Bunyamwera Orthobunyavirus Glycoprotein Precursor Is Processed by Cellular Signal Peptidase and Signal Peptide Peptidase

Xiaohong Shi\textsuperscript{a1}, Catherine H. Botting\textsuperscript{b}, Ping Li\textsuperscript{a}, Mark Niglas\textsuperscript{b}, Benjamin Brennan\textsuperscript{a}, Sally L. Shirran\textsuperscript{b}, Agnieszka M. Szemiel\textsuperscript{a} and Richard M. Elliott\textsuperscript{a2}

\textsuperscript{a}MRC-University of Glasgow Centre for Virus Research, University of Glasgow, Glasgow G61 1QH, United Kingdom

\textsuperscript{b}Biomedical Sciences Research Complex, University of St Andrews, St Andrews, KY16 9ST, United Kingdom

\textsuperscript{1}To whom Correspondence should be addressed. Email: xiaohong.shi@glasgow.ac.uk

\textsuperscript{2}This paper is dedicated to the memory of our colleague Richard M. Elliott who died on June 5 2015 whist this work was ongoing.

Running title: Bunyamwera virus glycoprotein precursor processing

Significance

Bunyamwera virus (BUNV) is the prototype of the Orthobunyavirus genus and Bunyaviridae family that contains important human and animal pathogens. The cleavage mechanism of orthobunyavirus glycoprotein precursor (GPC) and the host proteases involved have not been clarified. Here we found that NSm and Gc contain their own internal signal peptides (SPs) which mediate the GPC cleavage by host signal peptidase (SPase) and signal peptidase (SPP). Furthermore, the NSm domain-I (SP\textsuperscript{NSm}) plays an important post-cleavage role in cell fusion. Our data clarified the implication of host proteases in the processing of the orthobunyavirus GPC. This identifies SPP as a potential intervention target and the knowledge we gained will benefit the preventive strategies against other orthobunyavirus infections.
Abstract

The M genome segment of Bunyamwera virus (BUNV), the prototype of both the Bunyaviridae family and the Orthobunyavirus genus, encodes the glycoprotein precursor (GPC) that is proteolytically cleaved to yield two structural glycoproteins, Gn and Gc, and a nonstructural protein NSm. The cleavage mechanism of orthobunyavirus GPCs and the host proteases involved have not been clarified. In this study, we investigated the processing of BUNV GPC and found that both NSm and Gc proteins were cleaved at their own internal signal peptides (SPs), in which NSm domain-I functions as SP_{NSm} and NSm domain-V as SP_{Gc}. Moreover, the domain-I was further processed by a host intramembrane-cleaving protease, signal peptide peptidase (SPP) and is required for cell fusion activities. Meanwhile, the NSm domain-V (SP_{Gc}) remains integral to NSm, rendering the NSm topology as a two-membrane-spanning integral membrane protein. We defined the cleavage sites and boundaries between the processed proteins: Gn, from residue 17 to 312 or nearby residues; NSm, 332 to 477; and Gc, 478 to 1432. Our data clarified the mechanism of the precursor cleavage process, which is important for our understanding of viral glycoprotein biogenesis in the genus Orthobunyavirus and thus presents a useful target for intervention strategies.

\body

The family Bunyaviridae contains more than 350 named viruses that are classified into the five genera Orthobunyavirus, Hantavirus, Nairovirus, Phlebovirus, and Tospovirus, making it one of the largest families of RNA viruses. Several members of the family are serious human pathogens, such as La Crosse virus (LACV) and Oropouche virus (OROV) (Orthobunyavirus), Hantaan (HTNV) and Sin Nombre viruses (Hantavirus), Rift Valley fever virus (RVFV),
Phlebovirus) and severe fever with thrombocytopenia syndrome Virus (SFTSV) (Phlebovirus), Crimean-Congo hemorrhagic fever virus (CCHFV, Nairovirus) (1, 2). The characteristic features of bunyaviruses include a tripartite single-stranded RNA genome of negative- or ambireic polarity, cytoplasmic site of viral replication, and assembly and budding at membranes of the Golgi complex (1-3). Bunyamwera virus (BUNV), the prototype of both the family and the Orthobunyavirus genus, remains an important research model for many pathogens within this family.

The medium (M) genomic RNA segment of orthobunyaviruses encodes the glycoprotein precursor (GPC; in order Gn-NSm-Gc) that is cotranslationally cleaved to yield the mature viral glycoproteins Gn and Gc, and a nonstructural protein NSm. Gn and Gc form viral spikes that play a crucial role in virus entry (1, 2). Both Gn and Gc are type I integral transmembrane proteins and form a heterodimer in the endoplasmic reticulum (ER) prior to trafficking to, and retention in, the Golgi compartment where virus assembly occurs (2, 4). Bunyavirus glycoproteins are fusogenic and the fusion peptide is located on Gc, a class II fusion protein (5), but cell fusion requires the co-expression of both Gn and Gc glycoproteins (6). NSm, an integral membrane protein, comprises three hydrophobic domains (I, III and V) and two non-hydrophobic domains (II and IV) (Fig. S1A) and its N-terminal domain (I) is required for BUNV replication (7).

Cleavage of BUNV GPC is mediated by host proteases, but the details of which proteases are involved and the precise cleavage sites have not been clarified. Experimental data on GPC processing has only been reported for snowshoe hare orthobunyavirus (SSHV); the carboxyl terminus of SSHV Gn was determined by C-terminal amino acid sequencing to be an arginine (R) residue at position 299 (8) (Fig. S1B). Based on alignments of several orthobunyavirus GPC sequences, it was suggested that Gn-NSm cleavage occurs at a similar position to that defined for SSHV (9). This arginine (302R for BUNV) appears conserved in
GPCs of all orthobunyaviruses analyzed to date, and for most of the viruses lies in the sequence R-V/A-A-R (Fig. S1C), which has been believed as the site of Gn-NSm cleavage by furin-like proteases (10).

