Chlorella viruses contain genes encoding a complete polyamine biosynthetic pathway

Sascha Baumann[a], **Adrianne Sander**[a], **James R. Gurnon**[b], **Giane Yanai-Balser**[b], **James L. VanEtten**[b], and **Markus Piotrowski**[a]

aDepartment of Plant Physiology, Ruhr-Universität Bochum, 44780 Bochum, Germany

bDepartment of Plant Pathology and Nebraska Center for Virology, University of Nebraska, Lincoln, NE, 68583-0722, USA

Abstract

Two genes encoding the putative polyamine biosynthetic enzymes agmatine iminohydrolase (AIH) and *N*-carbamoylputrescine amidohydrolase (CPA) were cloned from the chloroviruses PBCV-1, NY-2A and MT325. They were expressed in *Escherichia coli* to form C-terminal $(His)_6$-tagged proteins and the recombinant proteins were purified by Ni^{2+}- binding affinity chromatography. The biochemical properties of the two enzymes are similar to AIH and CPA enzymes from *Arabidopsis thaliana* and *Pseudomonas aeruginosa*. Together with the previously known virus genes encoding ornithine/arginine decarboxlyase (ODC/ADC) and homospermidine synthase, the chloroviruses have genes that encode a complete set of functional enzymes that synthesize the rare polyamine homospermidine from arginine via agmatine, *N*-carbamoylputrescine and putrescine. The PBCV-1 *aih* and *cpa* genes are expressed early during virus infection together with the *odc/adc* gene, suggesting that biosynthesis of putrescine is important in early stages of viral replication. The *aih* and *cpa* genes are widespread in the chlorella viruses.

Keywords

Chlorella viruses; Phycodnaviridae; PBCV-1; polyamines; arginine decarboxylase; agmatine iminohydrolase; *N*-carbamoylputrescine amidohydrolase

Introduction

Polyamines are small cationic organic molecules of vital importance for all living organisms. The primary polyamines in eukaryotes are putrescine and the putrescine-derived spermidine and spermine (Fig. 1A). Putrescine is formed either directly by the decarboxylation of ornithine by ornithine decarboxylase (ODC) or indirectly from arginine by the sequential action of arginine decarboxylase (ADC), agmatine iminohydrolase (AIH, also known as agmatine deiminase), and *N*-carbamoylputrescine amidohydrolase (CPA). Both pathways occur in plants although the model plant *Arabidopsis thaliana* lacks an ODC gene (Hanfrey et al., 2001) and is therefore completely dependent on the ADC pathway. Putrescine is converted to spermidine by spermidine synthase (SPDS) and *S*-adenosylmethionine decarboxylase; the latter enzyme

Corresponding author: Dr. Markus Piotrowski, Ruhr-Universität Bochum, Department of, Plant Physiology, ND3, 44780 Bochum, Germany, e-mail: Markus Piotrowski@ruhr-uni-bochum.de, Phone: 0049-(0)234-3224290, Fax: 0049-(0)234-3214187

produces decarboxylated *S*-adenosylmethionine (dcSAM) which serves as the aminopropyl donor for SPDS. Spermine is synthesized in a similar way from spermidine and dcSAM by spermine synthase. Although putrescine, spermidine, and spermine are the main polyamines in eukaryotes, many other polyamines have been identified in nature that are either derived from the putrescine pathway, such as 1,3-diaminopropane and homospermidine, or from other precursors, e.g. cadaverine is derived from lysine (Cohen, 1998a).

Polyamines are not only essential for living cells but they also occur in the capsids of many viruses and, in some cases, are important for virus multiplication (reviewed in Cohen and McCormick, 1979;Walters, 2003). In other systems, host polyamines are important for resistance against viruses (reviewed in Walters, 2003).

The relationship between viruses and polyamines took a new turn with the discovery that chlorella virus *Paramecium bursaria* chlorella virus (PBCV-1) (family *Phycodnaviridae*) has genes encoding the two polyamine biosynthetic enzymes, ODC (Morehead et al., 2002) and homospermidine synthase (HSS) (Kaiser et al., 1999). HSS synthesizes the rare polyamine homospermidine from two molecules of putrescine (Fig. 1A). Although PBCV-1 ODC is able to directly convert ornithine to putrescine, its preferred substrate is arginine (Shah et al., 2004); consequently the PBCV-1 enzyme is referred to as ODC/ADC.

PBCV-1, the prototype chlorella virus, has a 331-kb dsDNA genome that contains 366 putative protein-encoding genes and a polycistronic gene that encodes 11 tRNAs (Van Etten, 2003;Yamada et al., 2006). PBCV-1 and another virus used in this study, NY-2A (370 kb genome), infect *Chlorella* NC64A (NC64A viruses); *Chlorella* NC64A is an endosymbiont of the protozoan *P. bursaria* that was originally isolated in North America. Other viruses, such as MT325 (314 kb genome) infect *Chlorella* Pbi (Pbi viruses), an endosymbiont of *P. bursaria* that was isolated in Europe (Reisser et al., 1986;Reisser et al., 1988).

The genes for the polyamine biosynthetic enzymes AIH and CPA from *A. thaliana* were recently identified (Illingworth et al., 2003;Janowitz et al., 2003;Piotrowski et al., 2003). This identification was due to their similarity to the corresponding enzymes from *Pseudomonas aeruginosa* PAO1, which were identified previously (Nakada et al., 2001;Nakada and Itoh, 2003). Database searches revealed that homologs of both enzymes are also encoded by virus PBCV-1. Here we report the characterization of these genes and their products from PBCV-1, as well as from chlorella viruses, NY-2A and MT325. Our results establish that chlorella viruses encode a complete polyamine biosynthetic pathway that allows the formation of homospermidine from arginine via agmatine, *N*-carbamoylputrescine, and putrescine.

Results

In the following sections we only show results obtained with recombinant AIH and CPA proteins from PBCV-1 genes; results obtained with recombinant AIH and CPA proteins from viruses NY-2A and MT325 will be mentioned in the text or, where appropriate, listed in the tables.

Identification of aih and cpa genes in chlorovirus genomes

BLAST (Altschul et al., 1997) searches using the AIH and CPA sequences from *A. thaliana* revealed homologous genes encoding both enzymes in the virus PBCV-1 genome. The predicted gene products from PBCV-1 open reading frames (ORF) A638R and A78R have 48% and 49% amino acid identity to *At*AIH and *At*CPA, respectively (Table 1). Two other chlorella viruses have been sequenced recently (Fitzgerald et al., 2006a;Fitzgerald et al., 2006b), virus NY-2A that infects the same host as PBCV-1 (*Chlorella* NC64A) and virus MT325 that infects *Chlorella* Pbi. Both of these viruses also contain putative AIH and CPA

encoding genes (Table 1) as well as PBCV-1 ODC/ADC and HSS homologs (Table 3). The PBCV-1 AIH has 97% and 63% amino acid identity to its NY-2A and MT325 homologs, respectively. Similar amino acid identities occur between PBCV-1 CPA and its homologs from NY-2A and MT325 (97% and 68% amino acid identity, respectively). These values are similar to the amino acid identities of PBCV-1 ODC/ADC (86% and 63%) and HSS (93% and 69%) to their respective NY-2A and MT325 homologs (see also Supplement Fig. S1). The viral AIH proteins are 5 to 24 amino acids shorter than the enzymes from *P. aeruginosa* and *A. thaliana*, while all of the CPA proteins are of similar size.

Expression and native mass of viral AIH and CPA proteins

We amplified the *aih* and *cpa* genes from PBCV-1, NY-2A and MT325 genomes by PCR. Recombinant proteins were expressed in *Escherichia coli* with C-terminal $(His)_6$-tags and were subsequently purified by metal affinity chromatography (Fig. 2). AIH proteins from PBCV-1 and NY-2A were obtained as soluble proteins in large amounts (up to 20 mg from 100 ml bacterial cultures), whereas the CPA proteins were expressed in lower amounts (up to 1 mg from 300 ml cultures). In contrast, only small amounts of soluble proteins were obtained from both MT325 genes (about 200 μg per 300 ml culture) because most of the expressed proteins appeared in inclusion bodies.

