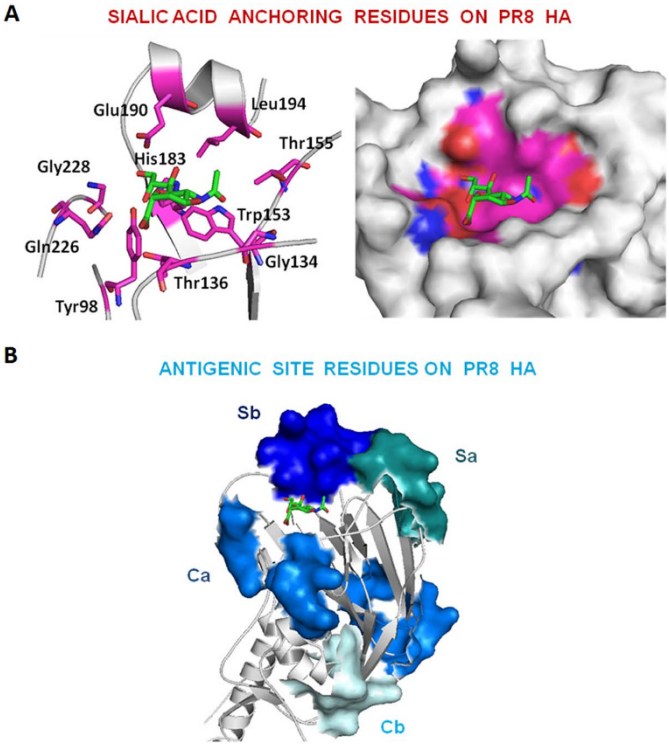

Figure 3 | Highlighting the amino acid residues constituting the sialic acid anchoring RBS residues and the antigenic site (Sa, Sb, Ca, Cb) residues of influenza H1N1 PR8 HA. (A) The RBS of PR8 HA is shown with the SA anchoring RBS residues highlighted in sticks (carbon atoms are colored pink, oxygen atoms are colored red and nitrogen atoms are colored blue). The SA moiety is also shown with carbon atoms colored green; oxygen and nitrogen atoms colored red and blue respectively. The left panel shows the cartoon depiction of HA-SA interaction, whereas the right panel shows the HA in its molecular surface rendering. (B) The antigenic site residues of the Sa, Sb, Ca, and Cb antigenic regions on PR8 HA are highlighted in molecular surface rendering colored aquamarine, dark blue, blue, and light cyan respectively. The exact amino acid residues and their numbers for these antigenic site residues are listed in the text.

almost always have no SA-anchoring RBS residues in their SIN (i.e. antigenic residues without any "RBS-link") (**Figure 5**; **Table 1**). The only examples of antigenic site mutations that occurred on "RBS-linked" antigenic residues are the N129Y, S160L, K163T, Q192L, and S140P mutations. However, none of these escape mutations has any effect on the SIN of SA-anchoring RBS residues.

Taken together, the above results demonstrate that antigenic site residues whose SIN contains SA-anchoring RBS residues (RBS-linked antigenic residues) may undergo specific types of mutations that influence SA-anchorage and, thus, receptor-binding affinity of HA. "RBS-linked" antigenic residues thus emerge as an important factor shaping the phenotype of escape mutants emerging from H1N1 virus evolution.

Towards understanding the immunological implications of our observations, we considered the known B-cell epitopes for PR8 HA from the immune epitope database (IEDB; www.immuneepitope.org). This analysis shows that many of the known B-cell epitopes (including epitope IDs 72805, 77507, 77508, 77509, 77510, 12285, and 76992) are constituted from one or more "RBS-linked" antigenic residues. The experimental data analyzed (*Table 1*)[10], shows that the specific RBS-linked antigenic residues contained within these epitopes are able to harbor mutations that modulate the SIN of key SA-anchoring RBS residues (**Table 2; Supplementary Figure S4**). This

