Cloning, Expression, and Characterization of a DNA Binding Domain of gpNu1, a Phage λ DNA Packaging Protein†

Qin Yang,‡ Tonny de Beer,§ Liping Woods,‡ Jeffrey D. Meyer,‡ Mark C. Manning,‡ Michael Overduin,§ and Carlos Enrique Catalano*,‡,||

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

Received May 28, 1998; Revised Manuscript Received September 16, 1998

ABSTRACT: Terminase is an enzyme from bacteriophage λ that is required for insertion of the viral genome into an empty pro-capsid. This enzyme is composed of the viral proteins gpNu1 (20.4 kDa) and gpA (73.3 kDa) in a holoenzyme complex. Current models for terminase assembly onto DNA suggest that gpNu1 binds to three repeating elements within a region of the λ genome known as *cosB* which, in turn, stimulates the assembly of a gpA dimer at the *cosN* subsite. This prenicking complex is the first of several stable nucleoprotein intermediates required for DNA packaging. We have noted a hydrophobic region within the primary amino acid sequence of the terminase gpNu1 subunit and hypothesized that this region constitutes a protein−protein interaction domain required for cooperative assembly at *cosB* and that is also responsible for the observed aggregation behavior of the isolated protein. We therefore constructed a mutant of gpNu1 in which this hydrophobic "domain" has been deleted in order to test these hypotheses. The deletion mutant protein, gpNu1ΔK, is fully soluble and, unlike full-length protein, shows no tendency toward aggregation; However, the protein is a dimer under all experimental conditions examined as determined by gel permeation and sedimentation equilibrium analysis. The truncated protein is folded with evidence of secondary and tertiary structural elements by circular dichroism and NMR spectroscopy. While physical and biological assays demonstrate that gpNu1ΔK does not interact with the terminase gpA subunit, the deletion mutant binds with specificity to *cos*-containing DNA. We have thus constructed a deletion mutant of the phage λ terminase gpNu1 subunit which constitutes a highly soluble DNA binding domain of the protein. We further propose that the hydrophobic amino acids found between Lys100 and Pro141 define a self-association domain that is required for the assembly of stable nucleoprotein packaging complexes and that the C-terminal tail of the protein defines a distinct gpA-binding site that is responsible for terminase holoenzyme formation.

Bacteriophage λ consists of a linear, 48.5 kb[1] double-stranded DNA genome tightly packaged within the viral capsid and a complex structure known as the tail whose function is to "inject" the genome into an *Escherichia coli* (*E. coli*) cell to initiate the infectious process (*1*−*3*). Once inside the bacterial cell, the genome circularizes through complementary 12 base single-stranded "sticky" ends forming a complex sequence known as the ***co***hesive end ***s***ite (*cos*) of the viral genome (*1*, *4*). During the latter stages of infection, the circular genome is replicated by a rolling-circle mechanism that yields linear concatemers of the viral genome, the preferred packaging substrate (*4*−*6*). The terminase enzyme, a major component of the multiprotein "machine" responsible for DNA packaging in phage λ, excises a single genome from the concatemer and inserts it into an empty pro-capsid (*6*−*9*). Similar packaging machines are also required for virus assembly in all of the tailed, double-stranded DNA bacteriophages (*10, 11*) and the eukaryotic herpes virus groups (*12*).

Packaging of λ DNA initiates with the assembly of the terminase subunits, gpA (73.3 kDa) and gpNu1 (20.4 kDa), onto the *cosN* and *cosB* subsites of *cos*, respectively, forming a stable prenicking complex (Figure 1) (*13*). gpA-mediated nicking of the duplex within *cosN* (*14*, *15*) followed by terminase-mediated strand separation (*16*−*18*) yields complex **I**, a stable packaging intermediate originally isolated from phage-infected cells (*13*, *19*, *20*). This nucleoprotein complex next binds to the portal vertex, a specialized vertex in an empty viral pro-capsid, and this triggers an ATP-

† This work was supported by National Institutes of Health Grant GM50328-02. T.B. was supported by a National Institutes of Health Fellowship No. CA74466-02.

* Address correspondence to this author.
‡ Department of Pharmaceutical Sciences, School of Pharmacy.
§ Department of Pharmacology, School of Medicine.
|| Molecular Biology Program, School of Medicine.

