Cell

Supplemental Information

Ebola Viral Glycoprotein Bound to Its Endosomal Receptor Niemann-Pick C1

Han Wang, Yi Shi, Jian Song, Jianxun Qi, Guangwen Lu, Jinghua Yan, and George F. Gao

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Protein expression and purification

The NPC1 domain C protein (NPC1-C) was expressed in *Escherichia coli* strain BL21 (DE3) as inclusion bodies and then refolded *in vitro* using the method as previously described (Shi et al., 2011), with some slight modifications. The cDNAs encoding residues 374-620 of NPC1 were cloned into pET21a vector using NdeI and XhoI sites without a stop condon which gave the protein a C terminal His_6-tag. The refolded NPC1-C protein was then concentrated and purified by gel filtration on a HiLoad 16/60 Superdex® 75 PG column (GE Healthcare). The NPC1-C mutants were constructed by site-directed mutagenesis kit (New England Biolabs), and then were expressed, refolded and purified following the procedures of wild-type NPC1-C protein.

The fully-glycosylated NPC1-C protein was expressed in 293T cells using the pCDNA4.0 vector, with the construct of an NGFR signal peptide, encoding residues 374-620 of NPC1 and a C terminal His6-tag. After 6 days of cell culture, the culture supernatant containing the protein was collected and purified by metal affinity chromatography using a HisTrap® HP 5-ml column (GE Healthcare), and then further purified by gel filtration on a HiLoad 16/60 Superdex®200 PG column (GE Healthcare).

The EBOV transmembrane-and-mucin-domains-removed GP (GP-ectoΔmucin) protein was expressed using the Bac-to-Bac baculovirus expression system (Invitrogen), using the method developed by Saphire's group (Lee et al., 2008), with some slight modifications. The gene sequence encoding the ectodomain of EBOV GP (GenBank: KJ660347.1) with the deletion of the native signal peptide, mucin-like and

transmembrane domains (residues 1-32, 312-463 and 633-676) and mutations of two N-linked glycosylation site (T42V and T230V) was cloned into the pFastBac1 vector (Invitrogen) as described by Saphire's group (Lee et al., 2008). For protein secretion and purification, a gp67 signal peptide and a His_6-tag were added at the N-terminus and C-terminus of the protein separately. Transfection and virus amplification were performed according to the Bac-to-Bac baculovirus expression system manual (Invitrogen) using Sf9 cells (Shi et al., 2013; Stevens et al., 2004; Zhang et al., 2010). Soluble GP-ectoΔmucin protein was produced in Hi5 cells after infected with high-titer recombinant baculovirus for 2 days and harvested from the culture supernatants by metal affinity chromatography using a HisTrap® HP 5-ml column (GE Healthcare), and whereafter purified by gel filtration on a HiLoad 16/60 Superdex®200 PG column (GE Healthcare).

To mimic endosomal protease cleavage and produce primed EBOV GP (GPcl) which was capable of NPC1 binding, 1 mg of EBOV GP-ectoΔmucin protein was incubated with 5 μg thermolysin (Sigma-aldrich) overnight at 18℃ in 20 mMTris (pH 8), 150 mM NaCl, 1 mM $CaCl_2$, and purified on a Superdex 200® column.

Preparation and crystallization of NPC1-C and NPC1-C-GPcl complex

All the crystals were obtained by using the sitting drop vapor diffusion method with 1μL protein mixing with 1 μL reservoir solution and then equilibrating against 100 μL reservoir solution at 18℃. Diffractive crystals of free NPC1-C protein were obtained in a condition consisting of 0.07 M sodium acetate trihydrate, pH 4.6, 5.6% (w/v)

polyethylene glycol 4,000 and 30% (v/v) glycerol at the protein concentration of 5mg/ml. Derivative crystals were obtained by soaking NPC1-C crystals overnight in mother liquor containing 1mM $KAuCl_4 \cdot xH_2O$.

To obtain the complex crystals of NPC1-C bound to GPcl, the individual proteins were in vitro mixed at a molar ratio of 1:1 and incubated at 4 ℃ for about 2 hours before concentrating and being used for crystallization. The complex crystals were initially formed in a commercially available kit, but diffracted to low resolutions, around 5.0 Å. The conditions that yield the better-diffractable crystals were then optimized. Eventually, diffractive crystals with the increasing resolution to 2.3 Å were obtained in a reservoir solution of 15 % (v/v) pentaerythritolpropoxylate, 0.2 M sodium chloride, pH 5.5, and 0.1 M MES-NaOH with a protein concentration of 10 mg/ml.

Data collection and structure determination

For data collection, all crystals were cryo-protected by briefly soaking in reservoir solution supplemented with 20% (v/v) glycerol before flash-cooling in liquid nitrogen. The native NPC1-C (wavelength, 0.979 Å) and Au derivative (wavelength, 1.038 Å) data sets were collected at the Shanghai Synchrotron Radiation Facility (SSRF) BL17U1, whereas the complex dataset (wavelength, 0.97930 Å) was collected at BL19U1. All the data sets were processed with HKL2000 software (Otwinowski and Minor, 1997).

