Published in final edited form as:
J Proteome Res. 2012 May 4; 11(5): 2968–2981. doi:10.1021/pr300041t.

Cross-linking measurements of the *Potato leafroll virus* reveal protein interaction topologies required for virion stability, aphid transmission, and virus-plant interactions

Juan D. Chavez[1,ξ], **Michelle Cilia**[2,3,ξ], **Chad R. Weisbrod**[1], **Ho-Jong Ju**[3,4], **Jimmy K. Eng**[1], **Stewart M. Gray**[2,3], and **James E. Bruce**[1,*]

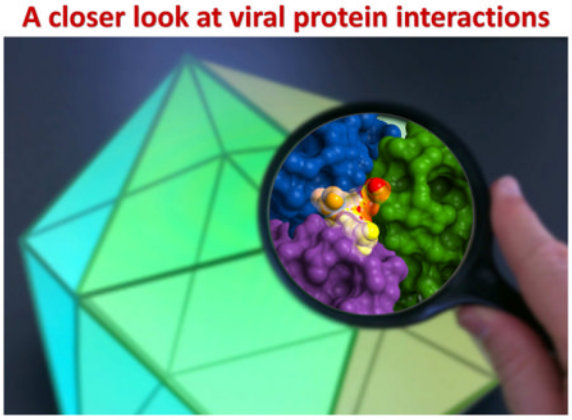
[1]Department of Genome Sciences, University of Washington, Seattle, Washington, 98109

[2]Robert W. Holley Center for Agriculture and Health, United States Department of Agriculture, Agricultural Research Service, Ithaca, New York, 14853

[3]Department of Plant Pathology and Plant-Microbe Biology, Cornell University, Ithaca, New York 14853

[4]Department of Agricultural Biology and Plant Medicinal Research Center, College of Agricultural & Life Sciences, Chonbuk National University, 664-14 Deokjin-Dong 1Ga Deokjin-Gu Jeonju Jeonbuk 561-756, South Korea

Abstract



Protein interactions are critical determinants of insect-transmission for viruses in the family *Luteoviridae*. Two luteovirid structural proteins, the capsid protein (CP) and the readthrough protein (RTP), contain multiple functional domains that regulate virus transmission. There is no structural information available for these economically important viruses. We used Protein Interaction Reporter (PIR) technology, a strategy that uses chemical cross-linking and high resolution mass spectrometry, to discover topological features of the *Potato leafroll virus* (PLRV) CP and RTP that are required for the diverse biological functions of PLRV virions. Four cross-linked sites were repeatedly detected, one linking CP monomers, two within the RTP, and one

Corresponding Author: James E. Bruce, Corresponding Author Address: Department of Genome Sciences, 815 Mercer St., Seattle, WA 98109, jimbruce@uw.edu, Lab phone number: 206.543.0220, Lab fax number: 206.616.0008.
ξThese authors contributed equally to the paper.

Supplementary Dataset
PDB files for the top monomer and trimer PLRV models generated in this study.

Supporting Information Available: This material is available free of charge via the Internet at http://pubs.acs.org.

linking the RTP and CP. Virus mutants with triple amino acid deletions immediately adjacent to or encompassing the cross-linked sites were defective in virion stability, RTP incorporation into the capsid, and aphid transmission. Plants infected with a new, infectious PLRV mutant lacking 26 amino acids encompassing a cross-linked site in the RTP exhibited a delay in the appearance of systemic infection symptoms. PIR technology provided the first structural insights into luteoviruses which are crucially lacking and that are involved in vector-virus and plant-virus interactions. These are the first cross-linking measurements on any infectious, insect-transmitted virus.

Keywords

Protein Interaction Reporter Technology; protein-protein interactions; Luteovirus; Polerovirus; *Potato leafroll virus*; cross-linking; aphids; circulative transmission; readthrough protein; protein disorder

INTRODUCTION

Viruses within the family *Luteoviridae*, collectively referred to as luteovirids, are icosahedral, positive sense, unicomponent, RNA viruses that are retained in the phloem of host plants (Figure 1A)[1]. They are economically important viruses in staple food crops and have been used extensively as a model system to study virus transmission by insects[2–10]. Luteovirids are exclusively transmitted by aphid vectors, phloem-feeding arthropods in the suborder Homoptera. Luteovirids share several common biological traits with circulative viruses transmitted by insects to animals, including humans, e.g., tissue tropisms in hosts and insect vectors, circulative transmission by insect vectors, and transmission by a limited number of vector species[11, 12]. Unlike animal-infecting arboviruses (arthropod-borne viruses), there is currently no evidence that luteovirids replicate in their insect vectors[13].

Protein interactions are critical determinants of virus infection and insect-transmission. Two luteovirid structural proteins, the capsid protein (CP) and the readthrough protein (RTP), function either *in cis* or *in trans* to direct aspects of virus movement within plant hosts and aphid vectors. The CP is required for local and systemic movement in plants and the RTP acts *in trans* to retain virus in the phloem where it is available to aphids[1]. How these two virus proteins regulate these different activities is unknown. However, at least three lines of evidence lead us to hypothesize that the accessible surface areas in the CP and the RTP are most likely to be involved in the plant-virus-vector interactions, specifically, the co-evolutionary arms race that drives pathogen adaption, host resistance, and vector specificity. First, transcapsidation (the packaging of one luteovirid RNA genome into a capsid comprised of structural proteins from a different luteovirid) alters the vector specificity that is the hallmark of luteovirids[14, 15]. Second, virions comprised of the structural proteins are the only vehicle for luteovirid intercellular and long-distance transport in plants and aphids[1, 4, 6–8, 11, 13, 16–19]. Third, mutations in the structural proteins of PLRV[1, 7, 8] and other luteovirids[16, 20, 21] prevent aphid transmission and alter tissue-distribution in host plants. Luteovirids share a conserved arrangement of three open reading frames in the 3′ half of their genome (Figure 1A), two of which encode the structural proteins[22] (Figure 1B). The virus capsid consists predominantly of the CP encoded by ORF 3 and a minor amount of the RTP translated via a readthrough of the CP stop codon[23–26]. The RTP is not required for particle assembly or plant infection[25, 27], but particles containing only the CP are not transmissible by aphids to plants[4, 20, 21, 23, 28]. These RTP-minus virions can pass from the aphid gut into hemocoel indicating that CP alone contains the capsid protein topological features, e.g., structural surface features, required for virus-gut interactions [4, 20, 21]. However, RTP-minus virions do not move as efficiently across the gut as wild-type (WT)

virus [21]. Therefore, although the RTP is not required for the virus to move across the gut membrane, it may facilitate virus uptake and is required for virus to cross into the aphid salivary gland [21].

Luteovirids are dependent upon aphids for plant-to-plant transmission and all share a common circulative pathway within their aphid vectors. Luteovirids evolved specific associations with vector tissues and all share a common circulative pathway within their aphid vectors. However, each virus species is transmitted efficiently by only one or a few aphid species. The circulative transmission pathway through the aphid and the biological factors contributing to "vector specificity" are better understood for luteovirids than any other circulative arthropod-borne virus, including arthropod-borne animal viruses [12, 29–35]. Ultrastructural studies demonstrated that acquisition of luteovirids into the aphid hemocoel occurs through the hindgut or midgut epithelial cells by endocytosis [18, 36–38]. Although there are exceptions[18, 39], the hindgut does not appear to be a major barrier to luteovirid acquisition as aphids can acquire many luteovirid species into the hemocoel that they cannot transmit. Other morphologically similar but unrelated viruses are not acquired into the hemocoel[18]. These observations have led to the hypothesis that common virion topological features mediate the movement of most luteovirids across the gut epithelial cells. Vector species-specific transmission is therefore largely determined by topologies in the virion capsid[7, 8, 12, 13, 16, 17, 40, 41]. Yet, the topologies of most arthropod-borne viruses remain largely uncharted, and luteovirids have remained recalcitrant to structural study using crystallography or other technologies.

One tool that has been gaining popularity for studies of virion protein interaction topology is chemical cross-linking coupled with mass spectrometry[42–45]. Cross-linking is a method used to covalently link one protein to another using a chemical (cross-linker). Reactive groups in the cross-linker (Figure 1C) interact with specific amino acid side chains that lie close together during cross-linker application. The new bonds that are formed provide information about the interfacial regions and topological features of proteins and protein complexes. Application of conventional cross-linking technology in the field of virology has been limited by: 1) dynamic range challenges due to the fact that the large majority of proteins are not cross-linked during these experiments[46]; 2) heterogeneity and complexity of cross-linking reaction product types[47]; and 3) spectral complexity of cross-linked peptides making it difficult to identify sequences, even if they are detectable[42, 43]. To address these challenges in our topological studies of luteovirids, we used a chemical cross-linking strategy referred to as Protein Interaction Reporter (PIR) technology[48]. The PIR strategy relies on novel cross-linker molecules (Figure 1C) with N-hydroxy succinamide ester (NHS) reactive groups and labile bonds that can be cleaved specifically within a mass spectrometer, releasing the intact peptides that were covalently linked during the reaction. Importantly, intact peptide release allows collision-induced dissociation (CID) of individual peptides to produce MS/MS spectra and peptide identification with traditional proteomics search engines like SEQUEST[49] or Mascot[50]. With non-cleavable cross-linkers, peptide and protein identification is often not possible since peptide precursor masses cannot be measured and nor can individual peptides be fragmented and measured. The cleavable feature of PIR technology makes proteome-wide application feasible where conventional cross-linking is not possible. Importantly, PIR technology allows identification of sites within interacting protein sequences *in vivo*[51, 52]. This yields unparalleled ability to gain structural data on proteins and protein complexes in their native conformations in cells[51, 52]. Finally, PIR technology also includes an affinity tag (Figure 1C) to enable purification of cross-linked peptides from the majority of unlabeled species that result from complex sample analysis such as *in vivo* cross-linking applications. This feature helps meet the dynamic range challenges mentioned above. Here, we used PIR and mass spectrometry to investigate the topology of *Potato leafroll virus* (PLRV), the type species in the genus

Polerovirus within the family *Luteoviridae*. We report the first dataset of cross-linked peptides from an infectious virus and provide the first topological data on protein interaction topologies in the PLRV RTP that are required for aphid transmission and virus-host interactions.