In eukaryotes, most secreted and membrane proteins contain cleavable N-terminal signal peptides (SPs), which are recognized by the signal recognition particles (SRP) when nascent polypeptide chains emerged from the ribosome at ER and translocate it into the ER lumen where they are usually cleaved by cellular signal peptidases (SPase) (11). The imbedded peptide remnant is usually subsequently released for degradation by the cellular signal peptide peptidase (SPP) or SPP-like proteases, which belong to the family of intramembrane-cleaving aspartyl proteases (I-CLiPs) (12, 13). SPP is an ER-resident I-CLiPs (14) and is implicated in other important biological functions, such as in generating C-terminal peptides for MHC class I presentation (15) and human lymphocyte antigen E (HLA-E) epitopes (16). SPP activities also involve in the intra-membrane cleavage of the core proteins of hepatitis C virus (HCV), GB virus B and classical swine fever virus (CSFV) (family Flaviviridae) (17, 18).

To investigate the cleavage events of orthobunyavirus GPC, we used several approaches including mutagenesis, virus assays, RNA interference, mass spectrometry and biological assays. We aimed to determine the cleavage sites between the boundaries of the mature proteins, e.g. Gn and NSm, and NSm and Gc. Our data revealed the implementation of the cellular SPase and SPP in the cleavage of BUNV GPC and clarified the mechanism of orthobunyaviruses GPC cleavage.
**Results**

Gn-NSm Cleavage Does Not Occur at the RVAR Motif by Furin-like Proteases. We first investigated whether the Gn-NSm cleavage takes place between residues 302R and 303R at the RVAR motif (Fig. S1). Five mutations were generated at this site in the M expression plasmid pTM1BUNM, including a deletion mutation (ΔRVAR) and four substitution mutations (MMKR, AAAA, RSLK and RRKR) (Fig. 1A). These plasmids were transfected into BSR-T7/5 cells and the radiolabelled viral proteins were immunoprecipitated with anti-BUN serum followed by SDS-PAGE fractionation. Interestingly, like the wild-type (wt) BUNM control, all mutated GPCs were cleaved into Gn, NSm and Gc (Fig. 1B), suggesting this location is not a cleavage site by furin-like proteases. Furthermore, the furin inhibitor I (dec-RVKR-cmk, Calbiochem) had no effect on the BUNV GPC cleavage (Fig. S2A) and yields of virus produced in the presence of the drug over 30 hrs (Fig. S2B).

There exist eight residues between residues 302R and 311S at Gn-NSm junction (Fig. S1 and Fig. 1C). To investigate whether these residues harbor the Gn-NSm cleavage site, we constructed six mutants that contain internal progressive deletions between residues 298L and 311S (Fig. 1C). As shown in Fig 1D, all mutated precursors were properly cleaved. Moreover, the deletions resulted in the increased migration of Gn bands on the gel, with a relative shift corresponding to the number of amino acids removed (lanes 3 to 8), suggesting that these residues still belong to the Gn cytoplasmic tail (Gn CT) and the Gn-NSm cleavage must occur at or within NSm domain-I.

NSm Domain-I Functions as an Internal Signal Peptide. After excluding the Gn-NSm cleavage at motif RVAR\textsuperscript{302}, we speculated that NSm domain-I, a type II transmembrane domain (TMD), could function as an internal SP for NSm (SP\textsuperscript{NSm}). Using SignalP 4.1 server (www.cbs.dtu.dk/services/SignalP/) (19) the domain was predicted as a cleavable internal SP that cleaves between residues 331G and 332T (Fig. S3A and Fig.2A). It has been reported...
that the residues at the -3 and -1 positions relative to the SPase cleavage site are most
critical for cleavage by cellular SPase complex (20). Therefore, we generated a series of
mutant GPCs, including six substitution mutations at the -3 and -1 positions and one
substitution mutation in the core region of the domain (SPm) (Fig. 2A). When the residues
at either -3 (329I) or -1 (331G), or at both positions, were changed to the charged arginine
(R), the NSm protein was not detected (Fig. 2B, lanes 2, 4 and 6). However, substitution with
alanine at these positions and substitution mutation within the domain (SPm) did not affect
Gn-NSm cleavage (lanes 3, 5, 7 and 8) in that NSm was clearly seen. Furthermore, we
purified NSm protein from the cells infected with recombinant virus, rBUNNSmV5, in which
the V5-epitope was inserted in the NSm cytoplasmic domain (Fig. S3B), for mass
spectrometric (MS) analysis and confirmed that the residue 332T is the first N-terminal
residues of NSm (Fig. S3C and S3D).

The effects of the mutations on glycoprotein function were also assessed with regard to
the Golgi trafficking and cell fusion activities. Consistent with the above observation, the
arginine substitutions at -3 and -1 position totally abolished Golgi targeting (Fig. 2C, panels c,
e, and g) and cell fusion (panels d, f and h), whereas alanine substitution had no effect on
either Golgi colocalization (panels l, o and r) or cell fusion (panels m, p and s). When these
mutations were tested for virus rescue, we were able to generate viruses from M segment
mutants containing alanine substitution mutations (rAEG, rIEA and rAEA), but not from
arginine substitution mutants. The rescued viruses showed similar plaque phenotypes to
the wt control (panels n, q and t). Taken together, our data confirmed that the NSm
domain-I functions as an internal SP^{NSm}.

Mapping the C-termini of Gn Protein. After SPase cleavage of SP^{NSm} at residue 332T,
the SP is still attached to the upstream Gn CT (preGn). To define the Gn end, we constructed
a series of individual Gn mutants that terminate between residues 298 to 332 (Fig. 3A).
Consistent with the earlier results from internal deletions, the deletions from residues 312 to 298 resulted in the linear reduction in the molecular weight of Gn proteins (Fig. 3B, lanes 4 to 7) and Gn312 is comparable in size to wt Gn (comparing lanes 3 and 7), indicating that Gn likely ends at residue 312 or nearby residues. However, deletions in SP\textsuperscript{NSm} did not cause a linear reduction in Gn molecular weights (lanes 8 to 10). It is noticeable that Gn317, which contains only five remnant hydrophobic residues at the C-terminus, produced a smaller band (about 25 kDa), which we believed to be a degradation product (lane 8). We also compared the size of the intracellular and virion Gn proteins and found no difference (Fig. 3C), confirming the intracellular and virion Gn proteins end at same position.

