SUPPLEMENT ARTICLE

High-resolution Crystal Structure of Dimeric VP40 From *Sudan ebolavirus*

Matthew C. Clifton,[1,4,a] Jessica F. Bruhn,[5,a] Kateri Atkins,[1,4] Terry L. Webb,[1,3] Ruth O. Baydo,[1,3] Amy Raymond,[1,3] Donald D. Lorimer,[1,3] Thomas E. Edwards,[1,3] Peter J. Myler,[1,2] and Erica Ollmann Saphire[5,6]

[1]Seattle Structural Genomics Center for Infectious Disease, [2]Seattle Biomedical Research Institute, and [3]Beryllium, Bainbridge Island, Washington; [4]Beryllium, Bedford, Massachusetts; [5]Department of Immunology and Microbial Science, and [6]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California

Ebolaviruses **cause severe hemorrhagic fever. Central to the Ebola life cycle is the matrix protein VP40, which oligomerizes and drives viral budding. Here we present the crystal structure of the Sudan virus (SUDV) matrix protein. This structure is higher resolution (1.6 Å) than previously achievable. Despite differences in the protein purification, we find that it still forms a stable dimer in solution, as was noted for other Ebola VP40s. Although the N-terminal domain interface by which VP40 dimerizes is conserved between Ebola virus and SUDV, the C-terminal domain interface by which VP40 dimers may further assemble is significantly smaller in this SUDV assembly.**

Keywords. Ebola virus; *Sudan ebolavirus*; VP40; matrix protein; virus budding.

Ebolaviruses cause hemorrhagic fever in humans with up to 90% case fatality [1]. There are 5 species in the genus of *Ebolavirus*, including *Zaire ebolavirus* (Ebola virus; EBOV) and *Sudan ebolavirus* (Sudan virus; SUDV). EBOV is the causative agent of the largest and ongoing outbreak of Ebola hemorrhagic fever, in West Africa in 2014 and 2015. SUDV shares approximately 70% amino acid sequence identity with EBOV and has caused 5 outbreaks, including the second largest, which infected 425 individuals in Uganda between 2000 and 2001 [2].

Ebola virions are filamentous in shape and contain a single strand of negative-sense RNA. The viral matrix protein VP40 is responsible for binding to the host membrane and budding new virions from infected cells [3, 4]. Indeed, VP40 alone is capable of assembling and budding filamentous virus-like particles from cells, in the absence of any other viral machinery [5–7]. In addition, VP40 has been shown to regulate viral replication and transcription [8].

VP40 exists in multiple structural states, including a dimer, an RNA-binding octameric ring, and a rearranged hexamer [9, 10]. Rearrangement of VP40 into one structure or the other is thought to confer multifunctionality, because each structure may drive a separate and essential biological function [9–11]. Of these structures, the VP40 dimer could be the precursor for both other structures: the RNA-bound ring and the matrix forming filamentous assemblies [10]. Better understanding of the dimer structure could lead to improved templates for drug design, against the dimer itself, or against sites in it that are required for the structural rearrangement. A previous structure of the SUDV dimer is available at 1.83 Å resolution [10], but in that structure, several regions of the C-terminal domain (CTD) are disordered. Here we report a higher-resolution (1.6 Å) crystal structure of the SUDV VP40 dimer in which additional regions of VP40 are now visualized.

METHODS

Cloning and Protein Expression

The expression construct for SUDV VP40 has been described elsewhere [10]. Briefly, VP40 from the Boniface strain of SUDV was truncated with a deletion of the N-terminal 43 residues to enhance crystal formation and produce higher-resolution diffraction. pET-46 Ek/LIC

[a]M. C. C. and J. F. B. contributed equally to this work.

Correspondence: Erica Ollmann Saphire, PhD, The Scripps Research Institute, 10550 N Torrey Pines Rd, IMM-21, La Jolla, CA 92037 (erica@scripps.edu).

The Journal of Infectious Diseases® 2015;212:S167–71

DOI: 10.1093/infdis/jiv090

vector (Novagen) was used as the expression vector. This construct will be referred to as VP40ΔN (Seattle Structural Genomics Center for Infectious Disease ID EbsuA.17247.a.AJ21). VP40ΔN was expressed in BL21 (DE3) Rosetta *Escherichia coli* cells in 2 L of Terrific Broth (Teknova) at 37°C containing ampicillin. Cells were grown to an optical density of 0.4 at 600 nm, at which point protein expression was induced by adding 0.5 mmol/L isopropyl-β-D-thiogalactopyranoside (VWR) and the cultures were grown at 25°C for 16 hours. Cells were harvested by centrifugation, and pellets were stored at −80°C.

