

# SCIENTIFIC REPORTS



OPEN

## Chikungunya virus infectivity, RNA replication and non-structural polyprotein processing depend on the nsP2 protease's active site cysteine residue

Received: 02 September 2016

Accepted: 24 October 2016

Published: 15 November 2016

Kai Rausalu<sup>1</sup>, Age Utt<sup>1</sup>, Tania Quirin<sup>2</sup>, Finny S. Varghese<sup>2</sup>, Eva Žusinaite<sup>1</sup>, Pratyush Kumar Das<sup>1,†</sup>, Tero Ahola<sup>2</sup> & Andres Merits<sup>1</sup>

Chikungunya virus (CHIKV), genus *Alphavirus*, family *Togaviridae*, has a positive-strand RNA genome approximately 12 kb in length. In infected cells, the genome is translated into non-structural polyprotein P1234, an inactive precursor of the viral replicase, which is activated by cleavages carried out by the non-structural protease, nsP2. We have characterized CHIKV nsP2 using both cell-free and cell-based assays. First, we show that Cys478 residue in the active site of CHIKV nsP2 is indispensable for P1234 processing. Second, the substrate requirements of CHIKV nsP2 are quite similar to those of nsP2 of related Semliki Forest virus (SFV). Third, substitution of Ser482 residue, recently reported to contribute to the protease activity of nsP2, with Ala has almost no negative effect on the protease activity of CHIKV nsP2. Fourth, Cys478 to Ala as well as Trp479 to Ala mutations in nsP2 completely abolished RNA replication in CHIKV and SFV *trans*-replication systems. In contrast, *trans*-replicases with Ser482 to Ala mutation were similar to wild type counterparts. Fifth, Cys478 to Ala as well as Trp479 to Ala mutations in nsP2 abolished the rescue of infectious virus from CHIKV RNA transcripts while Ser482 to Ala mutation had no effect. Thus, CHIKV nsP2 is a cysteine protease.

Chikungunya virus (CHIKV) belongs to genus *Alphavirus* (family *Togaviridae*). It is transmitted by *Aedes* mosquitoes and has caused several massive outbreaks since 2004<sup>1</sup>. CHIKV has a positive-strand RNA genome approximately 12 kb in length. A large 5' open reading frame (ORF), which covers 2/3 of the viral genome, is translated directly from the viral genomic RNA and encodes for a non-structural (ns) polyprotein designated as P1234. The second ORF encodes for the precursors of viral structural proteins and is translated from a specific subgenomic RNA synthesized in virus-infected cells<sup>2</sup>.

All virus-specific enzymatic activities, required for viral RNA synthesis, are present in P1234 and its cleavage products<sup>3</sup>. Although uncleaved P1234 possesses several enzymatic activities<sup>4</sup>, it is incapable of performing viral RNA replication at detectable levels<sup>5</sup>. To become active, P1234 must first be processed into the early replicase (P123 polyprotein + nsP4). This complex can, in principle, perform all the essential steps of viral RNA synthesis<sup>6</sup>. However, during alphavirus infection the early replicase is converted into the mature (nsP1 + nsP2 + nsP3 + nsP4) form<sup>7</sup>. All these cleavages are performed by a protease located at the C-terminus of nsP2<sup>8–10</sup> and the processing of P1234 is regulated at multiple levels<sup>11–14</sup>.

The protease part of nsP2 can be easily purified as an active recombinant protein<sup>10,15</sup>. In addition, the protease activity of nsP2 can be studied using *in vitro* translation and cell culture models<sup>9,16</sup>. Early studies showed that alphavirus nsP2 is similar to papaine-like proteases. The catalytic dyad of Sindbis virus (SINV) nsP2 is represented by Cys481 and His558 residues<sup>8,17,18</sup>; these correspond to Cys478 and His548 in CHIKV and Semliki Forest virus (SFV) nsP2. Mutation of these residues results in the complete loss of protease activity<sup>19</sup> and abolishes the

<sup>1</sup>Institute of Technology, University of Tartu, Tartu, Estonia. <sup>2</sup>Department of Food and Environmental Sciences, University of Helsinki, Helsinki, Finland. <sup>†</sup>Present address: Genome-Scale Biology Research program, Institute of Biomedicine, University of Helsinki, Finland. Correspondence and requests for materials should be addressed to A.M. (email: andres.merits@ut.ee)

infectivity of SINV genomes<sup>18</sup>. Different cleavage sites in P1234 are recognized in different ways. Specifically, SFV nsP2 cannot cleave short substrates representing the cleavage site between nsP2 and nsP3 (2/3 site). This cleavage requires both the native N-terminus of nsP2 and a long substrate comprising a few P-side residues followed by approximately 165 N-terminal residues of nsP3<sup>12,15</sup>.

The 3D structures of the proteases from Venezuelan equine encephalitis virus (VEEV)<sup>20</sup>, SINV<sup>21</sup> and CHIKV (Protein Data Bank code 3TRK) reveal that nsP2 protease is a member of MEROPS Clan CN and contains a papain-like protease linked to a C-terminal domain resembling FtsJ-like methyltransferases (MTL domain)<sup>20</sup>. Subsequent molecular modelling revealed the features of the catalytic site and the S1'-S4 subsites<sup>22</sup>. A peptidomimetic inhibitor of nsP2 protease binds at the interface of the protease and MTL domains resulting in conformational change that most probably assists in leaving group departure of either the amine or Cys thiolate during the catalytic cycle<sup>23</sup>; the binding mode of natural substrates is most likely similar. In addition to the catalytic Cys, Asn475 and Lys480 residues (correspond to residues 476 and 481 of CHIKV nsP2) have been shown to be important for the protease activity of VEEV nsP2<sup>23</sup>. Known 3D structure coupled with the functional importance of nsP2 have made this protein an attractive target for the development of inhibitors of alphavirus infection<sup>23-28</sup>.

CHIKV nsP2 possesses all the enzymatic activities known for nsP2 derived from other alphaviruses<sup>29,30</sup>. However, data concerning the protease activity of CHIKV nsP2 is conflicting. On one hand, studies performed using the purified protease part of nsP2<sup>31</sup> or full length nsP2 with a native N-terminus<sup>16</sup> have demonstrated that CHIKV nsP2 is functionally similar to SFV nsP2. On the other hand, Saisawang and co-authors observed that CHIKV nsP2, derived from East/Central/South African (ECSA) isolate from Thailand, differs from other alphavirus nsP2 proteases in the recognition of small peptide substrates<sup>32</sup>. Recently the same team also reported that the catalytic dyad Cys478 of CHIKV nsP2 could be interchangeable with a proximal Ser482 residue that also contributes to the protease activity of nsP2<sup>33</sup>. Thus, CHIKV nsP2 was reported to have properties very different from those revealed for the ns proteases of other alphaviruses. Indeed, functions of the same ns protein originating from different alphaviruses can be substantially different<sup>34,35</sup>. However, in this case the data was obtained using only a single short peptide substrate based assay and the conclusion was not directly confirmed by comparing the enzymes from different alphaviruses in the same experiment. Furthermore, this finding has not been verified using cell-based experiments.

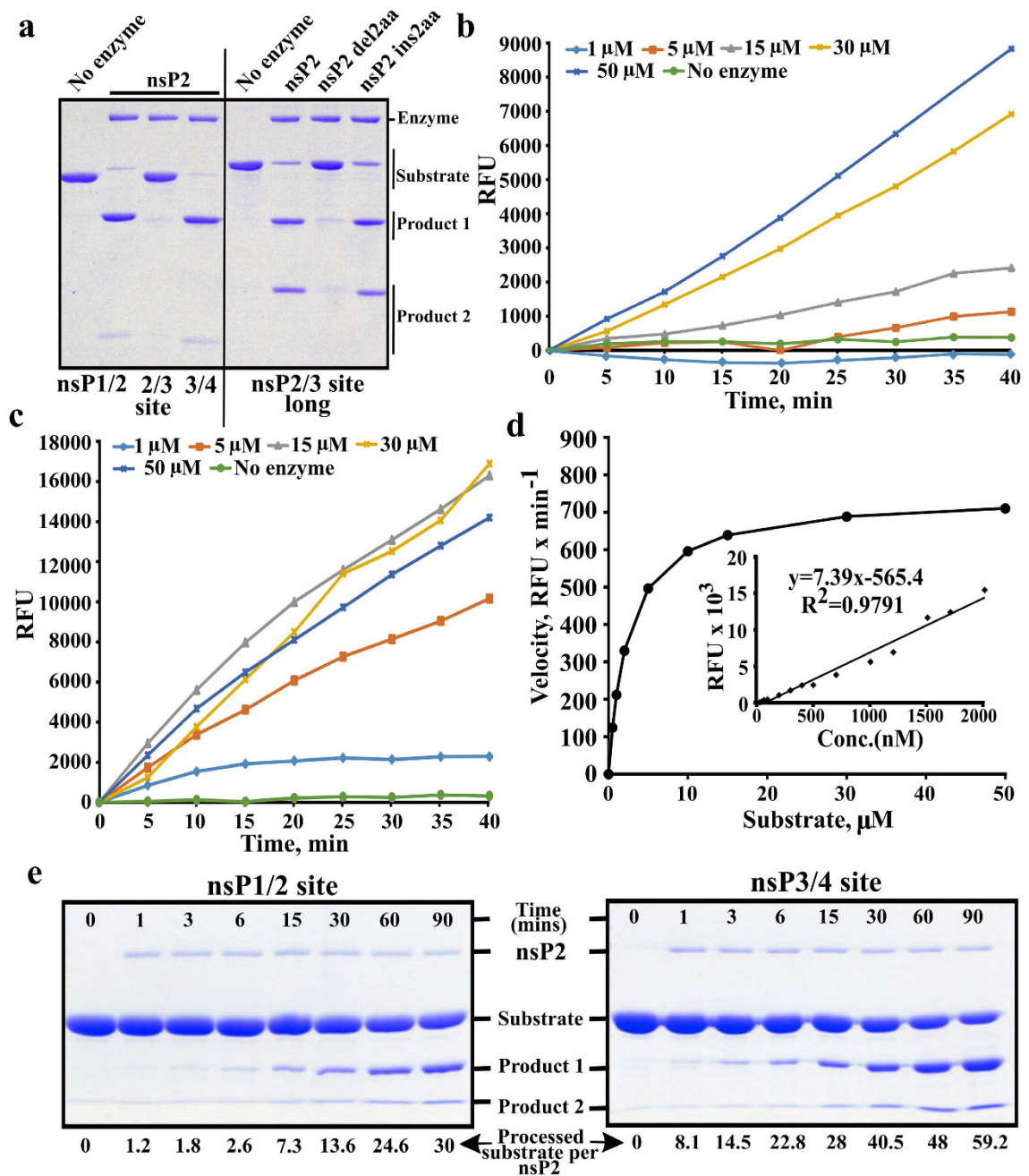
As unusual properties of CHIKV nsP2 would have implications for the molecular biology of this virus as well as for the development of inhibitors of CHIKV infection, we evaluated the properties of CHIKV nsP2 using multiple verified cell-free and cell-based assays. This analysis revealed that Cys478 residue is indispensable for CHIKV ns polypeptide processing, RNA replication and viral infectivity. Substitution of Ser482 residue did not affect any of these crucial functions. Thus, it must be asserted that CHIKV nsP2 is a classical alphavirus ns protease with properties very similar to those previously revealed for SFV nsP2.

## Results

**Substrate requirements for CHIKV nsP2 protease are similar to those of SFV nsP2.** CHIKV nsP2 readily cleaved GFP-10:5-Trx substrates containing short P10-P5' sequences originating from the 1/2 and 3/4 sites but not from the 2/3 site. Consistent with the requirement of nsP3 macro-domain for 2/3 site cleavage, the longer substrate GFP-10:170 was efficiently cleaved (Fig. 1a). Another critical determinant for 2/3 site cleavage, revealed in studies of SFV nsP2, is a requirement for the intact native N-terminus of nsP2<sup>12</sup>. It was found that the removal of just two amino acid residues from the N-terminus of CHIKV nsP2 almost completely abolished the processing of GFP-10:170. However, in contrast to SFV nsP2, the addition of two extra amino acid residues (Gly-Ala) to the N-terminus of CHIKV nsP2 had only a minor effect on the cleavage of GFP-10:170 (Fig. 1a).

