Supplementary Information for

A universal design of restructured dimer antigens: development of a superior vaccine against the Newcastle disease virus in transgenic rice

Fanshu Ma[1,2,4,9†], Qianru Xu[6†], Aiping Wang[4], Daichang Yang[7], Qingmei Li[5], Junqing Guo[5], Longxian Zhang[1,2], Jiquan Ou[8], Rui Li[5], Heng Yin[8], Kunpeng Li[8], Li Wang[5], Yanan Wang[5], Xiangyue Zhao[5], Xiangxiang Niu[1], Shenli Zhang[1], Xueyang Li[1], Shujun Chai[5], Erqin Zhang[1,2*], Zihe Rao[10*], Gaiping Zhang[3,1,2, 4*]

*Correspondence: zhanggaiping2003@163.com (G.Z.), zhangerqin76@163.com (E.Z.), raozh@mail.tsinghua.edu.cn (Z.R.).

This PDF file includes:

SI Materials and methods.
Fig. S1 to S12.
Table S1 to S2.

SI Materials and methods

Construction of the Plant Vector and Rice Genetic Transformation

The 35 HN sequences of NDV clinical strains were aligned in NCBI. The DNA sequence coding for the HN gene (GenBank accession No. JN618348.1) was synthesized by GenScript Corporation using the rice codon bias. The HN gene was subcloned in the *Mly* I–*Xho* I sites of the intermediate vector pMP3 containing a Gt13a promoter, a signal peptide, and a *t-nos* terminator (Healthgen Biotechnology Co., Ltd, China). The recombinant plasmid *pMP3-2HN* was digested by *EcoR I* and *Hind III* to clone into the plant vector pCAMBIA1300 (Healthgen Biotechnology Co., Ltd) that contained the *hpt II* (hygromycin resistance) gene as a selective marker, as well as the right and left borders necessary for T-DNA transmission. The plasmid pCAMBIA1300-2HN was transformed into the callus regenerated from rice cultivar TP309 by *Agrobacterium*-mediated transformation as described previously (1).

PCR Analysis

Total genomic DNA was extracted from young leaves of transgenic rice by the CTAB method (Doyle et al., 1987). PCR products were amplified with the forward primer HN1 (5′-CACATCCATCATTATCCATCCACC-3′) and reverse primer HN2 (5′-TGACGAACTGC TGCATCTTG-3′). The program of the PCR was one cycle of 94 °C for 5 min; 35 cycles of 94 °C for 40 s, 52 °C for 40 s, 72 °C for 50 s followed by a 72 °C incubation for 10 min. The PCR products were identified by 1 % agarose gel electrophoresis and then photographed.

Colloidal Gold-based Immunochromatographic Strip Test

To identify whether recombinant HN proteins were expressed, rice seed extracts of different transgenic lines were detected using an NDV colloidal gold immunochromatographic strip (Henan Provincial Key Laboratory of Animal Immunity, Henan Academy of Agricultural Sciences, China). Each sample extract was diluted 1:5000 with physiological saline, and 100 μL of the sample was added to the sample pool. The non-transgenic rice strain TP309 was used as a negative control. After 5 min, the results of the Test-line (T) and the Control-line (C) were recorded. Two red lines indicated a positive result, and only the C-line indicated a negative result.

ELISA

HN protein concentrations in the extracts were measured by an ELISA quantitative assay. Briefly, chicken anti-NDV IgG (stored in our lab) was diluted to 1:500 by carbonate buffer (pH 9.6) and then coated onto 96-well microplates (Corning, USA) overnight at 4 °C. The plates were blocked with 5 % skim milk for 2 h at 37 °C. The rice extract and pure HN protein standard were diluted and incubated for 1 h at 37 °C. After washing five times, a mouse monoclonal antibody against HN protein was added to the wells and incubated for 1 h. Subsequently, goat anti-mouse IgG/HRP antibody (Abbkine, Beijing, China) was added, and the reaction was carried out at 37 °C for 1 h. After washing five times, tetramethylbenzidine (TMB) substrate (Sigma-Aldrich) was added to each well. Ten minutes later, the OD_{450} value was read by a microplate reader. A standard curve was drawn according to the OD values and concentrations of the standard HN protein, and the content of total soluble protein (TSP) was determined according to the OD_{450} nm value of the sample.

HN Protein Purification

The rice seeds were ground into powder and extracted with phosphate buffer (25 mM PB, pH 5.7) in a 1:6 (wt/vol) ratio at room temperature for 1.5 h with constant stirring. The seed residue and precipitate were removed by centrifugation at 20,000 g for 30 min at 4 °C. The supernatant was aspirated, and the conductance value and pH were measured. The sediment was discarded by filtration with a 0.8 to 0.22 μm filter. Then, the clarified extract was loaded onto an SP HP purification column (GE Healthcare) with PB buffer (25 mM, pH 5.7). The impurities were removed with 8 % and 100 % buffer B (25 mM PB, 1 M NaCl), and the HN protein was eluted with 30 % buffer B (25 mM PB, 1 M NaCl). The enriched HN protein was loaded onto a 75 pg gel

filtration purification column, and the first peak was collected. The sample was loaded onto a Superdex 200 increase gel filtration purification column (GE Healthcare) to obtain a final sample for crystallization.

SDS-PAGE and Western Blot

Transgenic rice extracts and pure OsrHN protein were separated by 12 % SDS–PAGE and then transferred to PVDF membranes (Millipore). The membranes were blocked to prevent nonspecific reaction by phosphate-buffered saline-Tween 20 (PBST) buffer containing 5 % skim milk. After 2 h, Western blot membranes were washed three times with PBST buffer and incubated with monoclonal antibody (stored in our lab) for the HN protein at a 1:1,000 dilution (2). After 1 h, HRP-conjugated goat anti-mouse antibody at a 1:8,000 dilution (Abbkine, Beijing, China) was made to react with the monoclonal antibody for 1 h at room temperature. Finally, the hybridization reaction was detected by chemiluminescence detection.

Immune-electron Microscopy

To perform immune-electron microscopy analysis, immature transgenic rice seeds at 14 days after flowering were collected and fixed for 2 h at 4 °C in 0.1 M phosphate buffer containing 2.5 % glutaraldehyde (pH 7.3). After fixation, the samples were postfixed on ice for 1 h in 1% osmium tetroxide in 0.1 M phosphate buffer, dehydrated using a graded series of ethanols, and embedded in LR white resin. Ultrathin sections were prepared using an ultramicrotome, and then treated with primary anti-HN (dilution 1:1000) followed by secondary 20-nm gold particle-labeled goat anti-mouse IgG (dilution 1:200). Finally, the sections were stained with uranyl acetate for 15 min and examined using a transmission electron microscope (H-7100, Hitachi, Tokyo).

Crystallization

The NDV Osr2HN was concentrated to ~8 mg/mL in 25 mM PB, 100 mM NaCl, pH 5.7. The crystallization conditions were identified by screening using Hampton index crystallization kits (Hampton Research). The protein was crystallized at 25 °C by the sitting drop vapor diffusion method. The drops contained protein and precipitant of 60 % Tacsimate, pH 7.0, at a 1:1 ratio. After 3–4 days, crystals with a rectangular shape appeared. The crystals were harvested and frozen in 60 % Tacsimate (pH 7.0) and 15 % glycerol.

