Chlorovirus Skp1-Binding Ankyrin Repeat Protein Interplay and Mimicry of Cellular Ubiquitin Ligase Machinery

Eric A. Noel  
*University of Nebraska-Lincoln, eric.andrew.noel@gmail.com*

Ming Kang  
*University of Nebraska - Lincoln*

Jiri Adamec  
*University of Nebraska - Lincoln*

James L. Van Etten  
*University of Nebraska - Lincoln, jvanetten1@unl.edu*

George A. Oyler  
*University of Nebraska - Lincoln*

Follow this and additional works at: [https://digitalcommons.unl.edu/vanetten](https://digitalcommons.unl.edu/vanetten)

Part of the Genetics and Genomics Commons, Plant Pathology Commons, and the Viruses Commons

Noel, Eric A.; Kang, Ming; Adamec, Jiri; Van Etten, James L.; and Oyler, George A., "Chlorovirus Skp1-Binding Ankyrin Repeat Protein Interplay and Mimicry of Cellular Ubiquitin Ligase Machinery" (2014). James Van Etten Publications. 9.

[https://digitalcommons.unl.edu/vanetten/9](https://digitalcommons.unl.edu/vanetten/9)
Chlorovirus Skp1-Binding Ankyrin Repeat Protein Interplay and Mimicry of Cellular Ubiquitin Ligase Machinery

Eric A. Noel, Ming Kang, Jiri Adamec, James L. Van Etten, George A. Oyler

Nebraska Center for Virology, University of Nebraska, Lincoln, Nebraska, USA; School of Biological Sciences, University of Nebraska, Lincoln, Nebraska, USA; Department of Biochemistry, University of Nebraska, Lincoln, Nebraska, USA; Department of Plant Pathology, University of Nebraska, Lincoln, Nebraska, USA; Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, Maryland, USA

ABSTRACT

The ubiquitin-proteasome system is targeted by many viruses that have evolved strategies to redirect host ubiquitination machinery. Members of the genus Chlorovirus are proposed to share an ancestral lineage with a broader group of related viruses, nucleo-cytoplasmic large DNA viruses (NCLDV). Chloroviruses encode an Skp1 homolog and ankyrin repeat (ANK) proteins. Several chlorovirus-encoded ANK repeats contain C-terminal domains characteristic of cellular F-boxes or related NCLDV chloropox PRANc (pox protein repeats of ankyrin at C-terminal) domains. These observations suggested that this unique combination of Skp1 and ANK repeat proteins might form complexes analogous to the cellular Skp1-Cul1-F-box (SCF) ubiquitin ligase complex. We identified two ANK proteins from the prototypic chlorovirus Paramecium bursaria chlorella virus-1 (PBCV-1) that functioned as binding partners for the virus-encoded Skp1, proteins A682L and A607R. These ANK proteins had a C-terminal Skp1 interactional motif that functioned similarly to cellular F-box domains. A C-terminal motif of ANK protein A682L binds Skp1 proteins from widely divergent species. Yeast two-hybrid analyses using serial domain deletion constructs confirmed the C-terminal localization of the Skp1 interactional motif in PBCV-1 A682L. ANK protein A607R represents an ANK family with one member present in all 41 sequenced chloroviruses. A comprehensive phylogenetic analysis of these related ANK and viral Skp1 proteins suggested partnered function tailored to the host alga or common ancestral heritage. Here, we show protein-protein interaction between corresponding family clusters of virus-encoded ANK and Skp1 proteins from three chlorovirus types. Collectively, our results indicate that chloroviruses have evolved complementing Skp1 and ANK proteins that mimic cellular SCF-associated proteins.

IMPORTANCE

Viruses have evolved ways to direct ubiquitination events in order to create environments conducive to their replication. As reported in the manuscript, the large chloroviruses encode several components involved in the SCF ubiquitin ligase complex including a viral Skp1 homolog. Studies on how chloroviruses manipulate their host algal ubiquitination system will provide insights toward viral protein mimicry, substrate recognition, and key interactive domains controlling selective protein degradation. These findings may also further understanding of the evolution of other large DNA viruses, like poxviruses, that are reported to share the same monophyly lineage as chloroviruses.

The ubiquitin-proteasome proteolytic pathway is an attractive target for viruses in their battle to create an intracellular environment conducive for their replication. In fact, targeting of ubiquitin-associated enzymes is a reoccurring theme in permissive viral infections (1). This eukaryotic regulatory system mediates a variety of biological processes, including protein turnover, DNA repair, trafficking, and signal transduction (2). Through sequential reactions of three enzyme types, covalent attachment of ubiquitin chains to the substrate is achieved, which then targets the substrate protein for proteasomal degradation. This cascade is initiated by the ubiquitin-activating enzyme (E1), which forms an ATP-dependent high-energy thioester bond with an ubiquitin moiety, enabling passage to a ubiquitin conjugate (E2). Substrate-specific ubiquitin ligases (E3) then catalyze the transfer of ubiquitin from the E2 enzyme to the target protein, creating an isopeptide bond most commonly between a lysine of the substrate and the C-terminal glycine of ubiquitin (Fig. 1) (3). In this system the E3 ubiquitin ligase is the most critical component in determining the selection of substrate proteins for ubiquitin modification, which in turn has pleiotropic cell-regulatory effects, particularly by inducing substrate protein degradation.

Several ubiquitin-interfering viral proteins interact directly with E3 family enzymes, in particular the cullin-based RING (really interesting new gene) finger-type ubiquitin ligase (CRL) SCF (Skp1, cullin, F-box) complex (4, 5). The F-box protein is one of the four subunits of the SCF complex which mediates target recruitment for ubiquitination. F-box proteins are a large family of proteins present in all eukaryotes that are characterized by the presence of the F-box domain, a conserved sequence of approximately 50 amino acids that interacts with the Skp1 component of...
Virulence factors from bacteria and viruses have harnessed F-box proteins compatible with cellular SCF ligase components to mediate substrate recruitment and degradation for the benefit of the pathogen (8–13). Recent genomic analyses have identified a pair of F-box-like function, imitating the Skp1 and F-box components of the eukaryotic SCF complex. Most of the characterized cellular F-box proteins harbor substrate-binding repeat-containing motifs in the C-terminal region, such as leucine-rich repeats (LRR), WD40 repeats (WD), or Kelch repeats (7).