[1] Abbreviations used: 1D NMR, one-dimensional nuclear magnetic resonance; 2D NMR, two-dimensional nuclear magnetic resonance; β-ME, 2-mercaptoethanol; *cos*, cohesive end site, the junction between individual genomes in immature concatemeric λ DNA; DIPSI, decoupling in the presence of scalar interactions; FID, free induction decay; 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, the small subunit of phage λ terminase; gpNu1ΔK, a deletion mutant of gpNu1 lacking the C-terminal 81 amino acids; IHF, *E. coli* integration host factor; kb, kilobase; kDa, kilodalton; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; PAGE, polyacrylamide gel electrophoresis; the TER reaction, the endonuclease activity of the terminase enzyme.

10.1021/bi981271d CCC: $18.00
Published on Web 12/12/1998

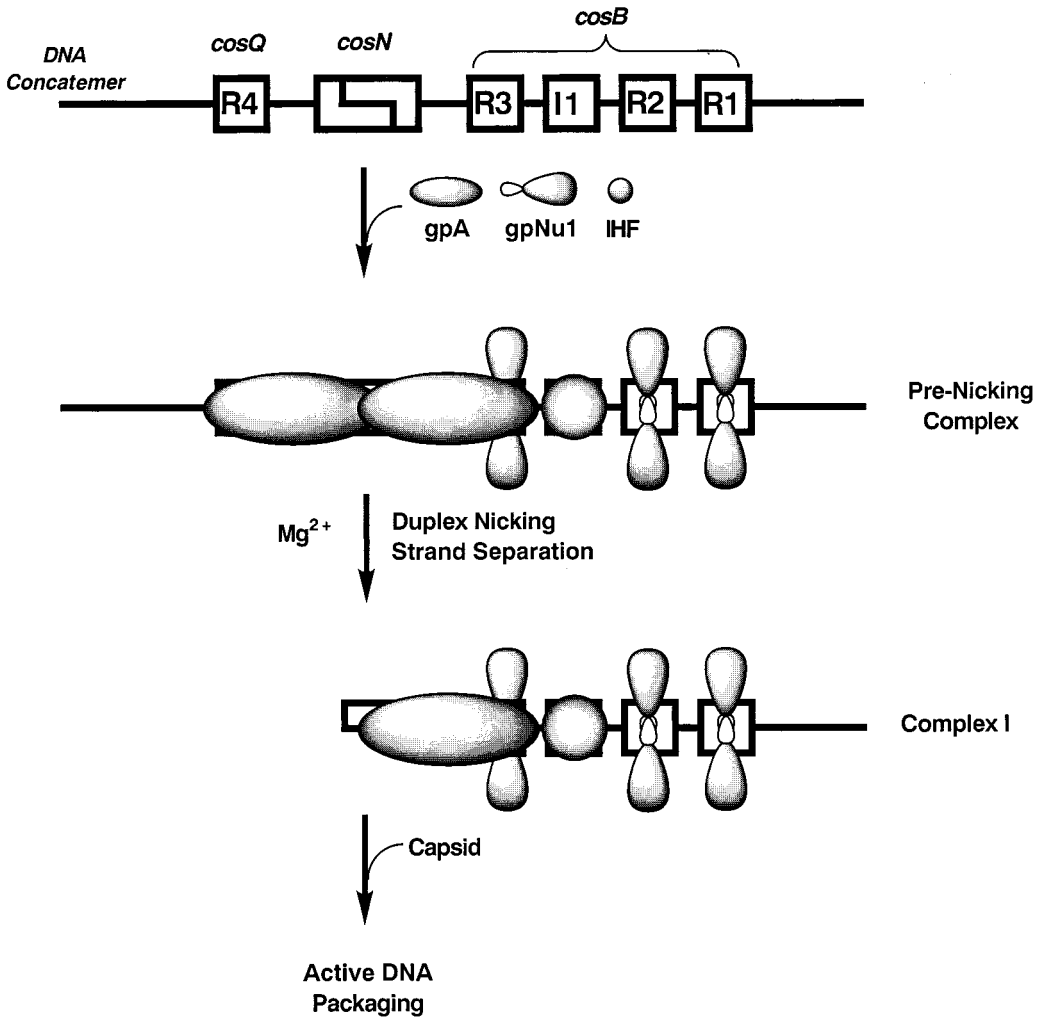

FIGURE 1: Model for Terminase Assembly at *cos*. The *cos* region of the λ genome is shown at top. The three subsites, *cosQ*, *cosN*, and *cosB* are indicated, as are the three gpNu1 binding elements (R1−R3) and the IHF binding element found within *cosB*. The terminase subunits and IHF assemble at *cos* forming a prenicking 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.

