

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Dissertations and Theses in Biological Sciences

Biological Sciences, School of

---

6-2014

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

Eric Andrew Noel

University of Nebraska-Lincoln, [eric.andrew.noel@gmail.com](mailto:eric.andrew.noel@gmail.com)

Follow this and additional works at: <http://digitalcommons.unl.edu/bioscidiss>



Part of the [Cell Biology Commons](#), [Pathogenic Microbiology Commons](#), and the [Virology Commons](#)

---

Noel, Eric Andrew, "Chlorovirus Skp1 and Core Ankyrin-Repeat Protein Interplay and Mimicry of Cellular Ubiquitin Ligase Machinery" (2014). *Dissertations and Theses in Biological Sciences*. 70.

<http://digitalcommons.unl.edu/bioscidiss/70>

This Article is brought to you for free and open access by the Biological Sciences, School of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Dissertations and Theses in Biological Sciences by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

**CHLOROVIRUS SKP1 AND CORE ANKYRIN-REPEAT PROTEIN  
INTERPLAY AND MIMICRY OF CELLULAR UBIQUITIN LIGASE  
MACHINERY**

by

Eric A. Noel

A THESIS

Presented to the Faculty of  
The Graduate College at the University of Nebraska  
In Partial Fulfillment of Requirements  
For the Degree of Master of Science

Major: Biological Sciences

Under the Supervision of Professor James L. Van Etten

Lincoln, Nebraska

June, 2014

**CHLOROVIRUS SKP1 AND CORE ANKYRIN-REPEAT PROTEIN  
INTERPLAY AND MIMICRY OF CELLULAR UBIQUITIN LIGASE  
MACHINERY**

Eric A. Noel, M.S.

University of Nebraska, 2014

Advisor: James Van Etten

The ubiquitin-proteasome system is a common target of several unrelated viruses that have evolved convergent strategies to redirect host ubiquitin machinery to serve their own needs. Members of the genus *Chlorovirus*, a group of large dsDNA viruses that infect certain freshwater chlorella-like green algae, encode a conserved Skp1 homolog and ankyrin-repeat (ANK) proteins, some of which contain C-terminal domains characteristic of cellular F-boxes or related viral PRANC domains. These observations suggested that this unique combination of chlorovirus proteins either interact with or imitate the key components of the SCF (Skp1-Cul1-F-box) ubiquitin ligases. Using mass spectrometry, we identified two functional classes of ANK proteins from prototype chlorovirus PBCV-1 as binding partners for the virus-encoded Skp1: (i) cellular F-box-like ANK proteins (e.g., ank-A682L); and (ii) core ANK proteins (e.g., ank-A607R). The former ANK class contains a C-terminal F-box-like motif that is recognized by cellular Skp1 proteins from widely divergent species. Yeast two-hybrid analysis confirmed putative F-box-dependent protein interaction of PBCV-1 ank-A682L by constructing serial domain deletions. The other class of viral ANK proteins constitutes a core ANK

family with one member represented in all 41 sequenced chloroviruses. A comprehensive phylogenetic analysis of these core 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 core ANK and Skp1 proteins across three chlorovirus genera. Our results indicate that chloroviruses have evolved complimenting Skp1 and ANK proteins that mimic cellular SCF-associated proteins.

*To my parents, Nick and Cindy Noel, for being my cornerstone.*

## ACKNOWLEDGMENTS

First, it is imperative I express my sincere gratitude to my advisor and friend, Dr. James L. Van Etten. The magnitude of opportunity you provided me as a graduate student cannot be measured or repaid in full value. Your enthusiasm for science is contagious and is only magnified by your warm personality and constant encouragement. It has been my absolute privilege to work beside a world class research scientist, and the Godfather of Chlorella viruses. Thank you for personally investing in me and introducing me to the fascinating world of giant viruses. You are an extraordinary person and an admirable professional I aspire to be.

Secondly, I would like to recognize my co-advisor and friend, Dr. George A. Oyler, who served as the genesis that seeded the foundation of my thesis. I have never encountered a person with such a raw passion for science combined with an unparalleled, creative mind. Tremendously courageous, your ambition to explore the unknown is empowering. You have been a constant resource for advice and strong advocate of mine. As your research technician and subsequent graduate student, it has been a genuine pleasure working with you.

My colleagues in the Van Etten lab offered invaluable input that contributed significantly to my thesis. I cannot thank enough James Gurnon, Cristian Quispe, Dr. David Dunigan, Dr. Irina Agarkova, Dr. Les Lane and Dr. Gary Duncan, for their influential advice, amusement and friendship. Most critically, Dr. Ming Kang from the Oyler lab was the catalyst to my bench top progress as he showed me firsthand several

molecular protocols that were instrumental in shaping my thesis and maturity as an analytical research scientist.

Special thanks to my committee members, Dr. James Van Etten, Dr. George Oyler and Dr. Audrey Atkin for their participation, support and suggestions that greatly improved my research. Thank you, Dr. Charles Wood, Director of the Nebraska Center for Virology, for your dedication to training young virologists and endorsing NCV as a premier research institution.

Finally, I thank my family, a continuous source of love and encouragement. Mom and Dad, thank you for your unwavering support and engraving in me early the values of relentless perseverance. You are the most selfless people I know. As a toddler, you provided me with everything I needed to flourish as a student, athlete, and now as an adult. And to my older brother Nicholas, war hero, thank you for letting me exist as your shadow for all those years.

## **GRANT INFORMATION**

This research was supported by the National Institutes of Health under Ruth L. Kirschstein National Research Service Award 1 T32 AIO60547 from the National Institute of Allergy and Infectious Diseases.



## TABLE OF CONTENTS

---

<b>TITLE PAGE</b> .....	i
<b>ABSTRACT</b> .....	ii
<b>DEDICATION</b> .....	iv
<b>ACKNOWLEDGMENTS</b> .....	v
<b>GRANT INFORMATION</b> .....	vii
<b>TABLE OF CONTENTS</b> .....	viii

### CHAPTER I

#### INTRODUCTION AND BACKGROUND

1. OVERVIEW .....	1
2. THE ALGAL HOST <i>CHLORELLA VARIABILIS</i> SPECIES NC64A .....	4
3. <i>PARAMECIUM BURSARIA</i> CHLORELLA VIRUS -1 .....	6
3.1. Virus structure.....	6
3.2. Genome .....	8
3.3. Protein-encoding genes .....	9
3.4. Chlorovirus gene colinearity .....	11
3.5. Infection cycle.....	14

### CHAPTER II

#### THE UBIQUITIN PROTEASOME SYSTEM

1. OVERVIEW .....	16
2. Components of the SCF Complex .....	20
2.1. Skp1 .....	21
2.2. Cullin-1 .....	22
2.3. F-box proteins .....	23
2.4. Rbx1 .....	24

## CHAPTER III

### CHLOROVIRUS SKP1 AND CORE ANKYRIN-REPEAT PROTEIN

### INTERPLAY AND MIMICRY OF CELLULAR UBIQUITIN LIGASE

### MACHINERY

1. OVERVIEW .....	26
2. MATERIALS AND METHODS.....	31
2.1. Cell Lines.....	31
2.2. cDNA Cloning and Plasmid Constructs.....	31
2.3. Yeast Two-Hybrid Screening.....	32
2.4. Cells and Viral Lysate Preparation.....	33
2.5. Transfection and Pull-down by Metal Chelation Chromatography.....	34
2.6. Western Blot Analyses.....	35
2.7. Antibody.....	36
2.8. Sequence identification and alignments.....	36
2.9. Phylogenetic Analysis.....	36
2.10. Nucleotide sequence accession numbers.....	37
3. RESULTS .....	37
3.1. Chlorovirus ankyrin-repeat protein displays a unique carboxyl terminal F-box-like domain.....	37
3.2. Highly conserved core ankyrin repeat protein family among chloroviruses constitutes monophyletic clusters according to host alga.....	40
3.3. Chloroviruses encode a conserved homolog of cellular Skp1 that resembles the same monophyletic origin as the viral core ankyrin-repeat protein family.....	44
3.4. Cellular Skp1 and vSkp1 <sub>PBCV-1</sub> proteins interact differently with two of eight ankyrin-repeat proteins from PBCV-1.....	48
3.5. PBCV-1 ank-A682L and Skp1 proteins interact in a putative F-box-dependent manner.....	51
3.6. Chlorovirus-encoded core ankyrin-repeat family members and vSkp1 proteins interact exclusively within corresponding monophyly clusters.....	54
3.7. vSkp1 <sub>PBCV-1</sub> interacts with endogenous cellular F-box proteins and PBCV-1 ankyrin-repeat proteins.....	55
4. DISCUSSION.....	57
5. REFERENCES .....	64

## FIGURES AND TABLES

---

### CHAPTER I

<b>Figure 1</b>	Phylogenetic tree of the DNA polymerases of the NCLDV.....	3
<b>Figure 2</b>	<i>Paramecium bursaria</i> and its symbiotic chlorella cells.....	5
<b>Figure 3</b>	The 5-fold-averaged cryoEM structure of PBCV-1.....	7
<b>Figure 4</b>	Dot-plot alignments of ten newly sequenced chlorovirus genomes.....	13
<b>Figure 5</b>	Infection of <i>Chlorella variabilis</i> strain NC64A by PBCV-1.....	15

### CHAPTER II

<b>Figure 1</b>	Ubiquitin–proteasome cascade.....	17
<b>Table 1</b>	Chlorovirus ubiquitin-associated proteins.....	20

### CHAPTER III

<b>Figure 1</b>	The SCF ubiquitin ligase complex.....	27
<b>Figure 2</b>	PBCV-1 ankyrin-repeat proteins contain a carboxyl-terminal F-box-like domain .....	38
<b>Figure 3</b>	Domain organization of PBCV-1 ankyrin-repeat proteins.....	40
<b>Figure 4</b>	Phylogenetic trees of the core ankyrin-repeat protein family encoded by chloroviruses .....	42
<b>Figure 5</b>	Sequence alignment of Skp1 homologs.....	46
<b>Figure 6</b>	Phylogenetic tree of vSkp1 homologs coded by chloroviruses.....	47
<b>Figure 7</b>	Early translation of vSkp1 <sub>PBCV-1</sub> protein .....	48
<b>Figure 8</b>	Interactions between PBCV-1 and cellular SCF-associated proteins detected by Y2H.....	50
<b>Figure 9</b>	PBCV-1 ank-A682L deletion mutant constructs.....	52
<b>Figure 10</b>	PBCV-1 encoded F-box-like protein ank-A682L and Skp1 proteins interact in a putative F-box-dependent manner .....	53

<b>Figure 11</b>	Ni <sup>2+</sup> pull down assay confirms vSkp1 <sub>PBCV-1</sub> interaction with F-box and core ANK proteins .....	56
<b>Table 1</b>	Protein sequence similarity of chlorovirus core ANK and vSkp1 proteins.....	44
<b>Table 2</b>	Summary of Y2H.....	51

## CHAPTER I

### INTRODUCTION AND BACKGROUND

#### 1. OVERVIEW

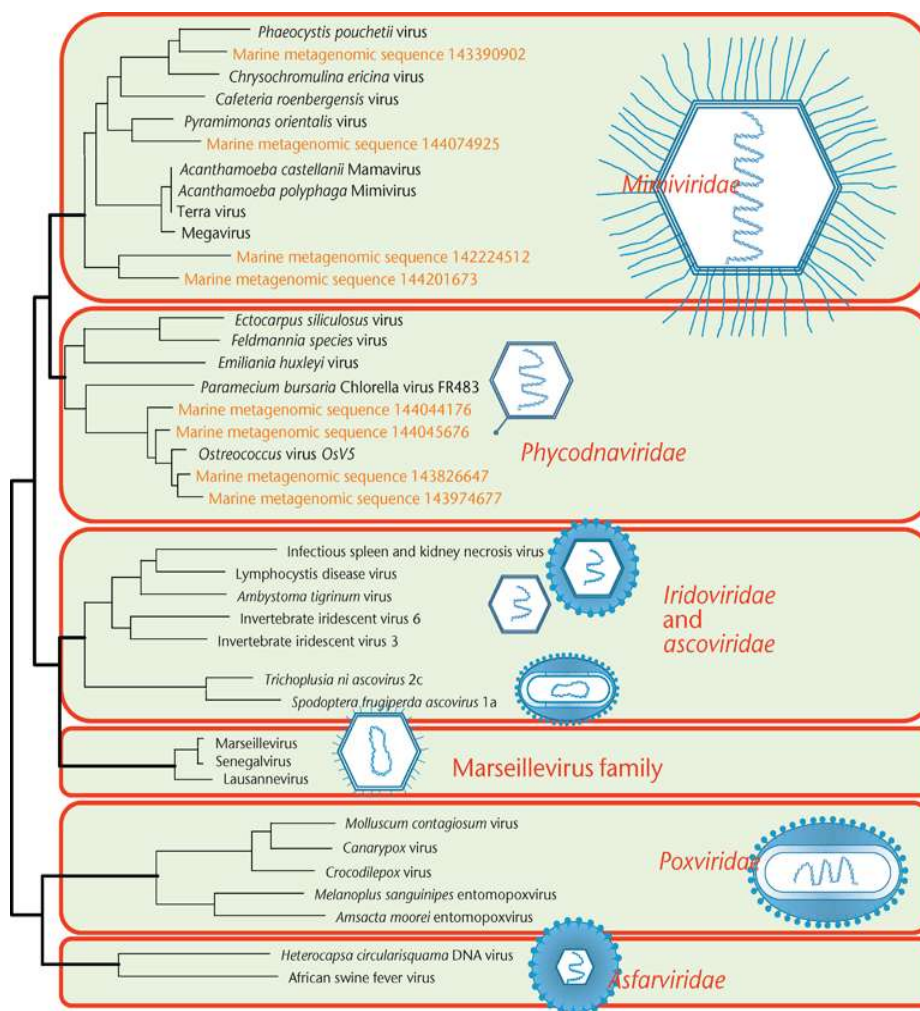
Massive viral genomes are continuously being uncovered that challenge the traditional size limitations of genetic material that can be packaged inside a virus particle. Not only do the genomes of several giant DNA viruses expand far beyond the smallest free-living bacterium (e.g., 2.5-Mb Pandoravirus vs. ~0.5-Mb *Mycoplasma genitalium*) (35, 83), but newly discovered virus genomic content exists unmatched to previously annotated organisms (83). Therefore it is not surprising that these viruses equipped with an arsenal of novel genes have been theorized to constitute the fourth domain of cellular life. Ubiquitous in nature, large DNA viruses exhibit extraordinary diversity in genome structure and composition further encompassing a unique network of virus families. Among these giant viruses with genomes in excess of 300-kb, the *Phycodnaviridae* family represents large icosahedral double stranded DNA (dsDNA) viruses that infect marine or freshwater eukaryotic green algae (114). To date, research efforts have identified phycodnaviruses as key elements in ecology with economic benefit as they regulate algal community structures such as red and brown algal blooms and serve as a source of commercially significant enzymes (115, 120).

Phycodnaviruses are divided into six genera based on their host range: *Chlorovirus*, which includes the proto-type species *Paramecium bursaria* chlorella virus-1 (PBCV-1), *Coccolithovirus*, *Phaeovirus*, *Prasinovirus*, *Prymnesiovirus*, and

*Raphidovirus* with genomes ranging in size from 154-kb to 406-kb (30). While other members of the family exist in marine environments, chloroviruses only infect the freshwater symbiotic green alga *Chlorella*. Often referred to as zoochlorellae, these eukaryotic algae are associated with the *Paramecium bursaria*, the coelenterate *Hydra viridis*, and the heliozoon *Acanthocystis turfacea* (115). Chloroviruses are abundant in freshwater throughout the world with average titers of 1-100 plaque-forming units (PFU) per ml of indigenous water and as high as 100,000 PFU per ml (112). Evolutionary genomic analysis supports the monophyletic lineage of phycodnaviruses with six other large dsDNA virus families: *Poxviridae*, *Asfarviridae*, *Iridoviridae*, *Ascoviridae*, *Mimiviridae* and *Marseillevirus* (Fig. 1) (49, 50). Collectively these viruses are referred to as nucleo-cytoplasmic large dsDNA viruses (NCLDV) (128). These viruses infect animals and diverse unicellular eukaryotes, and replicate either exclusively in the cytoplasm of the host cells, or possess both cytoplasmic and nuclear stages in their life cycle. The genome size of members of this group ranges from 100-kb to 1.2-Mb. A comprehensive phylogenetic analysis of NCLDV evolution identified approximately 50 conserved genes that mapped to the genome of a common ancestral virus while five genes were found to be conserved in all NCLDVs sequenced genomes (127). Interestingly, NCLDVs encode a diverse array of complex machinery spanning prokaryotic, eukaryotic and archaeal gene homologs, which further supports the suggested role of amoeba as the melting pot of NCLDV evolution.

The focus of this opening section serves to introduce the model host and virus system, unicellular alga *Chlorella variabilis* species NC64A (*C. variabilis*) and prototype chlorovirus PBCV-1. Moreover, this introductory chapter provides sufficient

background on the examined backdrop where cellular and viral proteins overlap the ubiquitin proteasome system (UPS). As the subject of this research, virus manipulation of cellular ubiquitin machinery is a widely used and effective means of disrupting cellular homeostasis and favoring permissive infection. In an effort to become more familiar with the UPS, this introduction reviews key enzymes involved in substrate ubiquitination.



**FIG. 1. Phylogenetic tree of the DNA polymerases of the NCLDV.** In addition to known viruses, the tree includes environmental sequences isolated by metagenomics that are denoted with their GenBank identifiers. For each virus family (and two subfamilies in

the case of iridoviruses), a schematic image of the virion (roughly to scale) is shown; the virus membrane is shown in blue. From (128).

## **2. THE ALGAL HOST *CHLORELLA VARIABILIS* SPECIES NC64A**

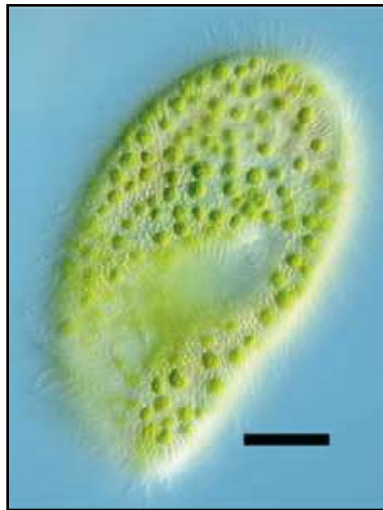
As one of the most abundant and ancient species on the planet, green algae have emerged as significant contributors in global energy reformation and biogeochemical recycling (44). Algae continue to gain worldwide attention regarding their multifaceted function in reducing harmful atmospheric carbon dioxide emissions, producing oxygen and yielding valuable biomass using land and water not suitable for traditional agriculture. Considering their ability to multiply rapidly, varieties of algae have learned to effectively incorporate external sources of fixed organic compounds in addition to gaining energy through photosynthesis to boost growth. Given algae require only light, simple salts, and a nitrogen source for growth, culture maintenance in the laboratory is relatively inexpensive with limited labor requirements. Algae form an ensemble of diverse photosynthetic organisms, ranging from multicellular giant kelp to single-celled genera such as *Chlorella*.

Members in The genus *Chlorella* (Phylum Chlorophyta) are small (~2 to 10  $\mu\text{m}$  in diameter), coccoid, nonmotile, unicellular green algae that exist as the most widely distributed algae in freshwater throughout the world (9). While most *Chlorella* species are free-living, *C. variabilis* exists as an endosymbiont of ciliated protozoan *P. bursaria* in nature (Fig. 2). In fact, *P. bursaria* cells are rarely isolated from natural environments without the green alga residing in its cytoplasm. By means not completely understood, the engulfed microalga establishes residence inside a protective perialgal vacuole within



*P. bursaria* preventing digestion by the protist (55). Fortunately, *C. variabilis* can be cultured independently of *P. bursaria* in the laboratory without compromising alga replication. When the alga is not enclosed by the protozoan, it becomes susceptible to viral infection.

The *C. variabilis* 46.2-Mb genome has recently been sequenced (9) and predicted to contain 9,791 protein-encoding genes (CDS) with the highest average GC content (67.2%) reported in sequenced eukaryotic genomes at this time. Interestingly, these included several genes encoding meiosis-specific proteins suggesting that *C. variabilis* may have diploid and putative cryptic sexual reproductive phases. Furthermore, gene annotation reveals signs of adaptation to endosymbiosis with respect to cellular proteins dedicated to protein-protein interaction and cell wall metabolism (43). The availability of both the host and virus sequences makes chloroviruses an attractive model system.