Virulence factors from bacteria and viruses have harnessed F-box proteins compatible with cellular SCF ligase components to mediate substrate recruitment and degradation for the benefit of the pathogen (8–13). Recent genomic analyses have identified a pair of F-box-like function, imitating the Skp1 and F-box components of the eukaryotic SCF complex. Most of the characterized cellular F-box proteins harbor substrate-binding repeat-containing motifs in the C-terminal region, such as leucine-rich repeats (LRR), WD40 repeats (WD), or Kelch repeats (7).

Virulence factors from bacteria and viruses have harnessed F-box proteins compatible with cellular SCF ligase components to mediate substrate recruitment and degradation for the benefit of the pathogen (8–13). Recent genomic analyses have identified a pair of F-box-like function, imitating the Skp1 and F-box components of the eukaryotic SCF complex. Most of the characterized cellular F-box proteins harbor substrate-binding repeat-containing motifs in the C-terminal region, such as leucine-rich repeats (LRR), WD40 repeats (WD), or Kelch repeats (7).

Virulence factors from bacteria and viruses have harnessed F-box proteins compatible with cellular SCF ligase components to mediate substrate recruitment and degradation for the benefit of the pathogen (8–13). Recent genomic analyses have identified a pair of F-box-like function, imitating the Skp1 and F-box components of the eukaryotic SCF complex. Most of the characterized cellular F-box proteins harbor substrate-binding repeat-containing motifs in the C-terminal region, such as leucine-rich repeats (LRR), WD40 repeats (WD), or Kelch repeats (7).

Virulence factors from bacteria and viruses have harnessed F-box proteins compatible with cellular SCF ligase components to mediate substrate recruitment and degradation for the benefit of the pathogen (8–13). Recent genomic analyses have identified a pair of F-box-like function, imitating the Skp1 and F-box components of the eukaryotic SCF complex. Most of the characterized cellular F-box proteins harbor substrate-binding repeat-containing motifs in the C-terminal region, such as leucine-rich repeats (LRR), WD40 repeats (WD), or Kelch repeats (7).

Virulence factors from bacteria and viruses have harnessed F-box proteins compatible with cellular SCF ligase components to mediate substrate recruitment and degradation for the benefit of the pathogen (8–13). Recent genomic analyses have identified a pair of F-box-like function, imitating the Skp1 and F-box components of the eukaryotic SCF complex. Most of the characterized cellular F-box proteins harbor substrate-binding repeat-containing motifs in the C-terminal region, such as leucine-rich repeats (LRR), WD40 repeats (WD), or Kelch repeats (7).
shaking in YPAD (10 g of yeast extract, 20 g of peptone, 20 g of dextrose, and 30 mg of adenine sulfate in 1 liter) medium at 30°C and 250 rpm and diluted to an optical density at 600 nm (OD_{600}) of 0.1 in fresh YPAD medium. Cells were then cultured for an additional 4 h to reach an OD_{600} of 0.4 to 0.6, pelleted by centrifugation at 4,000 × g for 5 min at room temperature, and washed in lithium acetate (LiAc) solution (100 mM LiAc, 10 mM Tris-Cl, 1 mM EDTA, pH 8.0) in a 1/2 volume of original culture. Finally, cells were suspended in LiAc solution in a 1/100 volume of original culture, dispersed into 100-μl aliquots for transformation or supplemented with 5% glycerol and 10% (vol/vol) dimethyl sulfoxide (DMSO), and stored at ~80°C as frozen competent cells. Synthetic complete (SC) medium contained 6.7 g/liter yeast nitrogen base, 20 g/liter dextrose, and 1.4 g/liter amino acid dropout mix (Sigma, St. Louis, MO) omitting leucine and tryptophan (SC-Leu-Trp).

**Gene cloning and plasmid construction.** Genes carried by the host and virus were synthesized with optimization according to the codon preference of *S. cerevisiae* by GenScript (Piscataway Township, NJ). Coding regions were amplified by PCR and incorporated into entry clone pENTR/D-TOPO based on Gateway cloning technology (Invitrogen, Carlsbad, CA) (data not shown). PCR products were cloned directionally into *E. coli* pENTR/D-TOPO vector (Invitrogen). For overexpression of pDEST32 and pDEST22. The structure of each vector was sequence veriﬁed. The *GAL4* DNA-binding domain and *GAL4* DNA activation domain, respectively, the gene of interest was fused in frame with the *GAL4* DNA-binding domain. Entry clones flanked with *attL1* and *attL2* sites underwent LR recombination (Gateway technology; Invitrogen) into the destination vectors. As a result, the gene of interest was fused in frame with the GAL4 DNA-binding domain (DBD) or the activation domain (AD) in the yeast two-hybrid (Y2H) yeast expression vectors pDEST32 and pDEST22 containing the GAL4 DNA-binding domain and GAL4 DNA activation domain, respectively. The *E. coli* cells containing positive cloned plasmids were selected with 50 μg/ml gentamicin and 100 μg/ml ampicillin, respectively, for pDEST32 and pDEST22. The structure of each vector was sequence verified. Plasmids were isolated with a QiAprep Spin Miniprep kit (Valencia, CA) according to the manufacturer’s instructions. For overexpression of the vSkp1 proteins in *E. coli*, a PET-23at (+) vector (Novagen Biosciences, San Diego, CA), which carries an N-terminal T7 tag sequence in addition to a C-terminal His, sequence tag, was used. Total plasmid concentration was quantiﬁed with a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE).