To confirm the subsequent processing of SP\textsuperscript{NSm} upon SPase cleavage, we constructed a mutated Gn (Gn308V5) with domain-I replaced with non-hydrophobic V5 epitope and a further 6 residues from the Gc CT (QEIKQK) (Fig. 3A). It is worth mentioning that the unprocessed Gn332 (preGn) would be similar in size to Gn308V5 (35.33 v 35.12 kDa). As anticipated, Gn308V5 runs at a higher molecular weight than wt Gn and the processed Gn332 (Fig. 3D), indicating that SP\textsuperscript{NSm} was further processed from preGn. The preGn was visible by Western blot (WB) analysis of V5-tagged Gn proteins (Gn-27V5 and Gn-86V5) as a faint band above the predominant processed Gn protein (Fig. 3E). Cell fusion assays with Gn mutants (in co-expression with Gc) indicated that the extensive syncytia were formed only from cells coexpressing Gc and Gn332 (PreGn), and any deletions in SP\textsuperscript{NSm} diminished the cell fusion (Fig 3F), suggesting that the liberated SP\textsuperscript{NSm} is required in cell fusion.

MS analysis of virion Gn protein identified 17S as the first residue of the processed Gn protein (Fig. S4B), but unable to confirm the Gn end, with the furthest C-terminal residue was mapped to 303R (Fig. S4C and SI Table ), similar to the previously determined SSHV Gn end (8). It is probably due to the fact that the newly identified Gn C-terminal residues (303
to 312) is rich in positively charged arginine and lysine residues being targets by trypsin-like proteases (21). The terminal residues identified by MS are summarized in Table S1.

NSm domain-V functions as an internal noncleavable SP\textsuperscript{Gc}. Our previous study suggested that NSm domain-V function as a SP\textsuperscript{Gc} (7). Indeed, the deletion of the domain abolished Gc processing as no Gc protein was detected from the mutated BUNV GPCs that contain whole or partial internal deletion in the domain (Fig. 4A, lanes 4 to 6). To study whether the domain undergoes any further processing as SP\textsuperscript{NSm}, we compared the size of C-terminal truncated NSm proteins (based on pTmBUNGn-NSmV5) (Fig. 4B). The deletions resulted in the increased migration shift of NSm bands (Fig. 4C, lanes 4 to 8). NSm476 that contains the intact domain-V was identical in size with the parental NSmV5 (comparing lanes 9 and 10), indicating that domain-V is not further cleaved. Furthermore, no size change was noticeable when the domain-V was swapped by either non-hydrophobic residues from EGFP (residues 1 to 20) or hydrophobic signal peptide of Hantaan virus (HTNV, strain 76-118) Gn protein (SP\textsuperscript{HGn}, residues 1 to 19) (Fig. 4D), whereas NSm456 that lacks the domain-V was obviously smaller (Fig. 4D, lane 6), a corroborating evidence that SP\textsuperscript{Gc} remains integral to the mature NSm. This renders the topology of mature NSm as a two-membrane-spanning protein (residues 332 to 477) that consists of the ectodomain (II), TMD (III), endoplasmic loop (IV) and C-terminal type-II TMD (V) (Fig. 4E).

Requirement of SP\textsuperscript{NSm} for Cell Fusion and Virus Replication. To further investigate the role of SP\textsuperscript{NSm} in virus replication, we made four mutant GPC constructs (Fig. 5A), and compared the impact of mutations in promoting cell fusion and virus viability. All four mutated GPCs were cleaved into Gn and Gc, and also NSm from the BUNM-SPm\textsuperscript{NSm} and -SP\textsuperscript{NSm}/HTNtm (Fig. 5B). A cell fusion assay revealed that only the GPC-\textsuperscript{ΔNSm-I}, which contained deletion of whole mature NSm (residues 332 to 477) but retained SP\textsuperscript{NSm} for Gc translocation, produced extensive syncytial formation (Fig. 5C). For other three GPC mutants, the domain-V, SPm\textsuperscript{NSm},
and SP\textsuperscript{NSm}/HTNtm functioned properly as SP\textsuperscript{GIC} for GPC processing, but these mutated GPCs failed to cause cell fusion. Consistently, from GPC-\Delta NSm.I we rescued a recombinant virus (rBUN\Delta NSm) that did not express NSm protein (Fig. 5D, lane 4). The rBUN\Delta NSm was attenuated, evidenced by smaller plaque phenotype (Fig. 5E) and reduced virus yield at late infection (Fig. 5F). The same strategy was used to create a viable recombinant OROV lacking NSm (22).

To characterize further SP\textsuperscript{NSm}, we assessed the effect of the deletion mutations in the domain on GPC processing and cell fusion. As we mentioned early, deletions in SP\textsuperscript{NSm} did not lead to a linear reduction in the molecular weight of Gn proteins (Fig. 5G). The Gn bands from Gn324 to Gn332 are comparable in size with wt Gn (lanes 6 to 11), but further deletion towards Gn312 (in the case of Gn317 and Gn322) would interrupt the proper processing of Gn protein (lanes 4 and 5), suggesting that the processing requires the stable SP TMD structure. When these mutants were tested in cell fusion, we noticed that even removal of two residues from SP\textsuperscript{NSm} (Gn330) had a significant impact on syncytia formation and that further deletions diminished the extent of cell fusion (Fig. 5H). The data together indicated that in addition to the role as SP, the liberated SP\textsuperscript{NSm} has post-cleavage function.