dependent translocation of the terminase subunits across the duplex and packaging of DNA within the capsid (Figure 1) (*9*, *21*, *22*). Though no direct evidence exists, it is likely that a combination of the portal vertex proteins and the DNA-bound terminase subunits constitute the packaging apparatus. Upon encountering the next downstream *cos* in the concatemer, terminase again symmetrically nicks the duplex at *cosN* and strand separation simultaneously releases the DNA-filled capsid and regenerates complex **I** (*6*, *8*, *9*). The addition of a tail to the DNA-filled capsid yields a fully infectious virus, while the regenerated complex **I** again captures an empty pro-capsid to initiate a second round of genome packaging (*23*).

While the process of DNA packaging by λ terminase is reasonably well-characterized, the nature of the nucleoprotein intermediates involved in the packaging process are ill-defined. Terminase assembly models presume that gpNu1 binds to the individual R-elements found within *cosB*, and that gpNu1 assembly at *cosB* is required for efficient assembly of a gpA dimer at *cosN* and duplex nicking (*6*, *9*, *19*) (Figure 1). These models are based on indirect evidence and speculation, however, and there are no data available on the protein composition of any of the packaging intermediates. Nevertheless, protein−protein and protein−DNA interaction "domains" have been deduced from elegant genetic studies by Feiss and co-workers as follows: (1) The C-terminal end of the gpA subunit is responsible for capsid binding (*24*, *25*) while the N-terminus is responsible for interactions with the smaller gpNu1 subunit of the enzyme (*26*, *27*). (2) A N-terminal domain of gpNu1 is required for specific DNA binding (*27*), and a putative helix-turn-helix DNA binding motif has been identified by primary sequence analysis (*28*; A. Becker, cited in ref *8*). This domain also contains an ATPase catalytic site (*29*, *30*). (3) The C-terminal $\approx$90 amino acids of gpNu1 are responsible for interactions with the larger gpA subunit (*27*).

We (*31*) and others (*32*) have constructed vectors for the expression of the individual terminase subunits in *E. coli* in order to examine their biochemical and biophysical properties. Expression of the smaller gpNu1 subunit of the enzyme invariably resulted in the formation of insoluble aggregates of the protein, however (*33*−*35*). We noted a hydrophobic stretch of amino acids within the primary sequence of the protein and postulated that this region might define a gpNu1 self-association domain responsible for the observed aggregation behavior of the protein. We further presumed that

this region might mediate protein−protein interactions required for the assembly of a stable gpNu1 nucleoprotein complex at *cosB*. In order to examine these possibilities and to define functional domains within the gpNu1 polypeptide, we have constructed a deletion mutant of gpNu1 (amino acids 1−100) which lacks the hydrophobic region of the protein. The construction, expression, and purification of this mutant protein and the characterization of its structural and functional features are presented here.