**Y2H screening.** Yeast strain MaV203 competent cells were transformed with bait and prey vectors according to a high-efficiency transformation protocol (30). Yeastmaker carrier DNA (salmon sperm DNA; Clontech, Mountain View, CA) was used to increase transformation efficiency. Selection criteria regarding yeast transformants containing bait or prey plasmid and an empty plasmid were set by resistance to carbenicillin and chloramphenicol. Batch culture was grown to an OD_{600} of 0.6 at 37°C. Induction of protein overexpression was achieved with the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside for 2 h at 25°C. Cells were harvested by centrifugation (5,000 × g) for 5 min at 4°C and then disrupted by sonication in PBS buffer with 2 mM PMSF for 5 min on ice. Cell lysate was centrifuged (16,000 × g) for 10 min at 4°C, and supernatant was collected. For pulldown experiments, the supernatant was collected and diluted with 4 ml of PBS before it was applied to a vSkp1PBCV-1 bound column.

**vSkp1PBCV-1 overexpression and candidate F-box protein pulldown by metal chelation chromatography.** BL21-Gold(DE3)pLysS cells were transformed with pET-23at (+) vSkp1PBCV-1, and transformants were selected by resistance to carbenicillin and chloramphenicol. Batch culture was grown to an OD_{600} of 0.6 at 37°C. Induction of protein overexpression was achieved with the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside for 2 h at 25°C. Cells were harvested by centrifugation (5,000 × g) for 5 min at 4°C and then disrupted by sonication in PBS buffer with 2 mM PMSF for 5 min on ice. Cell lysate was centrifuged for 10 min at 4°C, and the supernatant was applied to a Ni^{2+} column. Column resin and elution buffer were from a Ni-nitrilotriacetic acid (NTA) buffer kit (Novagen Biosciences, San Diego, CA). One milliliter of resuspended Ni-NTA agarose 4 ml of PBS was loaded on a 5-ml self-packing column with a 45- to 90-μm-pore-size polyethylenefiltration (frit) (Life Science Products, Frederick, CO), and the resin was allowed to settle. After the resin was equilibrated with 8 ml of PBS, 8 ml of sample lysate containing overexpressed vSkp1PBCV-1 was loaded on the column. The Ni^{2+} column was washed four consecutive times using 4 ml of PBS buffer with 10 mM imidazole. Three lysates were applied individually to the vSkp1PBCV-1 column: (i) uninfected *C. variabilis* cells, (ii) *C. variabilis* cells at 30 min p.i., and (iii) *C. variabilis* cells at 180 min p.i. Protein was eluted from the column with 0.5 ml of Ni-NTA elution buffer for 5 min. The eluate was collected and analyzed by SDS-PAGE using PAGEGold precast gels (Lonza, Rockland, ME) with Coomassie brilliant blue staining solution. Bands were excised from the gel and analyzed for protein identification by mass spectrometry.

**Protein identification by nano-LC-MS/MS.** The excised protein bands from the SDS-PAGE gel were washed with deionized water and destained by repeated cycles of dehydration (50 mM ammonium bicarbonate in 50:50 acetonitrile) and rehydration in 50 mM ammonium bicarbonate. Proteins were reduced with 10 mM dithiothreitol (DTT) in 50 mM ammonium bicarbonate buffer for 45 min. For alkylation, 40 mM iodacetamide in ammonium bicarbonate was added, and the sample was incubated in the dark for an additional 30 min at room temperature. Following two cycles of dehydration/rehydration as described above, sequencing-grade trypsin in 50 mM ammonium bicarbonate was added to dried gel pieces, which were incubated overnight at 37°C. Peptides were extracted with 100% acetonitrile–0.1% formic acid (FA) and dried using a SpeedVac. The dried mixture of tryptic peptides was dissolved in 5% acetonitrile–0.1% FA and analyzed by nano-liquid chromatography–tandem mass spectrometry (nano-LC-MS/MS) using an UltiMate 3000
MEGA, version 3.1, adapted from Kipreos and Pagano (36) using ClustalW2 (37) running 1-encoded ANK proteins were aligned with the F-box consensus sequence identified by a BLAST search using the National Center for Biotechnology Information. Extracted spectra were searched against NCBI and a custom-based viral database. The data acquisition was done with alternating collision-induced dissociation (CID), high-energy collision dissociation (HCD), and ETD; six MS/MS spectra were obtained for every full scan, and five MS/MS scans were averaged for full scans and MS/MS scans. The normal collision energy was set to 35% for CID, with an isolation width of 3.0 and activation time of 10 ms. Supplemental activation was enabled for ETD; the activation time was set to 100 ms with an isolation width of 3.0. The signal threshold for triggering an MS/MS event was set to 500 counts. Extracted spectra were searched against NCBI and a custom-based viral protein database using the MASCOT search engine.

**Antibody.** Polyclonal rabbit anti-vSkp1PBCV-1 was produced by Harlan Bioproducts for Science, Inc. (Indianapolis, IN). The antiserum was prepared using their standard 73-day protocol by immunizing one rabbit with 500 µg of vSkp1 peptide for three consecutive boosts.

**Western blot analyses.** Samples were separated on SDS-PAGE gels and transferred to polyvinyl difluoride (PVDF) membranes (0.45-µm pore size) using standard procedures (32). Membranes were blocked for 1 h at room temperature with 5% nonfat dry milk in Tris-buffered saline with 0.05% Tween 20 (TBS-T). Blots were probed using primary polyclonal antibody rabbit anti-vSkp1PBCV-1 (Harlan) overnight with agitation at 4°C in blocking buffer. After five washes with TBS-T, membranes were probed with goat anti-rabbit conjugated to horseradish peroxidase (HRP) (A6154; Sigma) as the secondary antibody. After an additional five washes with TBS-T, the membranes were developed using ECL SuperSignal West Pico chemiluminescent substrate (ThermoFisher Scientific). Bands were visualized using a ChemiDoc MP Imaging System with Imagine Lab Software (Bio-Rad, Hercules, CA).

**Protein modeling.** vSkp1PBCV-1 and PBCV-1 A682L structure predictions were constructed by PHYRE (protein homology/analogy recognition engine, version 2.0) (33). Molecular docking of viral proteins was performed using PatchDock (34), a molecular docking algorithm based on shape complementarity principles. Protein complexes were inputted into the molecular modeling system, Chimera, where three-dimensional protein models were constructed. Analysis of conformational ensembles of the Skp1-Skp2 complex (Protein Data Bank [PDB] accession number 1FQV) (35) and the vSkp1PBCV-1-A682L complex was completed using Chimera imagery.

