THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 11, pp. 7068–7074, March 17, 2006


# A Uracil-DNA Glycosylase Inhibitor Encoded by a Non-uracil Containing Viral DNA*

Received for publication, October 13, 2005, and in revised form, January 17, 2006 Published, JBC Papers in Press, January 18, 2006 DOI 10.1074/jbc.M511152200

**Gemma Serrano-Heras, Margarita Salas, and Alicia Bravo**[1]

*From the Instituto de Biología Molecular "Eladio Viñuela" (Consejo Superior de Investigaciones Científicas), Centro de Biología Molecular "Severo Ochoa" (Consejo Superior de Investigaciones Científicas-Universidad Autónoma de Madrid), Universidad Autónoma, Cantoblanco, 28049 Madrid, Spain*

**Uracil-DNA glycosylase (UDG) is an enzyme involved in the base excision repair pathway. It specifically removes uracil from both single-stranded and double-stranded DNA. The genome of the *Bacillus subtilis* phage φ29 is a linear double-stranded DNA with a terminal protein covalently linked at each 5′-end. Replication of φ29 DNA starts by a protein-priming mechanism and generates intermediates that have long stretches of single-stranded DNA. By using *in vivo* chemical cross-linking and affinity chromatography techniques, we found that UDG is a cellular target for the early viral protein p56. Addition of purified protein p56 to *B. subtilis* extracts inhibited the endogenous UDG activity. Moreover, extracts from φ29-infected cells were deficient in UDG activity. We suggested that inhibition of the cellular UDG is a defense mechanism developed by φ29 to prevent the action of the base excision repair pathway if uracil residues arise in their replicative intermediates. Protein p56 is the first example of a UDG inhibitor encoded by a non-uracil-containing viral DNA.**

Uracil in DNA may arise from spontaneous deamination of cytosine and from the occasional use of dUTP instead of dTTP during DNA replication. Hydrolytic deamination of cytosine generates G:U mismatches that cause G:C to A:T transition mutations. To maintain the genome integrity, most prokaryotic and eukaryotic cells rapidly eliminate uracil from DNA by the base excision repair (BER)[2] pathway, which is initiated by the uracil-DNA glycosylase (UDG) enzyme. UDGs have an unusually broad phylogenetic distribution. Some DNA viruses, such as herpesviruses and poxviruses, also encode a UDG activity, whereas the human immunodeficiency virus, type 1, packages cellular UDG (UNG2 enzyme) into virus particles. In these instances, the UDG activity appears to have an important role in virus replication (1).

The UDG enzyme hydrolyzes the *N*-glycosidic bond between the uracil residue and the deoxyribose sugar of the DNA backbone, generating an apurinic-apyrimidinic (AP) site. The first UDG activity reported was purified from *Escherichia coli* cells (2). Since then, enzymes highly homologous to the archetypal *E. coli* UDG have been identified in numerous organisms, including human cells (Family-1 UDGs) (3). Studies on substrate specificity showed that Family-1 UDGs efficiently remove uracil residues from both single-stranded and double-stranded DNAs, often with preference for the single-stranded substrates (4, 5). The AP site generated by the UDG enzyme is further recognized by an AP endonuclease, which cleaves the phosphodiester bond of the DNA backbone 5′ to the AP site. Several AP endonucleases are active not only on double-stranded DNAs but also on single-stranded DNAs (6–8). Further repair can be accomplished via two pathways that involve different subsets of enzymes and result in replacement of one (short patch pathway) or more (long patch pathway) nucleotides (9).

Over the course of evolution, bacteriophages have developed unique proteins that bind to and inactivate critical cellular proteins, shutting off key processes. The DNA genome of phage PBS2 contains uracil in place of thymine residues (10). Following infection of *Bacillus subtilis*, this phage induces various enzymatic activities that ensure the use of dUTP rather than dTTP during viral DNA synthesis. In addition, phage PBS2 encodes an inhibitor of the *B. subtilis* UDG enzyme, named Ugi, which is essential for the preservation of uracil residues incorporated into phage DNA (11–13). Some studies showed that Ugi (84 amino acids) specifically inactivates Family-1 UDGs (13, 14). The x-ray crystal structures of Ugi in complex with different UDGs revealed that Ugi mimics electronegative and structural features of duplex DNA (15–17).

Unlike phage PBS2, the genome of the *B. subtilis* phage φ29 does not contain uracil residues. It is a linear double-stranded DNA, with a terminal protein (TP) covalently linked at both 5′-ends. Replication of φ29 DNA starts nonsimultaneously at either DNA end, where the replication origins are located, by a protein-priming mechanism (18). Electron microscopy studies revealed two main types of φ29 replicative intermediates in infected cells, named type I and type II (19, 20) (Fig. 1*A*). Type I intermediates are unit-length linear double-stranded DNA molecules with one or more single-stranded branches of varying lengths. Type II intermediates are unit-length linear molecules in which a region of the DNA starting from one end is double-stranded, and the adjacent region containing the other end is single-stranded. Thus, both replicative intermediates have long stretches of single-stranded DNA. Most unexpectedly, during the functional characterization of the early viral protein p56 (56 amino acids), we found that p56 interacted with the *B. subtilis* UDG and inhibited its activity. Furthermore, we demonstrated that extracts from φ29-infected cells lacked UDG but not AP endonuclease activity. UDG inhibitors encoded by non-uracil containing viral DNAs were not identified previously. We propose that inhibition of the host UDG by protein p56 ensures the integrity of the φ29 replicative intermediates if uracil residues arise either by cytosine deamination or dUMP misincorporation.

* This work was supported in part by Ministerio de Ciencia y Tecnología Grant BMC2002-03818 (to M. S.) and by Comunidad Autónoma de Madrid Grant GR/SAL/0087/2004 (to A. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported by the Ministerio de Ciencia y Tecnología (Programa Ramón y Cajal). To whom correspondence should be addressed: Centro de Biología Molecular "Severo Ochoa," Universidad Autónoma, Cantoblanco, 28049 Madrid, Spain. Tel.: 34-91-497-8435; Fax: 34-91-497-8490; E-mail: abravo@cbm.uam.es.

[2] The abbreviations used are: BER, base excision repair; UDG, uracil-DNA glycosylase; AP, apurinic-apyrimidinic; $Me_2SO$, dimethyl sulfoxide; DSP, dithiobis(succinimidylpropionate); PDH, pyruvate dehydrogenase; TP, terminal protein; Tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine; ssDNA, single-stranded DNA; E2, dihydrolipoamide acetyltransferase; E3, dihydrolipoamide dehydrogenase; cfu, colony-forming units.

## EXPERIMENTAL PROCEDURES

*Bacterial Strains, Bacteriophages, and Plasmids*—*B. subtilis* 110NA (23), a nonsuppressor ($su^-$) strain, and *B. subtilis* MO-101-P (24), a suppressor strain ($su^{+44}$), were used. Phage $\phi$29 *sus*4 (56) (23) was used in cross-linking experiments. This phage carries a suppressor-sensitive mutation in gene 4, which encodes an activator of the late transcription (25). Phage $\phi$29 *sus*14(1242) (26) was used to measure UDG activity in extracts of infected cells. This phage contains a suppressor-sensitive mutation in gene 14, which encodes the holin protein (27). Thus, cell lysis is delayed. Phage stocks were prepared as reported (28).