Next, experiments were performed using fluorogenic peptide substrates. It has been demonstrated that residues P6 to P1 have a clear impact on the cleavage in the context of protein-based substrates and, most importantly, in the context of infectious virus<sup>11,15</sup>. Accordingly, peptide substrates used previously have been sub-optimal, as they contained only three or four P-side residues<sup>31-33</sup>. Based on the above considerations, peptide substrates representing the P6-P2' regions of 1/2 and 3/4 sites were designed. However, due to problems in peptide synthesis, we could not obtain a peptide representing the 1/2 site with required (>90%) purity. Therefore only the substrate with the sequence LDRAGG↓YI (designated as Short), representing the 3/4 site, was used in subsequent experiments. This substrate was processed by CHIKV nsP2, but the reaction velocity could not be saturated (Fig. 1b). This indicated that the Short substrate was still sub-optimal. Previously, we observed that SFV nsP2 cleaves a substrate having 10 amino acid residues from the P-side more efficiently than substrates having only 6 P-side residues (our unpublished data). Therefore a longer substrate, spanning positions P10 to P5' (DELRLDRAGG↓YIFSS; designated as Long), was used<sup>36</sup>. This substrate was processed more efficiently than Short (compare Fig. 1b,c). Importantly, the reaction velocity reached a plateau (Fig. 1d) indicating that Long peptide represents a preferred tool for functional studies.

The protease domain of SFV nsP2 is capable of hydrolysing a 400-fold molar excess of a protein-based substrate in 5 min indicating a cleavage rate of  $\geq 1.3 \text{ s}^{-1}$  or  $\geq 78 \text{ min}^{-1}$ <sup>10</sup>. This is largely comparable with data observed for many virus-encoded proteases<sup>37</sup>. In contrast, Saisawang and colleagues determined a  $k_{\text{cat}}$  for CHIKV nsP2 as  $0.0016 \text{ s}^{-1}$  or  $0.096 \text{ min}^{-1}$ <sup>33</sup>. Using a previously described assay<sup>10</sup> it was found that in 3 min CHIKV nsP2 was able to process approximately 1.8-fold molar excess of a substrate corresponding to the 1/2 site and approximately 14.5-fold molar excess of a substrate representing the 3/4 site (Fig. 1e). Thus, similar to SFV nsP2, the CHIKV enzyme prefers the 3/4 substrate for cleavage *in trans*. However, in CHIKV the difference in cleavage efficiencies for these two substrates was considerably smaller than previously observed for SFV<sup>10</sup>, probably because both 1/2 and 3/4 sites of CHIKV have preferred P4 Arg residues<sup>15</sup>. The cleavage rate for 3/4 substrate was  $\approx 0.08 \text{ s}^{-1}$  or  $4.8 \text{ min}^{-1}$ . This activity is very similar to that observed for VEEV protease domain using similar substrates<sup>23</sup> and approximately 50-fold higher than reported by Saisawang and colleagues using peptide substrates<sup>33</sup>.



**Figure 1. Processing of different substrates by CHIKV nsP2.** (a) Processing of recombinant protein substrates. Reaction mixtures were incubated for 60 min and reaction products were resolved using 12% SDS-PAGE. Substrates contained either short (10:5) peptides originating from 1/2, 2/3 and 3/4 sites (left side of the panel) or a longer (10:170) region representing 2/3 site and macro domain of nsP3 (right side of the panel). Enzymes used to perform cleavage are shown above the panel. (b) Processing of Short peptide substrate. The concentration of nsP2 was kept constant (78 nM) while concentration of the substrate ranged from 1 to 50  $\mu\text{M}$  (shown above graphs); in the control reaction 15  $\mu\text{M}$  of substrate was incubated without the enzyme. The fluorescence resulting from cleavage is shown as the function of time. (c) Processing of Long peptide substrate. The experiment was performed as in panel b. (d) Enzyme kinetic constant  $K_m$  was derived by plotting the initial reaction velocity on vertical axis against increasing Long peptide substrate concentration (0.5, 1, 2, 5, 10, 15, 30, 50  $\mu\text{M}$ ) on horizontal axis. Inset: representative plot of EDANS fluorescence generated by treating increasing concentrations of Long peptide substrate with proteinase K (50  $\mu\text{g}$  per reaction). (e) Processing kinetics of wt nsP2 revealed using recombinant protein based substrates. Wt nsP2 was mixed with substrates corresponding to 1/2 (left panel) and 3/4 (right panel) sites using molar ratio 1:100. Reactions were carried out at 30  $^\circ\text{C}$ ; aliquots were collected at 0 (before adding enzyme), 1, 3, 6, 15, 30, 60 and 90 min, and analysed using 10% SDS-PAGE. ImageJ program (NIH, USA) was used to quantify substrate band intensities. The values were further analysed using MS Excel software and expressed as pmol of processed substrate per pmol of nsP2 as indicated at the bottom of the figure. All experiments were repeated at least twice; data from one of reproducible experiment is shown.

Finally, we performed estimation of  $K_m$  and  $k_{cat}$  using Long substrate and CHIKV nsP2 with a native N-terminus. By carrying out a protease assay with increasing concentration of Long substrate, the apparent  $K_m$  was found to be  $2.5 \pm 0.1 \mu\text{M}$  (Fig. 1d). In order to estimate  $k_{cat}$  for this reaction an approach used for SARS coronavirus 3C protease<sup>38</sup> taking into account secondary quenching of EDANS (5-[(2'-aminoethyl)-amino] naphthalenesulfonic acid) fluorescence, was used. The treatment of Long peptide substrate with 20, 30 or 50  $\mu\text{g}$  of proteinase K reproducibly generated comparable intensity of fluorescence indicating that the substrate was cleaved close to 100% and that proteinase K itself doesn't affect the fluorescence (Fig. 1d). Using this calibration, the  $k_{cat}$  of CHIKV nsP2 for cleavage of Long substrate was found to be  $\approx 2.5 \text{ min}^{-1}$ , similar to that observed using recombinant protein substrates. Taken together this data also indicates that unlike sub-optimal peptides, a properly designed one (such as Long) does represent a suitable substrate for CHIKV nsP2 protease.

**Substitution of catalytic Cys478 residue blocks the ability of nsP2 to cleave substrates corresponding to 1/2 and 3/4 sites.** We mutated the proposed catalytic Cys478 residue of CHIKV nsP2 to Ala (C478A), the Ser482 residue, reported to be able to substitute for Cys478 residue<sup>33</sup> to Ala (S482A), and combined these mutations (C478A + S482A). We also mutated the Trp479 residue, adjacent to catalytic Cys478, to Ala (W479A). To obtain purified mutant nsP2 proteins, we first took advantage of the ability of CHIKV nsP2 to tolerate N-terminal tags. Using this approach recombinant wild type (wt) His-nsP2, His-nsP2<sup>C478A</sup>, His-nsP2<sup>S482A</sup> and His-nsP2<sup>C478A+S482A</sup> were all obtained at high purity (Fig. 2a). The identities of all of recombinant proteins were confirmed using mass-spectrometry. The purification of His-nsP2<sup>W479A</sup> turned out to be difficult; the protein contained a substantial amount of a faster migrating contaminating protein (Fig. 2a). No major *E. coli* derived protein was revealed by mass-spectrometry in His-nsP2<sup>W479A</sup> indicating that the faster migrating protein represents truncated form(s) of His-nsP2<sup>W479A</sup>. Next, the CD-spectra of all purified recombinant proteins were obtained and compared to each other. This analysis revealed that all these proteins were folded similarly with no detectable structural deformity from nsP2 or His-nsP2 (Fig. 2b,c). However, His-nsP2<sup>W479A</sup> exhibited a CD spectrum with abnormally high ellipticity (Fig. 2c). Together with altered size-exclusion chromatography profiles (not shown) this indicates that His-nsP2<sup>W479A</sup> most probably forms high molecular weight aggregates.

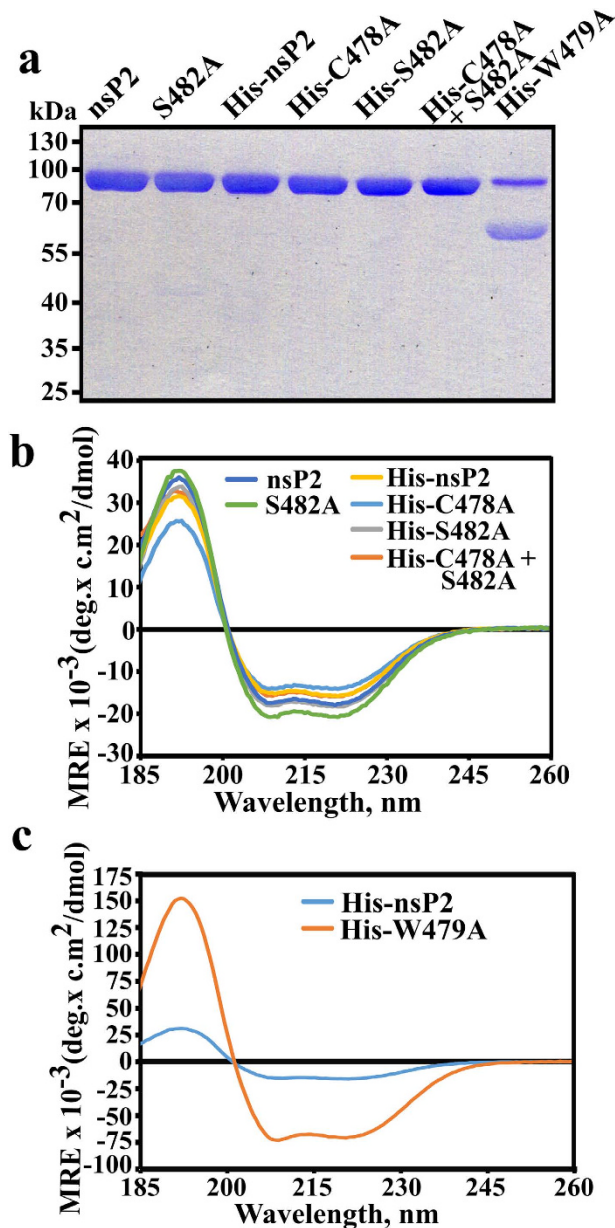
Next we compared the abilities of His-tagged nsP2 proteins to cleave recombinant protein-based substrates. None of the enzymes was able to cleave recombinant protein containing a short version of the 2/3 site with detectable efficiency. His-nsP2<sup>S482A</sup> was, however, capable of cleaving substrates containing 1/2, 3/4 and 2/3 (long) sites with efficiency similar to that of wt His-nsP2 (Fig. 3a). His-nsP2<sup>W479A</sup> was completely inactive, with the caveat of formation of inactive aggregates. Importantly, His-nsP2<sup>C478A</sup> and His-nsP2<sup>C478A+S482A</sup> were both also completely unable to cleave 1/2 and 3/4 substrates (Fig. 3a). Surprisingly, both of these mutants were able to cleave, albeit with much reduced efficiencies (compared to wt His-nsP2), 2/3 (long) substrate (Fig. 3a). This result highlights the unique nature of 2/3 site cleavage. Peptide substrates led to results that were identical to those obtained for the corresponding (3/4) recombinant protein based substrate: His-nsP2<sup>C478A</sup>, His-nsP2<sup>C478A+S482A</sup> and His-nsP2<sup>W479A</sup> were all defective in cleavage of both Long and Short substrates while His-nsP2 and His-nsP2<sup>S482A</sup> had similar activities (Fig. 3b).

As His-tag alters several enzymatic and non-enzymatic properties of nsP2<sup>29</sup> it was important to compare the activities of His-tagged nsP2 with those of non-tagged protein. Direct comparison of wt nsP2 and His-nsP2 confirmed that both proteins are active. Using recombinant protein substrates, we were unable to detect significant differences between these two enzymes (Fig. 3c). The assay using peptide substrates revealed that His-nsP2 was slightly (<2-fold) but reproducibly less active than wt nsP2 (Fig. 3e). Thus, the N-terminal His-tag has a small influence on protease activity of CHIKV nsP2.

Next, we introduced the S482A mutation into pET-nsP2 vector designed for production of nsP2 with authentic N-terminus *via* removal of the N-terminal tag by the protease activity of the protein itself<sup>29</sup>. nsP2<sup>S482A</sup> was successfully purified (Fig. 2a). This approach was not applicable for other mutant nsP2 proteins, which were clearly unable to process 1/2 and 3/4 substrates (Fig. 3a,b). Comparison of nsP2 and nsP2<sup>S482A</sup> in the same experiment revealed no differences in the cleavage of recombinant protein based substrates. Expectedly, both enzymes were also capable of cleaving peptide substrates; in these assays nsP2<sup>S482A</sup> was slightly, but reproducibly, less active than wt nsP2 (Fig. 3e).

