Polerovirus N-terminal readthrough domain structures reveal novel molecular strategies for mitigating virus transmission by aphids

Carl J. Schiltz[1,6‡], Jennifer R. Wilson[2,7‡], Christopher J. Hosford[1,8], Myfanwy C. Adams[1], Stephanie E. Preising[2], Stacy L. DeBlasio[2,3], Hannah J. MacLeod[3,9], Joyce Van Eck[4,5], Michelle L. Heck[2,3,5*] and Joshua S. Chappie[1,*]

[1] Department of Molecular Medicine, Cornell University, Ithaca, NY, 14853, USA
[2] Section of Plant Pathology and Plant-Microbe Biology, School of Integrative Plant Sciences, Cornell University, Ithaca, NY, 14853, USA
[3] USDA-Agricultural Research Service, Emerging Pest and Pathogen Research Unit, Ithaca, NY, 14853, USA
[4] Section of Plant Breeding and Genetics, School of Integrative Plant Sciences, Cornell University, Ithaca, NY 14853, USA
[5] Boyce Thompson Institute for Plant Research, Ithaca, NY, 14853, USA

Present address:
[6] Department of Biological Sciences, Vanderbilt University, Nashville, TN, 37232, USA
[7] USDA-Agricultural Research Service, Corn, Soybean & Wheat Quality Research Unit, Wooster, OH, 44691, USA
[8] New England Biolabs, Inc., Ipswitch, MA, 01938, USA
[9] Accelevir Diagnostics, Baltimore, MD 21202, USA

* To whom correspondence should be addressed. Email: chappie@cornell.edu, mlc68@cornell.edu

‡ Authors contributed equally

Abstract

Poleroviruses, enamoviruses, and lutoeviruses are icosahedral, positive sense RNA viruses that cause economically important diseases in food and fiber crops. They are transmitted by phloem-feeding aphids in a circulative manner that involves the movement across and within insect tissues. The N-terminal portion of the viral readthrough domain (^{N}RTD) has been implicated as a key determinant of aphid transmission in each of these genera. Here, we report crystal structures of the ^{N}RTDs from the poleroviruses turnip yellow virus (TuYV) and potato leafroll virus (PLRV) at 1.53-Å and 2.22-Å resolution, respectively. These adopt a two-domain arrangement with a unique interdigitated topology and form highly conserved dimers that are stabilized by a C-terminal peptide that is critical for proper folding. We demonstrate that the PLRV ^{N}RTD can act as an inhibitor of virus transmission and identify ^{N}RTD mutant variants that are lethal to aphids. Sequence conservation argues that enamovirus and luteovirus ^{N}RTDs will follow the same structural blueprint, which affords a novel approach to block the spread of these agricultural pathogens in a generalizable manner.

Introduction

Poleroviruses (Family: *Solemoviridae*), enamoviruses (Family: *Solemoviridae*) and luteoviruses (Family: *Tombusviridae*), formerly described as luteovirids but herein referred to as P/E/L viruses, are insect-vector borne, icosahedral viruses capable of infecting most major crop and biofuel plants. Their positive-sense RNA viral genomes are roughly 5.8 kb in size and share a conserved arrangement of the open reading frames that spawn five to nine known gene products[1] (**Fig. S1a**). These in turn orchestrate plant infection and insect transmission through a series of temporally and spatially regulated protein interactions[2]. P/E/L viruses are transmitted almost exclusively by sap-feeding aphid vectors[1,3]. P/E/L virions circulate throughout the aphid body, interacting with proteins in the aphid's gut and accessory salivary glands prior to transmission to a new host plant. The aphid gut represents the first barrier for transmission, providing selectivity for the uptake of

P/E/L viruses. Virus replication is limited to the plant phloem and no replication occurs in the insect vector[4]. This mode of transmission is deemed circulative, non-propagative.

P/E/L viruses encode two structural proteins[5] (**Fig. S1a**). The coat protein (CP), derived from ORF3, constitutes the major component of the viral capsid[6]. Stochastic ribosomal readthrough of the CP stop codon generates a second minor capsid component termed the readthrough protein (RTP), which contains an additional readthrough domain encoded by ORF5 (RTD) that is fused to the CP C-terminus[6-9]. The leakiness of the CP stop codon has been maintained throughout evolution and ensures that the RTP is incorporated into the capsid sub-stoichiometrically[10]: mutant viruses that lack the stop codon and make only the full-length RTP cannot assemble proper virions, infect plants, or be transmitted by aphid vectors[11-13]. A soluble form of the RTP that is not associated with the capsid plays a role in phloem limitation and movement within the plant host[13,14]. The readthrough domain itself can be subdivided into a globular N-terminal portion (^{N}RTD) and an unstructured C-terminal portion (^{C}RTD) that undergoes proteolytic processing as part of the normal viral lifecycle[15,16]. Mutant viruses lacking the RTD are not aphid transmissible but form functional capsids capable of protecting the RNA genome and can infect plants at a reduced titer[11,16-18]. In contrast, engineered RTP truncations that remove only the ^{C}RTD incorporate efficiently into virions[13], retain the ability to interact with aphid proteins[19] and can be transmitted to new hosts[10]. These observations implicate the ^{N}RTD as a key determinant of P/E/L virus transmission and necessary for traversing aphid gut epithelial cells during viral uptake.

Previous structural studies have detailed the underlying organization of P/E/L capsids. Cryo-electron microscopy (cryo-EM) and crystallographic characterization of polerovirus CP constructs lacking the RTD confirmed that P/E/L viruses assemble with T=3 icosahedral symmetry[20,21], which arranges 180 quasiequivalent monomers into closed particles that display two-fold, three-fold, and five-fold symmetry[22,23] (**Fig. S2a,b**). Despite these efforts, nothing is known about the structure of the RTD, how it is presented on the capsid, and why it is limited

within the mature virion. Deciphering these details is essential for understanding how P/E/L viruses interact with and are transmitted by their aphid vectors.

Here we present atomic-resolution crystal structures of polerovirus ^{N}RTDs, which define a two-domain architecture with a unique, interdigitated topology. Our structures rationalize phenotypes observed in previous mutagenesis studies and provide new insights into the organization of the RTD on the capsid surface and the factors limiting its incorporation within mature virions. We also uncover an unexpected evolutionary connection to non-aphid transmissible tombusviruses, which informs how the presence or absence of specific structural features correlate with different requirements for transmission to a new host. Functional experiments establish that the PLRV ^{N}RTD can act as an inhibitor of virus transmission and identify ^{N}RTD mutant variants that function as novel bioinsecticides. We demonstrate several effective methods for delivering the ^{N}RTD to aphids, paving the way for a generalized management strategy to prevent the spread of destructive P/E/L pathogens. Molecular approaches to block virus transmission are of major interest for the development of novel disease control technologies[2] and our findings represent a significant advance toward achieving this in an agricultural setting.

Results & Discussion

Structural organization of polerovirus ^{N}RTDs provides insights into evolution of aphid transmission

To understand the molecular interactions regulating polerovirus acquisition and transmission by aphids, we generated soluble versions of the ^{N}RTD regions from PLRV (residues 230-458 of the complete RTP fusion) and TuYV (residues 224-459) that could be expressed in *E. coli* and purified on the milligram scale for structural and biochemical studies (**Fig. S1b,c**). Size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) shows that these constructs

form stable dimers in solution (**Fig. S1d,e**). Both readily crystallized, and we solved the structure of the PLRV ^{N}RTD at 2.22 Å by single wavelength anomalous diffraction (SAD) phasing[24] using selenomethionine-labeled protein (**Fig. S3a,b and Table S1**). The TuYV ^{N}RTD structure was subsequently solved by molecular replacement[25], yielding a more complete model that was refined to 1.53-Å resolution (**Fig. 1 and Table S1**).