**Involvement of SPP in the Processing of BUNV GPC.** As SPP is the ER-resident I-Clips, we suspect that SPP is probably involved in the further processing of SP\textsuperscript{NSm}. To address the issue, we generated three lentiviruses that express small hairpin RNAs (shRNAs): two specific to human SPP mRNA and one to EGFP (shGFP) as negative control. Both shSPP1 and shSPP2 were able to inhibit the SPP expression in the transduced A549 cells (shSPP2 showed a better silencing effect) (Fig. 6A). We then examined the impact of SPP knockdown on BUNV replication following low multiplicity infection of A549V cells (MOI of 0.01 pfu/cell). WB analysis showed that the detection of BUNV N was delayed by 24 hrs in shSPP-silenced cells compared with shGFP control (Fig. 6B) and the virus titre in SPP-knockdown cells was over
10-fold lower than controls across the infection period (Fig. 6C). This indicates the likely involvement of SPP in BUNV replication. In order to investigate if SPP is implicated in replication of other bunyaviruses, we infected the transduced A549V cells with SBV (Orthobunyavirus genus), RVFV (Phlebovirus genus) and Puumala virus (PUUV, Hantavirus genus), and their N proteins and virus titres were determined by WB and plaque assay. A significant inhibitory effect of SPP silencing was observed for SBV (Fig. 6D and 6E). For RVFV, Inhibition was noticeable but to lesser extent than BUNV and SBV (not statistically significant) (Fig. 6F and 6G). Significant inhibited was observed for PUUV replication in SPP-knockdown cells (Fig. 6H and 6I), but PUUV growth was more inhibited in shGFP control cells than in SPP-knockdown cells, evidenced by the N protein detection in SPP-knockdown cells but not in shGFP control at 72 hrs p.i. (Fig. 6H, lane 7 at bottom panel of long exposure) and significantly lower virus titre in shGFP control (Fig. 6I). This pattern for PUUV is largely due to antiviral activity we found present in shGFP lentivirus preparation (Fig. S5A), to which PUUV is more sensitive than BUNV to the inhibition effect (Fig. S5B).

The inhibitory effect of SPP knockdown on BUNV infection was also visualized using a recombinant virus (rBUNGc-eGFP) with eGFP fused to Gc (3). Naive Huh7 cells and cells expressing shSPP2 were infected with rBUNGc-eGFP (MOI of 0.01), and cells were examined at 10, 24 and 48 hrs p.i. by fluorescence microscopy. Production of eGFP-tagged Gc, and hence production of progeny virus particles, was observed in naive cells at 10 hrs p.i. and virus spread to adjacent cells was clearly evident at 24 hrs p.i. All cells were infected by 48 hrs p.i. (Fig 6J, panels a to c). In SPP-knockdown cells, the eGFP-tagged Gc was not observed until 24 hrs p.i., and the spreading to neighboring cells was still limited at 48 hrs p.i. (panels d to f).
Discussion

The cleavage between BUNV Gn and NSm has long been thought to occur at the amino acid motif RV/AAR, which is conserved in several orthobunyaviruses (9, 10) and fits the minimum furin cleavage site (RXXR)(23). However, as the furin-like proprotein convertases (PCs) process substrates in the lumen of the Golgi complex and endosome or at the cell surface (24), it is unlikely that the RxxR motif in the Gn CT can be accessed by these proteases. Moreover, some members of the genus Orthobunyavirus, such as Wyeomyia virus (WYOV), SBV and OROV lack the RxxR motif (Fig. S1C). In fact, we have proven that the motif and the downstream eight residues (residues 303 to 310) are still part of Gn CT. In the Bunyaviridae family, the furin-like protease is involved in the GPC processing of CCHFV (Nairovirus genus) for generating a 38-kDa NSm protein, whereas the CCHFV furin site is located at the ectodomain of pre-Gn protein (25).

By using mutagenesis and MS analysis we confirmed that the NSm domain-I is $\text{SP}^{\text{NSm}}$ which is cleaved by SPase at residue 332T of mature NSm. The residual $\text{SP}^{\text{NSm}}$, which is still linked to the upstream Gn CT (as preGn), is further processed from Gn C-terminus by the ER-resident SPP. The implication of SPP in BUNV GPC process is validated by our observations: 1) The further processing of NSm domain-I ($\text{SP}^{\text{NSm}}$) upon SPase cleavage, 2) the detection of preGn by WB analysis of V5-tagged Gn protein, and 3) Inhibition of BUNV and SBV infection in SPP-knockdown cells. We also assessed the impact of SPP knockdown on two other bunyaviruses, RVFV (Phlebovirus) and PUUV (Hantavirus). SPP knockdown had no significant inhibitory effect on RVFV infection, whereas it inhibited PUUV infection. However, as PUUV is more subdued in the shGFP-induced cells, we were unable to draw a clear conclusion. Some lentivirus expressing shRNAs can trigger IFN activation(26) and the effect of siRNA on innate immunity is sequence and structure related (27). It should be mentioned that the coding strategies and sizes of products encoded by M segments of the
viruses in the family are very divergent and thus it is plausible that the precursor processing differs from genus to genus.

Besides its role as a SP, we provide evidence that the liberated $SP^{NSm}$ has post-cleavage function in cell fusion. We speculate that the liberated $SP^{NSm}$ is likely incorporated into the virion by interacting with one of viral glycoproteins upon cleavage by SPP, which is probably required for that interaction. However, we were unable to find the peptide from the purified virus particles by MS analyses, perhaps due to the technical challenge for the reason of the small size and hydrophobicity of the domain. Another possibility is that the Gn CT is modified during the processing of $SP^{NSm}$ by SPP and that modification might be crucial for glycoprotein activities. Whatever the cases, it seems that the sequence specificity of $SP^{NSm}$ is important. Several signal peptides of viral proteins have post-cleavage functions. For instance, $SP^{GP-C}$ of lymphocytic choriomeningitis virus and Junín virus (Arenaviruses) precursor glycoproteins C (GP-C) is an essential structural component of mature virions and is required in both glycoprotein maturation, cell fusion and virus infectivity (28-30).

Based on our findings, we propose a new model for the process and topology of orthobunyavirus glycoproteins (Fig. 7A). All three proteins, in precursor form, contain their own SPs. The N-terminal $SP^{Gn}$ targets the nascent Gn polypeptide through the ER membrane where it is cleaved by SPase at residue 17S. The internal $SP^{NSm}$ mediates the translocation of the nascent NSm chain into the ER membrane and it is subsequently cleaved at residue 322T by SPase from the NSm. $SP^{NSm}$ is further processed by SPP to free the Gn CT. NSm domain-V/$SP^{Gc}$ translocates the nascent Gc chain to the ER membrane and is cleaved at residue 478E to separate the mature NSm and the nascent Gc chain. However, unlike $SP^{NSm}$, the domain-V/$SP^{Gc}$ is not further processed and remains as the C-terminal domain of mature NSm. The new topology model of mature Gn, NSm and Gc proteins is illuminated in Fig. 7B.
Gn and Gc proteins are type I transmembrane protein and NSm is a two-membrane-spanning protein.