EXPERIMENTAL PROCEDURES

Materials and Methods. Tryptone, yeast extract, and agar were purchased from DIFCO. Restriction enzymes were purchased from Promega. Mono-Q HR5/5 FPLC columns, DEAE-sepharose FF, and SP-sepharose FF chromatography resins were purchased from Pharmacia. α^{32}P-ATP was purchased from ICN. Unlabeled nucleoside triphosphates were purchased from Boehringer Mannheim Biochemicals. Ni-NTA agarose was purchased from Qiagen. The ECL Western blotting kit was purchased from Amersham. Protein molecular weight standards for gel permeation chromatography were purchased from Pharmacia. 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 Shimadzu RF-1501 spectrofluorophotometer. A protein concentration of 10 μg/mL in 10 mM potassium phosphate buffer, pH 7.4, was used, and a buffer blank was subtracted from the fluorescence spectrum. 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. Near-UV CD spectra utilized a protein concentration of 1.1 mg/mL in a 1 cm strain-free cuvette. Data were collected between 250 and 350 nm at 0.5 nm intervals using a bandwidth of 1.0 nm and a dwell time of 5 s. Far-UV CD spectra utilized a protein concentration of 110 μg/mL in a 1 mm strain-free cuvette. Data were collected from 180 to 250 nm at 0.5 nm intervals using a bandwidth of 1.5 nm and a dwell time of 5 s. MOLDI-TOF mass spectral analysis was performed at the University of Colorado Health Sciences Center Macromolecular Resources Center. 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). Calculation of protein secondary structures based upon the far-UV CD data was performed using the SELCON program (*36*). 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. Antiterminase antibodies were also a generous gift of Dr. Feiss, and Western blot analysis was performed using the ECL nonradioactive method according to the manufacturer's protocol (Amersham). Construction of a plasmid which expresses the terminase gpA subunit containing a C-terminal hexa-histidine sequence (gpA-H6) and purification of this protein to homogeneity were performed by Woods and Catalano (in preparation). Purification of gpA and full-length gpNu1 was performed as previously described (*35*). All of our purified proteins were homogeneous as determined by SDS−PAGE and densitometric analysis using a molecular dynamics (MD) laser densitometer and the ImageQuant data analysis package. Unless otherwise indicated, protein concentrations were determined spectrally using millimolar extinction coefficients (*20*, *35*).

Construction of the gpNu1 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 (*37*). Primers were designed such that *Eco*RI and *Bam*HI restriction sequences were present at the 5′ and 3′ ends, 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** CTT CAG TTC CTG TGC GTC-[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 bold 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 100 amino acids of the protein. Amplification of the truncated gene, isolation of the PCR product, and construction of the overexpression plasmid (pNu1ΔK) was performed as described previously for the full-length protein (*35*). Colonies from BL21(DE3) cells transfected with this plasmid efficiently expressed the C-terminal deleted gpNu1 mutant protein, gpNu1ΔK, as determined from whole cell lysates analyzed by SDS−PAGE.

Expression and Purification of gpNu1ΔK. Four liters of 2X-YT media containing 50 μg/mL ampicillin, 25 mM potassium phosphate, pH 7.5, and 5 mM glucose was inoculated with a 40 mL overnight culture of BL21(DE3)-[pNu1ΔK] 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 30 °C for an additional 2.5 h and then harvested by centrifugation. Unless otherwise indicated, all subsequent steps were performed at 0−4 °C. The cell pellet was resuspended in ice cold buffer A (20 mM Tris, pH 8.0, 2 mM EDTA, 7 mM β-ME, and 10% glycerol) containing 100 mM NaCl, and the cells were disrupted by sonification. Insoluble cellular debris was removed by centrifugation (12000*g*, 30 min), and solid ammonium sulfate was added to the clarified supernatant to 50% saturation. Insoluble protein was removed by centrifugation (12000*g*, 30 min),

and gpNu1ΔK was then precipitated with the addition of ammonium sulfate to 75% followed by centrifugation. The gpNu1ΔK-containing pellet was taken into buffer A and, after dialysis against the same buffer, loaded onto a DEAE-sepharose column (200 mL) also equilibrated with buffer A. The column was developed with a salt gradient with gpNu1ΔK eluting at 180 mM NaCl. Column fractions were examined by SDS−PAGE, and the appropriate fractions were pooled, dialyzed against buffer A, and loaded onto a SP-sepharose column equilibrated with the same buffer. The column was developed with a salt gradient with gpNu1ΔK eluting at 200 mM NaCl. As before, column fractions were examined by SDS−PAGE; the appropriate fractions were pooled, dialyzed against buffer A, and loaded onto a MonoQ HR5/5 column equilibrated with the same buffer. The column was developed with a salt gradient with gpNu1ΔK eluting at 400 mM NaCl. Column fractions were examined by SDS−PAGE, and the appropriate fractions were pooled, dialyzed against 20 mM Tris buffer, pH 8.0, containing 1 mM EDTA, 7 mM β-ME, and 20% glycerol, and stored at −80 °C.