Each TuYV ^{N}RTD monomer folds into a two-domain protein comprised of a total of 16 β-strands. Eleven of these strands form two anti-parallel β-sheets – ordered β12-13-7-16-1-4-5 (sheet 1) and β5-6-14-10 (sheet 2) – that sandwich together into a jellyroll fold (**Fig. 1, orange**). β5 adopts a twisted conformation that runs orthogonal to the plane of the sandwich and connects the two sheets along one edge (**Fig. 1a**). The short, β11 strand connects the sheets on the opposite edge. The additional five strands form an anti-parallel β-sheet (β9-8-15-2-3) that curves into a small barrel with a short a helix (α1) flanking the edge of β3 (**Fig. 1, purple**). A series of well resolved loops (L1-L5) connect these segments, with L4 folding over and acting as a lid. We designate this barrel the 'cap domain', as it sits above the jellyroll base. The DALI alignment algorithm[26] indicates that the cap domain fold is present in a number of unrelated proteins, including the dimerization domains of aminopeptidases, the N-terminal region of the F1-ATPase rotary subunits, the *Aeropyrum pernix* IF5B initiation factor, and mammalian *Norovirus* spike proteins (**Fig. S4**). The topology differs in TuYV, however, as the individual secondary structure elements are distributed throughout the jellyroll rather than being clustered sequentially as a single globular unit (**Fig. 1b and S4**). PLRV ^{N}RTD monomers adopt the same specific domain arrangement and topology (**Fig. S3a,b**), suggesting this unique organization is a characteristic feature of poleroviruses and, due to sequence conservation in this region, likely enamo- and luteoviruses as well (**Data S1**).

Further structural comparison using DALI[26] reveals that the jellyroll folds of the PLRV and TuYV ^{N}RTDs are structurally related to the P domains of tombusviruses, with the nearest

structural homologs being tomato bushy stunt virus (TBSV) and cucumber necrosis virus. Tombusviruses share key biological properties with P/E/L viruses – including a small, positive-sense single-stranded RNA genome, similar host range, and a non-enveloped icosahedral capsid with T=3 symmetry comprised of 180 copies of the CP (**Fig. S2a and c**) – but are distinct in that they not aphid transmissible and lack an analogous readthrough domain. Instead, the viral CP contains two domains (S and P) that are constitutively expressed as a single polypeptide[27,28], with the S domain forming the icosahedral capsid shell and the P domain extending from each monomer via flexible linker at all points of two-fold rotational symmetry within the assembled virion (**Fig. S2c,d**). While previous cryo-EM studies demonstrated structural homology between polerovirus CPs and tombusvirus S domains (**Fig. S2e**), structural superposition here shows that the TBSV P domain aligns with the jelly roll domain of each ^{N}RTD but lacks the corresponding segments that make up the cap domain (**Figs. 1c,d and S3c**). The conserved topologies between both families (**Fig. 1b and d**) and the intricate distribution of cap domain segments throughout the primary sequence (**Data S1**) suggest that poleroviruses may be ancestral to tombusviruses, with tombusviruses likely evolving via the gradual loss of cap domain elements and truncation of loops L1-L5 rather than through the concerted acquisition of these segments in a manner that would be constrained by the proper folding of both domains.

We also observe a largely unstructured peptide (the 'C peptide') extending from β16 in the TuYV ^{N}RTD, which transverses sheet 1 and terminates in a final strand (β17) that packs against β12 in an antiparallel orientation (**Fig. 1a,b, marine**). From the electron density, we can define the sequence of this segment unambiguously as the C-terminal portion of the construct (residues 431-459) (**Fig. S5a**). A disconnected fragment of the C peptide (residues 442-445) is resolved in the PLRV structure (**Fig. S5b**), likely owing to partial proteolytic cleavage and dissociation of the liberated fragment during purification and/or crystallization.

The C peptide stabilizes the ^{N}RTD dimer interface

PLRV and TuYV ^{N}RTDs crystallize as dimers (**Fig. S1f,g**), consistent with their stoichiometry in solution. Individual monomers superimpose with an overall RMSD ranging from 1.12-1.42 Å across all atoms, with the L2, β4-β5, and β13-β14 loops and portions of the C peptide exhibiting the greatest degree of structural variability (**Fig. S3d,e**). Within each dimer, ^{N}RTD monomers are oriented parallel to the dimer symmetry axis with the sheet 1 side of the jelly roll facing inward (**Figs. 2a,b and S6a,b**). Cap domain loops L4 and L5 form the upper portion of the TuYV dimer interface, with main chain atoms and residues E315, H356, E360, N362, and Y410 (H362, E366, N368, S369, and Y415 in PLRV) making hydrogen bonds *in trans* (**Figs. 2c,d and S6c,d**). Interacting side chains from β12, L2, and β9 provide additional contacts at the edges of the dimer (**Figs. 2d and S6d**). The C peptides snake up from the bottom of the TuYV jelly roll, filling the large cavity beneath the cap domains before exiting in opposite directions to wrap around sheet 2 (**Fig. S5c**). The R440 and R443 side chains anchor an extensive network of stabilizing hydrogen bonds and hydrophobic interactions along interior of the structure while β17 serves a similar role on the exterior (**Figs. 2e and S5d**). Together, the C peptides increase the total buried surface area from 908 Å^2 to 3615 Å^2, constituting a major driving force of dimerization. Although we only resolve a partial fragment from one C peptide in the PLRV ^{N}RTD dimer structure (**Figs. S5b and S6a,b**), this piece forms similar stabilizing interactions with both monomers (**Figs. S5d and S6e**). Deletion of the C peptide from either ^{N}RTD expression construct renders the resulting proteins insoluble. ConSurf analysis[29] shows that residues directly contacting the TuYV C peptides are highly conserved across all P/E/L viruses (**Fig. S7 and Data S1**), signifying the general importance of these interactions as they are maintained throughout the evolution and adaptation of all three genera. Interestingly, a C-terminal truncation of the CABYV RTP terminating immediately after the C peptide can be efficiently incorporated into mature virions whereas mutants disrupting the anchoring arginines cannot[13] (**Fig. S5e**). Together these data underscore the critical role the C peptide plays in the proper folding and stability of the ^{N}RTD dimer.

Disruption of ^{N}RTD folding and stability impairs transmission of mutant viruses

Viral mutants have played an integral part in advancing our understanding of P/E/L virus biology, particularly with regard to movement, uptake, and transmission. Our data afford the opportunity to re-examine these perturbations in a structural context and, consequently, reinterpret the associated phenotypes. Systematic deletion of conserved residues throughout the PLRV ^{N}RTD previously yielded some mutants where the RTP was not incorporated into the assembled virion and other mutants that were incorporated but were not aphid transmissible[14] (**Table S2**). Many of the mutated side chains are buried and form stabilizing hydrogen bonding and hydrophobic interactions (**Fig. S8a,b**), suggesting that the triplet deletions interfere with the structural integrity and folding of the ^{N}RTD dimer. To test this hypothesis, we introduced a subset of the triple deletion and triple alanine mutations from previous studies into our PLRV ^{N}RTD construct (unincorporated: ^{241}PML^{243}, ^{409}YNY^{411}; incorporated: ^{233}RFI^{235}, ^{268}EDE^{270}, ^{315}SST^{317}) and examined the solubility of the resulting proteins. Western blot analysis was used to detect the ^{N}RTD in the soluble (supernatant) and insoluble (pellet) fractions following recombinant expression in *E. coli*. The two non-incorporated mutants (PML and YNY) were insoluble in this context, as neither was detected in the supernatant (**Fig. S8d,e**). The incorporated EDE mutants were similarly insoluble while the SST mutants were partially soluble. Only the ΔRFI mutant remained soluble though a triple alanine mutant at this same position was not. These observations argue that the *in vivo* defects associated with mutant viruses arise from disruption of the proper folding and/or stability of the ^{N}RTD, which is critical for aphid transmission.

Point mutations in the RFI, EDE, and YNY motifs produced similar transmission defects in TuYV[30] (**Table S2**). An alanine substitution at R227 (R233 in PLRV RFI triplet) reduced transmission while a double alanine mutant at E262 and D263 (E268 and D269 in PLRV EDE triplet) was only transmissible after microinjection, meaning the mutant virus was unable to traverse the aphid gut. As in PLRV, these side chains participate in a network of hydrogen bonds that stretches between β1, β4 and β6 and buttresses sheets 1 and 2 at the bottom of the jellyroll

domain (**Fig. S8b,c**). Interestingly, a compensatory second site mutation converting proline 235 to a leucine restores both the infectivity and transmission of the R227A and E262A/E263A TuYV mutants[30] (**Table S2**). P235 sits in an unstructured segment between β1 in the jelly roll domain and β2 in the cap domain, facing into a hydrophobic pocket lined with F259, I288, I421, and I424 (**Fig. S8c**). We speculate that a leucine substitution at this position would alter the overall secondary structure and/or strengthen the existing hydrophobic interactions to keep the cap domain in place, ultimately overcoming any instability in the distal portions of the fold.