This study revealed a new dimension for SPP in virus replication. The new knowledge will benefit vaccine development and help identify new antiviral drugs against pathogenic virus infections caused by viruses in the family. Indeed, with the knowledge we gained we have generated recombinant BUNV and SBV viruses that lack full mature NSm as well as NSs proteins (Shi; et al., unpublished data) and these viruses would be potential candidates or tools for vaccine development.

Materials and Methods

The materials and methods are described in SI Materials and Methods. They include cells and viruses, antibodies, plasmids and mutagenesis, transfection of cells, metabolic radiolabeling and immuno-precipitation, immunofluorescence staining, BUNV glycoprotein fusion assay, virus rescue by reverse genetics, virus infection, titration and purification, preparation of BUNV glycoprotein Gn, Gc and NSm proteins, mass spectrometric (MS) Analysis, SPP knockdown by lentivirus shRNA, western blotting (WB).

Statistical Analysis. Data were expressed as the mean and SD. The P value and statistical significance of difference was analyzed by using unpaired t test with GraphPad 6 software.

*P value < 0.05, significant; **P< 0.01, very significant; ***P <0.001, extremely significant.

ACKNOWLEDGEMENTS
We thank Drs. Klaus K. Conzelmann (Ludwig-Maximilians-Universität München), David Jackson and Richard Randall (University of St. Andrews) and Martin Lowe (University of Manchester) for providing reagents used in this work, Drs. Friedemann Weber (Justus-Liebig-Universität Gießen) and Alain Kohl (University of Glasgow) for critical reading of the manuscript. We also thank Angela Elliott and Edward Dornan and Dr. Ye Liu for technical support. This study was supported by Wellcome Trust grant (to RME) and grant (094476/Z/10/Z) that funded the purchase of the TripleTOF 5600 mass spectrometer at the BSRC of University of St Andrews.

**Figure Legend**

**Fig. 1.** Gn-NSm does not cleave at the previously predicted RVAR site. (A) Substitution and internal deletions at RVAR motif. (B) Effect of deletion and substitution mutations on GPC cleavage. (C) Internal deletions between residues 298L and 311S at Gn-NSm junction. (D) The processing of mutant GPCs. Transfected BSRT7/5 cells were radiolabelled with [35S]methionine. The viral proteins were immunoprecipitated with anti-BUN and analyzed by SDS-PAGE. The position of the viral proteins is marked.

**Fig. 2.** NSm domain-I acts as an internal SPNSm. (A) Mutations at -3 (329I) and -1 (331G) positions of the SPase cleavage site (on pTM1-BUNM-NSmV5). (B) Effect of substitution mutations on NSm processing. The radiolabeled viral proteins were immunoprecipitated with anti-V5 antibody. (C) Effect of mutations on the Golgi targeting of the viral glycoproteins, cell fusion and virus viability. For immunofluorescence assays, the transfected BSRT7/5 cells were stained with a mixture of anti-Gc MAb (M810, in red) and anti-GM130 (in green for the Golgi staining) antibodies and examined by confocal microscopy. Nuclei were stained in blue with DAPI. Cell fusion and virus rescue were performed as described in SI Materials and Methods.
Fig. 3. Mapping the C-termini of Gn protein by mutagenesis. (A) Mutations in the Gn CT and NSm domain-I (on pTM1BUN-Gn332). ‘*’ represents the stop codon. (B) Effect of deletions on the migration shift of Gn protein. (C) Profile of the radiolabelled intracellular and virion proteins. (D) The expression of Gn332 (preGn) and Gn308V5. (E) Western blot (WB) analysis of V5-tagged Gn332 at residue 27 or 86. (F) Cell fusion assay on BSRT7/S cells co-transfected with pTM1BUN-NSmGc (Gc) and one of the Gn mutants.

Fig. 4. NSm domain-V functions as a noncleavable SP\textsubscript{Gc}. (A) Deletion of Domain-V abolished the Gc processing. (B) Mutations at NSm domain-V. ‘*’ represents the stop codon. (C) Effect of deletions on the migration shift of the V5-tagged NSm protein. (D) WB analysis of V5-tagged NSm and its mutants. (E) The revised topology of NSm protein.

Fig. 5. Requirement of SP\textsuperscript{NSm} for GPC processing, cell fusion and virus replication. (A) Schematic showing either deletion of NSm or mutations in SP\textsuperscript{NSm}. (B) Effect of mutations on the GPC cleavage. (C) Cell fusion on BSRT7/S cells transfected with BUNM mutants. (D, E and F) The protein profile (D), plaque phenotype (E) and growth curve (F) of the recombinant virus lacking NSm (rBUN\textDeltaNSm). (G) Effect of deletion mutations in SP\textsuperscript{NSm} on the migration shift of Gn protein. (H) Cell fusion on BSRT7/S cells co-transfected with pTM1BUN-NSmGc and one of the Gn mutants.

Fig 6. SPP KO affects BUNV infection. (A) SPP-knockdown in A549 cells. (B and C) WB analysis (B) and Growth kinetics (C) of BUNV infection in shRNA expressing A549V cells. (D and E) Effects of SPP-knockdown on SBV infection. (F and G) Effects of SPP knockdown on RVFV infection. (H and I) Effects of SPP-knockdown on Puumala virus (PUUV) infection. The transduced A459V cells were infected with virus (MOI of 0.01). At each time point the supernatants were harvested for virus titration and cell lysates were collected for WB. The relevant proteins were probed with antibodies against SPP, tubulin (T) or viral N proteins. (J)
Effect of SPP-knockdown on BUNV spreading. Transduced Huh7 cells were infected with rBUNGc-eGFP (MOI of 0.01) and incubated until time points (10, 24 and 48 hrs p.i.) as indicated. Cells were stained with anti-GM130 and examined by confocal microscopy. EGFP-tagged Gc and virus particles were shown in green, GM130 were stained in red and the nuclei were stained in blue with DAPI.