Chlorella virus AIH and CPA proteins are functional enzymes

A colorimetric ammonia assay was used to test the recombinant proteins for activity because AIH and CPA each release ammonia from their presumed substrates in stoichometric amounts with their products (Fig. 1B). In addition, thin-layer chromatography was used to verify the products of the two reactions as *N*-carbamoylputrescine (for AIH) and putrescine (for CPA). PBCV-1 AIH produced ammonia from agmatine, while at the same time *N*-carbamoylputrescine was formed (Fig. 3A, B, 0-90 min). This *N*-carbamoylputrescine was subsequently converted to putrescine with a concomitant release of more ammonia after addition of PBCV-1 CPA (Fig. 3A, B, 90-240 min). Thus, PBCV-1 AIH and CPA were both functional enzymes displaying their predicted activities. Similar results were obtained with recombinant proteins from NY-2A and MT325 (results not shown). When PBCV-1 AIH was heat denatured before assaying, no release of ammonia and no conversion of agmatine occurred, even after prolonged incubation (Fig. 3B, lane C'). This experiment also establishes that agmatine is not a substrate for PBCV-1 CPA.

Enzyme characterization

The properties of both enzymes (pH and temperature optima, K_m and k_{cat}) were determined and the results are summarized in Table 2. The pH optima of the viral AIH enzymes (pH 7.5) were similar to those from *A. thaliana* and *P. aeruginosa*. In contrast, the viral CPA enzymes had highest activities in slightly acidic conditions (pH 5.5-6.5), when compared to the CPA enzymes from *A. thaliana* and *P. aeruginosa* (pH 8.0-9.0). The temperature optima for both enzymes from the Pbi virus MT325 (20 °C for AIH and 30 °C for CPA) were slightly lower than the two enzymes from the NC64A viruses and from *A. thaliana* and *P. aeruginosa* (Table 2).

The K_m values were similar within each group of enzymes with the exception of the *P. aeruginosa* enzymes, which had higher K_m values for both enzymes. The protein concentrations of the purified MT325 enzymes were so low that they could not be reliably determined; therefore, we did not calculate k_{cat} values for these two enzymes. The k_{cat} values for the other viral enzymes were similar to those obtained from *A. thaliana* and *P. aeruginosa* enzymes.

Substrate specificities

We tested the structurally similar compounds arginine, arginine methyl ester, argininamide, and arcaine (1,4-diguanidinobutane), a derivative of agmatine (1-amino-4-guandinobutane), as possible substrates for PBCV-1 AIH. Under standard conditions (0.5 µg protein/ml, 1 mM substrate, 30 °C, 30 min), no ammonia was released from arginine, arginine methyl ester or argininamide. However, increased amounts of enzyme (80 µg/ml) and prolonged incubation (one and two hr), resulted in very weak activity with arginine methyl ester (0.2 nkat/mg protein). In contrast, arcaine was a relatively good substrate for the viral AIH enzymes (Fig. 4); the specific activity of PBCV-1 AIH with arcaine was only 4 fold less than with agmatine (results not shown). Arcaine has been reported to inhibit plant AIH enzymes (Smith, 1969;Yanagisawa and Suzuki, 1981). These authors reported about 28% to 50% inhibition of maize AIH with 1 or 2 mM arcaine; 2 mM arcaine inhibited *A. thaliana* recombinant AIH by 39% (Janowitz and Piotrowski, unpublished results). Therefore, we were surprised that arcaine was a substrate for the viral AIHs. We tested the plant *At*AIH and also observed release of ammonia from arcaine by this enzyme (results not shown). Although arcaine has two guanidino groups only one molecule of ammonia per molecule of arcaine is released initially by the viral enzymes. However, prolonged incubation with the enzyme results in the release of ammonia that clearly exceeds the molar amount of arcaine (Fig. 4). This result can be explained if the product of the first reaction, 1-carbamoyl-4-guanidinobutane, is also a substrate for the viral AIH but the specific activity of the enzyme for this substrate is lower.

We tested *N*-carbamoyl-D,L-aspartic acid, *N*-carbamoyl-β-alanine, and L-citrulline as substrates for the viral CPA enzymes. However, no ammonia was released from any of these compounds indicating that they are not substrates for the viral CPA enzymes.

Expression of PBCV-1 polyamine biosynthesis genes during infection

The expression of the PBCV-1 *aih* and *cpa* genes during virus replication was monitored by northern blot (results not shown) and by microarray analyses (Fig. 5). Both genes are expressed only within the first 20 to 40 min of infection. Since PBCV-1 DNA replication begins between 60 and 90 min p.i. (Van Etten et al., 1984), *aih* and *cpa* are early expressed genes. Their expression overlaps with expression of the PBCV-1 *odc/adc* gene (ORF A207R) that is also expressed early (Fig. 5 and Morehead et al., 2002). However, the PBCV-1 *odc/adc* gene is also expressed late (from 90 to 360 min) and therefore, also overlaps with expression of the *hss* gene (ORF A237R) that is a typical late gene (Fig. 5). The *aih* and *cpa* transcripts are each about 1600 nt in size as revealed by the northern analyses (results not shown). These are appropriate sizes for the AIH and CPA proteins, which are 359 and 298 amino acids, respectively.

Immediate early genes (detectable expression within the first 5 to 10 min after infection) often have ATGACAA sequences in their predicted promoter regions (Kawasaki et al., 2004). However, this sequence was absent in the putative promoter regions of *aih* and *cpa* genes from all three viruses (data not shown).

Occurrence of aih and cpa genes in other chlorella viruses

To determine if the *aih* and *cpa* genes are common among the chlorella viruses, *a638r* and *a78r* were hybridized to DNA isolated from 37 additional viruses that infect *Chlorella* NC64A and 5 viruses that infect *Chlorella* Pbi. The *a638r* (*aih*) probe hybridized to some extent with all 37 NC64A viruses (Fig. 6A). The *a78r* (*cpa*) probe hybridized to 33 of the NC64A viruses (Fig. 6B). Neither probe hybridized to the host DNA or to the 5 Pbi virus DNAs. These results suggest that the *aih* gene is common in all of the NC64A viruses and that the *cpa* gene might be absent in a few NC64A viruses, i.e. NYs-1, IL-5-2s1, MA-1D, and NY-2B, or that the nucleotide sequence is sufficiently different from the PBCV-1 *cpa* gene so that hybridization

is not detected. The lack of hybridization to the 5 Pbi viruses by both genes is very likely due to differences in their nucleotide sequences.

New sequencing projects have identified *aih* and *cpa* genes, as well as *odc/adc* and *hss* genes, in the NC64A virus AR158 (Fitzgerald et al., 2006b) and in a new chlorella virus, ATCV-1, that infects endosymbiontic chlorella from the heliozoon *Acanthocystis turfacea* (Bubeck and Pfitzner, 2005) (Table 3, unpublished results). However, a newly sequenced Pbi virus, FR483, lacks the *aih* gene, but contains the other three polyamine biosynthetic genes (Fitzgerald et al., 2006a).

Phylogenetic analysis of viral polyamine biosynthesis genes

Neighbor-joining phylogenetic analyses of the polyamine biosynthetic gene products (ADC/ODC, AIH, CPA, and HSS) of PBCV-1, NY-2A, AR158, MT325, and FR483 and of 30-60 homologs from other species resulted in four trees in which the virus sequences formed monophyletic clades with bootstrap support of 100%, indicating that the last common ancestor of the NC64A and Pbi viruses already contained these genes. Additionally, all four trees supported the monophylly of the NC64A and the Pbi viruses with 100% bootstrap support (data not shown and Supplement Figure S2).