NMR Spectra. NMR samples contained 0.5−1.0 mM gpNu1ΔK in 10 mM sodium phosphate buffer, pH 7.2, and 10% (v/v) $^2H_2O/^1H_2O$. One-dimensional (1D) ^{1}H NMR experiments were recorded at 7.5, 15, 25, and 35 °C using a Varian INOVA 600 MHz spectrometer equipped with a shielded triple resonance 5 mm probe. For each 1D spectrum, 16 or 64 FIDs were acquired of 2048 data points each. A two-dimensional (2D) ^{1}H NOE spectrum (*38*) was recorded at 25 °C with a mixing time of 150 ms, collecting 512 t_1 experiments of 2048 complex data points each and acquiring 64 FIDs per t_1 experiment. 2D ^{1}H DIPSI spectra (*39*−*41*) were recorded at 25 °C with mixing times of 20 and 40 ms, collecting 860 t_1 experiments of 2048 complex data points each and acquiring 80 FIDs per t_1 experiment. The water resonance was suppressed by low-power saturation during the relaxation delay of 1.2 s. NMR spectra were processed using the NMRPipe software package (*42*). In short, time domain data were multiplied by phase-shifted squared-sine-bell functions, followed by zero-filling, Fourier transformation, and base-line corrections.

Analysis of Subunit Association. Analytical gel filtration analysis utilized a Superose 12 HR10/30 column equilibrated with 50 mM sodium phosphate buffer, pH 7.2, containing 150 mM NaCl running at 0.4 mL/min. A molecular weight standard curve was constructed by standard methods (*43*) using the following molecular weight standards: Blue dextran (FW > 2 000 000, RT = 19.8 min), bovine serum albumin (67 000, 34.1 min), ovalbumin (43 000, 35.7 min), chymotrypsinogen A (25 000, 39.6 min), and ribonuclease A (13 700, 41.1 min).

Sedimentation equilibrium experiments utilized a Beckmann XL-A analytical ultracentrifuge equipped with a Ti-60 four-place rotor. Protein samples, at the concentrations indicated in each individual experiment, were prepared by extensive dialysis against 10 mM sodium phosphate buffer, pH 7.2. The samples (100 μL) were analyzed in six-channel charcoal-epon centerpieces (12 mm) with 10 μL of FC-43 (Beckmann fluorocarbon-43) added to each sample as a base fluid. Absorbance optics were used to monitor the optical density of each sample at the indicated wavelengths. Data were collected at 0.001 cm radial increments and stored as the average of five replicate measurements. Data sets were collected until successive scans were superimposable to ensure that equilibrium had been achieved. The data were analyzed with the program NONLIN (*44*) using only those data within the absorbance range between 0.1 and 1.6 OD. The reduced apparent molecular weight (σ) for gpNu1ΔK was calculated from

$$\sigma = M_r*(1 - \nu*\rho)*\omega^2/RT$$

where M_r and ν are the molecular weight and partial specific volume, respectively, of gpNu1ΔK, ρ is the buffer density, ω is the radial velocity, R is the ideal gas constant, and T is the temperature (*45*, *46*). The buffer density was calculated as described by Laue et al. (*46*), and a value of $\nu = 0.7289$ was calculated from the amino acid composition of the protein (*46*).

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 gpNu1ΔK 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 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 conducted as described by Yang et al. (*13*) except that the radiolabeled DNA substrate (*cos*-containing or nonspecific) was added at a concentration of 20 pM. 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 by Tomka and Catalano using pAFP1 as a nuclease substrate (*20*). ATPase catalytic activity was examined as described previously (*47*) and utilized an ATP concentration of 50 μM and a DNA (Sca1-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 gpNu1ΔK. Purification of phage λ terminase has historically been frustrated by the insolubility of the overexpressed protein (*20*, *33*), likely due to self-association of gpNu1, the small terminase subunit (*34*, *35*). The hydropathy plot for gpNu1 shown in Figure 2A reveals a hydrophobic region located between amino acids ≈100 and 140 of the protein primary sequence. We reasoned that this region of the protein might, via hydrophobic interactions, be responsible for the self-association behavior of gpNu1, resulting in the insolubility of the isolated protein. We have therefore constructed a deletion mutant of gpNu1 which truncates the protein at lysine 100 in the primary sequence (gpNu1ΔK). Importantly, 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