X-ray Data Collection, Structure Determination, and Refinement

Plant-made 2HN X-ray diffraction data were collected at the Shanghai Synchrotron Radiation Facility (SSRF) BL17U1. The datasets were processed with HKL2000 software (3). Structures were determined by the molecular replacement method using the CCP4 suite. The NDV HN Victoria tertiary structure (PDB ID: 3t1e) was used as the search model to determine the initial phases. Manual rebuilding and structure refinement were performed with COOT (4). Figures of the structure were prepared by PyMol.

Small-angle X-ray Scattering (SAXS) and Structural Modeling of Osr2HN

Three different concentrations (1.1 mg/mL, 3 mg/mL, and 4.8 mg/mL) of Osr2HN were used for SAXS measurements. Data for buffers were collected, and the scattering data were then scaled and the average values for the buffers subtracted. The qualities of the scattering curves were analyzed by the program PRIMUS (5) to ensure that there was no obvious aggregation or radiation damage. The P(r) distribution was calculated with the program GNOM (6). The low-resolution envelope of the Osr2HN in solution was modeled by the DAMMIF program (7). The structure of Osr2HN obtained by crystallization in this study was superimposed onto the low-resolution model of the Osr2HN in solution using the program Chimera (8).

Glycoproteomics Analysis

The N-glycosylation profile of the HN protein was analyzed using mass spectrometry, following previously established methods (9). In brief, the purified Osr2HN protein was incubated in a 50 mM triethylammonium bicarbonate (TCEP) solution (Sigma-Aldrich) at 55 °C for 1 h. The reduced HN protein was then alkylated with iodoacetamide (IAA) and incubated at room temperature for 30 min. Subsequently, the HN protein was digested overnight at 37 °C using trypsin. The resulting peptides were desalted using a C18 microcolumn (Phenomenex, 15 μm, 300Å) and dried under vacuum. Glycopeptides were enriched using ZIC-HILC (Merck Millipore, 5 μm, 200Å). The samples were analyzed using C18-RPLC-MS/MS (HCD) by Wuhan GeneCreate Biological Engineering Co., Ltd.

Immunization Study in Mice

BALB/c mice of 6-8-week-old females were divided into three groups. The plant-derived HN vaccine was prepared by mixing the purified HN protein and Montanide™ ISA 71 VG adjuvant (Seppic, France). The purity of the Osr2HN and OsrHN protein after three purification steps reached 95 %. According to the vaccine dose in each group, the pure HN protein produced by plants was diluted in PBS and mixed with Montanide™ ISA 71 VG adjuvant at a ratio of 3:7. PBS mix with adjuvant as the negative control. And then the mice were immunized with 0.5 μg and 10 μg Osr2HN or OsrHN, intramuscularly, at weeks 0 and 4 weeks. The blood samples were collected following each immunization for analyzing HN-specific serum antibody responses.

Animal Immunization and Challenge Study in Chicken

All of the chickens in the experiments were four-week-old SPF chickens purchased from Beijing Boehringer Biotechnology Co., Ltd (Beijing, China). The chickens were acclimated for one week prior to the beginning of the vaccine study. The chickens were numbered individually and randomly assigned into eight treatment groups. To verify that the four-week-old SPF chickens had no prior exposure to NDV, twenty chickens randomly selected prior to the start of the trial were bled. A hemagglutination inhibition (HI) test was used to detect the HN-specific antibodies in the serum.

The plant-derived HN vaccine was prepared by mixing the purified HN protein and Montanide™ ISA 71 VG adjuvant (Seppic, France). The pure Osr2HN was diluted in PBS and mixed with Montanide™ ISA 71 VG adjuvant at a ratio of 3:7. The commercial LaSota attenuated vaccine, LaSota inactivated vaccine, and A-VII inactivated vaccine were purchased from the manufacturer. According to the label, the EID_{50} of the commercial live vaccine LaSota (HARVAC, China) is ≥ 10^7 per recommended dose (0.05 mL) for initial immunization or booster injections by the intramuscular route.

In the eight experimental groups, six groups were intramuscularly (IM) inoculated with different doses of Osr2HN (0.5 μg, 1.5 μg, 4.5 μg, 9 μg, 18 μg and 36 μg) of 100 μL. The seventh group was intramuscularly inoculated with the same volume of non-transgenic rice (TP309) or PBS as a negative control. The eighth group was vaccinated with 50 μL of commercial live vaccine by eye drop as a positive control. One day before the vaccination, 1 mL of blood was sampled from the wing vein of each of 10 randomly selected chickens to confirm that the SPF chickens had no prior exposure to NDV. All of the treatment groups were boosted with the respective vaccine 28 days after the initial vaccination. Sera were collected on days 7, 14, 21, and 28 post-inoculation after the first and booster immunizations to determine the level of antibody response to the treatment. Chickens were routinely monitored daily for the first seven days after immunization to assess safety. Lesions at the injection site, animal behaviour, and body weight were also recorded. The virus challenge experiment was performed 28 days after booster immunization. The chickens were challenged via the oculo-nasal route with $10^{6.5}$ EID_{50} of the highly virulent NDV (XX-08 strain of gene type VII). Chickens were observed daily throughout the trial for clinical signs of disease until 15 days post-challenge.

Flow Cytometry of T-cell Responses

PBMCs were stimulated with HN peptides along with anti-CD28/anti-CD49d (BD Biosciences) co-stimulatory antibodies (1 μg/mL) for 2 h at 37 °C in a CO_2 incubator. 1 μL of Golgi-plug and Golgi-

stop were added and incubated for another 4 h. PBMCs were then surface stained with antibodies against CD3, CD4, CD8, and Live-Dead-APC-Cy7. The stained cells were washed once and permeabilized with cytofix/cytoperm for 30 min, and then stained with anti-IFN-r, anti-TNF-a, anti-IL2, anti-IL-4, and anti-IL-17 antibodies for 30 min. Finally, the cells were washed and analyzed by flow cytometry.

Antibody Isotype, IgG Subclass of Mouse Sera

The antibody isotype was determined by ELISA. Briefly, purified HN protein was diluted by carbonate buffer (pH 9.6) and coated onto 96-well microplates (Corning, USA) overnight at 4 °C. After three washes with PBST, the plates were blocked with 5 % skim milk for 2 h at 37 °C. The immune sera from different vaccine groups were serially diluted and added into the plate. After incubation at 37 °C for 1 h, HRP labeled anti-mouse IgG1, IgG2, IgG3, IgG4, IgA, and IgM (Abcam, UK) were added. The plate was incubated at 37 °C for 1 h. After washing five times, tetramethylbenzidine (TMB) substrate (Sigma-Aldrich) was added to each well. Ten minutes later, the OD_{450} value was read by a microplate reader.

