Domain Structure of gpNu1, a Phage Lambda DNA Packaging Protein[†]

Qin Yang,[‡] Nancy Berton,[‡] Mark C. Manning,[‡] and Carlos Enrique Catalano*[,‡,§]

Department of Pharmaceutical Sciences and Molecular Biology Program, University of Colorado Health Sciences Center, Denver, Colorado 80262

Received June 18, 1999; Revised Manuscript Received September 3, 1999

ABSTRACT: The terminase enzyme from bacteriophage lambda is responsible for the insertion of a dsDNA genome into the confines of the viral capsid. The holoenzyme is composed of gpA and gpNu1 subunits in a $gpA_1 \cdot gpNu1_2$ stoichiometry. While genetic studies have described regions within the two proteins responsible for DNA binding, capsid binding, and subunit interactions in the holoenzyme complex, biochemical characterization of these domains is limited. We have previously described the cloning, expression, and biochemical characterization of a soluble DNA binding domain of the terminase gpNu1 subunit (Met1 to Lys100) and suggested that the hydrophobic region spanning Lys100 to Pro141 defines a domain responsible for self-association interactions, and that is important for cooperative DNA binding [Yang et al. (1999) *Biochemistry 38*, 465−477]. We further suggested that the genetically defined gpA-interactive domain in the C-terminal half of the protein is limited to the C-terminal ≈40 amino acids of gpNu1. Here we describe the cloning, expression, and biochemical characterization of gpNu1ΔP141, a deletion mutant of gpNu1 that comprises the DNA binding domain and the putative hydrophobic self-assembly domain of the full-length protein. Purified gpNu1ΔP141 shows a strong tendency to aggregate in solution; However, the protein remains soluble in 0.4 M guanidine hydrochloride, and circular dichroism (CD) and fluorescence spectroscopic studies demonstrate that the protein is folded under these conditions. Moreover, CD spectroscopy and thermally induced unfolding studies suggest that the DNA binding domain and the self-association domain represent independent folding domains of gpNu1ΔP141. The mutant protein interacts weakly with the gpA subunit, but does not form a catalytically competent holoenzyme complex, suggesting that the C-terminal 40 residues are important for appropriate subunit interactions. Importantly, gpNu1ΔP141 binds DNA tightly, but with less specificity than does full-length protein, and the data suggest that the C-terminal residues are further required for specific DNA binding activity. The implications of these results in the assembly of a functional holoenzyme complex are discussed.

Bacteriophage λ is a double-stranded DNA virus that infects *E. coli* (*1*, *2*). The virus is composed of a 48.5 kb[1] linear genome with complementary 12 base single-stranded "sticky" ends tightly packaged within the viral capsid. Upon injection into the cell, the sticky ends anneal, and the nicks are sealed by host ligase forming a circular DNA duplex (*3*, *4*). During the latter stages of infection, DNA replication occurs via a rolling-circle mechanism that yields linear concatemers of the genome linked in a head-to-tail fashion (*3*). One of the final steps in virus development is the excision of individual genomes from the concatemer, and their packaging into an empty procapsid (*4*−*6*).

Terminase enzymes are common to the double-stranded DNA bacteriophage and form, at least in part, a "machine" responsible for the DNA packaging process (*5*, *7*, *8*). The terminase enzyme from bacteriophage λ is composed of two virally encoded proteins, gpA (641 amino acids, 73.3 kDa) and gpNu1 (181 amino acids, 20.4 kDa), isolated as a $gpA_1 \cdot gpNu1_2$ holoenzyme complex (*9*, *10*). A model for the assembly of a catalytically competent packaging complex has been proposed (*5*, *6*, *11*) and is shown in Figure 1. In this model, gpNu1 dimers bind in a cooperative manner to three repeating recognition elements (R1−R3) found within the *cosB* region of the viral genome concatemer. Assembly of a gpNu1 nucleoprotein complex at *cosB* stimulates the assembly of a gpA dimer at *cosN*, and is responsible for the unusual stability of the resulting prenicking complex. The endonuclease activity of the gpA subunit introduces symmetrical nicks into the duplex, a reaction whose activity and fidelity are modulated by ATP. Terminase-mediated strand separation of the nicked, annealed duplex yields complex I, which next binds an empty procapsid to initiate terminase translocation along the concatemer and active DNA packaging. Both the strand-separation and translocase activities of

[†] This work was supported by National Science Foundation Grant MCB-9728550.

* Address correspondence to this author. Phone: (303) 315-8561; Fax: (303) 315-6281; Internet: carlos.catalano@uchsc.edu.

[‡] Department of Pharmaceutical Sciences, School of Pharmacy.

[§] Molecular Biology Program, School of Medicine.

[1] Abbreviations: β-ME, 2-mercaptoethanol; *cos*, cohesive end site, the junction between individual genomes in immature concatemeric λ DNA; GDN, guanidine hydrochloride; gpA, the large subunit of phage λ terminase; gpA-H6, a gpA subunit containing a hexa-histidine purification tag at the C-terminus of the protein; gpNu1-FL, the full-length small subunit of phage λ terminase; gpNu1ΔK100, a deletion mutant of gpNu1 (Met1−Lys100) lacking the C-terminal 81 amino acids; gpNu1ΔP141, a deletion mutant of gpNu1 (Met1−Pro141) lacking the C-terminal 40 amino acids; IHF, *E. coli* integration host factor; kb, kilobase(s); kDa, kilodalton(s); PAGE, polyacrylamide gel electrophoresis.

10.1021/bi991408f CCC: $18.00
Published on Web 10/05/1999

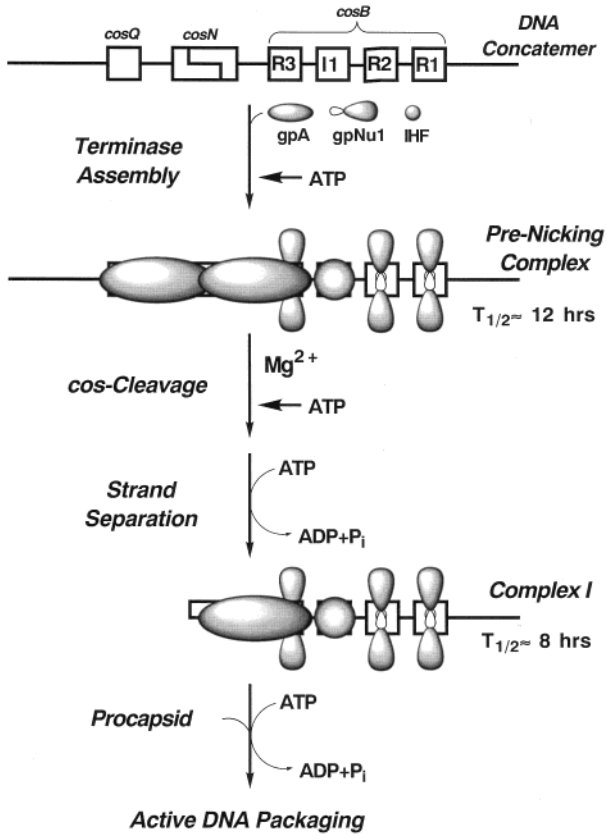

FIGURE 1: Model for terminase assembly at *cos*. The *cos* region of the λ genome is shown at the top. The three subsites, *cosQ*, *cosN*, and *cosB*, are indicated, as are the three gpNu1 binding elements (R1−R3) and the IHF binding element found within *cosB*. The terminase subunits and IHF assemble at *cos*, forming a pre-nicking complex which, in the presence of Mg^{2+}, nicks the duplex, ultimately yielding the stable packaging intermediate complex I. Complex I binds an empty pro-capsid which initiates an ATP-dependent insertion of viral DNA into the capsid. We note that the stoichiometry of the terminase subunits assembled at *cos* remains speculative. Details are presented in the text.

the enzyme are presumed to be driven by ATP hydrolysis. Upon encountering the next downstream *cos* site in the concatemer (the end of the genome), terminase again nicks the duplex at *cosN* and separates the strands, simultaneously releasing a DNA-filled capsid and regenerating complex I.