Purification

Cells were resuspended for 30 minutes at 4°C in 100 mL of 50 mmol/L monosodium phosphate (NaH_2PO_4; pH 8.0), 300 mmol/L sodium chloride (NaCl), 10 mmol/L imidazole, 500 U of benzonase (Novagen), 1 mg/mL lysozyme (Sigma Aldrich), and 1 cOmplete protease inhibitor tablet (Roche). Cells were subsequently lysed by means of sonication and clarified by means of centrifugation. The lysate was filtered and applied to two 5 mL nickel ion (Ni^{2+})–charged HiTrap Chelating HP (GE Healthcare) columns. The protein was eluted with a 20 column-volume gradient of 50 mmol/L NaH_2PO_4 (pH 8.0), 300 mmol/L NaCl, and 500 mmol/L imidazole. The fractions of interest were pooled and dialyzed into 40 mmol/L NaH_2PO_4 (pH 8.0) and 50 mmol/L NaCl containing 50 U of benzonase. The protein was further purified using size exclusion chromatography over a Sephacryl S-100 16/60 column in 10 mmol/L Tris-hydrochloride (HCl) (pH 8.0) and 300 mmol/L NaCl. Because the molecular weight of the eluted protein was observed to be larger than anticipated, another size exclusion chromatographic analysis was performed using a Superdex S200 30/100 GL column (GE Healthcare). The fractions containing VP40ΔN were pooled, syringe filtered (0.45 μm; Pall), and concentrated to 11.04 mg/mL for crystallography.

Dynamic Light Scattering and High-Performance Liquid Chromatography

Dynamic light scattering experiments were performed using a Zetasizer Nano S ZEN1600 Instrument (Malvern) at 20°C. Samples were tested in triplicate and analyzed for dispersity and mass distribution. High-performance liquid chromatographic experiments were carried out at using an Agilent 1100 liquid chromatography system (Agilent Technologies) outfitted with a multiwavelength detector and a refrigerated 4°C thermostatted autosampler. The VP40ΔN sample and standards were tested over a Superdex 200 10/300 GL column (GE Healthcare). All chromatograms were evaluated with ChemStation software version A.01.04 (Agilent Technologies). Triplicate 200 μg injections of purified VP40ΔN were performed at room temperature using 10 mmol/L Tris-HCl (pH 8.0) and 300 mmol/L NaCl as the mobile phase. Gel filtration standards (Bio-Rad) were applied in duplicate as part of the assay, and protein molecular weights were determined based on linear regression analysis of those standards.

Dynamic Scanning Fluorimetry

For buffer optimization, the protein was diluted to 2 mg/mL in 10 mmol/L Tris-HCl (pH 8.0) and 300 mmol/L NaCl and tested against the pHat Buffer Screen (Rigaku) consisting of 11 buffers, ranging in pH from 3.4 to 11.6. Individual wells included 2 μL of SYPRO dye (1:200 dilution in water), 2 μL of protein at 2 mg/mL, and 25 μL of buffer from the specified screen (Supplementary Table 1*A*). Salt screens were carried out in individual wells containing 2 μL of SYPRO dye, 2 μL of protein at 2 mg/mL, 12.5 μL of 100 mmol/L HEPES (pH 7.0), and 12.5 μL of salt buffer (Supplementary Table 1*B*). Melting temperatures were measured by fluorescence intensity as the temperature was increased from 20°C to 90°C over 60 minutes using the DNA Engine Opticon 2 System (BioRad). Analysis was completed using Opticon Monitor 3 software, showing normalized fluorescence intensity only.

Crystallization

VP40ΔN was crystallized at 11 mg/mL by vapor diffusion in 10% polyethylene glycol 4000, 20% glycerol, and 0.1 mol/L 2-(N-morpholino)ethanesulfonic acid (MES) (pH 6.5) in 2 weeks using 0.4 μL of protein and 0.4 μL of reservoir solution at 16°C. Crystals were directly flash-frozen in liquid nitrogen. Data were collected at the Advanced Photon Source beamline 23-ID-D. The structure was determined by molecular replacement using the EBOV VP40 (Protein Data Bank [PDB] ID 1ES6) structure as a model.