To construct plasmid pCR2.1-TOPO.p56, the TOPO TA cloning kit (Invitrogen) was used. Basically, a $\phi$29 DNA region, which contains gene 56, was amplified by the PCR using the oligonucleotides 5′-CGCATTGTATGAGCTTTCTAGGATGG-3′ and 5′-GCAGGGAATT**CTGCAG**TCAAAGGACTTTATC-3′ as primers. The 267-bp PCR-synthesized fragment was cloned into the *E. coli* expression vector pCR2.1-TOPO (Invitrogen), which is based on the T7 promoter. Transformed cultures of the *E. coli* TOP10 strain were plated on LB plates containing 50 $\mu$g/ml of kanamycin. To construct plasmid pPR53.p56, the 272-bp PstI restriction fragment of plasmid pCR2.1-TOPO.p56, which carries gene 56, was inserted into the PstI site of the *B. subtilis* constitutive expression vector pPR53, which is based on the $P_R$ promoter of phage λ (29). The *B. subtilis* strain YB886 (30) was used for cloning, and transformants were selected for phleomycin resistance (0.8 $\mu$g/ml). Gene 56 was engineered to encode a FLAG-tagged p56 protein (p56FLAG). To this end, a two-step mutagenesis method was used. In the first step, gene 56 was amplified by PCR using plasmid pPR53.p56 as template and the oligonucleotides A (5′-CCTCTAGAGTCGAC**CTGCAG**-3′) and B (5′-GTCATCGTCATCCTTATAGTCAGGACTTTATCCAACCTTAG-3′) as primers. In the second step, the 298-bp PCR-synthesized fragment was used as template and the oligonucleotides A and C (5′-CCCTCAGGG**CTGCAG**TTATTACTTGTCATCGTCATCCTTATAGTC-3′) as primers. Oligonucleotides A and C include a PstI restriction site (boldface). The 322-bp PCR-amplified fragment was further digested with PstI, and the 293-bp digestion product was inserted into the PstI site of the *B. subtilis* vector pPR53 (plasmid pPR53.p56FLAG).

*Phage Growth under One-step Conditions*—*B. subtilis* 110NA cells were exponentially grown at 30 °C in LB medium supplemented with 5 mM $MgSO_4$ to an absorbance at 560 nm ($A_{560}$) equivalent to $\sim 10^8$ colony-forming units (cfu) per ml. The culture was then infected with $\phi$29 at a multiplicity of 5–10. After 10 min of incubation with gentle shaking, unadsorbed phages were eliminated by centrifugation of the infected culture. Cells were resuspended in the same volume of medium and incubated with vigorous shaking for the indicated time.

*Immunoblotting*—Gel-separated proteins were transferred electrophoretically to Immobilon-P membranes (Millipore) using a Mini Trans Blot (Bio-Rad) at 100 mA and 4 °C for 60 min. Transfer buffer contained 25 mM Tris, 192 mM glycine, 20% methanol. Membranes were probed with anti-p56 serum for 60 min. Antigen-antibody complexes were detected using anti-rabbit horseradish peroxidase-conjugated antibodies and ECL Western blotting detection reagents (Amersham Biosciences). For quantitative immunoblotting (29), cell extracts were prepared from a known number of viable cells, which was determined before phage addition. Increasing amounts of the cell extract and known amounts of purified protein p56 were run in the same gel.

*In Vivo Chemical Cross-linking*—Bacteria were washed with 50 mM Hepes, pH 8.0, and concentrated 20-fold in buffer P (50 mM Hepes, 10 mM EDTA, 20% sucrose, pH 8.0). The cross-linker dithiobis(succinimidylpropionate) (DSP) (Pierce) was dissolved in $Me_2SO$ just before use. DSP was added to the culture at the indicated concentration. After incubation at room temperature for 20 min, Tris-HCl, pH 7.5, was added at a final concentration of 150 mM to quench the reaction. Cells were harvested by centrifugation, resuspended in loading buffer without $\beta$-mercaptoethanol (60 mM Tris-HCl, pH 6.8, 2% SDS, 30% glycerol), and disrupted by sonication.

*Isolation of p56FLAG Complexes*—*B. subtilis* 110NA cells carrying plasmid pPR53.p56FLAG were exponentially grown in LB medium containing phleomycin (0.8 $\mu$g/ml) at 30 °C to an $A_{560}$ equivalent to $\sim 10^8$ cfu/ml. Cells were concentrated 10-fold in buffer TBS (50 mM Tris-HCl, pH 7.5, 150 mM NaCl) and disrupted by French pressure treatment (20,000 p.s.i.). The whole-cell extract was centrifuged at 7,000 rpm and 4 °C in a Sorvall SS.34 rotor for 10 min. The supernatant was loaded onto an anti-FLAG M2 affinity column (Sigma) under gravity flow. Later, the column was extensively washed with buffer TBS. Proteins bound to the column were eluted with buffer TBS containing FLAG-peptide (500 $\mu$g/ml) (Sigma). Eluted proteins were precipitated with acetone and resuspended in loading buffer (60 mM Tris-HCl, pH 6.8, 2% SDS, 5% $\beta$-mercaptoethanol, 30% glycerol).

*Peptide Mass Fingerprinting*—Gel-separated proteins were stained with SyproRuby and subjected to *in situ* digestion with trypsin, as described (31). Peptide masses were measured with a matrix-assisted laser desorption/ionization time of flight mass spectrometer Autoflex (Bruker Daltonic; Bremen, Germany) equipped with a reflector and employing 2,5-dihydroxybenzoic acid as matrix and a Anchor-Chip surface target. The mass spectra were fitted to data bases by the program Mascot (32).

*Purification of Protein p56*—*E. coli* BL21(DE3) cells carrying plasmid pCR2.1-TOPO.p56 were grown in LB medium containing kanamycin (50 $\mu$g/ml) at 34 °C. When the culture reached an $A_{560}$ of 0.9, isopropyl 1-thio-$\beta$-D-galactopyranoside was added to a final concentration of 0.5 mM. After 30 min, cells were incubated with rifampicin (120 $\mu$g/ml) for 75 min. Cells were collected by centrifugation and frozen at −70 °C before being used. Protein p56 was purified under ice-cold conditions using the following protocol. Cells were ground with alumina in buffer A (50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 7 mM $\beta$-mercaptoethanol, 5% glycerol) containing 0.65 M NaCl. After removal of alumina and cell debris by centrifugation, the cleared lysate was mixed with polyethyleneimine (0.3%), incubated on ice for 20 min, and centrifuged at 12,000 rpm in a Sorvall GSA rotor for 10 min. The supernatant was made 0.3 M NaCl with buffer A and centrifuged as before for 20 min. The resulting pellet was washed with buffer A containing 0.7 M NaCl. After centrifugation (12,000 rpm in a Sorvall SS34 rotor for 20 min), the supernatant was processed by stepwise ammonium sulfate precipitation at 65, 45, and finally 30%. Proteins were recovered from the 30% ammonium sulfate supernatant by raising it to 50% ammonium sulfate. This last pellet was resuspended in buffer A to a final salt concentration of 55 mM (estimated by conductivity measurements). The protein preparation was then loaded onto a Mono Q column equilibrated with buffer A containing 55 mM NaCl. Protein p56 was eluted from the column with 0.3 M NaCl. This fraction was further loaded onto a 15–30% glycerol gradient and subjected to centrifugation at 62,000 rpm in a Beckman SW.65 rotor for 20 h. Fractions containing p56 were pooled, precipitated with ammonium sulfate to 70% saturation, resuspended in buffer A containing 50% glycerol, and stored at −70 °C.