Virus Neutralization Assay

BHK21 cells were plated into 96-well cell plates and incubated at 37 °C in a CO_2 incubator for 16 h. Subsequently, 50 μL of 100 $TCID_{50}$ NDV F48E8 strain and XX-08 strain were mixed with an equal volume of two-fold serially diluted chicken sera in a 96-well plate. Both the virus and serum were diluted in serum-free DMEM. After incubation at 37 °C for 1 h, virus and serum mixtures were added to 96-well cell plates in sequence. The 96-well plates were incubated for 2 h to allow the virus to fully infect the cells, and then the solution in the well was discarded, and the cell plate was gently washed twice with sterile PBS. Then, 100 μL of DMEM containing 2 % FBS was added to each well and cultured at 37 °C in a CO_2 incubator. The cell plate was removed at 48 h. The cells were fixed by adding 100 μL of pre-cooled absolute ethanol to each well. Immunocytochemical staining of NDV in cells was performed with anti-HN protein monoclonal antibody 5F2 and goat anti-mouse IgG / HRP antibody (Abbkine, Wuhan, China). The plate was incubated with antibodies for 1 h at 37 °C, and was washed three times with PBST. The virus staining results were observed under the microscope. When the antibodies in the serum completely neutralized the virus, there was no cell staining in the wells. Conversely, when neutralizing antibodies did not completely neutralize the virus, red-stained cells appeared in the wells. The neutralizing titer in the serum sample was determined as the highest dilution that causes 100 % neutralization, and expressed as the reciprocal of this dilution.

Hemagglutination Inhibition (HI) Assay

Chicken serum or purified antibodies were two-fold serially diluted. Four HA units of NDV were mixed with diluted serum in a V-shaped 96-well plate. The plate was then incubated at 37 °C for 1 h. The washed chicken red blood cells were subsequently added to each well at a concentration of 1 % and incubated for 30 min at room temperature. The HI titer was recorded and read as the highest serum dilution to prevent hemagglutination.

Peptides Map Assay

The peptide map assay was used to analyze the binding of antibodies to the overlapping peptides of the HN protein. The peptides were artificially synthesized based on the full-length HN protein sequence of the F48E8 strain (GenBank accession No. ACK57499.1), and conjugated to BSA using sulfo-SMCC according to the manufacturer's instructions. Briefly, 20 mg of BSA was reacted with 10 mg of sulfo-SMCC for 1 h at room temperature, followed by the addition of 20 mg of peptide and incubation for 2 h at room temperature. Unconjugated peptides were removed by purification using a desalting column (Thermo), and the conjugated peptides were stored at -80 °C.

For the ELISA assay, ELISA plates (Thermo) were coated overnight with BSA-peptide at a concentration of 100 ng/well, while BSA alone served as the negative control. After blocking with 5 % skim milk for 2 h, the immunized sera from Osr2HN, OsrHN, and LaSota inactive vaccine

were added at a dilution of 1:50 and incubated at 37 °C for 1 h. After washing, HRP-labeled anti-chicken antibodies were added at a dilution of 1:10,000 and incubated for another hour. The plate was then stained with TMB substrate (Sigma-Aldrich) and read using a microplate reader at OD_{450} nm.

Virus NA Assay

The antibody was purified by protein G columns (GE Healthcare, USA). The NA activity of viruses was determined by a fluorescence-based assay using a neuraminidase assay kit (Beyotime Biotechnology, China) according to the manufacturer's instructions. Briefly, 70 μL of detection buffer and 10 μL of allantoic fluid infected with NDVs were added to 96-well polystyrene plates, followed by the addition of 10 μL Milli-Q water. The solution was mixed gently for 2 min. Then, 10 μL of NA fluorescent substrates were added to the solution and mixed well. After incubation at 37 °C for 30 min, the fluorescence with an excitation wavelength of 322 nm and an emission wavelength of 450 nm was measured by a fluorescence spectrophotometer (BioTek). NA activity is shown as the fluorescence intensity of the samples minus the background values of non-infected allantoic fluid.

Inhibition Effects on Virus Entry and Release from the Cells

To study the inhibitory effects of antibodies on virus entry and release from cells, we conducted co-incubation experiments by using BHK-21 cells.

For virus entry inhibition, we incubated serially diluted immune serum or purified antibodies with 100 $TCID_{50}$/well of XX-08 strain of NDV for 1 h. The antibody-virus mixture was then added to the BHK-21 cells and cultured for 48 h. Cells were fixed with cold absolute ethanol and labeled with antibodies against HN and HRP-conjugated anti-mouse secondary antibody (Abbkine, Wuhan, China). AEC (Solarbio, China) staining was performed, and the number of lesions was quantified using Image J.

To assess virus release inhibition, BHK-21 cells were pre-infected with 100/well $TCID_{50}$ of the virus for 2 h. Serially diluted immune serum or purified antibodies were added to the wells, and cells were cultured for 48 hours. Cells were fixed, labeled with antibodies against HN and HRP-conjugated anti-mouse secondary antibody, and subjected to AEC staining. Lesion quantification was performed using Image J.

Statistical Analyses

Statistically significant differences between groups in the animal experiments were determined by pairwise analysis of variance (ANOVA) comparisons with Tukey's multicomparison test in GraphPad Prism version 7.00.

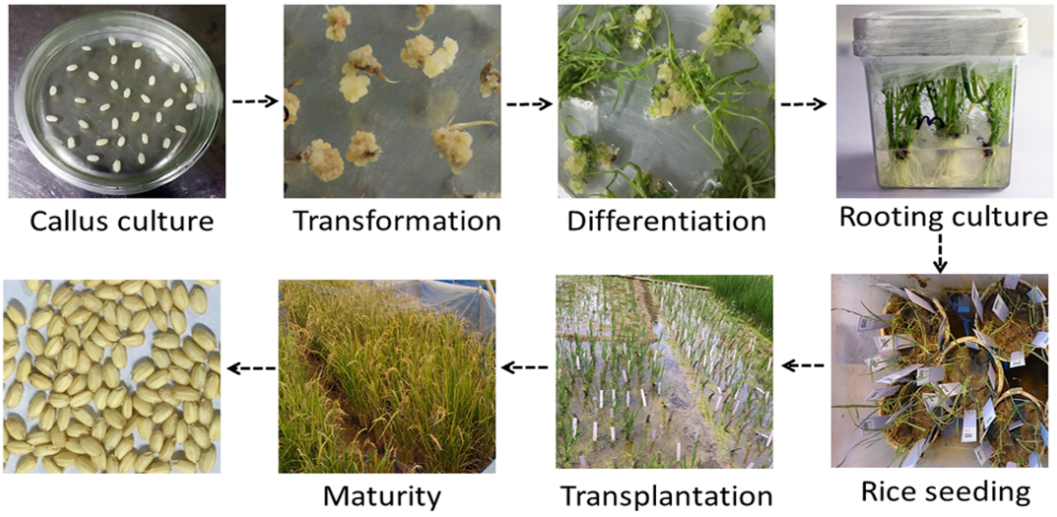

Fig. S1. An outline for the process of Osr2HN from transgenic rice. The process of obtaining Osr2HN from transgenic rice involves callus formation, Agrobacterium transformation, dark culture, hygromycin selection and differentiation, rooting, transferring plantlets to the soil, selecting positive transgenic rice plants, and large-scale cultivation for seed harvesting.