The K403A/Y404D (Y409 in PLRV YNY triplet) double mutant also showed reduced TuYV transmission, which could be improved by microinjection[30]. These conserved side chains lie on the edge of the cap domain in both structures where K403 helps anchor L4 and Y404 lines the wall of a cavity on the surface (**Figs. S7d and S8c**). Revertant mutations switching Y404D back to a tyrosine or structurally similar phenylalanine rescue the transmission defect, suggesting a pi-cation interaction with Q399 is important for the stabilization of this region of the protein[31].

^{N}RTD architecture does not limit stoichiometry in the context of the mature virion

Stochastic ribosomal readthrough of the CP stop codon sub-stoichiometrically limits the amount of RTD that is translated, and, hence, present for incorporation into mature, infectious virions[32]. Why the stop codon has been evolutionarily maintained in the virus genome despite the critical role for the ^{N}RTD in aphid transmission is unknown. Leveraging the observed homology with tombusvirus S and P domains (**Figs. 1, S2, and S3c**), we modelled the organization of the ^{N}RTD on the capsid surface to identify possible restraints on virion assembly (**Fig. 3**). Tombusvirus P domains are constitutively translated and tethered to each S domain via an unstructured linker (**Fig. 3a**). When assembled, the P domains occupy every two-fold symmetry axis in the T=3 icosahedral capsid (**Figs. 3b and S2c**). We anticipate that intact RTPs encoded by P/E/L viruses will follow the same architectural design but with the added constraint of head-to-head ^{N}RTD dimerization imposed. A composite model combining the TuYV ^{N}RTD and CP (PDB: 6RTK)

coordinates suggests a similar overall connectivity (**Fig. 3c**), with the ^{N}RTD dimer situated about the two-fold symmetry axis but rotated approximately 15° relative to the position of the TBSV P domains (**Fig. 3d-f**). This organization confirms previous predictions that special interactions stabilize the association of icosahedral asymmetric units (**Fig. 3f**, **black triangles**) across the two-fold axis of symmetry and may contribute to the overall pathway of virion assembly[21]. Importantly, we note no steric clashing if this RTP model is placed at each position in the T=3 icosahedral asymmetric unit (**Fig. 3g**). This implies that the ^{N}RTD could feasibly occupy every two-fold position in a polerovirus capsid and that the architecture of the ^{N}RTD itself does not intrinsically limit its stoichiometry. The proximity of this arrangement, however, might be problematic in that it could promote aggregation and/or collision between the disordered C-terminal region of the RTD in neighbouring subunits, ultimately destabilizing the structure or masking segments of the RTD that may interact with aphid receptors. We speculate that the leaky CP stop codon is therefore preserved to ensure a low concentration of this bulky C-terminal extension on the virion surface. ^{N}RTD dimerization might also impose kinetic constraints that further limit RTP incorporation into the capsid if the timescale of folding is slower than the rate of CP assembly.

^{N}RTD can function as an inhibitor of viral transmission

A soluble version of the tomato spotted wilt virus (Genus: *Orthosospovirus*; Family:*Tospoviridae*) membrane surface glycoprotein G_N was shown in feeding experiments to inhibit viral transmission by its insect vector, the western flower thrips, *Frankliniella occidentalis*[33,34]. Given the important role of the ^{N}RTD in aphid transmission, we asked whether the purified PLRV ^{N}RTD dimer could similarly compete with the mature virus for binding to aphid tissues and subsequently hinder its uptake and transmission by its primary vector, the green peach aphid, *Myzus persicae*. To test this, we first exposed *M. persicae* to the purified ^{N}RTD in an artificial diet feeding system prior to PLRV acquisition and then monitored the subsequent transmission to healthy plants (**Fig. S9**).

Transmission of PLRV to potato was significantly decreased under these conditions as compared to the no protein control (**Fig. 4a, Table S3, Table S4,** $P = 0.046$, likelihood ratio test), despite the fact that oral delivery of the same concentration of the purified bovine serum albumin (BSA) control significantly increased transmission ($P = 0.035$ compared to the no protein control; $P < 0.0001$ compared to ^{N}RTD, likelihood ratio test), a well-described phenotype in the literature observed for proteins unrelated to aphids or virus transmission, including BSA, casein, lysozyme and cytochrome C[35].

Next, we tested whether transient expression of the ^{N}RTD *in planta* using *Agrobacterium tumefaciens* would also block virus transmission (**Fig. S10**). Expression tests and western blot analysis showed that the PLRV ^{N}RTD requires a small protein tag to facilitate folding *in planta* (**Fig. S10a-d**), an unsurprising finding as native ^{N}RTD would be fused to the CP on its N-terminus and the ^{C}RTD on its C-terminus. *M. persicae* were allowed to feed on *Nicotiana benthamiana* leaves transiently expressing YFP-tagged PLRV ^{N}RTD before testing their ability to transmit virus (**Fig. S10e**). We tested both N-terminal and C-terminal YFP tags and found that aphids pre-exposed to the YFP-^{N}RTD by this delivery method also had a decreased ability to transmit virus (**Fig. S4b, Tables S5 and S6,** $P = 0.011$ compared to uninfiltrated control, likelihood ratio test).

Considering the promising results of transient *in planta* expression, we generated transgenic potato plants constitutively expressing YFP-^{N}RTD under the control of the cauliflower mosaic virus (CaMV) 35S constitutive promoter. Expression of YFP-^{N}RTD in these plants was confirmed via western blot analysis, RT-PCR, and fluorescence confocal microscopy (**Fig S11a-c**). YFP signal was observed along the cell periphery and in the nucleus of the YFP-^{N}RTD transgenics but not in the empty vector controls (**Fig. S11c**). *M. persicae* were exposed to transgenic leaves for 48 hours as in previous experiments and then PLRV titer in aphids was quantified by droplet digital PCR (**Fig. S11d**) and their ability to transmit PLRV was assessed (**Fig. S11e**). Aphids exposed to YFP-^{N}RTD acquired significant fewer copies of PLRV (**Fig. 4c,** $P = 0.038$, unpaired one-sided Student's *t* test) and had a reduced ability to transmit PLRV (**Fig. 4d,**

Tables S7 and S8, P = 0.044, likelihood ratio test). These results are consistent with the outcomes from the other delivery strategies and show that the reduction in PLRV transmission is likely due to the reduced ability of aphids to acquire virions across the midgut barrier.

A meta-analysis of our data found that pre-treatment of aphids with the PLRV ^{N}RTD significantly reduced the chances of a plant from becoming infected by nearly half (risk ratio of 0.55) with the 95% confidence interval ranging from a 25% reduction in infection (risk ratio of 0.75) to a 60% reduction (risk ratio of 0.40, **Fig. S12**) regardless of delivery route. There was remarkably low heterogeneity between experiments and even among delivery methods (I^2 = 0%, τ^2 = 0, P = 0.42, Cochran's Q test), indicating that this effect is highly reproducible even when using different delivery strategies. These results argue that the isolated ^{N}RTD can function as an inhibitor of viral transmission.

To test whether the ^{N}RTD interferes with virus transmission when delivered to the accessory salivary glands, aphids were microinjected with the purified PLRV ^{N}RTD prior to performing the transmission assay (**Fig. S13**). Microinjection bypasses the gut and delivers ^{N}RTD directly to the aphid body cavity, which facilities possible interactions with the accessory salivary glands[30]. No change in transmission was observed when two different concentrations of purified ^{N}RTD were microinjected into the hemocoel (**Fig. 4e, Tables S9 and S10**, P = 0.694 compared to no protein control; P = 0.820 compared to BSA control, likelihood ratio test), supporting the hypothesis that the ability of the ^{N}RTD alone to interfere with virus transmission occurs at the gut and not the accessory salivary glands.

Cap domain mutants are lethal to aphids

Our modelling suggests that the cap domain is poised to make direct contact with receptors in the aphid gut. We reasoned that mutating surface-exposed side chains within this domain would disrupt critical interactions needed for viral uptake and thus could impair the ability of the ^{N}RTD to function as an inhibitor in our transmission assays. To test this hypothesis, we introduced a series

of alanine substitutions into the PLRV ^{N}RTD, including point mutations at non-conserved residues H321, E366, H371, E374, and a "cluster" mutant containing three mutations at N368, C370, and Y411, which form a highly conserved pocket (**Fig. 5a**). These positions were chosen as they do not interfere with ^{N}RTD dimerization and folding. ^{N}RTD mutants were purified and delivered to aphids via artificial diet feeding prior to PLRV acquisition and then viral transmission to healthy plants was measured by ELISA (**Fig. S9b**). As predicted, the H321A mutation interfered with the inhibitor function of the ^{N}RTD and did not significantly decrease PLRV transmission compared to the no protein control (P = 0.817; likelihood ratio test) (**Fig. 4a**). This behavior was distinct from the increase in transmission observed with the BSA control (**Fig. 4a,** compared to BSA, P = 0.140, likelihood ratio test), providing further support that the observed phenotypes are specific for the ^{N}RTD and not a general consequence of ingesting protein.