Phage λ terminase is thus an integral part of a number of nucleoprotein complexes that are required for packaging of the viral genome. Our laboratory is interested in the biochemical and biophysical properties of the terminase subunits, the protein·protein and protein·DNA interactions responsible for the assembly of these nucleoprotein complexes, and the molecular basis for the transitions between them. With respect to the enzyme subunits, genetic studies by Feiss and co-workers have defined functional domains within the primary sequence of the proteins. Specifically, gpA·gpNu1 subunit interactions have been localized to the N-terminal ≈40 amino acids of gpA and the C-terminal half (≈90 amino acids) of gpNu1 (see Figure 2) (*12*, *13*).

We have previously described the cloning, expression, and biochemical characterization of a deletion mutant of the terminase gpNu1 subunit consisting of residues Met1 to Lys100 (gpNu1ΔK100) (*14*). The mutant protein forms a folded dimer in solution but, unlike full-length gpNu1, shows no tendency toward aggregation. While gpNu1ΔK100 binds with modest specificity to *cos*-containing DNA, physical and biochemical assays demonstrate the protein does not interact with the terminase gpA subunit (*14*). Based on this work, we suggested that the hydrophobic amino acids found between Lys100 and Pro141 define a self-association domain of the protein that is responsible for the observed aggregation behavior of the full-length protein (see Figure 2). We further speculated that the genetically defined gpA-interactive domain in the C-terminal half of gpNu1 is, in fact, limited to the C-terminal ≈40 amino acids of the protein. We have thus constructed gpNu1ΔP141, a mutant protein with these C-terminal residues deleted, to test this model and to more clearly define the protein·protein interactions required for holoenzyme complex formation and cooperative DNA binding. The results of these studies and their implications with respect to the formation of a catalytically competent DNA packaging machine are discussed.

EXPERIMENTAL PROCEDURES

Materials and Methods. Tryptone, yeast extract, and agar were purchased from DIFCO. Restriction enzymes were purchased from Promega. Guanidine hydrochloride was purchased from Mallinckrodt. Mono-Q HR5/5 FPLC columns and DEAE-Sepharose FF chromatography resins were purchased from Pharmacia. Radiolabeled nucleotides were purchased from ICN. Unlabeled nucleoside triphosphates were purchased from Boehringer Mannheim Biochemicals. All other materials were of the highest quality commercially available.

Bacterial cultures were grown in shaker flasks utilizing a New Brunswick Scientific series 25 incubator−shaker. All protein purifications utilized a Pharmacia FPLC system which consisted of two P500 pumps, a GP250-plus controller, a V7 injector, and a Uvicord SII variable-wavelength detector. UV−Vis absorbance spectra were recorded on a Hewlett-Packard HP8452A spectrophotometer. Fluorescence spectra were recorded at room temperature on a PTI QuantaMaster spectrofluorometer. Circular dichroism (CD) spectra were recorded on an Aviv model 62DS circular dichroism spectropolarimeter equipped with a Brinkmann Lauda RM6 circulating water bath and a thermostated cell holder. Unless otherwise indicated, all spectra were recorded in 10 mM sodium phosphate buffer, pH 7.2, containing 0.4 M guanidine hydrochloride (GDN).

Near-UV CD spectra utilized a protein concentration of 1 mg/mL in a 1 cm strain-free cuvette. Data were collected between 250 and 350 nm at 1 nm intervals using a bandwidth of 1.5 nm and a dwell time of 60 s. Far-UV CD spectra utilized a protein concentration of 100 μg/mL in a 0.1 cm strain-free cuvette. Data were collected from 202 to 260 nm at 0.5 nm intervals using a bandwidth of 1.5 nm and a dwell time of 30 s. Thermally induced protein denaturation experiments were performed as described previously (*15*), and the fraction of protein in the denatured state (F_D) was determined using the equation:

$$F_D = \frac{\phi_N - \phi_T}{\phi_N - \phi_D} \quad (1)$$

where ϕ_T is the ellipticity at temperature T, and ϕ_N and ϕ_D represent the ellipticity for the native and denatured protein,

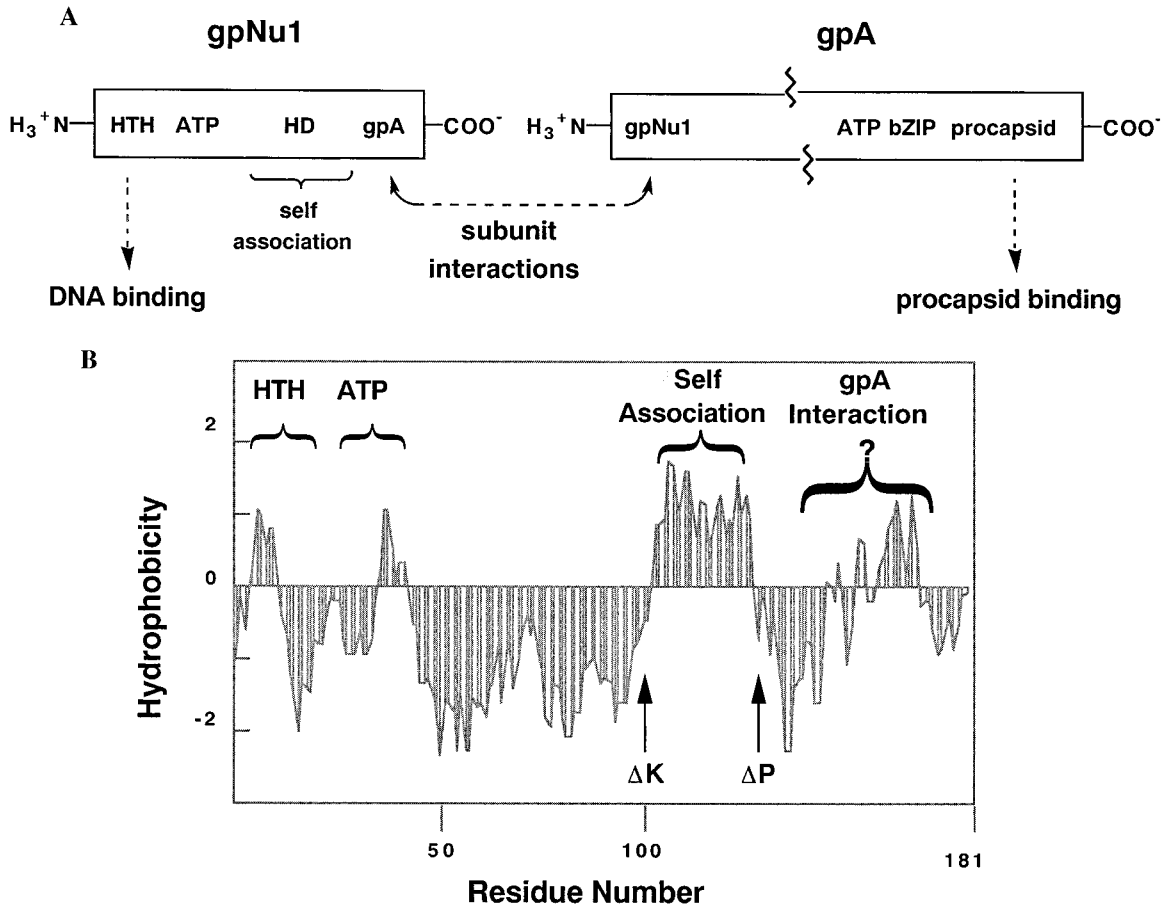

FIGURE 2: Panel A: Proposed domain structure of the terminase subunits. The locations of the putative helix−turn−helix DNA binding motif (HTH, Lys5− Glu24), the Walker A sequence (ATP, Val29−Gly36), and the putative hydrophobic self-association domain (HD, Lys100−Pro141) are indicated in gpNu1. The locations of the Walker A sequence (ATP, Gly491−Pro498) and the putative bZIP protein dimerization sequence (Glu560−Gly621) are indicated in gpA. Based on genetic studies, subunit interactions are mediated by the C-terminal ≈90 amino acids of gpNu1 and the N-terminal ≈40 amino acids of gpA. See text for details. Panel B: Kyte−Doolitle hydropathy plot of full-length gpNu1. The locations of the helix−turn−helix DNA binding motif (HTH) and the Walker A (ATP) sequences are indicated. The putative hydrophobic self-association and gpA-interactive domains are also indicated. The locations of the C-terminal ends in the gpNu1ΔK100 (ΔK) and gpNu1ΔP141 (ΔP) deletion mutant proteins are indicated with arrows.