**Cys478 and Trp479 residues of nsP2 are required for processing of CHIKV P1234 polyprotein.** Our analysis carried out using recombinant proteins and different types of substrates shows that unlike reported recently<sup>33</sup>, Cys478 residue is indispensable and irreplaceable for the ability of CHIKV nsP2 to process any substrate corresponding to 1/2 and 3/4 sites. However, the used systems are artificial: in alphavirus infection the polyprotein context may affect both the activity of the enzyme and the presentation of its substrates<sup>11</sup>.

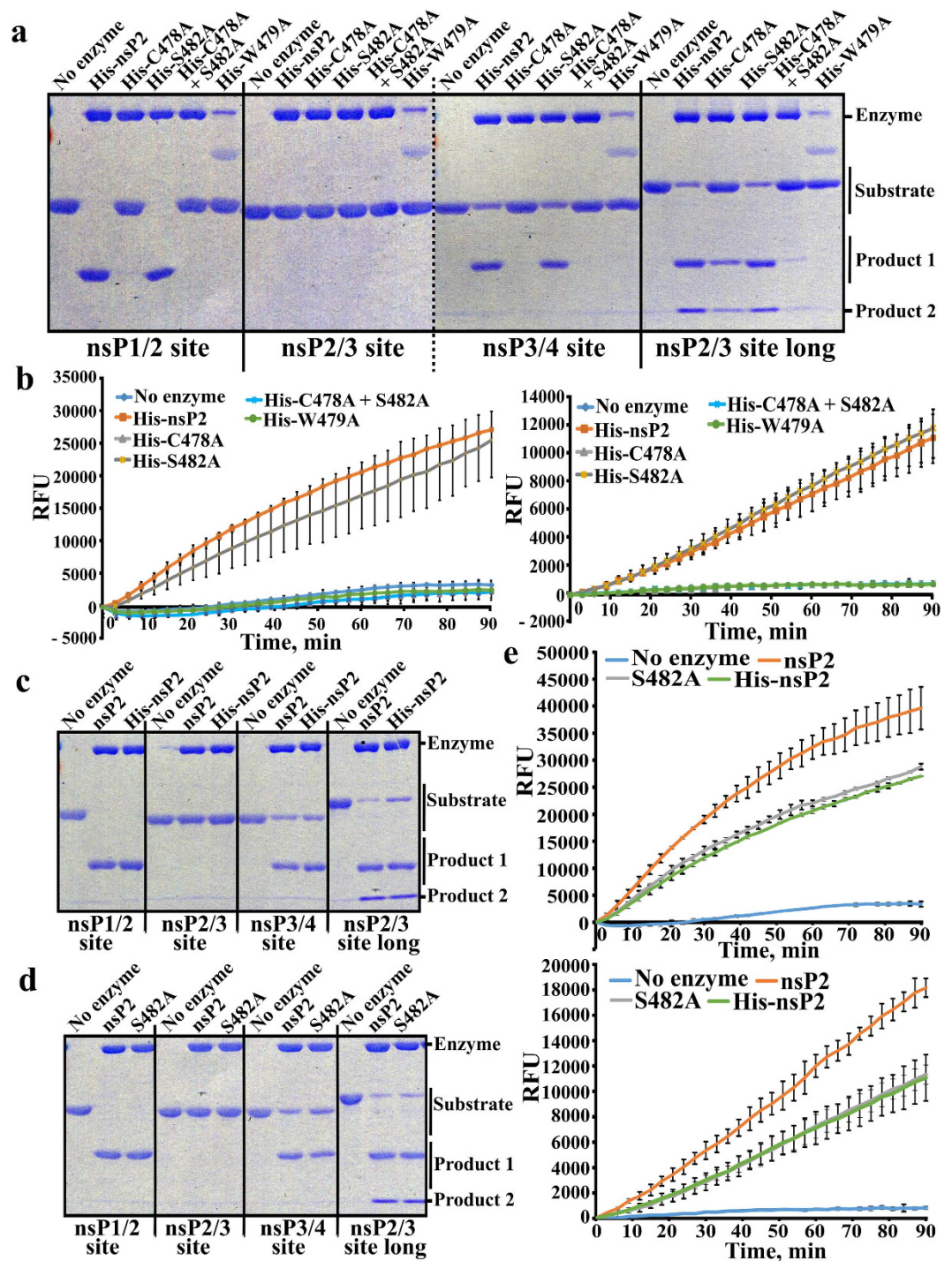
To analyse the effects of mutations in the nsP2 active site region on the processing of CHIKV P1234 polyprotein C478A, W479A, S482A and C478A + S482A substitutions were introduced into T7-P1234 plasmid (Fig. 4a)<sup>39</sup>. The plasmids, together with wt T7-P1234 and T7-P1234<sup>GAA</sup> polymerase mutant controls<sup>39</sup>, were transcribed/translated *in vitro*, CHIKV ns proteins were precipitated using antibodies against nsP1, nsP2, nsP3 and nsP4 and resolved by SDS-PAGE. The only specific translation product of T7-P12<sup>CA</sup>34, T7-P12<sup>WA</sup>34 and T7-P12<sup>CA+SA</sup>34 was a large polyprotein; precipitation by all four antibodies confirmed that it was P1234 (Fig. 4b). We could not reliably detect products that could correspond to P12 and P34 polyproteins that would result from the residual processing of 2/3 site by mutant proteases. Thus, it should be concluded that protease present in P12<sup>CA</sup>34, P12<sup>WA</sup>34 and P12<sup>CA+SA</sup>34 polyproteins were completely inactivated. This data confirms that Ser482 cannot substitute for Cys478. The lack of processing of T7-P12<sup>WA</sup>34 encoded polyprotein indicates that Trp479 is also crucial for nsP2 protease activity. In contrast, all mature ns proteins were detected among the products obtained from T7-P1234, T7-P12<sup>SA</sup>34 and T7-P1234<sup>GAA</sup> (Fig. 4b). The amounts of P1234 polyprotein were typically too low



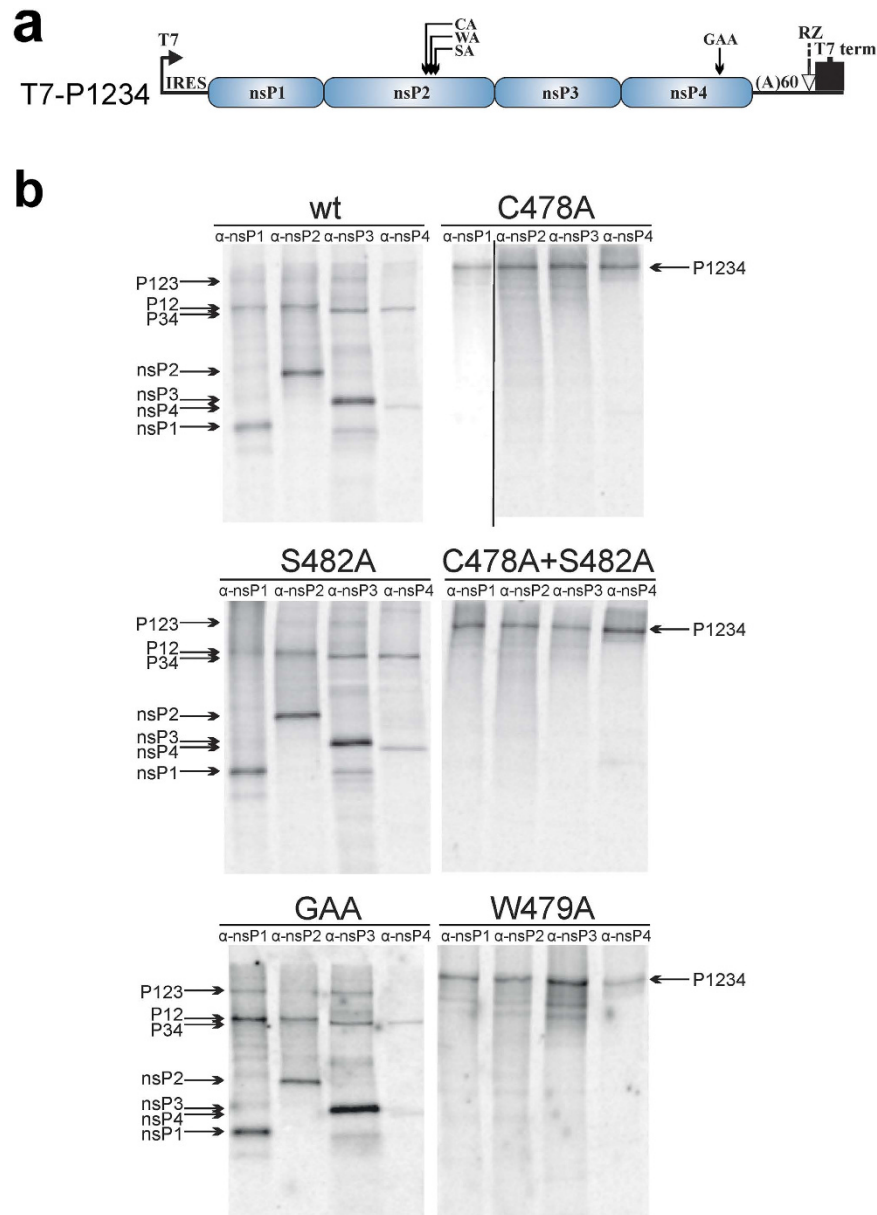
**Figure 2. Purification of recombinant nsP2 and its mutant forms.** (a) SDS-PAGE analysis of purified recombinant enzymes. The proteins are shown above the panel; each lane corresponds to approximately 1  $\mu\text{g}$  of purified protein. (b) CD spectroscopy analysis of the nsP2, nsP2<sup>S482A</sup>, His-nsP2, His-nsP2<sup>C478A</sup>, His-nsP2<sup>S482A</sup> and His-nsP2<sup>C478A+S482A</sup>. All of these recombinant proteins were found to be properly and comparably folded. (c) Comparison of CD spectra of His-nsP2 and His-nsP2<sup>W479A</sup>. The shapes of the spectra are identical (compare also with panel b) but His-nsP2<sup>W479A</sup> produces a strongly enhanced signal.

to be detected; however, several processing intermediates (P123, P12, P34) were clearly identified (Fig. 4b). Thus, the S482A substitution had no detectable effect on P1234 processing.

**C478A and W479A but not S482A substitution block processing of P1234 in transfected cells and inhibit CHIKV *trans*-replicase.** Mutations that have a strong effect on the processing of CHIKV P1234 in *in vitro* transcription/translation system do not always have the same effect in living cells<sup>16,39</sup>. Therefore BSR T7/5 (hereafter BSR) cells, a derivative of BHK-21 cells stably expressing T7 RNA polymerase<sup>40</sup>, were transfected with wt polyprotein T7-P1234 expression construct and its mutant derivatives (Fig. 5a). Cells were lysed at 18 h post transfection (h p.t.) and ns polyproteins and their processing products were analysed using Western blotting. A full set of mature ns proteins was detected in cells transfected with T7-P1234, T7-P12<sup>SA</sup>34 or T7-P1234<sup>GAA</sup> (Fig. 5b); coherent with previous observations<sup>39</sup> no ns polyproteins could be detected. In contrast, large P1234 polyprotein was the only distinct product detected in T7-P12<sup>CA</sup>34, T7-P12<sup>WA</sup>34 or T7-P12<sup>CA+SA</sup>34 transfected cells; no mature ns proteins or polyproteins that could be identified as P12 or P34 were observed (Fig. 5b). Thus,