While attempting to assay the other cap domain mutants, we noted that aphids died at a significant rate. To quantify this mortality, aphids were fed the ^{N}RTD mutants via an artificial diet and then moved to either PLRV-infected or uninfected HNS leaves for 24 hours after which the numbers of live and dead insects were counted (**Fig. S9c, Table S11**). E366A, H371A, E374A, and the cluster mutant all caused significant mortality (**Fig. 5b, Table S12,** $P < 0.001$ for all four mutants, respectively, as compared to no protein control), with E374A and the cluster mutant having the strongest effects. Aphids died even when transferred from the ^{N}RTD mutant laden diet treatments to uninfected leaves (**Fig. 5b**), showing that the observed mortality was linked exclusively to the purified ^{N}RTD variants and independent of infectious virus.

The H321A mutant did not cause significant aphid death when compared to no protein and BSA controls ($P > 0.44$) (**Fig. 5b**). The fact that this mutation localizes to a different region of the cap domain (**Fig. 5a**) and has different effects with respect to viral transmission and aphid mortality hints that the ^{N}RTD interacts with the aphid gut in a structure-specific manner. Differential binding of the WT ^{N}RTD and the mutants to various aphid proteins is one hypothesis that may explain both the transmission assay and aphid mortality data. Regardless of the

mechanism, the ability of these mutants to kill aphids means they can be deployed as novel biopesticides, either through transgenic plant delivery as described above for the wild-type ^{N}RTD or via some other delivery strategy.

Conclusions

Our work here defines the basic structural organization of polerovirus ^{N}RTDs and provides a model for how the RTP is incorporated into the mature virion. Polerovirus ^{N}RTDs adopt an interdigitated two-domain architecture and form dimers that are stabilized by a C-terminal peptide. The requirement for dimerization suggests that the ^{N}RTD is situated on the two-fold symmetry axis of the icosahedral viral capsid. Our modeling, however, suggests that this arrangement does not intrinsically limit the stoichiometry of the RTD within the mature virus, which instead may be a consequence of potential ^{C}RTD aggregation and/or kinetic constraints imposed by ^{N}RTD folding and dimerization. Our data also rationalize the effects of various RTD mutants that have been reported over the last several decades[16,18,36-38]. We now can attribute the observed transmission defects associated with different deletions and truncations to the production of insoluble forms of the ^{N}RTD.

Further structural comparisons revealed an unexpected evolutionary connection to tombusviruses, intimating that poleroviruses more closely resemble a common ancestor and that loss of the cap domain decouples a virus from its obligate vector and coincides with the ability to be transmitted in other ways. In 2021, the International Committee on Taxonomy of Viruses abolished the previous designation of P/E/L viruses as a single family, *Luteoviridae*, recategorizing luteoviruses as *Tombusviridae* and poleroviruses and enamoviruses as *Solemoviridae* based solely on differences in their respective RNA-dependent RNA polymerases[39]. Our data argue that P/E/L viruses are in fact structurally and mechanistically distinct from other members of these families and should therefore be treated as a separate group when considering the assembly and organization of the capsid and the mode of transmission.

There are presently no treatments to cure plants of polerovirus infections and current methods to breed viral disease-resistant crops or to control aphid vector populations have proven ineffectual to manage these viruses in the field. Our functional experiments demonstrate that the purified PLRV ^{N}RTD can act as an inhibitor to reduce its vector transmission, not only in the context of artificial diet feeding in the laboratory but also when but delivered transgenically *in planta* in the greenhouse. This provides a novel strategy for controlling these insect-transmitted viruses that can be deployed for widespread field application pending regulatory approval. The high degree of sequence conservation across the jelly roll and cap domains, particularly in regions contacting the C peptide, implies that enamovirus and luteovirus ^{N}RTDs will follow the same structural blueprint and that this inhibitor-based approach could be extended to mitigate these agricultural pathogens as well. Moreover, we show that certain point mutations in the cap domain are lethal to the insect vector, indicating we have discovered a novel biopesticide with many possible permutations and the use of transgenic plants as an already proven delivery strategy. The deployment of genetically-encoded insecticidal and viral transmission-blocking proteins to crop plants lies at the forefront of agricultural technology[40,41] and may one day eliminate the need for environmentally harmful and costly pesticide applications. Our work here represents a major step forward toward achieving this end.

Taken together, our findings advance our fundamental understanding of plant virology and vector biology and impart new tools that can be used to thwart both vector-borne phytoviruses and an economically damaging group of insects.

Data Availability

The atomic coordinates of the TuYV and PLRV ^{N}RTD structures are deposited in the Protein Data Bank with accession numbers 7ULN and 7ULO respectively. All other data, results, and reagents are available from the corresponding authors upon reasonable request.

Acknowledgements

We thank the Northeastern Collaborative Access Team (NE-CAT) beamline staff at the Advanced Photon Source (APS) for assistance with remote X-ray data collection, Dr. Lynn Johnson from the Cornell Statistical Consulting Unit for assistance with statistical methodologies and the meta-analysis of ^{N}RTD inhibitor functions, Maria Gutierrez-Feliciano (BTI) for maintenance of plants and insects, and Dr. Mamta Srivastava (BTI) for technical assistance with imaging. We also thank Kerry Swartwood (BTI) for potato transformation and Dr. Mariko Alexander (Cornell) for helping to initiate the collaboration that led to this work. Research conducted at the NE-CAT beamlines (24-ID-C and 24-ID-E) supported by NIH [P41 GM103403, S10 RR029205] and carried out under the general user proposal GUP-51113 (PI: J.S.C). This research also used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Confocal microscopy images were taken at the BTI Plant Cell Imaging Center, which is supported by the NSF (DBI-0618969). This work was supported by United States Department of Agriculture Grant 2019-05200 (to M.H. and J.S.C.) and USDA ARS CRIS Project 8062-22410-007-000-D (to M.H.). J.S.C. is a Meinig Family Investigator in the Life Sciences. J.R.W. was supported by a NSF Graduate Research Fellowship (DGE-1650441) and a NIFA Predoctoral Fellowship (2019-67011-29610). M.C.A. is supported by a NIFA Predoctoral Fellowship (2020-67034-31750). Funding for this research was also provided by the Schmittau Novak Small Grants Program from the Cornell School of Integrative Plant Sciences (to J.R.W. and C.J.S.).

Author Contributions

C.J.S., J.R.W., M.H. and J.S.C. designed the study and analysed data. C.J.S., C.J.H., and M.C.A. purified and crystallized all ^{N}RTD constructs and collected X-ray diffraction data. C.J.S. solved the ^{N}RTD structures and built the models with assistance from M.C.A. and J.S.C. J.S.C. carried

out computational modelling. C.J.S. generated and purified all N-RTD mutants. M.C.A performed additional N-RTD protein purifications. J.R.W. conducted nearly all aphid experiments, including microinjection, artificial diet, and transient *in planta* delivery transmission assays as well as the mortality assays. J.R.W. generated all expression constructs for transient *in planta* delivery and potato transformation. J.V.E. generated the transgenic potato plants. S.E.P. validated and performed experiments on the transgenic potato plants with microscopy assistance from S.L.D. H.J.M. completed and analysed data for additional transmission assays. J.R.W. and M.L.H. conducted statistical analysis on all aphid experiments. C.J.S., J.R.W., M.H., and J.S.C. wrote the manuscript. M.H., J.S.C., J.R.W., C.J.S. and M.C.A obtained funding to support the research.

Ethics Declarations

M.L.H., J.S.C., J.R.W., C.J.S., and M.C.A have a patent filing related to the technologies described in this paper (U.S. Provisional Patent Application Serial No. 63/289,790).

REFERENCES

1 Wilson, J. R., DeBlasio, S. L., Alexander, M. M. & Heck, M. Looking through the lens of 'omics technologies: Insights into the transmission of insect vector-borne plant viruses. *Curr Issues Mol Biol* **34**, 113-144, doi:10.21775/cimb.034.113 (2019).