RESULTS

Size Exclusion Chromatography and Dynamic Light Scattering

It was thought that VP40 was primarily monomeric [12], until VP40 produced in bacterial cells was shown to be primarily dimeric by size exclusion chromatography coupled to multiangle light scattering [10]. In this work, SUDV VP40 was purified using sonication instead of microfluidization. Size exclusion chromatography indicates that this SUDV VP40ΔN is also a dimer, with an elution volume of 29.1 mL on a Superdex 200 10/300 GL column (Figure 1*A* and Supplementary Figure 1). In addition, dynamic light scattering confirms that VP40 is a dimer in solution with a mass of 72.0 ± 1.6 kDa (99.7% of total mass), compared with a calculated 65.1 kDa for the dimer (Figure 1*B*). The remaining 0.3% of the sample consists of a large aggregate with a molecular weight of 5.97×10^9 Da. No monomers or discrete oligomers larger than dimer were observed.

Dynamic Scanning Fluorimetry

VP40ΔN is fairly thermostable, with melting temperatures ranging from 32°C to 55°C in a variety of buffers (Supplementary

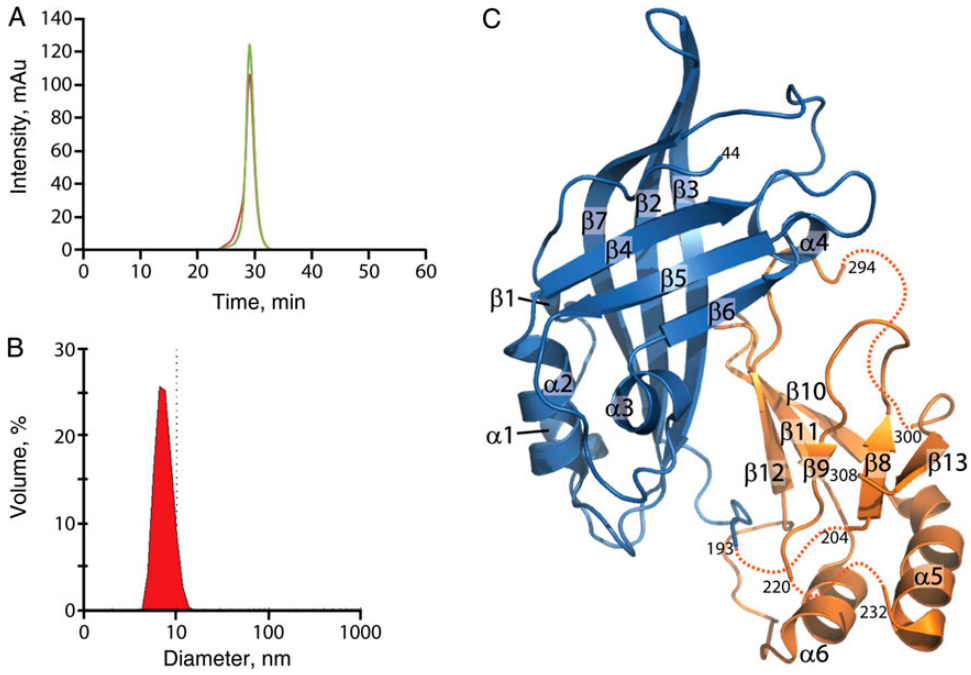

Figure 1. *A, B,* High-performance liquid chromatography analysis (*A*), performed in triplicate, and dynamic light scattering (*B*) confirm that *Sudan ebolavirus* VP40ΔN forms a dimer in solution. *C*, Structure of the VP40ΔN protomer at 1.6 Å, with secondary structure elements and residue numbers labeled. The N-terminal domain (*blue*) is generally well ordered, forming an antiparallel β-sandwich, whereas the C-terminal domain (*orange*) has multiple disordered regions (*dashed lines*), as seen in similar VP40 structures.

Table 1). The highest melting temperatures were achieved in sodium/potassium phosphate, and bicine buffers with a pH between 7 and 8. In general, the addition of salt has very little effect on the melting temperature of this protein, although higher concentrations of salt slightly destabilize it.