**Sequence identification and alignments.** Gene sequences were identified by a BLAST search using the National Center for Biotechnology Information nonredundant database. The C termini of the eight PBCV-1-encoded ANK proteins were aligned with the F-box consensus sequence adapted from Kipreos and Pagano (36) using ClustalW2 (37) running MEGA, version 3.1, with default settings (Gonnet matrix; gap opening penalty, 10; gap extension penalty, 0.1 for pairwise alignments and 10 and 0.2 for multiple alignments). Known residues of Skp2 that contact cellular Skp1 were reported by Sonnberg et al. (19). Sequence alignment of Skp1 homologs was created by the Phylogeny.fr platform (38, 39) using default settings (MUSCLE, BLOSSUM62 substitution matrix, and Gblocks for alignment curation).

**Phylogenetic analysis.** To construct a robust phylogenetic tree from chlorovirus proteins, sequence alignments were created using the Phylologen.fr platform (38, 39) with default settings (MUSCLE, version 3.7, for multiple alignment; Gblocks, version 0.91b, for alignment curation; PhyML, version 3.0, approximate likelihood ratio test [aLRT], for phylogeny; and TreeDyn, version 198.3, for tree rendering). The following additional parameters were enforced: the minimum number of sequences for a flank position was 85%; all segments with contiguous nonconserved positions bigger than 8 were rejected; minimum block length was 10; no gaps were allowed in the final alignment. Phylogenetic trees were computed using the Whelan and Goldman (WAG) substitution model. We show only bootstrap values of <90%. Branches with bootstrap support of less than 50 were collapsed.

**Identification of proteins used for analysis.** All viral genes investigated in this article correspond to their gene locus tag identifiers in the National Center for Biotechnology Information (NCBI) database. The following are PBCV-1 ANK proteins: A672R (GenBank accession number NP_049028), A005R (GenBank NP_048353), A682L (GenBank NP_049038), A247R (GenBank NP_049038), A607R (GenBank NP_048963), A330R (GenBank NP_048668), A7/8 (GenBank NP_048353), and A429L (GenBank NP_048786). The following vSkp1 proteins were investigated: A039L (PBCV-1; GenBank accession number NP_048387), Z339L (Acanthocystis turfacea chlorovirus-1 [vSkp1ACTV-1]; GenBank YP_001426820), and N799R (Paramecium bursaria chlorovirus FR483 [vSkp1FR483]; GenBank YP_001426431). Skp1-binding ANK proteins from ATCV-1 and FR483 are Z568R (GenBank accession number YP_001427049) and N710R (GenBank YP_001426342), respectively. Skp1 proteins correspond to GenBank accession numbers EFNS4040 and EFNS51326 (C. variabilis), P63208 (Homo sapiens), P52285 (Dictyostelium discoideum), and P52286 (S. cerevisiae). F-box proteins from C. variabilis were GenBank EFNS59680, UniProt E1ZSC7, GenBank EFNS4870, and GenBank EFN51283.

**RESULTS**

Chlorovirus ANK repeat proteins have a possible carboxyl-terminal Skp1 interactional motif. The combination of a substrate-binding domain, often containing protein sequence repeats, and an F-box domain is responsible for the bipartite functionality of cellular F-box proteins. Chlorovirus ANK proteins are composed of similar domain elements. ANK regions, located toward the N-terminal region, are presumed to be substrate binding. Interestingly, alignment of the eight PBCV-1-encoded ANK proteins revealed a common motif in the C-terminal region (Fig. 2). Manual inspection of this region revealed a conserved domain with some key binding residues, including some with Leu-Pro residues characteristic of the 50-amino-acid cellular F-box motif. Although sharing some common residues with the F-box motif, the putative Skp1 interactional domains in the viral ANK proteins are truncated (30 to 40 amino acids) and are positioned at the opposite end of cellular F-boxes. However, they resemble the C-terminal location of the F-box-like PRANC domain of ANK proteins in chordopox viruses (23, 40).

When aligned with a consensus F-box sequence (36) derived from the Pfam F-box profile, the C-terminal regions of the eight PBCV-1-encoded ANK proteins contain a limited number of residues that are present in characterized cellular F-boxes (Fig. 2). Among these shared residues are amino acids known to mediate contact between cellular proteins Skp1 and F-box protein Skp2 (19). Based on domain sequence similarity, PBCV-1 ANK proteins A5R, A682L, and A247R contain a C-terminal motif that most closely resembles a canonical F-box protein. Specifically, these three ANK proteins possess Leu-Pro residues common to most cellular F-box proteins. The viral ANK proteins are highly

December 2014 Volume 88 Number 23 jvi.asm.org

Chlorovirus Skp1 and Ankyrin Repeat Proteins
conserved in the chloroviruses. While their C-terminal sequences are somewhat divergent from those of cellular F-boxes, their locations in the protein are similar to the location of the PRANC domain from poxvirus-encoded ANK proteins (41). We hypothesize that the combination of the ANK domain and conserved C-terminal motif in chlorovirus ANK proteins may function as the substrate-binding and Skp1-binding domains, respectively. This hypothesis was supported by secondary and tertiary structure predictions of the carboxyl-terminal portion of the eight PBCV-1 ANK proteins, which indicated protein folding and exposed active sites for protein interactions similar to those that occur in F-boxes. The predicted structure of the interaction of the C terminus of PBCV-1 ANK protein A682R with the viral Skp1 protein compared with the crystal structure of human Skp1 interacting with Skp2 is reported in Fig. 3 and demonstrates the similarity of the Skp1 and F-box-like box interactions.

