JOURNAL OF VIROLOGY, Mar. 1997, p. 1931–1937 Vol. 71, No. 3
0022-538X/97/$04.00+0

Characterization of an ATP-Dependent DNA Ligase Encoded by *Chlorella* Virus PBCV-1

C. KIONG HO,[1] JAMES L. VAN ETTEN,[2] AND STEWART SHUMAN[1]*

Molecular Biology Program, Sloan-Kettering Institute, New York, New York 10021,[1] and Department of Plant Pathology, University of Nebraska, Lincoln, Nebraska 68583-0722[2]

Received 4 November 1996/Accepted 10 December 1996

We report that *Chlorella* virus PBCV-1 encodes a 298-amino-acid ATP-dependent DNA ligase. The PBCV-1 enzyme is the smallest member of the covalent nucleotidyl transferase superfamily, which includes the ATP-dependent polynucleotide ligases and the GTP-dependent RNA capping enzymes. The specificity of PBCV-1 DNA ligase was investigated by using purified recombinant protein. The enzyme catalyzed efficient strand joining on a singly nicked DNA in the presence of magnesium and ATP (K_m, 75 μM). Other nucleoside triphosphates or deoxynucleoside triphosphates could not substitute for ATP. PBCV-1 ligase was unable to ligate across a 2-nucleotide gap and ligated poorly across a 1-nucleotide gap. A native gel mobility shift assay showed that PBCV-1 DNA ligase discriminated between nicked and gapped DNAs at the substrate-binding step. These findings underscore the importance of a properly positioned 3′ OH acceptor terminus in substrate recognition and reaction chemistry.

The ATP-dependent DNA ligases catalyze the joining of 5′ phosphate-terminated donor strands to 3′ hydroxyl-terminated acceptor strands via three sequential nucleotidyl transfer reactions (10). In the first step, nucleophilic attack on the α-phosphate of ATP by ligase results in liberation of pyrophosphate (PP_i) and formation of a covalent intermediate in which AMP is linked to the ε-amino group of a lysine. The nucleotide is then transferred to the 5′ end of the donor polynucleotide to form DNA-adenylate—an inverted (5′)-(5′) bridge structure, AppN. Attack by the 3′ OH of the acceptor strand on the DNA-adenylate joins the two polynucleotides and liberates AMP.

Animal cells contain multiple DNA ligase isozymes encoded by separate genes (1, 28, 29). ATP-dependent DNA ligases are also encoded by animal DNA viruses, e.g., the poxviruses and African swine fever virus (ASFV), by the T-odd and T-even bacteriophages (T4, T6, T3, and T7), by yeasts, plants, and archaea (7). The ATP-dependent DNA ligases belong to a superfamily of covalent nucleotidyl transferases that includes the GTP-dependent eukaryotic mRNA capping enzymes (20, 22). The ligase-capping enzyme superfamily is defined by a set of six short motifs (Fig. 1). The lysine within motif I (KxDG) is the active site of AMP transfer by the ligases (6, 25, 28) and GMP transfer by the capping enzymes (2, 4, 15, 18). Conserved residues within motifs I, III, IV, and V are critical for covalent nucleotidyl transfer, as shown by mutational analyses (2–4, 6, 8, 18, 20, 21). The recently reported crystal structure of T7 DNA ligase shows that the ATP-binding site is made up of conserved motifs I, III, IIIa, IV, and V (24).

The bacteriophage T7 and T3 enzymes (359- and 346-amino-acid polypeptides, respectively) are the smallest ATP-dependent ligases described to date (7). Cellular DNA ligases are much larger; for example, human ligases I, III, and IV are 919-, 922-, and 844-amino-acid polypeptides, respectively (1, 29). Vaccinia virus and ASFV ligases are intermediate in size (552 and 419 amino acids, respectively) (5, 23). Sequence comparisons of cellular and virus-encoded proteins suggest that a catalytic domain common to all ATP-dependent ligases is embellished by additional isozyme-specific protein segments situated at their amino or carboxyl termini. Virus-encoded enzymes, by virtue of their small size, may well define the catalytic core of the ligase family.

Here, we report a new viral DNA ligase encoded by *Paramecium bursaria Chlorella* virus 1 (PBCV-1). PBCV-1 is the prototype of a family of large, polyhedral DNA viruses that replicate in unicellular eukaryotic *Chlorella*-like green algae (27). The PBCV-1 genome, like the genomes of the poxviruses and ASFV, is a linear, double-strand DNA molecule with inverted terminal repeats and covalently closed hairpin telomeres. The sequence of the 330-kbp PBCV-1 genome has been determined (9, 11–13); PBCV-1 encodes ~380 polypeptides.

An open reading frame encoding a ligase-like protein was encountered between nucleotide coordinates 264,000 and 265,000 of the PBCV-1 genome (11a; Genbank accession number is U77663). The predicted gene product includes the six motifs shared among the cellular and DNA virus-encoded ATP-dependent DNA ligases (Fig. 2). The order and spacing of these motifs in the PBCV-1 ligase-like protein are similar to those seen in other ligase family members (Fig. 1). The PBCV-1 polypeptide, at 298 amino acids, is smaller than any known ligase or capping enzyme. We have expressed the PBCV-1 protein in *Escherichia coli* and purified it to apparent homogeneity. We report that the recombinant protein is indeed an ATP-dependent DNA ligase. A biochemical characterization of the PBCV-1 ligase is presented.

MATERIALS AND METHODS

T7-based vector for expression of PBCV-1 DNA ligase. Oligonucleotide primers complementary to the 5′ and 3′ ends of the putative PBCV-1 ligase gene were used to amplify the 298-amino-acid open reading frame. Total PBCV-1 genomic DNA was used as the template for a PCR catalyzed by *Pfu* DNA polymerase (Stratagene). The sequence of the 5′ flanking primer was 5′-CATGAAGTTAC GTGTGTCATATGGCAATCACAAAGCC; that of the 3′ flanking primer was 5′-CAAGACTTCGTAAAAACGGATCCTACATGGGATGA. These primers were designed to introduce *Nde*I and *Bam*HI restriction sites at the 5′ and 3′ ends, respectively, of the ligase gene. The 0.9-kbp PCR product was digested with *Nde*I and *Bam*HI and then cloned into the *Nde*I and *Bam*HI sites of T7-based expression plasmid pET3c (Novagen). The resulting plasmid, pET-ChV-ligase, was transformed into *Escherichia coli* BL21(DE3). Dideoxy sequencing of the entire insert of pET-ChV-ligase confirmed that no alterations of the genomic DNA sequence were introduced during PCR amplification and cloning of the ligase gene.

* Corresponding author.