EXPERIMENTAL SECTION

Generation of PLRV Mutant

A full-length cDNA of PLRV containing a deletion of 26 amino acids in the C-terminus of the RTP (RTC 12) was generated to examine the effect of removing an important topological feature of the RTP that was identified using PIR. The following primers were manually designed using the sequence of the PLRV cDNA clone[7] as a template: P1, nucleotides (nt) 3874-3900 TGTCCGGCATTCAAGGATGGAATACTC, P2, nt 5025-5008/4929-4909 CGTTTGTATCGGGGTTTC/CCGCCGCACTCTGTTAGCCGA, P3, nt 4912-4929/5008-5028 GCTAACAGAGTGCGGCGG/ GAAACCCCGATACAAACGCAG, and P4 5883-5860, CTCCACTACACAACCCTGTAAGAGGATC, and the constructs were generated according to the yeast recombination method[53]. To remove nt positions 2474-5817 and create a linear fragment with blunt ends from the plasmid pBPY[53], a yeast-bacteria shuttle vector containing full-length cDNA of PLRV, the plasmid was digested with 10-fold excess of the enzymes Stu1 (New England Biolabs, site: AGC/CCT) and Bsa1 (New England Biolabs, site GGTCTC$(N)_1$/), in New England Biolabs Buffer 4 (20 mM tris-acetate, 50 mM potassium acetate, 10 mM magnesium acetate, 1mM dithiothreitol) containing 100 μg/mL Bovine Serum Albumin (BSA) at 37 °C, overnight. After linearization, the long fragment was dephosphorylated with calf intestinal alkaline phosphatase (New England Biolabs) to inhibit fragment-end ligation. The short insertion fragments containing the mutation region were PCR-amplified from pBNUP110[53] with primers homologous to their junctions (Supplementary Figure 1). The linearized pBYP and the PCR-generated small fragments were mixed with yeast grown on YPD plates (1% yeast extract, 2% peptone, 2% dextrose, 2% agar) for yeast transformation based on the Duplex-A-yeast two hybrid system manual (Origene Thechnologies, Rockville, MD). Total DNAs including recombinant plasmid DNAs were isolated from pooled yeast colonies grown on the SD/-trp media agar plate (SD media, trpyptophan-drop out amino acid, Clontech, Mountain View, CA) at 3 days after plating [53] and were used to transform *Escherichia coli* strain DH5α. Plasmid DNAs recovered from DH5α were verified via polymerase chain reaction using high fidelity Taq polymerase (Ex Taq, TaKaRa) and Sanger sequencing in the mutation region. *Agrobacterium tumefaciens* strain LBA4404 was then transformed using plasmids extracted from DH5 and used for agroinoculation as previously described[54, 55].

Virus Preparation

The plasmid pBPY containing the full-length wild-type (WT) cDNA of the PLRV-Canada strain (NCBI Accession #: D13954 D00734) in the pBIN19 backbone for plant transformation with *A. tumefaciens* has been described previously [6]. *A. tumefaciens* (strain LBA4404) carrying pBPY or RTC 12 was grown in YEB (0.5% Difco beef extract broth, 0.5% peptone, 0.1% yeast extract, 0.5% sucrose, 2 mM MgSO4, pH 7.2) broth with 50 mg/mL kanamycin for 48 hr. The cultures were centrifuged at 5000 ×g to pellet the *A. tumefaciens* and the pellets were resuspended to 0.4 OD_{600nm} in water. For production of PLRV or PLRV RTC12 mutant, fully-expanded leaves in *Nicotiana benthamiana* plants at the nine to ten leaf stage were infiltrated with pBPY in *A. tumefaciens*. Tissue was harvested for virus isolation five to six days post-inoculation and frozen at −80 °C for one week..

Using a pre-chilled blender (Waring), infected *N. benthamiana* leaves were homogenized for 3 min in a 0.1 M sodium citrate buffer (pH adjusted to 6.5 using 0.5 M sodium phosphate containing 0.5% 2-mercaptoethanol, 1 g tissue to 5 ml buffer). Virus was partially purified using a modified version[33] of the protocol of Hammond et al.[56] and all centrifugation steps were completed at 4 °C. Virions were collected after centrifugation in a 30% sucrose cushion and resuspended in 1.0 mL phosphate buffered saline (PBS). Virus concentration was determined by reading the A_{260}, A_{280}, A_{320} using a NanoDrop spectrophotometer (Thermo Scientific) and by the following calculation: $(A_{260} - A_{320})/8.0$. Purified virus was stored at −80 °C for up to six months.

Aphid Transmission Assays

Non-viruliferous *Myzus persicae* were fed on a finely-stretched parafilm membrane pouch containing a 1:3 dilution of virus preparation (approximately 100 ng virus) added to 200 μl of sterile 20% sucrose solution for 24 hr. After the 24-hr period, typical feeding behavior was observed for most aphids on the purified virus preparation and aphids were viable. The aphids (5 per plant) were transferred to caged, healthy hairy night shade plants (*Solanum sarrachoides*) (6 plants total) for a 72 hr inoculation-access period and then killed by fumigation. The accumulation of PLRV was assessed using DAS-ELISA 3-weeks post inoculation as described [6].

PIR Cross-linking

The biotinylated-Rink (BRink) PIR cross-linker was synthesized in-house using solid phase FMOC peptide synthesis chemistry as described previously [51]. BRink (1 mM final concentration) was added to 1.0 mg purified PLRV and reacted with the virus at 25°C for 25 min. Reactions were performed in PBS at pH 7.0; therefore, maximizing reactivity for exposed lysine residues within the PLRV capsid.

LC-MS/MS

Disulfide bonds in the cross-linked virion proteins were reduced with 5 mM *tris*(2-carboxyethyl)phosphine (TCEP) for 30 min, followed by alkylation with 10 mM iodoacetamide (IAA) for 30 min, in the dark. Reduction and alkylation were performed at 25°C. PLRV virions were then digested with trypsin (1:200 ratio, Promega) at 37°C for 16 hr with gentle agitation. The peptide mixture was desalted using a C18 sep-pak (Waters). Strong cation exchange (SCX) was performed using spin columns packed with polysulfotethyl aspartamide (Nest Group Inc.) to remove excess hydrolyzed cross-linker. After extensive washing with 25 mM potassium phosphate (KH_2PO_4), 25% acetonitrile (ACN) pH 2.4, peptides were eluted from the SCX stationary phase with 25 mM KH_2PO_4, 1 M potassium chloride, 25% ACN, pH 2.4. The ACN was evaporated off using vacuum centrifugation and the pH was adjusted to 7.0 with 1 M NaOH. BRink cross-linked peptides were enriched using avidin affinity chromatography (Ultralink monomeric avidin, Pierce) as described previously [52]. The enriched fraction was concentrated by vacuum centrifugation and stored at −80°C until analysis by liquid chromatography-mass spectrometry (LC-MS/MS).

Peptide samples were analyzed by LC-MS/MS using a Waters NanoAcquity Ultra-high Performance Liquid Chromatography (UPLC) coupled to a Thermo Scientific Velos ion trap-Fourier transform ion cyclotron resonance (FTICR) mass spectrometer. Samples were loaded onto a trap column (3 cm × 100 μm inner diameter (i.d.)) packed with 200 Å pore size, 5 μm Magic C-18AQ or Magic C8 particles (Michrom Bioresources Inc.) using a flow rate of 2 μl/min of 99% solvent A (H_2O, 0.1% formic acid), and 1% solvent B (ACN, 0.1% formic acid) where they were washed for 10 min. Peptides were then separated by reversed-phase chromatography over an analytical column (30 cm × 75 μm i.d.) packed with 200 Å

pore size, 5 μm Magic C-18AQ or Magic C4 particles using a binary phase, linear gradient from 80% solvent A 20% solvent B, to 60% solvent A, 40% solvent B over 120 min at a flow rate of 300 nL/min.

Eluting peptides were ionized by electrospray ionization (ESI) using a voltage of 1.7 kV and were analyzed with a Thermo Scientific Velos-FTICR mass spectrometer. Mass spectrometry analysis of cross-linked peptides was performed using a new analytical method based on an extension from our previously published strategy using the program X-links[47]. This new analytical method for the analysis of cross-linked peptides in real-time will be described in detail in a forthcoming manuscript (Weisbrod et al. in preparation.). Briefly, this method allows for the identification of cross-linked peptide pairs by validating the PIR mass relationship (mass precursor = mass peptide #1 + mass peptide #2 + mass reporter) used in X-links, in real-time, as the spectra are acquired by the mass spectrometer. Accurate mass measurements of PIR cross-linked peptide ions were made in the ICR cell at a resolving power of 50,000 at m/z 400. Precursor ions with a charge state $\geq 4^+$ were selected for MS^2 with CID performed in the ion trap using a normalized collision energy of 30 V and fragment ions were analyzed in the ICR cell at a resolving power of 25,000 at m/z 400. Fragment ions generated at the MS^2 stage, consisting of the released peptides, were isolated, fragmented by CID using a normalized collision energy of 35 V, and the product ions analyzed in the Velos ion trap in an MS^3 scan. No charge state exclusion criteria are applied when selecting released peptides for MS^3. Cross-linked peptides were further verified manually using chromatographic retention time.

Database Searching

Two MS search engines were used to maximize discovery in the MS data since the use of different search algorithms can result in the identification of different peptides[57], even with the same peptide scoring parameters applied. Tandem mass spectra were converted into mzXML and mascot generic format (MGF) peak list files using tools in the Trans-Proteome Pipeline[58]. The MGF files were submitted to Mascot v.2.3 (Matrix Science, Boston, MA) [50] while mzXML files were submitted to Sequest v.2011.02.0 for database interrogation against a 6-frame translation of the refseq PLRV genome (GI:9629160) and against the entire NCBI non-redundant database (download date July 6th, 2011) to eliminate precursor ions derived from common contaminants. Since covalent attachment of the cross-linker to lysine residues interferes with trypsin cleavage, three possible missed tryptic cleavages were allowed during the search. Carbamidomethylation was considered as a fixed modification on cysteine. The remaining tag from the PIR cross-linker (mass= 99.0320 Da) was treated as a variable modification on residues lysine, serine, threonine, and the protein N-terminus. Oxidation of methionine and deamidation of residues asparagine and glutamine were also considered as variable modifications. The mass error tolerances were set at 30 parts per million (ppm) in the MS mode and 0.8 Daltons (Da) in the MS/MS mode. Peptide assignments from Mascot or Sequest with an expect value (E-value) less than 0.05 were accepted. All MS/MS spectra were manually interpreted by the investigators for accuracy, completeness of b and y ion series, and mass shift indicating the addition of the PIR tag on the lysine residue.

Structural Modeling

Structural modeling of the monomer subunit of the PLRV coat protein was performed using Phyre2[59]. The amino acid sequence consisting on residues 1-207 for PLRV CP (UniProt MCAPS_PLRV1) was submitted using the web interface of the Phyre2 server. The model for the PLRV CP was constructed using the 20 top sequence aligning template structures consisting of PDB structures and known homologous folds. The top template structure used was from the coat protein for ryegrass mottle virus (PDB = C2IZW). SymmDock[60] was

used to model the PLRV trimer. The 3D atomic coordinates for the PLRV monomer model from Phyre2 were used as input into Symmdock with a symmetry order of 3. Cross-linking constraints were used to define the interface region including residues N186, G187, K188, and S189. The pdb files for the top monomer and trimer models of PLRV are available in the Supplementary Dataset.