**Fig. 7.** The model of BUNV GPC (Gn-NSm-Gc) cleavage. (A) Schematic showing BUNV GPC processing. (B) Topology of Gn, Gc and NSm proteins. See discussion for details.
Supporting Information

SI Materials and Methods

**Cells and Viruses.** A549, A459V (31), A549-NPro (32), Vero E6, BHK-21, HEK 293T, Huh7 and BSR-T7/5 (33) cells were maintained as described previously (4). BUNV, SBV, RVFV (strain MP12) and Puumala virus (PUUV, strain CG1820) were used as representative strains for genera of *Orthobunyavirus, Phlebovirus and Hantavirus* in SPP knockdown assays. rBUNGc-eGFP is a recombinant BUNV with the truncated Gc tagged by eGFP (3).

**Antibodies.** MAb810, a monoclonal antibody against BUNV Gc, anti-BUN and anti-NSm, the rabbit antisera against BUNV particles and NSm peptide TDQKYTLDEIADVLQA (residues 338 to 353 of BUNV GPC), have been described previously (34). Rabbit anti-BUN N, anti-SBV N and anti-RVFV N antisera were raised against nucleoproteins of BUNV, SBV and RVFV, respectively (35-37). A rabbit antiseraum against GM130, a cis-Golgi matrix protein was provided by Dr M. Lowe (University of Manchester, UK). Other antibodies were purchased commercially: rabbit polyclonal antibody against SPP (anti-SPP) (Abcam); mouse monoclonal antibody against the V5 epitope (anti-V5) (Serotec); goat anti-rabbit antibody conjugated with fluorescein isothiocyanate (FITC) (Sigma), and goat anti-mouse antibody conjugated with Cy5 (Chemicon International Inc).

Production of Anti-PUU N. The 6His-tagged Puumala virus (PUUV) N protein was expressed in BL21-CodonPlus (DE3)-RP *E. coli* (Agilent Technologies) under IPTG induction at 18-20 °C for 20 hrs, purified with Ni-NTA resin and its identity confirmed by mass spectrometry. The purified PUU N was used for generating rabbit antisera against PUUV N protein (Eurogentec).

**Plasmids and Mutagenesis.** Plasmids that generate full-length BUNV antigenome RNA transcripts, pT7riboBUNL(+), TVT7R-BUNM(+), and pT7riboBUNS(+), or express the full
length GPC, pTM1-BUNM, have been described previously (7, 38); pT7riboBUNM-NSmV5(+) and pTM1BUNM-NSmV5 contain the V5 epitope (GKIPNPLLGLDST) inserted between residues 403 and 420 of the NSm coding region (Fig. S3B). The substitution and internal deletion mutations in BUNV GPC (in the backbone of either pTM1-BUNM for high expression or pT7riboBUNM for virus rescue) were generated by PCR-directed mutagenesis on the appropriate parental templates.

A series of individual Gn expression plasmids (pTM1BUNGn298 to Gn332) were derived from pTM1BUNM (Fig. 3A) and Gn-NSm expression plasmids (pTM1BUNGn-NSmV5-454 to 476) were derived from pTM1BUNM-NSmV5 (Fig. 4A), in which the V5 epitope is inserted in the NSm cytoplasmic domain (IV) between residues 403 and 420 (Fig. S4B). The V5 tagged NSm was also introduced into TVT7RBUNM[+] for generating recombinant virus containing V5-tagged NSm. Two mutated Gn-NSm constructs (pTmBUNGnNSmV5-SPGC/SPHGn and pTmBUNGnNSmV5-SPGC/eGFP1-20) were created by replacement of the domain-V with either signal peptid of Hantaan virus Gn protein (resides 1 to 19) or N-terminal 20 residues of eGFP (Fig. 4A). Three mutated GPC constructs (pTM1BUNMΔ446-476, Δ457-476 and Δ467-476) contain internal deletions in the coding region of NSm domain IV and V. pTM1BUNM-NSmGc expresses NSm and Gc proteins (residues 309 to 1433). pTmBUNMΔNSm _V and pTmBUNMΔNSm _I contain internal deletions in NSm coding region, in which ΔNSm _V has deletion of NSm domains I to IV (resides 311 to 456) except domain-V (SPGC) and pTmBUNMΔNSm _I has deletion of NSm domains II to V (resides 332 to 477) except domain-I (SPNSm) as SP for Gc. pTMBUNM-SPNSm/HTNtm has 17 residues of SPNSm (residues 311 to 327) swapped with those of HTNV Gn TMD (TFCFGWVLAPITFIIL, residues 490 to 506) (Fig. 5A). pTmBUNM-SPmNSm contains four point mutations in the coding region of NSm domain-I (Fig. 2A). All the constructs were confirmed by DNA sequencing.
Transfection of Cells. BSRT7/5 cells were grown either in 35-mm dishes for immunoprecipitation experiments or in 12-well plate for cell fusion assay, or on 13-mm glass coverslips for immunofluorescence assays. Cells were transfected with the mixtures of plasmid DNA and transfection reagent (Lipofectamine® 2000 Transfection Reagent, Invitrogen) and incubated at 37°C for 24 hrs.

Metabolic Radiolabeling and Immunoprecipitation. At 24 hrs post transfection or post infection (p.i.), cells were labeled with [35S]methionine (PerkinElmer) for 3 hrs. For labelling of virus particles, the infected BHK21 cells grown in 175-cm² flask were labelled with [35S]methionine for 6 hrs p.i. and virus particles were purified by ultracentrifugation (4). Cells were then lysed on ice in 300 µl of non-denaturing RIPA buffer (50 mM Tris-HCl [pH7.4], 1% Triton X-100, 300 mM NaCl, 5 mM EDTA) containing a cocktail of cOmplete protease inhibitors (Roche). BUNV proteins were immunoprecipitated with anti-BUN, anti-NSm or anti-V5 antibodies conjugated to Dynabeads Protein A (Novex, Life Technologies). The beads bound to immune complexes were washed with RIPA buffer containing 0.1% Triton X-100 and with cold PBS, and bound viral proteins were analyzed by SDS-PAGE under reducing conditions.

immunofluorescence Staining. The transfected or infected cells were fixed with 4% formaldehyde-2% Sucrose-PBS and permeabilized with 0.2 % Triton X-100 in PBS before staining with specific primary antibodies and secondary antibody conjugates. Cells were examined using a Zeiss LSM confocal microscope.