Highly conserved ANK repeat protein families among chloroviruses group into monophyletic clusters according to the host alga. Annotated genomes of 41 chloroviruses, which are classified by their ability to infect one of three different symbiotic chlorella-like host algae, contain between 319 and 416 predicted protein-encoding genes per genome (57). Collectively, the 41 chloroviruses encoded 178 ANK-containing proteins. One ANK protein family contains 51 members and, after likely gene duplications are eliminated, all 41 annotated chloroviruses contain at least one member of this protein family. The PBCV-1 family member is A607R. This family of ANK proteins was analyzed by a maximum-likelihood reconstruction method in two ways: (i) only the ANK domains (Fig. 4A) and (ii) only the C-terminal 40 amino acids which contain the putative Skp1-binding domains (Fig. 4B). Sequences were aligned with MUSCLE (version 3.7), and phylogenetic trees were computed using the WAG substitution model. Phylogenetic analysis of the ANK domain from the 51-member ANK protein family revealed three strongly supported clades clustered by the virus host alga. Bootstrap branch supports of 100, 100, and 97 expected-likelihood weights (ELW) provide convincing evidence in support of the monophyly of the NC64A, SAG, and Pbi viruses, respectively (including Pbi virus NE-JV-1; 63 ELW). Surprisingly, the ANK domains from this protein family are well preserved among the chloroviruses as several of these ANK domains have 100% sequence identity (Table 1). These ANK domains are most diverse when viruses that infect different host algae are compared (e.g., the highest degree of sequence identity of the ANK domains between any two virus types is less than 50%).

Analysis of the last 40 residues of the C-terminal region (putative Skp1-binding domain) from the 51-member ANK family again showed three strongly supported clades grouped by virus host alga with equivalent bootstrap branch supports (Fig. 4B). The highest degree of sequence identity of the terminal region between any two virus types was only 35% while several of these C-terminal regions are identical among viruses that infect the same host (Table 1). As expected, the clustering of the C-terminal region from Pbi virus NE-JV-1 was a close outlier, as it was for the ANK domain (42 ELW).

Chloroviruses encode an Skp1. Genome analysis revealed one vSkp1 protein in all sequenced chloroviruses except for NC64A virus KS-1B. Skp1 is a core component of the eukaryotic SCF complex that mediates proteosome-dependent protein degradation. Cellular Skp1 also functions in cell cycle progression, transcriptional regulation, and signal transduction (42). Multiple-sequence alignment of cellular Skp1 homologs, including ones from the host C. variabilis, were aligned with vSkp1 proteins from NC64A, SAG, and Pbi virus members PBCV-1, ATCV-1, and FR483, respectively (Fig. 5). Protein alignment showed 70% conserved amino acid composition among the cellular and viral proteins. Phylogenetic analysis of the 40 vSkp1 proteins revealed three strongly supported clades grouped by the virus host alga (Fig. 6). Bootstrap branch supports of 99, 100, and 99 ELW support the monophyly of the NC64A, SAG, and Pbi viruses, respectively (including Pbi virus NE-JV-1; 97 ELW). While several chloroviruses infecting the same host have identical Skp1 proteins, the highest sequence identity between Skp1 proteins from chloroviruses that infect different hosts is less than 40% (Table 1).

Neither the chlorovirus ANK proteins nor vSkp1 proteins were
associated with the PBCV-1 particle proteome, which contains 148 unique virus-encoded proteins (about 35% of the virus coding capacity) and 1 host protein (43). The genes encoding these virion-associated proteins are dispersed throughout the virus genome, and most are transcribed late or early-late in the infection cycle. In contrast, the PBCV-1 ANK and vSkp1 genes are expressed as early genes (44, 45). Western blotting confirmed the appearance of the vSkp1PBCV-1 protein at 45 min p.i. (Fig. 7).

We performed Y2H analysis to test the hypothesis that PBCV-1 ANK proteins interacted with either cellular or vSkp1PBCV-1. Sequences encoding the ANK proteins and the Skp1 targets were yeast codon optimized and cloned into the pDEST32 bait and pDEST22 prey vectors to allow testing for interactions. The Y2H analyses revealed selective binding interactions between PBCV-1 ANK and Skp1 proteins. The results established that vSkp1PBCV-1 interacted strongly with two of the eight ANK proteins, PBCV-1 A682L and A607R (Fig. 8A). PBCV-1 A607R interacted with one of the two host C. variabilis Skp1 proteins (GenBank accession number EFN54040), which is the more highly conserved Skp1 member. The remaining six PBCV-1 ANK proteins did not interact with any of the Skp1 proteins. The results established that vSkp1PBCV-1 interacted strongly with two of the eight ANK proteins, PBCV-1 A682L and A607R (Fig. 8A). PBCV-1 A607R interacted with one of the two host C. variabilis Skp1 proteins (GenBank accession number EFN54040), which is the more highly conserved Skp1 member. The remaining six PBCV-1 ANK proteins did not interact with any of the Skp1 proteins. To further test the Skp1-binding potential of the ANK proteins A682L and A607R, we assayed these ANK proteins in the bait vector for binding with the cellular Skp1 proteins from human, Dictyostelium discoideum, and S. cerevisiae, as well as vSkp1ATCV-1 and vSkp1FR483 proteins in addition to

FIG 4 Phylogenetic trees of the 51-member ANK repeat protein family coded by chloroviruses. (A) ANK domain of the ANK family proteins. (B) Putative Skp1-binding domain (last 40 residues of the C terminus) of the ANK family proteins. A maximum-likelihood reconstruction method was based on the separate alignments of the ANK and Skp1-binding domains from the ANK family proteins present in 41 chloroviruses that infect three algal hosts (10 gene duplicates were omitted from the analysis). The ANK and Skp1-binding domains are located in the amino and carboxyl termini of the ANK protein, respectively. The following chloroviruses are represented (corresponding geographical sampling locations): NC64A viruses NY-2A (NY, USA), NY-2B (NY, USA), Ar158 (Buenos Aires, Argentina), NY-s-1 (NY, USA), MA-ID (MA, USA), IL-5-2s1 (IL, USA), An-69-C (Canberra, Australia), IL-3A (IL, USA), PBCV-1 (NC, USA), NE-JV-4 (NE, USA), MA-1E (MA, USA), CvsA1 (Sawara, Japan), and GviKI (Kyoto, Japan); SAG viruses Can0610SP (British Columbia, Canada), OR0704.3 (OR, USA), NE-JV-3 (NE, USA), NE-JV-2 (NE, USA), ATCV-1 (Stuttgart, Germany), WD060 (WI, USA), MOO655PH (MO, USA), NTS-1 (NE, USA), Br6064L (Sao Paolo, Brazil), TN603.4.2 (TN, USA), GM0701.1 (Guatemala), MNO810.1 (MN, USA), and Canal-1 (NE, USA); Pbi viruses NE-JV-1 (NE, USA), Mt325 (MT, USA), Can18-4 (Canada), Fr-5-L (France), C-2 (Czech Republic), OR0704.2.2 (OR, USA), CVB-1 (Berlin, Germany), Fr483 (France), NW665.2 (Norway), AP110A (Unknown), CVA-1 (Amônau, Germany), CVG-1 (Göttingen, Germany), CVM-1 (Marburg, Germany), and CVR-1 (Rauschenberg, Germany). Numbers following virus names on the figure correspond to the total number of amino acid residues.