RESULTS and DISCUSSION

PLRV Particles were Infectious

To confirm the WT virus preparation was infectious and aphid-transmissible, non-viruliferous aphids (*M. persicae*) were fed on 100 ng of purified PLRV in a 20% sucrose solution for 24 hr and transferred to healthy hairy night shade plants. Five out of 6 plants became infected when tested by ELISA (absorbance values >2.0 at 405nm). Therefore, the topological data presented here represents *in vivo* virion topology because the virions were purified from plant cells, contain packaged viral genomic RNA, and were transmissible by their aphid vectors.

The Advanced Design of PIR was Essential for Detecting PLRV Cross-linked Peptides

(1) The biotin affinity tag enriched for cross-linked peptide pairs—PIR technology enables measurements of protein interaction topologies using mass spectrometry[61–64] and has been successfully applied on *Shewanella oneidensis* [62], *E.coli* [52] on proteome-wide scales. The advanced molecular design of the PIR molecule (Figure 1C) was essential to the success of detecting PLRV cross-linked peptides. Specifically, the enrichment step using the PIR affinity-tag was critical to detect cross-linked peptides because LC-MS/MS analysis of the trypsin digest before avidin-enrichment yielded no detectable cross-linked products (data not shown). Following avidin-affinity enrichment, four cross-linked sites within PLRV virions were repeatedly identified (Table 1). To our knowledge, this study represents the first successful demonstration of any cross-linking technology to infectious virus particles yielding the first dataset of cross-linked virus peptides originating from intact, infectious virions.

(2) Mass spectrometry-cleavable bonds produced predictable mass relationships and enabled identification of cross-linked PLRV peptides—The two mass spectrometry-labile RINK bonds in the PIR molecule enable release of the cross-linked peptides in the mass spectrometer [48] and the bonds are specifically cleaved at lower energies than are required for cleaving peptide bonds during mass spectrometry analysis. The mass spectrometry-labile bonds therefore release the cross-linked peptides without fragmentation of the peptide backbone. All species of cross-linked product, such as dead-end, intra- and intercross-linked peptides are cleaved at these labile RINK bonds, generating product type-specific mass relationships. Dead-end cross-links occur when one reactive NHS-ester interacts with a peptide and the other with water, intra-cross-linked products link two reactive amino acids within the same peptide, and inter-cross-links are formed when the NHS-esters link two different peptides together (Supplementary Figure 2). Our analysis focused on the identification of intercross-linked peptides within PLRV virions as these typically yield the most useful protein topological information. Inter-cross-linked species are cleaved to yield three components (Figure 2). The parent ion of the inter-cross-linked peptide pair is fragmented to yield a reporter ion and two released peptides (Figure 2, inter-cross-link). The sum of these three fragment masses matches the inter-cross-linked parent mass (parent ion mass = reporter ion mass + released peptide A + released peptide B). In total we identified 92 unique mass relationships that fulfill the PIR mass relationship criteria for inter-peptide pairs within a mass error of less than 20 ppm, with more than 50% having a mass error of less than 5 ppm. These PIR mass relationships are listed in Supplementary

Table 1. Having high mass accuracy measurements is critical for successful identification of cross-linked peptides using the PIR approach.

Application of PIR Technology to Purified PLRV Yields Unique Topology Data in the CP and RTP

An advantage of the PIR approach over traditional cross-linking methods is that peptide sequence identification can be performed in the same manner as traditional proteomics experiments. Mass spectra containing the peptide backbone fragment ion information of the released peptides can be searched against a PLRV protein database using search engine tools such as MASCOT [50] and SEQUEST [49] (Figure 2B). A residual mass tag from the PIR cross-linking reagent, with a mass of 99.032 Da, generates an observable shift in the fragmentation pattern of the peptide backbone and enables one to unambiguously assign the cross-linked sites within those peptides after database searching. Confident identification of both released peptides to the PLRV CP and RTP was achieved for four unique cross-linked peptide pairs (Table 1). One of these pairs was an unambiguous homodimer (Table 1, Figure 3). The homodimer cross-linked peptide pair was identified with a mass-to-charge ratio (*m/z*) = 970.196. The quadruply-charged precursor ion was measured with a PIR relationship mass error of 5.6 ppm relative to the calculated theoretical mass (Figure 3A). The spectrum of the released peptides consisted of a singly charged ion at *m/z* = 1377.638 and the doubly charged reporter ion at *m/z* = 561.747 (Figure 3B). Database searches with MASCOT and SEQUEST resulted in identification of the peptide as GNGKSSDPAGSFR, which resides in the C-terminal region of the CP (amino acids 185–197); the underlined K indicates the cross-linked site (Figure 3C). This cross-linked peptide could only have formed via interactions between two or more CP monomers, as this peptide exists only once within each CP sequence. This unambiguous cross-linked CP homodimer is herein referred to as XL-CP (Figure 1).

The additional three cross-linked pairs which were fully identified involved peptides from the RTP domain of PLRV. Currently there is no structural information about the RTP in the literature. Furthermore, previous attempts by our group failed to detect intermolecular interactions within the RTP or between the CP and RTP using a yeast-2-hybrid approach (data not shown). Thus, these are the first data reporting any protein interaction topology in a luteovirid RTP and the first experimental validation that these regions of the RTP are exposed on the surface of virions. These will be referred to as XL-RTP-1 (Figure 4), XL-RTP-2 (Figure 5), and XL-RTP-CP (Figure 6). The peptides that were cross-linked in XL-RTP-1 included EVDSGSEPGPSPQPTPTPTPQKHER and ISKLR, spanning amino acids 209–233 and 349–353, respectively (Figure 4C). XL-RTP-2 (Figure 5, Table 1) was defined by two tryptic peptides FFLVGPAIQKTAK and EPEGKPVGNKPR, amino acids 396–408 and 461–472, respectively (Figure 5, Table 1). These data do not allow one to distinguish whether these cross-linked peptides could be from within a single RTP monomer or via interactions between different RTP monomers. Nevertheless, the identification of cross-links between these two sites of the RTP shows that in virions, these regions are important points of contact that define the topology of the incorporated RTP. A third peptide pair, XL-CP-RTP, revealed a cross-link between the RTP peptide, FFLVGPAIQKTAK, amino acids 396–408, with the CP peptide GNGKSSDPAGSFR, amino acids 185–197, (Table 1, Figure 6). Epitope mapping of PLRV placed amino acids 185–193 on the exposed face of the virion where it would be accessible for interaction with the RTP[65]. Thus, the cross-linking data in the RTP suggest two hypotheses for the RTP. One hypothesis is that there is a single conformation of the RTP that places the lysine residues of all three peptides (Figure 1B, K188, K405 and K470) in spatial proximity. Alternatively, interaction of the RTP with itself and with a region of the CP may also result from multiple distinct conformations for the RTP. Variable conformations in the RTP may in part provide the structural basis which

allows the RTP to interact with multiple binding partners: for example interactions between the RTP and the multitude of aphid and plant proteins encountered by virions during their journey from host to host. The latter hypothesis is also supported by the fact that no cross-link was observed between K188 and K470. Remarkably, the peptide EPEGKPVGNKPR, with mass 1405.734, was found in additional mass relationships for which confident sequence identification of the second peptide was not obtained (Supplementary Table 1), suggesting that this region of the RTP is involved in binding other proteins that we could not identify (perhaps plant proteins) or PLRV peptides with additional modifications (either co-analytical *or* biologically-relevant) that we were unable to identify using our database searching strategy.

PIR Results are Useful for Structure Prediction

Luteovirids have icosahedral symmetry ubiquitous among many plant and animal viruses[66]. The geometry of icosahedral virions is generated by interlocking multimers of CP subunits[67, 68]. Luteovirids are thought to incorporate 180 copies of the CP into virions with T=3 quasi-equivalence. A minor proportion of the CPs in the virion will be replaced with the full-length RTP [7, 35], but even the structure of the PLRV CP monomer is not known. Therefore, we modeled the monomer structure using Phyre2. Phyre2 accurately predicts protein structures by comparing the linear amino acid sequence against a database of approximately 10 million proteins with the PSI-BLAST algorithm to identify patterns of mutation across evolutionary time[59]. Hidden Markov modeling is used to create a protein fingerprint, which is compared against a similarly-derived database of proteins with known structures (Protein Data Bank). Thus, highly accurate structural predictions can be generated from proteins with remote amino acid homology. Phyre2 produced a PLRV CP monomer structure similar to our previous report [7] (Figure 7). 86% of residues modeled at >90% confidence with the crystal structure of *Ryegrass mottle virus* coat protein [69] (pdb=c2izw) being used as the top aligned model. The PLRV CP can be divided into at least two domains, the N-terminal random (R) domain and the shell (S) domain shared by most positive-sense RNA icosahedral coat proteins[70, 71]. The topology of the S-domain of the PLRV CP monomer model resembles the jelly roll [72] that is conserved in the S-domains of most icosahedral viruses with solved structures. The R-domain protrudes from the jelly roll S domain (Figure 7) in this PLRV monomer model.

Structural modeling of the full-length RTP using Phyre2 or I-Tasser[73] produced no models matching with significant statistical confidence to any previously solved structures. This is consistent with our finding that the RTD of luteovirids are predicted to be highly disordered using the disorder prediction algorithms PONDR[74] and IUPred[75] (data not shown).

A top-down view of intact icosahedral virus particles infecting humans[76–78] and plants[67] would reveal interactions of CP trimers along multiple axes of symmetry, including at the β-annulus. We imported the PLRV CP monomer structure into Symmdock to model the structure of CP trimers as they would appear in assembled virion particles. Symmdock is an algorithm for prediction of complexes with C_n symmetry by geometry-based docking given the structure of the asymmetric monomer unit [60, 79]. Modeling the trimer without the cross-linking constraints from the homodimer interaction produced more than 1780 possible structures. After using the single cross-linking constraint provided by the homodimer peptide (Figure 3), only four possible trimer structures remained. The top model (Figure 8A) converged with previously reported structures of other T=3 quasi-symmetry icosahedral viruses[67], placing interactions within the S-domain between N186, K188, and D191 in proximity to the β-annulus of the icosahedrons. These interactions are identical to those reported to provide structural integrity at the homologous site in *Rice yellow mottle virus* (RYMV) [67], a plant icosahedral virus in the family *Sobemoviridae*. The model is also consistent with the data that link K188 in the CP to K405 in the RTP (Table 1, XL-CP-

RTP). The model presented here places the GNGKSSDPAGSFR peptide on the external surface of the virion where it would be available to interact with incorporated RTP at this site. Additionally, the trimer model places the protruding positively-charged N-terminal R-rich domains towards the internal face of the virions, where they would be available for interactions with viral RNA (Figure 8). This observation is supported by *in vitro* sequence-independent R-domain-viral RNA interactions in the *Southern cowpea mosaic virus* [71] and by the crystal structure of the *Tomato bushy stunt virus* [80].