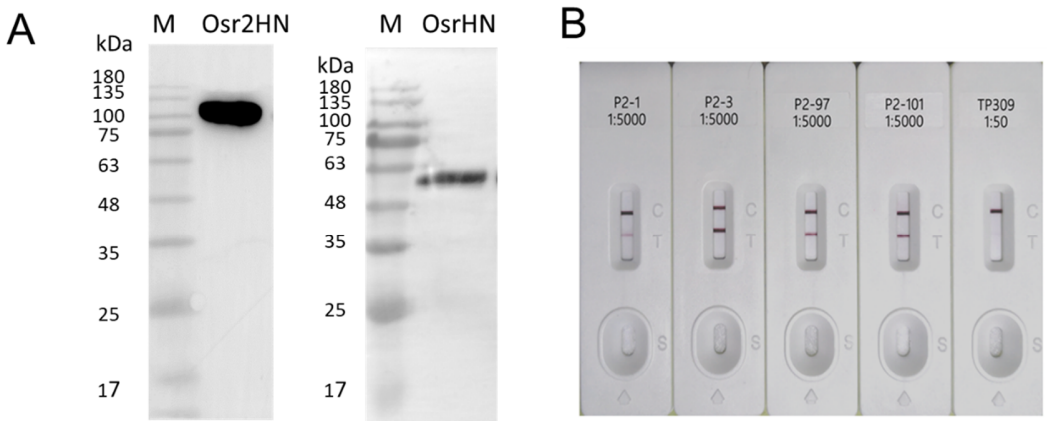

Fig. S2. (A) The Western blot result of Osr2HN and OsrHN. **(B)** Detection results of different Osr2HN transgenic lines by immunochromatographic strip.

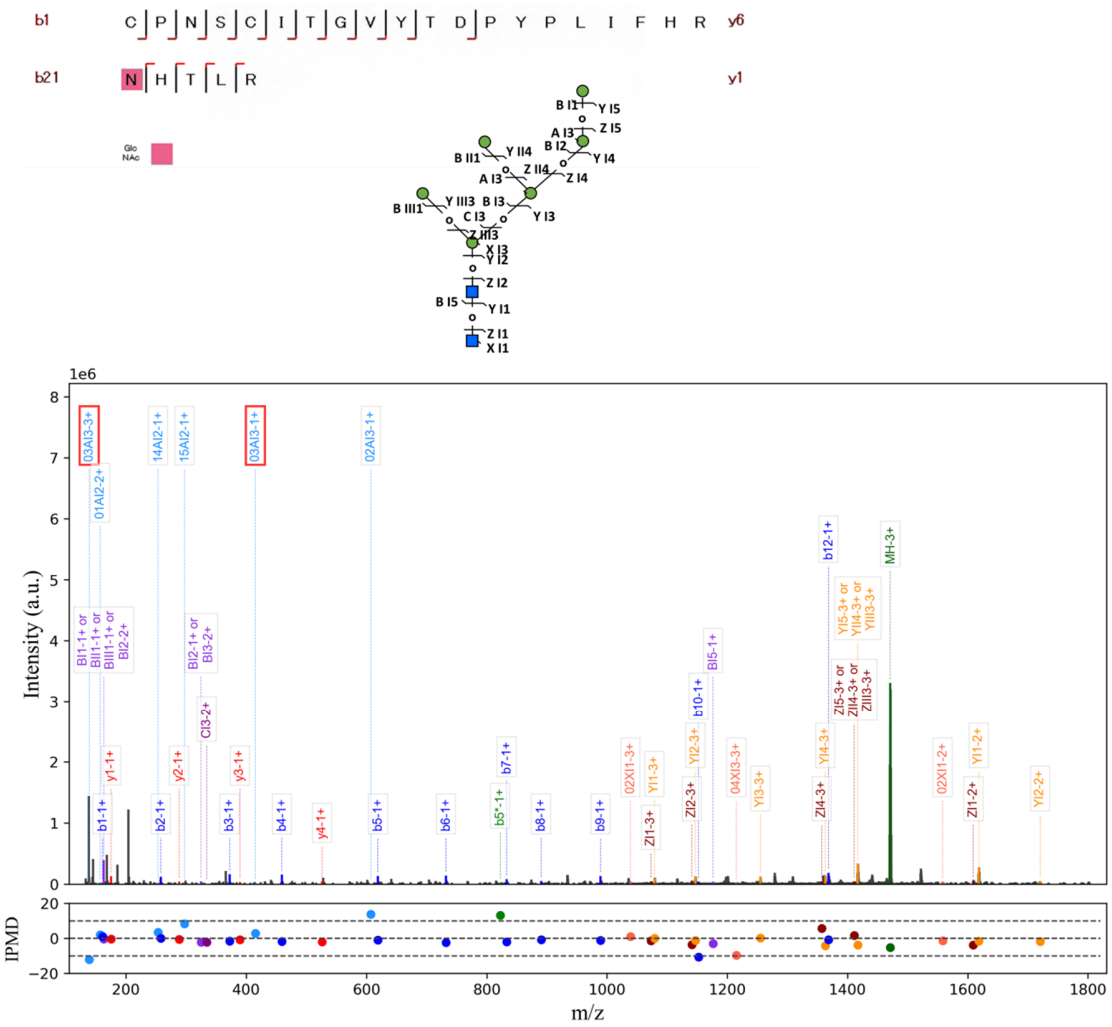

Fig. S3. The first type of glycan structure at Asn 481 and 1018.