Structural Analysis

We solved the crystal structure of dimeric SUDV VP40 at 1.6 Å (PDB 3TCQ). The data collection and refinement statistics are given in Supplementary Table 2. There is a single copy of VP40 in the asymmetric unit with potential oligomerization interfaces observed in the crystal packing (Figure 1*C*). The SUDV matrix protein comprises 2 domains separated by a disordered linker. The N-terminal domain (NTD; residues 44–193) adopts a 7-stranded β-sandwich. The CTD (residues 194–308) forms a smaller stack between 2 highly bent β-sheets. The CTD is more disordered than the NTD, resulting in only short stretches of sequence (203–222, 232–294, and 300–308) that can be modeled into the electron density. The higher resolution of this structure (1.6 Å) compared with another structure of SUDV VP40, 4LD8 (1.83 Å) [10] allowed for the assignment of residues 276–278, which form part of a long loop connecting α6 to β12, and residues 232–234, which are involved in the CTD-to-CTD interface seen in the crystal packing. It should also be noted that L203 and K221 were not built in this model but are present in the 4LD8 model.

DISCUSSION

SUDV VP40 forms a stable dimer in solution, as was reported by Bornholdt et al [10] (Figure 2*A*). Analysis of the dimer by dynamic light scattering did not detect any monomers or larger, nonaggregated, oligomers (Figure 1*B*). There is only one VP40 in the asymmetric unit of this crystal structure. An analysis of the protein-protein interfaces in this structure by the program PISA (Proteins, Interfaces, Structures and Assemblies) [13] similarly identifies the interface between the NTDs in this crystal packing as the basis for the dimer (Figure 2*A*, Supplementary Figure 2*A*, and Supplementary Table 3). This interface buries 1670 Å^2 of surface area and has a positive free energy of dissociation (0.3 kcal/mol), indicating that it is stable in solution. This interface has a limited hydrogen-bond network but is primarily stabilized by hydrophobic interactions. Hydrogen bonding does occur between R52 and the carboxyl groups of I59 and D56, and also between S110 and D109. The larger hydrophobic network involves residues P53, V54, A55, I59, H61, T62, F108, T112, A113, A114, M116, L117, and L181. This interface is highly conserved across *Ebolaviruses*, with 88% sequence identity for residues involved in this interface between SUDV and EBOV, compared with 75% for the entire protein.

Several structures of VP40 from EBOV have been previously determined [9, 10, 12]. Structural alignments between this SUDV dimer and EBOV dimers show a high degree of structural