BUNV Glycoprotein Fusion Assay. BSR-T7/5 cells, grown on 12-well plates, were transfected with 1 µg of either pTM1-BUNM or one of the mutant M expression plasmids. At 24 hrs post transfection, cells were treated with low pH medium (pH 5.3) for 15 min and then the medium was replaced by GMEM supplemented with 10% foetal calf serum (FCS). After incubation for 4 hrs at 37 °C the extent of cell fusion was observed following Giemsa
staining. Cells that contain more than 3 nuclei are considered as syncytia and the nuclei in 30 syncytia in total were counted. The fusion is represented as average nuclei per syncytium (n = 30).

**Virus Rescue by Reverse Genetics.** Rescue experiments were performed essentially as described previously with minor modifications (38). Sub-confluent monolayers of BSR-T7/5 cells grown in T25 were transfected with a mixture of plasmids comprising 1.0 µg each of pT7riboBUNL(+), pT7riboBUNS(+) and TVT7RBUNM(+) or one of the M cDNA mutants, and 7.5 µl Lipofectamine 2000 (Life Technologies) in 0.7 ml Opti-MEM media (Life Technologies). At 4 hrs post transfection, 4 ml growth medium was added and incubation continued at 33°C until CPE was evident.

**Virus Infection, Titration and Purification.** Cells seeded in dishes or flasks were infected at an MOI of 0.01 or 1.0 PFU/cell. The inoculum was removed after 1hrs, and cells were washed with PBS to remove the unattached viruses. Supernatants were harvested at the indicated time points. BUNV and SBV were titrated by plaque assay on BSRT7/5 cells and RVFV on BHK21 cells. The PUUV was titrated on Vero E6 cells by using the Avicel immunofocus assay. Briefly, the virus infected cells were covered with 0.6% Avicel (FMC BioPolymer) overlay medium and incubated for 7 days, fixed with 4% formaldehyde-PBS and permeabilized with 0.5% Triton-100, 20mM glycine-PBS. Cell monolayer was blocked with 4% skimmed milk, reacted with anti-PUUN and probed with peroxidase-conjugated anti-rabbit IgG. The foci were developed after incubation with True Blue peroxidase substrate (KPL). The purification of BUNV particles was described previously (4).

**Preparation of BUNV Glycoprotein Gn, Gc and NSm Proteins.** BUNV NSm protein was purified from lysate of the rBUNV-NSmV5 infected cells using anti-V5 Affinity gel (biotool.com). Briefly, BHK-21 cells were infected with rBUNV-NSmV5 and cells were harvested 48 hrs p.i. when cells showed obvious cytopathic effect (CPE). The cell pellet was
lysed with non-denaturing RIPA buffer and mixed with anti-V5 Affinity bead. The V5-tagged
NSm protein (NSmV5) was separated by 15% Tris-glycine SDS-PAGE (Fig. S3C). For
preparation of BUNV Gn and Gc proteins, the purified BUNV particles were separated by
12.5% Tris-glycine SDS-PAGE and stained with Coomassie Brilliant Blue G-250 solution (Fig.
S3E). Gn and Gc bands were excised for mass spectrometric analysis.

**Mass Spectrometric (MS) Analysis.** Protein bands were cut into 1mm cubes, which
were then subjected to in-gel digestion with either glutamyl endopeptidase (GluC) (NEB),
AspN (NEB) or trypsin (Promega) by a ProGest Investigator in-gel digestion robot (Genomic
Solutions, Ann Arbor, MI) using standard protocols (39). Briefly the gel cubes were de-
stained by washing with acetonitrile and subjected to reduction and alkylation before
digestion with endoproteinase at 37°C. The peptides were then extracted with 10% formic
acid and concentrated down to 20 μL using a SpeedVac (ThermoSavant). In some GluC
digestions of Gn protein PMSF (1mM) was added to prevent any trypsin-like proteolytic
cleavage. The peptides were then separated on an Acclaim PepMap 100 C18 trap and an
Acclaim PepMap RSLC C18 column (Thermo Fisher Scientific), using a nanoLC Ultra 2D plus
loading pump and nanoLC as-2 autosampler (Eksigent). The peptides were eluted with a
gradient of increasing acetonitrile, containing 0.1 % formic acid (5-40% acetonitrile in 5 min,
40-95% in a further 1 min, followed by 95% acetonitrile to clean the column, before
reequilibration to 5% acetonitrile). The eluate was sprayed into a TripleTOF 5600+
electrospray tandem mass spectrometer (Sciex) and analyzed in Information Dependent
Acquisition (IDA) mode, performing cycles of 250 msec of MS followed by 100 msec MSMS
analyses on the 15 most intense peaks seen by MS. The MS/MS data file generated via the
‘Create mgf file’ script in PeakView (Sciex) was analyzed using the Mascot algorithm (Matrix
Science), against an internal database to which the amino acid sequence of the BUNV
glycoprotein precursor (Accession number P04505), had been added, with GluC, AspN, or
trypsin as the cleavage enzyme and carbamidomethyl as a fixed modification of cysteines
and methionine oxidation and deamidation of glutamines and asparagines as a variable
modifications, followed by an ‘error tolerant’ search, to look for peptides where only one
end conforms to the enzyme specificity criteria, in order to delineate the C-terminal peptide
sequence.

**SPP Gene (HM13) Knockdown by Lentivirus shRNA.** Short hairpin RNA (shRNA) cDNAs
specific for human SPP (nt 777 – 795 for shSPP1 and 1034 – 1052 for shSPP2, Accession No:
AJ420895) and eGFP (nt 441 – 463, Accession No: U57607) as control were cloned into
pLKO.1 puro (Addgene). Lentivirus stocks were prepared in HEK 293T cells by transfection of
constructs: pLKO, pMD-G (pVSVG), a plasmid expressing the vesicular stomatitis virus
glycoprotein (VSV-G)(40), and pCMV∆R8.91 (pCMV-R), a plasmid expressing the gag/pol rev
genes of HIV-1(41), with Lipofectamine 2000 (Invitrogen). shRNA expression in A549 or
A549V cells was achieved by transducing cells with shRNA-expressing lentiviruses for 48 hrs
in the presence of 8 µg/ml polybrene (Santa Cruz Biotechnology).