respectively (*16*, *17*). The unfolding curves were analyzed using a double-sigmoidal curve function as described previously (*18*):

$$F_D = \left(\left(\frac{F_{D1} - F_{N1}}{1 + \left(\frac{T}{T_{m1}}\right)^{m_{T1}}}\right) + F_{N1}\right) + \left(\left(\frac{F_{D2} - F_{N2}}{1 + \left(\frac{T}{T_{m2}}\right)^{m_{T2}}}\right) + F_{N2}\right) \quad (2)$$

where F_{N1}/F_{N2} and F_{D1}/F_{D2} represent the pre-transition and post-transition base lines for the individual transitions, respectively, and m_{T1}/m_{T2} and T_{m1}/T_{m2} represent the slopes and melting temperature of each transition, respectively. All data sets were fit to the above equations by nonlinear regression methods using the IGOR graphics/analysis package (WaveMetrics, Lake Oswego, OR).

Electrospray mass spectral analysis was performed on a Finnigan LCQ ion trap mass spectrometer using a mobile phase consisting of 32.5% acetonitrile and 17.5% methanol in water containing 0.01% trifluoroacetic acid at 50 μL/min. Prediction of protein secondary structures and hydrophobicity based upon primary sequence data was performed by the methods of Chou and Fasman and Kyte and Doolittle, respectively, using the DNASIS program (Macintosh version 2.0). Automated DNA sequence analysis was performed by the University of Colorado Cancer Center Macromolecular Resources Core facility. Both strands of the duplex were examined to ensure the expected DNA sequence.

Bacterial Strains, DNA Preparation, and Protein Purification. E. coli BL21(DE3) cells were a generous gift of D. Kroll (University of Colorado Health Sciences Center, Denver, CO). All synthetic oligonucleotides used in this study were purchased from Gibco/BRL and were used without further purification. Plasmids pSF1 and pAFP1, kindly provided by M. Feiss (University of Iowa, Iowa City, IA), were purified from the *E. coli* cell lines C600[pSF1] and JM107[pAFP1], respectively, using Qiagen DNA prep columns. Purification of gpA and full-length gpNu1 was performed as previously described (*15*). Purification of gpA-H6 was performed as described by Hang et al. (*19*). All of our purified proteins were homogeneous as determined by SDS−PAGE and densitometric analysis using a Molecular Dynamics laser densitometer and the ImageQuant data analysis package. Unless otherwise indicated, protein concentrations were determined spectrally using millimolar extinction coefficients (*10*, *15*).

Construction of the gpNu1ΔP141 Deletion Mutant Overexpression Plasmid. A truncated *Nu1* gene was amplified

by PCR using pSF1 as a DNA template. This plasmid contains the wild-type *Nu1* gene cloned into a pBR322 background (*20*). Primers were designed such that *Eco*RI and *Bam*HI restriction sequences were present at the 5′ and the 3′ end, respectively, of the PCR product. The primer sequences were as follows: forward primer: $^{5'}$-CCT CTC CCT TTC TCC *GAA TTC* **ATG** GAA GTC AAC AAA AAG C-$^{3'}$; reverse primer: $^{5'}$-CTT CCT *GGA TCC* **TTA** CGG AAA ACG CCG CTG CAC-$^{3'}$. The *Eco*RI and *Bam*HI restriction sequences are indicated in italics while the f-MET (forward primer) and stop (reverse primer) codons are shown in boldface type. Sequences complementary to the *Nu1* gene are underlined. The stop codon present in the reverse PCR primer yields, upon amplification, a truncated gpNu1 gene which expresses only the first 141 amino acids of the protein. Amplification of the truncated gene, isolation of the PCR product, and construction of the overexpression plasmid (pNu1ΔP) were performed as described previously for wild-type gpNu1 (*15*). Colonies from BL21(DE3) cells transfected with this plasmid efficiently expressed the C-terminal deleted gpNu1 mutant protein, gpNu1ΔP141, as determined from whole cell lysates analyzed by SDS−PAGE and Western blotting (not shown).

Expression and Purification of gpNu1ΔP141. One liter of 2×-YT media containing 50 μg/mL ampicillin, 25 mM potassium phosphate, pH 7.5, and 5 mM glucose was inoculated with a 10 mL overnight culture of BL21(DE3)-[pNu1ΔP] derived from an isolated colony. The culture was maintained at 37 °C until an OD of 0.45 (600 nm) was obtained, at which point IPTG (1.2 mM) was added. The cells were maintained at 37 °C for an additional 2 h and then harvested by centrifugation. Unless otherwise indicated, all subsequent steps were performed at 0−4 °C. Purification of gpNu1ΔP141 from inclusion bodies was performed essentially as described for the full-length gpNu1 subunit (*15*) with minor modification. Briefly, gpNu1ΔP141 was solubilized from the cell lysis pellet with 100 mL of buffer A (50 mM Tris·HCl, pH 8.0, 5 mM EDTA) containing 6 M guanidine hydrochloride (GDN). The sample was then diluted to 4 L with buffer B (20 mM Tris·HCl, pH 8.0, 1 mM EDTA, 7 mM β-ME) containing 100 mM NaCl and 2 M GDN, and the protein was refolded by dialysis against 40 L of buffer B for 20 h with one buffer change. The protein was purified by DEAE-Sepharose and MonoQ HR5/5 chromatography as previously described (*15*) and stored at 4 °C. As necessary, the protein was concentrated using Centricon10 centrifugal concentrators after dialysis into 10 mM sodium phosphate buffer, pH 7.2, containing at least 0.5 M GDN.

Interaction of the Terminase Subunits. C-Terminal hexa-HIS-containing gpA (gpA-H6, 2 μM) was preincubated with either full-length gpNu1 (4 μM) or the truncated gpNu1ΔP141 mutant (4 μM) in 50 mM sodium phosphate, pH 8.0, 100 mM NaCl buffer for 20 min on ice. Ni−NTA agarose resin (0.5 mL in the same buffer) was added to the protein mixture, and the incubation was continued for an additional 60 min with mild shaking. The resin was then batch-loaded into a column and washed twice with 300 μL of wash buffer (50 mM sodium phosphate, pH 8.0, 20 mM imidazole, and 500 mM NaCl), and gpA-H6 was finally eluted from the column with elution buffer (50 mM sodium phosphate, pH 8.0, 250 mM imidazole, and 500 mM NaCl). Fractions (200 μL) were collected and analyzed by SDS−PAGE.

Activity Assays. Gel mobility shift assays were performed and analyzed as described by Yang et al. (*14*). The concentration of protein and salt added to the binding reactions is indicated in each individual experiment. The *cos-cleavage assay* was performed as described previously using pAFP1 as a nuclease substrate (*10*, *21*). ATPase catalytic activity was examined as described previously (*22*, *23*) and utilized an [α-^{32}P]ATP concentration of 50 μM and a DNA (*Sca*I-linearized pAFP1) concentration of 300 μM (total nucleotide). The concentration of protein used in these assays is indicated in each individual experiment.