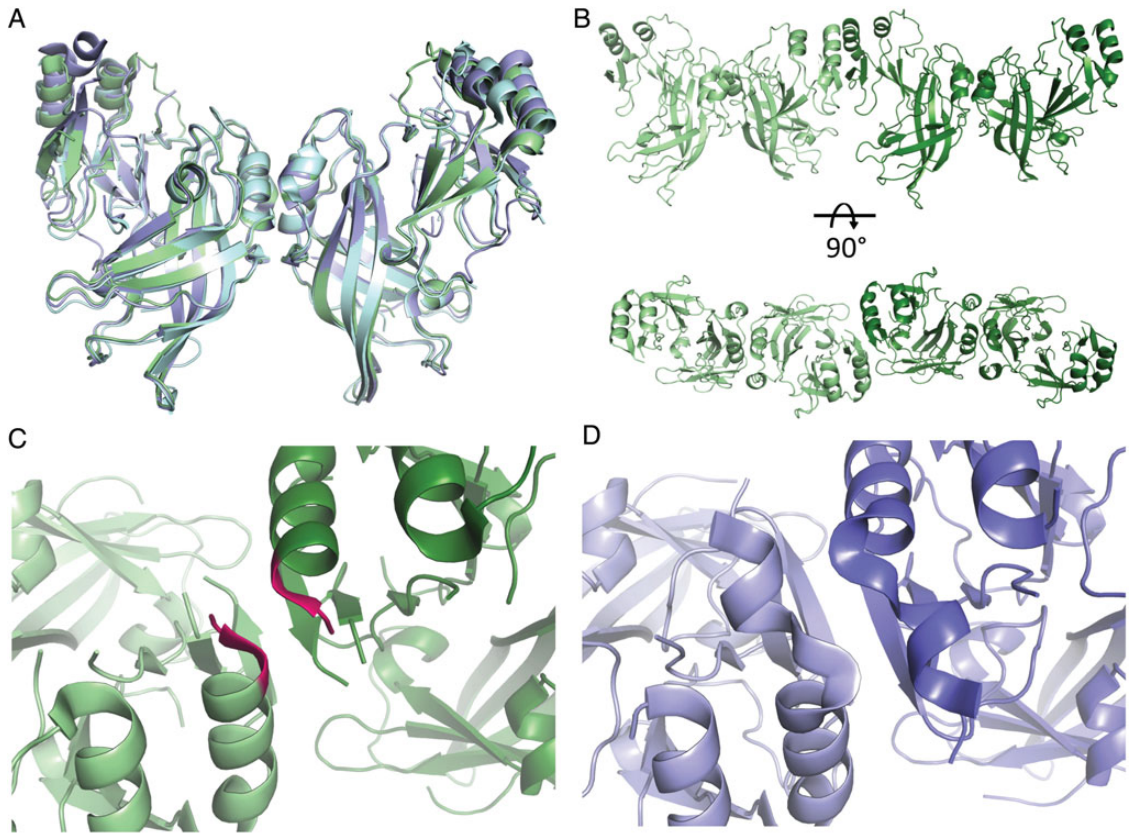

Figure 2. *A*, Alignment of VP40 dimers from Sudan virus (SUDV) and Ebola virus (EBOV). SUDV (Protein Data Bank [PDB] 3TCQ) is shown in green, and two EBOV structures in blue (PDB 1ES6) and cyan (PDB 4LDB). VP40 dimers are thought to further oligomerize via the conserved C-terminal domain (CTD)-to-CTD interface. *B*, CTD-to-CTD interface between two SUDV VP40 dimers is shown in two orientations. *C*, Residues 232–234 were not visible in a lower-resolution structure of SUDV VP40 (PDB 4LD8), but they are clearly visible in this crystal structure. They contribute to the CTD-to-CTD interface and are shown in magenta. *D*, Same interface in EBOV is shown for comparison.

conservation (Figure 2*A*) (root-mean-square deviations of 0.49–0.69 Å based on all common Cα for the entire dimer). It should be noted that in all alignments of these crystal structures, the position of the larger 3-stranded β-sheet in the CTD remains constant relative to the NTD. Taken together, this suggests strong selective pressures for maintaining this domain organization and the dimer interface in Ebola matrix proteins.

Newly formed Ebola virions bud out of infected cells owing to VP40 oligomerization at the plasma membrane [3, 4]. Bornholdt et al [10] identified a conserved CTD-to-CTD interface that is essential for budding virus-like particles (Figure 2*B* and Supplementary Figure 1*B*). This interface links VP40 dimers into filaments (Supplementary Figure 3) and is primarily hydrophobic. In the VP40ΔN structure, residues T232, P234, I237, M305, V306, and I307 form this interface with a buried surface area of 622 $Å^2$ (Supplementary Figure 1*B*). The higher resolution of this structure allowed for the assignment of residues 232–234, which are disordered in another SUDV structure (PDB 4LD8) (Figure 2*C*). These residues contribute substantially to the CTD-to-CTD interface, closely resembling the same interface in EBOV VP40 (Figure 2*D*). It should be noted that VP40 dimers are not predicted to oligomerize through this interface in solution, which is consistent with our dynamic light scattering data. Instead, polymerization via this interaction likely requires a biological trigger or other stabilizing factors to form, such as interactions with the membrane and/or other viral or cellular proteins.

The same CTD-to-CTD interface is found in EBOV VP40 (Figure 2*D* and Supplementary Figure 2*B*, 3*B* and 3*D*) [10, 12]. However, the EBOV interface is significantly larger than that of SUDV, with EBOV burying 1190 $Å^2$, almost twice the surface area. This larger interaction surface is partially accounted for by the presence of a larger methionine residue at position 241 instead of a valine, and a difference in the number of CTD residues (especially L203) that are ordered and could be built into the electron density map (even though SUDV crystals overall diffract to higher resolution than EBOV).