**FIG. 2. *Paramecium bursaria* and its symbiotic chlorella cells.** Chlorella resides in perialgal vacuoles inside single-celled protozoan *P. bursaria*. While enclosed by the protozoan, the alga is protected from virus infection (size bar: 20 $\mu$ m). Figure from (114).

### 3. *PARAMECIUM BURSARIA* CHLORELLA VIRUS -1

PBCV-1 is the prototype member of the chloroviruses (*Phycodnaviridae* family), a group of large, icosahedral, plaque-forming, dsDNA viruses that infect freshwater chlorella-like green algae. PBCV-1 infects the unicellular green alga *C. variabilis*, a natural endosymbiont of the protozoa *Paramecium bursaria*. The PBCV-1 host alga normally exists as a symbiont in the *Paramecium*. When the alga is in the *Paramecium*, it is resistant to PBCV-1 infection. While the role of chloroviruses in freshwater water ecology remains elusive, these viruses are found throughout the world with seasonal fluctuating titers (123). Considering susceptible hosts lyse within 6 to 16 h in the laboratory, and burst sizes typically exceed 100 PFU per cell, the influence of virus turnover in nature is potentially driving microbial diversity and evolution (117).

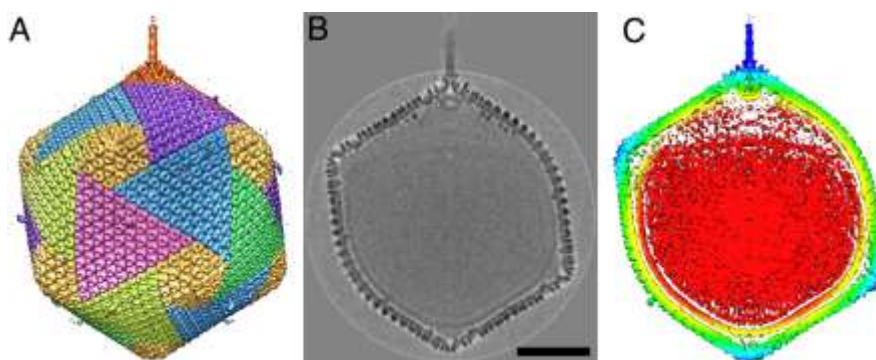
#### 3.1. Virus structure

Novel in structure, the architecture of PBCV-1 includes an outer glycoprotein capsid that surrounds a lipid bilayer membrane and a dsDNA genome (17). Vp54, a glycoprotein and the major component of the capsid, represents about 40% by weight of the total protein content in the virus. The crystal structure of Vp54 showed that the subunit consists of two 8-stranded, antiparallel  $\beta$ -barrel, “jelly-roll” domains. The monomeric Vp54 forms a trimeric capsomer (a “major capsomer”), which has pseudo-6-fold symmetry, relating to the jelly-roll structures (17).

Using cryoelectron microscopy (cryoEM) and assuming icosahedral symmetry (125), the virion has a diameter ranging from 1.650 Å, measured along the 2-fold and 3-fold axes, and 1,900 Å, measured along the 5-fold axes (17). The glycoprotein shell is

composed of 20 triangular units, termed trisymmetrons, and 12 pentagonal caps, termed pentasymmetrons at the 5-fold vertices. Each trisymmetron and pentasymmetron consists of a pseudo-hexagonal array of 66 and 30 trimeric capsomers, respectively. PBCV-1 has a triangulation number ( $T$ ) of 169d quasi-equivalent lattice (115).

Further 5-fold symmetry, 3D reconstruction analyses revealed that the icosahedral symmetry of PBCV-1 is compromised by a unique vertex containing a spike structure. The external portion of the cylindrical spike structure is 340-Å-long and the part of the spike structure that is outside the capsid has an external diameter of about 35 Å at the tip expanding to about 70 Å at the base (Fig. 3) (129). The unique vertex also contains a pocket on the inside adjacent to the cylindrical spike structure that is predicted to house enzymes involved in the initial stages of infection (e.g., cell wall degrading enzymes). Consequently, the virus DNA located inside the envelope is packaged nonuniformly in the particle. Although the function of the spike structure is unknown, it is predicted to aid in puncturing the host cell wall, similar to the ones described in several bacteriophages (80). The spike is too narrow to deliver DNA and so it must be moved aside during DNA release into the cell.



**FIG. 3. The 5-fold-averaged cryoEM structure of PBCV-1.** (A) Hexagonal arrays of major capsomers form trisymmetrons and pentasymmetrons (yellow). The unique vertex with its spike structure is at the top. Capsomers in neighboring trisymmetrons are related by a 60° rotation, giving rise to the boundary between trisymmetrons. (B) Central cross-section of the cryoEM density. (Scale bar: 500 Å.) (C) The same view as in B but colored radially, with red density being within 680 Å, yellow between 680 and 745 Å, green between 745 and 810 Å, light blue between 810 and 880 Å, and dark blue greater than 880 Å. Note the typical lipid low-density gap surrounding the red nucleocapsid density. Figure from (17).

### 3.2. Genome

The PBCV-1 genome is a linear, ~334-kb non-permuted dsDNA molecule with a GC content of approximately 40% (115). The ends of the PBCV-1 genome are hairpin loop-like structures consisting of a 35 nucleotide-long region that is covalently closed and unevenly base-paired (88). The 35-nucleotides at the two ends are complementary when the sequences are inverted. Adjacent to the hairpin ends are identical ~2.2-kb repeats that possibly house the DNA replication initiation site (122). The remainder of the PBCV-1 genome contains primarily single-copy DNA. PBCV-1 is predicted to encode approximately 410 proteins using 40 codons as the minimal length for a protein, and 11 tRNA genes. 148 virus-encoded proteins and one host-encoded protein are present in the mature virion (126).

Most chlorovirus genomes contain methylated bases. For example, genomes from 37 sampled chloroviruses have 5- methylcytosine (5 mC) in amounts ranging from 0.12

to 47.5% of the total cytosine. In addition, 24 of the 37 viral DNAs contain N6-methyladenine (6 mA) in amounts ranging from 1.5 to 37% of the total adenine (77, 78). The methylated bases occur in specific DNA sequences, which led to the discovery that the chlorella viruses encode multiple 5mC and 6mA DNA methyltransferases. About 25% of the virus-encoded DNA methyltransferases have a companion DNA site-specific (restriction) endonuclease, including some with unique cleavage specificities (116).

### 3.3. Protein-encoding genes

Initial sequencing of the PBCV-1 genome in 1995-1997 revealed 697 potential open reading frames (ORFs), of which 373 were predicted to be CDSs (59, 63-66). A CDS is defined as a gene sequence flanked by start and stop codons that presumably codes for a protein. The PBCV-1 genome was recently re-sequenced to correct errors (29). PBCV-1 is predicted to have 416 CDSs. The following criteria were used to define a CDS: (i) minimal size of 40 codons or more initiated by an ATG site; (ii) the largest CDS was chosen when CDSs overlapped; (iii) the putative promoter region (50 nucleotides upstream of the start codon) is AT-rich (>70%). The total number of CDSs from PBCV-1 has changed since the first publication describing the sequence because of sequencing mistakes. The PBCV-1 CDS density is 1.11 CDS/kb, which is similar to that reported for other large DNA viruses (e.g., *Ostreococcus tauri* virus-5 is 1.45 CDS/kb and *Feldmania* sp. virus-1 is 0.968 CDS/kb) (26, 94).

PBCV-1 CDSs are evenly distributed on both DNA strands of the virus genome with minimal intergenic spaces (>100 nucleotides). Exceptions to this rule include a 1788-nucleotide sequence in the middle of the PBCV-1 genome that encodes 11 tRNA

genes, co-transcribed as a large precursor and then processed to mature tRNAs (124). In addition, this area has 900 nucleotides that are GC rich (>60%) and the observed-to-expected ratio of CpG dinucleotides is 1.45, characteristics of a CpG island. The ~2.2-kb inverted terminal repeat region contains 4 duplicated CDSs (103).

Of the 416 predicted PBCV-1 CDSs, ~40% resemble proteins of known function, including many that are novel for a virus. In fact, the majority of the predicted CDSs in giant viruses similar to PBCV-1 do not match anything in gene databases, indicating these viruses are a rich source of novel biochemical functions yet to be discovered (115).

Chloroviruses are unusual because they encode enzymes involved in sugar metabolism. Two PBCV-1 encoded enzymes synthesize GDP-L-fucose from GDP-D-mannose, GDP-D-mannose dehydratase (A118R), and fucose synthase (A295L) (36, 109) and three enzymes, glucosamine synthetase (A100R), UDP-glucose dehydrogenase (A609L), and hyaluronan synthase (A98R), contribute to the synthesis of hyaluronan, a linear polysaccharide composed of alternating  $\beta$ -1,4-glucuronic acid and  $\beta$ -1,3-N acetylglucosamine residues (25, 61). Hyaluronan synthesis occurs at the plasma membrane and is simultaneously extruded through the membrane to the exterior of the cell (25). More unusual CDSs from PBCV-1 include six enzymes involved in lipid metabolism and another five enzymes involved in polyamine metabolism.

CDSs from PBCV-1 can be classified into a cluster of orthologous groups (COG) according to their putative functions. The CDSs resemble genes from all three domains of life, with the majority being bacterial-like. Prokaryotic CDSs tend to be clustered in islands towards the extremity of the virus genome and are co-localized with bacterial-like insertion sequences (32). Interestingly, a few PBCV-1 proteins revealed unusual

characteristics. For example, some proteins are bifunctional (e.g., dCMP deaminase can also function as a dCTP deaminase) (130); other virus-encoded enzymes have extremely high activity (e.g., virus topoisomerase II activity is 30x faster than human topoisomerase II) (34, 62); and a few chlorovirus proteins were the first characterized from a virus [e.g. K<sup>+</sup> ion channel protein (84) and translation elongation factor enzyme (EF-3) (121)].

Cluster analysis of PBCV-1 genes based on deep RNA sequencing was performed to classify genes into groups based on the similarity of their expression profiles (10). The clusters essentially defined two broad temporal mRNA abundance patterns: i) genes whose transcripts peaked before 60 min p.i. and began to decrease and ii) genes with increasing transcript accumulation throughout the first hour of infection. Overall early gene clusters contained 28% (47/166) of the virion associated genes vs. 40% (100/249) for late gene clusters. By 20 min post infection (p.i.), transcripts are detected for most PBCV-1 genes while transcript levels continue to increase globally up to 60 min p.i., at which time 41% of the poly (A<sup>+</sup>)-containing RNAs in the infected cells map to the PBCV-1 genome (89).

### **3.4. Chlorovirus gene colinearity**

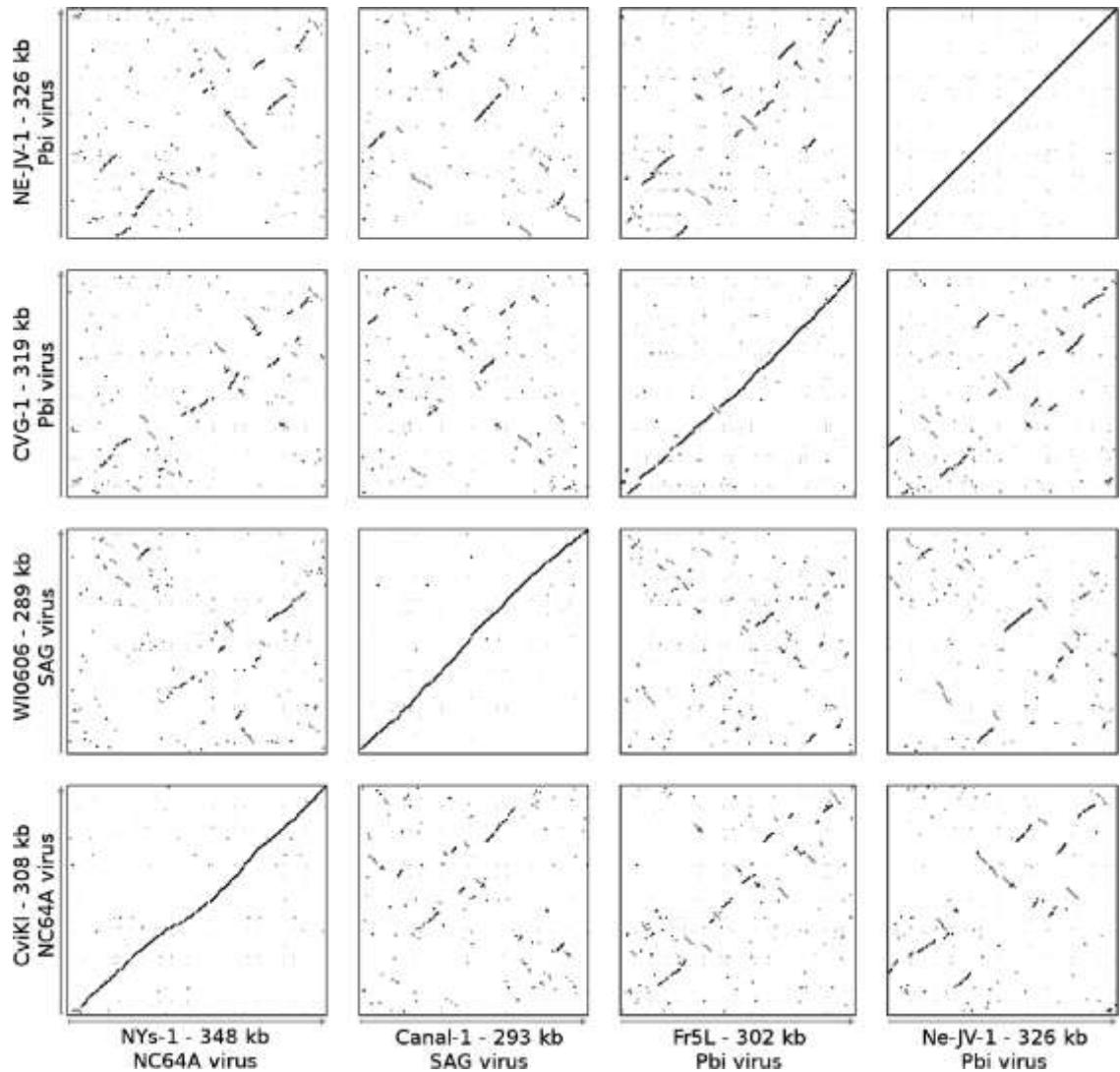
In total, 41 chloroviruses have been sequenced and annotated including 14 viruses that infect the same host as PBCV-1, *C. variabilis*, 14 viruses that infect *M. conductrix*, and 13 viruses that infect *C. heliozoae* (NC64A-, Pbi-, and SAG viruses, respectively). The 41 chloroviruses share in common a core protein family set that consists of 155 protein families which represents 56% of the average protein family content of chloroviruses; the majority (66%) of those proteins have an annotated function (51).

However, the sum total of chlorovirus-encoded genes is much larger than those present in any one virus. Not surprisingly, homologs from viruses infecting the same host are the most similar. For example, the average amino acid identity between homologs from PBCV-1 and other NC64A viruses NY-2A or AR158 is ~73%. In contrast, PBCV-1 and Pbi-viruses MT325 or FR483 orthologs have ~50% amino acid identity while PBCV-1 and SAG-virus ATCV-1 orthologs have ~49% amino acid identity (115). Using PBCV-1 as a model, there is good synteny among all viruses that infect *C. variabilis* whereas PBCV-1 has only slight synteny with the 14 Pbi viruses and the 13 SAG viruses (33).

It is clear that gene order is highly conserved among viruses infecting the same algal host, with only a few readily identifiable localized rearrangements (Fig. 4). The high conservation of gene order contrasts strongly with the low residual gene colinearity between genomes from viruses infecting different algal hosts. The largest conserved genomic regions between chloroviruses infecting different hosts encompass 32 colinear genes. This observation is consistent with the reported high level of gene colinearity between the genomes of NC64A viruses PBCV-1 and NY-2A, as well as between those of Pbi viruses MT325 and FR483, but not between NC64A- and Pbi viruses. The only exception to this genomic trend is Pbi virus, NE-JV-1; its gene order is different from that of other Pbi infecting viruses. This lack of gene colinearity is consistent with the basal phylogenetic position of NE-JV-1 within the Pbi virus clade (51). NE-JV-1 also has no long-range conserved gene colinearity with NC64A viruses or SAG viruses. In contrast, although the NC64A viruses also form two separate phylogenetic sub-groups (one sub-group contains PBCV-1 and the other NY-2A) genomes from both sub-groups share



almost perfect gene colinearity as exemplified by the dot-plot comparison between CviKI (PBCV-1 sub-group) and NYs-1 (NY-2A sub-group) (Fig. 4).

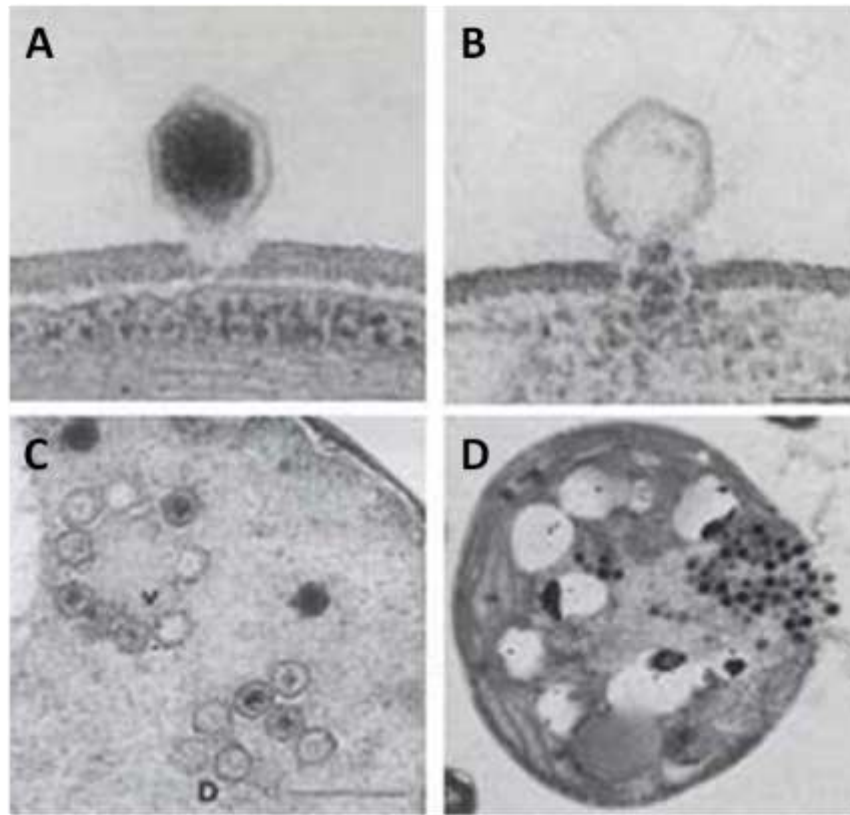


**FIG. 4. Dot-plot alignments of ten newly sequenced chlorovirus genomes.** Each dot represents a protein match between two viruses (BLASTP e-value < 1e-5) from genes in the same orientation (black) or in reverse orientation (gray). Best BLAST matches are shown with larger dots. Figure from (51).

### 3.5. Infection cycle

Virus attachment to the cell wall occurs rapidly and probably involves hair-like external fibers that extend from some of the trisymmetron capsomers. To initiate infection, PBCV-1 uses viral-encoded enzymes to digest the cell wall around the point of attachment (Fig 5A). Following digestion of the alga cell wall, the internal viral membrane fuses with the host membrane facilitating entry of the viral DNA and virion-associated proteins into the cell (Fig 5B), leaving an empty virus capsid on the cell surface (72). This fusion process initiates rapid depolarization of the host membrane and the rapid release of  $K^+$  from the cell (1). The rapid loss of  $K^+$  and associated water fluxes from the host reduce its turgor pressure, which may aid ejection of viral DNA and virion-associated proteins into the host. Depolarization may also prevent infection by a second virus (42).