RESULTS

Construction, Expression, and Purification of gpNu1ΔP141. Based on the domain model presented in Figure 2, we constructed gpNu1ΔP141, a mutant of gpNu1 in which the C-terminal 40 amino acids of the protein have been deleted. Of note is that limited proteolysis studies have demonstrated a relatively stable N-terminal digestion intermediate of gpNu1 with a molecular weight expected from a similar deletion of the C-terminus of the protein (C. E. Catalano and A. Hanagan, unpublished). Similar to wild-type, full-length gpNu1 (gpNu1-FL), virtually all of the expressed gpNu1ΔP141 protein was present in the insoluble pellet fraction of the crude cell lysate. Nevertheless, the protein could be purified to homogeneity as described under Experimental Procedures. SDS−PAGE analysis yielded an apparent molecular mass of ≈15.2 kDa, consistent with the M_r = 15 917 Da predicted from the gene sequence. Moreover, electrospray mass spectral analysis of the purified protein yielded a molecular mass of 15 919 Da. We note that purification of gpNu1ΔP141 was extremely difficult due to the tendency of the protein to aggregate at each of step of the purification protocol, especially during the dialysis steps. Once purified, however, the protein could be concentrated (2.4 mg/mL) in the presence of 0.5 M guanidine hydrochloride (GDN) with no evidence of aggregation during prolonged storage at 4 °C. The protocol described here yields 2 mg of highly purified protein per liter of cell growth.

Preliminary Characterization of gpNu1ΔP141. The UV spectrum of gpNu1ΔP141 shown in Figure 3A (inset) is typical of a globular protein that is essentially devoid of contaminating DNA (*24*, *25*), and an extinction coefficient (ϵ_{280}) of 15.3 $mM^{-1}\cdot cm^{-1}$ (in 10 mM sodium phosphate, pH 7.2, 0.5 M GDN) was determined by the method of Gill and von Hipple (*26*, *27*). The figure further demonstrates that while gpNu1ΔP141 remains soluble in the presence of 0.4 M GDN, significant aggregation is observed with further dilution of GDN from the sample as evidenced by increased light scattering in the spectrum (increased absorbance at 320 nm).

The fluorescence spectrum of the protein taken in 0.4 M GDN (Figure 3B) displays an emission maximum of 332 nm that remains unchanged using excitation frequencies between 260 and 285 nm. Excitation at longer wavelengths results in significant deterioration of the spectrum (not shown). Importantly, the fluorescence spectrum of the protein taken in 6 M GDN displays a significant red-shift in the emission maximum (349 nm) and a 2-fold increase in the fluorescence intensity. These data suggest that gpNu1ΔP141 retains a folded conformation in the presence of 0.4 M GDN but, as expected, denatures at elevated concentrations of GDN.

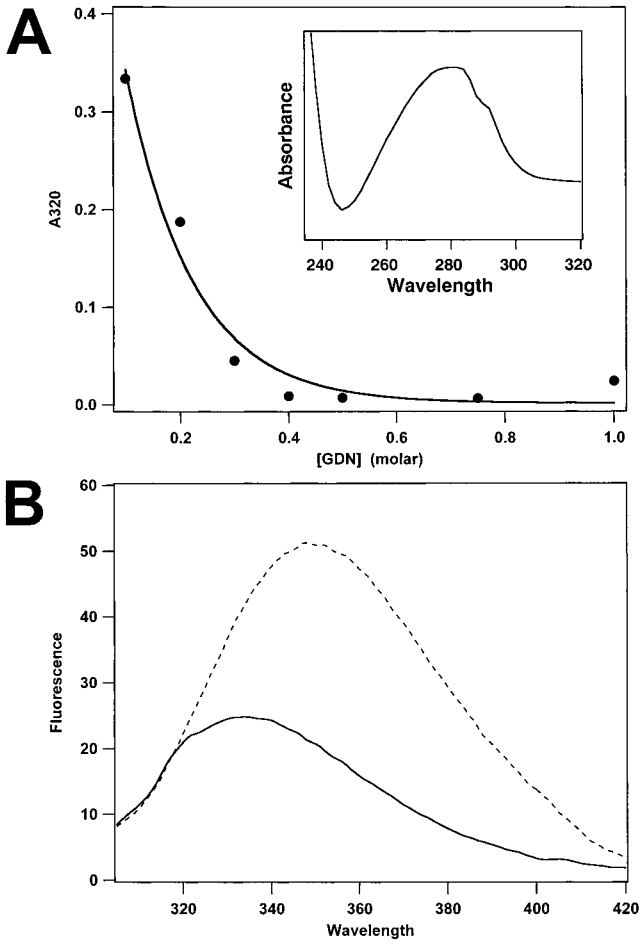

FIGURE 3: Panel A: Aggregation of gpNu1ΔP141 as a function of guanidine hydrochloride concentration. Light scattering due to protein aggregation was monitored at 320 nm. Inset: UV spectrum of the protein in 0.5 M GDN. Panel B: Fluorescence spectrum of gpNu1ΔP141 in the presence of 0.4 M GDN (solid line) and 6 M GDN (dashed line). All spectra were recorded as described under Experimental Procedures.

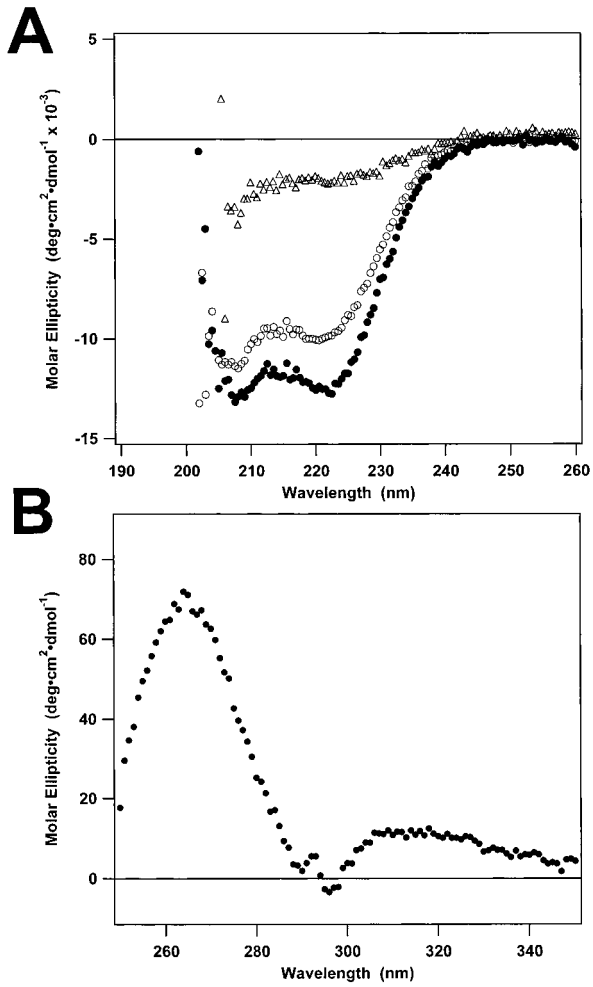

FIGURE 4: Panel A: Far-UV CD spectra of gpNu1ΔP141 recorded at 4 °C before (●) and after (○) heating to 95 °C for 15 min, and the spectrum recorded at 95 °C (△). Panel B: Near-UV CD spectrum recorded at 4 °C in 0.4 M guanidine hydrochloride.