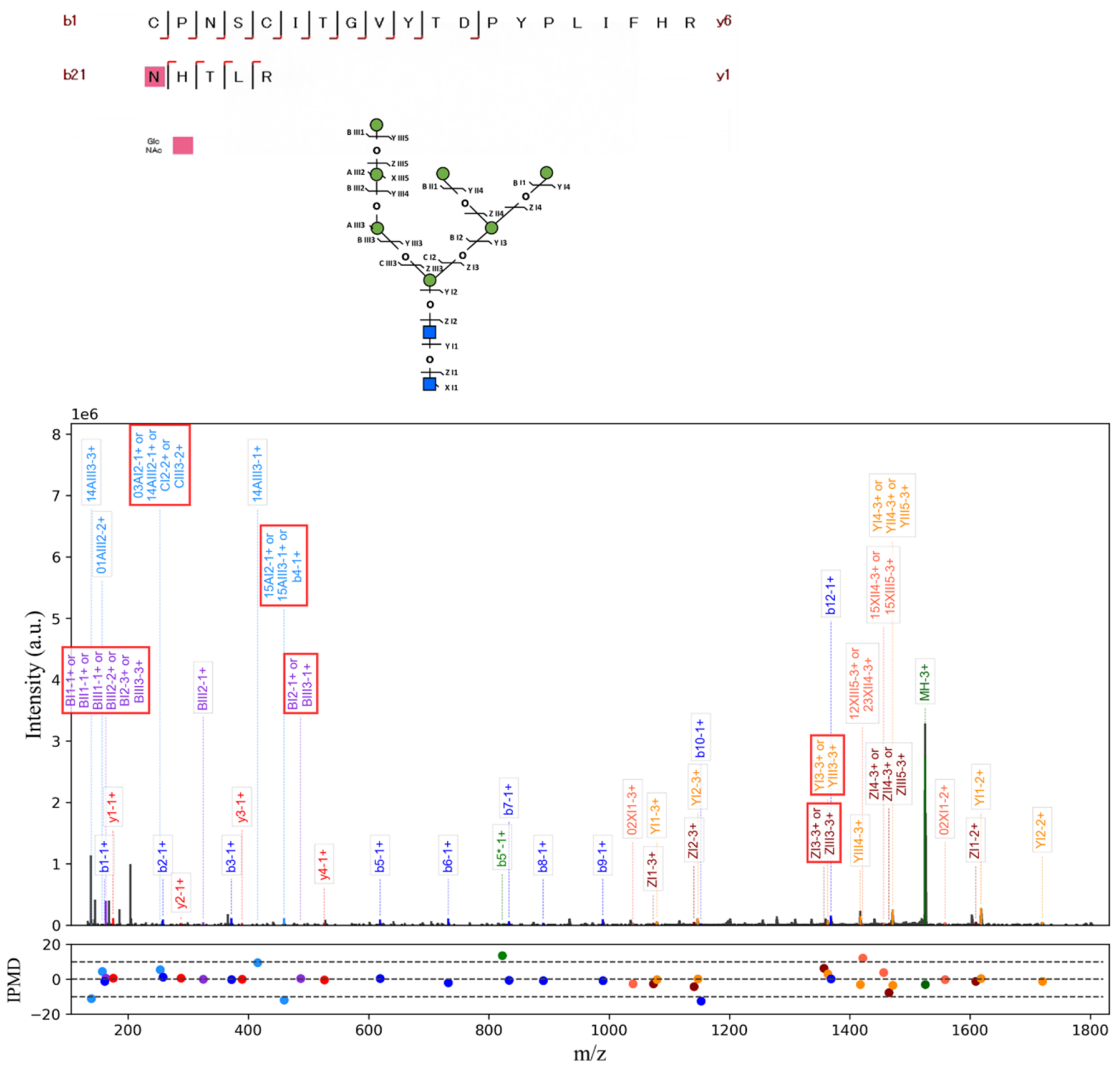

Fig. S4. The second type of glycan structure at Asn 481 and 1018.

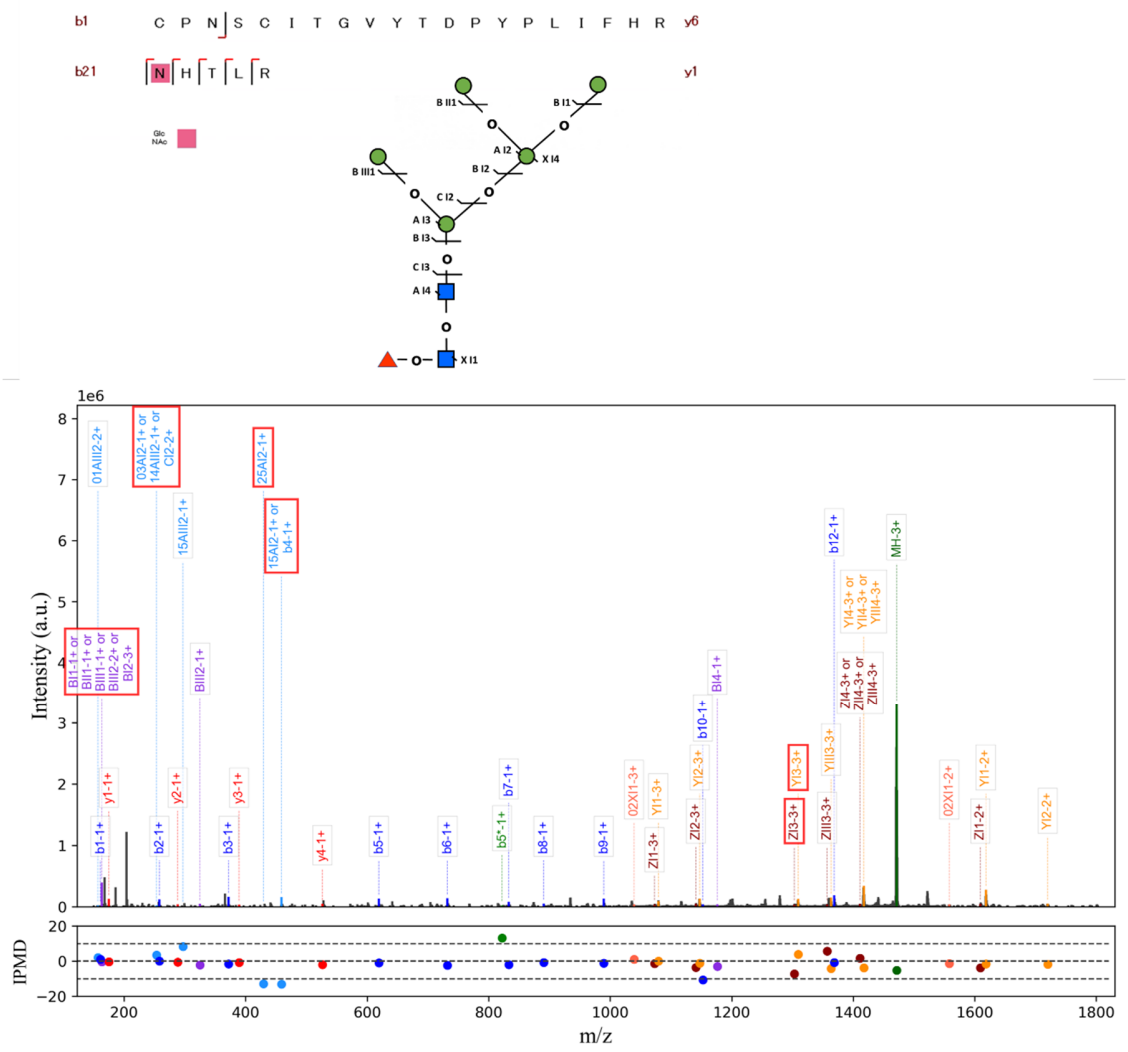

Fig. S5. The third type of glycan structure at Asn 481 and 1018.

