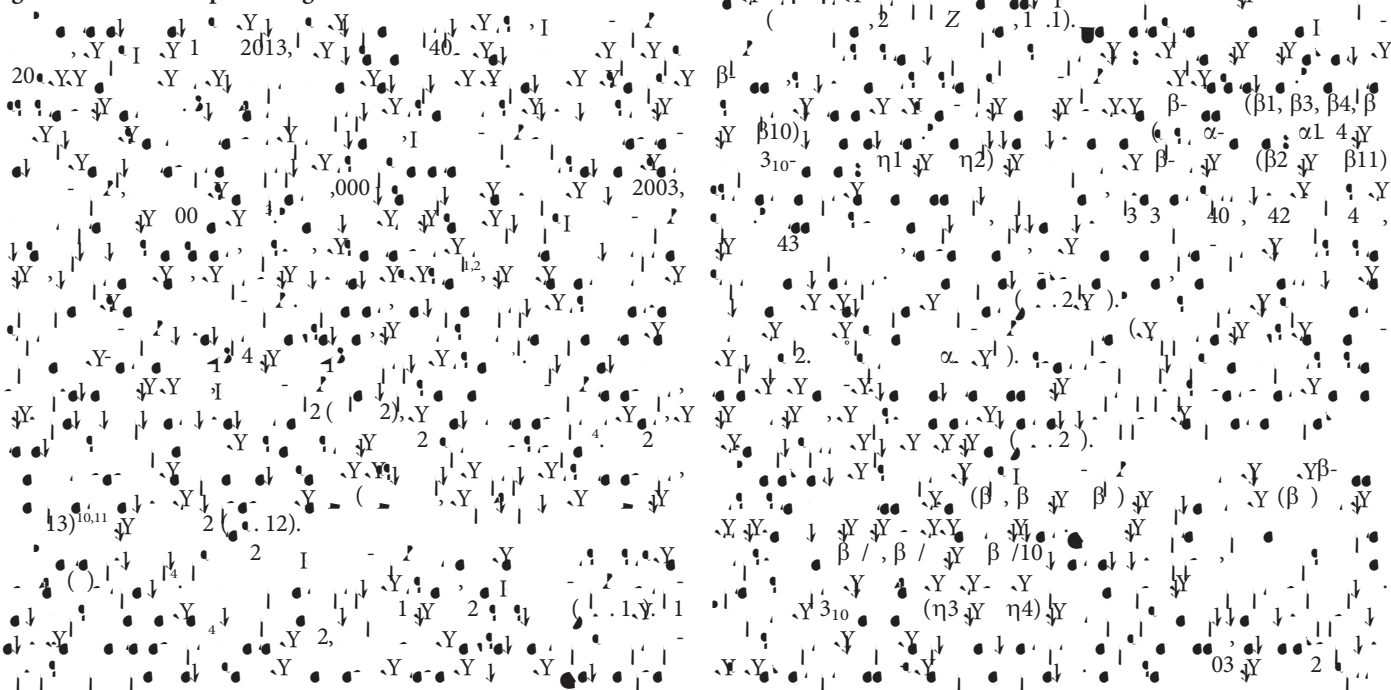


# Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26

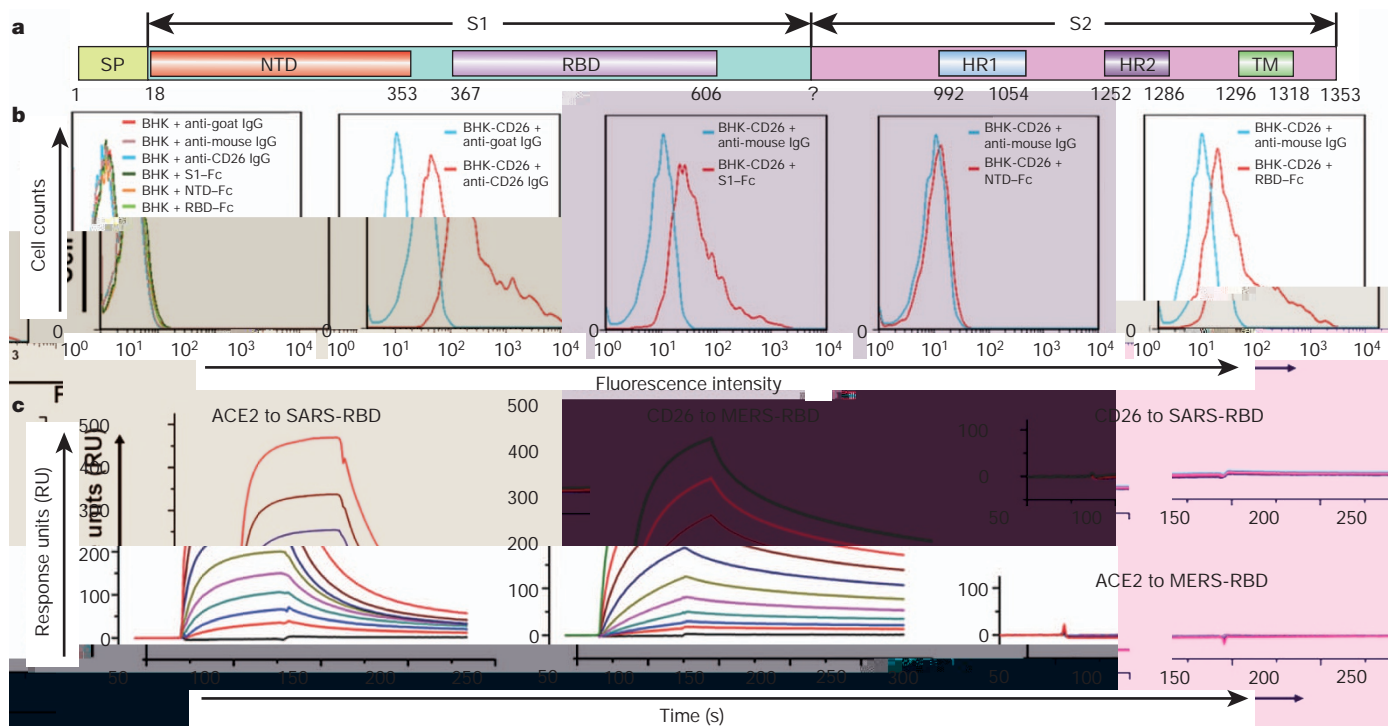
Guangwen Lu<sup>1\*</sup>, Yawei Hu<sup>2\*</sup>, Qihui Wang<sup>1\*</sup>, Jianxun Qi<sup>1\*</sup>, Feng Gao<sup>3,4\*</sup>, Yan Li<sup>1</sup>, Yanfang Zhang<sup>1,5</sup>, Wei Zhang<sup>1</sup>, Yuan Yuan<sup>1,6</sup>, Jinku Bao<sup>4</sup>, Buchang Zhang<sup>2</sup>, Yi Shi<sup>7</sup>, Jinghua Yan<sup>1</sup> & George F. Gao<sup>1,5,6,7,8</sup>