*Interaction between p56 and E. coli UDG in Vitro*—The reaction mixture (10 $\mu$l) contained 20 mM Tris-HCl, pH 8.0, 1 mM dithiothreitol, 1 mM EDTA, 20 mM NaCl, 20% glycerol, 2 $\mu$g of protein p56, and 0.2 $\mu$g of *E. coli* UDG (New England Biolabs). After incubation at room temperature for 15 min, the sample was kept at 4 °C for 15 min and analyzed by

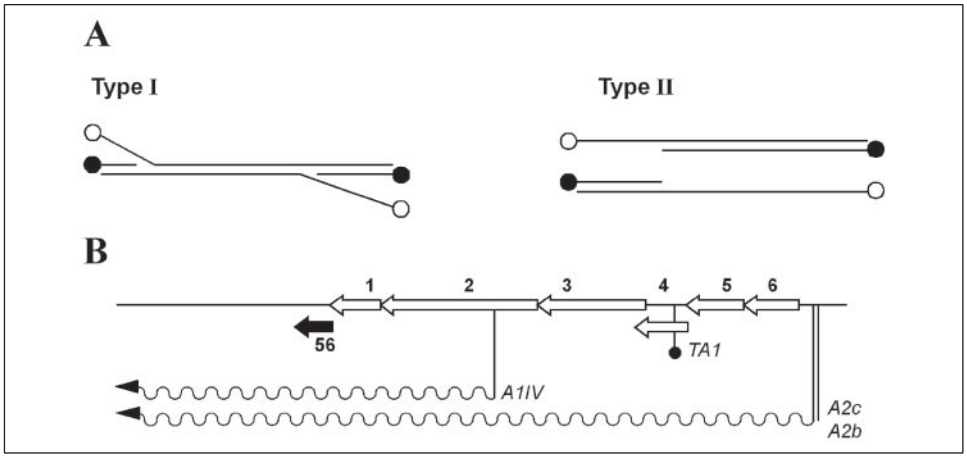


FIGURE 1. *A,* types of $\phi$29 replicative intermediates. The genome of $\phi$29 is a linear double-stranded DNA with a TP covalently linked at each 5′-end (parental TP; *white circle*). A free molecule of the TP (primer TP; *black circle*) provides the hydroxyl group needed by the viral DNA polymerase to start DNA synthesis at both $\phi$29 DNA ends (protein-priming replication mechanism) (21). During elongation, $\phi$29 DNA polymerase catalyzes highly processive polymerization coupled to strand displacement (22), and consequently, complete replication of both strands proceeds continuously from each priming event. See Introduction for more details. *B,* genetic and transcriptional map of the left end of the $\phi$29 genome. Only relevant features are shown. *White arrows* indicate the location of genes previously identified. The *black arrow* indicates the position of gene 56, which has been identified in this work. Genes 6 (double-stranded DNA binding protein), 5 (single-stranded DNA binding protein), 3 (TP), and 2 (DNA polymerase) are essential for *in vivo* $\phi$29 DNA replication (18). Gene 1 encodes a membrane-localized protein involved in the membrane association of $\phi$29 DNA replication. This protein enhances the rate of viral DNA replication *in vivo* (33). *Wavy lines* represent transcripts from the indicated early promoters. The position of the transcriptional terminator *TA1* is indicated.

nondenaturing PAGE (16% polyacrylamide). Gel electrophoresis was performed at 4 °C.

*UDG and AP Endonuclease Activities in B. subtilis Extracts—B. subtilis* strain 110NA was grown to mid-log phase in LB medium at 30 °C ($\sim 10^8$ cfu/ml). Cells were concentrated 10-fold in buffer U (50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 12 mM $\beta$-mercaptoethanol, 1 mM EDTA) containing a protease inhibitor mixture (1 tablet of Complete, Mini, EDTA-free per 10 ml) from Roche Applied Science. Cells were disrupted by French pressure treatment (20,000 p.s.i.). After centrifugation at 7,000 rpm and 4 °C in a Sorvall SS.34 rotor for 10 min, the supernatant (extract) was kept at 4 °C for up to 2 weeks. Total protein concentration (1.35 mg/ml) was determined by the Lowry method using a NanoDrop ND-1000 spectrophotometer. To measure UDG activity, a 34-mer oligonucleotide containing a single uracil residue at position 16 (ssDNA-$U^{16}$) (from Isogen) was used as substrate. It was 5′-labeled with [$\gamma$-$^{32}$P]ATP (3,000 Ci/mmol) (Amersham Biosciences) and T4 polynucleotide kinase (New England Biolabs). Reaction mixtures (20 $\mu$l) contained the indicated amount of extract and radiolabeled substrate in buffer U carrying protease inhibitors (see above). After incubation at 37 °C for 10 min, samples were treated with NaOH to a final concentration of 0.2 M and heated at 90 °C for 30 min. Samples were then dried in a SpeedVac, resuspended in 10 $\mu$l of formamide loading buffer (95% formamide, 20 mM EDTA, 0.05% xylene cyanol, 0.05% bromphenol blue), and subjected to electrophoresis in 8 M urea, 20% polyacrylamide gels. To measure AP endonuclease activity, the radiolabeled ssDNA-$U^{16}$ oligonucleotide was incubated with *E. coli* UDG (New England Biolabs), generating a 34-mer oligonucleotide with an AP site at position 16 (ssDNA-$AP^{16}$). This product was used as substrate. Reaction mixtures (20 $\mu$l) contained 50 mM Tris-HCl, pH 8.0, 12 mM NaCl, 12 mM $\beta$-mercaptoethanol, 10 mM $MgCl_2$, and the indicated amount of extract and radiolabeled substrate. After incubation at 37 °C for 15 min, 12 $\mu$l of the formamide loading buffer was added to the reactions. Purified human AP endonuclease (APE1) was purchased from Trevigen.