PBCV-1 lacks a gene encoding a recognizable RNA polymerase or a subunit of it and RNA polymerase activity is not detected in PBCV-1 virions. Therefore, viral DNA and virion-associated proteins are predicted to migrate to the nucleus and early viral transcription is detected 5-10 min post-infection (p.i.), presumably by hijacking a host RNA polymerase(s) (possibly RNA polymerase II) (53, 96). Host chromosome degradation caused by restriction endonucleases packaged in the PBCV-1 virion occurs within 2-5 min p.i. (1). Virus DNA replication begins between 60-90 min p.i. followed by transcription of late genes. At 240 min, PBCV-1 infected cells are covered with hyaluronan (41). Virus assembly occurs at 3-4 h p.i. in localized regions of the cytoplasm called virus assembly centers (Fig 5C). Finally, at 6-8 h p.i., host cell lysis occurs and progeny virions are released (Fig. 5D). Each cell releases ~1000 particles, of which ~30% are infectious.



**FIG. 5. Infection of *Chlorella variabilis* strain NC64A by PBCV-1. A.** Digestion of the wall at the point of attachment (1-3 min p.i.). **B.** The contents of the virus are released into the cell (3-5 min p.i). **C.** Cytoplasmic virus assembly centers at 3-5 h p.i. **D.** Release of progeny virus by lysis of host cell at 6-8 h p.i. Figure adjusted from (116).

## CHAPTER II

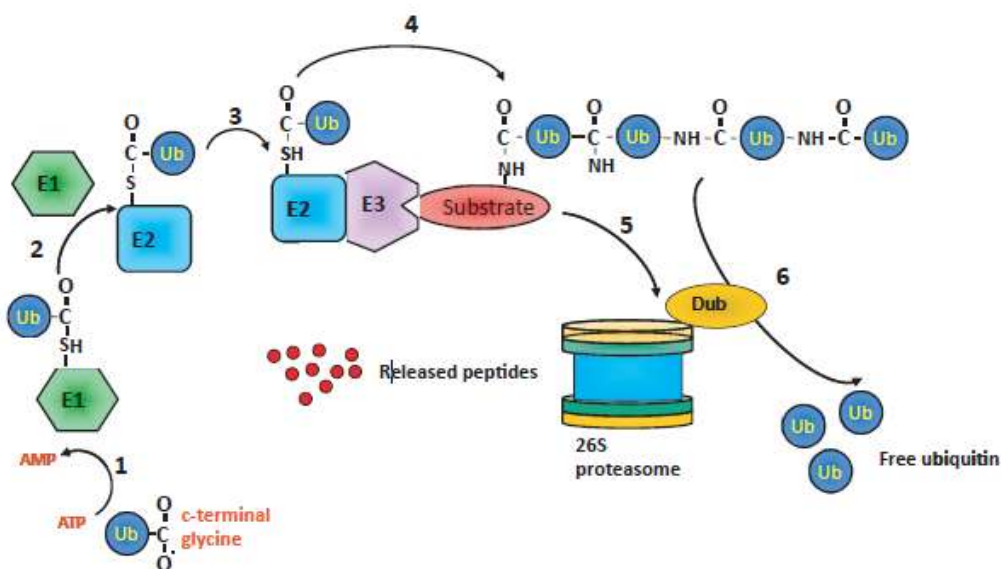
### THE UBIQUITIN PROTEASOME SYSTEM

#### 1. OVERVIEW

Discovery of another post-translational modification, protein ubiquitylation, is leading to a radical revision of our understanding of cell physiology. It is becoming increasingly more evident that protein ubiquitylation is as important as protein phosphorylation in regulating cellular activities. One consequence of protein ubiquitylation is protein degradation by the 26S proteasome. The Ubiquitin Proteasome System (UPS) mediates polyubiquitination and degradation of proteins in eukaryotic cells and thereby controls several biological processes via protein turnover. Therefore its role is imperative in a variety of basic cellular pathways during both cell life and cell death, and hence in both health and disease (18). The ubiquitin system is now known to be involved in basic biological processes, such as the control of cell division, signal transduction, regulation of transcription, DNA repair, quality control in the endoplasmic reticulum, stress response, induction of immune response and inflammation, apoptosis, embryonic development, and circadian clocks, to mention but a few (47). Considering the number of protein substrates and pathways involved, it is not surprising that UPS alterations underlie the pathogenesis of several diseases, including neurodegeneration (Parkinson's, Alzheimer's and Huntington's diseases) and various malignancies (19). However, despite the current model of ubiquitin-mediated proteolysis, much of the ubiquitin system remains unknown. We are only beginning to understand the full

physiological consequences of covalent modifications of proteins, not only ubiquitin, but also ubiquitin-related proteins.

UPS activity involves a three-step enzymatic cascade that results in the covalent attachment of ubiquitin to its target protein (Fig. 1). The ubiquitin-activating enzyme E1 carries out an ATP-dependent activation of the carboxyl-terminal glycine residue of ubiquitin by the formation of ubiquitin adenylate, followed by the transfer of activated ubiquitin to a thiol site of E1 with the formation of a thioester linkage (58). Activated ubiquitin is next transferred to a thiol site of ubiquitin-conjugating enzyme E2 by transacylation, and is then further transferred to an amino group of the protein substrate in a reaction that requires the ubiquitin-protein ligase E3 for target protein recognition (67).



**FIG. 1. Ubiquitin–proteasome cascade.** (1) ATP-dependent activation of ubiquitin (Ub) by E1 (ubiquitin activating enzyme). Activated Ub binds to a conserved cysteine in E1 via a thioester linkage from a carboxyl group in its terminal glycine. (2) Transfer of Ub to

ubiquitin conjugating enzyme (E2) forming an E2-Ub thiolester linkage. (3) The E2 carries the activated ubiquitin to the ubiquitin ligase (E3) which facilitates the transfer of the ubiquitin from the E2 to a lysine residue in the target protein, often by forming an intermediate complex with the E3 and the target. (4) Initial ubiquitination of target forming an Ub–protein conjugate linked by an isopeptide bond. Additional Ubs are ligated to form poly-Ub chains. (5) Proteins tagged with K48 linked poly-Ub chains are targeted to the proteasome. (6) Poly-Ub chains are disassembled by a proteasome associated deubiquitinating activity (DUB) and free Ub moieties are released. Proteasome-localized substrates are then unfolded, imported and degraded into peptide fragments. Figure from (91).

In this sequential reaction, the E3 ubiquitin ligases determine substrate specificity by selectively recruiting target proteins to the ubiquitin machinery (85). Many ubiquitin-interfering viral proteins act on cellular pathways that rely on the SCF (Skp1-Cullin-1-F-box) system, an E3 ubiquitin ligase (11). Among the E3 families, the SCF complex is among the best-characterized and most diverse class of E3 ligases. The SCF complex is comprised of S-phase kinase-associated protein 1 (Skp1), cullin-1 (Cul1), F-box protein and RING-box1 (Rbx1). Several recent studies have reported a number of proteins encoded by both animal and plant viruses that redirect host ubiquitination components to serve their own needs (40, 68, 87, 105). This is accomplished by a diverse group of pathogens that have evolved virulence factors that target and manipulate the function of cellular machinery by mimicking host cell functions (31). This molecular mimicry is indispensable for pathogens to exploit or subvert the host cellular processes for efficient

infection. With a total of 670 ubiquitin-associated proteins in the human genome (60–70 predicted members of the ubiquitin-specific protease family), it is almost certain that many more diseases will be found to arise from genetic errors in the UPS or by pathogen subversion of the system (69, 81).

Chloroviruses encode their own group of ubiquitin-associated enzymes contributing to the complexity of pathogenesis (Table 1). Transcriptome analysis of PBCV-1 (126) revealed early transcribed enzymes putatively involved in protein degradation including ubiquitin C-terminal hydrolase (A105L) and SCF-E3 ubiquitin ligase (A481L). Another NC64A virus, NY-2A, and 11 SAG viruses encode an ubiquitin-60S ribosomal protein (Z203L). 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. Considering its host's ubiquitin machinery encompasses 139 ubiquitin-associated protein-encoding genes, many opportunities for exploitation by the virus exist. Determining the true physiological substrates of virally encoded enzymes can be challenging, as is determining their effect on viral replication *in vivo*. While this chapter serves as an introduction to key enzymes and their interactive-domains involved in the UPS pathway, the focus of Chapter III will concentrate on similar chlorovirus-encoded proteins associated with the cellular ubiquitin ligase enzyme family and members of the SCF complex.

**TABLE 1 Chlorovirus ubiquitin-associated proteins**

<b>Protein</b>	<b>PBCV-1</b>	<b>NC64A viruses</b>	<b>SAG viruses</b>	<b>Pbi viruses</b>
Ubiquitin-60S ribosomal protein L40	0	1	11	0
Ubiquitin carboxyl-terminal hydrolase	1	15	13	14
Ubiquitin-protein ligase	0	6	1	2
Ankyrin-repeat	8	100	29	74
Skp1	1	13	13	14

## 2. Components of the SCF Complex

The essential components of the SCF ubiquitin E3 ligase include Skp1, Cul1, one of many F-box proteins, and the RING-H2-finger protein Rbx1. Thus the SCF complex contains three invariable components (Skp1, Cul1 and Rbx1) which provide a core structure to which one of the many substrate-specific subunits (F-box proteins) binds. The F-box protein itself is modular in structure with one domain of the protein containing a canonical F-box motif which mediates the binding to Skp1 and the other domain mediating substrate recognition. The substrate recognition domain frequently is formed from repeats such as WD40, LRR, and ankyrin repeats. The Skp1-Cul1-Rbx1 core also independently interacts with the ubiquitin E2-conjugating enzyme to couple ubiquitin transfer to the substrate (45). One of the F-box proteins binds to a specific substrate and such interaction facilitates the polyubiquitination of the substrate by ubiquitin transfer from the ubiquitin-charged E2. Since there are multiple F-box proteins, this mechanism illustrates how the same core complex can control a diverse spectrum of substrates. Related E3 ligases built with a conserved cullin core protein employ similar strategies to expand the substrate specificity (133).



## 2.1. Skp1

Skp1 (S-phase kinase associated protein 1) serves as an adaptor protein that provides a molecular link between Cul1/Rbx1 and the F-box proteins. The Skp1 protein contains two separate protein-interaction domains that are conserved among its family members between species. The N-terminal region of Skp1 (~1-70 a.a.) interacts with Cul1 while the C-terminal half (~100-163 a.a.) binds the F-box proteins (132). The use of Skp1 as an adaptor to link the core ubiquitin E3 ligase components of Cul1/Rbx1 with numerous and diverse substrate-targeting subunits, the F-box proteins, is a strategy to specifically target many proteins for ubiquitination. The role of Skp1 is to bring the substrate-targeting subunit, the F-box protein into proximity with the Cul1/Rbx1/E2 complex to promote ubiquitin transfer from the E2-ubiquitin to the F-box protein-bound substrates (20). In addition to targeting substrate proteins for degradation, Skp1 has been associated with certain activities that remain to be further characterized. For example, yeast Skp1 also binds to the kinetochore Cbf3 complex and is essential for yeast kinetochore/centromere function in G2 (6). However, the precise roles of Skp1 in these biological processes still remain unclear.

The crystal structure of Skp1 has been resolved. Skp1 contains a BTB/POZ-like domain in the N-terminal region (70). The BTB/POZ domain mediates the direct binding of Skp1 proteins to Cul1 and in another class of BTB/POZ proteins to Cul3. Thus the BTB/POZ domain appears to be a Cullin interaction motif. The structure of Skp1 also confirmed its similarity with Elongin C, which also has a BTB-like motif which interacts with Cul2. Both Skp1 and Elongin C also share structural resemblance to the potassium channel tetramerization domain, which also belongs to the BTB/POZ superfamily (70).

Thus the BTB/POZ-like structures determine the interaction between cullins and their adaptor proteins such as Skp1, Elongin C, and other BTB/POZ-containing proteins. Interestingly, although in mammals and single-cell organisms such as yeast there is only one conserved Skp1 homolog, other multi-cellular organisms encode multiple Skp1-like proteins (52). In *Arabidopsis thaliana*, 21 Skp1-related genes, called Arabidopsis-Skp1-like (ASK) are predicted and genetic studies suggest that ASK1 is part of SCF complexes that regulate the responses to the plant hormones auxin and jasmonate (131). In addition, six Skp1 homologs have been identified in *Drosophila melanogaster* and at least 21 Skp1-related proteins have been identified in *C. elegans* (75, 76). The presence of such large families of Skp1-related proteins in these organisms suggest that selective expression of these Skp1-related proteins during development or in a particular tissue may represent an additional level of regulation for their protein substrates. Modified forms of Skp1 also exist, for example, in the social amoeba Dictyostelium. Skp1 activation in the protist requires hydroxylation and subsequent glycosylation in regulating environmental oxygen which drives cyst formation and sporulation. Therefore the post-translational modification of Skp1 differentially influences sequential developmental transitions via polyubiquitination and degradation of F-box proteins and their respective regulatory factor substrates.

## **2.2. Cullin-1**

In the SCF complex, Cul1 forms the core scaffold that associates with Rbx1 at the extreme C-terminal region. The amino terminal region of Cul1 interacts with Skp1. The Rbx1-binding domain exhibits the highest conservation among cullins and was initially

identified as the cullin homolog domain (4). The conservation of this domain is consistent with the observation that Rbx1 can bind to almost all cullins (Cul1-5 and Cul7) and this binding couples cullin E3 ligases to the ubiquitin E2-conjugating enzymes (8). The N-terminal region for Skp1 binding is conserved between orthologs of Cul1 from different species, and also displays significant homology in the equivalent regions among other cullins such as Cul2, Cul3, Cul4 and Cul5 (24). In Cul2 and Cul5 this region has been shown to interact with Elongin C, a protein that shares considerable homology with Skp1 (132). Further, the N-terminal region of Cul3 has been shown to interact with BTB/POZ proteins that display a similar three-dimensional crystal structure to that of Skp1 (70).

### **2.3. F-box proteins**

F-box proteins serve as the substrate-targeting subunit of the SCF ubiquitin E3 ligase. They are structurally diverse however they all contain a relative conserved signature motif of about 45-50 amino acids. This motif, the F-box, was initially identified among human Skp2, yeast Cdc4, and human cyclin F, which all bind to Skp1 (21). F-box proteins also contain a separate protein-protein interaction domain that mediates the binding to various substrates. The binding of F-box proteins to their selected substrates usually targets the respective substrate for polyubiquitination and subsequent proteolysis through the 26S proteasome. However, F-box proteins can also mediate the processing of certain protein precursors to their cleavage products through limited ubiquitin-dependent proteolysis (22).

The existence of a large repertoire of F-box proteins means that SCF E3 ligase is one of the largest E3 ligase families. The number of reported F-box proteins is diverse

among organisms: human (69), yeast (11) *C. elegans* (326), rice (600) and *Arabidopsis* (700) (67). However, many F-box proteins may contain a less canonical F-box motif (99). In such cases, identification of the candidate F-box proteins in the protein databases using the standard sequence search algorithms is quite difficult. The classification of such a protein as a member of the F-box family relies on confirming its association with the other components of the SCF complex and its activity towards a particular protein substrate. The prototypical F-box proteins such as Skp2, FBXW7, or  $\beta$ -TrCP have been relatively well studied with their corresponding target substrates P27 (inhibitor of cell cycle G1/S progression) NOTCH (activator of cell cycle G2/M progression) and I $\kappa$ B (inhibitor of  $\kappa$ B), respectively (54). These studies clearly indicate that F-box proteins act as the substrate-targeting subunit of the SCF ubiquitin E3 ligases.

In addition to the F-box motif, many conserved F-box proteins contain either leucine-rich repeats (LRR) or WD40 repeats. However, a large number of F-box proteins contain other protein-protein interaction modules or unknown domains (21). The LRR or WD40 repeats of F-box proteins mediate the interaction between F-box proteins and their respective substrates through phosphorylated serines or threonines (73). The differential binding specificities of protein-protein interaction domains found in F-box proteins confers the substrate specificity of the SCF ubiquitin E3 ligase.

#### **2.4. Rbx1**

The RING-H2 protein, Rbx1 was identified through its interaction with mammalian Cul1, Cul2 and yeast Cdc53 (104). It was found that addition of Rbx1 stimulates the polyubiquitination activity of SCF complexes and its homologs are highly

conserved. Rbx1 contains a cysteine-containing and zinc-binding RING-finger domain in its C-terminal half that is distinct from but related to other RING-finger E3 ligases (132). A close homolog, Rbx2 also exists and can bind to Cul1 and related cullins (Cul2-5, and Cul7) through a highly conserved C-terminal region, called the cullin homology domain (101). A more distant Rbx1 homolog, Apc11, is a component of the Anaphase promoting complex/Cyclosome (Apc/C), an ubiquitin E3 ligase complex that regulates mitosis (134). Apc11 binds to a distant cullin-related protein, Apc2, in the Apc/C complex.

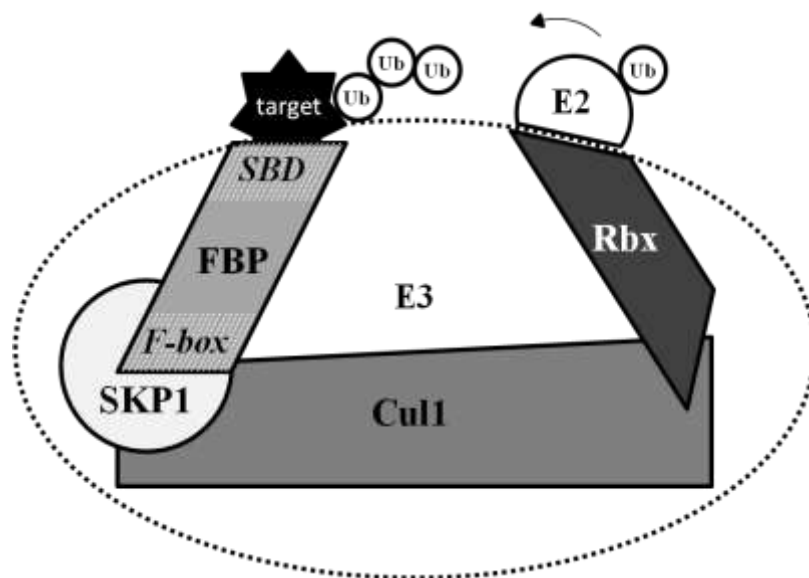
Rbx1 also binds to ubiquitin E2 enzymes such as Cdc34, and thus serves as the link between the E2 conjugating enzyme and Cul1. The RING-H2 domain of Rbx1 has been shown to be required for the E2-binding and ubiquitin-ligation reaction (92). However, although biochemical studies suggest that Rbx1 homologs share the same biochemical properties of cullin binding and act as a link between the cullin E3 ligases and the E2 enzymes, the physiological roles between Rbx1 homologs appear to differ in *Drosophila melanogaster* (79). These studies suggest that an additional mechanism may exist to distinguish between Rbx subtypes and various SCF substrates.

## CHAPTER III

# CHLOROVIRUS SKP1 AND CORE ANKYRIN-REPEAT PROTEIN INTERPLAY AND MIMICRY OF CELLULAR UBIQUITIN LIGASE MACHINERY

### 1. OVERVIEW

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 (38). This eukaryotic regulatory system mediates a variety of biological processes including protein turnover, DNA repair, and signal transduction. Through sequential reactions using three classes of enzymes, covalent attachment of ubiquitin to the substrate is achieved followed by protein degradation. This highly regulated ubiquitin cascade is initiated by the ubiquitin-activating enzyme (E1), which forms an ATP-dependent high energy bond with an ubiquitin moiety enabling passage to an ubiquitin-conjugase (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). Ubiquitin ligases serve as the enzyme platform that mediates polyubiquitination of proteins enabling cellular homeostasis by means of cell cycle regulation.



**FIG. 1. The SCF ubiquitin ligase complex.** Schematic representation of the cellular RING E3 ligase SCF complex mediating the catalytic transfer of ubiquitin (Ub) from the recruiting thioester-bound E2 conjugating enzyme to the target substrate forming a polyubiquitin chain. As the target recognition component, the F-box protein (FBP) harbors two domains: (1) the F-box domain in which resides in the amino-terminal of the protein where it binds linker protein Skp1, and (2) the substrate binding domain (SBD) responsible for bringing specific target proteins in close proximity to the catalytic core of the SCF complex.

Several ubiquitin-interfering viral proteins interact directly with E3 family enzymes, in particular the RING (Really Interesting New gene) finger-type ubiquitin ligase SCF (Skp1, Cullin, F-box) complex (106, 107). The F-box protein is one of the four subunits of the SCF complex and participates in the coordination of target recruitment for ubiquitination and degradation by the ubiquitin proteasome system. Structurally and functionally diverse, 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 interact with core members of the ligase complex (21). Substrate recruitment is dependent on this highly conserved motif located at the N-terminus and is required for interaction with the linker protein Skp1 (S-phase kinase associated protein 1). Most of the characterized cellular F-box proteins harbor additional substrate-binding motifs related to protein-protein interactions in the C-terminal region, such as leucine rich repeats (LRR), WD40 repeats (WD), or Kelch repeats (67).