The newly emergent Middle East respiratory syndrome coronavirus (MERS-CoV) can cause severe pulmonary disease in humans<sup>1,2</sup>, representing the second example of a highly pathogenic coronavirus, the first being SARS-CoV<sup>3</sup>. CD26 (also known as dipeptidyl peptidase 4, DPP4) was recently identified as the cellular receptor for MERS-CoV<sup>4</sup>. The engagement of the MERS-CoV spike protein with CD26 mediates viral attachment to host cells and virus–cell fusion, thereby initiating infection. Here we delineate the molecular basis of this specific interaction by presenting the first crystal structures of both the free receptor binding domain (RBD) of the MERS-CoV spike protein and its complex with CD26. Furthermore, binding between the RBD and CD26 is measured using real-time surface plasmon resonance with a dissociation constant of 16.7 nM. The viral RBD is composed of a core subdomain homologous to that of the SARS-CoV spike protein, and a unique strand-dominated external receptor binding motif that recognizes blades IV and V of the CD26  $\beta$ -propeller. The atomic details at the interface between the two binding entities reveal a surprising protein–protein contact mediated mainly by hydrophilic residues. Sequence alignment indicates, among betacoronaviruses, a possible structural conservation for the region homologous to the MERS-CoV RBD core, but a high variation in the external receptor binding motif region for virus-specific pathogenesis such as receptor recognition.