**Western Blotting (WB).** Lysates of plasmid-transfected or virus-infected cells were prepared
by the addition of 100 µl NuPAGE LDS sample buffer (Life Technologies) containing 1µl
Benzonase nuclease (Merck Millipore). Proteins were separated on 4 - 12% Bis-Tris protein
gel (Life technologies) and transferred to a Nitrocellulose Blotting Membrane (GE
Healthcare). The membrane was probed with appropriate antibodies and signals detected
by chemiluminescence.

**SI Figure Legends**

**Fig. S1.** Schematic of BUNV GPC and mutations at RVAR motif and Amino acid sequence
alignment at the BUNV Gn-NSm junction. (A) BUNV GPC with positions of amino acid
residues marking the predicted boundaries (Gn, NSm and Gc) indicated. The amino acid
sequence and the secondary structure prediction at the Gn-NSm junction (residues 295 to
(http://www.enzim.hu/hmmtop/) with ‘H’ indicating hydrophobic transmembrane helix, ‘i’ residue on inside and ‘O’ residues outside of transmembrane domains (TM/TMD). The NSm domains (I to V) (7) are shown under the NSm coding region. Signal peptide (SP) and TM are shown as grey and black boxes respectively. (B) Alignment between BUNV and SSHV GPCs. The previously predicted last C-terminal residues are arginine (R) at residue 302 for BUNV and residue 299 for SSHV (in bold). (C) Amino acid sequence alignment of 13 orthobunyavirus GPCs (aligned to residues 294 to 335 of BUNV GPC). The alignment was created by using EMBL-EBI Clustal Omega program (www.ebi.ac.uk/Tools/msa/clustalo/). The conserved arginine residues at residue 302 (for BUNV) are in bold. The Accession numbers of orthobunyavirus glycoprotein precursors used in alignment: Bunyamwera virus (BUNV, P04505), Maguari virus (MAGV, AAQ23639), Kairi virus (KRIV, ACV89517), Wyelope virus (WYOV, AGA54137), Cachoeira Porteira virus (CPV, AEZ35284), Snowshoe hare virus (SSHV, ABX47014), La Crosse virus (LACV, AAB62804), California encephalitis virus (CEV, AAD53039), Inkoo virus (INKV, AAB93841), Schmallenberg virus (SBV, AGC93514), Douglas virus (DOUV, HE795091), Simbu virus (SIMV, YP_006590085), Oropouche virus (OROV, AGH07923) and Oya virus (OYAV, AGS56983).

**Fig. S2.** Effect of Furin inhibitor on BUNV infection. BUNV-infected Vero E6 cells (MOI of 0.01 pfu/cell) were incubated for 30 hours in the presence of different concentrations (0, 12.5, 25, 50 and 100 µM) of furin inhibitor I (dec-RVKR-cmk, Calbiochem) and viral replication was evaluated by immunoprecipitation and virus titration. (A) Effect of the furin inhibitor I on GPC cleavage. Lane 2 is mock infected. (B) Yield of virus in the presence of Furin inhibitor I, the concentrations of which are similar to those that inhibit furin-mediated cleavage of the respiratory syncytial virus (RSV) fusion protein (42).
**Fig. S3.** Mass spectrometric (MS) analysis of BUNV NSm. (A) SP Prediction of NSm domain-I using SignalP 4.1 ([www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/)) (19) with cleavage between residues 331G and 332T. (B) The V5-tagged NSm (NSmV5). V5 epitope is inserted in the NSm cytoplasmic domain (IV) between residues 403 and 420. (C) Purification of NSm from cells infected with recombinant rBUNV-NSmV5. The V5-tagged NSm gel band was excised, subject to in-gel digestion and mass spectrometric analysis. (D) MS/MS fragment pattern for the NSm N-terminal residues.

**Fig. S4.** Mass spectrometric (MS) analysis of BUNV Gn and Gc. (A) Coomassie Blue stained viral proteins of purified virus particles. (B to D) MS/MS fragment patterns for the terminal residues at Gn N-terminus (B), Gn C-terminal region (C) and Gc N-terminus (D).

**Fig. S5.** shGFP lentivirus contains antivirus activity. (A) Biological IFN assay. A549/Npro cells were treated with series-diluted IFNα (pbI assay science), lentivirus or cell medium for 24 hrs and then challenged with EMCV (at moi of 0.1) for 48 hrs or until obvious CPE occurs. Cells were fixed and stained 4% formaldehyde-0.1% crystal violet solution. (B) Comparison of BUNV and PUUV infections responsive to shGFP lentivirus. The induced and naive cells were treated with different dilutions of shGFP lentivirus and infected with either BUNV or PUUV at moi of 0.1. The result was shown as fold inhibition of virus titres between the induced and non-induced cells.
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Amino acid sequence and position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gn</td>
<td></td>
</tr>
<tr>
<td>N-terminal</td>
<td>GluC +PMSF</td>
</tr>
<tr>
<td>Gn</td>
<td>GluC +PMSF</td>
</tr>
<tr>
<td>C-terminal</td>
<td>AspN</td>
</tr>
<tr>
<td></td>
<td>Tryp</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>NSm</td>
<td>GluC</td>
</tr>
<tr>
<td>N-terminal</td>
<td></td>
</tr>
<tr>
<td>NSm</td>
<td>GluC</td>
</tr>
<tr>
<td>C-terminal</td>
<td>Tryp</td>
</tr>
<tr>
<td>Gc</td>
<td>Tryp</td>
</tr>
<tr>
<td>N-terminal</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation of Endoproteinases: GluC, Glutamyl endopeptidase; AspN, Peptidyl-Asp metalloendopeptidase, Tryp, trypsin.

REFERENCES