**TABLE 1** Protein sequence similarity of chlorovirus A607R ANK family and vSkp1 proteins

<table>
<thead>
<tr>
<th>Chlorovirus group</th>
<th>ANK domain</th>
<th>ANK protein C terminus</th>
<th>vSkp1 full-length protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC64A</td>
<td>72.7–100</td>
<td>62–100</td>
<td>64.9–100</td>
</tr>
<tr>
<td>SAG</td>
<td>60.7–100</td>
<td>62.5–100</td>
<td>74.5–100</td>
</tr>
<tr>
<td>Pbi</td>
<td>73.2–100</td>
<td>85–100</td>
<td>89–100</td>
</tr>
</tbody>
</table>

Percent protein sequence similarity of Pbi viruses was calculated excluding NE-JV-1.
vSkp1<sub>PBCV-1</sub>. ANK protein PBCV-1 A682L interacted with cellular Skp1 proteins from human and <i>D. discoideum</i> while weakly interacting with vSkp1<sub>ATCV-1</sub> (Fig. 8B). PBCV-1 A682L did not interact with any additional Skp1 homologs. A summary of Y2H interactions involving ANK proteins PBCV-1 A682L and A607R is reported in Table 2.

ANK protein PBCV-1 A682L interacts with Skp1 proteins via a carboxyl Skp1-interactional motif. We constructed truncated versions of ANK protein PBCV-1 A682L by making a series of deletions from the N terminus (Fig. 9A) in order to identify the minimum domain required for binding to the Skp1 proteins. According to ExPASy’s PROSITE motif scan (46), A682L has seven ANK domains, designated by the following residue numbers: 38 to 70, 71 to 103, 104 to 136, 137 to 169, 170 to 202, 203 to 235, and 236 to 268. We constructed six mutants with different combinations of ANK repeats and Skp1-binding region deletions. One mutant contained only the predicted Skp1-binding domain excluding all ANK repeat motifs while another construct contained all the ANK repeats without the C-terminal putative Skp1-binding domain. An additional double point mutation of A682L was constructed by replacing the highly conserved C-terminal dipeptide L335-P336 of the putative Skp1-binding motif with two alanine residues. Y2H screening was performed using the series of truncated A682L mutants with Skp1 proteins (Fig. 9B). The results showed that deletion of the Skp1-binding domain (∆F-box) eliminated binding ability to the PBCV-1, human, and <i>D. discoideum</i> Skp1 proteins. The site-directed A682L mutant (L335A-P336A) did not bind to any Skp1 homologs, further supporting its role as a key binding dipeptide. Similar mutations in yeast F-box proteins also disrupt Skp1 protein interactions in vitro (47). Evidence of binding of the mutant lacking ANK domains 1 to 5 (∆ANKR1–5) demonstrated that at least two ANK domains adjacent to the C-terminal putative Skp1-binding region (approximately the last 40 residues) are sufficient for virus and heterogenous Skp1 protein interaction. Thus, the interactive behavior of A682L requires the equivalent obligatory domains regardless of origin of the Skp1 binding partner.

Chlorovirus-encoded ANK repeat family members and vSkp1 proteins interact exclusively within corresponding monophony clusters. To test the interactive boundaries of the 51-member ANK protein family, we performed additional Y2H trials using ANK family members from the remaining two chlorovirus types established by their algal host: ATCV-1 Z568R (a SAG virus) and FR483 N710R (a Pbi virus). We also cloned vSkp1 proteins from chloroviruses ATCV-1 and FR483. The results indicated positive protein interactions between all vSkp1 proteins tested and their partner ANK family proteins (Fig. 8C). These data confirm the binding of vSkp1<sub>PBCV-1</sub> and PBCV-1 A607R. To examine binding exclusivity, we tested the same ANK proteins against <i>C. variabilis</i>, human, and <i>D. discoideum</i> Skp1 proteins. There was no interaction with human and <i>D. discoideum</i> Skp1 proteins although PBCV-1 protein A607R interacted with <i>C. variabilis</i> Skp1 (GenBank accession number EFN54040). These results validate the specificity that chlorovirus-encoded ANK family proteins interact with their partnered vSkp1 exclusively and one of two Skp1 proteins from the host alga. 