	I		III		IIIa		IV		V		VI
ChV DNA	KIDG	-31-	EGSDGEIS	-24-	SYYWFDYVTDDP	-55-	EGVMI	-18-	LLKMKQFKDAEAT	-90-	PVFIGIRHEE
	*		*				*		*		
VAC DNA	KYDG	-43-	IVLDSEIV	-27-	CLFVFDCLYFDG	-52-	EGLVL	-13-	WLKIKR--DYLNE	-120-	PRFTRIREDK
SFV DNA	KYDG	-43-	FILDAELV	-28-	CLFVFDCLFYND	-52-	EGLML	-13-	WLKIKK--DHLKT	-120-	PRCTKIRDDK
FPV DNA	KYDG	-43-	MILDGEII	-28-	CIFIFDCLYFND	-52-	EGFVL	-13-	WLKIKK--DYLDG	-119-	PRCSRIREDK
Sc DNA	KYDG	-44-	LILDCEAV	-32-	CLFAFDILCYND	-52-	EGLMV	-18-	WLKLKK--DYLEG	-118-	PRFLRIREDK
Sp DNA	KYDG	-44-	FILDCEAV	-32-	CLFAFDILYLNG	-52-	EGLMV	-18-	WLKVKK--DYLSG	-122-	PRFIRIREDK
At DNA	KYDG	-44-	FILDCEVV	-32-	CIFAFDMLYLNG	-52-	EGLII	-17-	WLKLKK--DYMDS	-120-	PRLLRVREDK
Hu1 DNA	KYDG	-44-	FILDTEAV	-32-	CLYAFDLIYLNG	-52-	EGLMV	-17-	WLKLKK--DYLDG	-120-	PRFIRVREDK
Hu3 DNA	KYDG	-43-	MILDSEVL	-28-	CLFVFDCIYFND	-52-	EGLVI	-13-	WLKVKK--DYLNE	-121-	PRCTRIRDDK
Hu4 DNA	KLDG	-49-	CILDGEMM	-29-	CLCVFDVLMVNN	-52-	EGIMV	-14-	WLKIKP--EYVSG	-121-	PRIEKIRDDK
Xe DNA	KYDG	-44-	CILDTEAV	-32-	CVYAFDMLYLNG	-52-	EGLMV	-17-	WLKLKK--DYLEG	-120-	PRFLRIRDDK
Ce DNA	KYDG	-44-	FIVDAEVV	-30-	VVFLFDLLYFNG	-52-	EGLMI	-17-	WLKMKK--DYVDG	-120-	PRFLRIRDDK
Ca DNA	KMDG	-49-	VILDGEMV	-40-	FFLVFDILFLNG	-52-	EGLVL	-16-	WIKVKP--EYLEK	-113-	NHCRKIREDK
ASF DNA	KRNG	-43-	VYLDGELY	-22-	HFYVFDCFWSDQ	-51-	EGAIV	-18-	LAKLKPLLDAEFI		
Dam DNA	KYDG	-41-	FIVEGEII	-33-	NVFLFDLMYYEG	-51-	EGVMV	-17-	WIKFKR--DYQSE	-117-	PRFIRWRPDK
Mja DNA	KYDG	-40-	LIVEGECV	-34-	RVYLFDILYKDG	-68-	EGVMI	-15-	MYKFKPTLESLDV	-93-	PRVVRFRFDK
T4 DNA	KADG	-49-	VLIDGELV	-58-	KFQVWDYVPLVE	-54-	EGIIL	-14-	LYKFKEVIDVDLK	-83-	PIAIRLREDK
T3 DNA	KYDG	-50-	FMLDGELM	-55-	SVRLYAVMPIHI	-60-	EGLIV	-14-	WWKLKP--ECEAD	-77-	PSFEKFRGTE
T7 DNA	KYDG	-50-	FMLDGELM	-49-	HIKLYAILPLHI	-60-	EGLIV	-14-	WWKMKP--ENEAD	-96-	PSFVMFRGTE
T4 RNA	KEDG	-4-	TYLDGDEI		-112-		EGYVA	-6-	HFKIKS--DWYVS		
ChV CE	KTDG	-40-	SIFDGELC	-8-	AFVLFDAVVVSG	-59-	DGLII	-14-	LFKLKPGTHHTID	-44-	WKYIQGRSDK
ASF CE	KADG	-32-	TILDGEFM	-4-	EFYGFDVIMYEG	-62-	DGIIL	-11-	TFKWKPTWDNTLD	-104-	WEIVKIREDR
Sp CE	KSDG	-50-	TLLDGELV	-11-	RYLVFDCLACDG	-67-	DGLIF	-14-	LLKWKPKEMNTID	-70-	WRFLRFRDDK
Sc CE	KTDG	-53-	TLLDGELV	-12-	RYLMFDCLAING	-66-	DGLIF	-15-	LLKWKPEQENTVD	-105-	WEMLRFRDDK
VAC CE	KTDG	-34-	VVVFGEAV	-3-	NWTVYLIKLIEP	-54-	EGVIL	-10-	DFKIKK--ENTID	-86-	GEILKPRIDK
SFV CE	KTDG	-34-	VTLYGEAV	-4-	VWQIYLIKLITP	-52-	EGVVL	-9-	DYKIKL--DNTTD	-86-	GEILKPRIDK
MCV CE	KTDG	-34-	VALFGEAV	-6-	QLTVYLIKLMAP	-52-	EGVVL	-9-	DLKLKR--DNTVD	-86-	GRLLRPRLAK

FIG. 1. Conserved sequence elements define a superfamily of covalent nucleotidyl transferases. Six colinear sequence elements, designated motifs I, III, IIIa, IV, V, and VI, are conserved in polynucleotide ligases and mRNA capping enzymes as shown. The aligned amino acid sequences are those of the DNA ligases (DNA) encoded by *Chlorella* virus PBCV-1 (ChV), vaccinia virus (VAC), Shope fibroma virus (SFV), fowlpox virus (FPV), *Saccharomyces cerevisiae* (Sc), *Schizosaccharomyces pombe* (Sp), *Arabidopsis thaliana* (At), human ligase I (Hu1), human ligase 3 (Hu3), human ligase 4 (Hu4), *Xenopus laevis* (Xe), *Caenorhabditis elegans* (Ce), *Candida albicans* (Ca), ASFV (ASF), *Desulfurolobus ambivalens* (Dam), *Methanococcus jannaschii* (Mja), and bacteriophages T4, T3, and T7. Also included is T4 RNA ligase. Grouped below the ligases are the aligned sequences of capping enzymes (CE) encoded by *Chlorella* virus PBCV-1 (ChV), ASFV (ASF), *S. cerevisiae* (Sc), *S. pombe* (Sp), vaccinia virus (VAC), Shope fibroma virus (SFV), and molluscum contagiosum virus (MCV). The number of intervening amino acid residues separating the motifs is indicated between the motifs. Residues in the vaccinia virus DNA ligase that were found by mutational analysis to be essential for activity (21) are indicated by asterisks.