2 Heck, M. & Brault, V. Targeted disruption of aphid transmission: a vision for the management of crop diseases caused by *Luteoviridae* members. *Curr Opin Virol* **33**, 24-32, doi:10.1016/j.coviro.2018.07.007 (2018).

3 Gray, S., Cilia, M. & Ghanim, M. Circulative, "nonpropagative" virus transmission: an orchestra of virus-, insect-, and plant-derived instruments. *Adv Virus Res* **89**, 141-199, doi:10.1016/B978-0-12-800172-1.00004-5 (2014).

4 Pinheiro, P. V. *et al.* Plant viruses transmitted in two different modes produce differing effects on small RNA-mediated processes in their aphid vector. *Phytobiomes* (2019).

5 Miller, W. A., Dinesh, K. S. P. & Paul, C. P. Luteovirus gene expression. *Crit. Rev. Pl. Sci.* **14**, 179-211 (1995).

6 Dinesh-Kumar, S. P., Brault, V., Allen & Miller, W. Precise mapping and *in vitro* translation of a trifunctional subgenomic RNA of barley yellow dwarf virus. *Virology* **187**, 711-722, doi:10.1016/0042-6822(92)90474-4 (1992).

7 Bahner, I., Lamb, J., Mayo, M. A. & Hay, R. T. Expression of the genome of potato leafroll virus: readthrough of the coat protein termination codon *in vivo*. *Journal of General Virology* **71**, 2251-2256, doi:10.1099/0022-1317-71-10-2251 (1990).

8 Cheng, S. L., Domier, L. L. & D'Arcy, C. J. Detection of the readthrough protein of barley yellow dwarf virus. *Virology* **202**, 1003-1006, doi:10.1006/viro.1994.1427 (1994).

9 Filichkin, S. A., Lister, R. M., McGrath, P. F. & Young, M. J. *In vivo* expression and mutational analysis of the barley yellow dwarf virus readthrough gene. *Virology* **205**, 290-299, doi:10.1006/viro.1994.1645 (1994).

10 Wang, J. Y., Chay, C., Gildow, F. E. & Gray, S. M. Readthrough protein associated with virions of barley yellow dwarf luteovirus and its potential role in regulating the efficiency of aphid transmission. *Virology* **206**, 954-962, doi:10.1006/viro.1995.1018 (1995).

11 Liang, D., Gray, S. M., Kaplan, I. & Palukaitis, P. Site-directed mutagenesis and generation of chimeric viruses by homologous recombination in yeast to facilitate analysis of plant-virus interactions. *Mol Plant Microbe Interact* **17**, 571-576, doi:10.1094/MPMI.2004.17.6.571 (2004).

12 Lee, L. *et al.* A surface loop of the potato leafroll virus coat protein is involved in virion assembly, systemic movement, and aphid transmission. *Journal of Virology* **79**, 1207-1214, doi:10.1128/JVI.79.2.1207-1214.2005 (2005).

13 Boissinot, S., Erdinger, M., Monsion, B., Ziegler-Graff, V. & Brault, V. Both structural and non-structural forms of the readthrough protein of cucurbit aphid-borne yellows virus are

essential for efficient systemic infection of plants. *PLoS ONE* **9**, e93448, doi:10.1371/journal.pone.0093448 (2014).

14 Peter, K. A., Liang, D., Palukaitis, P. & Gray, S. M. Small deletions in the potato leafroll virus readthrough protein affect particle morphology, aphid transmission, virus movement and accumulation. *J Gen Virol* **89**, 2037-2045, doi:10.1099/vir.0.83625-0 (2008).

15 Alexander, M. M. *et al.* Insights in luteovirid structural biology guided by chemical cross-linking and high resolution mass spectrometry. *Virus Res* **241**, 42-52, doi:10.1016/j.virusres.2017.05.005 (2017).

16 Brault, V. *et al.* Aphid transmission of beet western yellows luteovirus requires the minor capsid read-through protein P74. *Embo J.* **14**, 650-659 (1995).

17 Reutenauer, A. *et al.* Identification of beet western yellows luteovirus genes implicated in viral replication and particle morphogenesis. *Virology* **195**, 692-699 (1993).

18 Chay, C. A., Gunasinge, U. B., Dinesh-Kumar, S. P., Miller, W. A. & Gray, S. M. Aphid transmission and systemic plant infection determinants of barley yellow dwarf luteovirus-PAV are contained in the coat protein readthrough domain and 17-kDa protein, respectively. *Journal of Virology* **219**, 57-65, doi:10.1006/viro.1996.0222 (1996).

19 Mulot, M. *et al.* Transmission of turnip yellows virus by *Myzus persicae* is reduced by feeding aphids on double-stranded RNA targeting the ephrin receptor protein. *Front Microbiol* **9**, 457, doi:10.3389/fmicb.2018.00457 (2018).

20 Byrne, M. J. *et al.* Combining transient expression and cryo-EM to obtain high-resolution structures of luteovirid particles. *Structure* **27**, 1761-1770.e1763, doi:10.1016/j.str.2019.09.010 (2019).

21 Adams, M. C., Schiltz, C. J., Heck, M. L. & Chappie, J. S. Crystal structure of the potato leafroll virus coat protein and implications for viral assembly. *Journal of Structural Biology* **214**, 107811 (2022).

22 Johnson, J. E. & Speir, J. A. Quasi-equivalent viruses: a paradigm for protein assemblies. *Journal of molecular biology* **269**, 665-675 (1997).

23 Prasad, B. & Schmid, M. F. Principles of virus structural organization. *Viral Molecular Machines*, 17-47 (2012).

24 Hendrickson, W. A. Anomalous diffraction in crystallographic phase evaluation. *Q Rev Biophys* **47**, 49-93, doi:10.1017/S0033583514000018 (2014).

25 Bunkóczi, G. *et al.* Phaser. MRage: automated molecular replacement. *Acta Crystallographica Section D: Biological Crystallography* **69**, 2276-2286 (2013).

26 Holm, L. & Rosenstrom, P. DALI server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545-549, doi:10.1093/nar/gkq366 (2010).

27 Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. Tomato bushy stunt virus at 2.9 A resolution. *Nature* **276**, 368-373 (1978).

28 Olson, A. J., Bricogne, G. & Harrison, S. C. Structure of tomato busy stunt virus IV. The virus particle at 2.9 A resolution. *J Mol Biol* **171**, 61-93 (1983).

29 Ashkenazy, H. *et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* **44**, W344-350, doi:10.1093/nar/gkw408 (2016).

30 Brault, V. *et al.* Effects of point mutations in the readthrough domain of the beet western yellows virus minor capsid protein on virus accumulation in planta and on transmission by aphids. *J Virol* **74**, 1140-1148 (2000).

31 Burley, S. K. & Petsko, G. A. Amino-aromatic interactions in proteins. *FEBS Lett* **203**, 139-143, doi:10.1016/0014-5793(86)80730-x (1986).

32 Xu, Y. *et al.* A stem-loop structure in potato leafroll virus open reading frame 5 (ORF5) is essential for readthrough translation of the coat protein ORF stop codon 700 bases upstream. *J Virol* **92**, doi:10.1128/JVI.01544-17 (2018).

33 Whitfield, A. E., Ullman, D. E. & German, T. L. Expression and characterization of a soluble form of tomato spotted wilt virus glycoprotein G_N. *Journal of virology* **78**, 13197-13206 (2004).

34 Whitfield, A. E. *et al.* A soluble form of the tomato spotted wilt virus (TSWV) glycoprotein G_N (G_N-S) inhibits transmission of TSWV by *Frankliniella occidentalis*. *Phytopathology®* **98**, 45-50, doi:10.1094/phyto-98-1-0045 (2008).

35 Bencharki, B. *et al.* Phloem protein partners of Cucurbit aphid borne yellows virus: possible involvement of phloem proteins in virus transmission by aphids. *Mol Plant Microbe Interact* **23**, 799-810, doi:10.1094/MPMI-23-6-0799 (2010).

36 Gildow, F. E. *et al.* Aphid acquisition and cellular transport of potato leafroll virus-like particles lacking P5 readthrough protein. *Phytopathology* **90**, 1153-1161, doi:10.1094/PHYTO.2000.90.10.1153 (2000).