Previous studies indicate that the PBCV-1 HSS enzyme is bacterial-like (Kaiser et al., 1999) and phylogenetic analyses depict the PBCV-1 ODC/ADC protein arising near the ancestral origin of the ODC clade (Morehead et al., 2002;Shah et al., 2004). *Aih* and *cpa* genes are common in bacteria and plants; in *Pseudomonas aeruginosa* PAO1 both genes are organized in an operon (Nakada et al., 2001) and they are also clustered in many other bacteria (Supplement Figure S2). However, determining the ancestors of the viral *aih* and *cpa* genes was difficult because of apparent horizontal gene transfer events, indicated by the unusual groupings of sequences from distant bacterial species (e.g. Firmicutes together with γ-Proteobacteria), and by the absence of a clear plant-bacteria split (Supplement Figure S2). Therefore, we cannot unequivocally determine if the viral *aih* and *cpa* genes are of bacterial or plant origin.

Polyamine biosynthesis genes have only been found in chlorella viruses and are absent in viruses from the other five *Phycodnaviridae* genera. Therefore, the chlorella viruses either acquired these genes after separation from the other phycodnaviruses or the genes were lost by the other phycodnaviruses during evolution. It is interesting that the G+C content of the four genes encoding the polyamine biosynthetic enzymes is always higher than the mean G+C content of their respective viral genomes (Table 4). This higher G+C content is especially striking for the *cpa* gene. Deviation of the G+C content of a gene from the mean G+C is often interpreted as a sign of a recent horizontal gene transfer (Lawrence and Ochman, 1997). Therefore, maybe the chlorovirus polyamine biosynthesis genes were acquired after separation from the other phycodnaviruses.

Discussion

The importance of polyamines in virus replication has often been discussed in the literature (reviewed in Cohen, 1998b;Cohen and McCormick, 1979) and these discussions can be summarized as follows: i) Polyamines, primarily putrescine, spermidine and spermine, are present in large amounts in the capsids of some viruses where they can neutralize up to 50% of the viral nucleic acid (e.g. Ames and Dubin, 1960;Cohen, 1998b;Gibson and Roizman, 1971;Tyms et al., 1990). In contrast, other viruses are devoid of polyamines (e.g. Nickerson and Lane, 1977). ii) Virus infection often leads to an increase in polyamine biosynthesis and consequently, in the polyamine content of infected cells (e.g. Goldstein et al., 1976;Torget et al., 1979). iii) Depletion of intracellular polyamine pools (e.g. by using inhibitors of polyamine

biosynthesis) inhibits the replication of some viruses but not others (e.g. Gibson et al., 1984;Greco et al., 2005;Pohjanpelto et al., 1988;Raina et al., 1981).

These previous studies assumed that the polyamines associated with virus replication are synthesized by host enzymes. However, this assumption is no longer valid for all viruses because two genes encoding functional polyamine biosynthetic enzymes were recently discovered in the chlorella virus PBCV-1, an ODC/ADC (Morehead et al., 2002;Shah et al., 2004) and a HSS (Kaiser et al., 1999). As reported here, PBCV-1 and at least two other chlorella viruses, NY-2A and MT325, have two additional genes that encode functional AIH and CPA enzymes. Thus, PBCV-1, and most likely NY-2A and MT325, encode a complete ADC pathway that converts arginine to putrescine (ODC/ADC proteins from NY-2A and MT325 have not been tested for activity). Consequently, these viruses are not dependent on their hosts for polyamine biosynthesis.

The fact that all four polyamine biosynthetic encoding genes are present in most chloroviruses, as shown by Southern blot analyses (Fig. 6; Kaiser et al., 1999;Morehead et al., 2002) and genome-sequencing projects (Table 3), suggests that these proteins are important for chlorella virus replication. However, at least one chlorella virus, FR483, which infects *Chlorella* Pbi lacks an *aih* gene (Fitzgerald et al., 2006a). The other three genes encoding polyamine biosynthetic enzymes are present in all chlorella viruses sequenced to date. *Chlorovirus* is the only genus within the family *Phycodnaviridae*, currently consisting of 6 genera (Wilson et al., 2005), with genes encoding polyamine biosynthetic enzymes; thus, either polyamines are especially important for chlorella virus replication or the chlorella host, in contrast to other algal hosts, is unable to provide the necessary amounts of polyamines for virus replication. The results also indicate that the chlorella viruses either acquired these enzymes after separation from the other phycodnaviruses or that all of the phycodnaviruses originally had the polyamine biosynthetic genes and that they have been selectively lost by the other phycodnaviruses during evolution. However, one should keep in mind that only a few other phycodnavirus genomes have been sequenced (Dunigan et al., 2006) and so the concept that only chlorella viruses encode these enzymes could change as more algal virus genomes are sequenced. It may not be relevant, but one distinction between the chlorella viruses and members of the five other phycodnavirus genera is that the chlorella viruses exist in fresh water and the other viruses occur in marine environments.

Unfortunately, not much is known about polyamine metabolism in chlorellae. ADC and ODC activities were reported in *Chlorella vulgaris* Beijerinck (Cohen et al., 1983;Cohen et al., 1984). The related green alga *Chlamydomonas reinhardtii* only uses the ODC pathway to synthesize putrescine (Theiss et al., 2002). However, although *Chlamydomonas* lacks an *adc* gene it has *aih* and *cpa* genes (available at: http://genome.jgi-psf.org/Chlre3/Chlre3.home.html), but, according to EST data, they are not expressed. It is unknown if *Chlorella* NC64A and *Chlorella* Pbi possess a complete ADC pathway consisting of ADC, AIH and CPA; if they do the host chlorella AIH may functionally replace the absent viral AIH in FR483 during infection. The genome of *Chlorella* NC64A, the host for PBCV-1 and NY-2A, is currently being sequenced and annotation of its sequence will provide information on its polyamine biosynthetic capability. However, it is unlikely that the four PBCV-1 polyamine biosynthetic genes are acquired from *Chlorella* NC64A because none of the four virus genes hybridize with *Chlorella* NC64A DNA (Fig. 6 A, B;Kaiser et al., 1999;Morehead et al., 2002).

PBCV-1 *aih* and *cpa* genes are transcribed as early genes, appearing within the first 20 min p.i. Their expression overlaps with transcription of the PBCV-1 *odc/adc* gene so that all three putrescine biosynthetic enzymes are presumably present early in infection. Therefore, one might expect the polyamine levels to increase during the first 60 min of PBCV-1 infection.

However, little change occurs in either the polyamine concentration or its composition (putrescine, cadaverine, spermidine, and homospermidine are present) during the first 60 min of virus infection (Kaiser et al., 1999). By 240 min p.i., the concentration of putrescine increases about 3.5 times, whereas the other polyamines decrease during this time. The net result is that the total polyamine concentration decreases slightly during virus replication (Kaiser et al., 1999). It is unlikely that polyamines play a role in the neutralization of PBCV-1 DNA in the capsid because the number of polyamine molecules per PBCV-1 virion is so low that they can only neutralize ~0.2% of the virus phosphate residues (Kaiser et al., 1999). Furthermore, the physiological significance of the polyamines in the PBCV-1 particles must be limited because they are only loosely associated with the virions, i.e. they can be replaced by Tris or displaced by washing the particles in a polyamine-free buffer without affecting virus infectivity (Kaiser et al., 1999).

The dual expression pattern (early and late) of the viral genes encoding the polyamine biosynthetic enzymes may indicate that putrescine (early in viral replication) and homospermidine (late in viral replication) have different functions in virus replication. It should be mentioned that the HSS protein, but not the other three polyamine biosynthetic enzymes, is a major component of PBCV-1 virions (Dunigan et al., manuscript in preparation). Therefore, the HSS protein may also serve a structural role in the virus life cycle.