Matched peptides shown in **Bold Red**

```
   1 STPHDLAGIS TVISKTEDKV TSLLSSSQDV IDRIYKQVAL ESPLALLNTE
  51 SIIMNAITSL SYQINGAANN SGCGAPVHDP DYIGGIGKEL IVDDISDVTS
 101 FYPSAYQEHL NFIPAPTTGS GCTRIPSFDM STTHYCYTHN VILSGCRDHS
 151 HSHQYLALGV LRTSATGRVF FSTLRSINLD DTQNRKSCSV SATPLGCDML
 201 CSKVTETEEE DYKSVAPTSM VHGRLGFDGQ YHEKDLDTTV LFKDWVANYP
 251 GVGGGSFIDD RVWFPVYGGL KPNSPSDTAQ EGKYVIYKRH NNTCPDEQDY
 301 QIRMAKSSYK PGRFGGKRVQ QAILSIKVST SLGKDPVLTI PPNTITLMGA
 351 EGRILTVGTS HFLYQRGSSY FSPALLYPMT VNNKTATLHS PYTFNAFTRP
 401 GSVPCQASAR CPNSCITGVY TDPYPLIFHR NHTLRGVFGT MLDDEQARLN
 451 PVSAVFDNIS RSRVTRVSSS STKAAYTTST CFKVVKTNKA YCLSIAEISN
 501 TLFGEFRIVP LLVEILKDDR VGGGGSGGGG SGGGGSSTPH DLAGISTVIS
 551 KTEDKVTSLL SSSQDVIDRI YKQVALESPL ALLNTESIIM NAITSLSYQI
 601 NGAANNSGCG APVHDPDYIG GIGKELIVDD ISDVTSFYPS AYQEHLNFIP
 651 APTTGSGCTR IPSFDMSTTH YCYTHNVILS GCRDHSHSHQ YLALGVLRTS
 701 ATGRVFFSTL RSINLDDTQN RKSCSVSATP LGCDMLCSKV TETEEEDYKS
 751 VAPTSMVHGR LGFDGQYHEK DLDTTVLFKD WVANYPGVGG GSFIDDRVWF
 801 PVYGGLKPNS PSDTAQEGKY VIYKRHNNTC PDEQDYQIRM AKSSYKPGRF
 851 GGKRVQQAIL SIKVSTSLGK DPVLTIPPNT ITLMGAEGRI LTVGTSHFLY
 901 QRGSSYFSPA LLYPMTVNNK TATLHSPYTF NAFTRPGSVP CQASARCPNS
 951 CITGVYTDPY PLIFHRNHTL RGVFGTMLDD EQARLNPVSA VFDNISRSRV
1001 TRVSSSSTKA AYTTSTCFKV VKTNKAYCLS IAEISNTLFG EFRIVPLLVE
1051 ILKDDRV
```

Fig. S6. Results of the LC-MS/MS in-gel tryptic digestion covering 83% of the HN protein sequence and flexible linker sequence (red sequence).

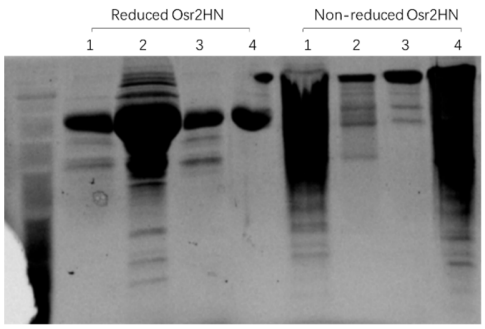

Fig. S7. The SDS-PAGE of reduced and non-reduced Osr2HN in different purifications.

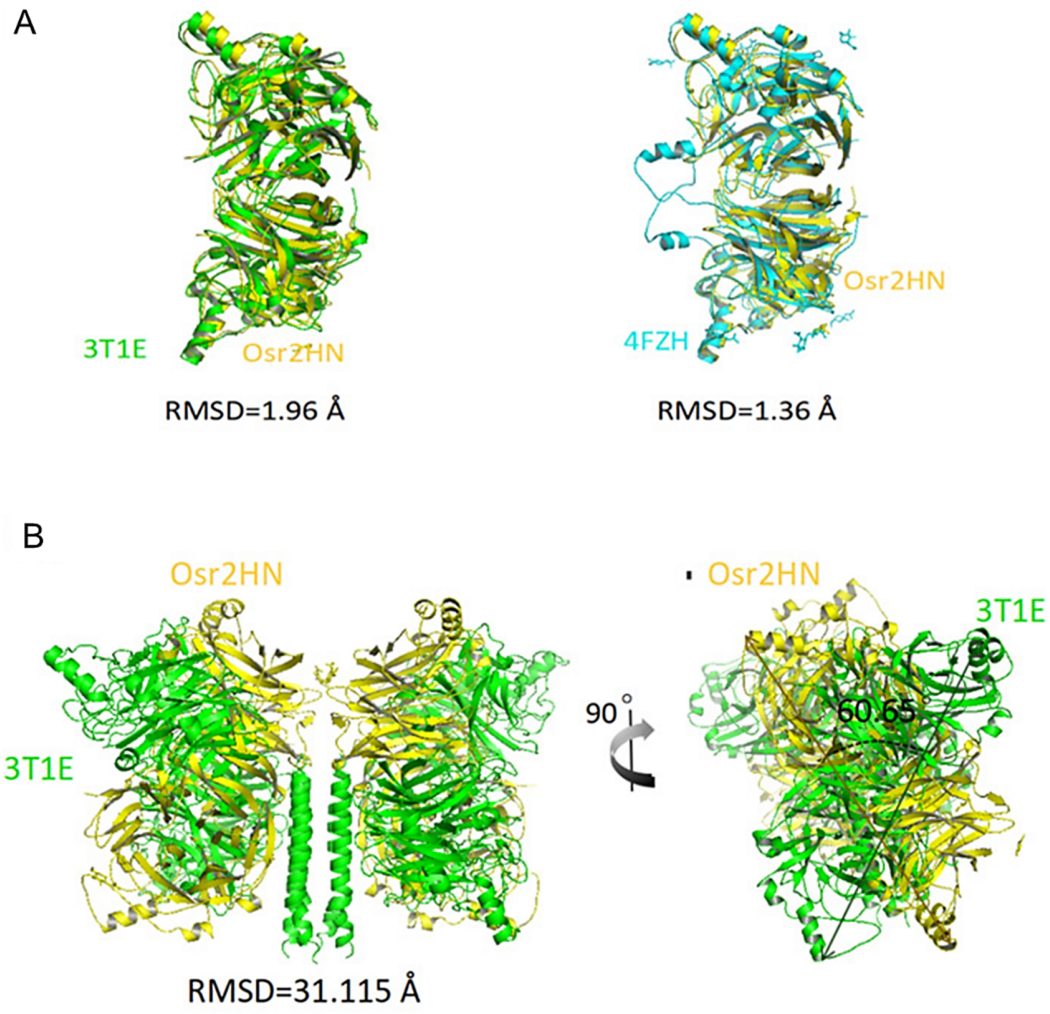

Fig. S8. Comparison of the structure of Osr2HN with the HN expressed in other expression systems. Superposition of Osr2HN (yellow) with HN from insect cell (3t1e with green) and mammalian cell expression system (4fzh with cyan) in the forms of HN dimer (**A**) and tetramer (**B**).