vSkp1<sub>PBCV-1</sub> interacts with endogenous cellular F-box proteins and PBCV-1 ANK repeat proteins. Although the results from the Y2H screen were convincing, it cannot be assumed that an interaction observed with the yeast system also occurs in a functional assay. To address this concern, we used absorption chromatography to test endogenous protein interactions. We constructed a plasmid to overexpress the His-tagged vSkp1<sub>PBCV-1</sub> protein and immobilized it on a Ni<sup>2+</sup> column. Three cell lysates were applied separately to either the viral Skp1-charged resin or an uninfected resin: (i) uninfected <i>C. variabilis</i> cells, (ii) <i>C. variabilis</i> cells at 30 min p.i., and (iii) <i>C. variabilis</i> cells at 180 min p.i. The His-tagged vSkp1<sub>PBCV-1</sub> and the binding partners contained in the lysates were subsequently eluted from the Ni<sup>2+</sup> column. Eluted proteins were resolved by SDS-PAGE with Coomassie blue staining. A subtle band was detected in the viral lysate precipitate that was absent in the uninfected control precipitate, which contained four bands native to the host <i>C. variabilis</i> (Fig. 10). No proteins were detected in the lysates applied to the uncharged resin. The viral lysate-specific band, as well as the four bands from the uninfected host cells, was excised, digested, and analyzed by nano-LC-MS/MS. All detected proteins corresponded to either cellular F-box proteins from <i>C. variabilis</i> or PBCV-1-encoded ANK proteins. A protein of 17 kDa in size was identified as the recombinant vSkp1<sub>PBCV-1</sub> protein in each lysate. In the uninfected lysate,
proteins of 35 kDa, 44 kDa, 67 kDa, and 76 kDa were identified as *C. variabilis* F-box proteins consisting of 331 (GenBank EFN51283), 393 (GenBank EFN54870), 621 (UniProt E1ZSC7), and 739 (GenBank EFN59680) amino acid residues, respectively. In the viral infected cell lysates, one protein below the 49-kDa marker appeared at 30 min p.i. and increased in intensity at 180 min p.i. This band corresponded to PBCV-1 ANK protein A607R. Interestingly, the proteins identified as *C. variabilis* F-box proteins of 67 kDa and 76 kDa decreased in the viral infected cell lysates as the viral ANK protein increased. There are two plausible explanations for this observation: (i) the chlorovirus ANK proteins are competing with cellular F-box proteins for vSkp1 binding, or (ii) the host transcriptional machinery and protein synthesis are downregulated later in infection.

**DISCUSSION**

The widespread prevalence of F-box proteins emphasizes their importance in intracellular regulation, particularly in substrate recruitment to the SCF complex (35, 48–50). They represent almost 2.3% of the protein-coding genes in the model plant *Arabidopsis thaliana* (50, 51), and increasing evidence suggests that plants use the ubiquitin-proteasome system (UPS) to recognize and combat pathogen invasion (7). Most of the characterized F-box proteins are associated with the cellular SCF complex and function as the substrate recognition domain for the SCF complex. Using protein sequence alignment and bioinformatics, we identified eight ANK proteins coded by chlorovirus PBCV-1 which contain N-terminal ankyrin repeats together with possible C-terminal Skp1-binding domains. Previous studies have shown that poxviruses, which have the same domain organization as the chloroviruses, encode ANK proteins. These poxvirus ANK proteins interact with the cellular Skp1 using a C-terminal domain similar to an F-box (9, 19–23). Considering that this domain arrangement does not exist in cellular F-box proteins (40), the conservation of these ANK genes in chloroviruses encouraged us to speculate that PBCV-1 ANK proteins interact with Skp1 and thus play an intimate role in modulating host cell SCF-mediated ubiquitination. Y2H screening and pulldown methods identified protein–protein interactions between Skp1 members and their interacting F-box-like-containing proteins. This approach established that vSkp1<sub>PBCV-1</sub> interacts with PBCV-1 ANK proteins with functional similarity to F-box domains.

As noted above, the chlorovirus Skp1-binding domains, like...
those of the poxviruses, are located at the C terminus of ANK proteins, which is the opposite of the location in cellular F-box proteins (19). Despite the lack of obvious homology of the C-terminal region of PBCV-1 ANK proteins to F-box domains, they share some key residues with human F-box protein Skp2 that contacts cellular Skp1 (Fig. 2). Here, we show that deleting the putative Skp1-binding domain in PBCV-1 A682L resulted in a protein that no longer interacted with its counterpart vSkp1. Serial domain deletions of A682L from the N terminus further support an essential Skp1-binding region required for protein interaction while in tandem with at least two ankyrin repeats. These results suggest that this domain combination is the minimum size required for binding to Skp1 family members. Additionally, a signature dipeptide of cellular F-box proteins, Leu-Pro, located at the beginning of the putative viral Skp1-binding domain, was present in three PBCV-1 ANK proteins. Replacement of this dipeptide with Ala-Ala in A682L prevented protein interaction with vSkp1PBCV-1. However, it is important to note that one of the viral ANK proteins (A607R) lacks the signature Leu-Pro residues, and it binds to vSkp1PBCV-1, even though its C terminus has fewer common residues with a canonical F-box domain.

A comprehensive phylogenetic analysis of PBCV-1 A607R identified a 51-member family of ANK proteins coded by all sequenced chloroviruses. This conserved class of ANK proteins constitutes a newly discovered protein family, with one member present in all 41 annotated chloroviruses. Using multiple-sequence

<table>
<thead>
<tr>
<th>Skp1 protein</th>
<th>A682L</th>
<th>A607R</th>
</tr>
</thead>
<tbody>
<tr>
<td>vSkp1PBCV-1</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>vSkp1ATCV-1</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>vSkp1FR483</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Skp1C. variabilis</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Skp1D. variabilis</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Skp1D. discoideum</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

a, +, proteins Skp1 and ANK interacted; –, proteins Skp1 and ANK did not interact.
b, The origins of Skp1 homologs are identified by subscripts.
PBCV-1-encoded ANK protein A682L deletion mutant constructs interact in a putative Skp1 interactional motif-dependent manner. (A) Five truncated versions of the full-length PBCV-1 A682L were constructed by making a series of deletions from the N terminus (deletion ANK domains 1 to 3 [ΔANKR1–3], ΔANKR1–5, ΔANKR1–6, and ΔANKR1–7). Another mutant incorporated all the ANK domains without the C-terminal putative Skp1-binding domain (ΔF-box). ANK domains 1 to 7 in A682L are located at the following residue numbers: 38 to 70, 71 to 103, 104 to 136, 137 to 169, 170 to 202, 203 to 235, and 236 to 268, respectively. The putative Skp1-binding domain is shown in dark gray, the ANK domain is shown in black, and functional unknown domains are shown in light gray. (B) Growth of yeast strains harboring versions of the A682L protein as prey and Skp1 as bait assayed with X-Gal, together with their -galactosidase activities, are shown. Y2H results indicate that truncated mutants missing the first three and five ANK domains (ΔANKR1–3 and ΔANKR1–5, respectively) retain their ability to interact with the Skp1 protein from PBCV-1 and slightly with Skp1 from human and D. discoideum. Deletion of the Skp1-binding domain (ΔF-box) from the viral ANK protein eliminated binding to all Skp1 members. The site-directed A682L double mutant (A682L-L335A-P336A) did not bind to any Skp1 homolog, further supporting the role of these residues as key binding residues in cellular F-box proteins. Empty vectors were used as negative controls. Values are means ± standard deviations (n = 3).
alignment methods, we showed high sequence identity among these ANK proteins according to the respective virus families. Further phylogenetic analysis of the ANK and putative Skp1-binding domains from this ANK protein family revealed three monophyletic clades separated according to the three host algae, suggesting phylogenetic analysis of the ANK and putative Skp1-binding domains from this ANK protein family revealed three monophyletic clades separated according to the three host algae, suggesting