In summary, most of the chlorella viruses have four genes that encode functional enzymes involved in polyamine biosynthesis. However, the biological function(s) of these enzymes and their products is currently an enigma.

Materials and methods

Cloning of viral genes for AIH and CPA

Isolation of viral DNAs has been described (Van Etten et al., 1983). The ORFs for the AIH and CPA proteins were amplified by PCR using the DyNAazyme EXT DNA polymerase (Finnzymes, Espoo, Finland). Each forward primer begins with an *Nde*I restriction site including an in frame start codon (CAT<u>ATG</u>), the reverse primers omit the stop codon and end with *Xho*I restriction sites for in frame cloning with the (His)$_6$-tag of the pET-21b(+) vector (Novagen, Nottingham, UK). If *Xho*I sites were present in the ORF, *Xho*I-compatible *Sal*I restriction sites were used instead. PCR products were T/A cloned into pGEM-T (Promega, Madison, WI). After establishing that the PCR products were free of error by sequencing (MWG Biotech, Ebersberg, Germany), the ORFs were cloned into the *Nde*I and *Xho*I sites of pET-21b(+).

The accession numbers for the virus genomes are **U42580**, **DQ491002**, **DQ491003**, **DQ491001**, **DQ890022**, and "bankit 840993" (provisional accession number) for viruses PBCV-1, NY-2A, AR158, MT325, FR483, and ATCV-1, respectively. The DNA sequences can also be found at http://greengene.uml.edu.

Expression of viral proteins in E. coli and purification of the (His)$_6$-tagged proteins

The proteins were expressed in *E. coli* strain CodonPlus (DE3) RIL (Stratagene, La Jolla, CA). Pre-cultures were grown overnight in 2YT medium (Ausubel et al., 2004) containing ampicillin (100 μg ml^{-1}) and chloramphenicol (30 μg ml^{-1}) in an orbital shaker (220 rpm) at 37°C. The expression cultures were inoculated 1:10 in 2YT medium (including ampicillin and chloramphenicol) with the pre-cultures. Optimized expression conditions to produce the recombinant proteins are described in the Supplement Table S3.

Bacteria were harvested by centrifugation and stored frozen at -80 °C. For protein purification the cell pellets were resuspended in 1/10 of the culture volume of lysis buffer (50 mM sodium

phosphate, pH 8.0, 300 mM NaCl, 5 mM 2-mercaptoethanol, 1 mg ml^{-1} lysozyme) and incubated for 15 min on ice. Cell lysis was completed by sonification with an ultrasound tip for 4×1 min at 50 W (Sonifier B17, Branson, USA). After removing the insoluble material by centrifugation (Sorvall GSA rotor, 10,000 rpm, 10 min), the supernatant was loaded onto a Ni^{2+}-NTA agarose (Qiagen, Hilden, Germany) column (bed volume 2 ml), which was equilibrated in 50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 10 mM imidazole. The column was washed with 50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 30 mM imidazole, and bound proteins were eluted in a total volume of 2.5 ml with 50 mM sodium phosphate, pH 8.0, 300 mM NaCl, and 250 mM imidazole. After desalting on PD-10 columns (GE Healthcare, Freiburg, Germany) into 50 mM potassium phosphate, pH 8.0 and 1 mM DTT (for the CPA proteins) or 50 mM potassium phosphate, pH 8.0 and 0.5 mM DTT (for AIH proteins), the purified recombinant proteins were frozen in 0.5 ml aliquots in liquid nitrogen and stored at -80 °C. Under these conditions the enzymes were stable for at least two months.

Enzyme assays

Enzymes assays were conducted in triplicates with a background control in which the protein was inactivated by boiling for 10 min. The typical CPA assay was performed in 50 mM MES, pH 6.0, 1 mM DTT, 1 mM *N*-carbamoylputrescine and 1 μg of protein in a total volume of 0.5 ml at 37 °C. The AIH assays consisted of 50 mM potassium phosphate, pH 7.5, 0.5 mM DTT, 1 mM agmatine and 1 μg protein in a total volume of 1 ml and were incubated at 30 °C.

Activity was measured as the release of ammonia from the substrates (Fig. 1B). Ammonia was detected with the indophenolblue method: a 100 μl aliquot of sample was mixed with 100 μl of 0.33 M sodium phenolate, 0.02 M sodium hypochlorite (prepared fresh daily), and 0.01% (w/v) sodium nitroprusside. The reaction was incubated for 3 min in a boiling water bath and then 600 μl of water was added. The absorbance at 640 nm was determined and compared to a calibration curve made with NH$_4$Cl.

Thin-layer chromatography was conducted on silica plates (Polygram SIL 201 G/UV254, Macherey-Nagel, Düren, Germany) that were developed in ethanol/NH$_4$OH (25%) 1:1. Five to 10 μl of sample was applied to the plates. Amines were detected by spraying with ninhydrin (0.2% in ethanol).

RNA isolation

Infected chlorella cells (m.o.i. of 5) were collected at 20, 40, 60, 90, 120, 240, 360 min p.i. Cells were disrupted with glass beads in the presence of Trizol (Invitrogen, Carlsbad, CA) and RNA was isolated using an Absolutely RNA Miniprep Kit (Stratagene, LaJolla, CA) according to the manufacturer's instructions. RNA integrity was verified in denaturing 1% agarose gels where intact host cytoplasmic and chloroplast rRNAs were monitored.

Microarray and dot blot analyses

A microarray containing 50-mer probes representing each ORF in the PBCV-1 genome was constructed by MWG Biotech (Ebersberg, Germany) and the Microarray Core Facility (University of Nebraska Medical Center). Four replicas of the entire genome are present in each array. For each time point, 20 μg of total RNA was reverse-transcribed using oligo(dT) as primers and cDNA was labeled with Cy3 or Cy5-dUTP (GE Healthcare, Piscataway, NJ) with the aid of a SuperScript Indirect cDNA Labeling System (Invitrogen, Carlsbad, CA) following the supplier's directions. Competitive hybridization experiments were conducted for each time point against a pool of transcripts representing every gene present in the time course

Results from 3 independent biological hybridizations were analyzed using the GenePix Pro v. 6.0 software (Molecular Devices, Sunnyvale, CA) and TIGR microarray software suite (TM4)

(Saeed et al., 2003). A number of transformations were performed in order to eliminate low quality data, to normalize the measured intensities using the Lowess algorithm, and to regularize the standard deviation of the intensity of the Cy3/Cy5 ratio across the blocks. Genes that displayed statistically significant modulation were identified by a one-way analysis of variance, using P values of <0.01 as a cutoff. Genes with similar expression profiles were grouped into 10 different clusters using a K-means algorithm.

Viral DNAs used for dot blots were denatured, applied to nylon membranes fixed by UV cross-linking, and hybridized with ^{32}P-labelled gene probes as described previously (Sun et al., 2000).

Phylogenetic analysis

The amino acid sequences of PBCV-1 polyamine biosynthesis enzymes were used to search the "non-redundant" dataset of GenBank using blastp (Altschul et al., 1997). All sequences showing similarities greater than 50% over the full sequence were collected (between 30 and 130 entries) and multiple entries from the same species were discarded. The remaining 30-70 sequences were aligned using the DNAMAN program (version 5.2.2, Lynnon Corp., Vaudreuil-Dorion, Canada) and sequences which obviously altered the alignment were discarded. A new alignment was produced with the remaining sequences, extremely diverse N- and C-terminal sequences were deleted, and neighbor-joining trees using Kimura distances and 1000 bootstrap replicates were constructed with the DNAMAN software. These trees were used to discard additional sequences in overrepresented monophyletic clades with high bootstrap support. To root the trees paralogous sequences were included (arginine deiminase for the AIH sequences and *N*-carbamoyl-β-alanine amidohydrolase for the CPA sequences). Again a new alignment was created with the remaining sequences (43 and 57 for the AIHs and CPAs, respectively). From these alignments phylogenetic inference was performed by Bayesian analysis using the Mr.Bayes software (version 3.1.2; Ronquist and Huelsenbeck, 2003) (gamma rates, model: mixed, 250.000 generations, burnin: 25% of saved trees).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–3402. [PubMed: 9254694]

Ames BN, Dubin DT. The role of polyamines in the neutralization of bacteriophage deoxyribonucleic acid. J Biol Chem 1960;235:769–775. [PubMed: 13793161]

Ausubel, FM.; Brent, R.; Kingston, RE.; Moore, DD.; Seidman, JG.; Struhl, K. Current protocols in molecular biology. John Wiley & Sons; New York: 2004.