37 Reinbold, C. *et al.* Studies on the role of the minor capsid protein in transport of Beet western yellows virus through *Myzus persicae*. *J. Gen. Virol.* **82**, 1995-2007 (2001).

38 Bruyere, A. *et al.* Effects of mutations in the beet western yellows virus readthrough protein on its expression and packaging and on virus accumulation, symptoms, and aphid transmission. *Virology* **230**, 323-334 (1997).

39 Walker, P. J. *et al.* Changes to virus taxonomy and to the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2021). *Archives of Virology* **166**, 2633-2648, doi:10.1007/s00705-021-05156-1 (2021).

40 Shatters, R. G. *et al.* Compositions and methods for modifying a plant characteristic without modifying the plant genome. United States of America patent (2021).

41 Torti, S. *et al.* Transient reprogramming of crop plants for agronomic performance. *Nature Plants* **7**, 159-171, doi:10.1038/s41477-021-00851-y (2021).

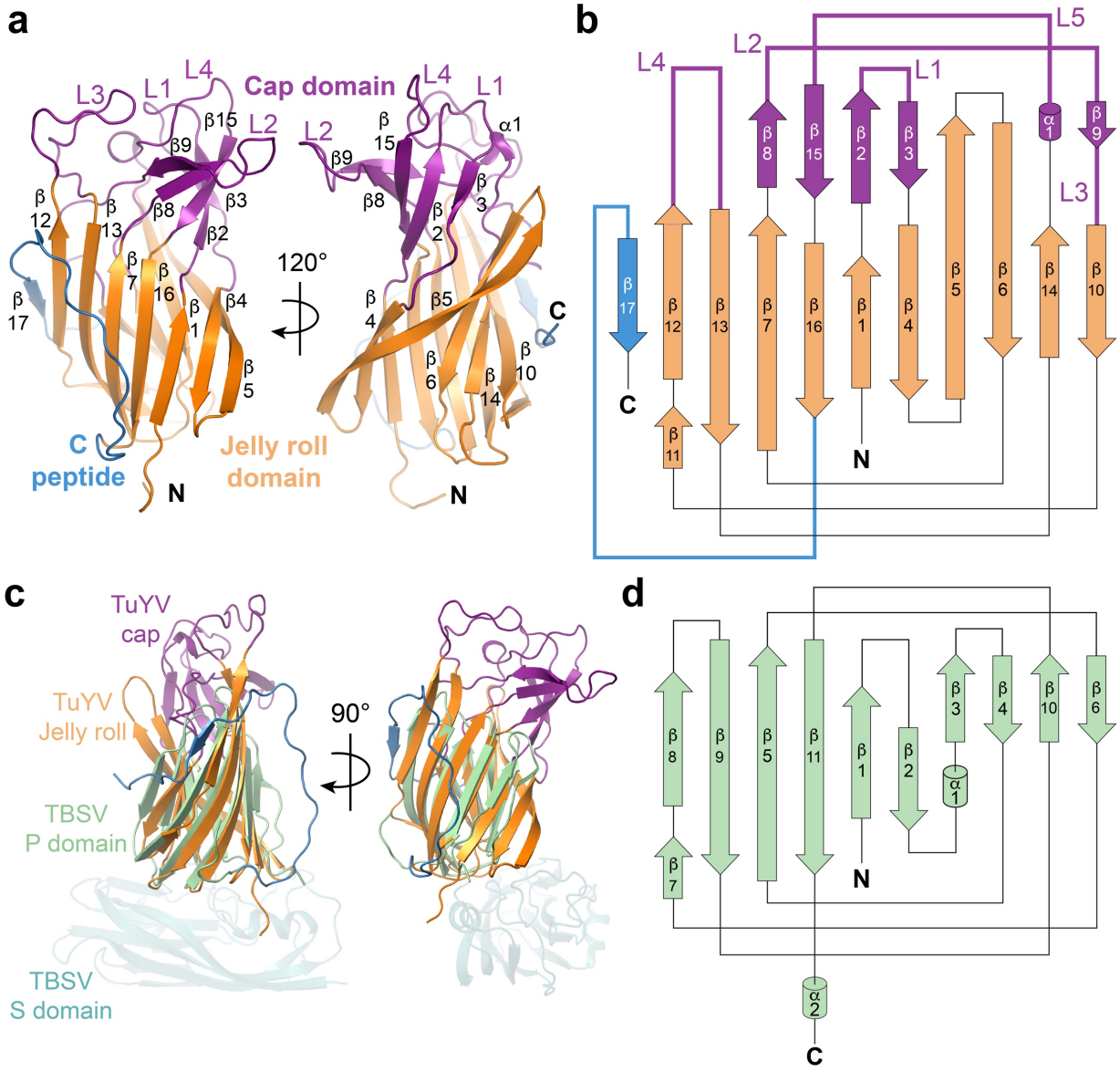

Fig. 1. Structure and topology of the TuYV ^{N}RTD. a-b, Structure (**a**) and topology (**b**) of TuYV ^{N}RTD with jelly roll domain (orange), cap domain (purple) and C peptide (marine) labeled. Cap domain loops are labeled L1-L5. **c,** Superposition of TuYV ^{N}RTD with tomato bushy stunt virus (TBSV) coat protein (PDB: 2TBV; sequence identity: 9% (across the P domain); DALI[26] Z score: 7.0; RMSD: 2.7 Å; P domain, light green; S domain, teal). **d,** Topology of TBSV P domain.