While the protein sequences of the C-terminal domains of the viral ANK proteins have weak homology to a canonical F-box domain, they do have bona fide F-box-like function based on protein-protein interactions with Skp1, predicted structural similarities to Skp1-canonical F-box interactions, the presence of core conserved Leu-Pro residues in some of the proteins, and a domain organization similar to that of the chordopox NCLDV member that has ANK proteins containing a C-terminal PRANCl domain. The weak sequence homology to the canonical F-box consensus is a significant finding because it reveals a much wider range of primary amino acid sequences that can fold into domains that mediate Skp1 binding than was previously appreciated. Rather than weakening the evidence for the F-box-like function of the chlorovirus ANK proteins, the low sequence homology suggests that Skp1-interacting motifs bearing F-box-like function exist in a wider range of repeat-containing proteins, such as ANK-containing proteins, than previously envisioned. This observation is likely to enhance the discovery of new F-box-like proteins.

In addition to encoding F-box-like ANK proteins, chloroviruses encode a candidate virulence factor, a Skp1 homolog, for intercepting the ubiquitin-proteasome system. While multiple Skp1 homologs exist in cellular organisms (28) including C. variabilis, virus-encoded Skp1 genes have only been reported in algal viruses, with the exception of the recently sequenced pandoraviruses (53). Sharing the domain architecture conserved among Skp1 family members (Fig. 5), chlorovirus vSkp1 homologs contain two separate protein interaction motifs: a POZ (pox virus and zinc finger) domain located in the N-terminal region (1 to 61 amino acids) and an Skp1 domain that resides in the C-terminal region (68 to 145 amino acids). In cellular Skp1 proteins, the POZ domain mediates contact with SCF subunit Cul1 while the Skp1 domain is responsible for binding the F-box-containing proteins. Phylogenetic analyses clearly demonstrate host-based clusters of the conserved cellular homolog, further suggesting specialized functionality that caters to their host-virus interactions. Together, the presence of bipartite domains and monophyly clusters according to virus host suggests that chlorovirus vSkp1 proteins are uniquely expressed to complement a counterpart ANK family protein, thereby mimicking cellular Skp1–F-box binding properties.

As the only chlorovirus that does not encode a vSkp1 protein, NC64A virus KS-1B represents an outlier of the 41 sequenced and annotated chloroviruses. The absence of the vSkp1 in this single example questions the gene’s importance in pathogenesis. Further speculation, four spontaneously derived, large deletion mutants of PBCV-1 suggest that the vSkp1 gene is not an essential gene for virus replication in a laboratory environment, according to the study of Landstein et al. (54). In this later study, one PBCV-1 mutant contained a 37-kb deletion at the left end of its 331-kb genome, which encoded 26 single-copy open reading frames including vSkp1. Interestingly, replication of the deletion mutants is attenuated as their burst sizes were half that of the parent virus burst size, and their plaque sizes were consequently smaller. This reduced virulence and plaque phenotype suggest a potential importance for vSkp1 in chlorovirus replication within the limits of the mutant analysis that included several other genes in the deletion.

Adding to the complexity of pathogenesis, some viruses, including the chloroviruses, encode their own ubiquitin homologs to enable replication and evade cellular immune defenses (55, 56). Depending on their needs, viruses have developed means to enhance or inhibit ubiquitination of specific substrates. In fact, 11 SAG viruses, including ATCV-1, and one NC64A virus, NY-2A, encode their own viral ubiquitin (Z203L and B699L, respectively), sharing up to 97% sequence identity with human ubiquitin. Besides encoding ubiquitin, many viruses encode putative E3 ligases and deubiquitinating enzymes, as well as adaptor proteins. For example, in addition to PBCV-1-encoded protein vSkp1 (A39L), PBCV-1 encodes putative enzymes involved in protein degradation, including a ubiquitin C-terminal hydrolase (A105L) and RING finger E3 ubiquitin ligase (A481L) (43). Together, this ensemble of virus-encoded ubiquitin-related genes creates an arsenal of proteins that threaten the host’s UPS machinery. These viral proteins may engage the ubiquitin system through ubiquitin-binding motifs, another important specificity-providing component of the ubiquitin system. Determining the true physiological substrates of virally encoded enzymes will be challenging, as is determining their effect on viral replication in vivo.
ACKNOWLEDGMENTS
We thank James Gurnon and Irina Agarkova for their technical help and maintaining the culture conditions and David Dunigan and Chris West for helpful discussions.

This research was supported by the National Institutes of Health under Ruth L. Kirschstein National Research Service Award 1 T32 AI065047 (E.A.N.) from the National Institute of Allergy and Infectious Diseases, NSF-EPSCoR grant EPS-1004094 (J.L.V.E.), Stanley Medical Research Institute 11R-0003 (J.L.V.E.), NIH grant P20 RR15635 from the COBRE Program of the National Center for Research Resources (J.L.V.E.), and the Department of Energy (DE-EE0003373) Consortium for Algae Biofuels Commercialization (G.A.O.).

REFERENCES

Chlorovirus Skp1 and Ankyrin Repeat Proteins

December 2014 Volume 88 Number 23
jvi.asm.org 13809