Chou and Fasman secondary structure analysis predicts a significant amount (38%) of α-helical structure for the protein, and, consistently, the far-UV circular dichroism (CD) spectrum of gpNu1ΔP141 shows strong negative maxima at 222 and 208 nm (Figure 4A). These data suggest that the mutant protein indeed possesses a significant amount of α-helical structure (*25*, *28*). Unfortunately, however, deconvolution analysis of the spectrum was not possible due to the strong absorbance at low wavelengths by the GDN in the sample. The near-UV CD spectrum of gpNu1ΔP141 shown in Figure 4B displays strong signals consistent with the existence of significant tertiary structure in the presence of 0.4 M GDN (Figure 4B). Importantly, the negative bands observed at 296 and 289 nm, representing the 0→0 and 0→1 vibronic components, respectively, of the ^1L_b band of tryptophan (*29*), are unperturbed from their positions in the gpNu1ΔK100 mutant protein (*15*). This suggests that the folded conformation of residues 1−100 is not significantly altered in the larger gpNu1ΔP141 mutant protein. Conversely, the positive band at 264 nm is significantly red-shifted relative to the band seen in gpNu1ΔK100 (252 nm) (*15*). Most likely, this signal arises from the ^1L_a transition from one of the tryptophan residues in the protein, and the red-shift in gpNu1ΔP141 could be due to stronger coupling to other aromatic groups, or slight changes in the dihedral angles of the tryptophan side chain relative to the peptide backbone (*29*). Of note is the broad positive band at 318 nm present in the spectrum of gpNu1ΔP141 (Figure 4B) but absent from that of gpNu1ΔK100 (*15*). This band indicates the presence of a disulfide bond in the protein that is conformationally strained, deviating from the normal ±90° dihedral angle (*30*). Importantly, gpNu1ΔP141 possesses a single cysteine residue (C114) that is absent from the gpNu1ΔK100 mutant protein. Consistently, while gpNu1ΔK100 migrates as a monomer in denaturing, nonreducing PAGE, gpNu1ΔP141 migrates as a dimer under these conditions but collapses to a monomer in the presence of β-ME (Q. Yang and C. E. Catalano, unpublished).

Stability of gpNu1ΔP141. Unlike full-length gpNu1 (*15*) and gpNu1ΔK100 (Q. Yang and C. E. Catalano, unpublished), thermal denaturation of gpNu1ΔP141 is only partially (≈80%) reversible (see Figure 4A). We have previously demonstrated that thermally induced denaturation of full-length gpNu1 shows a single, broad, and uncooperative unfolding transition when monitored by far-UV CD spectroscopy (*15*). Moreover, the melting temperature (T_m = 307 K) is significantly lower than that observed for either the isolated gpA subunit (T_m = 317 K) or the holoenzyme complex (T_m = 317 K), both of which also exhibit single,

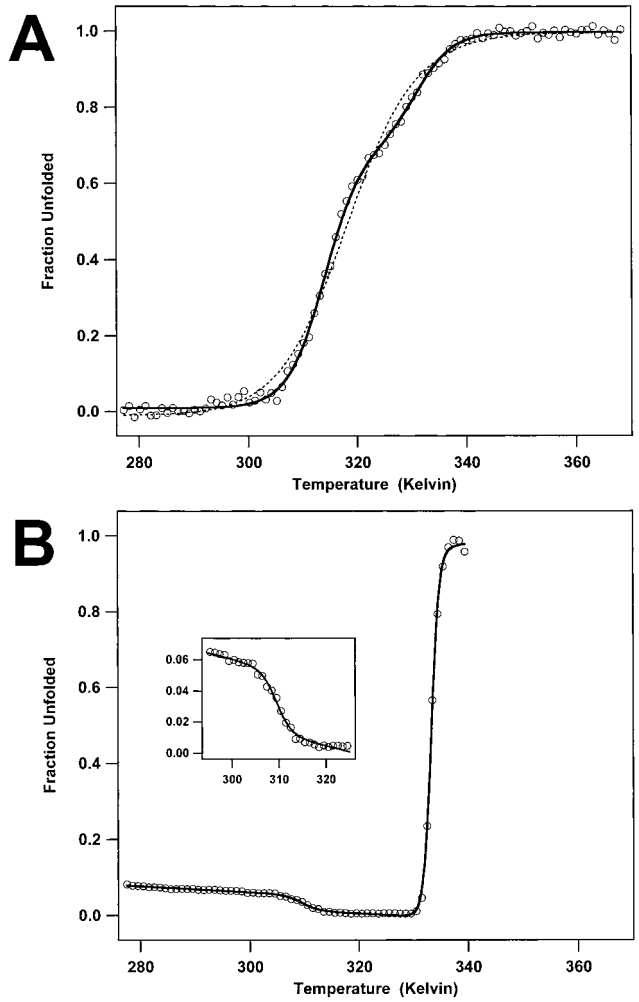

FIGURE 5: Panel A: Thermally-induced unfolding of gpNu1ΔP141 monitored at 222 nm (far-UV). The data were analyzed using a single sigmoidal curve function (dashed line) and a double sigmoidal curve function (solid line) as described under Experimental Procedures. Panel B: Thermally-induced unfolding of gpNu1ΔP141 monitored at 260 nm (near-UV). The data were analyzed using a double sigmoidal curve function as described under Experimental Procedures.

Table 1: Thermally-Induced Unfolding of gpNu1ΔP141[a]

signal monitored	T_{m1} (K)	T_{m2} (K)
far-UV CD (222 nm)	314.1 ± 0.3	331.3 ± 0.5
near-UV CD (260 nm)	309.5 ± 0.9	333.3 ± 0.1

[a] The data presented in Figure 5 were analyzed as described under Experimental Procedures.

though more cooperative, unfolding transitions (*15*). Interestingly, however, thermally induced unfolding of gpNu1ΔP141 reveals two discrete transitions that are observed in both the far-UV (Figure 5A) and the near-UV (Figure 5B) regions (Table 1). These data are consistent with the posit that gpNu1ΔP141 is composed of two domains, presumably represented by residues 1−100 and 100−140, respectively, that unfold independently (see Table 1). We note that the single uncooperative unfolding transition observed with the full-length protein is consistent with a molten globule or compact intermediate state (*15*). In this case, the putative domains are not fully folded, and discrete unfolding transitions are not expected.

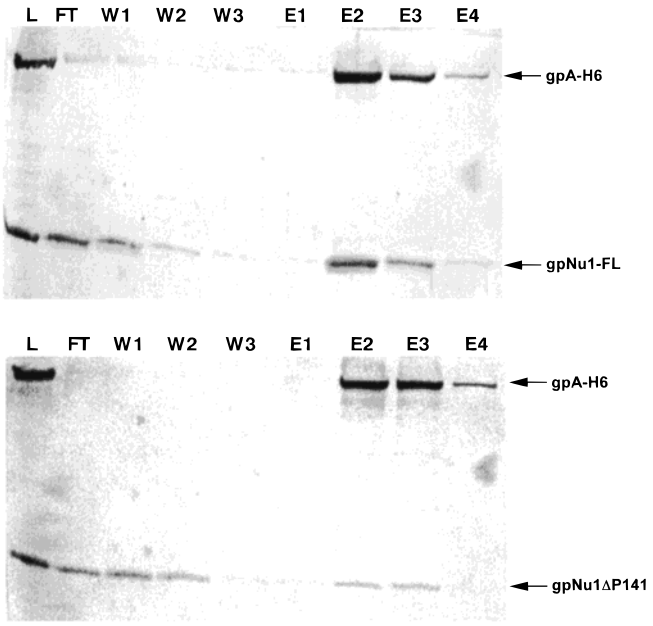

FIGURE 6: gpNu1ΔP141 interacts weakly with the terminase gpA subunit. gpA-H6 was preincubated with either full-length gpNu1 (panel A) or gpNu1ΔP141 (panel B) as described under Experimental Procedures and loaded onto a nickel-chelate column. The column was washed and protein eluted as described. L, load fraction; FT, column flow through; W1−W2, column wash fractions; E1−E4, column elution fractions. The positions of gpA-H6 and the relevant gpNu1 protein bands are indicated with arrows.