Expression and purification of recombinant PBCV-1 ligase. A 500-ml culture of *E. coli* BL21(DE3)/pET-ChV-ligase was grown at 37°C in Luria-Bertani medium containing 0.1 mg of ampicillin per ml until the A_{600} reached 0.5. The culture was adjusted to 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and incubation was continued at 37°C for 2 h. Cells were harvested by centrifugation, and the pellet was stored at −80°C. All subsequent procedures were performed at 4°C. Thawed bacteria were resuspended in 50 ml of buffer A (50 mM Tris HCl [pH 7.5], 2 mM dithiothreitol [DTT], 1 mM EDTA, 10% sucrose) containing 0.2 M NaCl. Lysozyme was added to a final concentration of 2 μg/ml, and the sample was sonicated for 30 s. Triton X-100 was added to a 0.1% final concentration. The suspension was frozen on dry ice and then allowed to thaw at 4°C. Sonication was repeated. Insoluble material was removed by centrifugation for 45 min at 18,000 rpm in a Sorvall SS34 rotor. The soluble extract (60 mg of protein) was adjusted to 50 mM NaCl by addition of 3 volumes of buffer A, and this material was applied to a 25-ml column of DEAE-cellulose that had been equilibrated with buffer A containing 50 mM NaCl. The flowthrough fraction (21 mg of protein) was applied to a 10-ml column of phosphocellulose that had been equilibrated in buffer A containing 50 mM NaCl. The column was washed with the same buffer and then eluted stepwise with buffer B (50 mM Tris HCl [pH 8.0], 2 mM DTT, 10% glycerol) containing 0.1, 0.15, 0.2, 0.3, 0.5, and 1.0 M NaCl. The polypeptide composition of the column fractions was monitored by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE). The recombinant PBCV-1 protein was retained on the column and was recovered predominantly in the 0.3 M fraction (7 mg of protein). An aliquot of the 0.3 M phosphocellulose fraction was applied to a 4.8-ml 15 to 30% glycerol gradient containing 50 mM Tris-HCl (pH 8.0), 2 mM DTT, 0.5 M NaCl, and 0.1% Triton X-100. The gradient was centrifuged at 50,000 rpm for 24 h at 4°C in a Beckman SW50 rotor. Fractions were collected from the bottom of the tube. Marker proteins—bovine serum albumin, soybean trypsin inhibitor, and cytochrome *c*—were sedimented in a parallel gradient. Enzyme fractions were stored at −80°C and thawed on ice just prior to use. Protein concentrations were determined by the Bio-Rad dye-binding assay with bovine serum albumin as the standard.

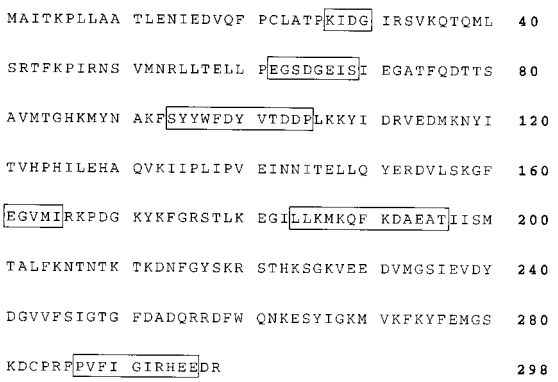

FIG. 2. Amino acid sequence of a putative PBCV-1 DNA ligase. Conserved motifs I, III, IIIa, IV, V, and VI are boxed.

Enzyme-AMP complex formation. Standard reaction mixtures (20 μl) containing 50 mM Tris HCl (pH 7.5), 5 mM DTT, 5 mM $MgCl_2$, 10 μM [α-^{32}P]ATP, and enzyme were incubated for 5 min at 37°C and then halted by adding SDS to a 1% final concentration. The samples were electrophoresed through a 12% polyacrylamide gel containing 0.1% SDS. Label transfer to the 34-kDa PBCV-1 polypeptide was visualized by autoradiographic exposure of the dried gel and was quantitated by scanning the gel with a FUJIX BAS1000 Bio-Imaging Analyzer.

Ligase substrate. The standard substrate used in ligase assays was a 36-bp DNA duplex containing a centrally placed nick (see Fig. 5). This DNA was formed by annealing two 18-mer oligonucleotides to a complementary 36-mer strand. The 18-mer constituting the donor strand was 5′ ^{32}P labeled and gel purified as previously described (19, 21). The labeled donor was annealed to the complementary 36-mer in the presence of a 3′ OH-terminated acceptor strand in 0.2 M NaCl by heating at 65°C for 2 min, followed by slow cooling to room temperature. The molar ratio of the 18-mer donor to 36-mer complement to 18-mer acceptor strands in the hybridization mixture was typically 1:3:6.

DNA ligation. Reaction mixtures (20 μl) containing 50 mM Tris HCl (pH 7.5), 5 mM DTT, 10 mM $MgCl_2$, 1 mM ATP, 1 pmol of 5′ ^{32}P-labeled DNA substrate, and enzyme were incubated at 22°C. Reactions were initiated by enzyme addition and halted by addition of 1 μl of 0.5 M EDTA and 5 μl of formamide. The samples were heated at 95°C for 5 min and then electrophoresed through a 12.5% polyacrylamide gel containing 7 M urea in 0.5× TBE (45 mM Tris borate, 1 mM EDTA). The labeled 36-mer ligation product was well resolved from the 5′-labeled 18-mer donor strand. The extent of ligation [36-mer/(18-mer + 36-mer)] was determined by scanning the gel with a PhosphorImager.

RESULTS

Expression of the PBCV-1 ligase-like protein in bacteria. The PBCV-1 open reading frame encoding a ligase-like protein was PCR amplified from viral genomic DNA and cloned into a T7 RNA polymerase-based bacterial expression vector.

The pET-ChV-ligase plasmid was introduced into *E. coli* BL21(DE3), a strain that contains the T7 RNA polymerase gene under the control of a *lac*UV5 promoter. A prominent 34-kDa polypeptide was detectable by SDS-PAGE in whole-cell extracts of IPTG-induced bacteria (Fig. 3A, lane 1). This polypeptide was not present when bacteria containing the pET vector alone were induced with IPTG (data not shown). After centrifugal separation of the crude lysate, the PBCV-1 protein was recovered in the soluble supernatant fraction (Fig. 3A, lane 2).

Recombinant PBCV-1 34-kDa protein forms a covalent enzyme-adenylate complex. The initial step in DNA ligation involves formation of a covalent enzyme-adenylate intermediate, EpA. EpA formation by DNA ligases can be detected with high sensitivity and specificity by label transfer from [α-^{32}P]ATP to the enzyme. To assay the adenylyltransferase activity of the expressed PBCV-1 protein, we incubated either whole-cell or soluble extracts of IPTG-induced BL21(DE3)/pET-ChV-ligase cells in the presence of [α-^{32}P]ATP and a divalent cation. This resulted in the formation of a nucleotidyl-protein adduct that migrated as a single 34-kDa species during SDS-PAGE (Fig. 3B, lanes 1 and 2). Labeling of this polypeptide was not detected in extracts prepared from bacteria that lacked the PBCV-1 gene (data not shown). We concluded that the expressed 34-kDa PBCV-1 protein is active in nucleotidyl transfer.

Purification of recombinant PBCV-1 ligase. The 34-kDa polypeptide was purified from the soluble bacterial extract by sequential DEAE-cellulose and phosphocellulose column chromatography steps. The recombinant protein did not bind to DEAE-cellulose at low ionic strength (50 mM NaCl). SDS-PAGE analysis of the DEAE flowthrough fraction (Fig. 3A, lane 3) showed that many of the bacterial polypeptides were eliminated during this step. The PBCV-1 protein adsorbed to phosphocellulose and was recovered during step elution with 0.3 M NaCl (Fig. 3A, lane 8). The phosphocellulose preparation was highly enriched with respect to the 34-kDa polypeptide; approximately 7 mg was obtained from a 500-ml culture of IPTG-induced bacteria. The adenylyltransferase activity coincided with the 34-kDa polypeptide during column chromatography (Fig. 3B and other data not shown).

When the phosphocellulose fraction was centrifuged through a 15 to 30% glycerol gradient in 0.5 M NaCl, a single peak of adenylyltransferase activity was detected that coincided with the 34-kDa polypeptide (Fig. 4). We estimated a sedimentation coefficient of 3.1S relative to marker proteins sedimented in a parallel gradient. This suggested that the PBCV-1 adenylyltransferase is a monomer of the 34-kDa protein.