The structure of NPC1-C was determined by the single-wavelength anomalous diffraction (SAD) method. The Au sites were first located by SHELXD (Uson and

Sheldrick, 1999) for the Au-SAD data. The identified position were then refined and the phases were calculated with SAD experimental phasing module of Phaser (Read, 2001). The real space constraints were further applied to the electron density map in DM (Cowtan and Zhang, 1999). After using Autobuild in Phenix package (Adams et al., 2010) to build the initial model, additional missing residues were added manually in COOT (Emsley and Cowtan, 2004). Rounds of refinement were performed using phenix.refine in Phenix (Adams et al., 2010) with energy minimization, isotropic ADP refinement, and bulk solvent modeling. The complex structure was solved by molecular replacement method using Phaser from the CCP4 program suite (Collaborative Computational Project, 1994)with the solved NPC1-C structure and previously reported GP structure (PDB code, 3CSY) as the search models. The atomic model was completed with Coot and refined with phenix.refine and the stereochemical qualities of the final models were assessed with PROCHECK (Laskowski et al., 1993). Data collection, processing and refinement statistics are summarized in Table S1. All structural figures were generated using Pymol software (http://www.pymol.org).

SPR Analysis

The SPR (surface plasmon resonance) analysis was performed using a BIAcore® 3000 machine with CM5 chips (GE Healthcare) at room temperature (25℃). All the proteins using in SPR analysis were exchanged to BIAcore® buffer, consisting of 10 mM HEPES, pH 7.5, 150 mM NaCl and 0.005% (v/v) Tween-20, via gel filtration. Around 2,000 response units of the GP or GPcl protein was immobilized on the CM5

chip and a blank channel was used as the negative control. The WT and mutant NPC1-C proteins or antibody 2G4 (which was used as a positive analyte for active GP-ectoΔmucin or GPcl protein) were serially diluted to the concentrations ranging from 6.25 to 400 μM or from 1.56 to 50 nM individually. The analytes were then used to flow over the chip surface with the response units measured. At the end of each cycle, the sensor surface was re-generated via a short treatment using 10 mM NaOH. The binding kinetics was analyzed with the software BIAevaluation® Version 4.1 using 1:1 Langmuir binding model.

Table S1. Crystallographic data collection and refinement statistics, Related to Figure 1 and 3.

	NPC1-C	Au derivative NPC1-C	GPcl-NPC1-C
Data collection			
Space group	P6122	P6122	P63
Cell dimensions			
a, *b*, *c* (Å)	60.01, 60.01, 270.35	59.95, 59.95, 270.26	107.26, 107.26, 93.78
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120
Wavelength (Å)	0.979	1.038	0.979
Resolution (Å)	50-2.00(2.07-2.00)	50-2.80(2.90-2.80)	50-2.30(2.38-2.30)
R_{merge}	0.123(0.501)	0.215(0.968)	0.102(0.758)
$I / \sigma I$	29.0(6.9)	32.1(8.3)	32.6(11.5)
Completeness (%)	99.9(99.7)	100.0(100.0)	100.0(100.0)
Redundancy	20.6(20.9)	54.5(55.8)	19.6(13.3)
Refinement			
Resolution (Å)	34.04-2.00		41.86-2.30
No. reflections	20604		27310
R_{work} / R_{free}	0.1821/0.2190		0.1837/0.2301
No. atoms			
Protein	1792		3657
Ligand/ion	0		0
Water	177		171
B-factors			
Protein	31.9		51.4
Ligand/ion			
Water	42.1		47.2
R.m.s. deviations			
Bond lengths (Å)	0.007		0.008
Bond angles (°)	1.030		1.152
Ramachandran analysis			
Most favoured (%)	87.2		85.0
Additional allowed (%)	12.3		13.2
Generally allowed (%)	0.5		1.8
Disallowed (%)	0		0

*Values in parentheses are for highest-resolution shell.

Supplemental references

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W*., et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica Section D, Biological crystallography *66*, 213-221.

Collaborative Computational Project, N. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr *50*, 760-763.

Cowtan, K.D., and Zhang, K.Y. (1999). Density modification for macromolecular phase improvement. Progress in biophysics and molecular biology *72*, 245-270.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta crystallographica Section D, Biological crystallography *60*, 2126-2132.

Laskowski, R.A., Macarthur, M.W., Moss, D.S., and Thornton, J.M. (1993). Procheck - a Program to Check the Stereochemical Quality of Protein Structures. J Appl Crystallogr *26*, 283-291.

Lee, J.E., Fusco, M.L., Hessell, A.J., Oswald, W.B., Burton, D.R., and Saphire, E.O. (2008). Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. Nature *454*, 177-182.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Method Enzymol *276*, 307-326.

Read, R.J. (2001). Pushing the boundaries of molecular replacement with maximum likelihood. Acta crystallographica Section D, Biological crystallography *57*, 1373-1382.

Shi, Y., Qi, J., Iwamoto, A., and Gao, G.F. (2011). Plasticity of human CD8alphaalpha binding to peptide-HLA-A*2402. Mol Immunol *48*, 2198-2202.

Shi, Y., Zhang, W., Wang, F., Qi, J., Wu, Y., Song, H., Gao, F., Bi, Y., Zhang, Y., Fan, Z*., et al.* (2013). Structures and receptor binding of hemagglutinins from human-infecting H7N9 influenza viruses. Science *342*, 243-247.

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.

Uson, I., and Sheldrick, G.M. (1999). Advances in direct methods for protein crystallography. Current opinion in structural biology *9*, 643-648.

Zhang, W., Qi, J., Shi, Y., Li, Q., Gao, F., Sun, Y., Lu, X., Lu, Q., Vavricka, C.J., Liu, D*., et al.* (2010). 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.