Linking PLRV topologies to Virion Biological Function Using Reverse Genetics

PLRV is an ideal virus for investigating interaction topologies that regulate virion assembly, systemic movement in plants, and aphid transmission. A suite of PLRV mutants[1, 6–8] with defects in capsid assembly, aphid transmission, and host local and systemic movement were described previously. The virus mutants can be tested for defects in the assembly of virus particles in infected host plants using DAS-ELISA and TAS-ELISA[7]. RNase-protection assays [6] can also be used to examine further whether CP mutations interfere with assembly of viral RNA into virions. Figure 9 illustrates a summary of a panel of assays performed with previously described mutants, where mutations were induced in sequences nearby the cross-linked sites identified in the present report. ELISA and RNAse protection assays proved a triple amino acid mutant, GNG, was unable to assemble stable virions[6]. The deleted residues in the GNG mutant lie immediately adjacent to the cross-linked lysine within the CP peptide **GNG**KSSDPAGSFR that we detected in XL-CP (Figure 3). Until now, linking the GNG deletion (or any luteovirus mutation) to a topological defect in luteovirid capsid monomer interactions was not possible. On the basis of only the linear amino acid sequence, at least two hypotheses to explain why deleting GNG prevents virion formation are plausible, (a) the GNG residues for an interface for CP multimer interactions, or (b) GNG residues are at a site distal to the monomer interaction interface and their deletion has an allosteric effect that changes the topology of the multimer interaction interface. The discovery of the cross-linked homodimer peptide provides strong evidence that close interactions between CP monomers at the GNG interface (<30Å between K188 residues in each monomer, Figure 8) are involved in maintaining the integrity of CP multimers and clearly explains why this deletion mutant fails to assemble stable virions, perhaps via disrupting hydrogen bonding among the GNG residues that contribute to virion stability.

Our previous studies have assigned several different functions to the readthrough domain (RTD), i.e. that portion of the RTP encoded by ORF 5 [1, 8] (Figure 1). It is likely that all virions contain some similar number of CP monomers that are replaced with full-length RTP (CP+RTD). The CP domain of the RTP fits within an icosahedral virion structure with the RTD projecting outward from the surface. Virions can assemble using only CP monomers and they are stable in plants and aphids, but their movement is impaired in both. The N-terminal portion of the RTD must be incorporated into the virion for virus transmission to occur [8, 28]. Virus without the N-terminal portion of the RTD does move across the aphid gut, albeit less efficiently than WT virus. However, it does not move across the accessory salivary gland barrier and it appears to be degraded more rapidly in the aphid hemolymph. PLRV three amino acid deletion mutants in the N-terminal portion differ in their transmission efficiency [8]. One of these mutants, RFI, has the RFI deletion immediately adjacent to the cross-linked K230 in XL-RTP-1 peptide QNPK.EVDSGSEPGPSPQPTPTPTPQKHE**R.FI**AY (Table 1). This mutant incorporates RTP into virions at a low level and systemically infects host plants including *N. benthamiana, N. clevelandii, P. floridana* and *S. sarrachoides* [8]. However, the RFI deletion mutant is not aphid transmissible when aphids are fed on purified mutant virus or when purified mutant virus is injected directly into the aphid hemocoel [8]. The XL-RTP-1 cross-

link strongly suggests that there are topological features in this region that are structurally important for aphid transmission to occur via interactions with residues distal in the RTP, near K351 and critical for incorporation of the RTP into virions at levels that are required for efficient transmission by aphids. We identified additional interactions in the RTP that are critical for RTP incorporation into virions. In PLRV, deletion of amino acids GPA within the peptide FFLV**GPA**IQKTAK that was identified in XL-RTP-2 abolishes RTP incorporation into virions (and by default, aphid transmission)[8], although the mutant virus was able to replicate and accumulate virus to WT-levels [8] in the host plants. The N-terminal 200 residues of the RTD is sensitive to mutagenesis [8] and this feature of the RTD is not unique to PLRV. Deletions in the N-terminal region of the RTD of the luteovirus *Beet western yellows virus* also abolish RTP incorporation and aphid transmission [28]. Our results demonstrate that the defects in RTP incorporation and aphid transmission in these various luteovirid mutants are due to disruptions in protein-interaction topologies *within* the RTP N-terminal domain. Such disruptions themselves, may preclude RTP incorporation by these mutants due to the perturbation of structural features required for RTP incorporation. Alternatively, the altered capsid topology of these mutants, in particular changes in the topology of the XL-RTP-1 interaction, may disrupt interactions with aphid vector proteins that are required for aphid transmission. The XL-RTP-1 interaction interface is a very exciting target, the first of its kind to be experimentally derived, for future studies on the molecular and biophysical basis of PLRV-aphid interactions.

PLRV is retained in phloem tissue where it is accessible to aphids for dispersal [1]. Phloem tissue tropism is mediated by a trans-acting effect of the RTP [1]. A deletion of the entire C-terminal portion of the RTP allows PLRV to move out of the phloem and infect mesophyll tissues [1]. Using PIR, we identified key topological features in the C-terminal domain of the RTP. Based on the phenotype of the previously described C-terminal deletion mutant, we hypothesized that the topology of the XL-RTP2 cross-linked site may be responsible for interactions with host proteins. To test whether the K405-K470 cross-linked site was specifically involved in PLRV-host interactions, we created a PLRV mutant, RTC12, which lacked a 26 amino acid stretch encompassing K470 in the C-terminal domain of the RTP (Supplementary Figure 3A). RTC12 was infectious and efficiently inoculated the host hairy nightshade (*S. sarrachoides*). A truncated form of the RTP was produced in plants (Supplementary Figure 3B). Two out of 2 plants became infected when tested by ELISA (absorbance values >3.0 at 405nm). Strikingly, the RTC12 deletion mutant showed WT symptoms of infection as well as a previously described PLRV mutant lacking the entire RTP C-terminal domain[1]. Additionally, symptoms in RTC12 infected plants were significantly delayed compared to infection with WT virus, 5.6 (+/− 1.4) weeks to 3.7 (+/− 0.3) weeks to symptom expression in plants, respectively. The remarkable similarity in phenotype between RTC12 and the previously described C-terminal deletion mutant implicate the topology of the interface defined by the crosslink K405-K470 as a critical feature of the RTP involved in virus-host interactions. Undoubtedly, discovering the interacting host proteins in this region of the RTP will be the exciting next step.

PLRV mutant analysis proves that there are critical virion topologies that we did not yet identify. Other structural mutants lie in tryptic peptides that we would not have been able to identify due to their size or accessibility of the PIR to the interior of the virion. However, all the cross-linked sites overlapped with mutants defective in virion assembly and RTP incorporation. Several mutations distal to the cross-linked sites conferred no loss of function, and we did not identify cross-links in these otherwise identifiable regions. This experimentally-derived overlap of structure and function is striking. These mutants were made over years in effort to characterize luteovirid CP sequences and function. PIR technology enabled us to confirm the bioinformatic predictions and biological assays by

showing these regions interact in functional capsids and provided us with distance constraints for structural modeling to observe how these interactions occur in virus particles.

CONCLUSIONS

Mass spectrometry is rapidly gaining popularity as a tool to explore virus structure, dynamics, and function[42–45, 81]. Here, we applied PIR technology for detection of cross-linked peptide pairs to examine PLRV topology. The unique features of PIR technology were essential to our ability to identify the cross-linked sites within the cross-linked PLRV peptides. By focusing our analysis on those ions which fulfill the expected mass relationships of PIR technology we were able to identify cross-linked peptide pairs and obtained peptide sequence information from the released peptides in a single LC-MS run, which is a benefit over previous strategies such as that described in Zheng et al [52]. The topologies in PLRV were critical to capsid function and may be conserved among members of *Luteoviridae*. Akin to classical reverse genetics experiments that link gene knock-out to changes in phenotype, virus mutants enabled us to test predictions about protein complex formation and virion topology based on the mutation location and the movement phenotypes in aphids and plants. By linking PIR and molecular virology, we identified an interface required for virion stability and topologies in the RTP required for RTP incorporation, aphid transmission, and host systemic movement. We also identified many mass relationships for which we could not establish peptide identifications. Due to the inherent biochemical complexity of proteins and their vast array of potential modifications, the identification of all peptides that are cross-linked in a given experiment remains a challenge moving forward; however, this limitation applies to all types of proteomics experiments. However, we were quite successful in elucidating the topologies of PRLV that are required for the major functions of the capsid (Figure 9). Since luteovirids move throughout cells in aphids and from cell-to-cell in plants as virions, the virion topology alone is guiding the circulative transmission process. For the first time, we provided information about interactions in the recalcitrant and highly disordered RTP that guide these processes. We hypothesize that the functional roles and quite possibly the structures of the RTP and virions with incorporated RTP will only be fully realized when studied as part of the multi-protein complexes that exist in aphids and plants due to the high level of intrinsic disorder. PIR technology will certainly be very useful to address this exciting new frontier by revealing the general principles that describe how the topology of arthropod-borne viruses change throughout their movement from host to host via insect vectors.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors graciously thank Dawn Smith (Cornell), Jason Ingram (USDA, ARS) for assistance rearing the aphids, and Kevin Howe (USDA, ARS) for comments on the manuscript, Jackie Mahoney (Cornell) for assistance with data formatting, Chunxiang Zheng (University of Washington) for running the disorder prediction programs, and Tom Hammond (Cornell) for care of the plants. The work was graciously funded by USDA NRI project 1907-22000-018-13, NSF-BREAD project 1109989, and NIH grants 2R01GM086688-03, 5R01RR023334-04, and 7S10RR025107.