Bubeck JA, Pfitzner AJ. Isolation and characterization of a new type of chlorovirus that infects an endosymbiotic *Chlorella* strain of the heliozoon *Acanthocystis turfacea*. J Gen Virol 2005;86:2871–2877. [PubMed: 16186243]

Cohen E, Arad SM, Heimer YH, Mizrahi Y. Polyamine biosynthetic enzymes in the cell cycle of *Chlorella*: correlation between ornithine decarboxylase and DNA synthesis at different light intensities. Plant Physiol 1984;74:385–388. [PubMed: 16663427]

Cohen E, Arad SM, Heimer YM, Mizrahi Y. Polyamine biosynthetic enzymes in *Chlorella*: characterization of ornithine and arginine decarboxylase. Plant Cell Physiol 1983;24:1003–1010.

Cohen, SS. A guide to the polyamines. Oxford University Press Inc; New York: 1998a.

Cohen, SS. A guide to the polyamines. Oxford University Press Inc; New York: 1998b. Viruses; p. 366–395.

Cohen SS, McCormick FP. Polyamines and virus multiplication. Adv Virus Res 1979;24:331–387. [PubMed: 389006]

Dunigan DD, Fitzgerald LA, Van Etten JL. Phycodnaviruses: a peek at genetic diversity. Virus Res 2006;117:119–132. [PubMed: 16516998]

Fitzgerald LA, Graves MV, Li X, Feldblyum T, Hartigan J, Van Etten JL. Sequence and annotation of the 314-kb MT325 and the 321-kb FR483 viruses that infect *Chlorella* Pbi. Virology. 2006ain press

Fitzgerald LA, Graves MV, Li X, Feldblyum T, Nierman WC, Van Etten JL. Sequence and annotation of the 369-kb NY-2A and the 345-kb AR158 viruses that infect *Chlorella* NC64A. Virology. 2006bin press

Gibson W, Roizman B. Compartmentalization of spermine and spermidine in the herpes simplex virion. Proc Natl Acad Sci U S A 1971;68:2818–2821. [PubMed: 5288261]

Gibson W, van Breemen R, Fields A, LaFemina R, Irmiere A. D,L-alpha-difluoromethylornithine inhibits human cytomegalovirus replication. J Virol 1984;50:145–154. [PubMed: 6321786]

Goldstein DA, Heby O, Marton LJ. Biphasic stimulation of polyamine biosynthesis in primary mouse kidney cells by infection with polyoma virus:uncoupling from DNA and rRNA synthesis. Proc Natl Acad Sci U S A 1976;73:4022–4026. [PubMed: 186774]

Greco A, Calle A, Morfin F, Thouvenot D, Cayre M, Kindbeiter K, Martin L, Levillain O, Diaz JJ. S-adenosyl methionine decarboxylase activity is required for the outcome of herpes simplex virus type 1 infection and represents a new potential therapeutic target. FASEB J 2005;19:1128–1130. [PubMed: 15863396]

Hanfrey C, Sommer S, Mayer MJ, Burtin D, Michael AJ. *Arabidopsis* polyamine biosynthesis: absence of ornithine decarboxylase and the mechanism of arginine decarboxylase activity. Plant J 2001;27:551–560. [PubMed: 11576438]

Illingworth C, Mayer MJ, Elliott K, Hanfrey C, Walton NJ, Michael AJ. The diverse bacterial origins of the *Arabidopsis* polyamine biosynthetic pathway. FEBS Lett 2003;549:26–30. [PubMed: 12914918]

Janowitz T, Kneifel H, Piotrowski M. Identification and characterization of plant agmatine iminohydrolase, the last missing link in polyamine biosynthesis of plants. FEBS Lett 2003;544:258–261. [PubMed: 12782327]

Kaiser A, Vollmert M, Tholl D, Graves MV, Gurnon JR, Xing W, Lisec AD, Nickerson KW, Van Etten JL. Chlorella virus PBCV-1 encodes a functional homospermidine synthase. Virology 1999;263:254–262. [PubMed: 10544099]

Kawasaki T, Tanaka M, Fujie M, Usami S, Yamada T. Immediate early genes expressed in chlorovirus infections. Virology 2004;318:214–223. [PubMed: 14972549]

Lawrence JG, Ochman H. Amelioration of bacterial genomes: rates of change and exchange. J Mol Evol 1997;44:383–397. [PubMed: 9089078]

Morehead TA, Gurnon JR, Adams B, Nickerson KW, Fitzgerald LA, Van Etten JL. Ornithine decarboxylase encoded by chlorella virus PBCV-1. Virology 2002;301:165–175. [PubMed: 12359457]

Nakada Y, Itoh Y. Identification of the putrescine biosynthetic genes in *Pseudomonas aeruginosa* and characterization of agmatine deiminase and *N*-carbamoylputrescine amidohydrolase of the arginine decarboxylase pathway. Microbiology 2003;149:707–714. [PubMed: 12634339]

Nakada Y, Jiang Y, Nishijyo T, Itoh Y, Lu CD. Molecular characterization and regulation of the *aguBA* operon, responsible for agmatine utilization in *Pseudomonas aeruginosa* PAO1. J Bacteriol 2001;183:6517–6524. [PubMed: 11673419]

Nickerson KW, Lane LC. Polyamine content of several RNA plant viruses. Virology 1977;81:455–459. [PubMed: 898668]

Piotrowski M, Janowitz T, Kneifel H. Plant C-N hydrolases and the identification of a plant *N*-carbamoylputrescine amidohydrolase involved in polyamine biosynthesis. J Biol Chem 2003;278:1708–1712. [PubMed: 12435743]

Pohjanpelto P, Sekki A, Hukkanen V, von Bonsdorff CH. Polyamine depletion of cells reduces the infectivity of herpes simplex virus but not the infectivity of Sindbis virus. Life Sci 1988;42:2011–2018. [PubMed: 2835567]

Raina A, Tuomi K, Mantyjarvi R. Roles of polyamines in the replication of animal viruses. Med Biol 1981;59:428–432. [PubMed: 6279979]

Reisser W, Becker B, Klein T. Studies on ultrastructure and host range of a *Chlorella* attacking virus. Protoplasma 1986;135:162–165.

Reisser W, Burbank DE, Meints SM, Meints RH, Becker B, Van Etten JL. A comparison of viruses infecting two different *Chlorella*-like green algae. Virology 1988;167:143–149. [PubMed: 2847410]

Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 2003;19:1572–1574. [PubMed: 12912839]

Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J. TM4: a free, open-source system for microarray data management and analysis. Biotechniques 2003;34:374–378. [PubMed: 12613259]

Shah R, Coleman CS, Mir K, Baldwin J, Van Etten JL, Grishin NV, Pegg AE, Stanley BA, Phillips MA. *Paramecium bursaria* chlorella virus-1 encodes an unusual arginine decarboxylase that is a close homolog of eukaryotic ornithine decarboxylases. J Biol Chem 2004;279:35760–35767. [PubMed: 15190062]

Smith TA. Agmatine iminohydrolase in maize. Phytochemistry 1969;8:2111–2117.