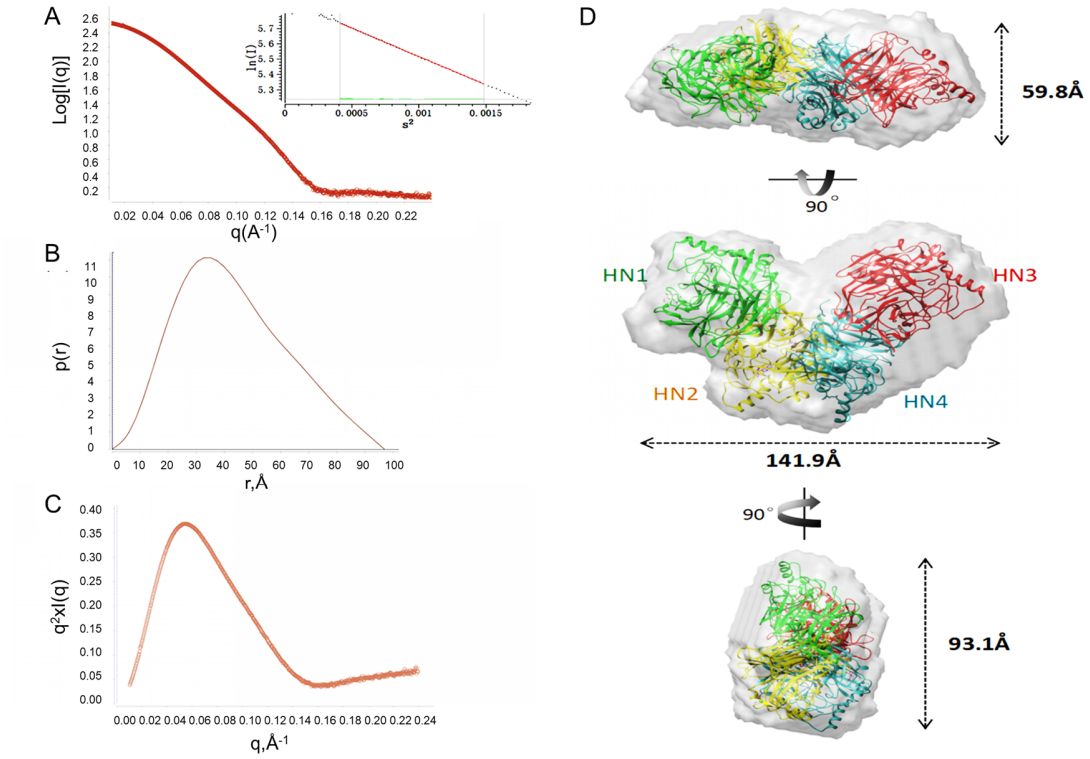

Fig. S9. SAXS data and the model of Ors2HN reveal the 2HN dimer in the solution state. (**A**)The Guinier region and the corresponding linear fitting are shown in the inset. (**B**) Pair-distance distribution (P(r)) function curves from SAXS. (**C**) Kratky plot calculated by Gnom from the experimental data. (**D**) Superposition of the envelope obtained by SAXS onto the crystallographic model of Osr2HN.

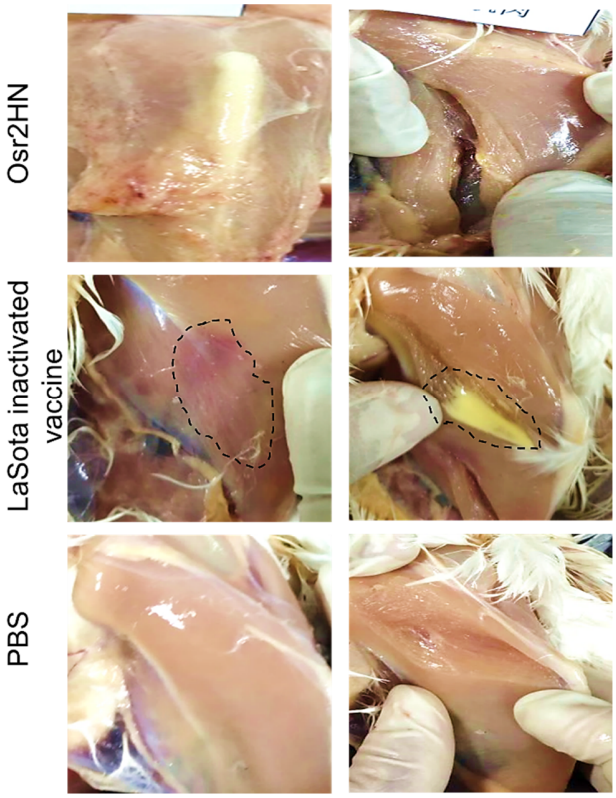

Fig. S10. Safety experiments were conducted using a high dose. Twenty-day-old chickens were individually administered 1 mL of Osr2HN vaccine (25 μg) and 1 mL of LaSota inactivated vaccine. After a period of 21 days, the chickens were dissected for further analysis.

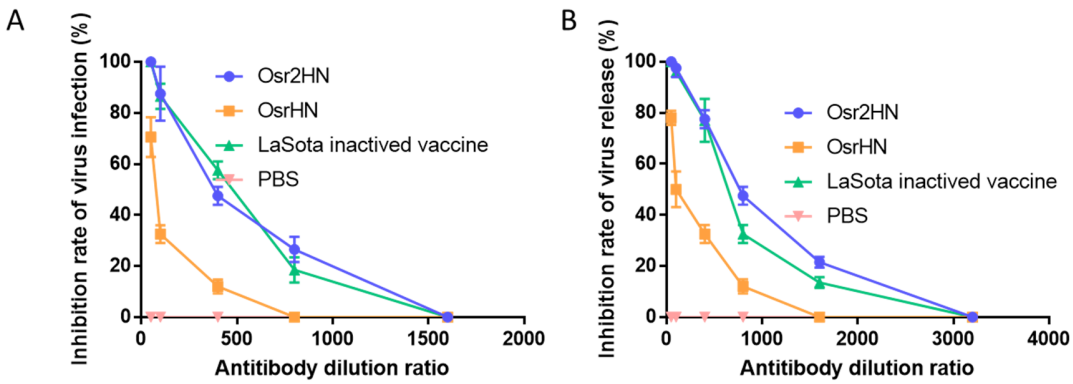

Fig. S11. The neutralizing effects of different immune sera during virus entry and release from the cells. (A) Inhibitory effects of antibodies induced by three vaccines on virus infection. **(B)** Inhibitory effects of antibodies induced by three vaccines on virus release.

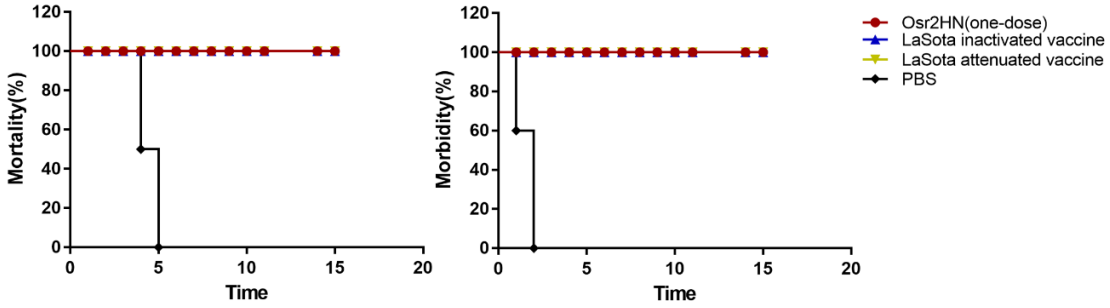

Fig. S12. Morbidity and mortality curves of one-dose immunization after challenge with wild-type NDV. None of the immune groups had clinical symptoms. In the PBS group, all of the chickens died.