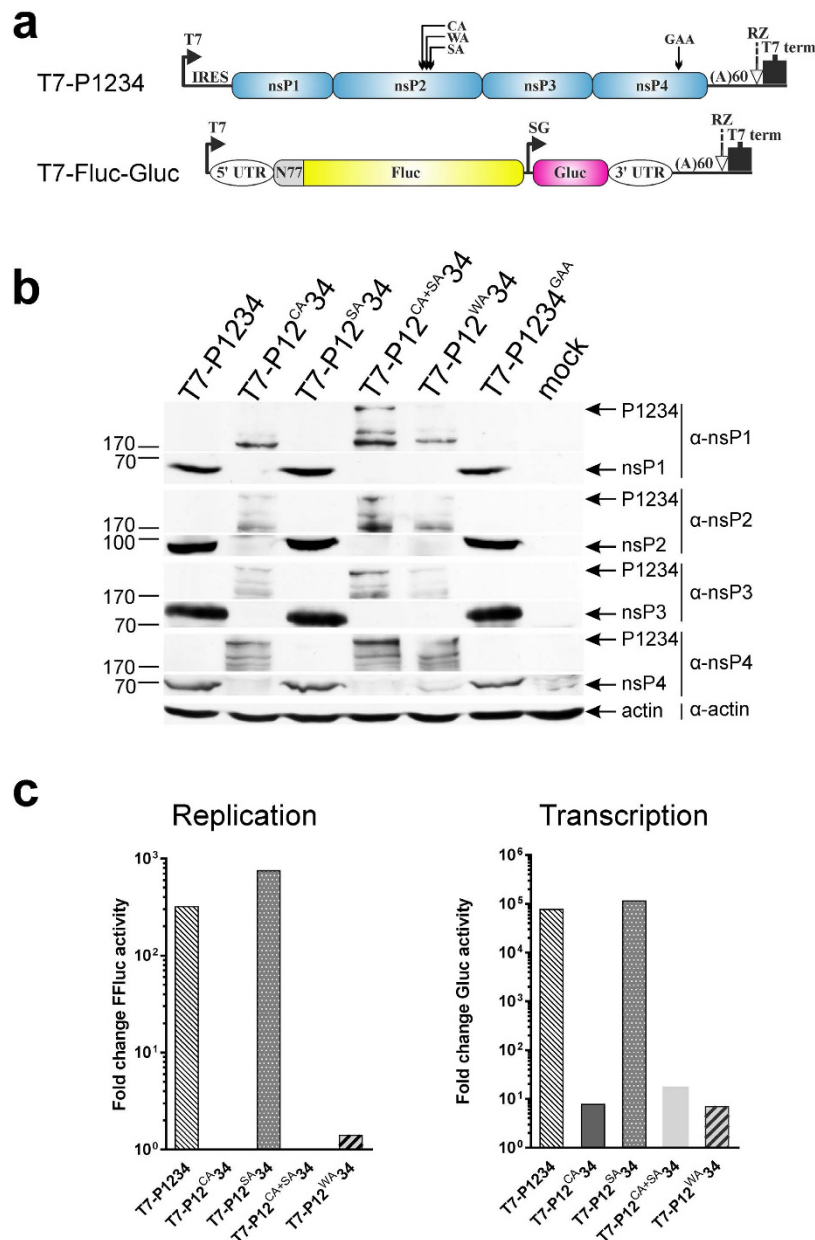
**Figure 3.** Effects of C478A, W479A and S482A substitutions on the protease activity of CHIKV nsP2. (a) Comparison of protease activities of His-nsP2, His-nsP2<sup>C478A</sup>, His-nsP2<sup>S482A</sup>, His-nsP2<sup>W479A</sup> and His-nsP2<sup>C478A+S482A</sup> using recombinant protein as substrates. The assay was performed and results analysed as described for Fig. 1a. The dotted vertical line on panel indicates merging of two different gels. (b) Comparison of protease activities of His-nsP2, His-nsP2<sup>C478A</sup>, His-nsP2<sup>S482A</sup>, His-nsP2<sup>W479A</sup> and His-nsP2<sup>C478A+S482A</sup> using the continuous FRET based assay. 1  $\mu$ g of His tagged nsP2 or its mutants was mixed with 15  $\mu$ M Long (left panel) or Short (right panel) substrate in the final reaction volume of 200  $\mu$ L/well. The 96-well plate was placed in the Synergy MX reader and the reaction was carried out at 30  $^{\circ}$ C for 90 min. Recorded EDANS fluorescence (RFU, relative fluorescence units) is shown at the vertical axes. (c) Comparison of wt nsP2 and His-nsP2 using recombinant protein based substrates. Assays were performed and data is presented as described for Fig. 1a. (d) Comparison of wt nsP2 and nsP2<sup>S482A</sup> using recombinant protein based substrates. Assays were performed and data is presented as described for Fig. 1a. (e) Comparison of wt nsP2, nsP2<sup>S482A</sup> and His-nsP2 using Long (upper panel) and Short (lower panel) peptide substrates. Assays were performed and data is presented as described for panel b. Experiments were repeated at least twice; data from one of reproducible experiment is shown.



**Figure 4. C478A and W479A substitutions in nsP2 region block the processing of CHIKV P1234 in cell-free system.** (a) Schematic presentation of the T7-P1234 plasmid used for the assay; mutations in nsP2 or nsP4 present in T7-P12<sup>CA</sup>34, T7-P12<sup>WA</sup>34, T7-P12<sup>SA</sup>34, T7-P12<sup>CA+SA</sup>34 and T7-P1234<sup>GAA</sup> are indicated above the drawing. IRES—internal ribosome entry site of encephalomyocarditis virus; RZ—antigenomic strand ribozyme from hepatitis D virus. (b) 2  $\mu$ g of a T7-P1234, T7-P12<sup>CA</sup>34, T7-P12<sup>WA</sup>34, T7-P12<sup>SA</sup>34, T7-P12<sup>CA+SA</sup>34 or T7-P1234<sup>GAA</sup> plasmids were used as templates for *in vitro* transcription and translation carried out in the presence of 20  $\mu$ Ci of [<sup>35</sup>S]methionine-cysteine mixture. Reactions were stopped by adding cycloheximide, CHIKV ns proteins and polyproteins were immunoprecipitated using antisera against nsP1, nsP2, nsP3 and nsP4. The precipitated proteins were separated by 8% SDS-PAGE and visualized using a Typhoon imager. Mutations introduced into ns polyproteins are shown above the panels, antibodies used for immunoprecipitation are indicated at the top of the respective lanes. Positions of ns proteins and polyproteins are indicated by arrows. The vertical line on C478A panel indicates merging of images obtained by different exposure, for full images see Supplementary Information. The experiment was repeated twice, data from one of reproducible experiment is shown.

effects of C478A, W479A and S482A mutations on CHIKV ns polyprotein processing observed in cell-free assays were also reproduced in cell-based assay.

The functional importance of any mutation in the alphavirus replicase is best assessed by the ability of the viral replicase to perform the amplification and transcription of suitable RNA templates. Here we took advantage of a recently developed CHIKV *trans*-replicase system<sup>39</sup>. BSR cells were co-transfected with the T7-Fluc-Gluc plasmid



**Figure 5.** C478A and W479A substitutions in nsP2 region abolish processing of P1234 and CHIKV *trans*-replication system mediated activation of reporter expression. **(a)** Schematic presentation of T7-P1234 plasmid encoding for CHIKV replicase and T7-Fluc-Gluc plasmid encoding for replication competent template RNA carrying sequences encoding for Fluc and Gluc reporter proteins. Mutations introduced into nsP2 or nsP4 are shown as in Fig. 4a. IRES – internal ribosome entry site of encephalomyocarditis virus; RZ – antigenomic strand ribozyme from hepatitis D virus, UTR – untranslated region from CHIKV genome, SG – CHIKV subgenomic promoter spanning residues –78 to +69 with respect to the start position of subgenomic RNA, N77 – region encoding for the 77 N-terminal amino acid residues of nsP1. **(b)** BSR cells were transfected with T7-P1234, T7-P12<sup>CA</sup>-34, T7-P12<sup>SA</sup>-34, T7-P12<sup>CA+SA</sup>-34 or T7-P1234<sup>GAA</sup>. At 18 h p.t. cells were lysed with Laemmli buffer, proteins were separated by 8% SDS-PAGE, transferred to membrane and detected using Western blotting and antibodies against CHIKV nsP1, nsP2, nsP3 and nsP4. Parts of the same gels corresponding to mature ns proteins and P1234 polyproteins are shown, as due to lower efficiency of detection P1234 and products of its degradation cannot be revealed using the same exposure used for detection of mature ns proteins; for full-length images see Supplementary Information. Positions of ns proteins and P1234 polyprotein are indicated by arrows on the right, positions of molecular mass markers (in kDa) are shown on the left. Actin was used as loading control. **(c)** BSR cells were co-transfected with T7-P1234, T7-P12<sup>CA</sup>-34, T7-P12<sup>SA</sup>-34, T7-P12<sup>CA+SA</sup>-34 (indicated below the graph) or T7-P1234<sup>GAA</sup> (control) and T7-Fluc-Gluc. At 18 h p.t. cells were lysed, activities of Fluc and Gluc were measured and normalized to those measured for control cells. Normalized activities of Fluc were used to estimate synthesis of full-length template RNA (Replication) by CHIKV *trans*-replicase; normalized Gluc activities were used to estimate synthesis of subgenomic RNA (Transcription). Data from one out of three reproducible independent experiments is shown.



(Fig. 5a) producing a reporter-containing replication-competent RNA<sup>39</sup> and plasmids expressing different polyprotein variants. Expression of the firefly luciferase (Fluc) marker, encoded by the 5' ORF of the replicating RNA template, was amplified more than 300-fold in the presence of T7-P1234 or T7-P12<sup>S482A</sup>34 (Fig. 5c, left panel). In contrast, no amplification of Fluc expression was observed in the presence of T7-P12<sup>CA</sup>34, T7-P12<sup>WA</sup>34 or T7-P12<sup>CA+S482A</sup>34 (Fig. 5c, left panel).

The *Gaussia* luciferase (Gluc) reporter is translated exclusively from replicase-generated subgenomic RNAs; thus its expression is strictly dependent on the ability of the replicase to synthesize RNA<sup>39</sup>. Compared to the inactive polymerase mutant, wt replicase activated Gluc expression nearly 100,000-fold. Similar activation was observed for T7-P12<sup>S482A</sup>34 (Fig. 5c, right panel) confirming that fully functional replicase complexes were assembled. Somewhat surprisingly, extremely low, but still detectable ( $\approx 10$ -fold) activation of Gluc expression was also observed in the presence of constructs harbouring C478A, W479A or C478A + S482A substitutions (Fig. 5c, right panel). The reason(s) for this small activity remain currently unknown though it is possible that unprocessed P1234 may possess minimal RNA synthesis ability. As these effects are extremely minor (activation of Gluc expression is reduced  $\approx 10,000$ -fold compared to wt P1234), they were only noticed using extremely sensitive CHIKV *trans*-replicase reporter assay. It is likely that this activity is too low to have biological significance.

### C478A and W479A mutations in P1234 polyprotein abolish RNA synthesis by CHIKV and SFV replicases.

Unlike positive-strand RNA, synthesis of negative-strand RNA cannot be revealed using reporter proteins<sup>41</sup>. Therefore the activity of CHIKV replicases harbouring nsP2 mutations was also analysed using Northern blotting. SFV replicase, harbouring the same set of mutations, was used for comparison. For unknown reasons, CHIKV replicase was rather inefficient in using template RNA derived from T7-Rluc-Tom vector (Fig. 6a); in contrast SFV replicase efficiently utilized template RNA derived from Tmed<sup>42</sup>. Nevertheless, it was clearly and consistently detected that in addition to wt replicases, only replicases harbouring S482A substitution were capable of negative-strand RNA synthesis (Fig. 6b). In fact, the synthesis of negative strands was slightly but reproducibly increased by this mutation, most likely due to slight reduction of protease activity of nsP2 (Fig. 3e) leading to increased half-life of early replicase complex.

It was also found that levels of positive-strand RNAs exceeded background levels (detected in the absence of enzymatically active replicase) only in the presence of wt replicases and those harbouring S482A substitution (Fig. 6c). Together with data obtained for negative-strand RNA synthesis, this confirms that only these replicases are active. This data complements our previous findings and clearly indicates that replicases harbouring C478A, W479A or C478A + S482A substitutions are incapable of both negative- and positive-strand RNA synthesis. Importantly, the activities of SFV replicase were affected by all protease active site mutations in the same way as those of CHIKV replicase (Fig. 6b,c). This is consistent with our data highlighting that the properties of CHIKV nsP2 protease are very similar to those previously revealed for SFV nsP2<sup>10,12,14,15</sup>.