References

1. Peter KA, Gildow F, Palukaitis P, Gray SM. The C terminus of the polerovirus p5 readthrough domain limits virus infection to the phloem. J Virol. 2009; 83(11):5419–29. [PubMed: 19297484]

2. Bourdin D, Rouze J, Tanguy S, Robert Y. Variation among clones of Myzus persicae and Myzus nicotianae in the transmission of a poorly and a highly aphid- transmissible isolate of potato leafroll luteovirus (PLRV). Plant Pathology. 1998; 47(6):794–800.
3. Garret A, Kerlan C, Thomas D. Ultrastructural study of acquisition and retention of potato leafroll luteovirus in the alimentary canal of its aphid vector, Myzus persicae Sulz. Arch Virol. 1996; 141(7):1279–92. [PubMed: 8774687]
4. Gildow FE, Reavy B, Mayo MA, Duncan GH, Woodford JA, Lamb JW, Hay RT. Aphid Acquisition and Cellular Transport of Potato leafroll virus-like Particles Lacking P5 Readthrough Protein. Phytopathology. 2000; 90(10):1153–61. [PubMed: 18944480]
5. Hogenhout SA, van der Wilk F, Verbeek M, Goldbach RW, van den Heuvel JF. Identifying the determinants in the equatorial domain of Buchnera GroEL implicated in binding Potato leafroll virus. J Virol. 2000; 74(10):4541–8. [PubMed: 10775590]
6. Kaplan IB, Lee L, Ripoll DR, Palukaitis P, Gildow F, Gray SM. Point mutations in the potato leafroll virus major capsid protein alter virion stability and aphid transmission. J Gen Virol. 2007; 88(Pt 6):1821–30. [PubMed: 17485544]
7. Lee L, Kaplan IB, Ripoll DR, Liang D, Palukaitis P, Gray SM. A surface loop of the potato leafroll virus coat protein is involved in virion assembly, systemic movement, and aphid transmission. J Virol. 2005; 79(2):1207–14. [PubMed: 15613347]
8. Peter KA, Liang D, Palukaitis P, Gray SM. Small deletions in the potato leafroll virus readthrough protein affect particle morphology, aphid transmission, virus movement and accumulation. J Gen Virol. 2008; 89(Pt 8):2037–45. [PubMed: 18632976]
9. Terradot L, Simon JC, Leterme N, Bourdin D, Wilson ACC, Gauthier JP, Robert Y. Molecular characterization of clones of the Myzus persicae complex (Hemiptera: Aphididae) differing in their ability to transmit the potato leafroll luteovirus (PLRV). Entomological Research. 1999; 89(4):355–363.
10. van den Heuvel JF, Verbeek M, van der Wilk F. Endosymbiotic bacteria associated with circulative transmission of potato leafroll virus by Myzus persicae. J Gen Virol. 1994; 75 (Pt 10): 2559–65. [PubMed: 7931143]
11. Gray S, Gildow FE. Luteovirus-aphid interactions. Annu Rev Phytopathol. 2003; 41:539–66. [PubMed: 12730400]
12. Hogenhout SA, Ammar el D, Whitfield AE, Redinbaugh MG. Insect vector interactions with persistently transmitted viruses. Annu Rev Phytopathol. 2008; 46:327–59. [PubMed: 18680428]
13. Gray SM, Banerjee N. Mechanisms of arthropod transmission of plant and animal viruses. Microbiol Mol Biol Rev. 1999; 63(1):128–48. [PubMed: 10066833]
14. Creamer R, Falk BW. Direct detection of transcapsidated barley yellow dwarf luteoviruses in doubly infected plants. Journal of General Virology. 1990; 71:211–217.
15. Rochow WF. Barley yellow dwarf virus: phenotypic mixing and vector specificity. Science. 1970; 167(3919):875–8. [PubMed: 17742616]
16. Brault V, Bergdoll M, Mutterer J, Prasad V, Pfeffer S, Erdinger M, Richards KE, Ziegler-Graff V. Effects of point mutations in the major capsid protein of beet western yellows virus on capsid formation, virus accumulation, and aphid transmission. J Virol. 2003; 77(5):3247–56. [PubMed: 12584348]
17. Brault V, Mutterer J, Scheidecker D, Simonis MT, Herrbach E, Richards K, Ziegler-Graff V. 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. 2000; 74(3): 1140–8. [PubMed: 10627524]
18. Gildow FE. Evidence for receptor-mediated endocytosis regulating luteovirus acquisition by aphids. Phytopathology. 1993; 83:270–277.
19. Gildow FE, Rochow WF. Role of accessory salivary glands in aphid transmission of barley yellow dwarf virus. Virology. 1980; 104(1):97–108. [PubMed: 18631658]
20. Chay CA, Gunasinge UB, DineshKumar SP, Miller WA, Gray SM. 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. Virology. 1996; 219(1):57–65. [PubMed: 8623554]

21. Reinbold C, Gildow FE, Herrbach E, Ziegler-Graff V, Goncalves MC, van Den Heuvel JF, Brault V. Studies on the role of the minor capsid protein in transport of Beet western yellows virus through Myzus persicae. J Gen Virol. 2001; 82(Pt 8):1995–2007. [PubMed: 11458007]
22. Miller WA, Dinesh KSP, Paul CP. Luteovirus gene expression. Critical Reviews in Plant Sciences. 1995; 14(3):179–211.
23. Brault V, van den Heuvel JF, Verbeek M, Ziegler-Graff V, Reutenauer A, Herrbach E, Garaud JC, Guilley H, Richards K, Jonard G. Aphid transmission of beet western yellows luteovirus requires the minor capsid read-through protein P74. EMBO J. 1995; 14(4):650–9. [PubMed: 7882968]
24. Cheng SL, Domier LL, D'Arcy CJ. Detection of the readthrough protein of barley yellow dwarf virus. Virology. 1994; 202(2):1003–6. [PubMed: 8030200]
25. Filichkin SA, Lister RM, McGrath PF, Young MJ. In vivo expression and mutational analysis of the barley yellow dwarf virus readthrough gene. Virology. 1994; 205(1):290–9. [PubMed: 7975225]
26. Wang JY, Chay C, Gildow FE, Gray SM. Readthrough protein associated with virions of barley yellow dwarf luteovirus and its potential role in regulating the efficiency of aphid transmission. Virology. 1995; 206(2):954–62. [PubMed: 7856106]
27. Reutenauer A, Ziegler GV, Lot H, Scheidecker D, Guilley H, Richards K, Jonard G. Identification of beet western yellows luteovirus genes implicated in viral replication and particle morphogenesis. Virology. 1993; 195(2):692–699. [PubMed: 8337839]
28. Bruyere A, Brault V, Ziegler-Graff V, Simonis MT, Van den Heuvel JF, Richards K, Guilley H, Jonard G, Herrbach E. 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. 1997; 230(2):323–34. [PubMed: 9143288]
29. Burrows ME, Caillaud MC, Smith DM, Benson EC, Gildow FE, Gray SM. Genetic Regulation of Polerovirus and Luteovirus Transmission in the Aphid Schizaphis graminum. Phytopathology. 2006; 96(8):828–37. [PubMed: 18943747]
30. Burrows ME, Caillaud MC, Smith DM, Gray SM. Biometrical genetic analysis of luteovirus transmission in the aphid Schizaphis graminum. Heredity. 2007; 98(2):106–13. [PubMed: 17021612]
31. Cilia M, Howe K, Fish T, Smith D, Mahoney J, Tamborindeguy C, Burd J, Thannhauser TW, Gray S. Biomarker discovery from the top down: protein biomarkers for efficient virus transmission by insects (Homoptera: Aphididae) discovered by coupling genetics and 2-D DIGE. Proteomics. 2011; 11:2440–2458. [PubMed: 21648087]
32. Cilia M, Tamborindeguy C, Fish T, Howe K, Thannhauser TW, Gray S. Genetics coupled to quantitative intact proteomics links heritable aphid and endosymbiont protein expression to circulative polerovirus transmission. J Virol. 2011; 85(5):2148–66. [PubMed: 21159868]
33. Yang X, Thannhauser TW, Burrows M, Cox-Foster D, Gildow FE, Gray SM. Coupling genetics and proteomics to identify aphid proteins associated with vector-specific transmission of polerovirus (luteoviridae). J Virol. 2008; 82(1):291–9. [PubMed: 17959668]
34. Gray SM, Gildow F. Luteovirus-aphid interactions. Ann Rev Phytopath. 2003; 41:539–566.
35. Callaway A, Giesman-Cookmeyer D, Gillock ET, Sit TL, Lommel SA. The multifunctional capsid proteins of plant RNA viruses. Annu Rev Phytopathol. 2001; 39:419–60. [PubMed: 11701872]
36. Garret A, Kerlan C, Thomas D. Ultrastructural study of acquisition and retention of potato leafroll luteovirus in the alimentary canal of its aphid vector, Myzus persicae Sulz. Archives of Virology. 1996; 141(7):1279–1292. [PubMed: 8774687]
37. Gildow FE, Damsteegt VD, Stone AL, Smith OP, Gray SM. Virus-vector cell interactions regulating transmission specificity of soybean dwarf luteoviruses. Journal of Phytopathology-Phytopathologische Zeitschrift. 2000; 148(6):333–342.
38. Reinbold C, Herrbach E, Brault V. Posterior midgut and hindgut are both sites of acquisition of Cucurbit aphid-bome yellows virus in Myzus persicae and Aphis gossypii. J Gen Virology. 2003; 84:3473–3484. [PubMed: 14645929]
39. Rouze-Jouan J, Terradot L, Pasquer F, Tanguy S, Ducray-Bourdin DG. The passage of Potato leafroll virus through Myzus persicae gut membrane regulates transmission efficiency. Journal of General Virology. 2001; 82:17–23. [PubMed: 11125153]

40. Briddon RW, Pinner MS, Stanley J, Markham PG. Geminivirus coat protein gene replacement alters insect specificity. Virology. 1990; 177(1):85–94. [PubMed: 2353465]
41. Kaplan IB, Lee L, Ripoll DR, Palukaitis P, Gildow F, Gray SM. Point mutations in the potato leafroll virus major capsid protein alter virion stability and aphid transmission. J Gen Virol. 2006:87.
42. Fu CY, Uetrecht C, Kang S, Morais MC, Heck AJ, Walter MR, Prevelige PE Jr. A docking model based on mass spectrometric and biochemical data describes phage packaging motor incorporation. Mol Cell Proteomics. 2010; 9(8):1764–73. [PubMed: 20124351]
43. Kang S, Hawkridge AM, Johnson KL, Muddiman DC, Prevelige PE Jr. Identification of subunit-subunit interactions in bacteriophage P22 procapsids by chemical cross-linking and mass spectrometry. J Proteome Res. 2006; 5(2):370–7. [PubMed: 16457603]
44. Lanman J, Lam TT, Barnes S, Sakalian M, Emmett MR, Marshall AG, Prevelige PE Jr. Identification of novel interactions in HIV-1 capsid protein assembly by high-resolution mass spectrometry. J Mol Biol. 2003; 325(4):759–72. [PubMed: 12507478]
45. Uetrecht C, Heck AJ. Modern Biomolecular Mass Spectrometry and its Role in Studying Virus Structure, Dynamics, and Assembly. Angew Chem Int Ed Engl. 2011
46. Tang X, Bruce JE. A new cross-linking strategy: protein interaction reporter (PIR) technology for protein-protein interaction studies. Mol Biosyst. 2010; 6(6):939–47. [PubMed: 20485738]
47. Anderson GA, Tolic N, Tang X, Zheng C, Bruce JE. Informatics strategies for large-scale novel cross-linking analysis. J Proteome Res. 2007; 6(9):3412–21. [PubMed: 17676784]
48. Tang X, Munske GR, Siems WF, Bruce JE. Mass spectrometry identifiable cross-linking strategy for studying protein-protein interactions. Anal Chem. 2005; 77(1):311–8. [PubMed: 15623310]
49. Eng JK, Fischer B, Grossmann J, Maccoss MJ. A fast SEQUEST cross correlation algorithm. J Proteome Res. 2008; 7(10):4598–602. [PubMed: 18774840]
50. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999; 20(18):3551–67. [PubMed: 10612281]
51. Zhang H, Tang X, Munske GR, Tolic N, Anderson GA, Bruce JE. Identification of protein-protein interactions and topologies in living cells with chemical cross-linking and mass spectrometry. Mol Cell Proteomics. 2009; 8(3):409–20. [PubMed: 18936057]
52. Zheng C, Yang L, Hoopmann MR, Eng JK, Tang X, Weisbrod CR, Bruce JE. Crosslinking measurements of in vivo protein complex topologies. Mol Cell Proteomics. 2011
53. Liang D, Gray SM, Kaplan I, Palukaitis P. Site-directed mutagenesis and generation of chimeric viruses by in yeast to facilitate an recombination it-virus interactions. Molecular Plant-Microbe Interactions. 2004; 17(6):571–576. [PubMed: 15195939]
54. Kaplan IB, Lee L, Ripoll DR, Palukaitis P, Gildow F, Gray SM. Point mutations in the potato leafroll virus major capsid protein alter virion stability and aphid transmission. Journal of General Virology. 2007; 88:1821–1830. [PubMed: 17485544]
55. Lee L, Kaplan IB, Ripoll DR, Liang D, Palukaitis P, Gray SM. A surface loop of the Potato leafroll virus coat protein is involved in virion assembly, systemic movement, and aphid transmission. Journal of Virology. 2005; 79(2):1207–1214. [PubMed: 15613347]
56. Hammond J, Lister RM, Foster JE. Purification, Identity and Some Properties of An Isolate of Barley Yellow Dwarf Virus from Indiana. J Gen Virol. 1983; 64:667–676.
57. Balgley BM, Laudeman T, Yang L, Song T, Lee CS. Comparative evaluation of tandem MS search algorithms using a target-decoy search strategy. Mol Cell Proteomics. 2007; 6(9):1599–608. [PubMed: 17533222]
58. Deutsch EW, Mendoza L, Shteynberg D, Farrah T, Lam H, Tasman N, Sun Z, Nilsson E, Pratt B, Prazen B, Eng JK, Martin DB, Nesvizhskii AI, Aebersold R. A guided tour of the Trans-Proteomic Pipeline. Proteomics. 2010; 10(6):1150–9. [PubMed: 20101611]
59. Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc. 2009; 4(3):363–71. [PubMed: 19247286]
60. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res. 2005; 33(Web Server issue):W363–7. [PubMed: 15980490]