Table 1 | Link between mutations in antigenic site and RBS residues for HA-escape mutants[10]. SA-anchoring RBS residues contained within the SIN of antigenic residues whose SIN score is generally higher than corresponding score in wild-type HA as a result of antigenic mutations that increase glycan-binding affinity are highlighted in yellow. Highlighted in cyan are SA-anchoring residues contained within the SIN of antigenic residues whose SIN score is generally lower than corresponding score in wild-type HA as a result of antigenic mutations that decrease glycan-binding affinity. In the case of mutations that do not alter the glycan-binding affinity, most of these residues do not have any SA-anchoring RBS residues in their SIN. The SA-anchoring RBS residues (whose SIN scores are not affected) are in the network of these antigenic mutations are shown in red text

Emergent Mutation	SIN Score	RBS Residues in SIN
Higher Affinity than WT		
L74P	LOW	Q226
L75P	LOW	Y98
E119G	LOW	Y98, T136
P128S	LOW	G134, T136
N129K	LOW	H183, T155, W153
G143R	LOW	Y98, T136
E156G	LOW	Y98, W153, T155, H183
E156K	LOW	Y98, W153, T155, H183
E158V	LOW	W153, T155
L164Q	LOW	W153, H183, Y98, Q226
K165E	LOW	H183, E190
N166K	LOW	H183, L194
Q192R	LOW	W153, T155
N193K	LOW	T136
Q196R	LOW	T155, W153, H183, Y98, Q226
E198G	LOW	H183, L194
R224I	LOW	Y98, W153, H183, Q226
I244T	LOW	H183, L194
Lower Affinity than WT		
I93T	LOW	Y98, T136, W153, H183, Q226
S145G	LOW	Y98
S145N	LOW	Y98, T136
N193S	LOW	W153
Similar Affinity as WT		
V77E	LOW	–
R78G	LOW	–
S79P	LOW	–
E119K	LOW	**Y98, T136**
N129Y	LOW	**H183, T155, W153**
S140P	LOW	–
G159D	LOW	–
S160L	LOW	**W153, H183**
K163T	LOW	**W153, H183**
S167Y	LOW	–
S167F	LOW	–
V169A	LOW	–
G173E	LOW	–
G173R	LOW	–
Q192L	LOW	**W153, T155**
G240E	LOW	–
G240R	LOW	–

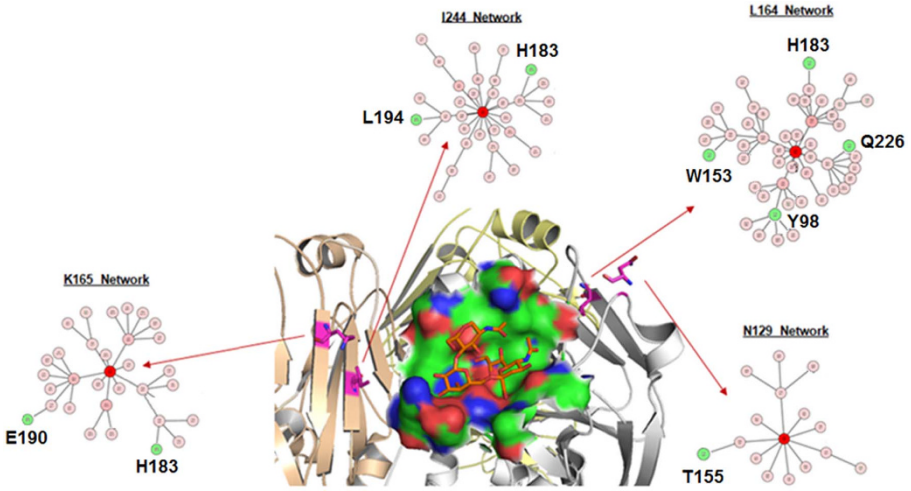

Figure 4 | Illustrative examples of "RBS-linked" antigenic residues on PR8 HA and their implications for influenza evolution. The SIN is shown for illustrative examples of antigenic site mutations on residues (*pink*) that are distant from the RBS (*green*), but unexpectedly found experimentally to impinge on glycan receptor (*orange*) binding affinity. Each of these SINs is shown as a network (residues are nodes and inter-residue interactions are edges with) with the following coloring scheme: the residue whose SIN is shown is colored blood red, SA-anchoring RBS residues contained within these SINs are colored green, and all other residues are colored light pink. It is clear that all residues whose SINs are highlighted have networks that contain one or more SA-anchoring RBS residues. Thus, these examples illustrate "RBS-linked" antigenic residues.

suggests that B-cell targeting of influenza PR8 HA (and potentially HA in general) — involving host antibodies recognizing surface epitopes within the Sa/Sb/Ca/Cb antigenic sites of HA — may be contributing to the emergence of potentially "fitter" influenza strains associated with increased host receptor-binding affinity.