gpNu1ΔP141 Interacts Weakly with the Terminase gpA Subunit. Reconstitution of terminase holoenzyme from a C-terminal hexaHIS-tagged gpA subunit [gpA-H6 (*19*)] and either wild-type gpNu1 or gpNu1ΔP141 is shown in Figure 6. The mixtures were batch-loaded onto nickel chelate columns which retained the gpA-H6 subunit. After the columns were washed to remove any unbound protein, gpA-H6 was eluted, and all of the fractions were analyzed by SDS−PAGE. Contrary to our prediction, the figure clearly demonstrates that gpNu1ΔP141 binds to the gpA subunit. This interaction is weaker than that observed with the full-length protein, however, and analysis of the data suggests that the apparent affinity of gpNu1ΔP141 for gpA is roughly 50% that of full-length gpNu1 under these conditions.

gpNu1ΔP141 Binds Tightly to DNA. Figure 7 demonstrates that gpNu1ΔP141 binds to *cos*-containing DNA and, surprisingly, binds more tightly than does full-length gpNu1 (see Table 2). Analysis of the data reveals that the mutant protein displays only modest discrimination between *cos*-containing and nonspecific DNA sequences, however (Figure 7A, Table 2). To more fully examine specificity in DNA binding by gpNu1ΔP141, two additional experiments were performed. (i) While it has been shown that an increase in ionic strength strongly affects all nucleoprotein complexes, nonspecific interactions are more severely affected than specific protein·DNA interactions (*31*). We thus examined the effect of NaCl on the stability of the gel-retarded gpNu1ΔP141·DNA complexes. Figure 7B shows that while salt disrupts both *cos*-containing and nonspecific protein·DNA complexes, significantly lower concentrations of salt are required to disrupt the nonspecific complexes ($C_{1/2} \approx 300$ mM vs ≈100 mM). Moreover, the curves are much steeper with nonspecific DNA complexes (slope = −0.627 ± 0.07) than with complexes formed with *cos*-containing DNA substrates

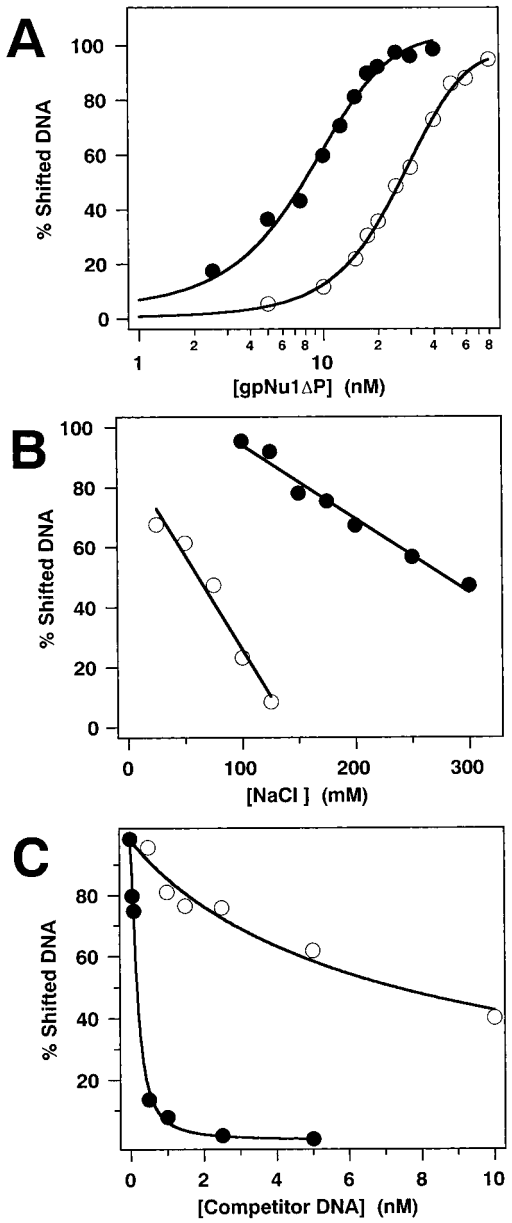

FIGURE 7: gpNu1ΔP141 binding to DNA. Panel A: Binding of *cos*-containing and nonspecific DNA by gpNu1ΔP141. The gel retardation experiments were conducted as described and included 50 mM NaCl and 20 pM *cos*-containing (●) or nonspecific (○) DNA. Panel B: Salt disrupts gpNu1ΔP141·DNA complexes. The gel retardation experiments were conducted as described under Experimental Procedures using 30 nM gpNu1ΔP141 and 20 pM *cos*-containing (●) or nonspecific (○) DNA substrates. NaCl was added to the binding buffer as indicated in the figure. Panel C: Competitive discrimination between specific and nonspecific DNA. The experiments were conducted as described under Experimental Procedures and included 30 nM gpNu1ΔP141, 20 pM radiolabeled *cos*-containing DNA, and 50 mM NaCl. Cold specific (●) or nonspecific (○) competitor DNA was added as indicated in the figure.

(slope = -0.247 ± 0.02). (ii) The experiment presented in Figure 7A suggests that gpNu1ΔP141 binds to *cos*-containing DNA with only modest specificity. This is not entirely unexpected, however, as discrimination between specific and nonspecific DNA is poorly demonstrated in gel retardation experiments performed in this manner. DNA binding specificity is better demonstrated by direct competition between the substrates (*32*), and the data presented in Figure 7C show the results of such a study. In this experiment, gpNu1ΔP141 was added to a binding mixture containing a ^{32}P-radiolabeled *cos*-containing DNA substrate and increasing concentrations of either specific (*cos*-containing) or nonspecific cold competitor DNA. The figure demonstrates that a *cos*-containing DNA substrate more effectively competes with complex formation than does a duplex of random sequence. Interestingly, however, the degree of discrimination in this assay (≈40-fold) is similar to that observed for both the full-length subunit and the gpNu1ΔK100 deletion mutant (Table 3).

Table 2: Direct DNA Binding by Wild-Type and Mutant gpNu1 Subunits[a]

protein	$C_{1/2}$ (nM) (specific DNA)	$C_{1/2}$ (nM) (nonspecific DNA)	discrimination[b]
gpNu1-FL[c]	35.2 ± 2.5	$(>1.6 \times 10^3)$[d]	≫45
gpNu1ΔP141	11.5 ± 1.6	31.1 ± 2.5	2.7
gpNu1ΔK100[e]	$(50.5 \pm 0.4) \times 10^3$	$(75.9 \pm 0.7) \times 10^3$	1.5

[a] The data presented in Figure 7A were analyzed as described under Experimental Procedures. [b] Discrimination= $C_{1/2}$(nonspecific DNA)/$C_{1/2}$(specific DNA). [c] The values presented differ slightly from those previously published (*15*) and reflect differences in binding conditions. Specifically, the concentration of radiolabeled DNA was 20 pM, and 50 mM NaCl was included in the binding buffer as described under Experimental Procedures. [d] No detectable retarded band was observed in the presence of 1.6 μM protein, and the discrimination value thus represents a lower limit. [e] Data taken from (*14*).

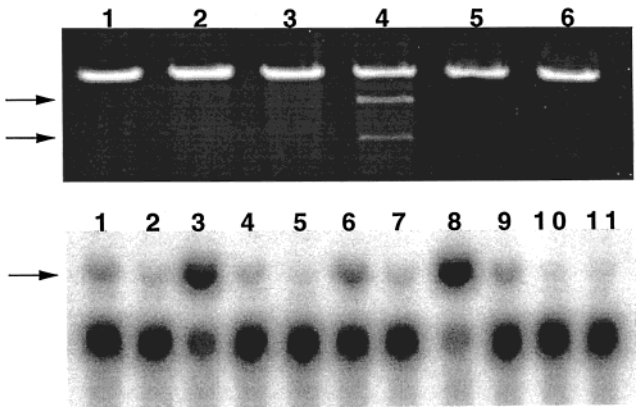

FIGURE 8: Catalytic activity of gpNu1ΔP141. Upper panel: The *cos*-cleavage assay was performed as described under Experimental Procedures with the following additions: lane 1, no additions; lane 2, gpA alone; lane 3, full-length gpNu1 alone; lane 4, gpA plus full-length gpNu1; lane 5, gpA plus gpNu1ΔP141; lane 6, gpNu1ΔP141 alone. gpA and gpNu1 (wild-type or mutant) were added to a final concentration of 1 μM and 2 μM, respectively. The positions of the DNA product bands are indicated with arrows. Lower panel: The ATPase assay was performed as described under Experimental Procedures in the absence (lanes 1−5) and presence (lanes 6−10) of DNA and with the following additions: lanes 1 and 6, gpA alone; lanes 2 and 7, full-length gpNu1 alone; lanes 3 and 8, gpA plus full-length gpNu1; lanes 4 and 9, gpA plus gpNu1ΔP141; lanes 5 and 10, gpNu1ΔP141 alone; lane 11, no additions. gpA and gpNu1 (wild-type or mutant) were added to a final concentration of 1 μM and 2 μM, respectively. The position of the ADP product band is indicated with an arrow. Quantitation of the ATP hydrolysis data is presented in Table 4.