### C478A and W479A but not S482A substitutions block the rescue of infectious CHIKV.

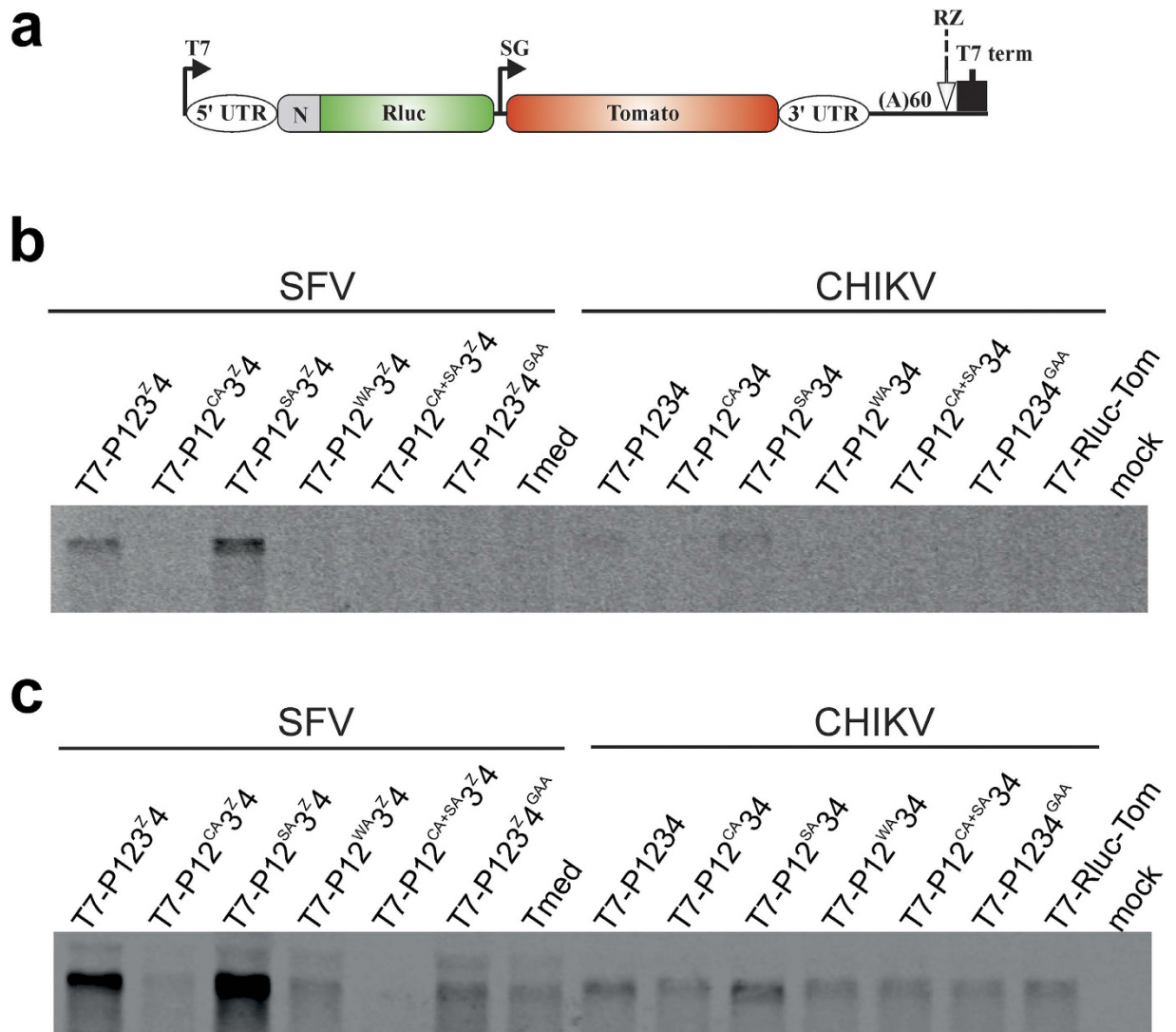
We verified the effects of the protease mutations also in the context of CHIKV RNAs obtained by transcription of infectious cDNA clones (Fig. 7a). Infectious centre assay (ICA) performed after RNA transfection revealed that the rescue efficiencies of wt CHIKV and CHIKV<sup>S482A</sup> were similar (Fig. 7a). No plaques were detected for cells transfected with the three other mutants. Next, cells were transfected with 10  $\mu$ g RNA transcripts (corresponding to  $\approx 10^6$  plaque forming units (PFU) for wt CHIKV). Cells transfected with wt CHIKV and CHIKV<sup>S482A</sup> transcripts displayed strong cytotoxic effects by 18 h p.t. In contrast, no cytotoxicity was observed for the other transcripts even at 144 h p.t. To study the presence of replicating virus genomes in the cells, all the transfected cultures were analysed for the presence of capsid protein (expression of which is completely dependent on viral RNA replication). High amounts of mature capsid protein (as well as some amounts of its precursors) were detected in cells transfected with transcripts of pSP6-CHIKV and pSP6-CHIKV<sup>S482A</sup>. In contrast, no capsid protein could be detected in cells transfected with pSP6-CHIKV<sup>C478A</sup>, pSP6-CHIKV<sup>W479A</sup> or pSP6-CHIKV<sup>C478A+S482A</sup> transcripts (Fig. 7b). Furthermore, while wt CHIKV and CHIKV<sup>S482A</sup> reached comparable high titers ( $7.5 \times 10^8$  and  $6.5 \times 10^8$  PFU/mL, respectively) and formed plaques of similar size (not shown) no infectious virus was detected in the other three samples. Taken together, this data unequivocally demonstrates that mutations C478A and W479A not only abolish ns polyprotein processing but are also lethal for CHIKV.

To analyse the effect of S482A mutation on ns polyprotein processing in the context of CHIKV infection, a pulse-chase experiment was performed. For wt CHIKV infected cells all ns proteins, P1234 precursor and P123, P12 and/or P34 processing intermediates were detected in pulsed samples (Fig. 7c). In samples chased for 15 min the amounts of P1234 were reduced; after 60 min chase all ns polyproteins were poorly detectable. In contrast, the amounts of mature ns proteins (with the exception of unstable nsP4) increased during the chase. No significant differences from this pattern were observed for CHIKV<sup>S482A</sup> (Fig. 7c). This data confirms that S482A substitution has no negative effect on ns polyprotein processing in CHIKV infected cells.

Finally, it was analysed whether the S482A substitution affects the replication complex (spherule), replication organelle (cytopathic vacuoles, CPV) or progeny virion formation. Comparison of wt CHIKV and CHIKV<sup>S482A</sup> revealed no detectable differences in the formation of any of these virus-specific structures (Fig. 8) confirming that Ser482 residue is dispensable for the CHIKV infectious cycle.

## Discussion

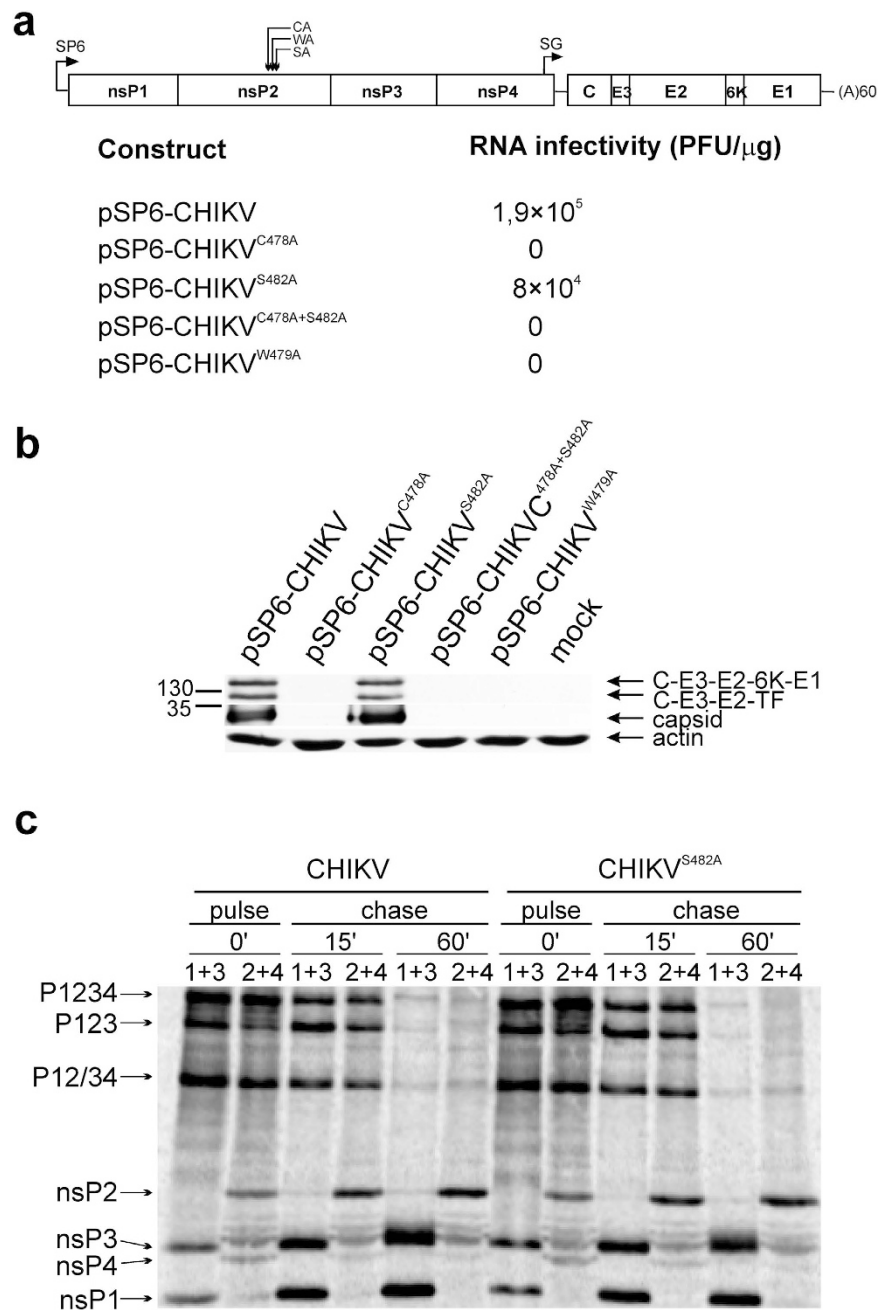
Thus far the data regarding the properties of CHIKV nsP2 protease has been controversial. As this enzyme is central for CHIKV infection and represents a favoured target for anti-CHIKV inhibitor development, revealing the processing requirements of the enzyme and the significance of residues located in its active site region was essential.



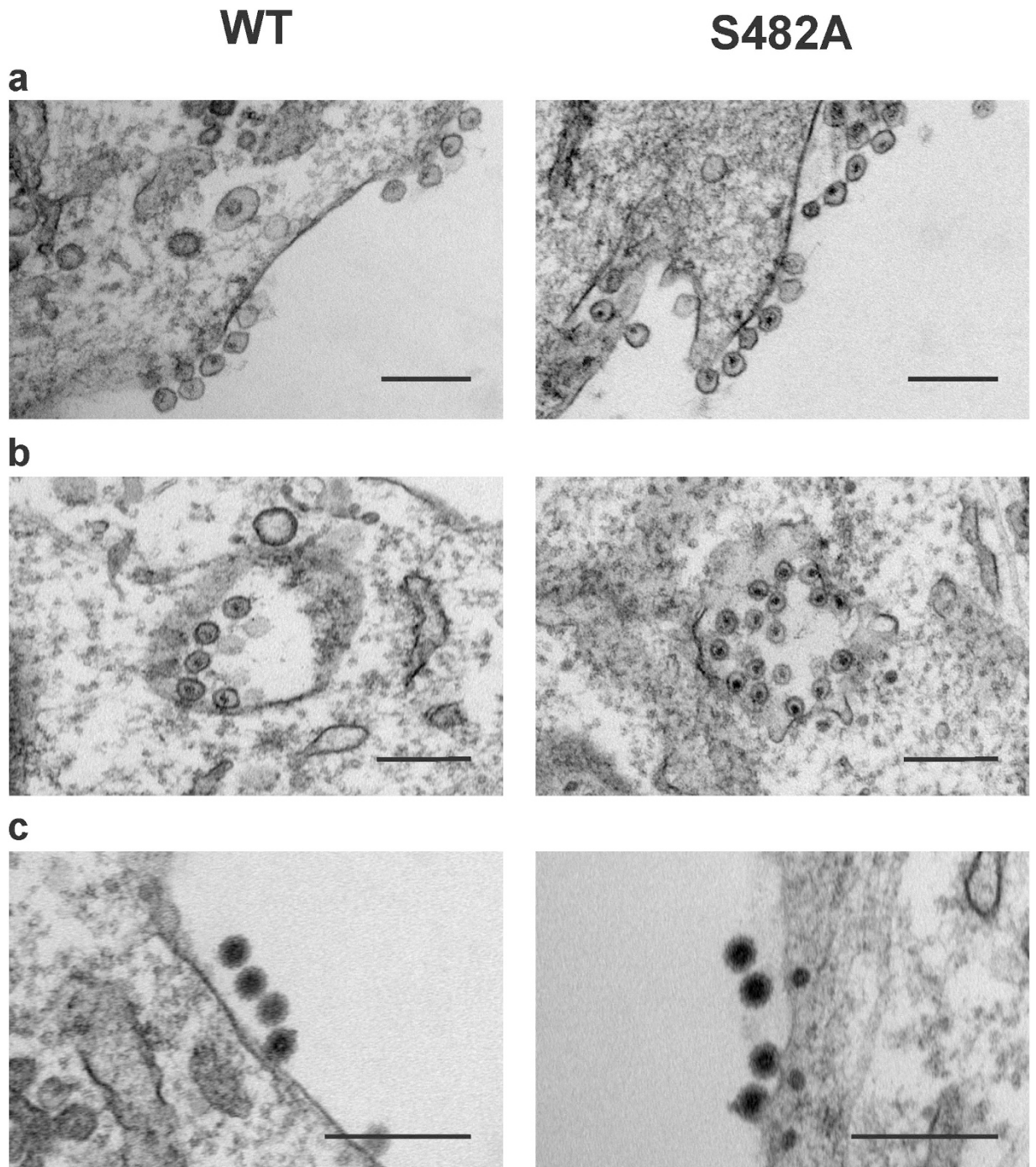
**Figure 6. RNA synthesis of CHIKV and SFV *trans*-replicases.** BSR cells were co-transfected with plasmid encoding for wt or mutant CHIKV replicase and T7-Rluc-Tom encoding template for CHIKV replicase. Alternatively, cells were transfected with plasmids expressing SFV replicase and Tmed encoding for SFV specific template. Total RNA was isolated at 18 h p.t. Eight micrograms total RNA was fractionated and transferred to Hybond-N+ nylon filter. Negative- and positive-strand RNAs were detected by hybridization with probes targeting the Rluc gene present in the template constructs. (a) Schematic presentation of T7-Rluc-Tom and Tmed plasmids encoding for replication competent template RNA of CHIKV and SFV. Plasmids have similar design except that number of codons from ORF1 of corresponding virus (designated as N) is 77 for T7-Rluc-Tom and 74 for Tmed. Other designations are the same as used for Fig. 5a. (b) Synthesis of negative-strand RNAs by SFV and CHIKV *trans*-replicases. (c) Synthesis of positive-strand full-length RNAs by SFV and CHIKV *trans*-replicases. Parts of the blots corresponding to template RNA are shown on panels b and c; for full-length images see Supplementary Information. The experiment was repeated three times with similar results.