Discussion

In addition to identification of "RBS-linked" antigenic residues as an important determinant of H1N1 evolution, SIN analysis of PR8 HA identified a remarkably high number of stabilizing atomic interactions between the highly networked SA-anchoring amino acids in the RBS of PR8 HA (many of which are indicated in **Figure 2**). Perhaps, a higher SIN profile of the RBS may be an evolutionary solution to limit the vibration entropy of HA RBS residues — an important consideration for enhancing protein-glycan interaction affinity[1], and in-turn influenza infection and transmission eficiency[23–25].

While the results presented in this study were derived by analyzing PR8 H1N1 HA as an illustrative model system, SIN analysis may be readily applied to analyze HA structures regardless of strain/subtype. More broadly, the results obtained here would suggest that an effective "antigenic-RBS linkage density" is a critical determinant of the evolutionary abilities of different H1N1 or for that matter other influenza strains. The analysis of known B-cell epitopes on PR8 HA suggests that immunologic targeting of influenza HA by host antibodies can select for escape mutants with increased receptor-binding properties, thus aiding in the propagating potentially fitter strains. As more antibody structures complexed to H1N1 HA are determined in future, the database of B-cell epitopes will expand, thus permitting a more comprehensive analysis of how B-cell targeting of influenza HA contributes to the evolution of receptor-binding properties.

This study emphasizes that SIN analysis may be a valuable tool to factor into mapping of the influenza antigenic site mutants that may be likely to emerge from global influenza circulation under herd immunologic pressure — particularly across the heavily pre-vaccinated communities during each influenza season. Indeed, SIN analysis of influenza HA structures provides a new perspective on the link between the receptor-binding affinity of HA and antigenic site mutants. This new perspective will be valuable in complementing current methods that track global circulation of influenza strains, such as antigenic cartography. In this capacity, continual network analysis of all circulating H1N1 HA structures can potentially accelerate and optimize the selection of the ideal vaccine strains for each flu season.

Methods

All protein sequences of H1N1 subtypes were obtained from www.fludb.org/brc/home.do. Sequences were aligned with MATLAB multialign and Jalview muscle multiple sequence alignment algorithms. Phylogenetic analyses were performed as required using the Phylowidget tool (www.phylowidget.org/full/index.html). Protein modeling was performed using Accelrys Discovery Studio (DS) by employing the build multiple homology models protocol (www.accelrys.com/products/discovery-studio/). The PR8 crystal structure (PDB ID:1RVZ) obtained from Protein Data Bank (www.rcsb.org) was chosen to build homology models of the antigenic mutant forms described throughout the manuscript. Pymol and python scripting were used for visualization of the modeled molecular structures. Modeled protein structures were analyzed with our significant interactions network (SIN) computation MATLAB protocols in the following manner. Using the coordinates of each protein structure (PDB file), 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 (each of these chemical and physical atomic interactions are described extensively in the literature; see references *S1-S45* for further information on these atomic interactions).

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 inter-chain interactions and long-range over short-range interactions (the relative energies of atomic interactions are described in the literature extensively; *Supplementary References S1–S45*). The resulting inter-residue energetic interaction matrix describes all first-order interactions for the analyzed molecular structure. 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 (henceforth termed *SIN score*) for each residue was computed by summing across the rows of the matrix, which was meant to correspond the extent of "networking" for each residue. The degree of networking scores were normalized with the maximum score for each protein so that the scores varied from 0 (minimum) to 1 (maximum) for each protein analyzed. MATLAB was used to develop the analytical methods outlined here. An R script was used to visualize the *SIN diagram* of each protein to visually appreciate the degree of *networks* constituting each protein structure (**Figure 1**). SIN scores were calculated for representative crystal structures at different resolutions (for the same protein) to demonstrate that small variations in the resolution of the structures did not alter the SIN of the residues in that structure (**Supplementary Table S1**).