- Considering its central role in substrate recruitment and degradation, it's not surprising virulence factors from bacteria and viruses have evolved functional F-box proteins compatible with cellular SCF ligase components (5, 13, 23, 82, 93, 111). Recent genomic analyses have identified F-box domains at the C-termini of some viral ankyrin-repeat (ANK)-containing proteins that are involved in the hijacking of the host cell ubiquitin machinery (2) while contributing to host cell tropism and permissiveness (118). ANK proteins are known to participate in protein-protein interactions and are involved in many functions, such as cell-cycle regulation, signal transduction, and intercellular trafficking (74, 98). Given the high number of domain combinations, ANK proteins have the ability to withstand considerable sequence variation and can accommodate additional structural and functional motifs. Although common in eukaryotes and bacteria, the 33-residue ANK motif is relatively rare among viruses, with the exception of poxviruses and their relatives. In fact, several members of the *Poxviridae* family encode ANK proteins that contain a unique C-terminal F-box domain which interacts with cellular SCF subunits (12, 15, 16, 100, 102, 111). These experiments established that the poxviral



motif at the C-terminus is a functional F-box that exploits the versatile ANK motif as a protein-interaction module to dictate targeting specificity. Further, these ankyrin-containing proteins from Orf virus, for example, interact *in vivo* with the SCF complex component Skp1 in an F-box-dependent but ankyrin-independent manner.

As a member of the *Phycodnaviridae* family, chloroviruses are included in the same seven member monophyletic group as *Poxviridae*, classified as Nucleo-Cytoplasmic Large DNA Viruses (NCLDV) (128). The NCLDV infect animals and diverse unicellular eukaryotes, and replicate either exclusively in the cytoplasm of their host cells, or possess both nuclear and cytoplasmic stages in their life cycle. Members of this group contain genome sizes that range from 100-kb to 1.2-Mb. A comprehensive genome analysis of the NCLDVs identified approximately 50 conserved genes that mapped to the genome of a common ancestral virus while five genes were conserved in all NCLDV sequenced genomes (127). Tracing the evolutionary lineage of this monophyletic group of viruses may reveal shared ancestor genes designed to optimize the subversion of cellular E3 ubiquitin ligase activity.

Collectively referred to as chloroviruses, NC64A-, SAG- and Pbi-viruses infect one of three different symbiotic chlorella-like green algal species: the sequenced and annotated *Chlorella variabilis* NC64A, *Chlorella* SAG 3.83 (renamed *Chlorella heliozoae*) and *Chlorella* Pbi (renamed *Micratinium conductrix*), respectively (86). The prototype chlorovirus, *Paramecium bursaria* chlorella virus-1 (PBCV-1) is a 331 kb double-stranded DNA, plaque forming virus that is predicted to encode 416 proteins and 11 tRNAs (114). The host of PBCV-1, the unicellular green alga *Chlorella variabilis* strain NC64A, is a natural endosymbiont of the protozoan *Paramecium bursaria*.

Chloroviruses encode many candidate proteins that have the potential to manipulate the SCF complex. For example, PBCV-1 encodes eight ANK proteins that appear to contain C-terminal domains characteristic of the SCF subunit F-box protein, imitating poxvirus ANK genes. The truncated C-terminal position of the ANK protein-binding domain raises the question of whether the extreme C-terminal domains of the PBCV-1 ANK proteins function as F-boxes. The 41 annotated NC64A-, SAG- and Pbi-viruses have several ANK proteins ranging from 5 to 9, 2 to 3 and 3 to 7 members, respectively.

PBCV-1 and 39 of the annotated chloroviruses also encode a putative Skp1 protein (vSkp1). Cellular Skp1 serves as a critical subunit of the SCF complex which governs cell cycle regulation. As the linker subunit, cellular Skp1 bridges cullin-1 (Cul1) and F-box proteins to form the E3 ligase catalytic core (Fig. 1). The *C. variabilis* algal host encodes two Skp1 homologs; however sequence analysis suggests that one gene (59724) is more likely to encode a functional protein than the other one (28112), as is the case of other species encoding multiple Skp1 or Skp1-like genes (57). We hypothesize that chloroviruses use their vSkp1s and putative F-box-like ANK proteins to hijack or divert the host alga SCF complex to favor viral infection. Here we present the results of protein interplay between cellular SCF proteins and viral-encoded proteins. Our results provide the first evidence, to the authors' knowledge, of a functioning vSkp1 with partnering virus and cellular genes. Furthermore, we identified the minimum viral ANK domain required for protein binding. We predict that viral genes, specifically vSkp1<sub>PBCV-1</sub> and its partnering ANK proteins are capable of mimicking the Skp1 and F-box components of eukaryotic SCF ubiquitin ligase machinery, respectively. These viral

proteins are predicted to alter the host's SCF complex and provide an environment more supportive of viral replication.

## **2. MATERIALS AND METHODS**

**2.1. Cell Lines.** *E. coli* DH5-alpha and BL21-Gold(DE3)pLysS (BGDS) competent cells were grown in LB (Stratagene, La Jolla, CA). Media were supplemented with 100 µg/ml carbenicillin and 34 µg/ml chloramphenicol. MaV203 *Saccharomyces cerevisiae* (Proquest Two-hybrid System, Invitrogen) served as the host yeast strain for the bait and prey plasmids. The yeast GAL4-based two-hybrid system includes deletions in its endogenous GAL4 and GAL80 genes. The strain also contains a set of non-reverting auxotrophic mutations (*leu2* and *trp1*) for the selection of bait and prey fusion vectors. MaV203 competent cells were cultivated in YPAD prior to transformation with bait and prey plasmids.

**2.2. cDNA Cloning and Plasmid Constructs.** Host and virus-encoded genes were designed and synthesized 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). PCR products were cloned directionally by adding four bases to the forward primer (CACC). The entry clone was selected with 50 µg/ml kanamycin. The tailor-made 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 domain in the vector backbone. The structure of each vector was sequence verified. The

same cloning technology was used to produce fusion hybrid proteins. The yeast expression vectors pDEST32 and pDEST22 were used to generate GAL4 DNA-binding domain and GAL4 DNA activation domain fusion proteins and were selected with 50 µg/ml gentamycin and 100 µg/ml ampicillin, respectively. Yeast expression clones were selected based on antibiotic markers. Plasmids were isolated with a QIAprep Spin Miniprep Kit (Venlo, Limburg) according to the manufacturer's instructions. The protein overexpression strategy used pET-23a vector (Novagen Biosciences, San Diego, CA), which carries an N-terminal T7-tag sequence in addition to a C-terminal His-6 sequence tag. Total plasmid concentration was quantified with a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE).

**2.3. Yeast Two-Hybrid Screening.** Yeast strain MaV203 competent cells were transformed with bait and prey vectors according to a high-efficiency transformation protocol described in (39). The carrier DNA used was optimized Yeastmaker Carrier DNA-Salmon sperm DNA (Clontech, Mountain View, CA). Selection criteria regarding yeast transformants containing bait and prey vectors were based on the ProQuest Two-Hybrid System (Invitrogen), according to the manufacture's instruction manual. Nutritionally selective medium deficient in tryptophan and leucine (SC-Leu-Trp) was used to test for positive transformation. To assess protein-protein interaction, induction of three unrelated GAL4 inducible reporter genes (*HIS3*, *URA3* and *lacZ*) downstream at separate loci in the yeast genome enabled the measure of interaction strength between bait and prey proteins. Reporter genes *HIS3* and *URA3* were tested on respective dropout selection media. X-gal colorimetric assays were used to test induction of the *lacZ* reporter

gene measuring  $\beta$ -galactosidase activity. To select against positive interactions, medium was supplemented with 5-Fluoroorotic acid (5FOA) which is toxic to yeast cells expressing *URA3*. This combination of reporter genes and sensitivity assays allowed the interpretation of four phenotypes to access true interactions. Clones containing bait and prey that induce the reporter genes without interacting in the two-hybrid system are defined as false positives. False positives in which bait and prey undergo self-activation were reduced by testing co-transformed yeast cells with bait or prey plasmid and an empty plasmid according to the manufacturers protocol (Invitrogen). To further test for self-activation, the threshold of resistance to 3-amino-1,2,4-triazole (3AT) was determined in a dose-dependent manner (10, 25, 50 and 100 mM), an enzyme encoded by *HIS3* involved in histidine biosynthesis. Including this concentration of 3AT in plates deficient in histidine, minor increases in *HIS3* reporter gene expression are detected. The empty vectors and pEXP32/Krev1 and pEXP22/RalGDS-m2 vectors were used as negative controls, whereas yeast strains which were co-transformed with the pEXP32/krev1 and pEXP22/RalGDS-wt vectors were used as positive controls (Invitrogen).

**2.4. Cells and Viral Lysate Preparation.** *C. variabilis* cells ( $2 \times 10^7$  cells/ml) from an actively growing culture in modified Bold's basal medium (113) were inoculated with purified PBCV-1 at an MOI of 5 with agitation to ensure uniform infection. Infected cells were harvested by centrifugation (5000 x g) for 5 min at 4°C at 0, 15, 30, 45, 60, 90, 120, 180, 240 and 360 min postinfection (p.i.). Cells were immediately harvested following virus addition at time zero. *C. variabilis* cells were also harvested without addition of

virus as a control. The cells were resuspended in 2 mM PMSF (phenylmethanesulfonylfluoride) PBS buffer. Glass beads (0.25-0.3 mm) were added to the infected cell suspension and vortexed for 5 min. Cells were then sonicated (5 sec pulse at 30% amplitude) for 5 min on ice. The cell lysate was centrifuged ( $16000 \times g$ ) for 10 min at  $4^{\circ}\text{C}$  and supernatant was collected. For pull-down experiments, the supernatant was collected and diluted with 4 ml PBS before applying to a vSkp1-bound column.

**2.5. Transfection and Pull-down by Metal Chelation Chromatography.** BGDS cells were transfected with pET-23a encoding vSkp1<sub>PBCV-1</sub> and stable cell lines were generated by carbenicillin and chloramphenicol selection. Batch culture was grown to OD<sub>600</sub> 0.6 at  $37^{\circ}\text{C}$ . Induction of protein overexpression was achieved with the addition of 0.5 mM IPTG for 2 h at  $25^{\circ}\text{C}$ . Cells were harvested by centrifugation ( $5000 \times g$ ) for 5 min at  $4^{\circ}\text{C}$  then disrupted by sonication (5 sec pulse at 30% amplitude) twice in 2 mM PMSF PBS buffer for 5 min on ice. Cell lysate was centrifuged ( $5000 \times g$ ) for 5 min at  $4^{\circ}\text{C}$  and the supernatant was applied to a Ni<sup>2+</sup> column. Column resin and elution buffer were from a Ni-NTA Buffer Kit (Novagen Biosciences, San Diego, CA). One ml of resuspended Ni-NTA agarose in 4 ml of PBS was poured into the column and the resin was allowed to settle. Following exit of the flow-through, the column was washed with PBS. Eight ml of sample lysate containing overexpressed vSkp1<sub>PBCV-1</sub> was poured into the column. The Ni<sup>2+</sup> column was washed four consecutive times with 10 mM imidazole PBS. Three separate lysates were applied to the vSkp1<sub>PBCV-1</sub> 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(s) was eluted from the column with 0.5 ml Ni-NTA Elution Buffer four

consecutive times. Collected protein was concentrated by Microcon (YM-10) centrifugal filter (Millipore, Bedford, MA). Protein concentration was determined by Coomassie Plus Bradford Assay (Thermo Fisher Scientific, Waltham, MA). Eluted proteins were resolved by SDS/PAGE using PAGER Gold 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. Endogenous cellular F-box and viral ANK proteins in complex with the vSkp1<sub>PBCV-1</sub> were identified using inductively coupled plasma mass spectrometry (Agilent 7500cx) located in the Redox Biology Center, University of Nebraska-Lincoln.

**2.6. Western Blot Analyses.** Samples were separated on SDS-PAGE gels and transferred to PVDF membranes (0.45  $\mu\text{m}$  pore) using standard procedures (110). 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-vSkp1<sub>PBCV-1</sub> (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) (Sigma, A6154) as the secondary antibody. After an additional five washes with TBS-T, the membranes were developed using ECL reagents SuperSignal West Pico chemiluminescence substrate (ThermoFisher Scientific). Bands were visualized using ChemiDoc MP Imaging System with Imagine Lab Software (Bio-Rad, Hercules, CA).

**2.7. Antibody.** Polyclonal rabbit-anti-vSkp1<sub>PBCV-1</sub> was produced by Harlan Bioproducts for Science, Inc. The antiserum was prepared using the Standard 73-day Protocol by immunizing one rabbit with 500 µg of peptide for three consecutive boosts.

**2.8. Sequence identification and alignments.** Gene sequences were identified by a BLAST search using the National Center for Biotechnology Information nonredundant database. Constructs were yeast codon optimized. The C-termini of the eight PBCV-1-encoded ANK proteins were aligned with the F-box consensus sequence adapted from (54) using ClustalW2 (71) running MEGA3.1 with default settings (Gonnet matrix, gap opening penalty 10, gap extension penalty 0.1 for pairwise alignments and 10 and 0.2, respectively, for multiple alignments). Known residues of Skp2 that contact cellular Skp1 were reported by (100). Sequence alignment of Skp1 homologs was created by the Phylogeny.fr platform (27, 28) using default settings (MUSCLE, BLOSUM62 substitution matrix, Gblocks for alignment curation).

**2.9. Phylogenetic Analysis.** To construct a robust phylogenetic tree from chlorovirus proteins, sequence alignments were created using the Phylogeny.fr platform (27, 28) with default settings (MUSCLE 3.7 for multiple alignment, Gblocks 0.91b for alignment curation, PhyML 3.0 aLRT for phylogeny, TreeDyn 198.3 for tree rendering). The following additional parameters were enforced: 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 WAG substitution model. We



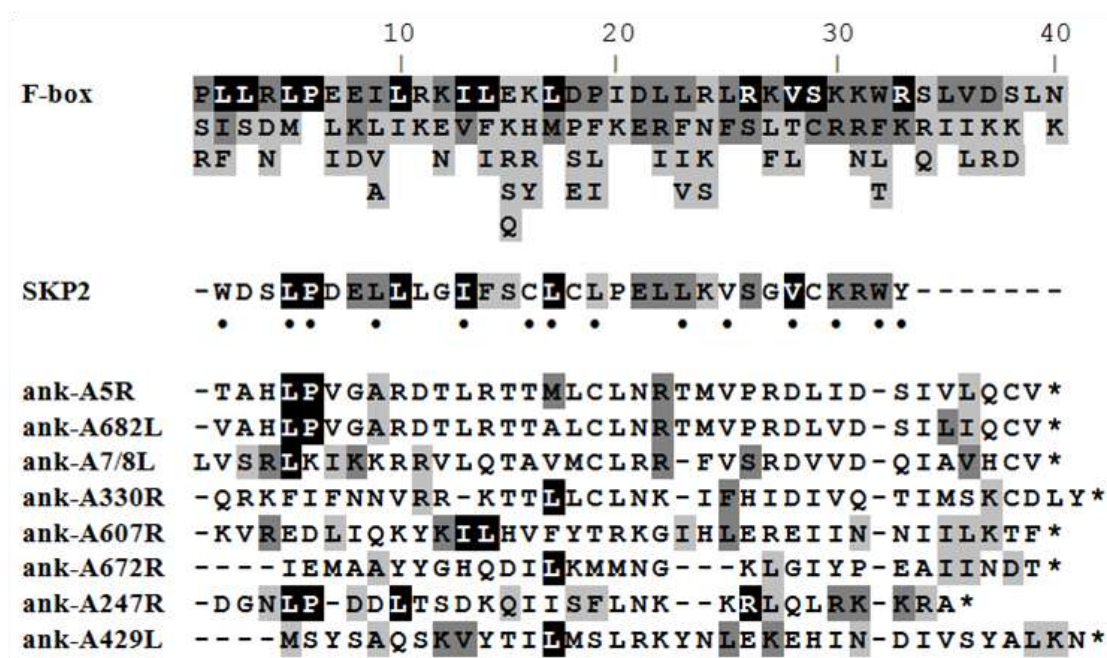
only show bootstrap values < 90%. Branches with bootstrap support less than 50 were collapsed.

**2.10. Nucleotide sequence accession numbers.** The locus-tag identifiers investigated in this article correspond to the National Center for Biotechnology Information (NCBI) reference sequences. PBCV-1 ANK proteins: A672R, A005R, A682L, A247R, A607R A330R, A7/8, A429L; Skp1 proteins: PBCV-1\_A039L, ATCV-1\_Z339L, Fr483\_N799R, CHLNCDRAFT\_59724 and CHLNCDRAFT\_28112 (*C. variabilis*). Accession numbers AAH09839, P52285 and AAC49492 correspond with *H. sapiens*, *D. discoidium*, and *S. cerevisiae* Skp1 proteins, respectively; *C. variabilis* F-box proteins: CHLNCDRAFT\_141219, CHLNCDRAFT\_134953, CHLNCDRAFT\_52604, CHLNCDRAFT\_133199.

### 3. RESULTS

**3.1. Chlorovirus ankyrin-repeat protein displays a unique carboxyl terminal F-box-like domain.** The combination of the substrate binding domain and F-box domain is responsible for the bipartite functionality of cellular F-box proteins. Chlorovirus ANK proteins are composed of the same domain ingredients. ANK regions concentrated in the N-terminus operate as substrate binding domains that mediate protein-protein interactions. Interestingly, protein alignment of the total eight PBCV-1 encoded ANK proteins revealed a common motif in the C-terminal region (Fig. 2). Further manual inspection of this widely shared region revealed a conserved domain characteristic of a cellular F-box, a protein structural motif of about 50 amino acids. Although sharing

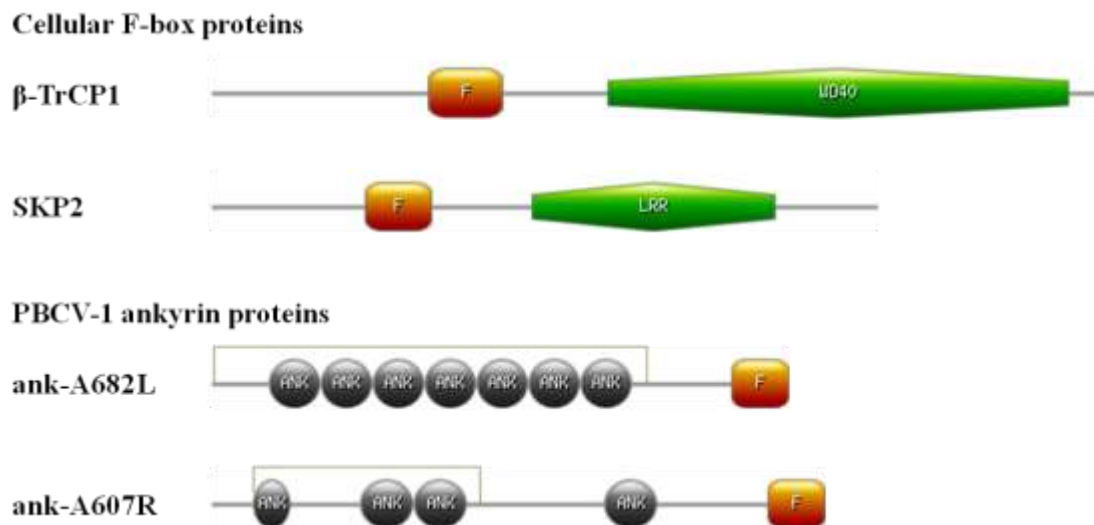
common residues of the F-box motif, the F-box-like domain in viral ANK proteins is truncated (30-40 amino acids) and positioned at opposite ends of cellular F-boxes. For example, the F-box-like domains of two ANK proteins coded by PBCV-1 are shown in comparison to the F-box domains from cellular F-box proteins,  $\beta$ -TrCP and Skp2 (Fig. 3).