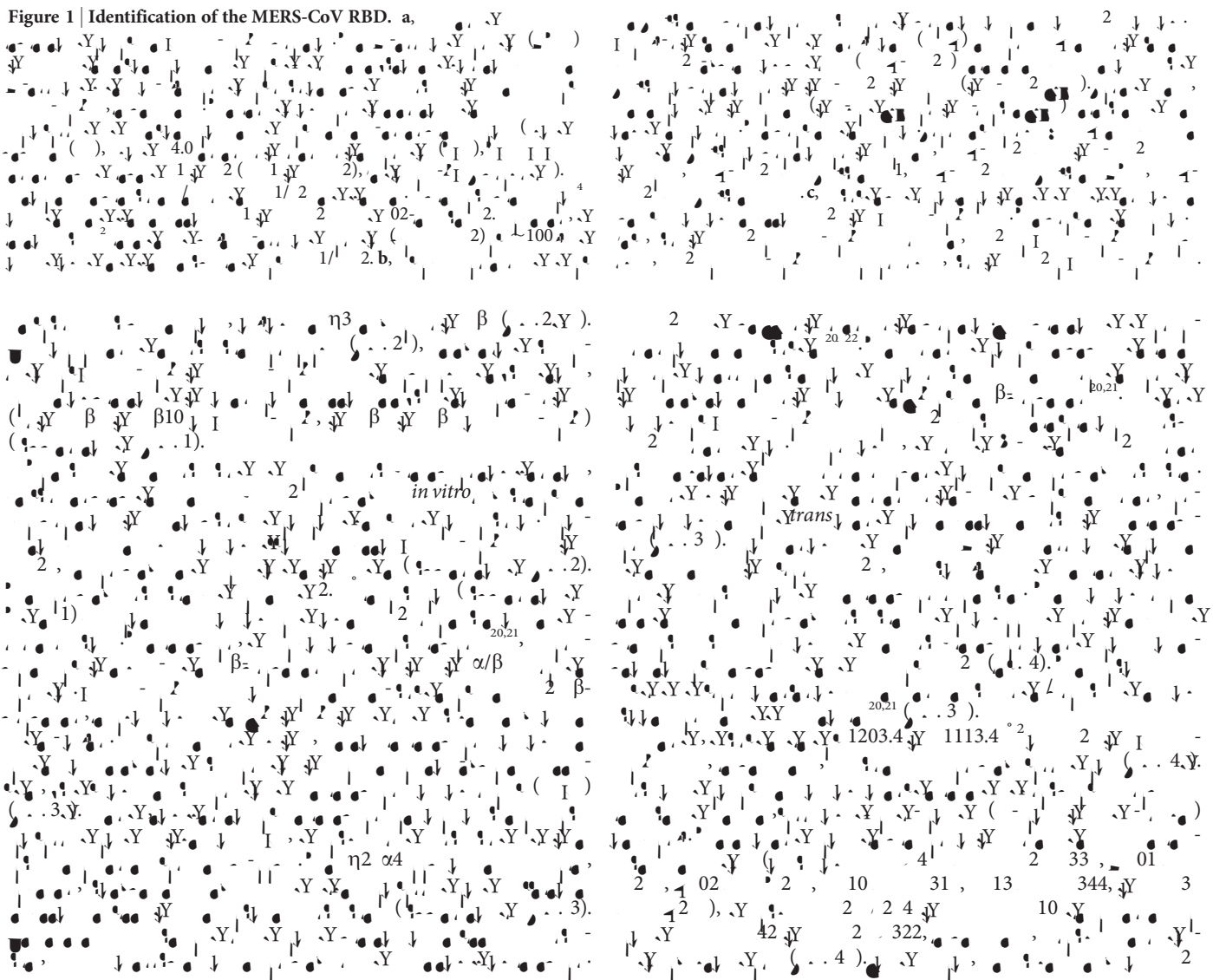


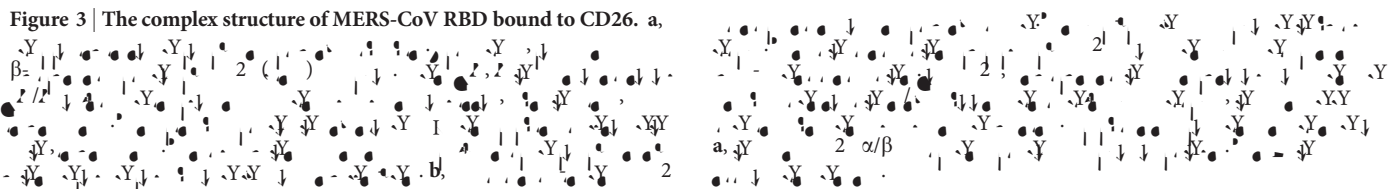
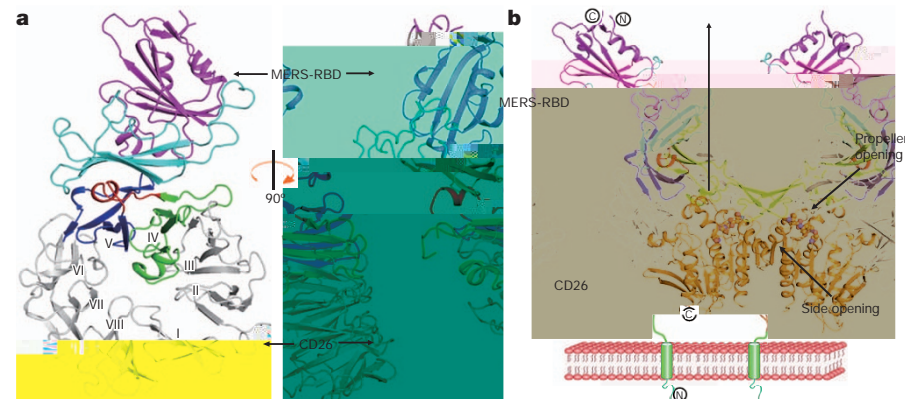
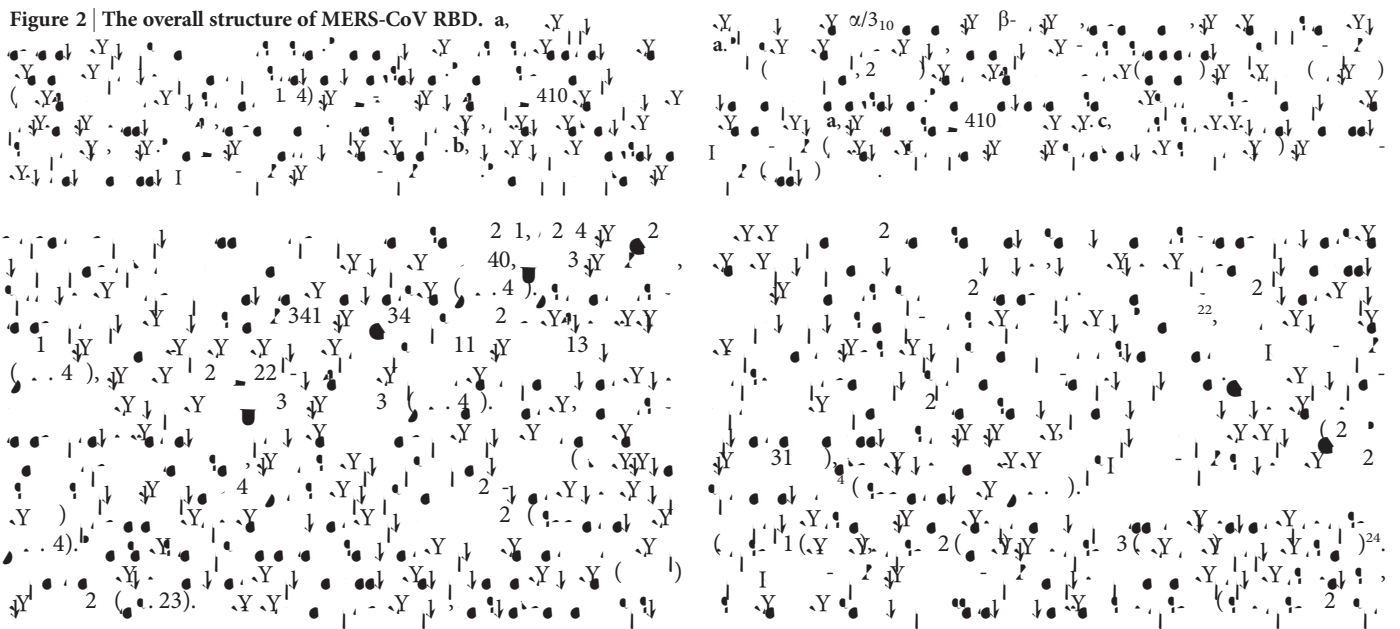
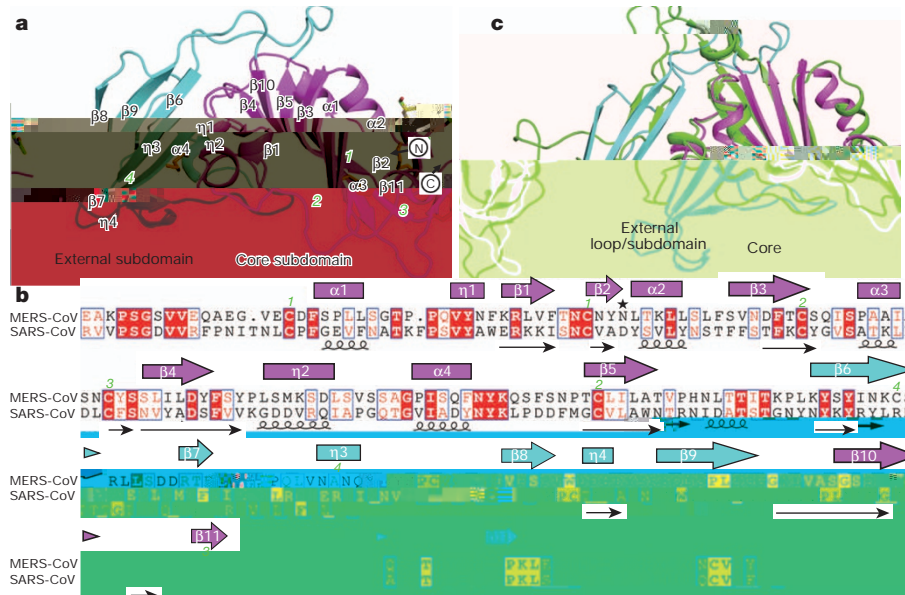
<sup>1</sup>CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China. <sup>2</sup>School of Life Sciences, Anhui University, Hefei 230039, China. <sup>3</sup>Laboratory of Non-coding RNA, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China. <sup>4</sup>School of Life Sciences, Sichuan University, Chengdu 610064, Sichuan, China. <sup>5</sup>Laboratory of Protein Engineering and Vaccines, Tianjin Institute of Industrial Biotechnology, Tianjin 300308, China. <sup>6</sup>School of Life Sciences, University of Science and Technology of China, Hefei 230026, China. <sup>7</sup>Research Network of Immunity and Health (RNIH), Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China. <sup>8</sup>Chinese Center for Disease Control and Prevention (China CDC), Beijing 102206, China.