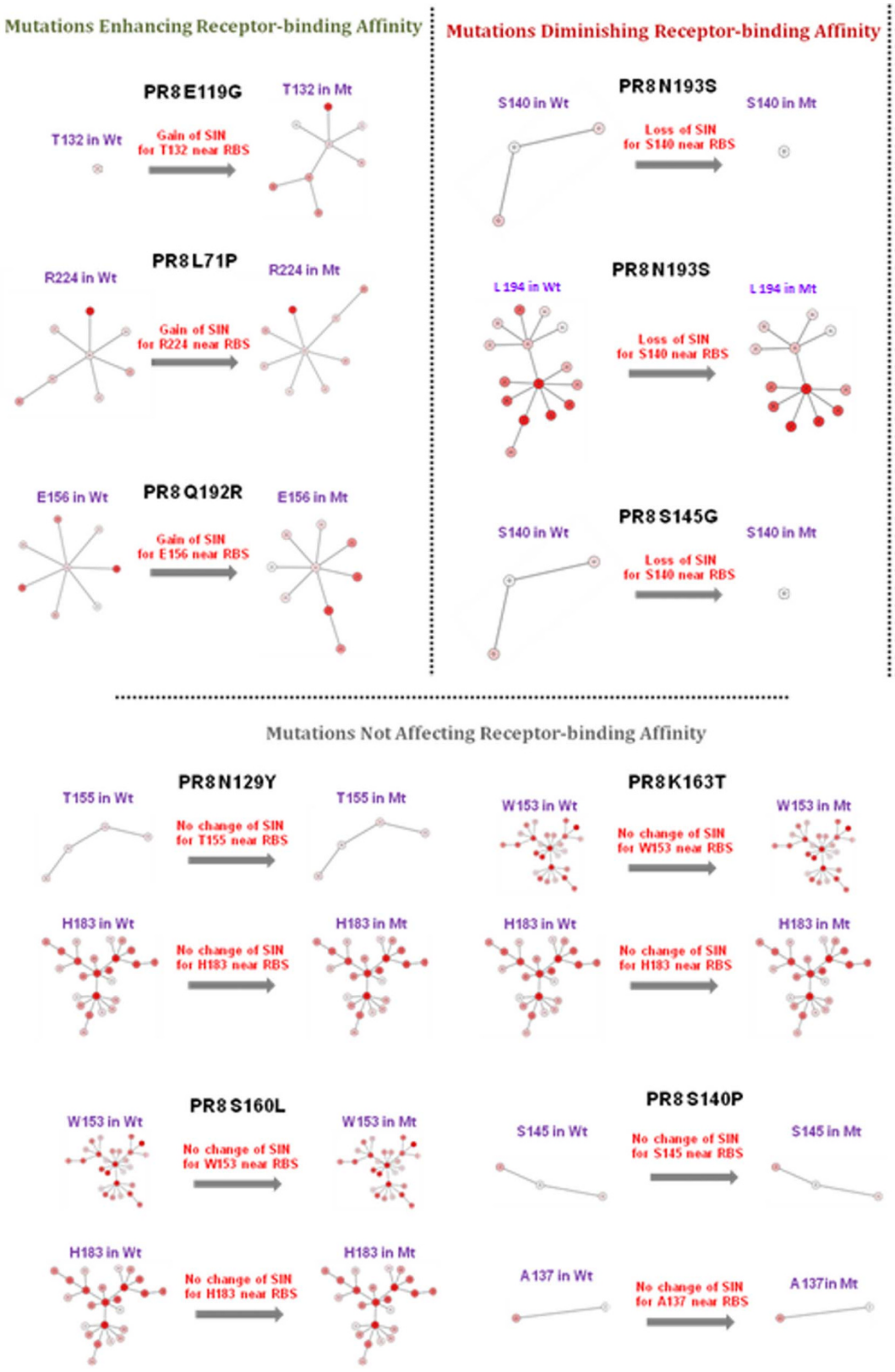

Figure 5 | Illustrative examples of SIN analysis for the HA antigenic mutants. As summarized comprehensively in *Table 1*, mutations at "RBS-linked" antigenic residues that enhance receptor-binding affinity also increase the SIN of one or more SA-anchoring RBS residues (*top left panel*). Conversely, antigenic mutations diminishing receptor-binding affinity are seen to decrease the SIN of SA-anchoring RBS residues (*top right panel*). On the other hand, antigenic mutations having negligible effect on receptor-binding affinity either have no SA-anchoring RBS residues in their SIN (*Table 1*), or more rarely do not modify the SIN of any SA-anchoring RBS residues that are contained within the SIN of such mutated antigenic residues (*bottom panel*). Taken together with the comprehensive results summarized in Table 1, these results show that specific types of mutation at "RBS-linked" antigenic residues have the potential to modulate receptor-binding affinity. This establishes "RBS-linked" antigenic residues as an important factor impinging on influenza genotype-phenotype relationship.