**FIG. 2. PBCV-1 ankyrin-repeat proteins contain a carboxyl-terminal F-box-like domain.** ClustalW2 was used to align the carboxyl-termini of the eight PBCV-1-encoded ANK proteins with the F-box consensus sequence adapted from (54) and F-box domain from Skp2, a cellular F-box protein. Residues are shaded according to frequency in the F-box consensus sequence (top) which was derived from the alignment of 234 sequences used to create the Pfam F-box profile; black signifies residues found in over 40% of the F-box sequences; dark gray signifies residues found in 20-40% of the F-boxes; light gray signifies residues found in 10-20% of the F-boxes. The dots indicate known residues of

Skp2 that contact cellular Skp1 according to (100). A dash indicates a space and an asterisk indicates the end of the protein.

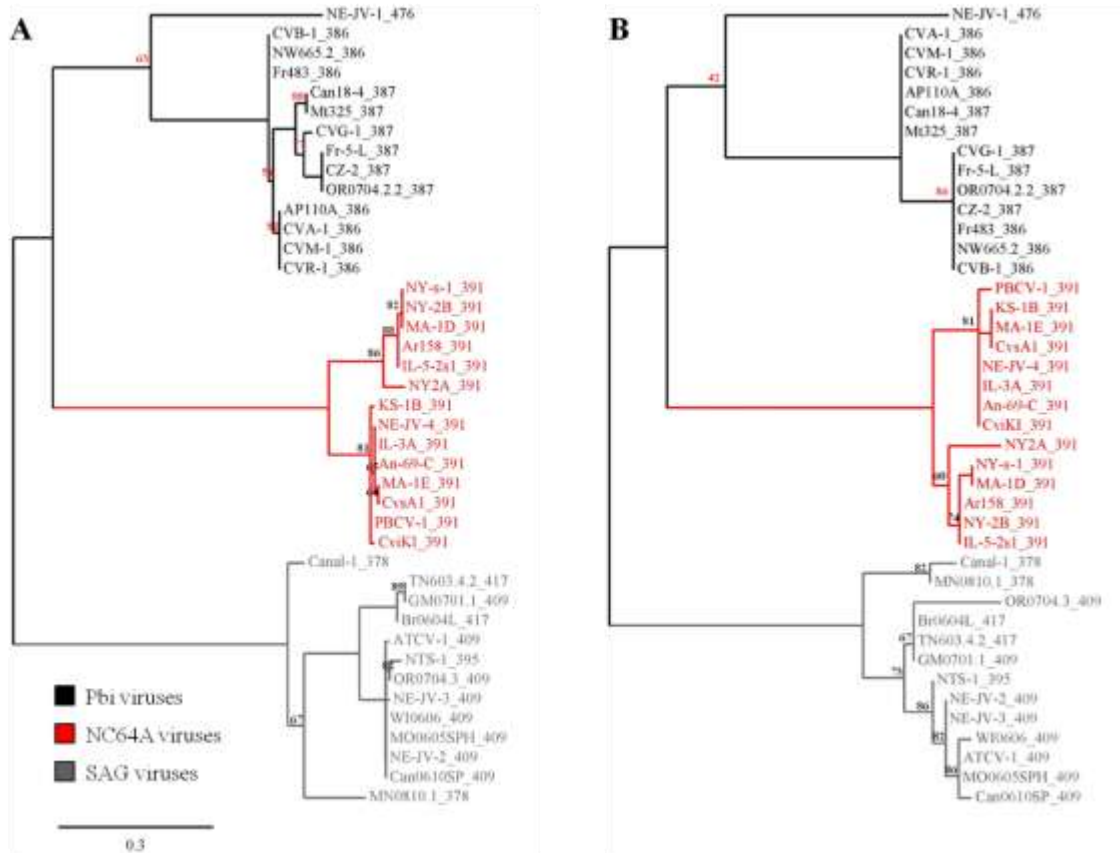
When aligned with a consensus F-box sequence (54) derived from the Pfam F-box profile, the C-terminal region of the eight PBCV-1-encoded ANK proteins contain several conserved residues in characterized cellular F-boxes (Fig. 2). Among these shared residues are known amino acids reported to mediate contact between cellular proteins Skp1 and F-box protein SKP2 (100). However, the number of conserved residues among the viral proteins varies in alignment accuracy with the established F-box motif. Based on domain sequence similarity, PBCV-1 ANK proteins ank-A5R, ank-A682L and ank-A247R form a class of ANK proteins that contain a C-terminal F-box-like motif that most closely resembles a canonical F-box protein. Specifically, the three ANK proteins possess leucine-proline residues positioned in line with the signature dipeptide of cellular F-box proteins. This class of F-box-like viral ANK proteins is highly conserved in the chloroviruses and its C-terminal position resembles the so-called PRANC (Pox protein Repeats of Ankyrin C-terminal) domain from poxvirus-encoded ANK proteins which interact with SCF proteins and possibly mediate substrate-specific degradation via its F-box domain (7). Hypothetically, the duality of the ANK domain and putative F-box motif in chlorovirus ANK proteins may function as substrate binding and F-box domains, respectively.



**FIG. 3. Domain organization of PBCV-1 ankyrin-repeat proteins.** Comparison of the domain location between cellular F-box proteins  $\beta$ -TrCP and Skp2 with two PBCV-1 encoded ANK proteins representing two classes of functional chlorovirus ANK proteins. The ank-A682L exemplifies a cellular F-box-like ANK protein whereas ank-A607R is a member of the core ANK family proteins. The putative F-box motif resides in the C-terminal region of the viral proteins opposite to cellular F-box domains. ANK regions are identified by ScanProsite and indicated by brackets. Domains are restricted by the first and final residue number in respective domains and sorted by the following color scheme: Cellular and predicted PBCV-1 ANK F-box domains are orange, cellular substrate binding domains WD40 and LRR (leucine rich repeats) are green. PBCV-1 ANK domain is black.

**3.2. Highly conserved core ankyrin repeat protein family among chloroviruses constitutes monophyletic clusters according to host alga.** Gene prediction algorithms from a chlorovirus genomics study (51) identified 319 to 416 protein-encoding genes (CDSs) per annotated genome of 41 chloroviruses that infect one of three

different symbiotic chlorella-like species. Genome surveillance identified 178 ANK-containing proteins distributed among the three types of chloroviruses. One ankyrin protein family contains 51 members that are distributed among all three virus families. After eliminating likely gene duplications, all 41 annotated chloroviruses contain at least one member of this core protein family of ANK genes. The member from PBCV-1 is ank-A607R. This core family of ANK proteins present in the three virus types was analyzed by a maximum likelihood reconstruction method in two ways: (i) core ANK domains only (Fig. 4A) and (ii) the C-terminal 40 amino acids (Fig. 4B). Sequences were aligned with MUSCLE (v3.7) and the phylogenetic trees were computed using the WAG substitution model. Phylogenetic analysis of the core ANK domains revealed three strongly supported clades clustered by the virus host alga. Bootstrap branch supports of 100, 100 and 97 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 core ANK domains are well preserved among the chlorovirus families as several of these ANK domains have 100% sequence identity (Table 1). Core ANK domains are most diverse among viruses that infect different host algae (e.g., the highest degree of sequence identity of the core ANK domains between any two virus families is less than 50%).



**FIG. 4. Phylogenetic trees of the core ankyrin-repeat protein family encoded by chloroviruses.** (A) ANK domain of core ANK family protein. (B) Putative F-box domain (last 40 residues of C-terminal) of core ANK family proteins. A maximum likelihood reconstruction method based on the separate alignments of the ANK and F-box-like domains from the core ANK family proteins native to 41 chloroviruses that infect 3 algal hosts. The ANK and F-box-like domains are located in the amino and carboxyl terminals of the ANK protein, respectively. Corresponding geographical sampling locations of chloroviruses: NC64A viruses NY-2A (New York, USA), NY-2B (New York, USA), Ar158 (Buenos Aires, Argentina), NY-s-1 (New York, USA), MA-1D (Massachusetts, USA), IL-5-2s1 (Illinois, USA), An-69-C (Canberra, Australia), IL-3A (Illinois, USA), PBCV-1 (North Carolina, USA), NE-JV-4 (Nebraska, USA), MA-1E (Massachusetts,

USA), CvsA1 (Sawara, Japan), CviKI (Kyoto, Japan); SAG viruses Can0610SP (British Columbia, Canada), OR0704.3 (Oregon, USA), NE-JV-3 (Nebraska, USA), NE-JV-2 (Nebraska, USA), ATCV-1 (Stuttgart, Germany), WI0606 (Wisconsin, USA), MO0605SPH (Missouri, USA), NTS-1 (Nebraska, USA), Br6064L (Sao Paulo, Brazil), TN603.4.2 (Tennessee, USA), GM0701.1 (Guatemala), MN0810.1 (Minnesota, USA), Canal-1 (Nebraska, USA); Pbi viruses NE-JV-1 (Nebraska, USA), Mt325 (Montana, USA), Can18-4 (Canada), Fr-5-L (France), CZ-2 (Czech Republic), OR0704.2.2 (Oregon, USA), CVB-1 (Berlin, Germany), Fr483 (France), NW665.2 (Norway), AP110A (Unknown), CVA-1 (Amönau, Germany), CVG-1 (Göttingen, Germany), CVM-1 (Marburg, Germany), CVR-1 (Rauschenberg, Germany).

Analysis of the last 40 residues of the C-terminal region (putative F-box domain) from the core ANK family proteins again showed three strongly supported clades grouped by virus host alga with equivalent bootstrap branch supports. Furthermore, the highest degree of sequence identity of the terminal region between any two different 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 core ANK domain (42 ELW).

**TABLE 1** Protein sequence similarity of chlorovirus core ANK and vSKP1 proteins

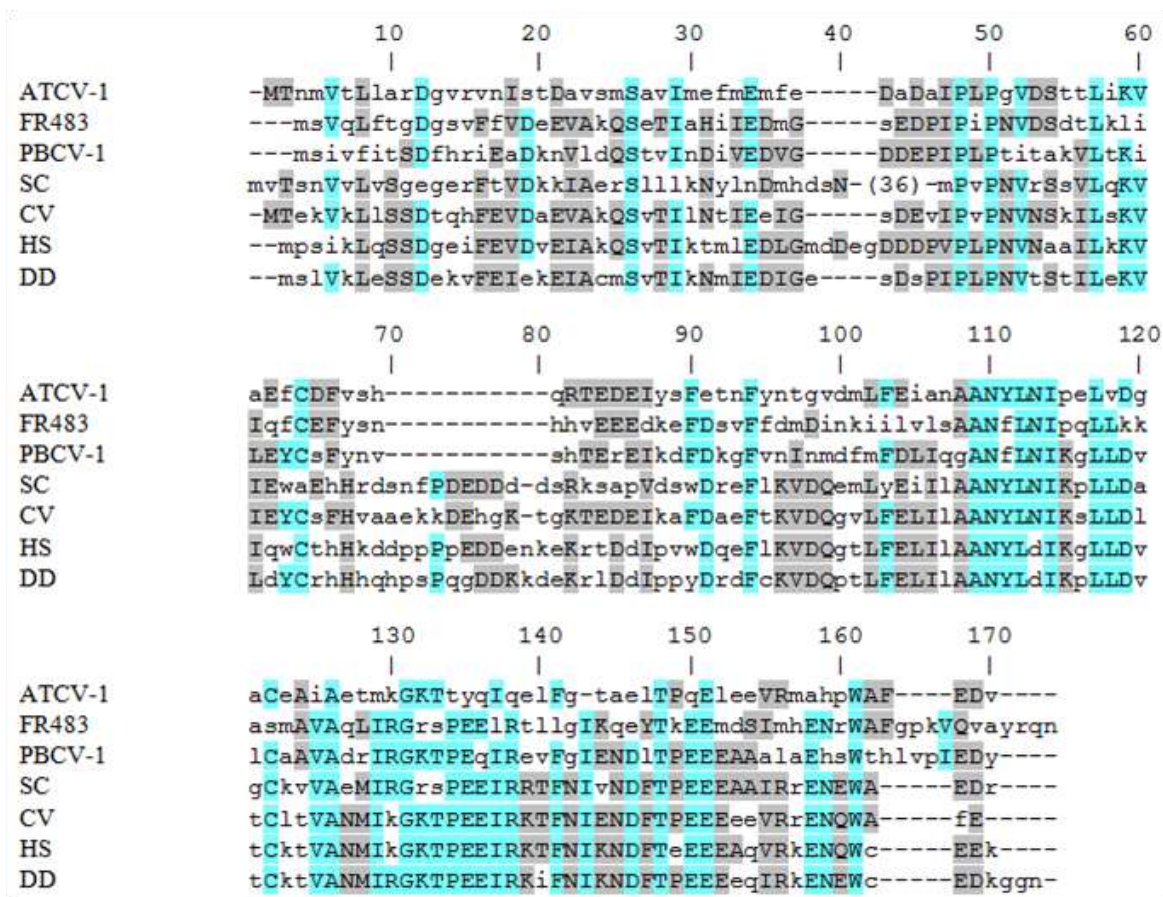
Chlorovirus	ANK		vSKP1
	ANK Domain	C-terminal	Full-length Protein
NC64A viruses	72.7-100%	62-100%	64.9-100%
SAG viruses	60.7-100%	62.5-100%	74.5-100%
Pbi viruses <sup>a</sup>	73.2-100%	85-100%	89-100%

<sup>a</sup>Percent protein sequence similarity of Pbi viruses was calculated excluding NE-JV-1.

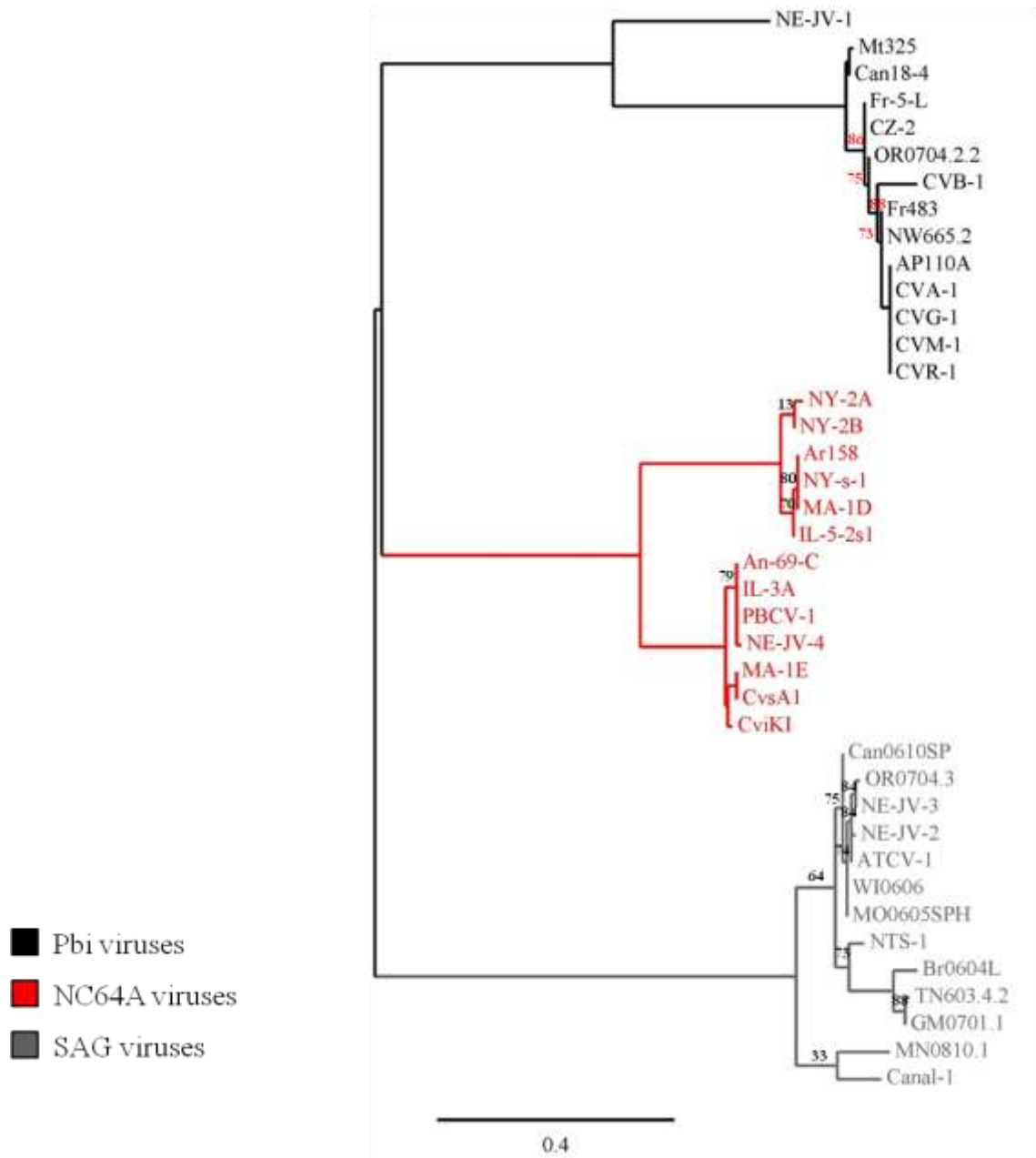
**3.3. Chloroviruses encode a conserved homolog of cellular Skp1 that resembles the same monophyletic origin as the viral core ankyrin-repeat protein family.** Genome analysis revealed one vSkp1 protein in all sequenced chloroviruses except for NC64A virus KS-1B. Skp1 is a core component of the eukaryotic SCF complex that mediates proteasome-dependent protein degradation. Cellular Skp1 also functions in cell-cycle progression, transcriptional regulation and signal transduction (46). Multiple sequence alignment of cellular Skp1 homologs including host *C. variabilis* were aligned with vSkp1 proteins from NC64A-, SAG- and Pbi virus members PBCV-1, ATCV-1 and FR486, respectively (Fig. 5). Protein alignment showed 70% conserved amino acid composition among the cellular and viral proteins. A maximum likelihood reconstruction method analyzed an alignment of vSkp1 genes native to 40 chloroviruses that infect the three algal hosts (Fig. 6). Sequences were aligned with MUSCLE (v3.7) and the phylogenetic tree was computed using the WAG substitution model. Phylogenetic analysis of vSkp1 proteins revealed three strongly supported clades grouped by the virus host alga. Bootstrap branch supports of 99, 100



and 99 (expressed in Expected-Likelihood Weights, ELW) support the monophyly of the NC64A-, SAG- and Pbi-viruses, respectively (including Pbi virus NE-JV-1, 97 ELW). While a number of chloroviruses of the same host share identical Skp1 proteins, the highest sequence identity is significantly reduced to less than 40% when comparing Skp1 proteins from chloroviruses that infect different hosts (Table 1). Similar to chlorovirus ANK proteins, vSkp1 proteins are not associated with the virus proteome which is comprised of 148 unique virus-encoded proteins (about 35% of the virus coding capacity) and 1 host protein (29). 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, which is consistent with the immediate early expression of chlorovirus ANK (53) and early vSkp1 proteins. Immunoblotting confirmed the appearance of the vSkp1<sub>PBCV-1</sub> protein during the time course of infection at 45 min p.i. (Fig. 7).

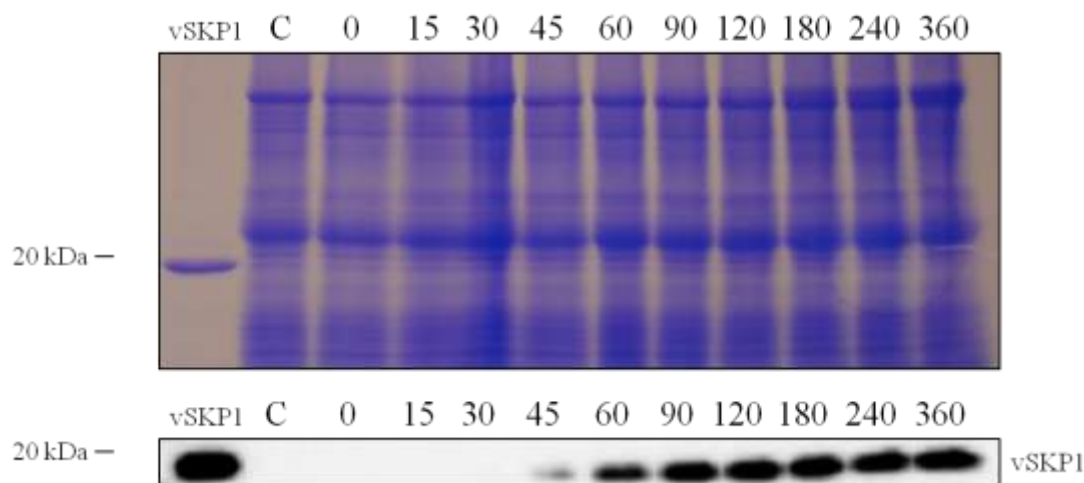


**FIG. 5. Sequence alignment of Skp1 homologs.** Amino acid identities are colored as the most conserved (according to BLOSUM62). Average BLOSUM62 scores represented as max: 3.0 (cyan) and low: 0.5 (gray). *S. cerevisiae* Skp1 has a 36 amino acid insertion of primarily N, D, and E residues in the N-terminal domain. SC is *S. cerevisiae*, CV is *C. variabilis*, HS is *Homo sapiens*, and DD is *Dictyostelium discoideum*. Protein alignment was created using (28).