DNA ligation. We assayed the ability of the recombinant PBCV-1 protein to seal a 36-mer synthetic duplex DNA substrate containing a single nick. The structure of the substrate is shown in Fig. 5. Ligase activity was evinced by conversion of the 5′ ^{32}P-labeled 18-mer donor strand into an internally labeled 36-mer product. The DNA ligase activity profile across the glycerol gradient paralleled that of enzyme-adenylate complex formation (Fig. 4B). These results demonstrate that the 34-kDa PBCV-1 protein is indeed a DNA ligase. Further characterization of the PBCV-1 ligase was performed by using the glycerol gradient preparation (peak fraction 19).

The extent of ligation of the nicked duplex (added at a 50 nM concentration with respect to the 5′-labeled donor strand) during a 10-min incubation in the presence of 1 mM ATP increased linearly with the concentration of PBCV-1 ligase from 50 to 500 pM (Fig. 6A). In the linear range of enzyme dependence in this experiment, the recombinant ligase joined about 100 to 120 fmol of DNA ends per fmol of enzyme. To estimate the ratio of product to enzyme, the enzyme molarity was calculated based on total protein concentration, assuming enzyme homogeneity. It was also assumed that all of the enzyme molecules in the preparation were catalytically active. The reaction saturated at >500 pM enzyme with 85% of the labeled donor strand converted to 36-mer in 10 min. This upper limit of ligation probably reflected incomplete annealing of all three component strands to form the nicked substrate.

Ligation could be detected in the absence of added ATP, but only at high levels of input enzyme (Fig. 6B). ATP-independent ligation was attributed to preadenylylated ligase in the enzyme preparation. The linear dependence of ATP-independent strand joining on enzyme indicated that about 0.9 mol of

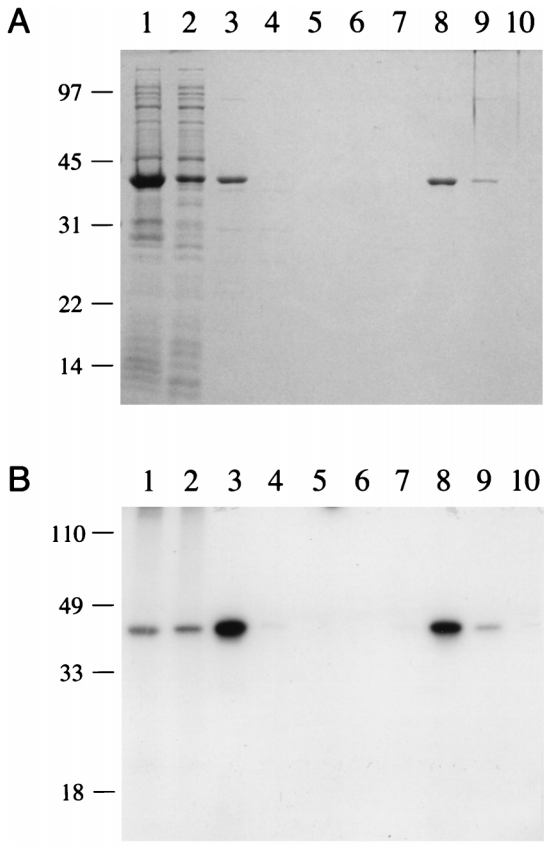

FIG. 3. Expression, purification, and adenylyltransferase activity of recombinant PBCV-1 ligase. (A) The polypeptide compositions of recombinant PBCV-1 protein at sequential stages of purification were analyzed by SDS-PAGE, as follows: whole-cell lysate of IPTG-induced BL21(DE3)/pET-ChV-ligase (lane 1); soluble lysate fraction (lane 2), DEAE-cellulose flowthrough fraction (lane 3), phosphocellulose flowthrough fraction (lane 4), and 0.1 M NaCl (lane 5), 0.15 M NaCl (lanes 6), 0.2 M NaCl (lane 7), 0.3 M NaCl (lane 8), 0.5 M NaCl (lane 9), and 1.0 M NaCl (lane 10) phosphocellulose eluates. The gel was fixed and stained with Coomassie blue dye. The positions and sizes (kilodaltons) of coelectrophoresed marker polypeptides are shown on the left. (B) Enzyme-adenylate complex formation. Reaction mixtures (20 μl) contained 50 mM Tris HCl (pH 7.5), 5 mM DTT, 5 mM $MgCl_2$, 10 μM [α-^{32}P]ATP, and recombinant PBCV-1 protein at the following stages of purification: whole-cell lysate (lane 1), soluble lysate fraction (lane 2), DEAE-cellulose flowthrough fraction (lane 3), phosphocellulose flowthrough fraction (lane 4), and 0.1 M NaCl (lane 5), 0.15 M NaCl (lanes 6), 0.2 M NaCl (lane 7), 0.3 M NaCl (lane 8), 0.5 M NaCl (lane 9), and 1.0 M NaCl (lane 10) phosphocellulose eluates. The reaction products were resolved by SDS-PAGE. An autoradiograph of the dried gel is shown. The positions and sizes (in kilodaltons) of prestained marker polypeptides are indicated on the left.

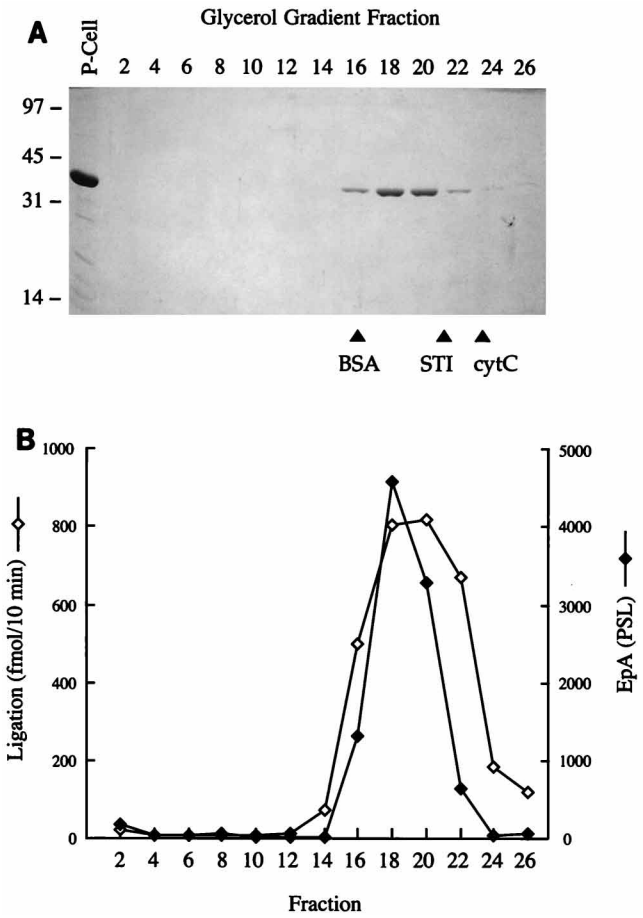

FIG. 4. Glycerol gradient sedimentation. (A) The phosphocellulose ligase preparation was sedimented in a 15 to 30% glycerol gradient as described in Materials and Methods. Aliquots (20 μl) of the phosphocellulose fraction (P-Cell) and the indicated glycerol gradient fractions were analyzed by SDS-PAGE. Polypeptides were visualized by staining with Coomassie blue dye. Marker proteins (bovine serum albumin [BSA], soybean trypsin inhibitor [STI], and cytochrome *c* [cytC]) were centrifuged in a parallel gradient; the marker peaks are indicated below the gel. (B) Aliquots of the glycerol gradient fraction were assayed for enzyme-adenylate complex formation and DNA strand-joining activities as described in Materials and Methods. The DNA ligation reaction mixtures contained 1 pmol of nicked duplex DNA and 0.1 μl (1 μl of a 1:10 dilution) of each of the indicated gradient fractions. Incubation was for 10 min at 22°C. Adenylyltransferase reaction mixtures contained 10 μM [α-^{32}P]ATP and 1 μl of the indicated fractions; incubation was for 5 min at 37°C. Adenylyltransferase activity was gauged by the signal intensity of the radiolabeled ligase polypeptide. PSL, photostimulatable luminescence.