Table S1. Crystallographic data collection and refinement statistics for the crystal structure of Osr2HN

Parameter	Value(s) for Osr2HN
PDB code	7BWU
X-ray source wavelength (Å)	0.97918
Resolution range (Å)	39.86–1.87 (1.937–1.87)
Space group	21 21 21
Unit cell	83.88 142.07 191.21 90 90 90
Unique reflections	188495 (18656)
Multiplicity	13.5
Completeness (%)	99.93 (99.96)
Mean I/sigma(I)	14.6
Wilson B-factor	29.59
R-merge	0.114
R-meas	0.118
R-pim	0.045
CC1/2	0.999
Reflections used in refinement	188465 (18652)
Reflections used for R-free	9289 (908)
R-work	0.2235 (0.3033)
R-free	0.2359 (0.3392)
Number of non-hydrogen atoms	13888
Macromolecules	13804
Ligands	84
Protein residues	1779
RMS (bonds)	0.012
RMS (angles)	1.47
Ramachandran favored (%)	94.52

Ramachandran allowed (%)	5.36
Ramachandran outliers (%)	0.11
Rotamer outliers (%)	0.26
Clash score	3.07
Average B-factor	38.80
Macromolecules	38.68
Ligands	58.23

Statistics for the highest-resolution shell are shown in parentheses.

Table S2. The sequences of the HN peptide library

Number	Sequence	Number	Sequence	Number	Sequence
P1	CMDRVVSQVALENDERE	P32	CHPHSHQYLALGVLRTS	P63	CEDPVLTVPPNTITLMG
P2	CQVALENDEREAKNTWR	P33	CYLALGVLRTSATGRVF	P64	CVPPNTITLMGAEGRVL
P3	CDEREAKNTWRLVFRTA	P34	CLRTSATGRVFFSTLRS	P65	CTLMGAEGRVLTVGTSH
P4	CNTWRLVFRTAVILLIV	P35	CGRVFFSTLRSINLDDN	P66	CGRVLTVGTSHFFYQRG
P5	CFSISAAALMYSMEAST	P36	CTLRSINLDDNQNRKS	P67	CGTSHFFYQRGSSYFSP
P6	CALMYSMEASTPGDLVG	P37	LDDNQNRKSCSVSATP	P68	CYQRGSSYFSPALLYPM
P7	CEASTPGDLVGILTAIS	P38	RKSCSVSATPLGCDML	P69	CYFSPALLYPMTVDNKT
P8	CDLVGILTAISRAEEKI	P39	SATPLGCDMLCSKVTE	P70	CLYPMTVDNKTATLHSP
P9	CTAISRAEEKITSALGS	P40	DMLCSKVTETEEEDY	P71	CDNKTATLHSPYAFNAF
P10	CEEKITSALGSNQDVVD	P41	CKVTETEEEDYNSVIPT	P72	CLHSPYAFNAFTRPGSV
P11	CALGSNQDVVDRIYKQV	P42	CEEDYNSVIPTPMVHGR	P73	FNAFTRPGSVPCQASA
P12	CDVVDRIYKQVALESPL	P43	CVIPTPMVHGRLGFDGQ	P74	PGSVPCQASARCPNS
P13	CYKQVALESPLALLNTE	P44	CVHGRLGFDGQYHEKDL	P75	QASARCPNSCVTGVYT
P14	CESPLALLNTESIIMSA	P45	CFDGQYHEKDLDVATLF	P76	PNSCVTGVYTDPYPLV
P15	CLNTESIIMSAITSLSY	P46	CEKDLDVATLFGDWVAN	P77	CGVYTDPYPLVFHRNHT
P16	CIMSAITSLSYQINRAA	P47	CATLFGDWVANYPGVGG	P78	CYPLVFHRNHTLRGVFG
P17	SLSYQINRAANNSGCG	P48	CWVANYPGVGGGSFIDN	P79	CRNHTLRGVFGTMLDDK
P18	NRAANNSGCGAPVHDP	P49	CGVGGGSFIDNRFWFPV	P80	CGVFGTMLDDKQARLNP
P19	SGCGAPVHDPDYIGGI	P50	CFIDNRVWFPVYGGLKP	P81	CLDDKQARLNPVSAVFD
P20	CVHDPDYIGGIGKELIV	P51	CWFPVYGGLKPNSPSDT	P82	CRLNPVSAVFDNISRSR
P21	CIGGIGKELIVDDASDV	P52	CGLKPNSPSDTAQEGRY	P83	CAVFDNISRSRITRVSS
P22	CELIVDDASDVTSFYPS	P53	CPSDTAQEGRYVIYKRY	P84	CSRSRITRVSSSSTKAA
P23	CASDVTSFYPSASQEHL	P54	EGRYVIYKRYNDTCPD	P85	CRVSSSSTKAAYTTST
P24	CFYPSASQEHLNFIPAP	P55	YKRYNDTCPDEQDYQI	P86	TKAAYTTSTCFKVVKT
P25	CQEHLNFIPAPTTGSG	P56	TCPDEQDYQIRMAKSS	P87	TSTCFKVVKTNKTYCL
P26	IPAPTTGSGCTRIPSF	P57	CDYQIRMAKSSYKPGRF	P88	VVKTNKTYCLSIAEIS
P27	GSGCTRIPSFDMSATH	P58	CAKSSYKPGRFGGKRVQ	P89	TYCLSIAEISNTLFGE

P28	IPSFDMSATHYCYTHN	P59	CPGRFGGKRVQQAILSI	P90	CAEISNTLFGEFRIVPL
P29	SATHYCYTHNVILSG	P60	CKRVQQAILSIKVSTSL	P91	CLFGEFRIVPLLVEILK
P30	YTHNVILSGCRDHPHS	P61	CILSIKVSTSLGEDPVL	P92	CFRIVPLLVEILKDDGI
P31	LSGCRDHPHSHQYLAL	P62	CSTSLGEDPVLTVPPNT		

Reference

1. He Y*, et al.* (2011) Large-scale production of functional human serum albumin from transgenic rice seeds. *Proc Natl Acad Sci U S A* 108(47):19078-19083.
2. Li Q*, et al.* (2019) Evaluation of an immunochromatographic strip for detection of avian avulavirus 1 (Newcastle disease virus). *J Vet Diagn Invest* 31(3):475-480.
3. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307-326.
4. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.
5. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, & Svergun DI (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography* 36(5):1277-1282.
6. Petoukhov MV & Svergun DI (2015) Ambiguity assessment of small-angle scattering curves from monodisperse systems. *Acta Crystallogr D Biol Crystallogr* 71(Pt 5):1051-1058.
7. Franke D & Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* 42(Pt 2):342-346.
8. Pettersen EF*, et al.* (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605-1612.
9. Zhang Q, Ma C, Chin LS, & Li L (2020) Integrative glycoproteomics reveals protein N-glycosylation aberrations and glycoproteomic network alterations in Alzheimer's disease. *Sci Adv* 6(40).