The easiest way to study protease activities is to use a purified enzyme and corresponding substrate(s). However, as during the infection process the enzymatic activities of viral enzymes are affected by multiple factors, using a single type of assay provides poor grounds to draw conclusions about the properties of nsP2, including quantitative efficacy. Furthermore, cell-free assay does not provide information about the biological relevance of the observed effects.

Here it was observed that the substrate requirements of CHIKV nsP2 protease were similar to those reported for its SFV counterpart. At the same time the protease activity of CHIKV nsP2 was considerably lower than that previously reported for the protease domain of SFV nsP2<sup>10</sup>. Most likely this reflects differences in the basic enzymatic activities of the proteases from different alphaviruses. CHIKV nsP2 harbours a Glu residue in position 515 that is predicted to be part of the S4 subsite of the enzyme<sup>22</sup>. In contrast, the SFV protease has Val515 that was shown to increase the ability of nsP2 to process P34 polyprotein in infected cells<sup>43</sup>. Thus, it is possible that in order



**Figure 7. S482A substitution in nsP2 region does not interfere with the rescue of CHIKV or with processing of ns polyproteins in infected cells.** (a) Schematic presentation of CHIKV genomes (SP6–SP6 RNA polymerase promoter present in the cDNA clone) and results of infectious centre assay. (b) BHK-21 cells were transfected with  $10 \mu\text{g}$  of *in vitro* synthesized capped transcripts and lysed either at 18 h p.t. (pSP6-CHIKV and pSP6-CHIKV<sup>S482A</sup> transcripts) or 144 h p.t. (pSP6-CHIKV<sup>C478A</sup>, pSP6-CHIKV<sup>W479A</sup> or pSP6-CHIKV<sup>C478A+S482A</sup> transcripts). Proteins were separated using 10% SDS-PAGE, transferred to membrane and capsid protein was detected using a polyclonal antiserum. Parts of the same gels corresponding to capsid protein and C-E3-E2-6K-E1/C-E3-E2-TF polyproteins are shown; due to lower efficiency of detection, polyproteins cannot be revealed using the same exposure as used for the detection of mature capsid protein; for full-length images see Supplementary Information. Arrows on the right indicate positions of detected proteins; positions of molecular mass markers (in kDa) are shown on the left. Actin was used as loading control. (c) BHK-21 cells were infected with wt CHIKV or CHIKV<sup>S482A</sup> at MOI 20. 3 h post-infection cells were starved in methionine-cysteine-free DMEM (Life Technologies) and labelled with  $50 \mu\text{Ci}$  of [ $^{35}\text{S}$ ]methionine-cysteine mixture (Perkin-Elmer) for 15 min (pulse, P). In the chased (C) samples the pulse was followed by a chase for 15 or 60 min. Ns proteins were immunoprecipitated using combinations of antibodies: against nsP1 and nsP3 or against nsP2 and nsP4. Immunoprecipitated proteins were separated using 8% SDS-PAGE and visualized using a Typhoon imager. Positions of ns proteins and ns polyproteins are indicated by arrows.



**Figure 8.** CHIKV<sup>S482A</sup> forms spherules, CPVs and virions similar to wt CHIKV. BHK-21 cells were infected with wt CHIKV and CHIKV<sup>S482A</sup> at MOI 10. At 8 h post-infection samples were fixed and processed for transmission electron microscopy. Scale bar corresponds to 200 nm. (a) Spherules (replication complexes) formed on the plasma membrane of infected cells. (b) Internalized CPVs containing multiple spherules. (c) Progeny virions budding out from the plasma membrane.

to compensate for the presence of preferred P4 Arg residue in the 1/2 site of P1234 polyprotein, CHIKV nsP2 must have lower basal protease activity; the same has previously been observed for the avirulent A7(74) strain of SFV<sup>43</sup>.

The current work also reinforces the crucial role of Cys478 residue in the catalytic site of nsP2 protease. We found no evidence that the enzyme harbouring C478A mutation is capable of cleaving any versions of substrates corresponding to 1/2 and 3/4 sites. Furthermore, this mutation blocked all the essential biological activities of the corresponding replicases and virus. Expectedly, no detectable cleavage of short 2/3 substrate by any recombinant enzyme was observed. However, with the exception of His-nsP2<sup>W479A</sup> that likely forms enzymatically inactive

aggregates, all the used enzymes were capable of inducing the cleavage of 2/3 long substrate. Again, Cys478 but not Ser482 residue was crucial for full efficiency of that cleavage reaction; however, when Cys478 was substituted with Ala, low level processing was still observed. The mechanism responsible for this process remains enigmatic though it certainly was not due to the proposed compensatory effect of Ser482 residue (Fig. 3a). Biophysical or perhaps even structural experiments are required to clarify mechanism(s) responsible for this cleavage. Interestingly, it has been previously observed that processing of 2/3 long substrate is also insensitive to the effect of Pro718 to Gly mutation in nsP2 that severely reduces processing of other substrates<sup>16</sup>. This indicates that due to its unique macromolecular assembly dependent mode of cleavage, the processing of 2/3 site is greatly enhanced by, but not totally dependent on, the Cys478 residue. It should, however, be emphasized that the unusual processing of 2/3 site was not detected upon more natural conditions, such as processing of P1234 in the test tube or in cells and that Cys478 to Ala mutation was irreparably lethal. Thus, it is possible that this residual cleavage activity may represent an artefact of cell-free assays.

The sequence of the active site motif of CHIKV nsP2 (476-NVCWAKS-482) is very similar to that of VEEV nsP2 (475-NVCWAKA-481). Interestingly, the significance of residues 475 and 480 of nsP2 of VEEV (corresponding to Asn476 and Lys481 in CHIKV nsP2) for protease activity has been recently demonstrated. The substitution of Asn residue with Ala led to 25-fold reduction in  $k_{cat}/K_m$  of the enzyme. Mutation of Lys residue to Ala led to a 9-fold decrease in  $k_{cat}$  and  $k_{cat}/K_m$ <sup>23</sup>. Thus, together with the data presented in this study, the importance of most residues in the active site motif of alphavirus nsP2 protease has been demonstrated. Two conserved residues (Cys478 and Trp479) are critically important for protease activity; residues N476 and K481 modulate the activity of the enzyme while substitution of Ser482 to Ala, a residue naturally occurring in the analogous position of VEEV and many other alphaviruses, has little to no effect. For different alphaviruses corresponding position can be occupied also by Cys or Thr residues indicating that considerable variation is tolerated in this position.

Alphavirus nsP2 proteases have not been systematically compared to each other using identical experimental conditions. However, similarities in 3D structures and data from the analysis of functionally important residues strongly indicate that there are no significant differences, at least regarding the functional significance of active site motif residues, between the enzymes encoded by different alphaviruses. CHIKV nsP2 is not an exception; it is a typical alphavirus cysteine protease. Contrary to the recent report, the Ser482 residue from the active site motif of CHIKV nsP2 cannot to any extent compensate for the loss of the catalytic Cys478 residue. The conserved nature of CHIKV nsP2 with its counterparts from different alphaviruses indicates that similar approaches can be used to study these enzymes and their roles in virus infection as well as develop and select for inhibitors of different alphavirus proteases. This opens the possibility that one compound could inhibit nsP2 proteases from different alphaviruses.

## Methods

**Cells and media.** All cells were grown at 37 °C in 5% CO<sub>2</sub> atmosphere; 10% fetal bovine serum, 100 U/mL penicillin and 0.1 mg/mL streptomycin. BHK-21 cells (ATCC CCL-10) were grown in Glasgow's Minimal Essential Medium (GMEM, Gibco) containing 2% tryptose phosphate broth and 200 mM HEPES. BSR cells were cultured in the same medium supplemented with 1 mg/mL of G418.

**Construction of plasmids for expression of mutant forms of CHIKV nsP2.** Plasmids pET-nsP2, designed for expression of CHIKV nsP2 (ECSA genotype, isolate LR-2006-OPY1) with native N-terminus, and pET-His-nsP2, encoding CHIKV nsP2 preceded by amino acid sequence MNHHHHHH-SGGGS-ENLYFQ<sup>29</sup>, were used for expression of wt nsP2 and wt His-nsP2 proteins, respectively. Expression vectors designated pET-nsP2<sup>S482A</sup>, pET-His-nsP2<sup>C478A</sup>, pET-His-nsP2<sup>C478A+S482A</sup>, pET-His-nsP2<sup>W479A</sup> and pET-His-nsP2<sup>S482A</sup> containing mutations resulting in the specified amino acid substitutions in nsP2 were generated *via* PCR-based site-directed mutagenesis. Plasmids pET-nsP2-delN2 and pET-nsP2-insN2<sup>29</sup> were used for expression of nsP2 with deletion of 1<sup>st</sup> and 2<sup>nd</sup> amino acid residues at the N-terminus or with addition of GA dipeptide to the N-terminus, respectively. The sequences of all the expression constructs were verified through DNA sequencing.

**Expression and purification of recombinant proteins.** EGFP-10:5-Trx substrates, containing short versions of the 1/2, 2/3 or 3/4 sites, as well as the longer version of the substrate corresponding to 2/3 site (EGFP-10:170) were purified as previously described<sup>16</sup>. The expression and purification of wt His-nsP2, His-nsP2<sup>C478A</sup>, His-nsP2<sup>C478A+S482A</sup> and His-nsP2<sup>S482A</sup> was performed as previously described<sup>29</sup>. Due to very low expression level purification of His-nsP2<sup>W479A</sup> protein was performed from 3 litres of induced *E. coli* culture. After loading the cleared cell lysate onto the Ni-affinity column an additional washing step with buffer containing 5 mM ATP:Mg and *E. coli* denatured proteins (0.1 mg/mL) was performed in order to remove the contaminating proteins<sup>44</sup>. It was also observed that His-nsP2<sup>W479A</sup> did not properly bind to the cation exchange column and remained mostly in the column flow-through, which was therefore used for its further purification using procedures applied for other His-nsP2 proteins. From Superdex 200 10/300GL size-exclusion column (GE Healthcare, UK) His-nsP2<sup>W479A</sup> eluted with much shorter retention time compared to other nsP2 proteins. Proteins wt nsP2, nsP2-delN2 and nsP2-insN2 were purified using previously described self-cleavage based procedure<sup>29</sup>; the same approach was used to obtain purified nsP2<sup>S482A</sup>. Protein concentrations were measured using Nanodrop spectrophotometer (Thermo Scientific, USA); proteins were flash-frozen and stored at -80 °C. The authenticity of all the purified proteins was verified by mass spectrometric analysis.