**FIG. 6. Phylogenetic tree of vSkp1 homologs coded by chloroviruses.** A maximum likelihood reconstruction method based on an alignment of chlorovirus-encoded vSkp1 genes native to 40 chloroviruses that infect 3 algal hosts. Sequences were aligned with MUSCLE (v3.7) configured with default settings. The following parameters were enforced: minimum number of sequences for a flank position was 85%; all segments with contiguous nonconserved positions bigger than 8 were rejected; minimum block length

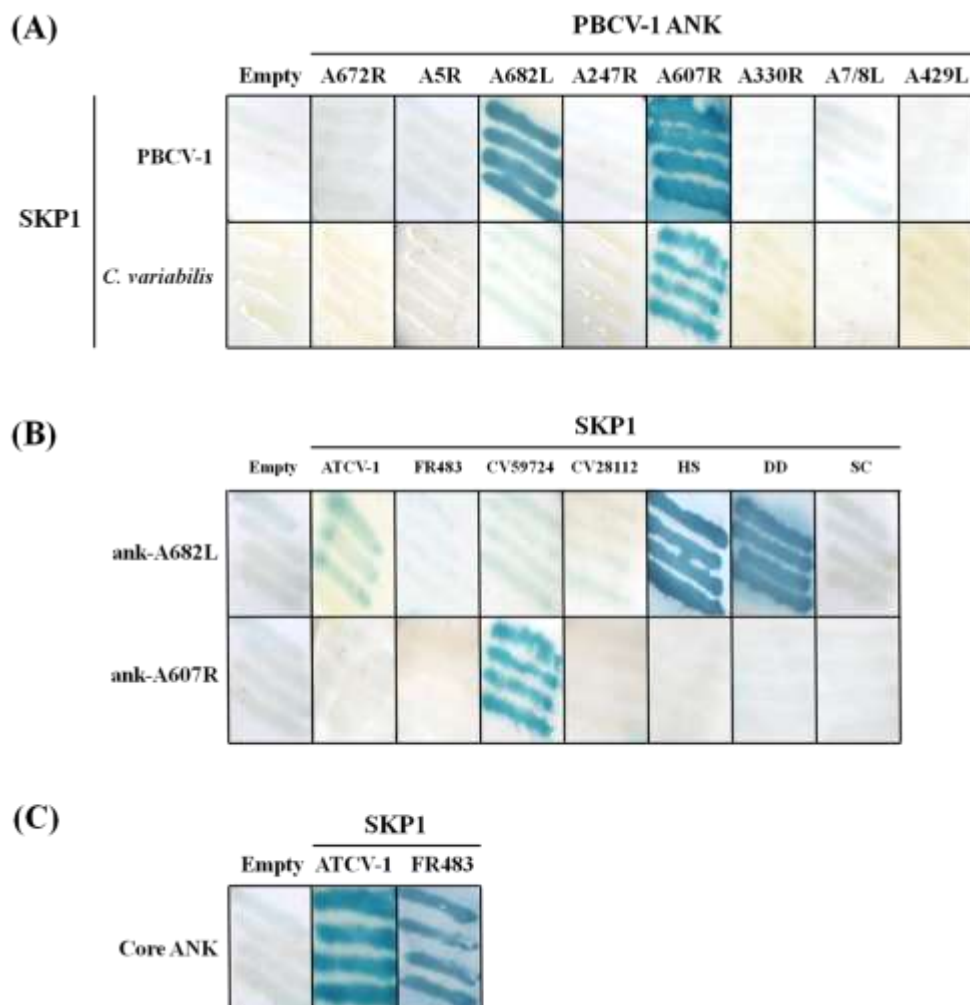
was 10; no gaps were allowed in the final alignment. The phylogenetic tree was computed using the WAG substitution model. We only show bootstrap values < 90%. Branches with bootstrap support less than 50 were collapsed.



**FIG. 7. Early translation of vSkp1<sub>PBCV-1</sub> protein.** Total protein was isolated from uninfected (control) and PBCV-1-infected *C. variabilis* cells at the indicated times (min.). Time zero signifies cells were immediately harvested after virus addition. Viral lysates were analyzed by SDS-PAGE (upper panel) and immunoblotting (lower panel). Detection of the vSkp1<sub>PBCV-1</sub> protein is shown with increasing intensity with time. The protein was probed with antiserum against vSkp1<sub>PBCV-1</sub>.

**3.4. Cellular Skp1 and vSkp1<sub>PBCV-1</sub> proteins interact differently with two of eight ankyrin-repeat proteins from PBCV-1.** The functional diversity of ANK-containing proteins is achieved through its co-existence of multiple interactive modules that mediate protein-protein interaction. To explore their interaction potentials, we performed yeast two-hybrid (Y2H) screens to identify binding partners for the viral ANK

proteins. We assayed all eight ANK proteins coded by PBCV-1 with either cellular or vSkp1<sub>PBCV-1</sub> proteins. Sequence constructs were yeast codon-optimized and cloned into the pDEST32 bait and pDEST22 prey vectors and tested for interactions. The Y2H screen showed selective binding interaction between PBCV-1 ANK and Skp1 proteins. Results confirmed vSkp1<sub>PBCV-1</sub> interacted strongly with two of the eight ANK proteins, PBCV-1 ank-A682L and ank-A607R (Fig. 8A). PBCV-1 ank-A607R interacted with one of the two Skp1 proteins (59724) from *C. variabilis*. The remaining six PBCV-1 ANK proteins did not interact with any of the Skp1 proteins. To further test the binding potential of ank-A682L and ank-A607R proteins, we assayed these with the cellular Skp1 proteins from host *C. variabilis*, human, Dictyostelium and yeast, in addition to vSkp1<sub>ATCV-1</sub> and vSkp1<sub>FR483</sub> proteins. PBCV-1 ank-A682L interacted with cellular Skp1 proteins from human and Dictyostelium while weakly interacting with vSkp1<sub>ATCV-1</sub> (Fig. 8B). PBCV-1 ank-A682L did not interact with any additional Skp1 homologs. A summary of Y2H interactions involving PBCV-1 ank-A682L and ank-A607R proteins is reported in Table 2.



**FIG. 8. Interactions between PBCV-1 and cellular SCF-associated proteins detected by Y2H.** PBCV-1 encoded F-box-like protein ank-A682L and core ank-A607R protein interact with vSkp1<sub>PBCV-1</sub>, however only ank-A607R interacts with Skp1 from host *C. variabilis* (CHLNCDRAFT\_59724) (A). PBCV-1 ank-A682L interacted with cellular Skp1 proteins from human (HS) and Dictyostelium (DD) while weakly interacting with vSkp1<sub>ATCV-1</sub>. Both PBCV-1 ANK proteins did not interact with the second Skp1 from host *C. variabilis* (CV28112) (B). Similar to the binding nature between vSkp1<sub>PBCV-1</sub> and ank-A607R, vSkp1<sub>ATCV-1</sub> and vSkp1<sub>FR483</sub> proteins also interact with their partner core ANK proteins, ank-Z568R and ank-N710R, respectively (C). Interacting expression

vectors induce the downstream colorimetric reporter gene lacZ (blue) indicating positive protein-protein interaction. The eight PBCV-1 ANK proteins are designated by locus tag.

**TABLE 2** Summary of Y2H

<b>SKP1</b>	<b>ank-A682L</b>	<b>ank-A607R</b>
PBCV-1	+	+
ATCV-1	+	—
FR483	—	—
<i>C. variabilis</i>	—	+
Human	+	—
Dictyostelium	+	—

**3.5. PBCV-1 ank-A682L and Skp1 proteins interact in a putative F-box-dependent manner.** We constructed truncated versions of PBCV-1 ank-A682L by making a series of deletions from the N-terminus (Fig. 9) in order to identify the minimum domain required for binding to the Skp1 proteins. According to ExPASy's PROSITE motif scan (97), ank-A682L has seven ANK domains designated by the following residue numbers: 38-70, 71-103, 104-136, 137-169, 170-202, 203-235 and 236-268. We constructed six mutants with different combinations of ANK repeats and F-box-like region deletions. One mutant contained only the F-box-like domain excluding all ANK repeat motifs while another construct contained all the ANK repeats without the C-terminal putative F-box domain. An additional double-point mutation of ank-A682L was constructed substituting the C-terminal dipeptide Leu-Pro (position 335 and 336) of the

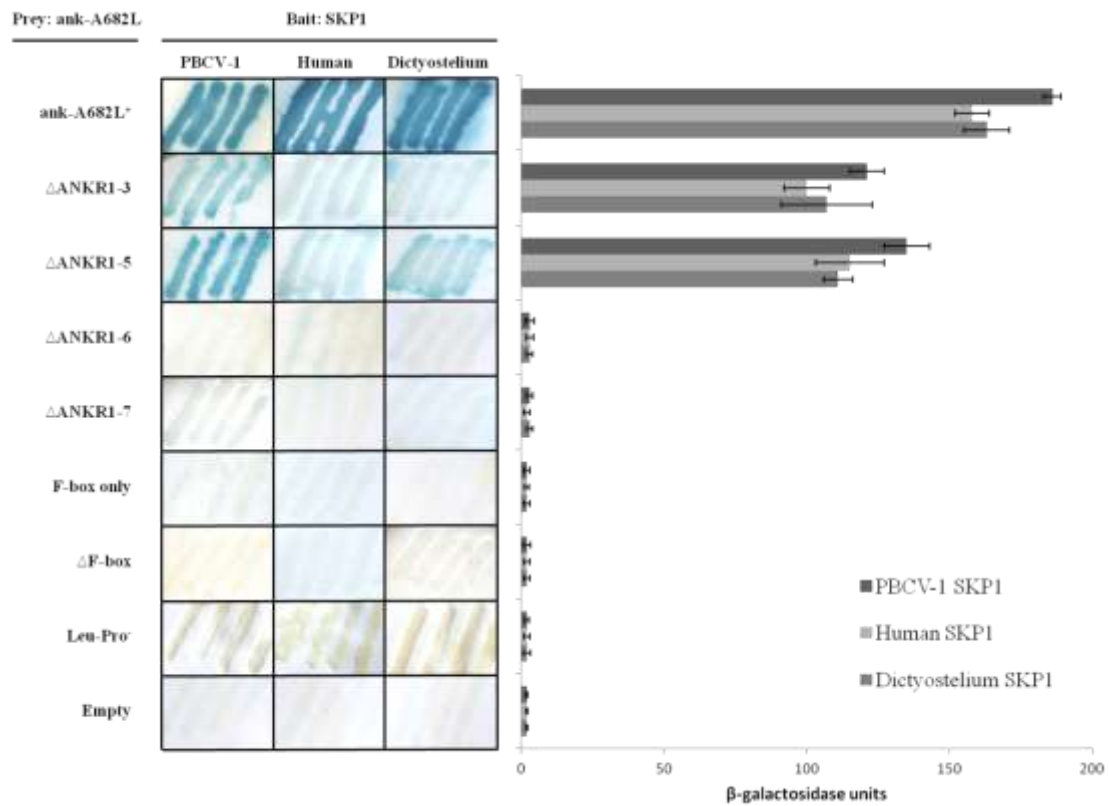
putative F-box motif with two alanine residues. Y2H screening was performed using the series of truncated ank-A682L mutants with Skp1 proteins (Fig 10). The results clearly showed that deletion of the F-box-like domain ( $\Delta$ F-box) eliminated binding by the PBCV-1, human and Dictyostelium SKP1 proteins. The site-directed ank-A682L mutant (Leu-Pro<sup>-</sup>) did not bind to any SKP1 homolog further supporting its role as a key binding dipeptide. Similar mutations in the F-box motifs of yeast F-box proteins have been shown to disrupt Skp1 protein interactions in vitro (90). Characteristic of the cellular F-box domain, Leu-Pro residues serve a critical role in the F-box motif, which is involved in Skp1 binding (56). Evidence of  $\Delta$ ANKR1-5 binding demonstrated that at least two ANK domains adjacent to the C-terminal putative F-box region (approximately last 40 residues) are sufficient for virus and heterogeneous Skp1 protein interaction. Thus the interactive behavior of ank-A682L requires the equivalent obligatory domains regardless of origin of the Skp1 binding partner.



**FIG. 9. PBCV-1 ank-A682L deletion mutant constructs.** Five truncated versions of the full length PBCV-1 ank-A682L (ank-A682L<sup>+</sup>) were constructed by making a series of



deletions from the N-terminus ( $\Delta$ ANKR1-3,  $\Delta$ ANKR1-5,  $\Delta$ ANKR1-6,  $\Delta$ ANKR1-7 and  $\Delta$ F-box). Another mutant incorporated all the ANK domains without the C-terminal putative F-box domain ( $\Delta$ F-box). The seven ANK domains of ank-A682L are located at the following residue numbers: 38-70, 71-103, 104-136, 137-169, 170-202, 203-235 and 236-268. The putative F-box domain is orange, ANK domain is black and functional unknown domain is gray.



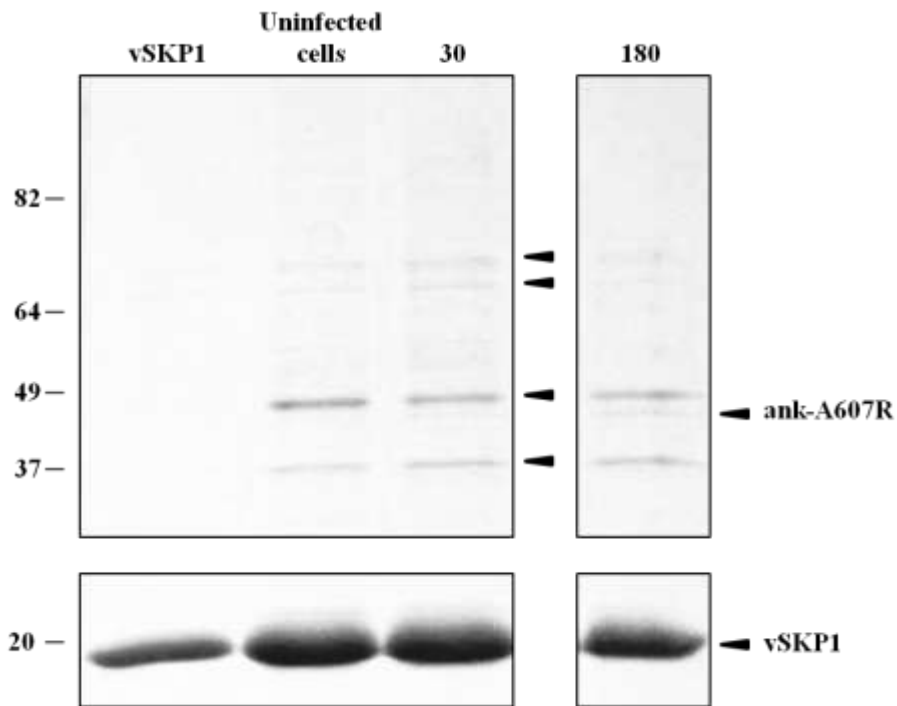
**FIG. 10. PBCV-1 encoded F-box-like protein ank-A682L and Skp1 proteins interact in a putative F-box-dependent manner.** Growth of yeast strains harboring versions of the indicated ank-A682L protein as prey and SKP1 as bait assayed with X-gal, together with their  $\beta$ -galactosidase activity. Truncated forms of the full length ANK protein (ank-

A682L<sup>+</sup>) were constructed by making a series of deletions from the N-terminal. Y2H results indicate that truncated mutants missing the first three and five ANK domains ( $\Delta$ ANKR1-3 and  $\Delta$ ANKR1-5, respectively) retain their ability to interact with the Skp1 proteins from PBCV-1, human and Dictyostelium. Deletion of the F-box-like domain ( $\Delta$ F-box) from the viral ANK protein eliminated binding to all Skp1 members. Site-directed mutant (Leu-Pro<sup>-</sup>) demonstrated its inability to bind to any Skp1 homolog further supporting its role as key binding residues and signature dipeptide of cellular F-box proteins. Empty vectors were used as negative controls. Values are means  $\pm$  SD ( $n=3$ ).

**3.6. Chlorovirus-encoded core ankyrin-repeat family members and vSkp1 proteins interact exclusively within corresponding monophyly clusters.** To test the interactive boundaries of the core ANK proteins, we performed additional Y2H trials using core ANK members from the remaining two chlorovirus types established by their algal host: ATCV-1 ank-Z568R (a SAG-virus) and FR483 ank-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 core ANK protein (Fig 8C). This data replicates the binding nature of vSkp1<sub>PBCV-1</sub> and PBCV-1 ank-A607R. To examine binding exclusivity, we tested the same core ANK proteins against *C. variabilis*, human and Dictyostelium Skp1 proteins. There was no interaction with human and Dictyostelium Skp1 proteins, although core protein ank-A607R from PBCV-1 interacted with *C. variabilis* Skp1 (59724). These results validate the specificity that chlorovirus-encoded core ANK proteins share with their partnered vSkp1 and one of two Skp1 proteins from the host alga.

**3.7. vSkp1<sub>PBCV-1</sub> interacts with endogenous cellular F-box proteins and PBCV-1 ankyrin-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 histidine-tagged vSkp1<sub>PBCV-1</sub> protein and immobilized it on a Ni<sup>2+</sup> column. Three sample lysates were applied separately to the viral Skp1-charged resin: (i) uninfected *C. variabilis* cells, (ii) *C. variabilis* cells at 30 min p.i. and (iii) *C. variabilis* cells at 180 min p.i. The pull-down experiment involved purification of the binding partners of vSkp1<sub>PBCV-1</sub> and their subsequent elution 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 host *C. variabilis* (Fig. 11). These bands were excised, digested and analyzed by inductively coupled plasma mass spectrometry (ICP-MS). All detected proteins corresponded to either cellular F-box proteins from *C. variabilis* or PBCV-1-encoded ANK proteins. A protein of 17 kDa 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 (CHLNCDRAFT\_141219), 393 (CHLNCDRAFT\_134953), 621 (CHLNCDRAFT\_52604) and 739 (CHLNCDRAFT\_133199) amino acid residues, respectively. In the viral infected cell lysates, one protein below marker 49 kDa appeared at 30 min p.i. and increased in intensity at 180 min p.i. This band corresponded to PBCV-1 core ANK protein, ank-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 core ANK protein increased. From this observation, there are two plausible explanations: (i) the chlorovirus ANK proteins are capable of competing with cellular F-box proteins for vSkp1 binding, or (ii) the host transcriptional machinery and protein synthesis is downregulated later in infection.



**FIG. 11.  $\text{Ni}^{2+}$  pull down assay confirms vSkp1<sub>PBCV-1</sub> interaction with F-box and core ANK proteins.** vSkp1<sub>PBCV-1</sub>-charged resin was subjected to three separate lysates from uninfected and PBCV-1-infected *C. variabilis* cells at 30 and 180 minutes p.i. Proteins were eluted from the column and resolved by SDS/PAGE. Binding proteins were excised and analyzed by MS. Pull-down results confirm that vSkp1<sub>PBCV-1</sub> interacts with four endogenous cellular F-box proteins from *C. variabilis* and viral core ANK protein, ank-A607R. Protein identity is designated by locus tag. In ascending molecular weight, the

identified host F-box proteins bound to vSkp1 have locus tag identifiers CHLNCDRAFT\_141219, CHLNCDRAFT\_134953, CHLNCDRAFT\_52604 and CHLNCDRAFT\_133199, respectively.