61. Tang X, Munske GR, Siems WF, Bruce JE. Mass spectrometry identifiable cross-linking strategy for studying protein-protein interactions. Analytical chemistry. 2005; 77(1):311–8. [PubMed: 15623310]
62. Yang L, Zhang H, Bruce JE. Optimizing the detergent concentration conditions for immunoprecipitation (IP) coupled with LC-MS/MS identification of interacting proteins. Analyst. 2009; 134(4):755–62. [PubMed: 19305927]
63. Tang X, Bruce JE. A new cross-linking strategy: protein interaction reporter (PIR) technology for protein-protein interaction studies. Molecular bioSystems. 2010; 6(6):939–47. [PubMed: 20485738]
64. Yang L, Tang X, Weisbrod CR, Munske GR, Eng JK, von Haller PD, Kaiser NK, Bruce JE. A photocleavable and mass spectrometry identifiable cross-linker for protein interaction studies. Analytical chemistry. 2010; 82(9):3556–66. [PubMed: 20373789]
65. Torrance L. Analysis of epitopes on potato leafroll virus capsid protein. Virology. 1992; 191(1): 485–9. [PubMed: 1384231]
66. Ivanovska IL, Miranda R, Carrascosa JL, Wuite GJ, Schmidt CF. Discrete fracture patterns of virus shells reveal mechanical building blocks. Proc Natl Acad Sci U S A. 2011; 108(31):12611–6. [PubMed: 21768340]
67. Qu C, Liljas L, Opalka N, Brugidou C, Yeager M, Beachy RN, Fauquet CM, Johnson JE, Lin T. 3D domain swapping modulates the stability of members of an icosahedral virus group. Structure. 2000; 8(10):1095–103. [PubMed: 11080631]
68. Crick FH, Watson JD. Structure of small viruses. Nature. 1956; 177(4506):473–5. [PubMed: 13309339]
69. Plevka P, Tars K, Zeltins A, Balke I, Truve E, Liljas L. The three-dimensional structure of ryegrass mottle virus at 2. 9 A resolution. Virology. 2007; 369(2):364–74. [PubMed: 17881031]
70. Lee SK, Dabney-Smith C, Hacker DL, Bruce BD. Membrane activity of the southern cowpea mosaic virus coat protein: the role of basic amino acids, helix-forming potential, and lipid composition. Virology. 2001; 291(2):299–310. [PubMed: 11878899]
71. Lee SK, Hacker DL. In vitro analysis of an RNA binding site within the N-terminal 30 amino acids of the southern cowpea mosaic virus coat protein. Virology. 2001; 286(2):317–27. [PubMed: 11485399]
72. Dolja VV, Koonin EV. Phylogeny of capsid proteins of small icosahedral RNA plant viruses. J Gen Virol. 1991; 72 (Pt 7):1481–6. [PubMed: 1856686]
73. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008; 9:40. [PubMed: 18215316]
74. Peng K, Radivojac P, Vucetic S, Dunker AK, Obradovic Z. Length-dependent prediction of protein intrinsic disorder. BMC Bioinformatics. 2006; 7:208. [PubMed: 16618368]
75. Dosztanyi Z, Csizmok V, Tompa P, Simon I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics. 2005; 21(16): 3433–4. [PubMed: 15955779]
76. Chen JZ, Settembre EC, Aoki ST, Zhang X, Bellamy AR, Dormitzer PR, Harrison SC, Grigorieff N. Molecular interactions in rotavirus assembly and uncoating seen by high-resolution cryo-EM. Proc Natl Acad Sci U S A. 2009; 106(26):10644–8. [PubMed: 19487668]
77. Fabry CM, Rosa-Calatrava M, Conway JF, Zubieta C, Cusack S, Ruigrok RW, Schoehn G. A quasi-atomic model of human adenovirus type 5 capsid. EMBO J. 2005; 24(9):1645–54. [PubMed: 15861131]
78. Luque D, Saugar I, Rodriguez JF, Verdaguer N, Garriga D, Martin CS, Velazquez-Muriel JA, Trus BL, Carrascosa JL, Caston JR. Infectious bursal disease virus capsid assembly and maturation by structural rearrangements of a transient molecular switch. J Virol. 2007; 81(13):6869–78. [PubMed: 17442720]
79. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. Geometry-based flexible and symmetric protein docking. Proteins. 2005; 60(2):224–31. [PubMed: 15981269]
80. Harrison SC, Olson AJ, Schutt CE, Winkler FK, Bricogne G. Tomato bushy stunt virus at 2. 9 A resolution. Nature. 1978; 276(5686):368–73. [PubMed: 19711552]

81. Uetrecht C, Barbu IM, Shoemaker GK, van Duijn E, Heck AJ. Interrogating viral capsid assembly with ion mobility-mass spectrometry. Nat Chem. 2011; 3(2):126–32. [PubMed: 21258385]

82. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997; 25(17):3389–402. [PubMed: 9254694]

83. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schaffer AA, Yu YK. Protein database searches using compositionally adjusted substitution matrices. FEBS J. 2005; 272(20): 5101–9. [PubMed: 16218944]

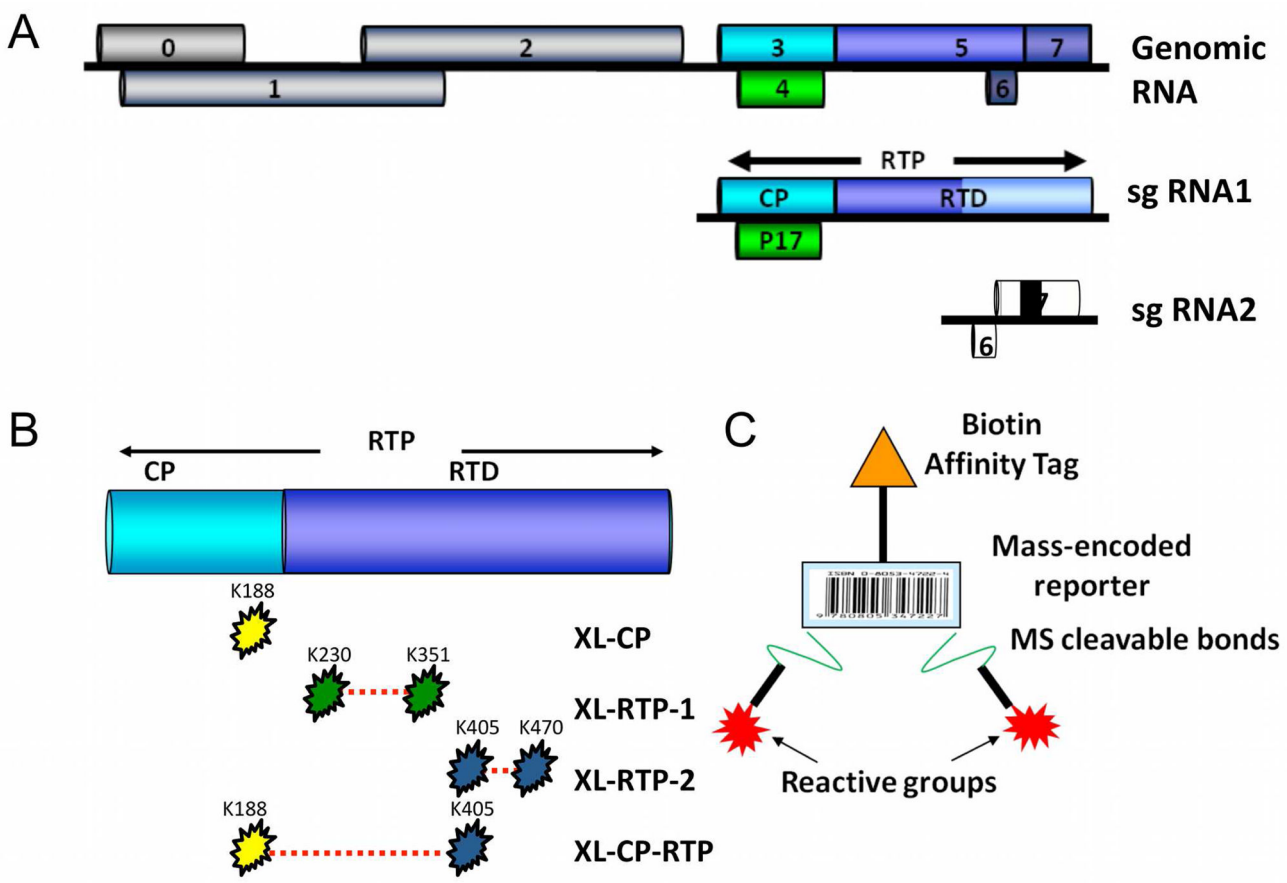


Figure 1. Summary of PLRV and PIR structural features

A Luteovirid genome organization of PLRV. Structural proteins, CP and RTP and the P17 movement protein are translated from subgenomic RNA1. The RTD is divided into two functional domains, the N-terminal and C-terminal. ORFs 0–2 function in replication and silencing suppression. ORF 6–7 function and translation efficiency from sgRNA2 are unknown. **B** Diagram indicating the cross-linked sites identified on the PLRV RTP protein sequence. XL-CP refers to the homodimer cross-link in coat protein, XL-RTP1 and XL-RTP2 are two distinct cross-links between different sites in the read through domain, and XL-CP-RTP is the cross-link between the coat protein and read through domains. **C** Cartoon schematic of the PIR molecule indicating important structural features.