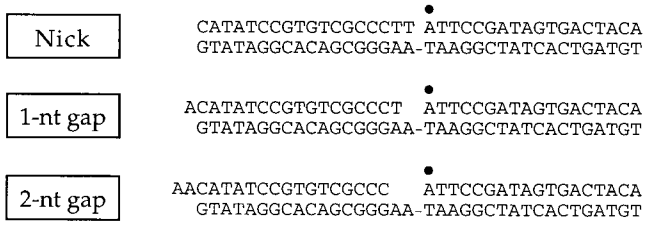

FIG. 5. Ligase substrates. Duplex substrates for PBCV-1 ligase were prepared by annealing a 5′-end-labeled 18-mer donor strand to a 36-mer complementary strand and an 18-mer acceptor strand. The structure of the standard nicked duplex substrate is shown at the top; the position of the 5′-end-labeled nucleotide is indicated by the dots. The structures of 1- and 2-nt-gapped duplexes are shown below that of the nicked duplex.

ends was sealed per mol of ligase (Fig. 6B), implying that 90% of the enzyme molecules had AMP bound at the active site.

Kinetics, ATP dependence, and nucleotide cofactor specificity of ligation. The kinetics of ligation were examined in DNA excess (50 nM) in the presence of 1 mM ATP and 10 mM $MgCl_2$. The initial rate was proportional to enzyme concentration in the range of 125 to 500 pM (Fig. 7A). In subsequent experiments, ligation assay mixtures contained 250 pM enzyme and activity was determined after a 2-min incubation at 22°C (unless otherwise specified). Under these conditions, ligase activity in 50 mM Tris-HCl buffer was optimal between pHs 7.0 and 7.5. Activity at pH 9.5 was ~50% of that at pH 7.5 (data not shown). The rate of ligation increased with ATP concentration from 10 to 200 μM and leveled off at 0.5 to 2 mM (Fig. 7B). A K_m of 75 μM ATP was calculated from a double-reciprocal plot of the data. Other ribonucleoside triphosphates or deoxynucleoside triphosphates at a 0.5 mM concentration could not substitute for ATP (Fig. 8).

Divalent cation dependence and specificity. Ligation depended on a divalent cation in excess of the input ATP; activity was enhanced steadily as Mg was increased from 2 to 20 mM (Fig. 9B). The divalent cation requirement was satisfied by Mn and, to a lesser extent, by Co but not by Ca, Cu, or Zn (Fig. 9A and B).

DNA substrate specificity—nicks versus gaps. The structure of the ligation substrate was altered such that the 3′ hydroxyl-terminated acceptor strand was separated from the 5′ phosphate donor terminus by a 2- or 1-nucleotide (nt) gap (Fig. 5). The specific activity of PBCV-1 ligase on a 1-nt gap substrate was 1% of the activity of a nicked duplex DNA (Fig. 10). PBCV-1 ligase was incapable of joining strands across the 2-nt gap. The implication is that the 3′ OH must be positioned fairly precisely relative to the 5′ phosphate donor terminus for ligation to occur.

Specificity of ligase binding to DNA—nick versus gap. A native gel mobility shift assay was employed to examine the binding of PBCV-1 ligase to the ^{32}P-labeled, nicked duplex DNA (19). Binding reactions were performed in the absence of magnesium and ATP to preclude conversion of substrate to product during the incubation. Control experiments verified that ATP-independent ligation of the nicked DNA substrate by stoichiometric amounts of PBCV-1 ligase required a divalent cation. No strand joining occurred under the reaction conditions employed in our gel shift experiments (data not shown). Mixing the ligase with nicked substrate resulted in the formation of a discrete protein-DNA complex that migrated more slowly than the free DNA during electrophoresis through a 6%

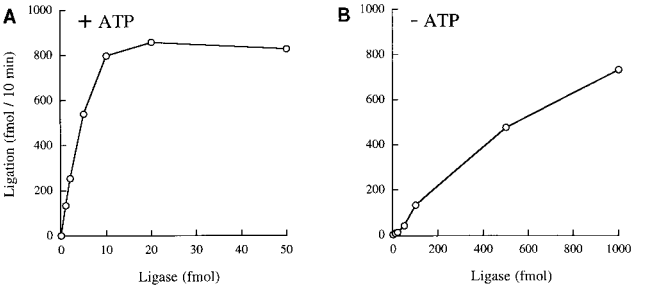

FIG. 6. ATP-dependent and ATP-independent ligation of duplex DNA containing a single nick. (A) Complete reaction mixtures (20 μl) containing 50 mM Tris HCl (pH 7.5), 5 mM DTT, 10 mM $MgCl_2$, 1 mM ATP, 1 pmol of nicked DNA substrate, and the indicated amounts of purified PBCV-1 ligase (glycerol gradient fraction) were incubated for 10 min at 22°C. (B) ATP was omitted from the reaction mixtures. Extent of ligation is plotted as a function of input ligase.

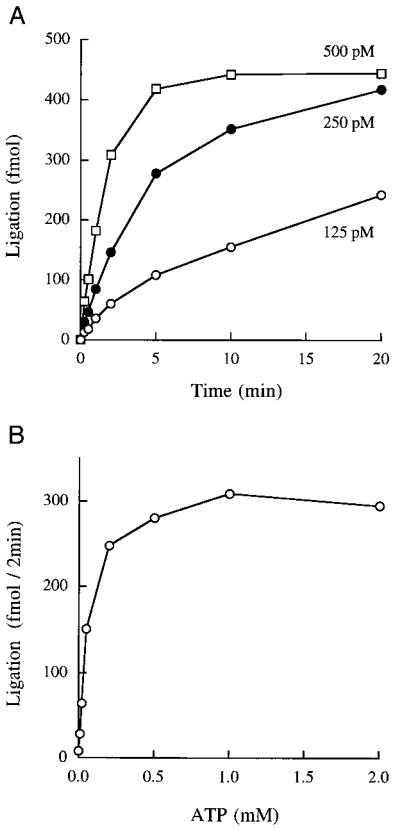

FIG. 7. Kinetics and ATP concentration dependence of strand joining. (A) Kinetics. Reaction mixtures (100 μl) contained 50 mM Tris HCl (pH 7.5), 5 mM DTT, 10 mM $MgCl_2$, 1 mM ATP, 5 pmol of nicked DNA, and either 500, 250, or 125 pM purified PBCV-1 ligase. The reaction was initiated by enzyme addition. Aliquots (10 μl) were withdrawn at the times indicated, and the reaction was quenched immediately. Extent of ligation is plotted as a function of incubation time. (B) ATP dependence. Reaction mixtures (20 μl) containing 50 mM Tris HCl (pH 7.5), 5 mM DTT, 10 mM $MgCl_2$, 1 pmol of nicked DNA, 5 fmol of PBCV-1 ligase, and ATP as indicated were incubated for 2 min at 22°C. Extent of ligation is plotted as a function of ATP concentration.

native polyacrylamide gel (Fig. 11). The abundance of this complex increased in proportion to the amount of input ligase. To estimate binding affinity, the gel was scanned with a PhosphorImager; the apparent dissociation constant, calculated as described by Riggs et al. (17), was 15 nM.