## RESULTS

*Protein p56 Accumulates throughout the $\phi$29-Infective Cycle*—The left end of the $\phi$29 genome contains genes 6, 5, 4, 3, 2, and 1 (Fig. 1*B*). With the exception of gene 4, which encodes a transcriptional regulator protein, the mentioned genes are involved in phage DNA replication (18, 33). Analysis of the nucleotide sequence downstream gene 1 revealed the existence of an open reading frame (ORF56) that would encode an acidic protein of 56 amino acids (protein p56). ORF56 would be mainly transcribed from two strong early promoters, named *A2c* and *A2b*, into a polycistronic RNA (Fig. 1*B*). Both promoters are partially repressed at late times of infection (25). In addition, ORF56 would be transcribed from the weak early promoter *A1IV*, which is located within gene 2 (34–36). To analyze whether protein p56 was synthesized during $\phi$29 infection, *B. subtilis* cells were infected with phage $\phi$29 under one-step growth conditions. At different times after infection, whole-cell extracts were obtained by sonication, and total proteins were analyzed by immunoblotting. A protein recognized by antibodies against p56 accumulated throughout the $\phi$29 lytic cycle (Fig. 2*A*). Overexposure of the blot shown in Fig. 2*A* allowed detection of p56 at 15 min of $\phi$29 infection (not shown). Quantitative immunoblotting indicated that protein p56 was present in $\sim 10^4$ molecules per cell at early stages of infection (15 min), and it increased up to $\sim 10^5$ molecules per cell at late stages of infection (50 min).

*Protein p56 Interacts with a Host Protein during $\phi$29 Infection*—To find out whether protein p56 interacted with a viral and/or cellular protein during $\phi$29 infection, we carried out *in vivo* chemical cross-linking experiments. Specifically, at 40 min of viral infection, cells were incubated with dithiobis(succinimidylpropionate) (DSP). This homobifunctional cross-linker reacts significantly with the $\epsilon$-amine of lysine residues. Protein p56 has four lysines. Following this treatment, whole-cell extracts were prepared by sonication. Total proteins were separated by SDS-Tricine-PAGE and analyzed by immunoblotting using antibodies against p56. As control, noninfected cells were incubated with DSP (Fig. 2*B*). The anti-p56 antibodies cross-reacted with a bacterial protein that migrated at ~20 kDa (Fig. 2*B, protein X*). In infected cells treated with DSP, in addition to protein p56, a major specific cross-linked product migrating at ~35 kDa (p56 complex) was detected. We further investigated whether the p56 complex was formed in the absence of other viral proteins. To this end, gene 56 was cloned into the *B. subtilis* expression vector pPR53 (29). Cells carrying the recombinant plasmid (pPR53.p56) constitutively synthesized protein p56 at levels slightly lower than those detected at 40 min of phage infection (Fig. 2*B*). When

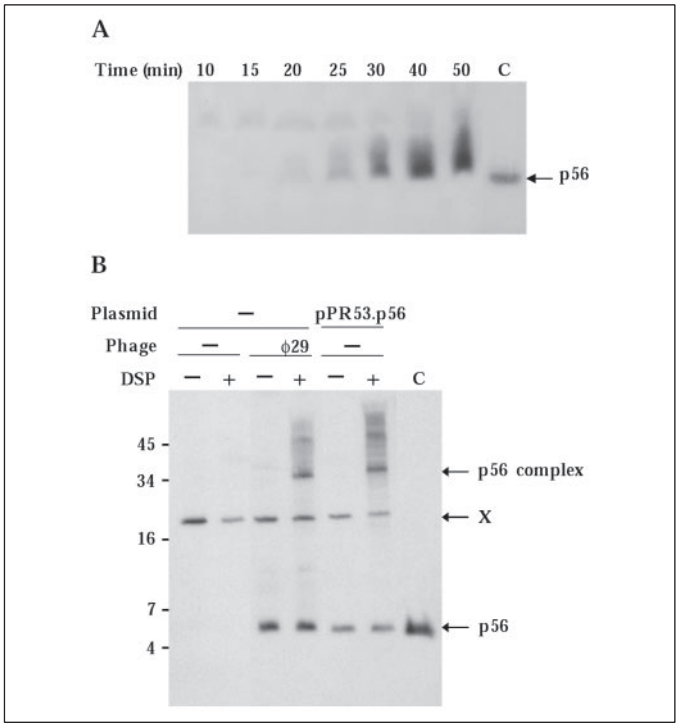


FIGURE 2. *A,* synthesis of protein p56 throughout the $\phi$29 lytic cycle. *B. subtilis* 110NA cells growing in LB medium at 30 °C were infected with the wild-type phage at a multiplicity of infection of 5. Phage addition marked the zero time of the experiment. Total lysis of the culture was observed 70 min after infection. At the indicated times, cell extracts were prepared, and total proteins were separated by SDS-Tricine-PAGE (37). The amount of extract loaded onto the gel corresponds to $10^8$ cells. The gel was processed for Western blotting using anti-p56 serum. Purified protein p56 was run in the same gel (*lane C*). *B*, interaction of protein p56 with an ~28 kDa host protein. *B. subtilis* 110NA cells growing in LB medium at 30 °C were infected with $\phi$29 *sus*4 (56) under one-step conditions. In these cells, the $\phi$29 regulatory protein p4 is not synthesized, and consequently, late viral transcription is not activated (25). The late genes encode components of the viral capsids, proteins involved in phage morphogenesis and those required for cell lysis. Hence, the cells remain intact at late stages of infection, which is needed to perform *in vivo* chemical cross-linking experiments (38). *B. subtilis* 110NA cells carrying plasmid pPR53.p56 were exponentially grown in LB medium at 30 °C to ~$10^8$ cfu/ml. DSP (200 $\mu$M) was used as cross-linker. Total proteins were separated by SDS-Tricine-PAGE (37). The gel was processed for immunoblotting using anti-p56 serum. The molecular mass of pre-stained proteins used as markers (Invitrogen) is indicated on the *left* in kDa. Purified protein p56 was run in the same gel (*lane C*).

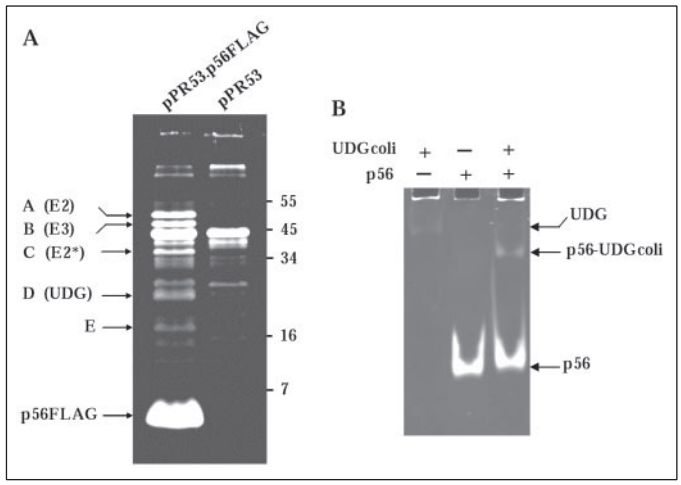


FIGURE 3. *A,* affinity purification of p56FLAG complexes. Anti-FLAG M2 affinity columns (Sigma) were used. Protein p56FLAG and associated proteins were eluted by competition with FLAG peptide. After acetone precipitation, eluted proteins were resolved by SDS-Tricine-PAGE and stained with SyproRuby (Molecular Probes). Molecular mass markers are indicated on the *right* in kDa. The peptide mass fingerprint search program Mascot and the NCBI data base were used for protein identification. Accession numbers are as follows: *A,* gi 16078524; *B,* gi 16078525; *C,* gi 16078524; and *D,* gi 16080848. *B*, interaction between protein p56 and *E. coli* UDG *in vitro*. Protein p56 (2 $\mu$g) and UDG purified from *E. coli* (0.2 $\mu$g; New England Biolabs) were mixed and analyzed by nondenaturing PAGE (16% polyacrylamide). The gel was stained with SyproRuby.

these cells were incubated with DSP, the ~35-kDa p56 complex was again detected. Hence, the viral protein p56 was able to interact with a host protein both during the infective process and in the absence of viral components. Taking into account the molecular mass of p56 (6.6 kDa) and DSP (0.4 kDa), the molecular mass of the host protein would be ~28 kDa.