#### 4. DISCUSSION

The widespread prevalence of F-box proteins reinforces their highly involved participation in intracellular regulation, particularly their important role in substrate recruitment to the SCF complex (6, 37, 95, 119). Representing almost 2.3% of the protein-coding genes in the model plant *A. thaliana* (37, 48), increasing evidence suggests that plants utilize the ubiquitin proteasome system to recognize and combat pathogen invasion (67). Most of the characterized F-box proteins are associated with the cellular SCF complex as the substrate specificity module. F-box proteins are subject to positive selection specifically in the C-terminal substrate-binding domain (108) while anchored to the ligase complex via its F-box domains. Using protein sequence alignment and bioinformatics, we identified eight ANK proteins coded by chlorovirus PBCV-1 which contain N-terminal ankyrin repeats together with C-terminal F-box-like domains. Perhaps more convincing, previous studies have shown that poxviruses encode ANK proteins with the same terminal domain organization as the chloroviruses and that they interact with cellular SCF subunits using this unique C-terminal F-box domain (12, 15, 16, 100, 102, 111). Considering this domain arrangement does not exist in cellular F-box proteins (73), the conservation of these ANK genes in chloroviruses encouraged us to speculate that PBCV-1 F-box-like ANK proteins play an intimate role in SCF manipulation. Y2H screening and pull-down methods identified protein-protein

interactions with SCF-associated homologs. This approach established that vSkp1<sub>PBCV-1</sub> interacts with PBCV-1 ank-A682L and ank-A607R proteins. Moreover, vSkp1<sub>ATCV-1</sub> and vSkp1<sub>FR483</sub> interacted with their respective PBCV-1 ank-A607R homologs, ank-Z568R and ank-N710R. Collectively, our data indicate that chloroviruses encode vSkp1 and ANK proteins that perform partnered-function interactions.

As noted above, the chlorovirus F-box-like domains, like the poxviruses, are located at the C-terminus of ANK proteins which is opposite of cellular F-box proteins (100). Despite the lack of obvious homology of the C-terminal F-box domains, PBCV-1 ANK proteins share key residues with human F-box protein Skp2 that contact cellular Skp1 (Fig. 2). Here, we showed that deleting the putative F-box domain in PBCV-1 ank-A682L resulted in a protein that no longer interacted with its counterpart vSkp1. Serial domain deletions of ank-A682L from the N-terminus further support putative F-box-dependent 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, we identified three ANK proteins which possess Leu-Pro residues in series at the beginning of the domain, a signature dipeptide of cellular F-box proteins. Significantly, replacement of this dipeptide with 2Ala in ank-A682L prevents protein interaction with vSkp1<sub>PBCV-1</sub>. However, it is important to note that one of the viral ANK proteins with putative F-box domains lack this signature Leu-Pro residues and they cannot be excluded as functional F-box proteins (111). Supporting this notion, we showed binding of the same vSkp1<sub>PBCV-1</sub> with another PBCV-1 ANK protein (ank-A607R) whose C-terminus had fewer common residues with a canonical F-box domain, including no Leu-Pro residues.

A comprehensive phylogenetic analysis of PBCV-1 ank-A607R identified a family of ANK proteins coded by all sequenced chloroviruses. This conserved class of ANK proteins, referred to as the core ANK proteins, constitute a newly discovered protein family with one member present in all 41 annotated chloroviruses. Using multiple sequence alignment methods, we showed high sequence identity among the core ANK proteins according to the respective virus families. Further phylogenetic analysis of the ANK and putative F-box domains of the core ANK proteins revealed three monophyletic clades separated according to the three host algae suggesting partnered function tailored to the host alga (Fig. 4). From this result, we speculate that this unique class of ANK genes serves as a repertoire of partnering proteins that function with chlorovirus-encoded vSkp1s. Our results indicate that a degenerative F-box motif resides in these core ANK proteins capable of recognizing the host and vSkp1 proteins exclusively.

In addition to encoding F-box-like ANK proteins, chloroviruses have evolved a candidate virulence factor, a Skp1 homolog, for intercepting the ubiquitin proteasome system. While multiple Skp1 homologs exist in higher organisms (57) including *C. variabilis*, virus-encoded Skp1 genes have only been reported in algal viruses to our knowledge. 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 (1-61 amino acids) and a Skp1 domain that resides in the C-terminal region (68-145 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 (Fig 6). Together, the presence of bipartite domains and monophyly clusters according to virus host suggests that chlorovirus vSkp1 proteins are uniquely expressed to compliment a counterpart core ANK protein thereby mimicking cellular Skp1-F-box binding properties.

As the only chlorovirus that doesn't 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. Furthering speculation, four spontaneously derived, large deletion mutants of PBCV-1 suggest that the vSkp1 is not an essential gene for virus replication in the laboratory (60). 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. However, replication of the deletion mutants was attenuated as their burst sizes were half that of the parent virus and their plaque sizes were consequently smaller. This reduced virulence and plaque phenotype suggests the probable importance of the conserved vSkp1 in chlorovirus genomes as a participating factor in the viral life cycle.

Adding to the complexity of pathogenesis, some viruses, including the chloroviruses, encode their own ubiquitin components to enable replication and evade cellular immune defenses (3, 52). Depending on their needs, viruses have developed means to enhance or inhibit ubiquitylation of specific substrates. Besides the coding of ubiquitin itself by some viral genomes, many viruses encode E3 ligases and de-ubiquitylating enzymes (DUBs), as well as adaptor proteins. For example, in addition to PBCV-1-encoded protein vSkp1 (A39L), transcriptome analysis of PBCV-1 (29) revealed early transcribed enzymes putatively involved in protein degradation including



ubiquitin C-terminal hydrolase (A105L) and SCF-E3 ubiquitin ligase (A481L). Another NC64A virus, NY-2A, and 11 SAG viruses including ATCV-1 also encode an ubiquitin-60S ribosomal protein (Z203L). 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 can be challenging, as is determining their effect on viral replication *in vivo*.

The molecular arms race between host and virus continues to evolve in complexity and pace as viruses uncover new strategies to gain leverage and outwit cellular lines of defense. Many ubiquitin-interfering viral proteins interact directly with ubiquitin ligase family enzymes, in particular the multisubunit SCF complex. Targeting cellular machinery by either avoiding or enhancing specific ubiquitination events has developed into a popular means of host exploitation. Massive viral genomes are continuously being uncovered that challenge the traditional size limitations of genetic material that can be packaged inside a virus particle. Not only do the genomes of several giant DNA viruses expand far beyond the smallest free-living bacterium [e.g., 2.5-Mb Pandoravirus vs. ~0.5-Mb *Mycoplasma genitalium* (35, 83)], the genomic content of recent discovered giant viruses is entirely novel compared to previously annotated organisms. For example, the 2.5 Mb genome of Pandoravirus was recently sequenced and more than 93% of its genes resemble nothing known. However, a comprehensive search of the NCBI nonredundant database for homologs to the 2556 CDSs returned 401 significant matches, of which more than half included ANK and F-box motifs (83).

It is not surprising that these viruses equipped with such an arsenal of novel genes have been theorized to constitute the fourth domain of cellular life (14). Ubiquitous in nature, large DNA viruses like chloroviruses exhibit extraordinary diversity in genome structure and composition further encompassing a unique network of virus families. By tracing gene origin, temporal expression, and endogenous protein-protein interaction, we have gained insight into this extensive pool of viral proteins by identifying interacting chlorovirus-encoded vSkp1 and ANK proteins. Extrapolating from this data, it is clear chloroviruses are equipped with unique viral homologs associated with the ubiquitin proteasome system. Furthermore, the presence of vSkp1 and multiple ANK proteins harboring F-box-like domains suggests that chloroviruses have also evolved a unique mechanism to potentially manipulate eukaryotic ubiquitin ligase machinery in chlorella-like algae. This study is the first to examine chlorovirus-encoded proteins regarding their potential role in manipulating or diverting core components of the ubiquitin ligase system in chlorella-like algae. Identifying the substrates that are degraded by the SCF complex during chlorovirus infection remains the goal of future endeavors.

## **ACKNOWLEDGEMENTS**

I thank James Gurnon and Irina Agarkova for their technical help and maintaining the culture conditions and David Dunigan for helpful discussions. I also thank Jiri Adamec and other members of the Redox Biology Center at UNL for MS analysis. Special thanks to the Nebraska Center for Virology and the School of Biological Sciences at UNL. This research was supported by the National Institutes of Health under Ruth L. Kirschstein National Research Service Award 1 T32 AIO60547 from the National Institute of Allergy and Infectious Diseases.

## 5. REFERENCES

1. **Agarkova, I., D. Dunigan, J. Gurnon, T. Greiner, J. Barres, G. Thiel, and J. L. Van Etten.** 2008. Chlorovirus-mediated membrane depolarization of *Chlorella* alters secondary active transport of solutes. *J Virol* **82**:12181-90.
2. **Al-Khodor, S., C. T. Price, A. Kalia, and Y. Abu Kwaik.** 2010. Functional diversity of ankyrin repeats in microbial proteins. *Trends Microbiol* **18**:132-9.
3. **Albrecht, M., Y. Kuhne, B. K. Ballmer-Weber, W. M. Becker, T. Holzhauser, I. Lauer, A. Reuter, S. Randow, S. Falk, A. Wangorsch, J. Lidholm, G. Reese, and S. Vieths.** 2009. Relevance of IgE binding to short peptides for the allergenic activity of food allergens. *J Allergy Clin Immunol* **124**:328-36, 336 e1-6.
4. **Angers, S., T. Li, X. Yi, M. J. MacCoss, R. T. Moon, and N. Zheng.** 2006. Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature* **443**:590-3.
5. **Aronson, M. N., A. D. Meyer, J. Gyorgyey, L. Katul, H. J. Vetten, B. Gronenborn, and T. Timchenko.** 2000. Clink, a nanovirus-encoded protein, binds both pRB and SKP1. *J Virol* **74**:2967-72.
6. **Bai, C., P. Sen, K. Hofmann, L. Ma, M. Goebel, J. W. Harper, and S. J. Elledge.** 1996. SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* **86**:263-74.
7. **Barry, M., N. van Buuren, K. Burles, K. Mottet, Q. Wang, and A. Teale.** 2010. Poxvirus exploitation of the ubiquitin-proteasome system. *Viruses* **2**:2356-80.

8. **Biedermann, S., and H. Hellmann.** 2011. WD40 and CUL4-based E3 ligases: lubricating all aspects of life. *Trends Plant Sci* **16**:38-46.
9. **Blanc, G., G. Duncan, I. Agarkova, M. Borodovsky, J. Gurnon, A. Kuo, E. Lindquist, S. Lucas, J. Pangilinan, J. Polle, A. Salamov, A. Terry, T. Yamada, D. D. Dunigan, I. V. Grigoriev, J. M. Claverie, and J. L. Van Etten.** 2010. The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell* **22**:2943-55.
10. **Blanc, G., M. Mozar, I. V. Agarkova, J. R. Gurnon, G. Yanai-Balser, J. M. Rowe, Y. Xia, J. J. Riethoven, D. D. Dunigan, and J. L. Van Etten.** 2014. Deep RNA sequencing reveals hidden features and dynamics of early gene transcription in *Paramecium bursaria* chlorella virus 1. *PLoS One* **9**:e90989.
11. **Blanchette, P., and P. E. Branton.** 2009. Manipulation of the ubiquitin-proteasome pathway by small DNA tumor viruses. *Virology* **384**:317-23.
12. **Blanie, S., J. Gelfi, S. Bertagnoli, and C. Camus-Bouclainville.** 2010. MNF, an ankyrin repeat protein of myxoma virus, is part of a native cellular SCF complex during viral infection. *Virol J* **7**:56.
13. **Both, G. W.** 2002. Identification of a unique family of F-box proteins in atadenoviruses. *Virology* **304**:425-33.
14. **Boyer, M., M. A. Madoui, G. Gimenez, B. La Scola, and D. Raoult.** 2010. Phylogenetic and phyletic studies of informational genes in genomes highlight existence of a 4 domain of life including giant viruses. *PLoS One* **5**:e15530.
15. **Buttigieg, K., S. M. Laidlaw, C. Ross, M. Davies, S. Goodbourn, and M. A. Skinner.** 2013. Genetic screen of a library of chimeric poxviruses identifies an

- ankyrin repeat protein involved in resistance to the avian type I interferon response. *J Virol* **87**:5028-40.
16. **Chang, S. J., J. C. Hsiao, S. Sonnberg, C. T. Chiang, M. H. Yang, D. L. Tzou, A. A. Mercer, and W. Chang.** 2009. Poxvirus host range protein CP77 contains an F-box-like domain that is necessary to suppress NF-kappaB activation by tumor necrosis factor alpha but is independent of its host range function. *J Virol* **83**:4140-52.
  17. **Cherrier, M. V., V. A. Kostyuchenko, C. Xiao, V. D. Bowman, A. J. Battisti, X. Yan, P. R. Chipman, T. S. Baker, J. L. Van Etten, and M. G. Rossmann.** 2009. An icosahedral algal virus has a complex unique vertex decorated by a spike. *Proc Natl Acad Sci U S A* **106**:11085-9.
  18. **Ciechanover, A.** 2003. The ubiquitin proteolytic system and pathogenesis of human diseases: a novel platform for mechanism-based drug targeting. *Biochem Soc Trans* **31**:474-81.
  19. **Ciechanover, A., and P. Brundin.** 2003. The ubiquitin proteasome system in neurodegenerative diseases: sometimes the chicken, sometimes the egg. *Neuron* **40**:427-46.
  20. **Connelly, C., and P. Hieter.** 1996. Budding yeast SKP1 encodes an evolutionarily conserved kinetochore protein required for cell cycle progression. *Cell* **86**:275-85.
  21. **Correa, R. L., F. P. Bruckner, R. de Souza Cascardo, and P. Alfenas-Zerbini.** 2013. The Role of F-Box Proteins during Viral Infection. *Int J Mol Sci* **14**:4030-49.

22. **Craig, K. L., and M. Tyers.** 1999. The F-box: a new motif for ubiquitin dependent proteolysis in cell cycle regulation and signal transduction. *Prog Biophys Mol Biol* **72**:299-328.
23. **Cunnac, S., A. Occhialini, P. Barberis, C. Boucher, and S. Genin.** 2004. Inventory and functional analysis of the large Hrp regulon in *Ralstonia solanacearum*: identification of novel effector proteins translocated to plant host cells through the type III secretion system. *Mol Microbiol* **53**:115-28.
24. **Dallaire, F., P. Blanchette, and P. E. Branton.** 2009. A proteomic approach to identify candidate substrates of human adenovirus E4orf6-E1B55K and other viral cullin-based E3 ubiquitin ligases. *J Virol* **83**:12172-84.
25. **DeAngelis, P. L., W. Jing, M. V. Graves, D. E. Burbank, and J. L. Van Etten.** 1997. Hyaluronan synthase of chlorella virus PBCV-1. *Science* **278**:1800-3.
26. **Delaroque, N., D. G. Muller, G. Bothe, T. Pohl, R. Knippers, and W. Boland.** 2001. The complete DNA sequence of the *Ectocarpus siliculosus* Virus EsV-1 genome. *Virology* **287**:112-32.
27. **Dereeper, A., S. Audic, J. M. Claverie, and G. Blanc.** 2010. BLAST-EXPLORER helps you building datasets for phylogenetic analysis. *BMC Evol Biol* **10**:8.
28. **Dereeper, A., V. Guignon, G. Blanc, S. Audic, S. Buffet, F. Chevenet, J. F. Dufayard, S. Guindon, V. Lefort, M. Lescot, J. M. Claverie, and O. Gascuel.** 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**:W465-9.

29. **Dunigan, D. D., R. L. Cerny, A. T. Bauman, J. C. Roach, L. C. Lane, I. V. Agarkova, K. Wulser, G. M. Yanai-Balser, J. R. Gurnon, J. C. Vitek, B. J. Kronschnabel, A. Jeanniard, G. Blanc, C. Upton, G. A. Duncan, O. W. McClung, F. Ma, and J. L. Van Etten.** 2012. Paramecium bursaria chlorella virus 1 proteome reveals novel architectural and regulatory features of a giant virus. *J Virol* **86**:8821-34.
30. **Dunigan, D. D., L. A. Fitzgerald, and J. L. Van Etten.** 2006. Phycodnaviruses: a peek at genetic diversity. *Virus Res* **117**:119-32.
31. **Elde, N. C., and H. S. Malik.** 2009. The evolutionary conundrum of pathogen mimicry. *Nat Rev Microbiol* **7**:787-97.
32. **Filee, J.** 2009. Lateral gene transfer, lineage-specific gene expansion and the evolution of Nucleo Cytoplasmic Large DNA viruses. *J Invertebr Pathol* **101**:169-71.
33. **Fitzgerald, L. A., P. T. Boucher, G. M. Yanai-Balser, K. Suhre, M. V. Graves, and J. L. Van Etten.** 2008. Putative gene promoter sequences in the chlorella viruses. *Virology* **380**:388-93.
34. **Fortune, J. M., O. V. Lavrukhin, J. R. Gurnon, J. L. Van Etten, R. S. Lloyd, and N. Osheroff.** 2001. Topoisomerase II from Chlorella virus PBCV-1 has an exceptionally high DNA cleavage activity. *J Biol Chem* **276**:24401-8.
35. **Fraser, C. M., J. D. Gocayne, O. White, M. D. Adams, R. A. Clayton, R. D. Fleischmann, C. J. Bult, A. R. Kerlavage, G. Sutton, J. M. Kelley, R. D. Fritchman, J. F. Weidman, K. V. Small, M. Sandusky, J. Fuhrmann, D. Nguyen, T. R. Utterback, D. M. Saudek, C. A. Phillips, J. M. Merrick, J. F.**



- Tomb, B. A. Dougherty, K. F. Bott, P. C. Hu, T. S. Lucier, S. N. Peterson, H. O. Smith, C. A. Hutchison, 3rd, and J. C. Venter.** 1995. The minimal gene complement of *Mycoplasma genitalium*. *Science* **270**:397-403.
36. **Fruscione, F., L. Sturla, G. Duncan, J. L. Van Etten, P. Valbuzzi, A. De Flora, E. Di Zanni, and M. Tonetti.** 2008. Differential role of NADP<sup>+</sup> and NADPH in the activity and structure of GDP-D-mannose 4,6-dehydratase from two chlorella viruses. *J Biol Chem* **283**:184-93.
37. **Gagne, J. M., B. P. Downes, S. H. Shiu, A. M. Durski, and R. D. Vierstra.** 2002. The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in *Arabidopsis*. *Proc Natl Acad Sci U S A* **99**:11519-24.
38. **Gao, G., and H. Luo.** 2006. The ubiquitin-proteasome pathway in viral infections. *Can J Physiol Pharmacol* **84**:5-14.
39. **Gietz, D., A. St Jean, R. A. Woods, and R. H. Schiestl.** 1992. Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res* **20**:1425.
40. **Graff, J. W., K. Ettayebi, and M. E. Hardy.** 2009. Rotavirus NSP1 inhibits NFkappaB activation by inducing proteasome-dependent degradation of beta-TrCP: a novel mechanism of IFN antagonism. *PLoS Pathog* **5**:e1000280.
41. **Graves, M. V., D. E. Burbank, R. Roth, J. Heuser, P. L. DeAngelis, and J. L. Van Etten.** 1999. Hyaluronan synthesis in virus PBCV-1-infected chlorella-like green algae. *Virology* **257**:15-23.