Here we present a higher-resolution structure of an Ebola matrix protein than previously available. This structure highlights

the conserved domain architecture and oligomerization interfaces found in Ebola VP40 proteins and provides an improved platform for structure-based drug design efforts targeted against VP40. VP40 is the most abundantly expressed Ebola protein and is essential for both formation of virions and regulation of RNA replication, making it an attractive target for anti-filovirus drug design [14].

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online (http://jid.oxfordjournals.org). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes

Acknowledgments. We thank the entire Seattle Structural Genomics Center for Infectious Disease team, and we thank Vanessa Anderson for technical assistance. We also appreciate the assistance of the 23-ID-D beamline scientists at the Advanced Photon Source in Argonne, Illinois. Coordinates for the SUDV VP40 have been deposited in the PDB (3TCQ). This is manuscript number 29029 from The Scripps Research Institute.

Financial support. This work was supported by the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH), US Department of Health and Human Services (federal contracts HHSN272 201200025C and HHSN272200700057C to P. J. M.); the Burroughs Wellcome Fund (Investigators in the Pathogenesis of Infectious Disease award to E. O. S.); the NIH (training grant T32 AI00760 to the TSRI Department of Immunology and Microbial Science and J. F. B.); and Achievement Rewards for College Scientists (ARCS) Foundation (J. F. B.). The General Medical Sciences and Cancer Institutes Structural Biology Facility at the Advanced Photon Source (GM/CA @ APS) is funded by the National Cancer Institute (grant ACB-12002) and the National Institute of General Medical Sciences (grant AGM-12006). This research used resources of the Advanced Photon Source, a US Department of Energy Office of Science User Facility operated by Argonne National Laboratory (under contract DE-AC02-06CH11357).

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Burke J, Declerq R, Ghysebrechts G, et al. Ebola haemorrhagic fever in Zaire, 1976. Bull World Health Organ **1978**; 56:271–93.
2. Kuhn JH, Dodd LE, Wahl-Jensen V, Radoshitzky SR, Bavari S, Jahrling PB. Evaluation of perceived threat differences posed by filovirus variants. Biosecur Bioterr **2011**; 9:361–71.
3. Harty RN, Brown ME, Wang G, Huibregtse J, Hayes FP. A PPxY motif within the VP40 protein of Ebola virus interacts physically and functionally with a ubiquitin ligase: implications for filovirus budding. Proc Natl Acad Sci U S A **2000**; 97:13871–6.
4. Panchal RG, Ruthel G, Kenny TA, et al. In vivo oligomerization and raft localization of Ebola virus protein VP40 during vesicular budding. Proc Natl Acad Sci U S A **2003**; 100:15936–41.
5. Geisbert TW, Jahrling PB. Differentiation of filoviruses by electron microscopy. Virus Res **1995**; 39:129–50.
6. Johnson RF, Bell P, Harty RN. Effect of Ebola virus proteins GP, NP and VP35 on VP40 VLP morphology. Virology J **2006**; 3:31.
7. Noda T, Sagara H, Suzuki E, Takada A, Kida H, Kawaoka Y. Ebola virus VP40 drives the formation of virus-like filamentous particles along with GP. J Virol **2002**; 76:4855–65.
8. Hoenen T, Jung S, Herwig A, Groseth A, Becker S. Both matrix proteins of Ebola virus contribute to the regulation of viral genome replication and transcription. Virology **2010**; 403:56–66.
9. Gomis-Ruth FX, Dessen A, Timmins J, et al. The matrix protein VP40 from Ebola virus octamerizes into pore-like structures with specific RNA binding properties. Structure **2003**; 11:423–33.
10. Bornholdt ZA, Noda T, Abelson DM, et al. Structural rearrangement of Ebola virus VP40 begets multiple functions in the virus life cycle. Cell **2013**; 154:763–74.
11. Hoenen T, Biedenkopf N, Zielecki F, et al. Oligomerization of Ebola virus VP40 is essential for particle morphogenesis and regulation of viral transcription. J Virology **2010**; 84:7053–63.
12. Dessen A, Volchkov V, Dolnik O, Klenk HD, Weissenhorn W. Crystal structure of the matrix protein VP40 from Ebola virus. EMBO J **2000**; 19:4228–36.
13. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol **2007**; 372:774–97.
14. Stahelin RV. Could the Ebola virus matrix protein VP40 be a drug target? Expert Opin Ther Targets **2014**; 18:115–20.