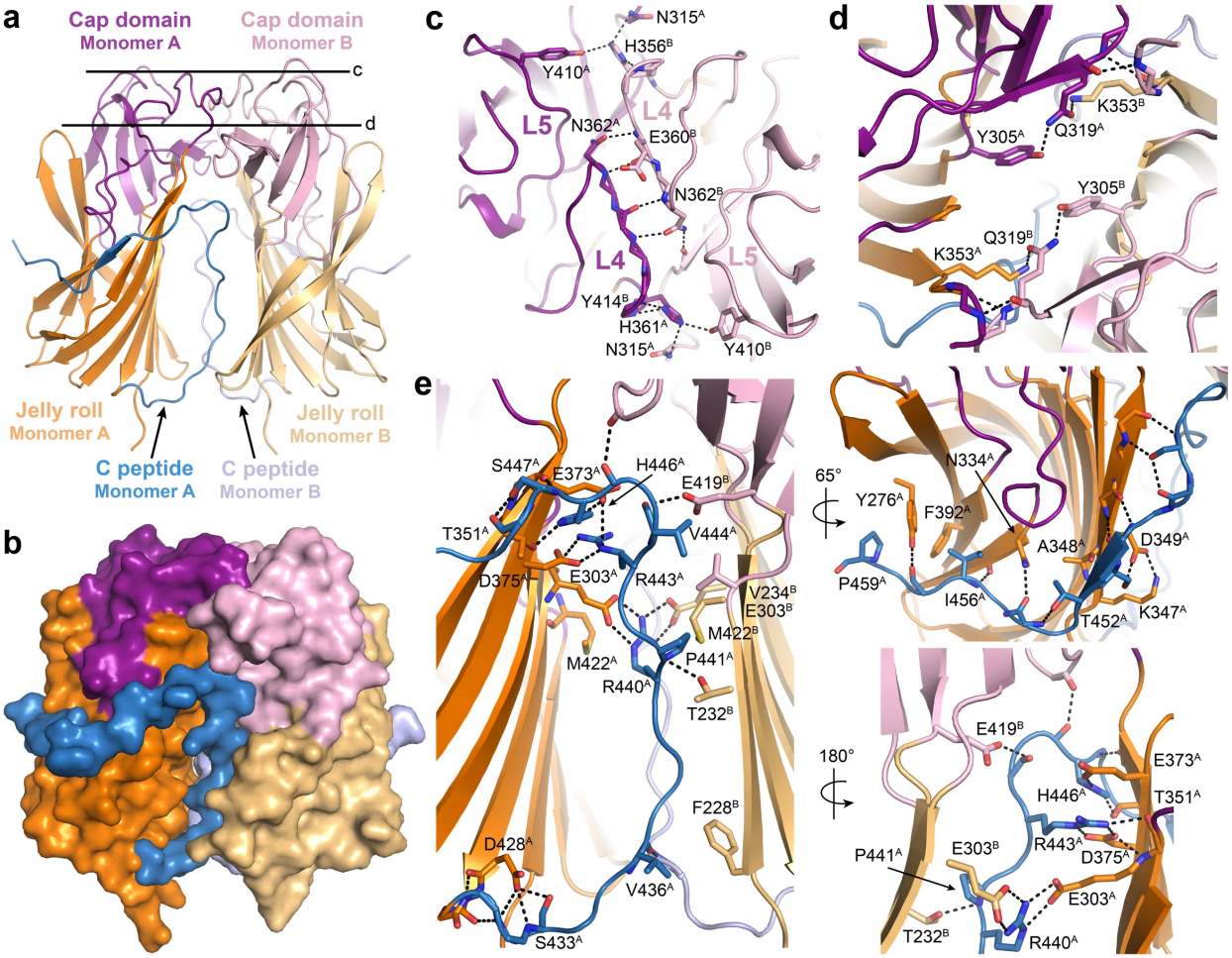

Fig. 2. Architecture of the TuYV ^{N}RTD dimer. a-b, Cartoon (**a**) and surface (**b**) representations of the TuYV ^{N}RTD dimer. Individual structural segments are labeled in each monomer and colored as follows: Jelly roll domains, orange and light orange; cap domains, purple and light pink; C peptides, marine and light blue. **c-d,** Slice sections through the dimer at the levels indicated by the solid lines in (**a**) highlighting stabilizing interactions at the dimer interface. Dashed black lines denote hydrogen bonds. Key residues are labeled with a superscript (A or B) to indicate from which monomer they originate. Secondary structure elements (see **Fig. 1**) are labeled where applicable. **e,** C peptide interactions. Residues contributing hydrogen bonding (dashed black lines) and hydrophobic contacts are labeled.

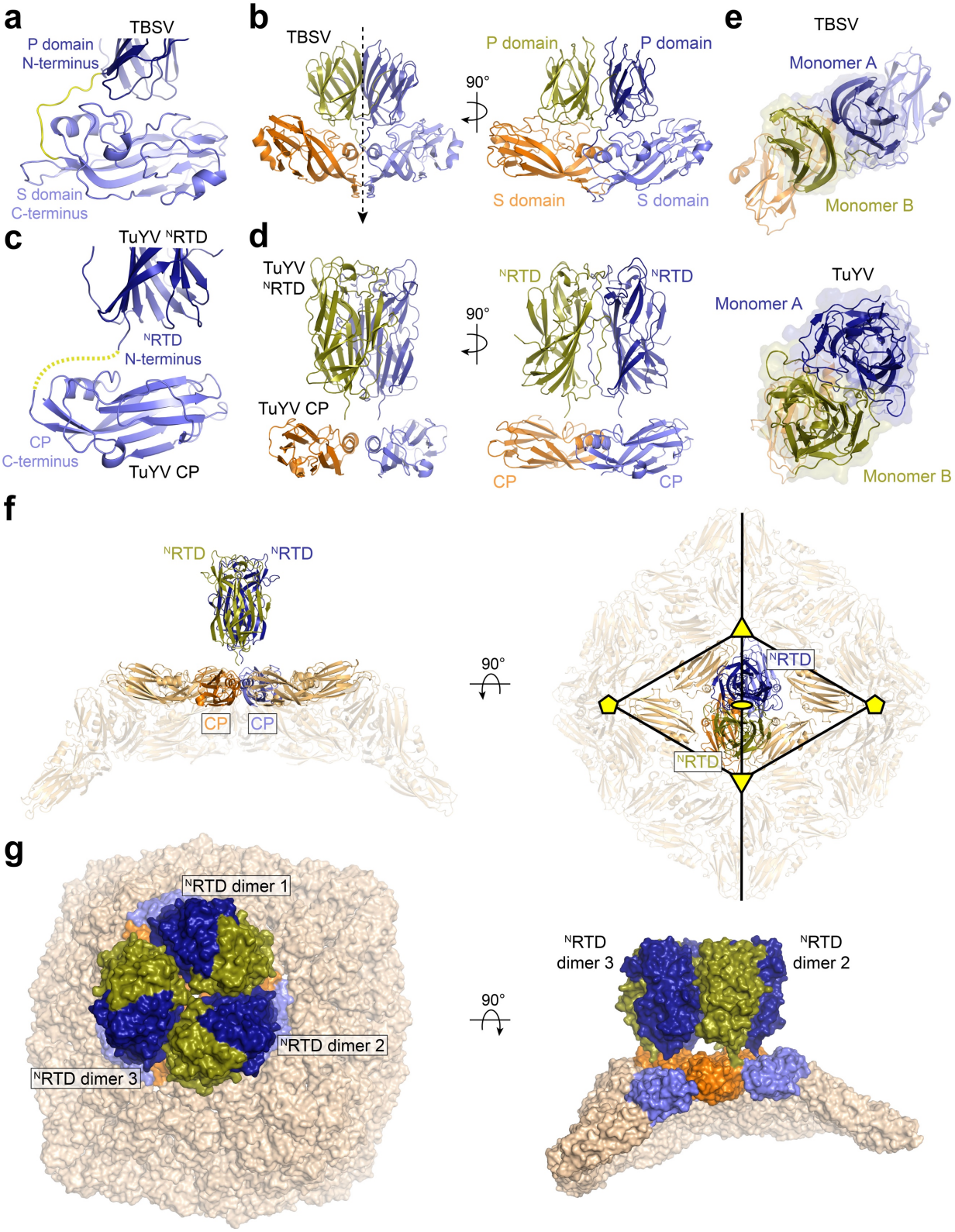

Fig. 3. NRTD architecture does not limit stoichiometry in the context of the mature virion.

a, Domain connectivity in TBSV capsid proteins. Unstructured linker that connects the C-terminus

of the S domain (light blue) to the N-terminus of the P domain (dark blue) highlighted in yellow. **b,** Arrangement of TBSV capsid proteins at the two-fold symmetry axis (dashed arrow) in the assembled virion (see **Fig. S2**) shown in two orientations. S and P domains associated with individual monomers are colored orange and olive (monomer A) and slate and dark blue (monomer B). **c,** Predicted connectivity in capsid proteins based on structural modeling. Dashed yellow line denotes the predicted trajectory linking the C-terminus of the TuYV CP (light blue, PDB: 6RTK) to the N-terminus of the TuYV ^{N}RTD (dark blue). **d,** Composite model of the polerovirus RTP built from the crystallized TuYV ^{N}RTD dimer and CP monomers taken from the cryo-EM reconstruction of modified TuYV virion devoid of the readthrough domain (PDB: 6RTK). RTP dimer is organized around two-fold symmetry axis analogous to the arrangement in (**b**) (see **Fig. S2**). **e,** View of subunit associations in (**b**) and (**d**) looking down the two-fold axis of symmetry in the direction of the dashed arrow in (**b**). **f,** Side (left) and top down (right) views of TuYV RTP modeled at the two-fold symmetry axis of the icosahedral virion. Two-, three-, and five-fold symmetry axes are marked with a yellow ellipse, yellow triangles, and yellow pentagons, respectively. RTP is colored as in (**d**) with the rest of the capsid subunits colored wheat. **g,** Model illustrating the feasible positioning of ^{N}RTD dimers (olive and dark blue) around icosahedral asymmetric unit of PLVR VLP assuming the structural organization in (**d**). Associated CP monomers are colored orange and slate with the reset of the capsid surface colored wheat.