*Protein p56 Interacts with UDG*—To identify the cell target for protein p56, gene 56 was engineered to encode a FLAG-tagged p56 protein (p56FLAG). This variant of p56 carries the peptide DYKDDDDK fused to its C-terminal end. The mutant gene was cloned into the *B. subtilis* expression vector pPR53 (plasmid pPR53.p56FLAG). By immunoblotting, we verified that *B. subtilis* cells carrying the recombinant plasmid constitutively synthesized protein p56FLAG at levels similar to those detected for p56 at late stages of $\phi$29 infection (not shown). Extracts from these cells were applied to an anti-FLAG affinity column. The pure FLAG peptide was then used to elute p56FLAG. As a negative control, extracts from *B. subtilis* cells harboring plasmid pPR53 were used. The eluted proteins were separated by SDS-Tricine-PAGE and stained with SyproRuby (Fig. 3*A*). Protein p56FLAG co-eluted with five proteins (named A to E) that were absent in the control sample. With the exception of protein E, such proteins were identified by peptide mass fingerprinting using the MASCOT search program (32). Proteins A and B were identified as the E2 (47.5 kDa; dihydrolipoamide acetyltransferase) and E3 (49.7 kDa; dihydrolipoamide dehydrogenase) subunits of the pyruvate dehydrogenase multienzyme complex (PDH), respectively. Their mobility in denaturing polyacrylamide gels was anomalous, as reported previously (39). Protein C was a degradation product of the PDH E2 subunit (E2*), whereas protein D was identified as the UDG DNA repair enzyme (26 kDa). Therefore, the affinity chromatography results suggested that protein p56 interacted with a subunit of the PDH complex (see "Discussion") and with UDG. Because the *in vivo* chemical cross-linking experiments (Fig. 2*B*) revealed that a protein of ~28 kDa was a cell target for p56, we performed additional experiments to confirm the potential interaction between p56 and UDG (see below).

The *B. subtilis* UDG is highly homologous to the *E. coli* UDG (Family-1 UDGs) (3, 40). Therefore, we analyzed whether protein p56 was able to interact with UDG purified from *E. coli*. To this end, protein p56 was incubated with *E. coli* UDG for 15 min at room temperature, and the reaction mixture was analyzed by nondenaturing PAGE. As shown in Fig. 3*B*, free protein p56, but not free UDG, was detected. Moreover, a protein band migrating faster than UDG was visualized (p56-UDG-coli). This band contained p56 and UDG, as confirmed by immunoblot analysis and peptide mass fingerprinting, respectively (not shown). In conjunction with the above, this result demonstrated that protein p56 forms a complex with UDG.

*Protein p56 Functions as an Inhibitor of the B. subtilis UDG*—Family-1 UDGs are able to excise uracil base efficiently from both single-stranded and double-stranded DNAs. Furthermore, they are specifically sensitive to Ugi (3), a UDG inhibitor encoded by phage PBS2. Removal of uracil by UDGs generates an AP site that can be processed by hydrolytic AP endonucleases. In the absence of an AP endonuclease activity, chemical cleavage of the DNA at the AP site can be achieved by treatment with heat and alkali.

To investigate whether protein p56 was an inhibitor of the *B. subtilis* UDG, we set up an assay to measure UDG activity in *B. subtilis* extracts. In the first experiment, a 34-mer single-stranded oligonucleotide containing a single uracil residue at position 16 (ssDNA-$U^{16}$) was incubated

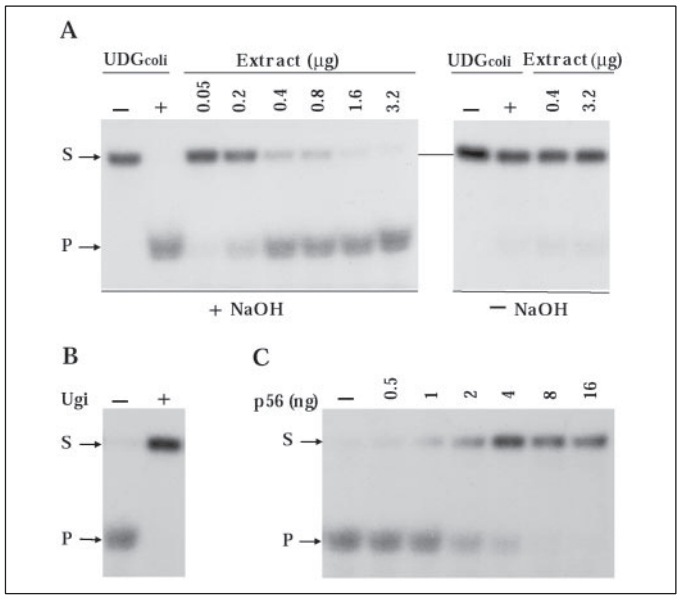


FIGURE 4. *A,* UDG activity in *B. subtilis* extracts. The 5′-end $^{32}$P-labeled ssDNA-U$^{16}$ substrate (*S*) (0.55 ng) was incubated with the indicated amount of extract in the absence of $Mg^{2+}$. After 10 min, the reaction mixtures were treated or not with NaOH. Formation of the cleavage product (*P*) was monitored by autoradiography after resolution on 8 M urea, 20% polyacrylamide gels. As internal control, UDG from *E. coli* (0.5 units; New England Biolabs) was used. *B,* inhibition of the *B. subtilis* UDG activity by Ugi. The substrate was incubated with 1.6 μg of extract in the absence or presence (0.1 units) of Ugi (New England Biolabs). Reactions were treated with NaOH. *C,* inhibition of the *B. subtilis* UDG activity by protein p56. The indicated amount of p56 was added to 1.6 μg of extract. Reactions were treated with NaOH.

with increasing amounts of a *B. subtilis* extract. After 10 min, NaOH was added to the reaction mixtures. As shown in Fig. 4*A*, a cleavage product was generated using 0.2 μg of extract. Moreover, the substrate was almost totally cleaved when 1.6 μg of extract were used. The same cleavage product was detected when the substrate was incubated with UDG purified from *E. coli*. Therefore, the *B. subtilis* extract was capable of removing uracil from single-stranded DNA generating an AP site (UDG activity). However, under the assayed conditions, the extract lacked AP endonuclease activity, because cleavage of the substrate did not occur when the reactions were not treated with NaOH (Fig. 4*A*). Furthermore, the UDG activity of the *B. subtilis* extract was sensitive to the presence of Ugi (Fig. 4*B*) and was able to excise uracil from double-stranded DNA bearing a G:U mismatch (not shown). Collectively, these results demonstrated the presence of Family-1 UDG in the *B. subtilis* extract.