Sun L, Gurnon JR, Adams BJ, Graves MV, Van Etten JL. Characterization of a β-1,3-glucanase encoded by chlorella virus PBCV-1. Virology 2000;276:27–36. [PubMed: 11021991]

Theiss C, Bohley P, Voigt J. Regulation by polyamines of ornithine decarboxylase activity and cell division in the unicellular green alga *Chlamydomonas reinhardtii*. Plant Physiol 2002;128:1470–1479. [PubMed: 11950995]

Torget R, Lapi L, Cohen SS. Synthesis and accumulation of polyamines and S-adenosylmethionine in Chinese cabbage infected by turnip yellow mosaic virus. Biochem Biophys Res Commun 1979;87:1132–1139. [PubMed: 465028]

Tyms, AS.; Clarke, JR.; Ford, L.; Eloranta, T. Polyamines and the growth of pathogenic viruses. In: Goldemberg, SH.; Algranati, ID., editors. The Biology and Chemistry of Polyamines. IRL Press; Oxford: 1990. p. 122-133.

Van Etten JL. Unusual life style of giant chlorella viruses. Annu Rev Genet 2003;37:153–195. [PubMed: 14616059]

Van Etten JL, Burbank DE, Joshi J, Meints RH. DNA synthesis in a *Chlorella*-like alga following infection with the virus PBCV-1. Virology 1984;134:443–449.

Van Etten JL, Burbank DE, Xia Y, Meints RH. Growth cycle of a virus, PBCV-1, that infects *Chlorella*-like algae. Virology 1983;126:117–125.

Walters DR. Polyamines and plant disease. Phytochemistry 2003;64:97–107. [PubMed: 12946408]

Wilson, WH.; Van Etten, JL.; Schroeder, DS.; Nagasaki, K.; Burssaard, C.; Delaroque, N.; Bratbak, G.; Suttle, C. Phycodnaviridae. In: Fauqet, CM.; Mayo, MAMJ.; Desselberger, U.; Ball, LA., editors. Virus Taxonomy, VIIIth Report of the ICTV. Elsevier/Academic Press; London: 2005. p. 163-175.

Yamada T, Onimatsu H, Van Etten JL. Chlorella viruses. Adv Virus Res 2006;66:293–336. [PubMed: 16877063]