Little or no specific complex was detected when PBCV-1 ligase was incubated with 1- or 2-nt gap DNA (Fig. 11). A diffuse smear of shifted material was detected at a 200 nM ligase concentration. Thus, PBCV-1 DNA ligase bound specifically at a DNA nick and was capable of discriminating between nicked and 2-nt-gapped DNAs at the substrate-binding step. This affirms the importance of the 3′ OH acceptor strand in substrate recognition.

Analysis of enzyme-AMP complex formation. The PBCV-1 ligase reacted specifically with [α-^{32}P]ATP. The amount of enzyme-AMP complex formed during a 5-min incubation at 37°C in the presence of 10 μM [α-^{32}P]ATP was proportional to the amount of added protein (data not shown). EpA formation increased as a function of ATP concentration and reached saturation near 20 μM [α-^{32}P]ATP (Fig. 12A). Half saturation was achieved at ~5 μM ATP. [α-^{32}P]dATP was an extremely poor substrate for EpA formation. We estimated from the NTP titration experiment in Fig. 12A that ATP was 100-fold more effective than dATP in EpA formation. Hence, PBCV-1

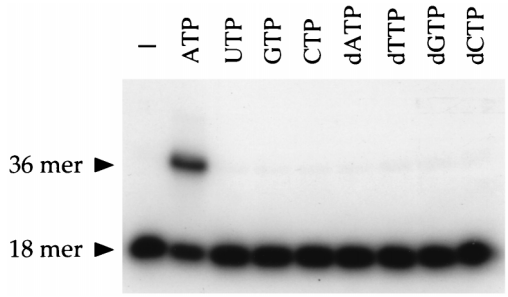

FIG. 8. Nucleotide specificity. Reaction mixtures (20 μl) containing 50 mM Tris HCl (pH 7.5), 5 mM DTT, 10 mM $MgCl_2$, 1 pmol of nicked DNA, 5 fmol of PBCV-1 ligase, and 0.5 mM nucleoside triphosphate or deoxynucleoside triphosphate, as indicated, were incubated for 2 min at 22°C. Nucleotide was omitted from a control reaction (—). The reaction products were electrophoresed through a 12% polyacrylamide gel containing 7 M urea in 0.5× TBE. An autoradiograph of the dried gel is shown. The positions of the 5′-end ^{32}P-labeled 18-mer donor strand and the 36-mer ligation product are indicated on the left.

ligase, which utilized ATP, but not dATP, as a cofactor in the strand-joining reaction, discriminated between ribose and deoxyribose sugars at the step of EpA formation. No protein-nucleotide complex was formed with [α-^{32}P]GTP (data not shown).

We calculated, based on the molar amount of AMP label transfer versus the molar amount of ligase added, that ~5% of the input protein was converted to ligase-[^{32}P]adenylate. This was consistent with the estimate (based on ATP-independent ligation activity) that 90% of the enzyme molecules in the preparation were preadenylated and, hence, incapable of EpA formation. We observed a fivefold increase in the extent of label transfer from [α-^{32}P]ATP to protein when a nicked DNA substrate was added to the reaction mixture (data not shown). This would be expected if transfer of unlabeled AMP from preadenylated enzyme to the 5′ phosphate of the donor strand (step 2 of the ligase reaction) regenerates unadenylated enzyme that can then react with [α-^{32}P]ATP.

EpA formation depended on a divalent cation cofactor. This requirement was satisfied by either 5 mM magnesium or 5 mM manganese and, to lesser extent, by 5 mM cobalt (Fig. 12B). Calcium, copper, and zinc were inactive at this concentration (Fig. 12B). The yield of EpA was proportional to the magne-

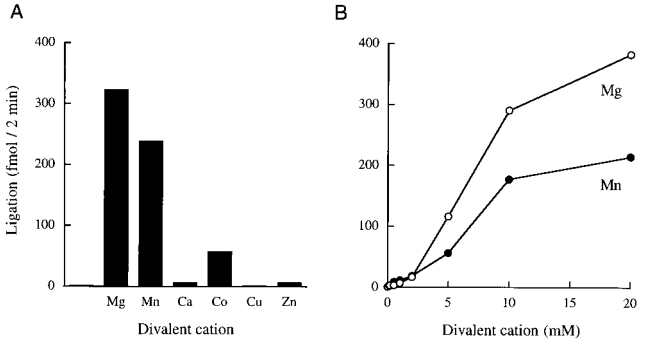

FIG. 9. Divalent cation specificity of strand joining. (A) Reaction mixtures (20 μl) containing 50 mM Tris HCl (pH 7.5), 1 mM ATP, 1 pmol of nicked DNA, 5 fmol of PBCV-1 ligase, and the indicated divalent cation at 10 mM were incubated for 2 min at 22°C. Divalent cation was omitted from a control reaction. Mg, Mn, Ca, and Co were added as chloride salts; Cu and Zn were added as sulfates. (B) Reaction mixtures (20 μl) containing 50 mM Tris HCl (pH 7.5), 5 mM DTT, 1 mM ATP, 1 pmol of nicked DNA, 5 fmol of PBCV-1 ligase, and $MgCl_2$ or $MnCl_2$, as indicated, were incubated for 2 min at 22°C. Extent of ligation is plotted as a function of divalent cation concentration.

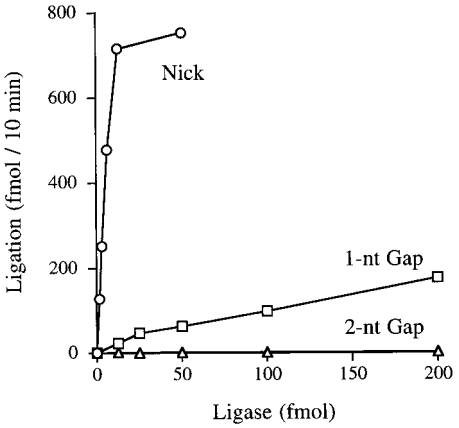

FIG. 10. DNA substrate specificity. Reaction mixtures (20 μl) containing 50 mM Tris HCl (pH 7.5), 5 mM DTT, 1 mM ATP, 1 pmol of either nicked DNA, 1- or 2-nt gap DNA substrate, and PBCV-1 ligase, as indicated, were incubated for 10 min at 22°C. Extent of ligation is plotted as a function of input enzyme.