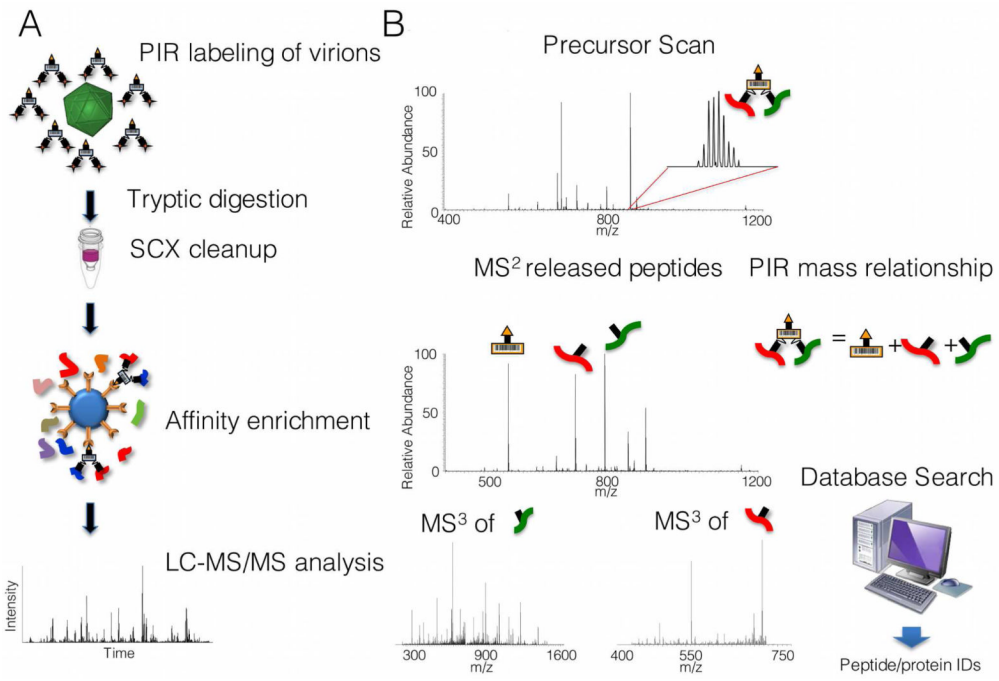


Figure 2. Experiment Overview

A Experimental workflow for PIR labeling of PLRV virons and preparation of the sample for LC-MS analysis, strong cation exchange (SCX). **B** Illustration of the identification of a cross-linked peptide pair using PIR mass relationships. First, a high resolution and mass accuracy measurement of the intact cross-linked peptide precursor ion is made and this is called the precursor scan. This ion is then subjected to CID using only enough collision energy to fragment the Rink bonds, and a second high resolution and mass accuracy measurement is made on the fragment ions, which consist of the reporter ion as well as the released peptide ions. If the masses of these fragment ions fulfill the specified mass relationship for the PIR molecule being used then the two released peptides are each isolated subjected to MS^3 CID at sufficient energy to fragment the peptide backbone and to obtain primary sequence information for each ion.

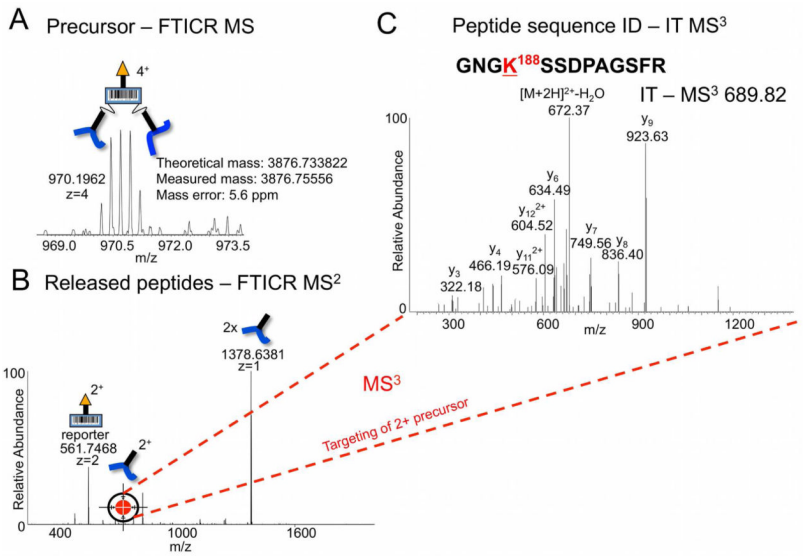


Figure 3. MS detection of the XL-CP cross-linked pair

A Precursor scan detects intact cross-linked species containing cross-linker reagent and two identical peptides from PLRV CP monomers at *m/z* = 970.916. **B** MS^2 scan of released peptide ions after CID cleavage of RINK bonds, at *m/z* = 1378.638 and z = 2 reporter ion at *m/z* = 561.746. **C** MS^3 scan of CID-generated peptide backbone fragment ions provides amino acid sequence information, E-value = $3.2e^{-3}$. If z = 1 ions are present, z = 2 ions of released peptides are also targeted for MS^3 fragmentation. FTICR indicates scans MS and MS^2 performed in Fourier transform ion cyclotron resonance cell of the mass spectrometer, IT indicates MS^3 scan performed in the ion trap.

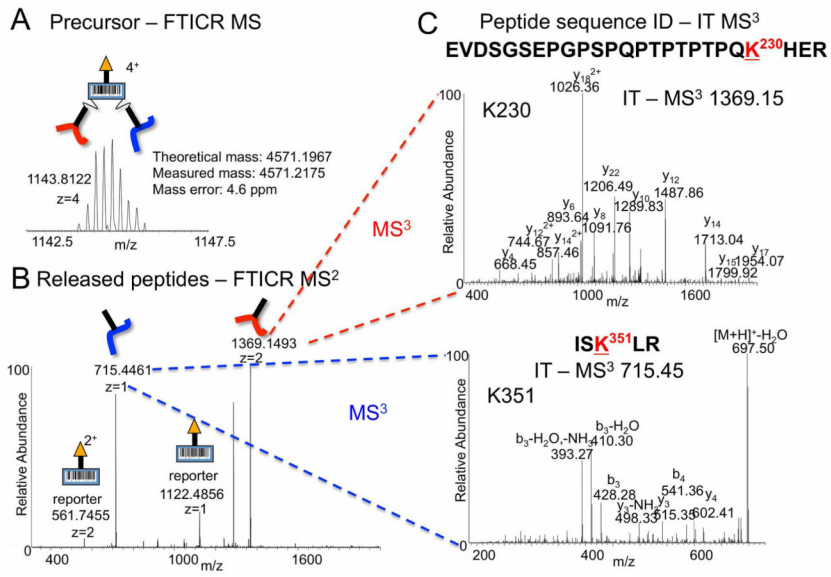


Figure 4. MS detection of the XL-RTP-1 cross-linked pair

A Precursor scan detects intact cross-linked species containing cross-linker reagent and two different peptides from PLRV RTP at *m/z* = 1143.812. **B** MS^2 scan of released peptide ions after CID cleavage of RINK bonds, at *m/z* = 1369.149, *m/z* = 715.4461, and z = 2 reporter ion at *m/z* = 561.745. **C** MS^3 scan of CID-generated peptide backbone fragment ions provides amino acid sequence information for two peptides matching the RTP, z = 2 ion CID fragmentation matches the peptide EVDSGSEPGPSPQPTPTPTPQKHER with the mass of the PIR stump on K(22) and z = 1 ion CID fragmentation matches the peptide ISKLR with the stump mass on K(3). FTICR indicates scans MS and MS^2 performed in Fourier transform ion cyclotron resonance cell of the mass spectrometer, IT indicates MS^3 scan performed in the ion trap.

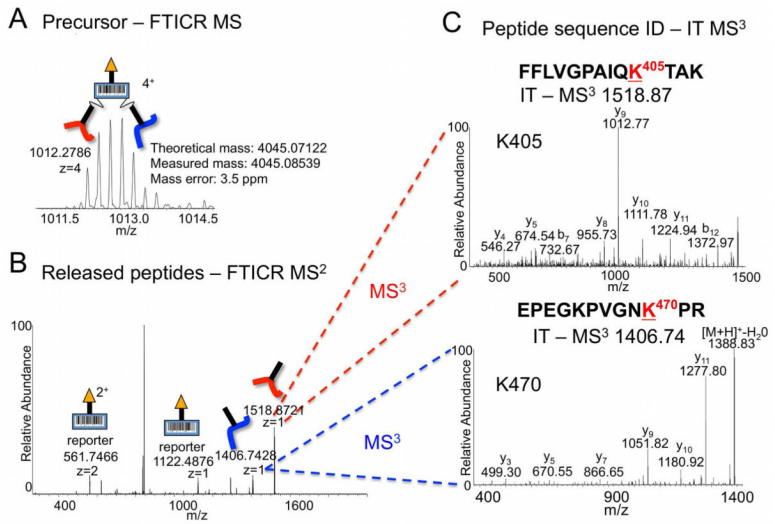


Figure 5. MS detection of the XL-RTP-2 cross-linked pair

A Precursor scan detects intact cross-linked species containing cross-linker reagent and two different peptides from PLRV RTP at *m/z* = 1012.278. **B** MS^2 scan of released peptide ions after CID cleavage of RINK bonds, at *m/z* = 1518.872, *m/z* = 1406.742, and z=2 reporter ion at *m/z* = 561.746. **C** MS^3 scan of CID-generated peptide backbone fragment ions provides amino acid sequence information for two peptides matching the RTP, z = 1 ion CID fragmentation of 1518.872 matches the peptide FFLGVPAIQKTAK, with the mass of the PIR stump on K(10) and z = 1 ion CID fragmentation matches the peptide EPEGKPVGNKPR with the stump mass on K(10). FTICR indicates scans MS and MS^2 performed in Fourier transform ion cyclotron resonance cell of the mass spectrometer, IT indicates MS^3 scan performed in the ion trap.