**CD spectroscopy.** Circular dichroism (CD) spectrum measurements for recombinant wt nsP2, nsP2<sup>S482A</sup>, His-nsP2, His-nsP2<sup>C478A</sup>, His-nsP2<sup>W479A</sup>, His-nsP2<sup>S482A</sup> and His-nsP2<sup>C478A+S482A</sup> were carried out as previously described<sup>16</sup>. Briefly, CD spectra were recorded on a Chirascan plus CD spectrometer (Applied Photophysics)

three consecutive times in the far-UV region from 185 to 260 nm at a protein concentration of 0.8  $\mu\text{M}$  (with the exception of His-nsP2<sup>W479A</sup> which was at 0.5  $\mu\text{M}$ ) in CD buffer (10 mM  $\text{KH}_2\text{PO}_4$ , pH 7.2 and 100 mM NaF). The final spectra were obtained after deducting the mean spectra of buffer from the mean spectra of protein samples.

**nsP2 protease assays.** A protease assay using recombinant proteins as substrates was performed as previously described<sup>16</sup>. Reaction mixtures were incubated at 30 °C for 60 min, followed by the addition of 2xSDS loading buffer. Samples were boiled and resolved using 10% SDS-PAGE. To analyse processing kinetics, wt nsP2 was mixed with recombinant proteins containing cleavage sequences (10:5) from 1/2 and 3/4 sites using molar ratio 1:100. The reaction was carried out in 150  $\mu\text{L}$  volume using previously described<sup>16</sup> conditions; 15  $\mu\text{L}$  aliquots were collected at 0 (before adding enzyme), 1, 3, 6, 15, 30, 60 and 90 min and analysed using 10% SDS-PAGE.

Continuous protease assays were performed using a fluorescence resonance energy transfer (FRET)-based approach. Two types of peptide substrates were used: peptide comprising 15 amino acid residues (DELRLDRAGG↓YIFSS) or 8 amino acid residues (LDRAGG↓YI) from the 3/4 site (arrow indicates the scissile bond). These substrates, designated respectively as Long and Short, were purchased from GenScript (USA). Both substrates contained 4-[[4-(dimethylamino)phenyl]-azo] benzoic acid (DABCYL) and EDANS at the amino and carboxy termini, respectively. The fluorogenic reactions were performed as previously described<sup>36</sup>.

**Determination of  $K_m$  and  $K_{cat}$  for wt nsP2 of CHIKV using FRET assay.** Michaelis-Menten plot (Fig. 1d) was generated by performing nsP2 protease reaction at various concentrations (0.5, 1, 2, 5, 10, 15, 30, 50  $\mu\text{M}$ ) of substrate. The velocity of the reaction was linear over several minutes indicating that the enzyme was working at its maximum velocity. Hyperbolic plot was generated by plotting initial velocity against the increasing substrate concentration.

EDANS standard curve was generated after considering the inner filter effect<sup>45</sup>. Briefly, increasing concentration of substrate (0 to 2000 nM as shown in Fig. 1d inset) were treated with 20, 30 or 50  $\mu\text{g}$  of proteinase K (Thermo Scientific) in a final reaction volume of 200  $\mu\text{L}$  at 37 °C for 1 h. The relative fluorescence units (RFU) of EDANS were measured on a Synergy MX microplate reader (BioTek, USA) at excitation wavelength of 340 nm and emission wavelength of 490 nm. The values were plotted using Microsoft Excel software and RFU to nM conversion were calculated from the equation of the line.

**Introduction of point mutations into CHIKV and SFV *trans*-replicases.** Point mutations, resulting in the following changes of encoded nsP2 protein, were made using T7-P1234 plasmid<sup>39</sup> and PCR-based site directed mutagenesis; designations of the obtained constructs are provided in parentheses: Cys478 to Ala (T7-P12<sup>CA</sup>34), Trp479 to Ala (T7-P12<sup>WA</sup>34), Ser482 to Ala (T7-P12<sup>SA</sup>34) and Cys478 and Ser482 to alanine residues (T7-P12<sup>CA+SA</sup>34). To introduce similar mutations into SFV *trans*-replicase plasmid P123Z4<sup>42</sup>, hereafter designated as SFV-T7-P123Z4, PCR-based site directed mutagenesis was used. The obtained plasmids were designated as SFV-T7-P12<sup>CA</sup>3Z4, SFV-T7-P12<sup>WA</sup>3Z4, SFV-T7-P12<sup>SA</sup>3Z4 and SFV-T7-P12<sup>CA+SA</sup>3Z4. The sequences of the constructs were verified through DNA sequencing.

***In vitro* transcription/translation and immunoprecipitation.** *In vitro* transcription and translation were carried out using the TNT-coupled T7 rabbit reticulocyte lysate system (Promega) according to the manufacturer's protocol. Reaction mixtures (50  $\mu\text{L}$ ) containing 20  $\mu\text{Ci}$  of [<sup>35</sup>S]methionine-cysteine mixture (Perkin-Elmer) and 2  $\mu\text{g}$  of a T7-P1234, T7-P12<sup>CA</sup>34, T7-P12<sup>WA</sup>34, T7-P12<sup>SA</sup>34, T7-P12<sup>CA+SA</sup>34 or T7-P1234<sup>GAA</sup><sup>39</sup> were incubated for 90 min at 30 °C; translation was stopped by adding cycloheximide to a final concentration of 1 mM. RNase A was added at a final concentration of 10 ng/ $\mu\text{L}$ , and the mixtures were incubated for another 5 min. Immunoprecipitation of CHIKV ns proteins and polyproteins was performed as previously described<sup>16</sup> using rabbit polyclonal antisera against nsP1, nsP2, nsP3 and nsP4 (all in-house). The precipitated proteins were denatured by heating in Laemmli buffer, separated by SDS-PAGE, and visualized using a Typhoon imager (GE Healthcare).

**Analysis of CHIKV *trans*-replicase activity.** Analysis of CHIKV *trans*-replicase activity was performed as previously described<sup>39</sup>. Briefly, BSR cells, grown on 35 mm plates to 90% confluence, were co-transfected with mixtures consisting of 1  $\mu\text{g}$  of plasmid encoding wt or mutant CHIKV replicase and 1  $\mu\text{g}$  of T7-Fluc-Gluc, a plasmid for expression of an RNA template<sup>39</sup>, using Lipofectamine LTX reagent according to the manufacturer's (Invitrogen) protocols. At 18 h p.t. cells were lysed with *Renilla* luciferase (Rluc) assay lysis buffer (Promega); activities of Fluc and Gluc were analysed using the Dual-Luciferase Reporter Assay kit and Glomax SIS luminometer (Promega).

The levels of replicase-generated positive- and negative-strand RNAs present in transfected cells were analyzed by Northern blotting as previously described in ref. 46, with some modifications. Briefly, BSR cells co-transfected with 1  $\mu\text{g}$  of plasmid encoding wt or mutant CHIKV replicase (T7-P1234, T7-P12<sup>CA</sup>34, T7-P12<sup>WA</sup>34, T7-P12<sup>SA</sup>34, T7-P12<sup>CA+SA</sup>34 or T7-P1234<sup>GAA</sup>) and 1.25  $\mu\text{g}$  of T7-Rluc-Tom<sup>39</sup> were lysed at 18 h p.t. with TRIreagent (Bioline), followed by RNA isolation according to the manufacturer's instructions. Eight micrograms total RNA was fractionated on a denaturing 1% agarose gel and transferred to Hybond-N+ nylon filter (GE Healthcare). RNA was cross-linked to the membrane with Stratilinker (Stratagene). <sup>33</sup>P-labeled RNA probes were made by *in vitro* transcription with T7 RNA polymerase from PCR-amplified DNA fragments containing sequences of the Rluc gene and the T7 promoter. The probe for positive-strand RNA detection corresponded to nucleotides 676 to 32, and the probe for negative-strand RNA detection corresponded to nucleotides 38 to 623 of the Rluc gene as present on template constructs. Hybridization was performed at 60 °C overnight, followed by several washes and 4 days exposure. Detection was performed using Typhoon Trio (GE Healthcare) instrument. Detection of positive- and negative-strand RNAs generated by SFV replicase was performed similarly except that Tmed<sup>46</sup>

was used as source of template RNA and plasmids SFV-T7-P123<sup>Z4</sup>, SFV-T7-P123<sup>Z4GAA</sup>, SFV-T7-P12<sup>CA3Z45</sup>, SFV-T7-P12<sup>WA3Z4</sup>, SFV-T7-P12<sup>SA3Z4</sup> and SFV-T7-P12<sup>CA+SA3Z4</sup> were used to express different forms of SFV replicase.

**Construction, rescue and analysis of recombinant viruses.** Mutations resulting in changes of Cys478 to Ala, Trp479 to Ala, Ser482 to Ala or Cys478 and Ser482 to alanine residues of nsP2 were introduced into pSP6-CHIKV (an infectious cDNA clone of CHIKV ECSA genotype, LR-2006-OPY1 isolate<sup>47</sup>); obtained clones were designated pSP6-CHIKV<sup>C478A</sup>, pSP6-CHIKV<sup>W479A</sup>, pSP6-CHIKV<sup>S482A</sup> and pSP6-CHIKV<sup>C478A+S482A</sup>. All these manipulations were done by using PCR-based site-directed mutagenesis and subcloning procedures; sequences were verified.

Plasmids containing CHIKV cDNA were linearized and transcribed *in vitro* using the mMESAGE mMACHINE SP6 Transcription kit (Ambion), to obtain capped RNAs. These procedures as well as ICA were performed as previously described<sup>16</sup>. Recombinant viruses (wt CHIKV and CHIKV<sup>S482A</sup>) were harvested at 18 h p.t.; for cells transfected with transcripts of pSP6-CHIKV<sup>C478A</sup>, pSP6-CHIKV<sup>W479A</sup> or pSP6-CHIKV<sup>C478A+S482A</sup> the cell culture supernatants were collected at 144 h p.t. The transfected cells were collected at the same time and lysed by boiling in Laemmli buffer.

**Western blot analysis.** Cells transfected with 1 µg of T7-P1234, T7-P12<sup>CA34</sup>, T7-P12<sup>WA34</sup>, T7-P12<sup>SA34</sup>, T7-P12<sup>CA+SA34</sup> or T7-P1234<sup>GAA</sup> were lysed at 18 h p.t. with SDS loading buffer; lysates from cells transfected with RNA transcripts were obtained as described above. Proteins were separated by SDS-PAGE, transferred to nitrocellulose membranes, and detected using primary antibodies against CHIKV ns proteins (*trans*-replicase transfected cells) or against CHIKV capsid protein (RNA transcript transfected cells). Antibody against β-actin (sc-47778; Santa Cruz Biotechnology) was used as loading control. The membranes were then incubated with appropriate secondary antibodies conjugated to horseradish peroxidase (LabAs Ltd, Estonia) and proteins were visualized using ECL Immunoblot Detection kit (GE Healthcare).

**Metabolic labelling.** Pulse-chase labelling of infected BHK-21 cells was carried out as described previously<sup>15</sup>. Briefly, BHK-21 cells were infected with wt CHIKV or CHIKV<sup>S482A</sup> at a multiplicity of infection (MOI) of 20 PFU/cell. At 3 h post infection the cells were starved in methionine-cysteine-free DMEM (Life Technologies) for 30 min and then labelled with 50 µCi of [<sup>35</sup>S] methionine-cysteine mixture (Perkin-Elmer) for 15 min (pulse). In the chase samples, the pulse was followed by a chase for 15 or 60 min in medium containing an excess of unlabelled methionine and cysteine. The cells were then lysed by boiling in 1% SDS, proteins in the cell lysates were immunoprecipitated using combinations of antibodies (nsP1 + nsP3 and nsP2 + nsP4), separated by SDS-PAGE, and visualized using a Typhoon imager (GE Healthcare).