We next examined whether the viral protein p56 was able to inhibit the UDG activity of the *B. subtilis* extract (Fig. 4*C*). To this end, the ssDNA-U$^{16}$ substrate was incubated with 1.6 μg of extract in the presence of different amounts of purified protein p56 (from 0.5 to 16 ng). After 10 min, NaOH was added to the reactions. A decrease in the amount of the cleavage product was detected with 2 ng of p56. Moreover, cleavage of the substrate did not take place in the presence of 8 ng of p56. Therefore, these results showed that protein p56 acts as an inhibitor of the *B. subtilis* UDG.

*Inhibition of UDG Activity in ϕ29-Infected Cells*—If protein p56 functions as a UDG inhibitor during ϕ29 infection, extracts from infected cells should be deficient in UDG activity. To test this hypothesis, *B. subtilis* cells were infected with ϕ29 under one-step growth conditions, and extracts were prepared at 30 min of infection. Then the ssDNA-U$^{16}$ substrate was incubated, in the absence of $Mg^{2+}$, with increasing amounts of the infected extract, and after 10 min NaOH was added to the reactions (Fig. 5*A*). As controls, total cleavage of the substrate was obtained with purified *E. coli* UDG and with extract from noninfected cells. On the contrary, nearly 90% of the substrate remained intact when it was incubated with 0.2 μg of the infected extract. This percentage did not significantly decrease when higher amounts of infected extract (up to 6.4 μg) were used. However, total cleavage of the substrate was observed when *E. coli* UDG was added to the infected extract (3.2 μg). Therefore, extracts from ϕ29-infected cells showed a drastic reduction in UDG activity.

AP sites generated by UDGs can be processed by hydrolytic AP endonucleases. The major AP endonuclease in *B. subtilis* is the ExoA protein. Like human AP endonuclease APE1, ExoA catalyzes the cleavage of the phosphodiester bond 5′ to the AP site, leaving a 3′-hydroxyl group. Moreover, it was shown that both AP endonucleases cleave single-stranded oligonucleotides containing an AP site (7, 8). To find out whether other base excision DNA repair enzymes were inactivated during ϕ29 infection, we measured AP endonuclease activity in extracts from noninfected and infected cells. In this assay, a 34-mer single-stranded oligonucleotide with a single AP site at position 16 (ssDNA-AP$^{16}$) was used as substrate. As positive control, the human APE1 enzyme was used. As shown in Fig. 5*B*, AP endonucleolytic activity was detected in the infected extract (1.6 μg), as well as in the noninfected extract (1.6 μg). From these results, we conclude that UDG activity, but not AP endonuclease activity, is inhibited by p56 during ϕ29 infection.

## DISCUSSION

Most prokaryotic and eukaryotic cells encode the UDG enzyme, which is involved in the BER pathway. It specifically removes uracil from DNA. In this work, we have demonstrated that the lytic phage ϕ29 encodes an inhibitor of the *B. subtilis* UDG. This inhibitor is an acidic protein of 56 amino acids, named p56. Hence, protein p56 is the first example of a UDG inhibitor encoded by a non-uracil containing viral DNA.

We have measured UDG activity in *B. subtilis* extracts. This enzymatic activity was $Mg^{2+}$-independent and excised uracil residues from both single-stranded and double-stranded DNAs. Addition of purified protein p56 to the *B. subtilis* extract inhibited the endogenous UDG activity. Moreover, a drastic reduction in UDG activity was observed in extracts from ϕ29-infected cells, whereas the AP endonuclease activity remained intact. Therefore, the early viral protein p56 functions as an inhibitor of the cellular UDG during ϕ29 infection. This conclusion was supported by protein-protein interaction experiments. A complex formed by p56 and a host protein of ~28 kDa was detected in infected cells using chemical cross-linking techniques. Moreover, UDG (26 kDa) co-eluted with p56FLAG when extracts of *B. subtilis* cells producing p56FLAG and anti-FLAG affinity columns were used. Protein p56 also formed a complex with *E. coli* UDG *in vitro*, as determined by native PAGE. *B. subtilis* UDG has strong sequence homology to *E. coli* UDG (40). Both enzymes belong to a family of highly conserved DNA glycosylases, which remove uracil from both single-stranded and double-stranded DNAs (3).

By affinity chromatography, we found that in addition to UDG, E2 and E3 co-eluted with protein p56FLAG. E2 and E3 are components of the PDH multienzyme complex, which catalyzes the irreversible oxidative decarboxylation of pyruvate to acetyl-coenzyme A. PDH complexes of Gram-positive bacteria have a core of 60 E2 subunits with icosahedral symmetry (41). Besides its role in oxidative metabolism, the E2 subunit appears to have DNA binding activity (42, 43). The low isoelectric point of protein p56 suggests that it might act as a double-stranded DNA mimic, like the UDG inhibitor Ugi and the highly acidic protein HI1450 from *Haemophilus influenzae* (15, 16, 44). If this were the case, co-

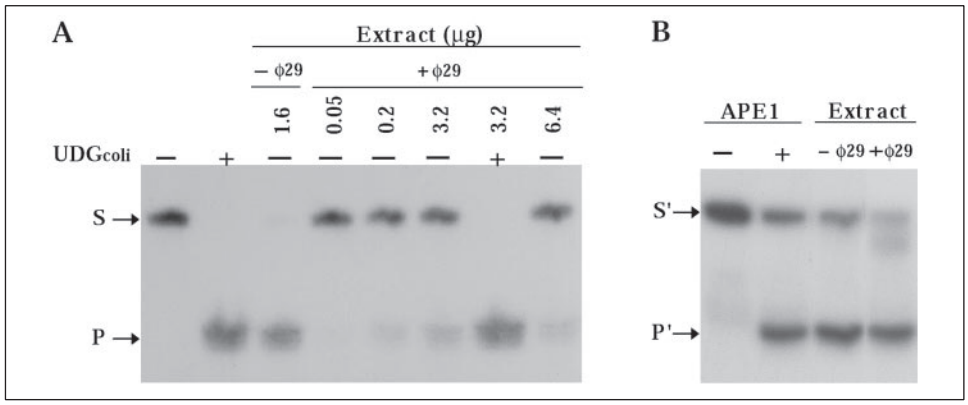


FIGURE 5. *A*, lack of UDG activity in extracts of $\phi$29-infected cells. The 5′-end $^{32}$P-labeled ssDNA-U$^{16}$ substrate (*S*) (0.55 ng) was incubated with the indicated amount of extract in the absence of $Mg^{2+}$. Reactions were treated with NaOH. See legend to Fig. 4*A* for more details. *B*, AP endonuclease activity in extracts of $\phi$29-infected cells. The 5′-end $^{32}$P-labeled ssDNA-AP$^{16}$ substrate (*S′*) (1.6 ng) was incubated with 1.6 $\mu$g of the indicated extract in the presence of $Mg^{2+}$. As internal control, human AP endonuclease (APE1; 0.1 units; from Trevigen) was used.

elution of the E2 subunit with p56FLAG could be related to the ability of E2 to interact with DNA. Nevertheless, at present, it is unknown whether a p56-E2 complex is formed during $\phi$29 infection.