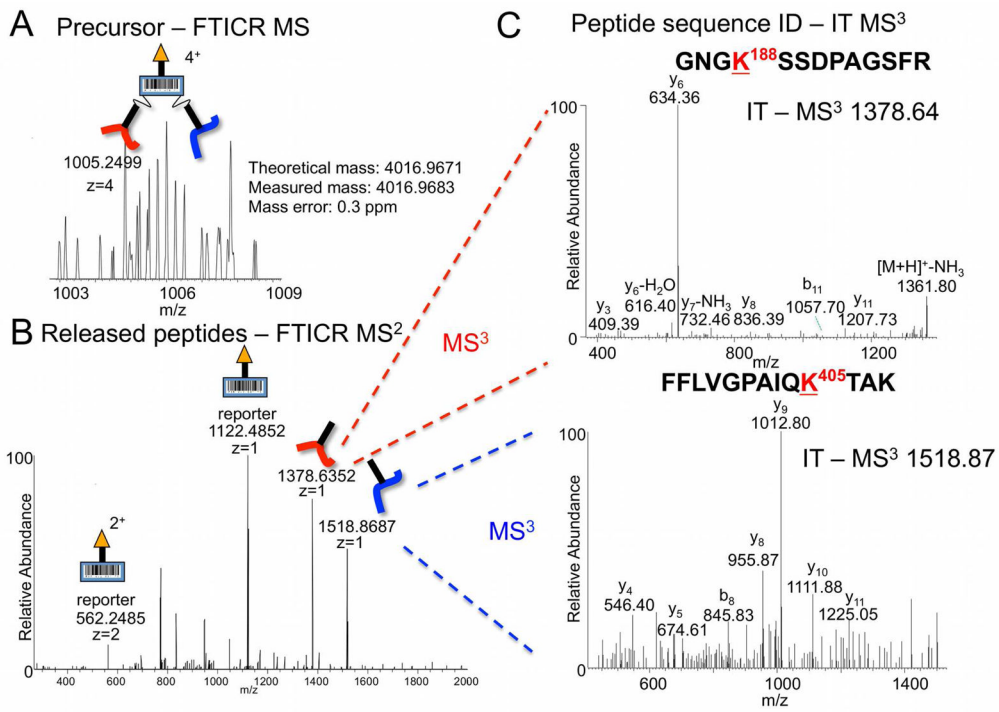


Figure 6. MS detection of the XL-CP-RTP cross-linked pair

K 188 is on the exterior of the virion and the same site identified in the XL-CP cross-link. **A** Precursor scan detects intact cross-linked species containing cross-linker reagent, one peptide from the CP and one peptide from the RTP at *m/z* = 1005.249. **B** MS^2 scan of released peptide ions after CID cleavage of RINK bonds, at *m/z* = 1518.872, *m/z* = 1378.64, and z = 2 reporter ion at *m/z* = 561.746. **C** MS^3 scan of CID-generated peptide backbone fragment ions provides amino acid sequence information for the two released peptides, z = 1 ion CID fragmentation of 1518.872 matches the peptide FFLGVPAIQKTAK, with the mass of the PIR stump on K(10) and z = 1 ion CID fragmentation of 1378.64 matches the peptide GNGKSSDPAGSFR with the stump mass on K(4). FTICR indicates scans MS and MS^2 performed in Fourier transform ion cyclotron resonance cell of the mass spectrometer, IT indicates MS^3 scan performed in the ion trap.

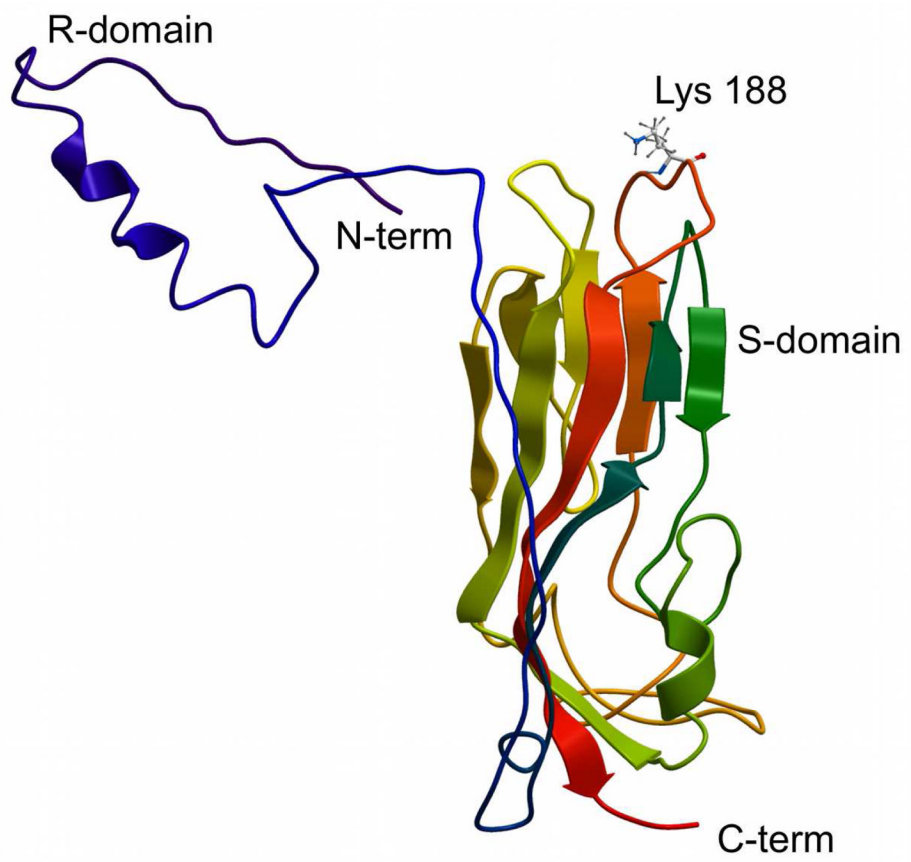


Figure 7. PLRV monomer modeled using Phyre^2, residues 1-208 of the CP

The disordered RTD is not included in the model, but would extend from the C-term. The S-domain of the PLRV CP monomer adopts the jelly roll configuration, the reactive K 188 is highly exposed on the monomer structure. The disordered N-terminal R-domain protrudes from the S-domain.

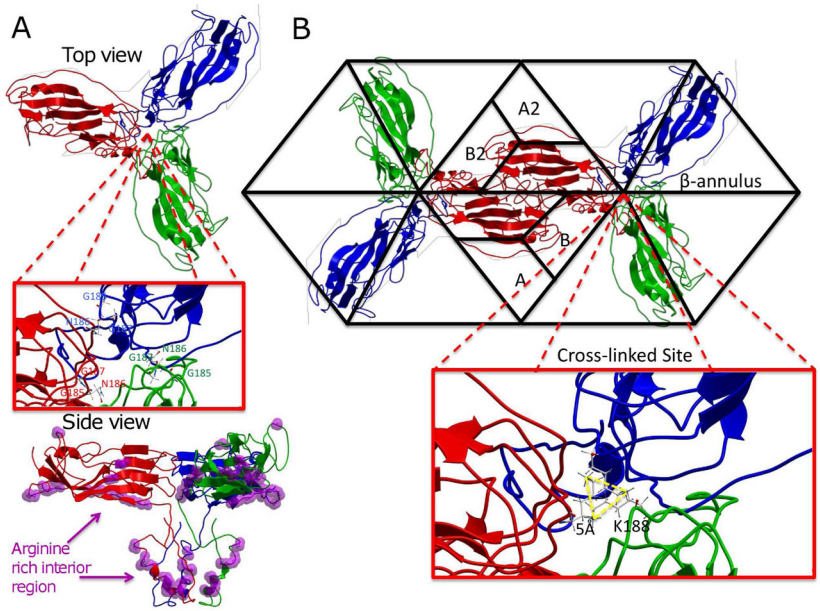


Figure 8. PLRV trimer modeled with SymmDock

A Interactions between CP monomers of PLRV modeled with cross-linking constraint between CP homodimer. Top view of symmetric trimer highlighting remarkable features of the model, including the conserved GNG interface required for virion formation and the arginine-rich R-Domain that is predicted to line the interior of the capsid. **B** Cages showing T=3 equivalent surface lattices with CP trimers modeled along the 2-fold axis of symmetry. Cross-linked K 188 at the β-annulus are within 5Å in each CP monomer.

Interaction Topology and Virion Function

	Virion Assembly	Systemic Infection	RTP Incorporation	Aphid Transmission
XL-CP	No	No	N/A	N/A
XL-RTP-1	Yes	Yes	Reduced	No
XL-RTP-2	Yes	Delayed[a]	No	N/A
XL-CP-RTP	N/A	Reduced	No	N/A

Figure 9. Reverse Genetics Places Topologies into Virion Biological Contexts

PLRV mutant analysis places the interaction topology (XL-CP, XL-RTP-1, XL-RTP-2, XL-CP-RTP) into a functional context. When virus mutants are made, they are characterized for 4 major biological functions: virion assembly, systemic infection, RTP incorporation into virions, and aphid transmission. “No” indicates a mutation at or near the cross-linked site disrupts one of these functions, “yes” indicates the mutant retains the function. XL-CP is defective in virions assembly/stability, XL-RTP-1 is reduced in aphid transmissibility, and XL-RTP-2 is defective in host systemic movement.

Table 1

Discovering *Potato leafroll virus* interaction topology using PIR technology and mass spectrometry.

Cross-link[a]	Precursor Mass[b]	Peptide Mass	Sequence[c]	Reactive K site[d]	Mascot Score	X-corr[e]	E-value[f]
XL-CP	3876.755	1377.630	ILWK.GNGKSSDPAGSFR.VTIR	188	35	1.300	$5.0e^{-4}$
XL-RTP-1	4571.205	714.439	DGVK.ISKLR.NDNT	351	13	1.015	$4.6e^{-2}$
		2735.282	QNPK.EVDSGSEPGPSPQPTPTPTPQKHE**R.FI**AY	230	41	N/A	$2.6e^{-4}$
XL-RTP-2	4045.088	1517.864	TDGR.FFLV**GPA**IQKTAK.YNYT	405	43	2.155	$4.5e^{-5}$
		1405.734	ASPR.EPEGKPVGNKPR.DETP	470	31	1.542	$3.2e^{-2}$
XL-CP-RTP	4016.967	1377.628	ILWK.**GNG**KSSDPAGSFR. VTIR	188	12	1.013	$2.0e^{-4}$
		1517.861	TDGR.FFLV**GPA**IQKTAK.YNYT	405	16	2.155	$4.0e^{-4}$

[a] Name of cross-linked peptide pair as diagrammed in Figure 1B.

[b] Precursor mass is the measured mass of the precursor from the MS1 scan event in the ICR cell and includes the mass of the PIR reporter ion.

[c] Amino acid sequence deduced from MS^3 fragmentation and database searching with Mascot and/or Sequest. Underlined K indicates position in sequence where the cross-linker stump mass was observed. **Bold** residues indicate a PLRV mutant previously described within the sequence. The four amino acid residues immediately preceding and following the identified peptide sequence are included for reference, and are separated by a "." from the identified peptide sequence.

[d] Amino acid residue number of the cross-linked lysine based on UniProt sequence MCAPS_PLRVW including the initial methionine

[e] Sequest cross-correlation score (X-corr)

[f] Lowest observed E-values are reported for peptides that were identified in Mascot or SEQUEST. In both search engines, the equation is based on the E-value used by BLAST[82, 83].