Catalytic Activity of gpNu1ΔP141. Binding of gpNu1 to *cosB* is critical to the assembly of gpA at *cosN* and the *cos*-cleavage reaction (Figure 1). We have shown that gpNu1ΔP141 binds, albeit with reduced affinity, to the terminase gpA subunit in solution (see Figure 6). The protein further retains DNA binding activity, and it was feasible that a DNA-bound holoenzyme complex might form and possess catalytic activity. Figure 8A demonstrates that while neither

Table 3: Competitive DNA Binding by Wild-Type and Mutant gpNu1 Subunits[a]

protein	$C_{1/2}$ (nM) (specific DNA)	$C_{1/2}$ (nM) (nonspecific DNA)	discrimination[b]
gpNu1-FL	0.12 ± 0.1	7.3 ± 1.6	60
gpNu1ΔP141	0.19 ± 0.03	7.4 ± 1.2	39
gpNu1ΔK100[c]	1.5 ± 0.15	>50	>30

[a] The data presented in Figure 7C were analyzed as described under Experimental Procedures. [b] Discrimination= $C_{1/2}$(nonspecific DNA)/$C_{1/2}$(specific DNA). [c] Data taken from (*14*).

Table 4: ATPase Activity of the Terminase Subunits[a]

additions	DNA (10 nM)	relative activity (%)
gpA	no	8
gpNu1	no	2
gpA + gpNu1	no	100
gpA + gpNu1ΔP141	no	4
gpNu1ΔP141	no	0.1
gpA	yes	20
gpNu1	yes	3
gpA + gpNu1	yes	100
gpA + gpNu1ΔP141	yes	9
gpNu1ΔP141	yes	0.1

[a] The data presented in Figure 8B were analyzed as described under Experimental Procedures. 100% relative activity corresponds to 34.5 and 45.5 μM ADP formed in 30 min in the absence and presence of DNA, respectively.

the isolated gpA subunit nor the full-length gpNu1 subunit alone possess any detectable nuclease activity under these experimental conditions,[2] terminase holoenzyme reconstituted from the individual subunits (gpA_1·$gpNu1_2$) is fully active. "Reconstitution" of terminase holoenzyme from gpA and the gpNu1ΔP141 deletion mutant does not, however, yield any detectable nuclease activity. It is important to note that no *cos*-cleavage activity is detected even though gpNu1ΔP141 binds tightly to DNA at these concentrations (see Figure 7).

We have previously characterized the ATPase activity of terminase holoenzyme and have identified catalytic sites within each subunit of the enzyme (*22*, *23*). The active site P-loop motif (*33*) is contained within gpNu1ΔP141, and it was possible that this mutant protein might retain ATPase activity. We thus directly examined the catalytic activity of the isolated gpNu1ΔP141 subunit and terminase enzyme "reconstituted" from this protein. Figure 8B demonstrates that while the isolated gpA subunit possesses modest ATPase activity, neither gpNu1-FL nor gpNu1ΔP141 exhibits significant catalytic activity under these experimental conditions. Figure 8B and Table 4 demonstrate that while reconstitution of terminase holoenzyme from gpA and the full-length gpNu1 subunit significantly stimulates ATP hydrolysis, enzyme prepared from gpNu1ΔP141 does not exhibit activity beyond that observed for gpA alone.

DISCUSSION

Terminase holoenzyme is purified as a gpA_1·$gpNu1_2$ holoenzyme complex that is catalytically active in vitro (*10*, *34*). This subunit stoichiometry is retained throughout the purification procedures, suggesting strong protein·protein interactions between the subunits. Genetic studies have demonstrated that a gpA-interactive domain of the gpNu1 subunit resides within the C-terminal half of the protein (*12*, *13*), but the precise location of this domain has not been defined. We previously constructed gpNu1ΔK100, a mutant subunit in which the C-terminal 81 amino acids have been deleted (*14*). Unlike the isolated full-length protein that appears to be in a molten globule state and shows a strong tendency to aggregate (*15*), gpNu1ΔK100 is a highly soluble dimer and is folded in solution (*14*). Based on genetic studies and our experiments with these two proteins, we suggested that (i) a region in the C-terminal half of gpNu1 interferes with proper folding of the subunit in the absence of gpA, (ii) deletion of this region allows the N-terminal DNA binding domain to adopt a folded conformation, (iii) the strongly hydrophobic region extending between amino acids Lys100 and Pro141 is responsible for gpNu1 self-association and is required for cooperative DNA binding, and (iv) the gpA-interactive domain of gpNu1 is restricted to the C-terminal 40 amino acids of the protein.

Consistent with the presence of a putative hydrophobic domain, gpNu1ΔP141 has a strong tendency to aggregate. Nevertheless, the mutant protein remains soluble in the presence of low concentrations of guanidine hydrochloride (GDN), and fluorescence and CD spectroscopic studies suggest that the mutant protein is folded under these buffer conditions. Thermally induced denaturation of gpNu1ΔP141 reveals two apparent unfolding transitions, and we posit that this represents the unfolding of two independent domains in the protein, presumably the N-terminal DNA binding domain (residues 1−100) and the hydrophobic self-association domain (residues 101−141). Moreover, CD spectroscopy suggests that the folded conformation of the N-terminal DNA binding domain, represented by the shorter deletion mutant gpNu1ΔK100, is not significantly perturbed by the self-association domain, consistent with independently folding domains.

While the spectroscopic data have shown that gpNu1ΔP141 is folded in solution, functional assays are required to verify that the protein domain(s) retain(s) the native fold found in full-length gpNu1. The mutant protein possesses the hydrophobic self-association domain postulated to be important in cooperative DNA binding, and the data presented here clearly demonstrate that this region of the protein is indeed important in high-affinity DNA binding interactions. Interestingly, however, gpNu1ΔP141 only modestly discriminates between *cos*-containing and nonspecific DNA substrates, suggesting that the C-terminal 40 amino acids of the protein are important in specific DNA binding interactions. We speculated that the gpA-interaction domain of gpNu1 was limited to this C-terminal region of the protein. Contrary to our prediction, gpNu1ΔP141 binds with modest affinity to the terminase gpA subunit, clearly indicating that at least part of the hydrophobic domain is involved in subunit interactions. The mutant protein possessed neither nuclease nor ATPase activity, however, suggesting that the observed interactions are insufficient to generate a catalytically competent holoenzyme complex.

The data presented here and in our previous publications (*14*, *15*) are consistent with the gpNu1 domain model presented in Figure 2A. The region spanning Met1−Lys100 defines a soluble DNA binding domain that also possesses ATP

[2] While *cos*-cleavage activity has been observed with the isolated gpA subunit, significantly higher concentrations than those used in these experiments are required.

binding (but not hydrolysis) activity (T. de Beer, N. Berton, and C. E. Catalano, unpublished). This domain binds DNA significantly less tightly than the full-length protein, however, and discriminates only modestly between *cos*-containing and nonspecific DNA substrates. The region spanning Lys100−Pro141 defines a hydrophobic self-association domain that is required for high-affinity DNA binding, and it is likely that protein·protein interactions mediated by this domain are required for the unusual stability of the gpNu1−nucleoprotein complexes assembled at *cosB* (Figure 1) (*35*). While this region of the protein is sufficient for high-affinity DNA binding, the region spanning Pro141−Glu181 is required discrimination between *cos*-containing and nonspecific DNA substrates. These C-terminal 40 amino acids are further required for appropriate subunit interactions leading to holoenzyme complex formation and catalytic activity, and modulate the folding behavior of the isolated gpNu1 subunit.