Inhibition of the cellular UDG activity after phage infection was established previously in two systems. The first example was the inhibition of the *B. subtilis* UDG after infection with phage PBS2 (11). The need for this inhibition was obvious, because the DNA genome of PBS2 contains uracil instead of thymine residues. Thus, phage PBS2 has developed a defensive mechanism against the action of the host UDG. The UDG inhibitor encoded by phage PBS2 was identified as an acidic protein of 84 amino acids (Ugi) (13). Less obvious was the inhibition of the *E. coli* UDG after infection with phage T5, because its DNA genome does not contain uracil residues. In this case, the UDG inhibitor has not yet been identified, although early studies suggested that it could be a pre-early phage-encoded protein (45). In fact, the pre-early region of phage T5 contains six small open reading frames that presumably encode host function inhibitors (46).

Why does phage $\phi$29 encode an inhibitor of the cellular UDG? We think that this inhibition is related to the mechanism of $\phi$29 DNA replication. As depicted in Fig. 6, replication of the linear $\phi$29 DNA starts nonsimultaneously at both ends, where the replication origins are located, using a free molecule of the TP as primer. Once initiated, replication proceeds by a strand displacement mechanism, and consequently, replicative intermediates (type I and type II) with long stretches of single-stranded DNA are generated (18). If uracil arises in the $\phi$29 genome, either by misincorporation of dUMP or by cytosine deamination, and the damage is not repaired before DNA replication, type I replicative intermediates carrying a uracil residue on single-stranded DNA could appear. The presence of uracil in the replicative intermediate could recruit components of the cellular BER pathway, such as UDGs and AP endonucleases. As shown in this work, *B. subtilis* cells synthesize an enzymatic activity that efficiently removes the aberrant base uracil from single-stranded DNA, generating an AP site. The subsequent action of an AP endonuclease activity would introduce a nick into the phosphodiester backbone with accompanying loss of the terminal DNA region. It has been shown that the ExoA protein of *B. subtilis* has AP endonuclease activity on single-stranded DNA (7). When two replication forks moving in opposite directions merge, type II replicative intermediates are formed. Then DNA synthesis would continue, and a shorter viral DNA molecule lacking one parental TP would be generated. Therefore, the action of the cellular UDG on single-stranded DNA regions of the $\phi$29 replicative intermediates would be harmful for viral replication.

Uracilation of DNA represents a constant threat to the survival of many organisms, including viruses. More than likely, phage $\phi$29 has developed alternative strategies to protect its genome from uracilation.

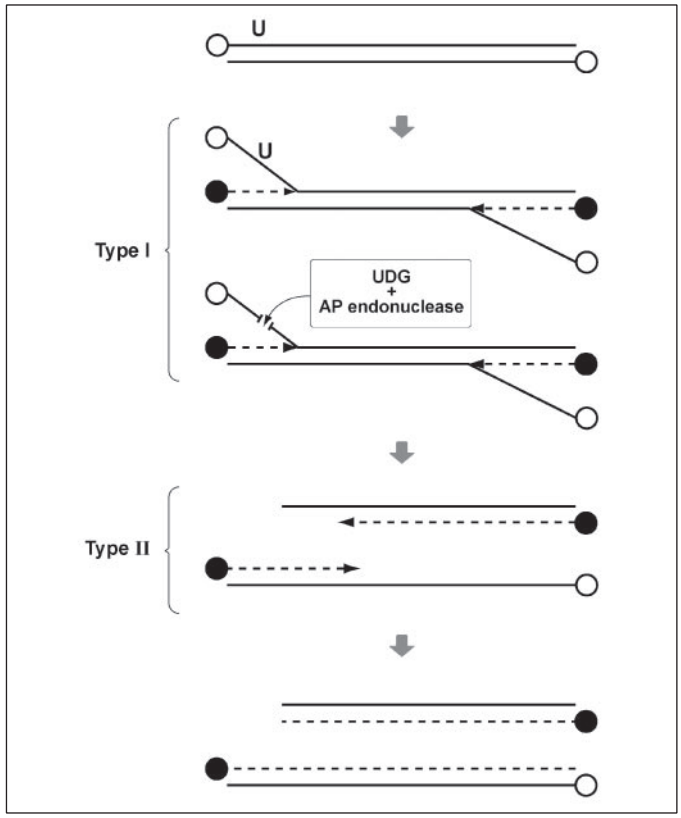


FIGURE 6. **Model of genome instability produced by the action of UDG on $\phi$29 replicative intermediates.** Parental TP (*white circle*) and primer TP (*black circle*) are indicated. *Dashed lines* indicate newly synthesized viral DNA. See "Discussion" for details.

One possibility may be the recruitment of the cellular deoxyuridine triphosphatase enzyme (dUTPase), which maintains a low dUTP:dTTP ratio and, consequently, minimizes the misincorporation of uracil into DNA. Most interestingly, herpesviruses, poxviruses, and certain retroviruses encode dUTPase, whose function is thought to be associated with the ability of these viruses to replicate in cells that produce low levels of dUTPase (1). Another possible strategy of $\phi$29 to counteract the accumulation of uracil in its genome may be the recruitment of a mismatch-specific UDG activity. For example, *E. coli* encodes a double-strand specific enzyme capable of removing uracil residues from G:U mismatches arising through spontaneous cytosine deamination. Such an enzyme is related to the human thymine-DNA glycosylase and is insensitive to inhibition by Ugi (47).

The generation of replicative intermediates containing long stretches of single-stranded DNA is a feature of the protein-primed mechanism

of DNA replication. Therefore, it is expected that gene 56 is conserved in the genome of other $\phi$29-related phages. In the case of phage B103, there is an open reading frame that would encode a 56-amino acid protein, whose deduced sequence has a high level of homology to protein p56 (64% identity and 75% similarity) (48). Furthermore, in phage GA-1, which is the most distantly $\phi$29-related phage, there is an open reading frame that would encode a 130-amino acid protein. The region of this putative protein spanning amino acids 27 and 82 shows 23% identity and 27% similarity with protein p56, suggesting that they may have similar functions.[3]

To conclude, protein p56 of phage $\phi$29 functions as an inhibitor of the cellular Family-1 UDG. This inhibition is likely a defense mechanism developed by $\phi$29 and $\phi$29-related phages to prevent the action of the BER pathway if uracil residues arise in the replicative intermediates.