\*These authors contributed equally to this work.



**Figure 1 | Identification of the MERS-CoV RBD. a,**







20. Rasmussen, H. B., Branner, S., Wiberg, F. C. & Wagtmann, N. Crystal structure of human dipeptidyl peptidase IV/CD26 in complex with a substrate analog. *Nature Struct. Biol.* **10**, 19–25 (2003).
21. Engel, M. *et al.* The crystal structure of dipeptidyl peptidase IV (CD26) reveals its functional regulation and enzymatic mechanism. *Proc. Natl Acad. Sci. USA* **100**, 5063–5068 (2003).
22. Gorrell, M. D., Gysbers, V. & McCaughan, G. W. CD26: a multifunctional integral membrane and secreted protein of activated lymphocytes. *Scand. J. Immunol.* **54**, 249–264 (2001).
23. Weihofen, W. A., Liu, J., Reutter, W., Saenger, W. & Fan, H. Crystal structure of CD26/dipeptidyl-peptidase IV in complex with adenosine deaminase reveals a highly amphiphilic interface. *J. Biol. Chem.* **279**, 43330–43335 (2004).
24. Lai, M. M., Perlman, S. & Anderson, L. J. in *Fields Virology* (ed. Knipe, D.M.) 1305–1336 (Lippincott Williams & Wilkins, 2007).
25. Reguera, J. *et al.* Structural bases of coronavirus attachment to host aminopeptidase N and its inhibition by neutralizing antibodies. *PLoS Pathog.* **8**, e1002859 (2012).
26. Li, F. Evidence for a common evolutionary origin of coronavirus spike protein receptor-binding subunits. *J. Virol.* **86**, 2856–2858 (2012).
27. Du, L. *et al.* The spike protein of SARS-CoV—a target for vaccine and therapeutic development. *Nature Rev. Microbiol.* **7**, 226–236 (2009).
28. Gierer, S. *et al.* The spike-protein of the emerging betacoronavirus EMC uses a novel coronavirus receptor for entry, can be activated by TMPRSS2, and is targeted by neutralizing antibodies. *J. Virol.* **87**, 5502–5511 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by the Ministry of Science and Technology of China (MOST) 973 Project (Grant no. 2011CB504703) and the National Natural Science Foundation of China (NSFC, Grant no. 81290342). Assistance by the staff at the Shanghai Synchrotron Radiation Facility (SSRF) of China and the High Energy Accelerator Research Organization (KEK) of Japan is acknowledged. We thank Z. Fan and T. Zhao for their technical assistance. G.F.G. is a leading principal investigator of the NSFC Innovative Research Group (Grant no. 81021003). We thank M. Yang from Tsinghua University for his help with data collection.

**Author Contributions** G.F.G. designed and coordinated the study. G.L., Y.H., Q.W. and Y.S. conducted the experiments. J.Q. and F.G. collected the data sets and solved the structures. Y.L., Y.Z., W.Z., Y.Y. and J.Y. assisted with the cell maintenance and protein preparations. G.L. and G.F.G. wrote the manuscript and J.Y., J.B. and B.Z. participated in the manuscript editing and discussion.

**Author Information** The coordinates and related structure factors have been deposited into the Protein Data Bank PDB under accession numbers 4KQZ for the free MERS-CoV RBD structure and 4KR0 for the RBD–CD26 complex structure. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.F.G. (gaof@im.ac.cn).