The domain organization of gpNu1 shows strong homology to a number of *E. coli* transcriptional regulatory proteins including the *lac* (lactose) repressor (*36*, *37*), *fru* (fructose) repressor (*38*), and *pur* (purine) repressor (*39*), as well as the phage λ cI repressor (*40*). Each of these proteins is composed of an N-terminal domain of ≈60 amino acids that contains a helix−turn−helix DNA binding motif, and a larger C-terminal domain involved in ligand binding and/or oligomerization interactions required for efficient DNA binding. Particularly striking are the similarities between gpNu1 assembly at *cosB* and cI repressor assembly at O_R and O_L. In each case, protein·DNA binding interactions (intrinsic binding) at three contiguous DNA elements are bolstered by protein·protein interactions (cooperative binding) in the complex (*41*). The roles of the nucleoprotein complexes assembled by cI and gpNu1 are quite different, however. The cI nucleoprotein complex acts as a "delicate switch" that must respond to cellular conditions and either maintain a lysogenic state or allow lytic development (*41*). Conversely, the gpNu1 nucleoprotein complex is responsible, at least in part, for the assembly of extremely stable packaging intermediates whose function is to protect the newly formed "sticky end" of the viral genome from cellular nucleases. We suggest that while intrinsic DNA binding energies are important in the specific recognition of *cosB* by gpNu1, cooperative DNA binding energies, presumably mediated by the hydrophobic domain of the protein, are equally important and play a major role in the assembly and stability of the resulting nucleoprotein complex. Furthermore, gpNu1 bound at *cosB* orchestrates the assembly of a gpA dimer at *cosN* and the formation of a catalytically competent nuclease complex (see Figure 1). The studies described here have defined domains of gpNu1 mediating intrinsic DNA binding interactions, self-assembly (cooperative) binding interactions, and gpNu1·gpA binding interactions required for holoenzyme complex formation. Thus, the work described here is a prelude to more sophisticated studies directed toward a mechanistic description of the complex interactions responsible for the assembly of this DNA packaging machine.

ACKNOWLEDGMENT

We are indebted to Dr. John Thompson for obtaining the mass spectral data and to Drs. Tonny de Beer and Robert Woody for helpful discussions and critical review of the manuscript.

REFERENCES

1. Hendrix, R., Roberts, J., Stsahl, F., and Weisberg, R. (1971) *Lambda II*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
2. Casjens, S. R. (1985) in *Virus Structure and Assembly* (Casjens, S. R., Ed.) pp 1−28, Jones and Bartlett Publishers, Inc., Boston, MA.
3. Furth, M. E., and Wickner, S. H. (1983) in *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W., and Weisberg, R. A., Eds.) pp 145−155, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
4. Murialdo, H. (1991) *Annu. Rev. Biochem. 60*, 125−153.
5. Catalano, C. E., Cue, D., and Feiss, M. (1995) *Mol. Microbiol. 16*, 1075−1086.
6. Becker, A., and Murialdo, H. (1990) *J. Bacteriol. 172*, 2819−2824.
7. Earnshaw, W. C., and Casjens, S. R. (1980) *Cell 21*, 319−331.
8. Black, L. W. (1989) *Annu. Rev. Microbiol. 43*, 267−292.
9. Gold, M., and Becker, A. (1983) *Methods Enzymol. 100*, 183−191.
10. Tomka, M. A., and Catalano, C. E. (1993) *J. Biol. Chem. 268*, 3056−3065.
11. Feiss, M. (1986) *Trends Genet. 2*, 100−104.
12. Frackman, S., Siegele, D. A., and Feiss, M. (1985) *J. Mol. Biol. 183*, 225−238.
13. Wu, W.-F., Christiansen, S., and Feiss, M. (1988) *Genetics 119*, 477−484.
14. Yang, Q., de Beer, T., Woods, L., Meyer, J., Manning, M., Overduin, M., and Catalano, C. E. (1999) *Biochemistry 38*, 465−477.
15. Meyer, J. D., Hanagan, A., Manning, M. C., and Catalano, C. E. (1998) *Int. J. Biol. Macromol. 23*, 27−36.
16. Pace, C. N., Shirley, B. A., and Thomson, J. A. (1989) in *Protein Structure: A Practical Approach* (Creighton, T. E., Ed.) pp 311−330, IRL Press, Oxford.
17. Pace, C. N. (1990) *Tibtech 8*, 93−98.
18. Hanagan, A., Meyer, J. D., Johnson, L., Manning, M. C., and Catalano, C. E. (1998) *Int. J. Biol. Macromol. 23*, 37−48.
19. Hang, Q., Woods, L., Feiss, M., and Catalano, C. E. (1999) *J. Biol. Chem. 274*, 15305−15314.
20. Feiss, M., Siegele, D. A., Rudolph, C. F., and Frackman, M. (1982) *Gene 17*, 123−130.
21. Woods, L., Terpening, C., and Catalano, C. E. (1997) *Biochemistry 36*, 5777−5785.
22. Tomka, M. A., and Catalano, C. E. (1993) *Biochemistry 32*, 11992−11997.
23. Hwang, Y., Catalano, C. E., and Feiss, M. (1996) *Biochemistry 35*, 2796−2803.
24. Dawson, R. M. C., Elliott, D. C., Elliot, W. H., and Jones, K. M. (1986) *Data for Biochemical Research*, Oxford University Press, New York.
25. Schmid, F. X. (1990) in *Protein Structure, a Practical Approach* (Creighton, T. E., Ed.) pp 251−285, IRL Press, New York.
26. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
27. Gill, S. C., and von Hippel, P. H. (1990) *Anal. Biochem. 189*, 283.
28. Sears, D. W., and Beychok, S. (1973) in *Physical Principles and Techniques of Protein Chemistry, Part C* (Leach, S. J., Ed.) pp 445−593, Academic Press, New York.
29. Kahn, P. C. (1979) *Methods Enzymol. 61*, 339−377.
30. Woody, R. W., and Dunker, A. K. (1996) in *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G. D., Ed.) pp 109−157, Plenum Press, New York.
31. Record, T., Ha, J.-H., and Fisher, M. A. (1991) *Methods Enzymol. 208*, 291−343.
32. Letovsky, J., and Dynan, W. S. (1989) *Nucleic Acids Res. 17*, 2639−2653.
33. Saraste, M., Sibbald, P. R., and Wittinghofer, A. (1990) *Trends Biochem. Sci. 15*, 430−434.

34. Gold, M., and Becker, A. (1983) *J. Biol. Chem. 258*, 14619−14625.
35. Yang, Q., Hanagan, A., and Catalano, C. E. (1997) *Biochemistry 36*, 2744−2752.
36. Geisler, N., and Weber, K. (1977) *Biochemistry 16*, 938−943.
37. Khoury, A. M., Nick, H. S., and Lu, P. (1991) *J. Mol. Biol. 219*, 623−634.
38. Scarabel, M., Penin, F., Bonod-Bidaud, C., Nègre, D., Cozzone, A. J., and Cortay, J.-C. (1995) *Gene 193*, 9−15.
39. Schumacher, M. A., Choi, K. Y., Zalkin, H., and Brennan, R. G. (1994) *Science 266*, 763−770.
40. Pabo, C. O., Sauer, R. T., Sturtevant, J. M., and Ptashne, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1608−1612.
41. Ptashne, M. (1986) *The Genetic Switch*, Cell Press, Cambridge, MA.

BI991408F