Yanagisawa H, Suzuki Y. Corn agmatine iminohydrolase. Purification and properties. Plant Physiol 1981;67:697–700. [PubMed: 16661738]
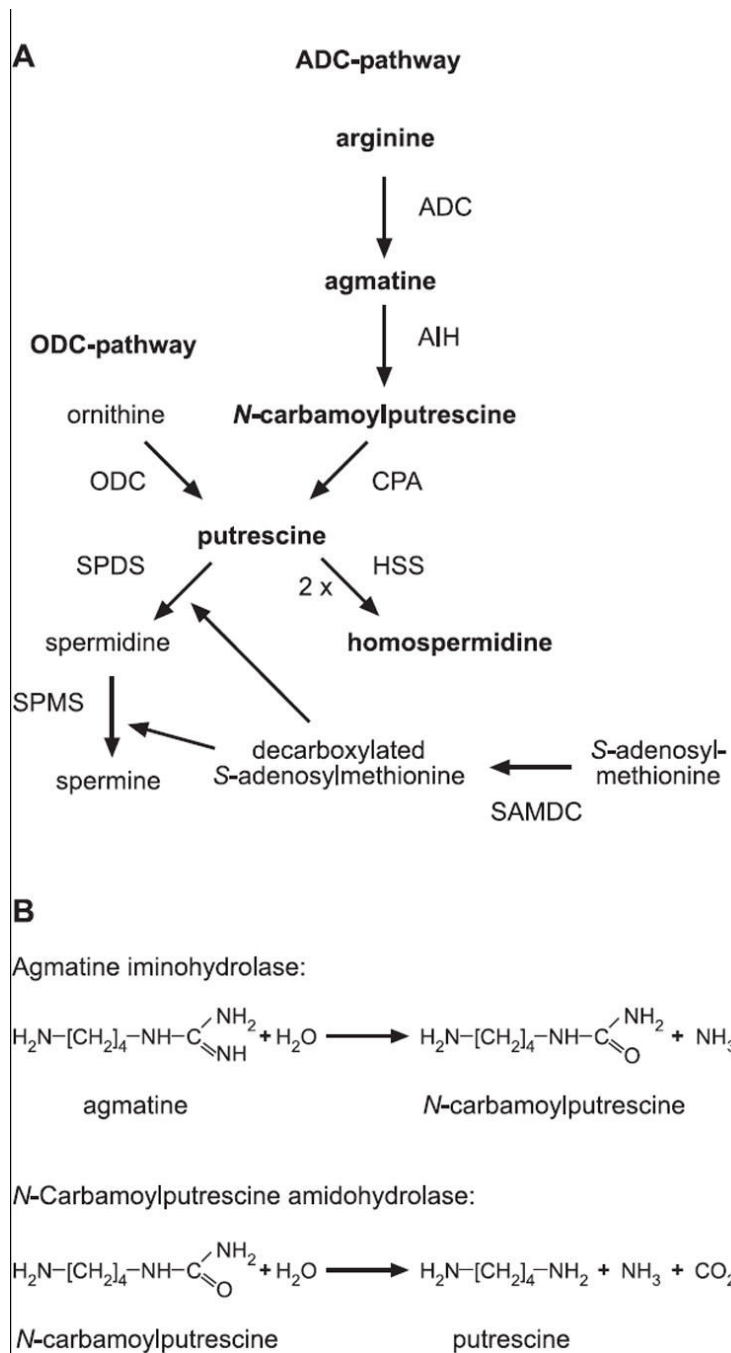
Fig 1. Polyamine biosynthesis pathways
(A) Overview. (B) Details of the reactions catalyzed by agmatine iminohydrolase and *N*-carbamoylputrescine amidohydrolase. ADC, arginine decarboxylase; AIH, agmatine iminohydrolase; CPA, *N*-carbamoylputrescine amidohydrolase; HSS, homospermidine synthase, ODC, ornithine decarboxylase; SAMDC, *S*-adenosylmethionine decarboxylase; SPDS, spermidine synthase; SPMS, spermine synthase.
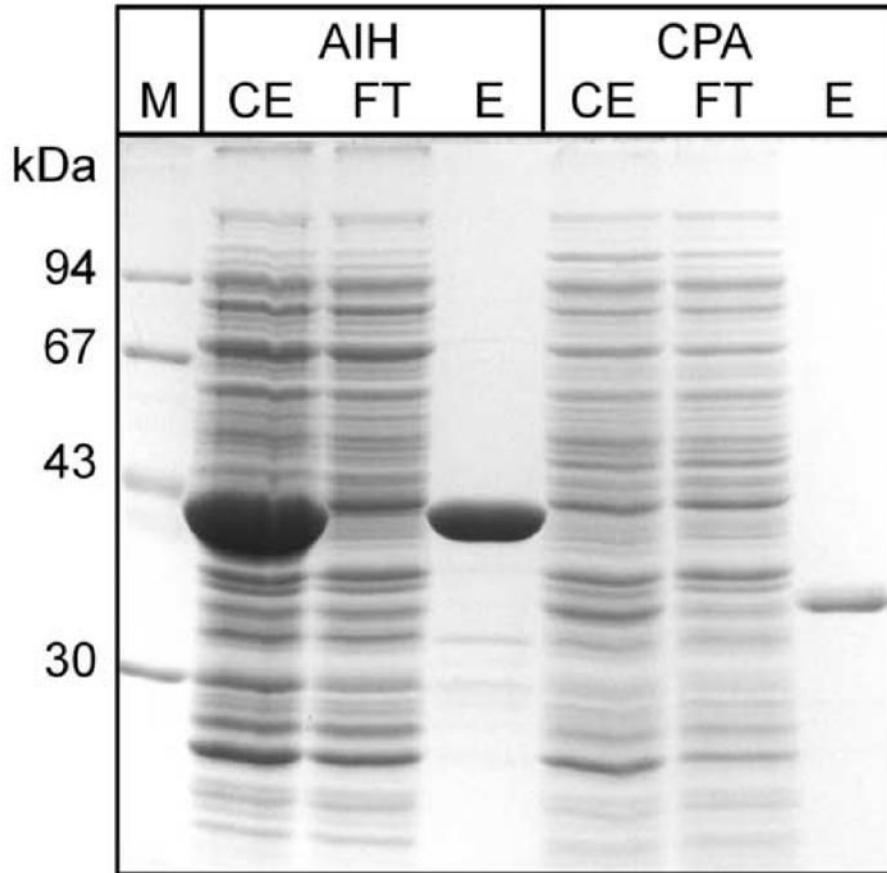
Fig 2. Expression and purification of (His)$_6$-tagged PBCV-1 AIH and CPA from *E. coli*
Ten μl each of bacterial crude extracts (CE) and the column flowthrough (FT) were loaded on the gel. Two and 10 μl of the AIH and CPA eluates (E) were loaded, respectively. M: low-molecular weight markers.
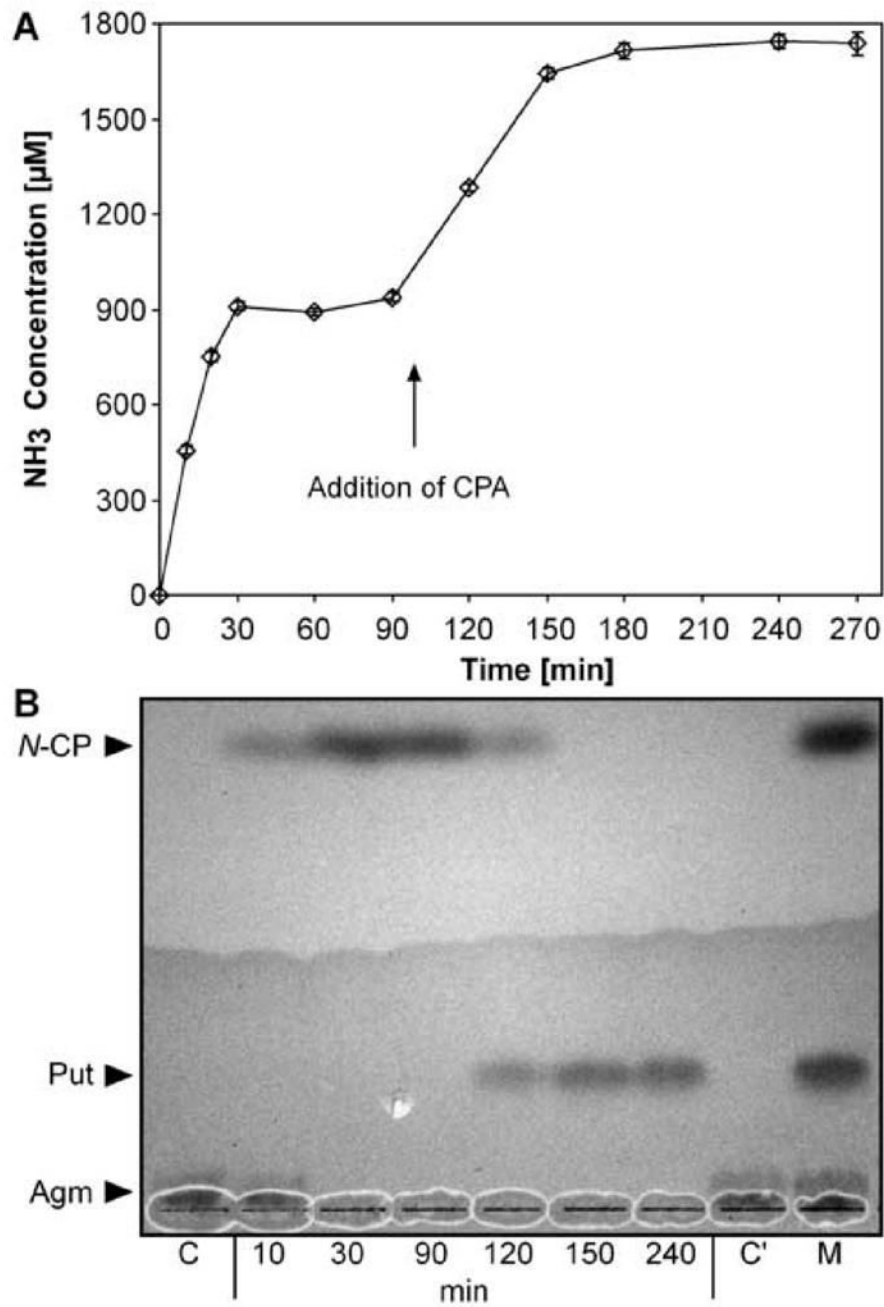
Fig 3. Enzymatic activities of PBCV-1 AIH and CPA
The release of ammonia (A) as well as the turnover of the substrates to the products (B) were analyzed. Agmatine (1 mM) was incubated with PBCV-1 AIH (20 μg) for 90 min, then 15 μg of PBCV-1 CPA was added. C, C': controls with heat denatured enzyme after 10 and 240 min incubation, respectively; M: standard mixture of 2 mM each of agmatine (Agm), *N*-carbamoylputrescine (*N*-CP) and putrescine (Put).
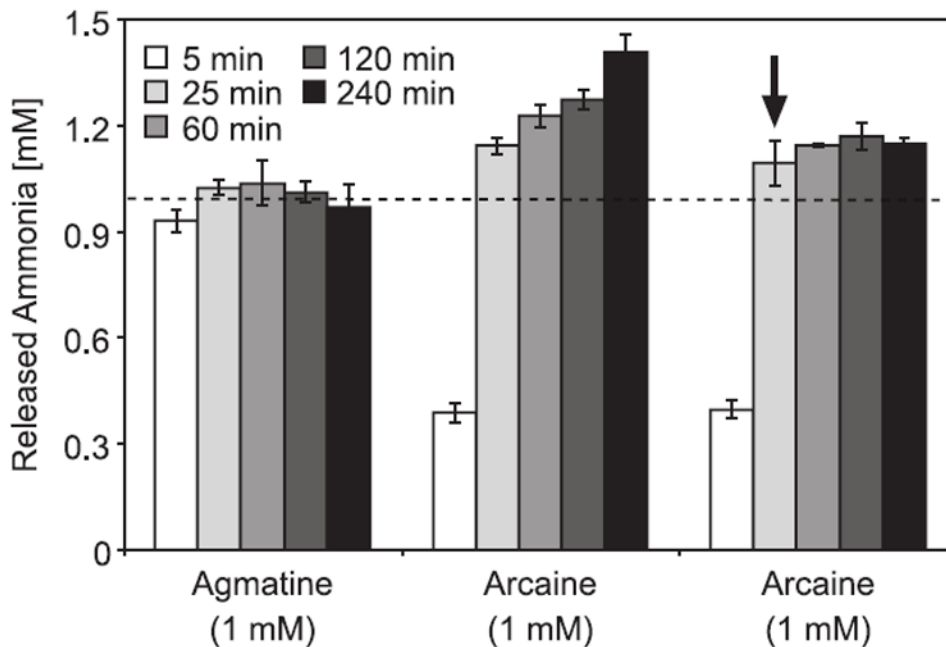
Fig 4. Release of ammonia from agmatine and arcaine by PBCV-1 AIH
Forty μg of protein was incubated with 1 mM of the substrates and the released ammonia was quantified at the indicated time points. In the second arcaine sample (third set of columns) the protein was heat inactivated after 30 min (indicated with an arrow). The dotted line shows the 1 mM level.
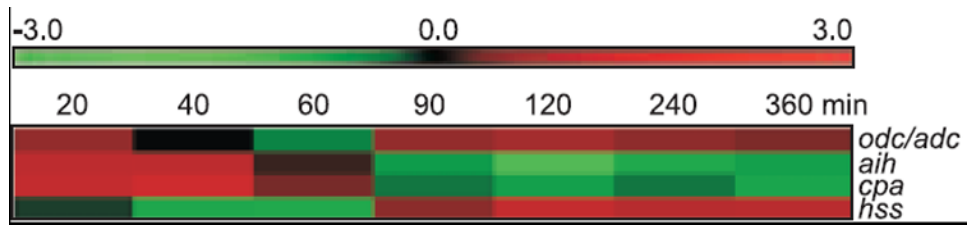
Fig 5.
Expression of PBCV-1 polyamine biosynthesis genes during the infection cycle as revealed by microarray analyses. Color code represents the log2 (Cy5/Cy3) ratio for each time point and has Cy3 as a reference.