*Acknowledgments—We are very grateful to Dr. G. del Solar for help with the TOPO TA cloning system and to J. M. Lázaro for advice in p56 purification. The institutional help of Fundación Ramón Areces to the Centro de Biología Molecular "Severo Ochoa" is acknowledged.*

## REFERENCES


1. Chen, R., Wang, H., and Mansky, L. M. (2002) *J. Gen. Virol.* **83,** 2339–2345
2. Lindahl, T. (1974) *Proc. Natl. Acad. Sci. U. S. A.* **71,** 3649–3653
3. Pearl, L. H. (2000) *Mutat. Res.* **460,** 165–181
4. Eftedal, I., Guddal, P. H., Slupphaug, G., Volden, G., and Krokan, H. E. (1993) *Nucleic Acids Res.* **21,** 2095–2101
5. Panayotou, G., Brown, T., Barlow, T., Pearl, L. H., and Savva, R. (1998) *J. Biol. Chem.* **273,** 45–50
6. Shida, T., Noda, M., and Sekiguchi, J. (1996) *Nucleic Acids Res.* **24,** 4572–4576
7. Shida, T., Ogawa, T., Ogasawara, N., and Sekiguchi, J. (1999) *Biosci. Biotechnol. Biochem.* **63,** 1528–1534
8. Marenstein, D. R., Wilson, D. M., III, and Teebor, G. W. (2004) *DNA Repair* **3,** 527–533
9. Dianov, G. L., Sleeth, K. M., Dianova, I. I., and Allinson, S. L. (2003) *Mutat. Res.* **531,** 157–163
10. Takahashi, I., and Marmur, J. (1963) *Nature* **197,** 794–795
11. Friedberg, E. C., Ganesan, A. K., and Minton, K. (1975) *J. Virol.* **16,** 315–321
12. Cone, R., Bonura, T., and Friedberg, E. C. (1980) *J. Biol. Chem.* **255,** 10354–10358
13. Wang, Z., and Mosbaugh, D. W. (1989) *J. Biol. Chem.* **264,** 1163–1171


[3] G. Serrano-Heras, M. Salas, and A. Bravo, unpublished results.


14. Karran, P., Cone, R., and Friedberg, E. C. (1981) *Biochemistry* **20,** 6092–6096
15. Mol, C. D., Arvai, A. S., Sanderson, R. J., Slupphaug, G., Kavli, B., Krokan, H. E., Mosbaugh, D. W., and Tainer, J. A. (1995) *Cell* **82,** 701–708
16. Savva, R., and Pearl, L. H. (1995) *Nat. Struct. Biol.* **2,** 752–757
17. Putnam, C. D., Shroyer, M. J., Lundquist, A. J., Mol, C. D., Arvai, A. S., Mosbaugh, D. W., and Tainer, J. A. (1999) *J. Mol. Biol.* **287,** 331–346
18. Salas, M. (1999) *Genet. Eng.* **21,** 159–171
19. Inciarte, M. R., Salas, M., and Sogo, J. M. (1980) *J. Virol.* **34,** 187–199
20. Sogo, J. M., García, J. A., Peñalva, M. A., and Salas, M. (1982) *Virology* **116,** 1–18
21. Blanco, L., and Salas, M. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 5325–5329
22. Blanco, L., Bernad, A., Lázaro, J. M., Martín, G., Garmendia, C., and Salas, M. (1989) *J. Biol. Chem.* **264,** 8935–8940
23. Moreno, F., Camacho, A., Viñuela, E., and Salas, M. (1974) *Virology* **62,** 1–16
24. Mellado, R. P., Viñuela, E., and Salas, M. (1976) *Eur. J. Biochem.* **65,** 213–223
25. Rojo, F., Mencía, M., Monsalve, M., and Salas, M. (1998) *Prog. Nucleic Acids Res. Mol. Biol.* **60,** 29–46
26. Jiménez, F., Camacho, A., de la Torre, J., Viñuela, E., and Salas, M. (1977) *Eur. J. Biochem.* **73,** 57–72
27. Steiner, M., Lubitz, W., and Blasi, U. (1993) *J. Bacteriol.* **175,** 1038–1042
28. Bravo, A., Alonso, J. C., and Trautner, T. A. (1990) *Nucleic Acids Res.* **18,** 2881–2886
29. Bravo, A., and Salas, M. (1997) *J. Mol. Biol.* **269,** 102–112
30. Yasbin, R. E., Fields, P. I., and Andersen, B. J. (1980) *Gene* (*Amst.*) **12,** 155–159
31. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) *Anal. Chem.* **68,** 850–858
32. Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999) *Electrophoresis* **20,** 3551–3567
33. Bravo, A., Serrano-Heras, G., and Salas, M. (2005) *FEMS Microbiol. Rev.* **29,** 25–47
34. Sogo, J. M., Lozano, M., and Salas, M. (1984) *Nucleic Acids Res.* **12,** 1943–1960
35. Barthelemy, I., Salas, M., and Mellado, R. P. (1986) *J. Virol.* **60,** 874–879
36. Bravo, A., Illana, B., and Salas, M. (2000) *EMBO J.* **19,** 5575–5584
37. Schagger, H., and von Jagow, G. (1987) *Anal. Biochem.* **166,** 368–379
38. Serrano-Heras, G., Salas, M., and Bravo, A. (2003) *J. Biol. Chem.* **278,** 40771–40777
39. Gao, H., Jiang, X., Pogliano, K., and Aronson, A. I. (2002) *J. Bacteriol.* **184,** 2780–2788
40. Glaser, P., Kunst, F., Arnaud, M., Coudart, M. P., Gonzales, W., Hullo, M. F., Ionescu, M., Lubochinsky, B., Marcelino, L., Moszer, I., Presecan, E., Santana, M., Schneider, E., Schweizer, J., Vertes, A., Rapoport, G., and Danchin, A. (1993) *Mol. Microbiol.* **10,** 371–384
41. Neveling, U., Bringe-Meyer, S., and Sahm, H. (1998) *Biochim. Biophys. Acta* **1385,** 367–372
42. Walter, T., and Aronson, A. (1999) *J. Biol. Chem.* **274,** 7901–7906
43. Stein, A., and Firshein, W. (2000) *J. Bacteriol.* **182,** 2119–2124
44. Parsons, L. M., Liu, F., and Orban, J. (2005) *Protein Sci.* **14,** 1684–1687
45. Warner, H. R., Johnson, L. K., and Snustad, D. P. (1980) *J. Virol.* **33,** 535–538
46. Wang, J., Jiang, Y., Vincent, M., Sun, Y., Yu, H., Wang, J., Bao, Q., Kong, H., and Hu, S. (2005) *Virology* **332,** 45–65
47. Gallinari, P., and Jiricny, J. (1996) *Nature* **383,** 735–738
48. Meijer, W. J., Horcajadas, J. A., and Salas, M. (2001) *Microbiol. Mol. Biol. Rev.* **65,** 261–287

# A Uracil-DNA Glycosylase Inhibitor Encoded by a Non-uracil Containing Viral DNA

Gemma Serrano-Heras, Margarita Salas and Alicia Bravo

*J. Biol. Chem. 2006, 281:7068-7074.*
*doi: 10.1074/jbc.M511152200 originally published online January 18, 2006*

---