## METHODS

## Protein expression and purification.

Protein expression and purification. The recombinant protein was expressed in *E. coli* BL21(DE3) cells using a pET28a(+) vector. The cells were grown in LB medium supplemented with 0.5% glucose and 100 µg/ml ampicillin. After reaching an optical density of 0.6 at 600 nm, the cells were induced with 1 mM IPTG. The culture was harvested after 16 h of induction. The cells were lysed by sonication in a lysis buffer containing 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 10 mM imidazole, and 1 mM EDTA. The lysate was centrifuged at 14,000g for 10 min. The supernatant was then loaded onto a Ni-NTA agarose column. The protein was eluted with a linear gradient of imidazole from 10 to 500 mM. The purified protein was dialyzed into a storage buffer containing 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 1 mM EDTA.

*In vitro* assays. The protein was assayed for activity using a standard assay. The reaction mixture contained 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT, and 1 mM substrate. The reaction was initiated by the addition of the protein. The absorbance was measured at 405 nm. The initial rate of the reaction was determined from the linear portion of the time course. The protein concentration was determined using a Bradford assay.

## Analytical gel filtration.

Analytical gel filtration. The protein was purified by size exclusion chromatography using a Superdex 200 column. The column was equilibrated with a buffer containing 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 1 mM EDTA. The protein was eluted with the same buffer. The elution volume was determined by monitoring the absorbance at 280 nm.

## Surface plasmon resonance assay.

Surface plasmon resonance assay. The protein was immobilized on a CM5 sensor chip using amine coupling. The chip was then injected with a solution of the protein. The binding was monitored by the change in refractive index. The binding affinity was determined from the dissociation constant (K<sub>d</sub>) using a 1:1 binding model. The data were fitted using the 1:1 binding model in the BIACore software.

## Flow cytometric assay.

Flow cytometric assay. The protein was labeled with a fluorescent dye. The cells were stained with the labeled protein. The fluorescence was measured using a flow cytometer. The data were analyzed using a histogram. The percentage of cells with fluorescence was determined. The protein concentration was determined using a Bradford assay.

## Crystallization.

Crystallization. The protein was crystallized using a sitting drop vapor diffusion experiment. The droplets were set up in a 96-well plate. The protein concentration was 10 mg/ml. The precipitant concentration was 1.0 M. The temperature was 20 °C. The crystals were harvested after 1 week. The crystals were then flash-frozen in liquid nitrogen.

## Data collection, integration and structure determination.

Data collection, integration and structure determination. The data were collected using a synchrotron X-ray source. The wavelength was 1.03 Å. The data were integrated using XDS. The structure was determined using PHENIX. The resolution was 1.4 Å. The R-factor was 0.18. The R-free factor was 0.22. The protein structure was refined using PHENIX. The final model was deposited in the Protein Data Bank.

## Secondary-structure determination.

Secondary-structure determination. The protein structure was determined using X-ray crystallography. The structure was refined using PHENIX. The final model was deposited in the Protein Data Bank.

29. Zhang, W. *et al.* Crystal structure of the swine-origin A (H1N1)-2009 influenza A virus hemagglutinin (HA) reveals similar antigenicity to that of the 1918 pandemic virus. *Protein Cell* **1**, 459–467 (2010).
30. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
31. Collaborative Computing Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
32. Usón, I. & Sheldrick, G. M. Advances in direct methods for protein crystallography. *Curr. Opin. Struct. Biol.* **9**, 643–648 (1999).
33. Read, R. J. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr. D* **57**, 1373–1382 (2001).
34. Cowtan, K. D. & Zhang, K. Y. Density modification for macromolecular phase improvement. *Prog. Biophys. Mol. Biol.* **72**, 245–270 (1999).
35. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
36. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
37. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
38. Gouet, P., Courcelle, E., Stuart, D. I. & Metz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308 (1999).