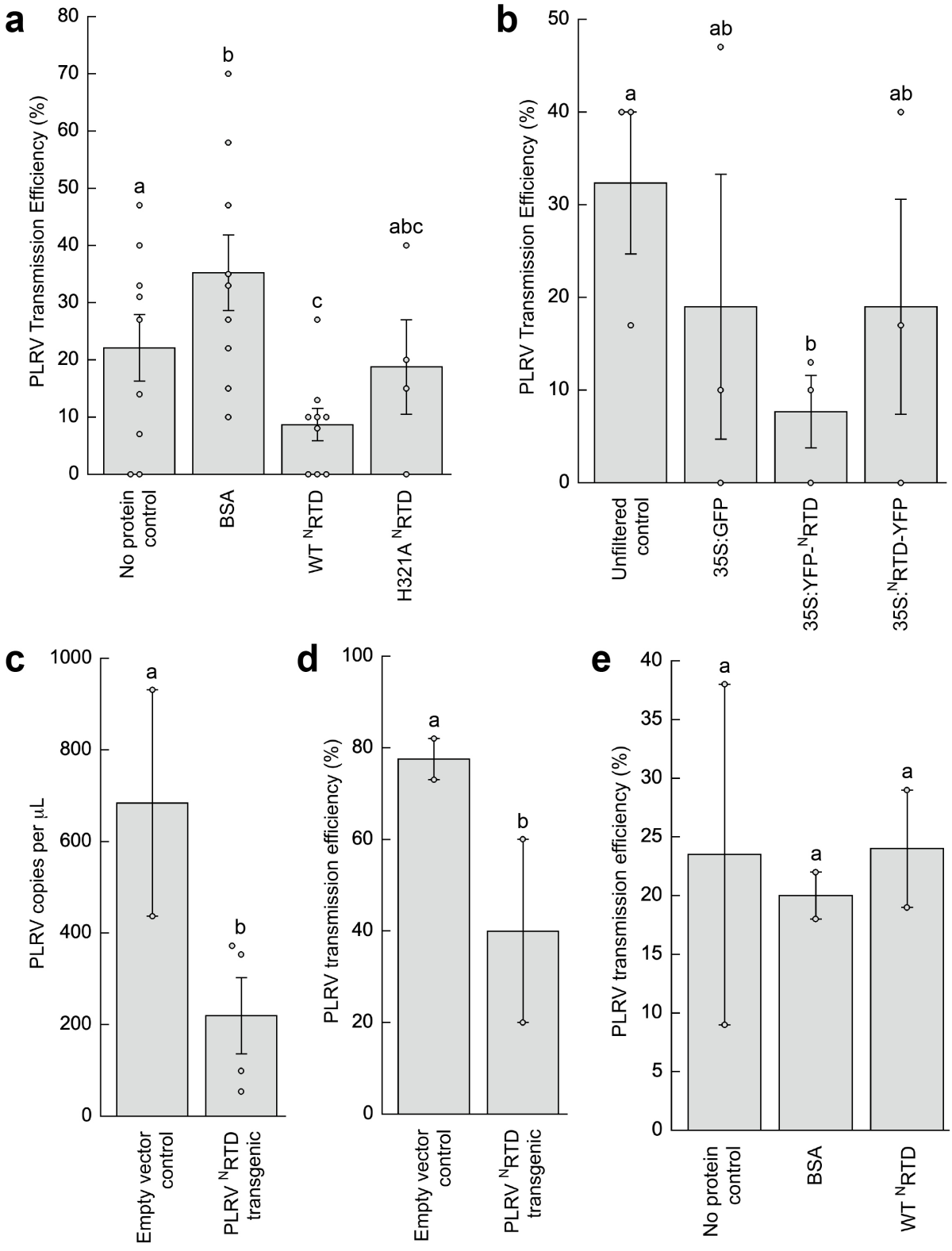

Fig 4. The ^{N}RTD can function as an inhibitor of viral transmission. a, The PLRV transmission efficiency of *Myzus persicae* is significantly different after feeding on various artificial diet

579 treatments (see **Fig. S9a**): no protein control (*n* = 101), BSA (*n* = 116), purified WT PLRV ^{N}RTD
580 *(n* = 124), or PLRV ^{N}RTD point mutant H321A (*n* = 43). **b,** PLRV transmission efficiency of *M.*
581 *persicae* aphids is reduced after transient *in planta* delivery of the PLRV ^{N}RTD (*n* = 37 for all
582 treatments, see **Fig. S10**). **c,** Aphid acquisition of PLRV is reduced after exposure to the PLRV
583 ^{N}RTD transgenic (*n* = 4) compared to empty vector control (*n* = 2) potato plants as quantified via
584 droplet digital PCR (see **Fig. S11**). Mean ± one standard error for all replicates (dots) is shown.
585 Letters above each bar represent significantly different treatments ($P < 0.05$) by Student's *t* test.
586 **d,** PLRV transmission efficiency of *M. persicae* aphids is reduced after exposure to the PLRV
587 ^{N}RTD transgenic (*n* = 60) compared to empty vector control (*n* = 26) potato plants (see **Fig. S11**).
588 **e,** The PLRV transmission efficiency of *M. persicae* is unaltered after microinjection with the no
589 protein (buffer) control (*n* = 19), or two concentrations each (0.1 mg/mL and 1 mg/mL) of BSA (*n*
590 = 40) or purified PLRV ^{N}RTD (*n* = 38). For panels (**a**), (**b**), (**d**), and (**e**), the mean ± one standard
591 error for all independent repeats of the experiment (dots) is shown. Each inoculated plant was
592 considered a replicate. Letters above each bar represent significantly different treatments ($P <$
593 0.05) by logistic regression analysis.

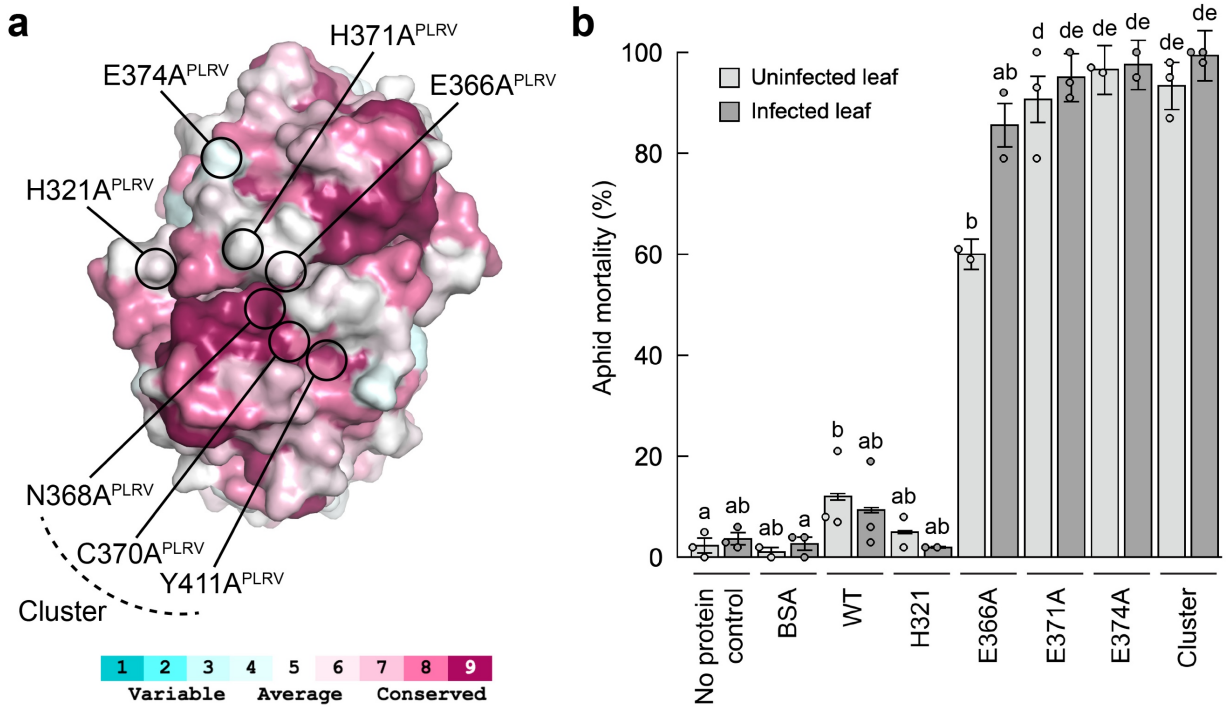

Fig. 5. Cap domain mutants are lethal to aphids. a, Top view of TuYV ^{N}RTD dimer illustrating positions of PLRV cap domain mutants. Coloring reflects sequence conservation among polero-, enamo-, and luteoviruses (legend below, see also **Data S1** and **Fig. S7**) and was generated using the ConSurf Server[29]. **b,** *Myzus persicae* mortality after feeding on 0.1 mg/mL of BSA (*n* = 232), purified WT PLRV ^{N}RTD (*n* = 247), PLRV ^{N}RTD point mutants H321A (*n* = 183), E366A (*n* = 118), H371A (*n* = 264), E374A (*n* = 161), a PLRV ^{N}RTD cluster mutant (containing mutations N368A, C370A, and Y411A; *n* = 256) or no protein controls (*n* = 244) for 48 hours and then moved to an uninfected (light gray) or PLRV-infected (dark gray) detached hairy nightshade leaf (See **Fig. S9b**). Mean ± one standard error for all independent repeats of the experiment (dots) is shown. Each individual aphid was considered a replicate. Different letters represent significantly different treatments ($P < 0.05$) by logistic regression analysis.