A

Agmatin Iminohydrolase

Isolate									Isolate
HOST					●	●	●	●	NY-s-1
NE-8D	●	●	●	●	●	●	●	●	IL-5-2s1
NYb-1	●	●	●	●	●	●	●	●	AL-2A
CA-4B	●	●	●	●	●	●	●	●	MA-1D
AL-1A	●	●	●	●	●	●	●	●	NY-2B
NY-2C	●	●	●	●	●	●	●	●	CA-4A
NC-1D	●	●	●	●	●	●	●		NY-2A
PBCV-1	●	●	●	●	●	●	●	●	XZ-3A
NC-1C	●	●	●	●	●	●	●	●	SH-6A
CA-1A	●	●	●	●	●	●	●	●	BJ-2C
CA-2A	●	●	●	●	●	●	●	●	XZ-6E
IL-2A	●	●	●	●	●	●	●	●	XZ-4C
IL-2B	●	●	●	●	●	●	●	●	XZ-5C
IL-3A	●	●	●	●	●	●	●	●	XZ-4A
IL-3D	●	●	●	●					CVA-1
SC-1A	●	●	●	●					CVB-1
SC-1B	●	●	●	●					CVG-1
NC-1A	●	●	●	●					CVM-1
NE-8A	●	●	●	●					CVR-1
AL-2C	●	●	●	●					
MA-1E	●	●	●	●					
NY-2F	●	●	●	●					
CA-1D	●	●	●	●					
NC-1B	●	●	●	●					

B

N-Carbamoylputrescine Amidohydrolase

Isolate									Isolate
HOST									NY-s-1
NE-8D	●	●	●	●					IL-5-2s1
NYb-1	●	●	●	●	●	●	●	●	AL-2A
CA-4B	●	●	●	●					MA-1D
AL-1A	●	●	●	●					NY-2B
NY-2C	●	●	●	●	●	●	●	●	CA-4A
NC-1D	●	●	●	●	●	●	●		NY-2A
PBCV-1	●	●	●	●	●	●	●	●	XZ-3A
NC-1C	●	●	●	●	●	●	●	●	SH-6A
CA-1A	●	●	●	●	●	●	●	●	BJ-2C
CA-2A	●	●	●	●	●	●	●	●	XZ-6E
IL-2A	●	●	●	●	●	●	●	●	XZ-4C
IL-2B	●	●	●		●	●	●	●	XZ-5C
IL-3A	●	●	●	●	●	●	●	●	XZ-4A
IL-3D	●	●	●	●					CVA-1
SC-1A	●	●	●	●					CVB-1
SC-1B	●	●	●	●					CVG-1
NC-1A	●	●	●	●					CVM-1
NE-8A	●	●	●	●					CVR-1
AL-2C	●	●	●	●					
MA-1E	●	●	●	●					
NY-2F	●	●	●	●					
CA-1D	●	●	●	●					
NC-1B	●	●	●	●					

Fig 6.
Hybridization of the PBCV-1 *aih* gene (A) and *cpa* gene (B) to DNAs isolated from the host *Chlorella* NC64A, 37 NC64A viruses and 5 Pbi viruses (CVA-1, CVB-1, CVG-1, CVM-1, and CVR-1). The DNAs were hybridized with a ^{32}P-labelled *a638r* (*aih*) gene probe (A) or a *a78r*(*cpa*) gene probe (B). The blots in each row contain 1, 0.5, 0.25, and 0.12 μg of DNA from left to right, respectively.

Table 1

Characteristics of AIH and CPA proteins from chlorella viruses in comparison to the enzymes from *Arabidopsis thaliana* and *Pseudomonas aeruginosa* PAO1

Enzyme	Length [amino acids]	Calculated MW [kDa]	Calculated IEP	Locus tag	Identity to *Arabidopsis*/ *Pseudomonas* enzymes [%]
AIH					
*At*AIH	383	43.1	4.92	At5g08170	100/60
*Pa*AIH	368	41.2	4.60	PA0292	60/100
PBCV-1 AIH	359	40.8	5.83	A638R	48/51
NY-2A AIH	359	40.9	5.46	B844R	47/50
MT325 AIH	363	41.1	5.42	M766L	45/51
CPA					
*At*CPA	299	33.5	5.93	At2g27450	100/63
*Pa*CPA	292	32.7	5.92	PA0293	63/100
PBCV-1 CPA	298	33.2	5.70	A78R	49/50
NY-2A CPA	298	33.2	5.42	B116R	49/51
MT325 CPA	296	32.9	7.04	M103L	47/47

Table 2

Enzymatic properties of AIH and CPA proteins from chlorella viruses, *Arabidopsis thaliana*, and *Pseudomonas aeruginosa* PAO1

Enzyme	pH optimum	Temperature optimum (°C)	K_m (mmol L^{-1})	k_{cat} (s^{-1})
AIH				
PBCV-1 AIH	7.5	25	0.043	10.2
NY-2A AIH	7.5	30	0.054	9.7
MT325 AIH	7.5	20	0.054	_[d]
*At*AIH[a]	7	35-40	0.112	18.0
*Pa*AIH[b]	8.0	45	0.6	4.2
CPA				
PBCV-1 CPA	5.5-6.0	40	0.250	3.1
NY-2A CPA	5.5-6.0	45	0.216	2.8
MT325 CPA	6.5	30	0.214	_[d]
*At*CPA[c]	8-9	40	0.135	2.9
*Pa*CPA[b]	8.0	40	0.5	3.3

[a] from Janowitz et al., 2003;

[b] from Nakada and Itoh, 2003,

[c] from Piotrowski et al., 2003,

[d] not determined (see text).

Table 3

Polyamine biosynthesis ORFs in several chlorella viruses. The information was obtained from the Greengene server (http://greengene.uml.edu)

| | NC64A Viruses | Chlorella Virus | | Pbi Viruses | |
	PBCV-1	NY-2A	AR158	FR483	MT325	ATCV-1
Enzyme						
ODC/ADC	A207R	B278R	C256R	N312L	M307L	Z760R
AIH	A638R	B844R	C759R	-	M766L	Z806R
CPA	A78R	B116R	C104R	N095L	M103L	Z169R
HSS	A237R	B305R	C286R	N232L	M233L	Z590L

Table 4

G+C content of the chlorella virus genes encoding polyamine biosynthetic enzymes

		Virus					
	PBCV-1	NY-2A	AR158	MT325	FR483	ATCV-1
Gene						
odc/adc	41%	45%	41%	46%	46%	51%
aih	43%	42%	43%	47%	-	51%
cpa	45%	46%	48%	51%	50%	55%
hss	44%	44%	44%	47%	46%	54%
Complete genome	40%	40.7%	40.7%	45.3%	44.6%	49%