Table 2 | SIN analysis of the B-cell epitopes on PR8 HA compiled from the Immune Epitope Database (IEDB) is shown

B-cell Epitope ID	Epitope residues	Numbering based on PR8 Crystal Structure	Does the epitope have SA-anchoring RBS residues with potential to modulate receptor-binding affinity upon antigenic mutation?
72805	WLTEKEGSYP	153–162	YES (P162, E156, E158)
77507	E169, G172, Q205, N206, Q209, E211	156, 159, 192, 193, 196, 198	YES (E156, Q192, N193, Q196, E198)
77508	L87, L88, V90, R91, S92, E132	74–75, 77–79, 119	YES (L74, L75, E119)
77509	P141, N142, E171, S173, P175, K176, K178, N179, S180	128–129, 158, 160, 162–163, 165–167	YES (P128, N129, E158, P162)
77510	S153, G156, S158, R237, D238	140,143,145,224–225	YES (G143, S145, R224)
77511	V182, G186, I195, S220, G253, S287	169 173 182 207 240 274	YES (I195)
12285	EGSYPKLKNSYVNK	158–171	YES (E158, P162, L164, K165, N166)
57751	SFERFEIFPKE	114–124	YES (E119)
71335	VTGLRNIPSIQSR	318–330	NO
71336	VTGLRNIPSIQSRGLFGAIAGFIEG	318–330 + 501–512	NO
76950	E170, E172	157, 159	YES (E157)
76990	I258	244	NO
76991	N245	231	NO
76992	P199	186	YES (P199)