**Electron microscopy.** BHK-21 cells were seeded onto 35 mm glass-bottom dishes (P35G-1.5-14-C; MatTek). The next day, cells were infected with wt CHIKV or CHIKV<sup>S482A</sup> at MOI 10 PFU/cell. 8 h post infection, the supernatant was discarded and cells were fixed with 2% glutaraldehyde in 0.1 M sodium cacodylate buffer for 30 min at room temperature and washed in the same buffer. Sample preparation for transmission electron microscopy was done as described earlier<sup>48</sup>. Sections were imaged in JEOL JEM-1400 electron microscope operated at 80 kV and images were acquired by a bottom-mounted camera Gatan Orius SC 1000B (Gatan Inc.).

## References

- Weaver, S. C. & Forrester, N. L. Chikungunya: Evolutionary history and recent epidemic spread. *Antiviral Res.* **120**, 32–39 (2015).
- Strauss, J. H. & Strauss, E. G. The alphaviruses: gene expression, replication, and evolution. *Microbiol. Rev.* **58**, 491–562 (1994).
- Rupp, J. C., Sokolowski, K. J., Gebhart, N. N. & Hardy, R. W. Alphavirus RNA synthesis and non-structural protein functions. *J. Gen. Virol.* **96**, 2483–2500 (2015).
- Salonen, A. *et al.* Properly Folded Nonstructural Polyprotein Directs the Semliki Forest Virus Replication Complex to the Endosomal Compartment. *J. Virol.* **77**, 1691–1702 (2003).
- Kallio, K., Hellström, K., Jokitalo, E. & Ahola, T. RNA Replication and Membrane Modification Require the Same Functions of Alphavirus Nonstructural Proteins. *J. Virol.* **90**, 1687–1692 (2016).
- Gorchakov, R. *et al.* A New Role for ns Polyprotein Cleavage in Sindbis Virus Replication. *J. Virol.* **82**, 6218–6231 (2008).
- Lemm, J. A., Rümenapf, T., Strauss, E. G., Strauss, J. H. & Rice, C. M. Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus- and plus-strand RNA synthesis. *EMBO J.* **13**, 2925–2934 (1994).
- Hardy, W. R. & Strauss, J. H. Processing the nonstructural polyproteins of sindbis virus: nonstructural proteinase is in the C-terminal half of nsP2 and functions both in cis and in trans. *J. Virol.* **63**, 4653–4664 (1989).
- Merits, A., Vasiljeva, L., Ahola, T., Kääriäinen, L. & Auvinen, P. Proteolytic processing of Semliki Forest virus-specific non-structural polyprotein by nsP2 protease. *J. Gen. Virol.* **82**, 765–773 (2001).
- Vasiljeva, L., Valmu, L., Kääriäinen, L. & Merits, A. Site-specific Protease Activity of the Carboxyl-terminal Domain of Semliki Forest Virus Replicase Protein nsP2. *J. Biol. Chem.* **276**, 30786–30793 (2001).
- Lulla, V., Karo-Astover, L., Rausalu, K., Merits, A. & Lulla, A. Presentation Overrides Specificity: Probing the Plasticity of Alphaviral Proteolytic Activity through Mutational Analysis. *J. Virol.* **87**, 10207–10220 (2013).
- Lulla, A., Lulla, V. & Merits, A. Macromolecular Assembly-Driven Processing of the 2/3 Cleavage Site in the Alphavirus Replicase Polyprotein. *J. Virol.* **86**, 553–565 (2012).
- de Groot, R. J., Hardy, W. R., Shirako, Y. & Strauss, J. H. Cleavage-site preferences of Sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing *in vivo*. *EMBO J.* **9**, 2631–2638 (1990).
- Vasiljeva, L. *et al.* Regulation of the Sequential Processing of Semliki Forest Virus Replicase Polyprotein. *J. Biol. Chem.* **278**, 41636–41645 (2003).
- Lulla, A., Lulla, V., Tints, K., Ahola, T. & Merits, A. Molecular Determinants of Substrate Specificity for Semliki Forest Virus Nonstructural Protease. *J. Virol.* **80**, 5413–5422 (2006).
- Utt, A. *et al.* Mutations Conferring a Noncytotoxic Phenotype on Chikungunya Virus Replicons Compromise Enzymatic Properties of Nonstructural Protein 2. *J. Virol.* **89**, 3145–3162 (2015).

17. Ding, M. X. & Schlesinger, M. J. Evidence that Sindbis virus NSP2 is an autoprotease which processes the virus nonstructural polyprotein. *Virology* **171**, 280–284 (1989).
18. Strauss, E. G., De Groot, R. J., Levinson, R. & Strauss, J. H. Identification of the active site residues in the nsP2 proteinase of Sindbis virus. *Virology* **191**, 932–940 (1992).
19. Golubtsov, A., Kääriäinen, L. & Caldentey, J. Characterization of the cysteine protease domain of Semliki Forest virus replicase protein nsP2 by *in vitro* mutagenesis. *FEBS Lett.* **580**, 1502–1508 (2006).
20. Russo, A. T., White, M. A. & Watowich, S. J. The Crystal Structure of the Venezuelan Equine Encephalitis Alphavirus nsP2 Protease. *Structure* **14**, 1449–1458 (2006).
21. Shin, G. *et al.* Structural and functional insights into alphavirus polyprotein processing and pathogenesis. *Proc. Natl. Acad. Sci. USA* **109**, 16534–16539 (2012).
22. Russo, A. T., Malmstrom, R. D., White, M. A. & Watowich, S. J. Structural basis for substrate specificity of alphavirus nsP2 proteases. *J. Mol. Graph. Model.* **29**, 46–53 (2010).
23. Hu, X. *et al.* Kinetic, Mutational, and Structural Studies of the Venezuelan Equine Encephalitis Virus Nonstructural Protein 2 Cysteine Protease. *Biochemistry (Mosc.)* **55**, 3007–3019 (2016).
24. Singh, K. D. *et al.* Homology modeling, molecular dynamics, e-pharmacophore mapping and docking study of Chikungunya virus nsP2 protease. *J. Mol. Model.* **18**, 39–51 (2011).
25. Agarwal, T., Asthana, S. & Bissoyi, A. Molecular Modeling and Docking Study to Elucidate Novel Chikungunya Virus nsP2 Protease Inhibitors. *Indian J. Pharm. Sci.* **77**, 453–460 (2015).
26. El-labbad, E. M. *et al.* Discovery of Novel Peptidomimetics as Irreversible CHIKV Nsp2 Protease Inhibitors Using Quantum Mechanical-Based Ligand Descriptors. *Chem. Biol. Drug Des.* **86**, 1518–1527 (2015).
27. Jadav, S. S. *et al.* Thiazolidone derivatives as inhibitors of chikungunya virus. *Eur. J. Med. Chem.* **89**, 172–178 (2015).
28. Bassetto, M. *et al.* Computer-aided identification, design and synthesis of a novel series of compounds with selective antiviral activity against chikungunya virus. *Antiviral Res.* **98**, 12–18 (2013).
29. Das, P. K., Merits, A. & Lulla, A. Functional Cross-talk between Distant Domains of Chikungunya Virus Non-structural Protein 2 Is Decisive for Its RNA-modulating Activity. *J. Biol. Chem.* **289**, 5635–5653 (2014).
30. Karpe, Y. A., Aher, P. P. & Lole, K. S. NTPase and 5'-RNA Triphosphatase Activities of Chikungunya Virus nsP2 Protein. *PLoS ONE* **6**, (2011).
31. Pastorino, B. A. M. *et al.* Expression and biochemical characterization of nsP2 cysteine protease of Chikungunya virus. *Virus Res.* **131**, 293–298 (2008).
32. Saisawang, C. *et al.* Full length and protease domain activity of chikungunya virus nsP2 differ from other alphavirus nsP2 proteases in recognition of small peptide substrates. *Biosci. Rep.* **35**, e00196 (2015).
33. Saisawang, C. *et al.* Chikungunya nsP2 protease is not a papain-like cysteine protease and the catalytic dyad cysteine is interchangeable with a proximal serine. *Sci. Rep.* **5**, 17125 (2015).
34. Tomar, S., Narwal, M., Harms, E., Smith, J. L. & Kuhn, R. J. Heterologous production, purification and characterization of enzymatically active Sindbis virus nonstructural protein nsP1. *Protein Expr. Purif.* **79**, 277–284 (2011).
35. Thaa, B. *et al.* Differential Phosphatidylinositol-3-Kinase-Akt-mTOR Activation by Semliki Forest and Chikungunya Viruses Is Dependent on nsP3 and Connected to Replication Complex Internalization. *J. Virol.* **89**, 11420–11437 (2015).
36. Stapleford, K. A. *et al.* Viral Polymerase-Helicase Complexes Regulate Replication Fidelity To Overcome Intracellular Nucleotide Depletion. *J. Virol.* **89**, 11233–11244 (2015).
37. Waxman, L. & Darke, P. L. The Herpesvirus Proteases as Targets for Antiviral Chemotherapy. *Antivir. Chem. Chemother.* **11**, 1–22 (2000).
38. Grum-Tokars, V., Ratia, K., Begaye, A., Baker, S. C. & Mesecar, A. D. Evaluating the 3C-like protease activity of SARS-Coronavirus: Recommendations for Standardized Assays for Drug Discovery. *Virus Res.* **133**, 63–73 (2008).
39. Utt, A. *et al.* Versatile Trans-Replication Systems for Chikungunya Virus Allow Functional Analysis and Tagging of Every Replicase Protein. *PLoS ONE* **11**, e0151616 (2016).
40. Buchholz, U. J., Finke, S. & Conzelmann, K.-K. Generation of Bovine Respiratory Syncytial Virus (BRSV) from cDNA: BRSV NS2 Is Not Essential for Virus Replication in Tissue Culture, and the Human RSV Leader Region Acts as a Functional BRSV Genome Promoter. *J. Virol.* **73**, 251–259 (1999).
41. Hellström, K., Kallio, K., Meriläinen, H.-M., Jokitalo, E. & Ahola, T. The ability of minus strand and modified plus strands to act as templates in Semliki Forest virus RNA replication. *J. Gen. Virol.* **97**, 1395–1407 (2016).
42. Spuul, P. *et al.* Assembly of Alphavirus Replication Complexes from RNA and Protein Components in a Novel trans-Replication System in Mammalian Cells. *J. Virol.* **85**, 4739–4751 (2011).
43. Saul, S. *et al.* Differences in Processing Determinants of Nonstructural Polyprotein and in the Sequence of Nonstructural Protein 3 Affect Neurovirulence of Semliki Forest Virus. *J. Virol.* **89**, 11030–11045 (2015).
44. Rial, D. V. & Ceccarelli, E. A. Removal of DnaK contamination during fusion protein purifications. *Protein Expr. Purif.* **25**, 503–507 (2002).
45. Liu, Y. *et al.* Use of a Fluorescence Plate Reader for Measuring Kinetic Parameters with Inner Filter Effect Correction. *Anal. Biochem.* **267**, 331–335 (1999).
46. Kallio, K. *et al.* Template RNA Length Determines the Size of Replication Complex Spherules for Semliki Forest Virus. *J. Virol.* **87**, 9125–9134 (2013).
47. Pohjala, L. *et al.* Inhibitors of Alphavirus Entry and Replication Identified with a Stable Chikungunya Replicon Cell Line and Virus-Based Assays. *PLoS ONE* **6**, (2011).
48. Hellström, K., Vihinen, H., Kallio, K., Jokitalo, E. & Ahola, T. Correlative light and electron microscopy enables viral replication studies at the ultrastructural level. *Methods* **90**, 49–56 (2015).

## Acknowledgements

The authors acknowledge Siret Usar for her technical assistance and Karl-Klaus Conzelmann for the BSR-T7/5 cell line. The electron microscopy unit of the Institute of Biotechnology, University of Helsinki is acknowledged for technical expertise. Margus Rätsep is acknowledged for help with CD spectroscopic analysis and Sergio Kasvandik for help with protein mass spectrometry. This work was supported by institutional research funding (IUT20-27) from Estonian Research Council (to AM) and by Academy of Finland grant number 1265997 (to TA).

## Author Contributions

P.K.D., T.A. and A.M. conceived the experiments. K.R., A.U., T.Q., F.S.V., E.Ž., P.K.D. and A.M. conducted the experiment. K.R., A.U., T.Q., F.S.V., E.Ž., P.K.D., T.A. and A.M. analysed the results. P.K.D., A.M. and T.A. wrote the manuscript. All authors reviewed the manuscript.



## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Rausalu, K. *et al.* Chikungunya virus infectivity, RNA replication and non-structural polyprotein processing depend on the nsP2 protease's active site cysteine residue. *Sci. Rep.* **6**, 37124; doi: 10.1038/srep37124 (2016).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016