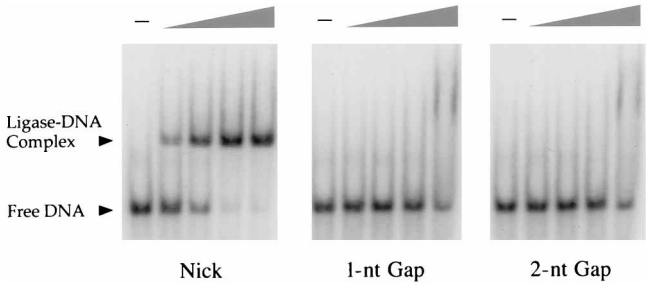

FIG. 11. Nucleic acid binding specificity. Reaction mixtures (10 μl) contained 50 mM Tris HCl (pH 7.5); 5 mM DTT; 5% glycerol; 50 nM ^{32}P-labeled, nicked DNA (left), 1-nt gap DNA (center), or 2-nt gap DNA (right); and 0, 25, 50, 100, or 200 nM purified PBCV-1 ligase (proceeding from left to right within each titration series). After incubation for 10 min at 22°C, the samples were electrophoresed at 100 V through a native 6% polyacrylamide gel in 0.25× TBE (22.5 mM Tris-borate, 0.6 mM EDTA). An autoradiogram of the dried gel is shown. Labeled species corresponding to free DNA and the ligase-DNA complex are indicated at the left.

sium concentration from 0.1 to 1 mM and plateaued at 2 to 10 mM (Fig. 12C). Manganese was a more effective cofactor than magnesium at 0.5 mM but was progressively less active at higher concentrations (Fig. 12C).

DISCUSSION

A *Chlorella* virus PBCV-1 gene encoding a putative DNA ligase was identified during sequencing of the viral DNA genome. We show that the 298-amino-acid gene product is an ATP-dependent DNA ligase. This was achieved by expressing the PBCV-1 protein in bacteria, purifying the protein to homogeneity, and characterizing its enzymatic properties. PBCV-1 ligase, like other cellular and virus-encoded DNA ligases, catalyzes strand joining via an enzyme-AMP intermediate. The PBCV-1 enzyme displays strict specificity for ATP as the nucleotide cofactor; dATP is inactive. PBCV-1 ligase thus resembles the T4, vaccinia virus, and cellular type I enzymes in its discrimination of the nucleoside triphosphate sugar moiety (14, 16, 19, 21, 26).

The high efficiency of PBCV-1 ligase in strand joining across a nick in duplex DNA contrasts sharply with the low efficiency of ligation across a 1-nt gap and the inability to seal strands across a 2-nt gap. Vaccinia virus ligase and yeast CDC9 ligase display similar properties. The latter two enzymes synthesize substantial levels of the DNA-adenylate intermediate on a 1-nt gap DNA substrate (19, 26). Similar results were obtained when stoichiometric levels of PBCV-1 ligase were reacted with the 1-nt gap DNA (6a).

A native gel mobility shift assay was used to demonstrate that formation of a stable complex between PBCV-1 ligase and DNA depends on a DNA nick. The PBCV-1 enzyme discriminates clearly at the DNA-binding step between nicked DNA molecules that can be sealed versus 2-nt-gapped molecules that cannot. Even a 1-nt gap significantly reduces the affinity of ligase for the DNA. This implies that the protein initially contacts both the donor and acceptor DNA strands on either side of the nick prior to any covalent modification of the DNA substrate. Similar specificity for binding to DNA nicks has been documented for the vaccinia virus ligase (19). Thus, both of these virus-encoded enzymes have an intrinsic ability to bind

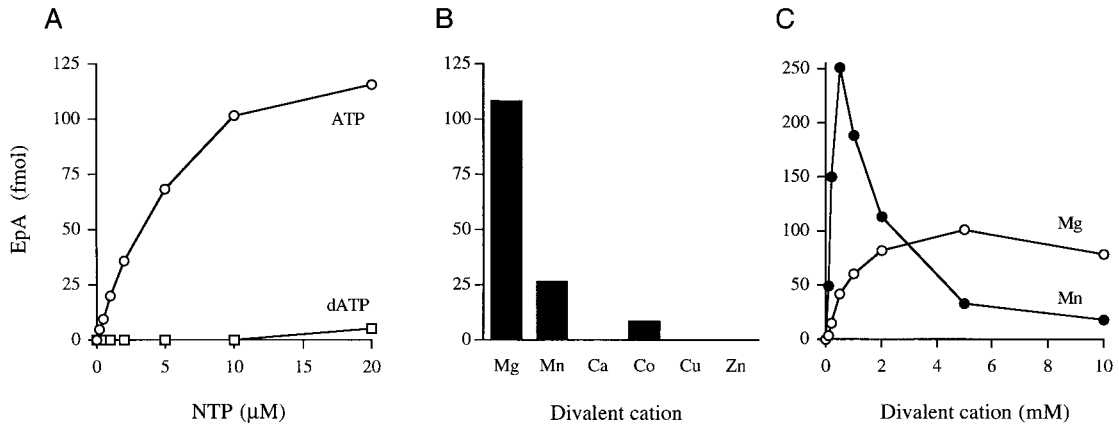

FIG. 12. Analysis of enzyme-AMP complex formation. (A) ATP dependence. Reaction mixtures (20 μl) containing 50 mM Tris HCl (pH 7.5), 5 mM $MgCl_2$, 5 mM DTT, 2 pmol of PBCV-1 ligase (glycerol gradient fraction), and [α-^{32}P]ATP or [α-^{32}P]dATP at the concentrations indicated were incubated for 5 min at 37°C. Extent of EpA formation is plotted as a function of NTP concentration. (B) Divalent cation specificity. Reaction mixtures (20 μl) containing 50 mM Tris HCl (pH 7.5), 10 μM [α-^{32}P]ATP, 2 pmol of PBCV-1 ligase, and the indicated divalent cation at 5 mM were incubated for 5 min at 37°C. Mg, Mn, Ca, and Co were added as chloride salts; Cu and Zn were added as sulfates. (C) Divalent cation dependence. Reaction mixtures (20 μl) containing 50 mM Tris HCl (pH 7.5), 5 mM DTT, 10 μM [α-^{32}P]ATP, 2 pmol of PBCV-1 ligase, and $MgCl_2$ or $MnCl_2$ at the concentrations indicated were incubated for 5 min at 37°C. Extent of EpA formation is plotted as a function of divalent cation concentration.

preferentially to DNA sites where their action is required. Electrophoretic resolution of the ligase-DNA complex from unbound DNA will permit the use of chemical and enzymatic footprinting techniques to delineate the interface between PBCV-1 ligase and the DNA substrate, a subject about which almost nothing is known for any DNA ligase. Of particular interest is identification of the region(s) of the ligase polypeptide that mediates nick recognition.

Because the *Chlorella* viruses are not amenable to genetic manipulation, it is not possible to determine the biological function of the DNA ligase during the PBCV-1 replication cycle. A role during viral DNA replication, repair, or recombination is plausible, although the molecular mechanisms of these transactions are largely unexplored for PBCV-1. Available insights come largely from analysis of the genomic DNA sequence, which identifies by sequence homology several potential replication and repair proteins. These include, in addition to the DNA ligase, a DNA polymerase, a polymerase processivity factor, a helicase, a DNA glycosylase-apyrimidinic lyase, and a type II topoisomerase (9, 11–13). The DNA ligase is the first instance in which it has been shown that a purified recombinant PBCV-1 DNA replication protein actually has the enzymatic properties attributed to it on the basis of sequence homology.