1. Skehel, J. J., Wiley, D. C. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu. Rev. Biochem.* **69**, 531–569 (2000).
2. Stevens, J., Blixt, O., Tumpey, T. M., Taubenberger, J. K., Paulson, J. C., Wilson, I. A. Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* **312**, 5772, 404–10 (2006).
3. Tumpey, T. M. *et al.* A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. *Science* **315**, 5812, 655–9 (2007).
4. Yewdell, J. W., Webster, R. G., Gerhard, W. U. Antigenic variation in three distinct determinants of an influenza type A haemagglutinin molecule. *Nature* **279**, 246 (1979).
5. Caton, A. J., Brownlee, G. G., Yewdell, J. W., Gerhard, W. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell* **31**, 417–27 (1987).
6. Gog, J. R. The impact of evolutionary constraints on influenza dynamics. *Vaccine* **26** Suppl 3, C15–24 (2008).
7. Ferguson, N. M., Galvani, A. P., Bush, R. M. Ecological and immunological determinants of influenza evolution. *Nature* **422**, 6930, 428–33 (2003).
8. Fleury, D., Barrère, B., Bizebard, T., Daniels, R. S., Skehel, J. J., Knossow, M. A complex of influenza hemagglutinin with a neutralizing antibody that binds outside the virus receptor binding site. *Nat Struct Biol.* **6**, 6, 530–4 (1999).
9. Wilson, I. A., Skehel, J. J., Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. *Nature* **289**, 5796, 366–73 (1981).
10. Hensley, S. E. *et al.* Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. *Science* **326**, 734–6 (2009).
11. Tokuriki, N., Oldfield, C. J., Uversky, V. N., Berezovsky, I. N., Tawfik, D. S. Do viral proteins possess unique biophysical features? *Trends Biochem Sci.* **34**, 53–9 (2008).
12. Domingo, E., Holland, J. J. RNA virus mutations and fitness for survival. *Annu Rev Microbiol.* **51**, 151–78 (1997).
13. Soundararajan, V. *et al.* Extrapolating from sequence--the 2009 H1N1 'swine' influenza virus. *Nat Biotechnol.* **27**, 6, 510–3 (2009).
14. Smith, D. J. *et al.* Mapping the antigenic and genetic evolution of influenza virus. *Science* **305**, 5682, 371–6 (2004).
15. Wei, C. J. *et al.* Cross-neutralization of 1918 and 2009 influenza viruses: role of glycans in viral evolution and vaccine design. *Sci Transl Med.* **2**, 24, 24ra21 (2010).
16. Das, S. R., Puigbò, P., Hensley, S. E., Hurt, D. E., Bennink, J. R., Yewdell, J. W. Glycosylation focuses sequence variation in the influenza A virus H1 hemagglutinin globular domain. *PLoS Pathog.* **6**(11), e1001211 (2010).
17. Kwong, P. D., Wilson, I. A. HIV-1 and influenza antibodies: seeing antigens in new ways. *Nat Immunol.* **10**, 573–8 (2009).
18. Soundararajan, V., Raman, R., Raguram, S., Sasisekharan, V., Sasisekharan, R. Atomic interaction networks in the core of protein domains and their native folds. *PLoS One* **5**(2), e9391 (2010).
19. Halabi, N., Rivoire, O., Leibler, S., Ranganathan, R. Protein sectors: evolutionary units of three-dimensional structure. *Cell* **138**, 4, 774–86 (2009).
20. Soundararajan, V., Patel, N., Subramanian, V., Sasisekharan, V., Sasisekharan, R. The many faces of the YopM effector from plague causative bacterium Yersinia pestis and its implications for host immune modulation. *Innate Immunity*, doi: 10.1177/1753425910377099 (2010).
21. Alon, U. Network motifs: theory and experimental approaches. *Nat Rev Genet.* **8**(6), 450–61 (2007).
22. Eisen, M. B., Sabesan, S., Skehel, J. J., Wiley, D. C. Binding of the influenza A virus to cell-surface receptors: structures of five hemagglutinin-sialyloligosaccharide complexes determined by X-ray crystallography. *Virology* **232**, 19–31 (1997).
23. Jayaraman, A. *et al.* A single base-pair change in 2009 H1N1 hemagglutinin increases human receptor affinity and leads to efficient airborne viral transmission in ferrets. *PLoS One* **6**(3), e17616 (2011).
24. Stevens, J., Blixt, O., Paulson, J. C. & Wilson, I. A. Glycan microarray technologies: tools to survey host specificity of influenza viruses. *Nat. Rev. Microbiol.* **4**, 857–864 (2006).
25. Xu, R., Xu, R., Ekiert, D. C., Krause, J. C., Hai, R., Crowe, J. E., Wilson, I. A. Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science* **328**, 357–360 (2010).

Acknowledgments

This work was supported by National Institutes of Health grant GM R37 GM057073-13 and Singapore–MIT Alliance for Research and Technology (SMART) to RS and the grant AI058113 to IAW. The authors would like to thank Mike Rooney for his help with developing a MATLAB and R based network visualization tool.

Author contributions

Venkataramanan S., S.Z., N.P., K.W., R.R. and S.R. carried out analyses and wrote various codes for the analysis. Venkataramanan, S., S.R., V.S. and R.S. designed the methods and the overall conceptual definitions. I.A.W. provided additional structural perspectives to the approach. Venkataramanan S. and R.S. did the major writing of the manuscript. R.R. and R.S. reviewed the manuscript.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

License: This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/

How to cite this article: Soundararajan, V. *et al.* Networks link antigenic and receptor-binding sites of influenza hemagglutinin: Mechanistic insight into fitter strain propagation. *Sci. Rep.* **1**, 200; DOI:10.1038/srep00200 (2011).