42. **Greiner, T., F. Frohns, M. Kang, J. L. Van Etten, A. Kasmann, A. Moroni, B. Hertel, and G. Thiel.** 2009. Chlorella viruses prevent multiple infections by depolarizing the host membrane. *J Gen Virol* **90**:2033-9.
43. **Grimsley, N., B. Pequin, C. Bachy, H. Moreau, and G. Piganeau.** 2010. Cryptic sex in the smallest eukaryotic marine green alga. *Mol Biol Evol* **27**:47-54.
44. **Grossman, A. R.** 2005. Paths toward algal genomics. *Plant Physiol* **137**:410-27.
45. **Haas, A. L., J. V. Warme, A. Hershko, and I. A. Rose.** 1982. Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *J Biol Chem* **257**:2543-8.
46. **Hellmann, H., and M. Estelle.** 2002. Plant development: regulation by protein degradation. *Science* **297**:793-7.
47. **Hershko, A., and A. Ciechanover.** 1998. The ubiquitin system. *Annu Rev Biochem* **67**:425-79.
48. **Hua, Z., C. Zou, S. H. Shiu, and R. D. Vierstra.** 2011. Phylogenetic comparison of F-Box (FBX) gene superfamily within the plant kingdom reveals divergent evolutionary histories indicative of genomic drift. *PLoS One* **6**:e16219.
49. **Iyer, L. M., L. Aravind, and E. V. Koonin.** 2001. Common origin of four diverse families of large eukaryotic DNA viruses. *J Virol* **75**:11720-34.
50. **Iyer, L. M., S. Balaji, E. V. Koonin, and L. Aravind.** 2006. Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Res* **117**:156-84.
51. **Jeanniard, A., D. D. Dunigan, J. R. Gurnon, I. V. Agarkova, M. Kang, J. Vitek, G. Duncan, O. W. McClung, M. Larsen, J. M. Claverie, J. L. Van**

- Etten, and G. Blanc.** 2013. Towards defining the chloroviruses: a genomic journey through a genus of large DNA viruses. *BMC Genomics* **14**:158.
52. **Kahloul, S., I. HajSalah El Beji, A. Boulaflous, A. Ferchichi, H. Kong, S. Mouzeyar, and M. F. Bouzidi.** 2013. Structural, expression and interaction analysis of rice SKP1-like genes. *DNA Res* **20**:67-78.
53. **Kawasaki, T., M. Tanaka, M. Fujie, S. Usami, and T. Yamada.** 2004. Immediate early genes expressed in chlorovirus infections. *Virology* **318**:214-23.
54. **Kipreos, E. T., and M. Pagano.** 2000. The F-box protein family. *Genome Biol* **1**:REVIEWS3002.
55. **Kodama, Y., and M. Fujishima.** 2009. Timing of perialgal vacuole membrane differentiation from digestive vacuole membrane in infection of symbiotic algae *Chlorella vulgaris* of the ciliate *Paramecium bursaria*. *Protist* **160**:65-74.
56. **Kondo-Okamoto, N., K. Ohkuni, K. Kitagawa, J. M. McCaffery, J. M. Shaw, and K. Okamoto.** 2006. The novel F-box protein Mfb1p regulates mitochondrial connectivity and exhibits asymmetric localization in yeast. *Mol Biol Cell* **17**:3756-67.
57. **Kong, H., L. L. Landherr, M. W. Frohlich, J. Leebens-Mack, H. Ma, and C. W. dePamphilis.** 2007. Patterns of gene duplication in the plant SKP1 gene family in angiosperms: evidence for multiple mechanisms of rapid gene birth. *Plant J* **50**:873-85.
58. **Kus, B. M., C. E. Caldon, R. Andorn-Broza, and A. M. Edwards.** 2004. Functional interaction of 13 yeast SCF complexes with a set of yeast E2 enzymes in vitro. *Proteins* **54**:455-67.

59. **Kutish, G. F., Y. Li, Z. Lu, M. Furuta, D. L. Rock, and J. L. Van Etten.** 1996. Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: map positions 182 to 258. *Virology* **223**:303-17.
60. **Landstein, D., D. E. Burbank, J. W. Nietfeldt, and J. L. Van Etten.** 1995. Large deletions in antigenic variants of the chlorella virus PBCV-1. *Virology* **214**:413-20.
61. **Landstein, D., M. V. Graves, D. E. Burbank, P. DeAngelis, and J. L. Van Etten.** 1998. Chlorella virus PBCV-1 encodes functional glutamine: fructose-6-phosphate amidotransferase and UDP-glucose dehydrogenase enzymes. *Virology* **250**:388-96.
62. **Lavrukhin, O. V., J. M. Fortune, T. G. Wood, D. E. Burbank, J. L. Van Etten, N. Osheroff, and R. S. Lloyd.** 2000. Topoisomerase II from Chlorella virus PBCV-1. Characterization of the smallest known type II topoisomerase. *J Biol Chem* **275**:6915-21.
63. **Li, Y., Z. Lu, D. E. Burbank, G. F. Kutish, D. L. Rock, and J. L. Van Etten.** 1995. Analysis of 43 kb of the Chlorella virus PBCV-1 330-kb genome: map positions 45 to 88. *Virology* **212**:134-50.
64. **Li, Y., Z. Lu, L. Sun, S. Ropp, G. F. Kutish, D. L. Rock, and J. L. Van Etten.** 1997. Analysis of 74 kb of DNA located at the right end of the 330-kb chlorella virus PBCV-1 genome. *Virology* **237**:360-77.
65. **Lu, Z., Y. Li, Q. Que, G. F. Kutish, D. L. Rock, and J. L. Van Etten.** 1996. Analysis of 94 kb of the chlorella virus PBCV-1 330-kb genome: map positions 88 to 182. *Virology* **216**:102-23.

66. **Lu, Z., Y. Li, Y. Zhang, G. F. Kutish, D. L. Rock, and J. L. Van Etten.** 1995. Analysis of 45 kb of DNA located at the left end of the chlorella virus PBCV-1 genome. *Virology* **206**:339-52.
67. **Magori, S., and V. Citovsky.** 2011. Hijacking of the Host SCF Ubiquitin Ligase Machinery by Plant Pathogens. *Front Plant Sci* **2**:87.
68. **Margottin, F., S. P. Bour, H. Durand, L. Selig, S. Benichou, V. Richard, D. Thomas, K. Strebel, and R. Benarous.** 1998. A novel human WD protein, h-beta TrCp, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Mol Cell* **1**:565-74.
69. **Matsumoto, M., S. Hatakeyama, K. Oyamada, Y. Oda, T. Nishimura, and K. I. Nakayama.** 2005. Large-scale analysis of the human ubiquitin-related proteome. *Proteomics* **5**:4145-51.
70. **Mayer, B. J.** 2006. Protein-protein interactions in signaling cascades. *Methods Mol Biol* **332**:79-99.
71. **McWilliam, H., W. Li, M. Uludag, S. Squizzato, Y. M. Park, N. Buso, A. P. Cowley, and R. Lopez.** 2013. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res* **41**:W597-600.
72. **Meints, R. H., K. Lee, D. E. Burbank, and J. L. Van Etten.** 1984. Infection of a Chlorella-like alga with the virus, PBCV-1: ultrastructural studies. *Virology* **138**:341-6.
73. **Mercer, A. A., S. B. Fleming, and N. Ueda.** 2005. F-box-like domains are present in most poxvirus ankyrin repeat proteins. *Virus Genes* **31**:127-33.

74. **Mosavi, L. K., T. J. Cammett, D. C. Desrosiers, and Z. Y. Peng.** 2004. The ankyrin repeat as molecular architecture for protein recognition. *Protein Sci* **13**:1435-48.
75. **Murphy, T. D.** 2003. *Drosophila* *skpA*, a component of SCF ubiquitin ligases, regulates centrosome duplication independently of cyclin E accumulation. *J Cell Sci* **116**:2321-32.
76. **Nayak, S., F. E. Santiago, H. Jin, D. Lin, T. Schedl, and E. T. Kipreos.** 2002. The *Caenorhabditis elegans* Skp1-related gene family: diverse functions in cell proliferation, morphogenesis, and meiosis. *Curr Biol* **12**:277-87.
77. **Nelson, M., D. E. Burbank, and J. L. Van Etten.** 1998. Chlorella viruses encode multiple DNA methyltransferases. *Biol Chem* **379**:423-8.
78. **Nelson, M., Y. Zhang, and J. L. Van Etten.** 1993. DNA methyltransferases and DNA site-specific endonucleases encoded by chlorella viruses. *EXS* **64**:186-211.
79. **Noureddine, M. A., T. D. Donaldson, S. A. Thacker, and R. J. Duronio.** 2002. *Drosophila* *Roc1a* encodes a RING-H2 protein with a unique function in processing the Hh signal transducer Ci by the SCF E3 ubiquitin ligase. *Dev Cell* **2**:757-70.
80. **Olia, A. S., S. Casjens, and G. Cingolani.** 2007. Structure of phage P22 cell envelope-penetrating needle. *Nat Struct Mol Biol* **14**:1221-6.
81. **Ovaa, H., B. M. Kessler, U. Rolen, P. J. Galardy, H. L. Ploegh, and M. G. Masucci.** 2004. Activity-based ubiquitin-specific protease (USP) profiling of virus-infected and malignant human cells. *Proc Natl Acad Sci U S A* **101**:2253-8.

82. **Pazhouhandeh, M., M. Dieterle, K. Marrocco, E. Lechner, B. Berry, V. Brault, O. Hemmer, T. Kretsch, K. E. Richards, P. Genschik, and V. Ziegler-Graff.** 2006. F-box-like domain in the poliovirus protein P0 is required for silencing suppressor function. *Proc Natl Acad Sci U S A* **103**:1994-9.
83. **Philippe, N., M. Legendre, G. Doutre, Y. Coute, O. Poirot, M. Lescot, D. Arslan, V. Seltzer, L. Bertaux, C. Bruley, J. Garin, J. M. Claverie, and C. Abergel.** 2013. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* **341**:281-6.
84. **Plugge, B., S. Gazzarrini, M. Nelson, R. Cerana, J. L. Van Etten, C. Derst, D. DiFrancesco, A. Moroni, and G. Thiel.** 2000. A potassium channel protein encoded by chlorella virus PBCV-1. *Science* **287**:1641-4.
85. **Price, C. T., and Y. A. Kwaik.** 2010. Exploitation of Host Polyubiquitination Machinery through Molecular Mimicry by Eukaryotic-Like Bacterial F-Box Effectors. *Front Microbiol* **1**:122.
86. **Proschold, T., T. Darienko, P. C. Silva, W. Reisser, and L. Krienitz.** The systematics of Zoochlorella revisited employing an integrative approach. *Environ Microbiol* **13**:350-64.
87. **Reviriego-Mendoza, M. M., and R. J. Frisque.** 2010. Interaction and co-localization of JC virus large T antigen and the F-box protein beta-transducin-repeat containing protein. *Virology* **410**:119-28.
88. **Rohozinski, J., L. E. Girton, and J. L. Van Etten.** 1989. Chlorella viruses contain linear nonpermuted double-stranded DNA genomes with covalently closed hairpin ends. *Virology* **168**:363-9.

89. **Rowe, J. M., A. Jeanniard, J. R. Gurnon, Y. Xia, D. D. Dunigan, J. L. Van Etten, and G. Blanc.** 2014. Global analysis of *Chlorella variabilis* NC64A mRNA profiles during the early phase of *Paramecium bursaria chlorella virus-1* infection. *PLoS One* **9**:e90988.
90. **Russell, I. D., A. S. Grancell, and P. K. Sorger.** 1999. The unstable F-box protein p58-Ctf13 forms the structural core of the CBF3 kinetochore complex. *J Cell Biol* **145**:933-50.
91. **Sadanandom, A., M. Bailey, R. Ewan, J. Lee, and S. Nelis.** 2012. The ubiquitin-proteasome system: central modifier of plant signalling. *New Phytol* **196**:13-28.
92. **Sasagawa, Y., S. Sato, T. Ogura, and A. Higashitani.** 2007. *C. elegans* RBX-2-CUL-5- and RBX-1-CUL-2-based complexes are redundant for oogenesis and activation of the MAP kinase MPK-1. *FEBS Lett* **581**:145-50.
93. **Schrammeijer, B., E. Risseuw, W. Pansegrau, T. J. Regensburg-Tuink, W. L. Crosby, and P. J. Hooykaas.** 2001. Interaction of the virulence protein VirF of *Agrobacterium tumefaciens* with plant homologs of the yeast Skp1 protein. *Curr Biol* **11**:258-62.
94. **Schroeder, D. C., Y. Park, H. M. Yoon, Y. S. Lee, S. W. Kang, R. H. Meints, R. G. Ivey, and T. J. Choi.** 2009. Genomic analysis of the smallest giant virus--*Feldmannia* sp. virus 158. *Virology* **384**:223-32.
95. **Schulman, B. A., A. C. Carrano, P. D. Jeffrey, Z. Bowen, E. R. Kinnucan, M. S. Finnin, S. J. Elledge, J. W. Harper, M. Pagano, and N. P. Pavletich.** 2000.



- Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature* **408**:381-6.
96. **Schuster, A. M., L. Girton, D. E. Burbank, and J. L. Van Etten.** 1986. Infection of a *Chlorella*-like alga with the virus PBCV-1: transcriptional studies. *Virology* **148**:181-9.
97. **Sigrist, C. J., E. de Castro, L. Cerutti, B. A. CuChe, N. Hulo, A. Bridge, L. Bougueleret, and I. Xenarios.** 2013. New and continuing developments at PROSITE. *Nucleic Acids Res* **41**:D344-7.
98. **Siozios, S., P. Ioannidis, L. Klasson, S. G. Andersson, H. R. Braig, and K. Bourtzis.** 2013. The diversity and evolution of *Wolbachia* ankyrin repeat domain genes. *PLoS One* **8**:e55390.
99. **Skowyra, D., K. L. Craig, M. Tyers, S. J. Elledge, and J. W. Harper.** 1997. F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. *Cell* **91**:209-19.
100. **Sonnberg, S., B. T. Seet, T. Pawson, S. B. Fleming, and A. A. Mercer.** 2008. Poxvirus ankyrin repeat proteins are a unique class of F-box proteins that associate with cellular SCF1 ubiquitin ligase complexes. *Proc Natl Acad Sci U S A* **105**:10955-60.
101. **Sood, R., M. Rao, S. Singhal, and A. Rattan.** 2005. Activity of RBx 7644 and RBx 8700, new investigational oxazolidinones, against *Mycobacterium tuberculosis* infected murine macrophages. *Int J Antimicrob Agents* **25**:464-8.

102. **Sperling, K. M., A. Schwantes, B. S. Schnierle, and G. Sutter.** 2008. The highly conserved orthopoxvirus 68k ankyrin-like protein is part of a cellular SCF ubiquitin ligase complex. *Virology* **374**:234-9.
103. **Strasser, P., Y. P. Zhang, J. Rohozinski, and J. L. Van Etten.** 1991. The termini of the chlorella virus PBCV-1 genome are identical 2.2-kbp inverted repeats. *Virology* **180**:763-9.
104. **Sun, Y., M. Tan, H. Duan, and M. Swaroop.** 2001. SAG/ROC/Rbx/Hrt, a zinc RING finger gene family: molecular cloning, biochemical properties, and biological functions. *Antioxid Redox Signal* **3**:635-50.
105. **Surjit, M., B. Varshney, and S. K. Lal.** The ORF2 glycoprotein of hepatitis E virus inhibits cellular NF-kappaB activity by blocking ubiquitination mediated proteasomal degradation of IkappaBalpha in human hepatoma cells. *BMC Biochem* **13**:7.
106. **Surjit, M., B. Varshney, and S. K. Lal.** 2012. The ORF2 glycoprotein of hepatitis E virus inhibits cellular NF-kappaB activity by blocking ubiquitination mediated proteasomal degradation of IkappaBalpha in human hepatoma cells. *BMC Biochem* **13**:7.
107. **Thiel, H., K. Hleibieh, D. Gilmer, and M. Varrelmann.** 2012. The P25 pathogenicity factor of Beet necrotic yellow vein virus targets the sugar beet 26S proteasome involved in the induction of a hypersensitive resistance response via interaction with an F-box protein. *Mol Plant Microbe Interact* **25**:1058-72.
108. **Thomas, J. H.** 2006. Adaptive evolution in two large families of ubiquitin-ligase adapters in nematodes and plants. *Genome Res* **16**:1017-30.

109. **Tonetti, M., D. Zanardi, J. R. Gurnon, F. Fruscione, A. Armirotti, G. Damonte, L. Sturla, A. De Flora, and J. L. Van Etten.** 2003. Paramecium bursaria Chlorella virus 1 encodes two enzymes involved in the biosynthesis of GDP-L-fucose and GDP-D-rhamnose. *J Biol Chem* **278**:21559-65.
110. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**:4350-4.
111. **van Buuren, N., B. Couturier, Y. Xiong, and M. Barry.** 2008. Ectromelia virus encodes a novel family of F-box proteins that interact with the SCF complex. *J Virol* **82**:9917-27.
112. **Van Etten, J. L.** 2003. Unusual life style of giant chlorella viruses. *Annu Rev Genet* **37**:153-95.
113. **Van Etten, J. L., D. E. Burbank, Y. Xia, and R. H. Meints.** 1983. Growth cycle of a virus, PBCV-1, that infects Chlorella-like algae. *Virology* **126**:117-25.
114. **Van Etten, J. L., and D. D. Dunigan.** 2012. Chloroviruses: not your everyday plant virus. *Trends Plant Sci* **17**:1-8.
115. **Van Etten, J. L., L. C. Lane, and D. D. Dunigan.** 2010. DNA viruses: the really big ones (giruses). *Annu Rev Microbiol* **64**:83-99.
116. **Van Etten, J. L., L. C. Lane, and R. H. Meints.** 1991. Viruses and viruslike particles of eukaryotic algae. *Microbiol Rev* **55**:586-620.
117. **Van Etten, J. L., C. H. Van Etten, J. K. Johnson, and D. E. Burbank.** 1985. A survey for viruses from fresh water that infect a eucaryotic chlorella-like green alga. *Appl Environ Microbiol* **49**:1326-8.

118. **Werden, S. J., and G. McFadden.** 2008. The role of cell signaling in poxvirus tropism: the case of the M-T5 host range protein of myxoma virus. *Biochim Biophys Acta* **1784**:228-37.
119. **Willems, A. R., M. Schwab, and M. Tyers.** 2004. A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. *Biochim Biophys Acta* **1695**:133-70.
120. **Wilson, W. H., J. L. Van Etten, and M. J. Allen.** 2009. The Phycodnaviridae: the story of how tiny giants rule the world. *Curr Top Microbiol Immunol* **328**:1-42.
121. **Yamada, T., T. Fukuda, K. Tamura, S. Furukawa, and P. Songsri.** 1993. Expression of the gene encoding a translational elongation factor 3 homolog of *Chlorella* virus CVK2. *Virology* **197**:742-50.
122. **Yamada, T., and T. Higashiyama.** 1993. Characterization of the terminal inverted repeats and their neighboring tandem repeats in the *Chlorella* CVK1 virus genome. *Mol Gen Genet* **241**:554-63.
123. **Yamada, T., T. Higashiyama, and T. Fukuda.** 1991. Screening of natural waters for viruses which infect *Chlorella* cells. *Appl Environ Microbiol* **57**:3433-7.
124. **Yamada, T., H. Onimatsu, and J. L. Van Etten.** 2006. *Chlorella* viruses. *Adv Virus Res* **66**:293-336.
125. **Yan, X., N. H. Olson, J. L. Van Etten, M. Bergoin, M. G. Rossmann, and T. S. Baker.** 2000. Structure and assembly of large lipid-containing dsDNA viruses. *Nat Struct Biol* **7**:101-3.

126. **Yanai-Balser, G. M., G. A. Duncan, J. D. Eudy, D. Wang, X. Li, I. V. Agarkova, D. D. Dunigan, and J. L. Van Etten.** 2010. Microarray analysis of *Paramecium bursaria* chlorella virus 1 transcription. *J Virol* **84**:532-42.
127. **Yutin, N., and E. V. Koonin.** 2012. Hidden evolutionary complexity of Nucleo-Cytoplasmic Large DNA viruses of eukaryotes. *Virol J* **9**:161.
128. **Yutin, N., Y. I. Wolf, D. Raoult, and E. V. Koonin.** 2009. Eukaryotic large nucleocytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. *Virol J* **6**:223.
129. **Zhang, X., Y. Xiang, D. D. Dunigan, T. Klose, P. R. Chipman, J. L. Van Etten, and M. G. Rossmann.** 2011. Three-dimensional structure and function of the *Paramecium bursaria* chlorella virus capsid. *Proc Natl Acad Sci U S A* **108**:14837-42.
130. **Zhang, Y., F. Maley, G. F. Maley, G. Duncan, D. D. Dunigan, and J. L. Van Etten.** 2007. Chloroviruses encode a bifunctional dCMP-dCTP deaminase that produces two key intermediates in dTTP formation. *J Virol* **81**:7662-71.
131. **Zhao, D., W. Ni, B. Feng, T. Han, M. G. Petrasek, and H. Ma.** 2003. Members of the Arabidopsis-SKP1-like gene family exhibit a variety of expression patterns and may play diverse roles in Arabidopsis. *Plant Physiol* **133**:203-17.
132. **Zheng, N., B. A. Schulman, L. Song, J. J. Miller, P. D. Jeffrey, P. Wang, C. Chu, D. M. Koepp, S. J. Elledge, M. Pagano, R. C. Conaway, J. W. Conaway, J. W. Harper, and N. P. Pavletich.** 2002. Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416**:703-9.

133. **Zheng, N., P. Wang, P. D. Jeffrey, and N. P. Pavletich.** 2000. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* **102**:533-9.
134. **Zielke, N., S. Querings, C. Rottig, C. Lehner, and F. Sprenger.** 2008. The anaphase-promoting complex/cyclosome (APC/C) is required for rereplication control in endoreplication cycles. *Genes Dev* **22**:1690-703.