The PBCV-1 ligase is the smallest DNA ligase described to date. It may well represent the minimum catalytic unit of an ATP-dependent ligase. Insofar as PBCV-1 ligase is also smaller than any known mRNA capping enzyme, it may constitute the catalytic core of the nucleotidyl transferase superfamily. PBCV-1 ligase includes the six conserved motifs that define the family but contains no additional sequence at the carboxyl terminus downstream of motif VI. It contains only 26 amino acids N terminal to the presumptive active site, Lys-27, within motif I. In this light, the PBCV-1 ligase emerges as an excellent model for further structural and functional studies of a eukaryotic ligase. The fact that the recombinant PBCV-1 ligase is purified in high yield as ligase-adenylate offers an opportunity to solve the structure of the covalent reaction intermediate. Realization of this goal would extend and complement the insights provided by the crystal structure of the bacteriophage T7 enzyme bound noncovalently to ATP (24).

ACKNOWLEDGMENTS

We thank Yu Li for help in sequencing the PBCV-1 DNA ligase gene.

This investigation was supported by NIH grants GM42498 (S.S.) and GM32441 (J.V.E.) and ACS grant FRA-432 (S.S.).

REFERENCES

1. **Barnes, D. E., L. H. Johnston, K. Kodama, A. E. Tomkinson, D. D. Lasko, and T. Lindahl.** 1990. Human DNA ligase I cDNA: cloning and functional expression in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **87:**6679–6683.
2. **Cong, P., and S. Shuman.** 1993. Covalent catalysis in nucleotidyl transfer: a KTDG motif essential for enzyme-GMP complex formation by mRNA capping enzyme is conserved at the active sites of RNA and DNA ligases. J. Biol. Chem. **268:**7256–7260.
3. **Cong, P., and S. Shuman.** 1995. Mutational analysis of mRNA capping enzyme identifies amino acids involved in GTP binding, enzyme-guanylate formation, and GMP transfer to RNA. Mol. Cell. Biol. **15:**6222–6231.
4. **Fresco, L. D., and S. Buratowski.** 1994. Active site of the mRNA-capping enzyme guanylyltransferase from *Saccharomyces cerevisiae*: similarity to the nucleotidyl attachment motif of DNA and RNA ligases. Proc. Natl. Acad. Sci. USA **91:**6624–6628.
5. **Hammond, J. M., S. M. Kerr, G. L. Smith, and L. K. Dixon.** 1992. An African swine fever virus gene with homology to DNA ligases. Nucleic Acids Res. **20:**2667–2671.
6. **Heaphy, S., M. Singh, and M. J. Gait.** 1987. Effect of single amino acid changes in the region of the adenylylation site of T4 RNA ligase. Biochemistry **26:**1688–1696.

6a. **Ho, C. K., and S. Shuman.** Unpublished data.

7. **Kletzin, A.** 1992. Molecular characterization of a DNA ligase gene of the extremely thermophilic archaeon *Desulfurolobus ambivalens* shows close phylogenetic relationship to eukaryotic ligases. Nucleic Acids Res. **20:**5389–5396.
8. **Kodama, K., D. E. Barnes, and T. Lindahl.** 1991. In vitro mutagenesis and functional expression in *Escherichia coli* of a cDNA encoding the catalytic domain of human DNA ligase I. Nucleic Acids Res. **19:**6093–6099.
9. **Kutish, G. F., Y. Li, Z. Lu, D. L. 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:**303–317.
10. **Lehman, I. R.** 1974. DNA ligase: structure, mechanism, and function. Science **186:**790–797.
11. **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:**134–150.

11a. **Li, Y., Z. Lu, S. Ropp, L. Sun, G. F. Kutish, D. L. Rock, and J. L. Van Etten.** Analysis of 72 kb of DNA at the right end of the *Chlorella* virus PBCV-1 genome. Unpublished data.

12. **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:**102–123.
13. **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:**339–352.
14. **Montecucco, A., M. Lestingi, G. Pedrali-Noy, S. Spadari, and G. Ciarrocchi.** 1990. Use of ATP, dATP and their alpha-thio derivatives to study DNA ligase adenylation. Biochem. J. **271:**265–268.
15. **Niles, E. G., and L. Christen.** 1993. Identification of the vaccinia virus mRNA guanylyltransferase active site lysine. J. Biol. Chem. **268:**24986–24989.
16. **Raae, A. J., and K. Kleppe.** 1978. Effect of ATP analogues on T4 polynucleotide ligase. Biochem. Biophys. Res. Commun. **81:**24–27.
17. **Riggs, A. D., H. Suzuki, and S. Bourgeois.** 1970. *lac* repressor-operator interactions: equilibrium studies. J. Mol. Biol. **48:**67–83.
18. **Schwer, B., and S. Shuman.** 1994. Mutational analysis of yeast mRNA capping enzyme. Proc. Natl. Acad. Sci. USA **91:**4328–4332.
19. **Shuman, S.** 1995. Vaccinia virus DNA ligase: specificity, fidelity, and inhibition. Biochemistry **34:**16138–16147.
20. **Shuman, S., Y. Liu, and B. Schwer.** 1994. Covalent catalysis in nucleotidyl transfer reactions: essential motifs in *Saccharomyces cerevisiae* RNA capping enzyme are conserved in *Schizosaccharomyces pombe* and viral capping enzymes and among polynucleotide ligases. Proc. Natl. Acad. Sci. USA **91:**12046–12050.
21. **Shuman, S., and X. M. Ru.** 1995. Mutational analysis of vaccinia DNA ligase defines residues essential for covalent catalysis. Virology **211:**73–83.
22. **Shuman, S., and B. Schwer.** 1995. RNA capping enzyme and DNA ligase: a superfamily of covalent nucleotidyl transferases. Mol. Microbiol. **17:**405–410.
23. **Smith, G. L., Y. S. Chan, and S. M. Kerr.** 1989. Transcriptional mapping and nucleotide sequence of a vaccinia virus gene encoding a polypeptide with extensive homology to DNA ligases. Nucleic Acids Res. **17:**9051–9062.
24. **Subramanya, H. S., A. J. Doherty, S. R. Ashford, and D. B. Wigley.** 1996. Crystal structure of an ATP-dependent DNA ligase from bacteriophage T7. Cell **85:**607–615.
25. **Tomkinson, A. E., N. F. Totty, M. Ginsburg, and T. Lindahl.** 1991. Location of the active site for enzyme-adenylate formation in DNA ligases. Proc. Natl. Acad. Sci. USA **88:**400–404.
26. **Tomkinson, A. E., N. J. Tappe, and E. C. Friedberg.** 1992. DNA ligase I from *Saccharomyces cerevisiae*: physical and biochemical characterization of the CDC9 gene product. Biochemistry **31:**11762–11771.
27. **Van Etten, J. L., L. C. Lane, and R. H. Meints.** 1991. Virus and virus-like particles of the eukaryotic algae. Microbiol. Rev. **55:**586–620.
28. **Wang, Y. C., W. A. Burkhart, Z. B. Mackey, M. B. Moyer, W. Ramos, I. Husain, J. Chen, J. M. Besterman, and A. E. Tomkinson.** 1994. Mammalian DNA ligase II is highly homologous with vaccinia DNA ligase. Identification of the DNA ligase II active site for enzyme-adenylate formation. J. Biol. Chem. **269:**31923–31928.
29. **Wei, Y. F., P. Robins, K. Carter, K. Caldecott, D. J. Pappin, G. L. Yu, R. P. Wang, B. K. Shell, R. A. Nash, P. Schar, et al.** 1995. Molecular cloning and expression of human cDNAs encoding a novel DNA ligase IV and DNA ligase III, an enzyme active in DNA repair and recombination. Mol. Cell. Biol. **1555:**3206–3216.