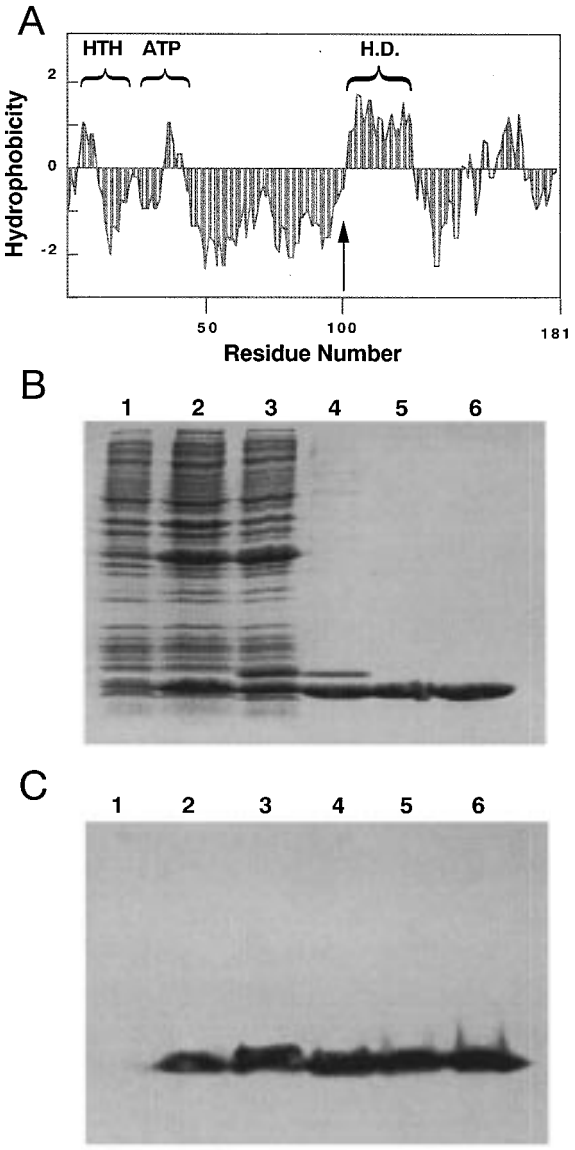

FIGURE 2: Cloning, Expression, and Purification of gpNu1ΔK. (A) Kyte−Doolitle hydropathy plot of full-length gpNu1. The locations of the hydrophobic domain (HD), the putative helix-turn-helix DNA binding motif (HTH, Lys3 to Glu22), and the Walker A (ATPase) sequence (ATP, Val29 to Asp48) are indicated. The location of the C-terminal end in the gpNu1ΔK deletion mutant protein (Lys100) is indicated with an arrow. (B) SDS−PAGE showing the purification of gpNu1ΔK: lane 1, preinduced crude cell lysate, lane 2, post induced crude cell lysate; lane 3, lysis supernatant; lane 4, pooled DEAE sepharose fractions; lane 5, pooled SP-sepharose fractions; lane 6, pooled Mono-Q fractions. (C) Western blot analysis of the gel shown in panel B.

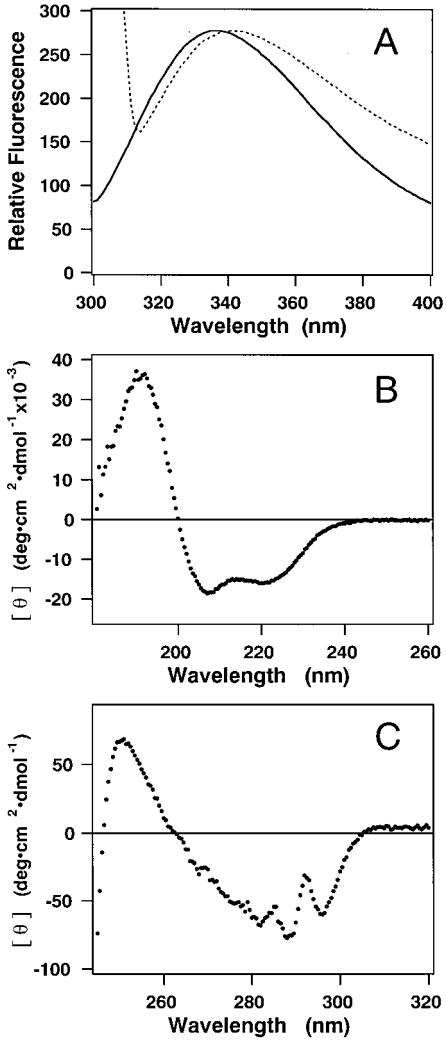

FIGURE 3: Fluorescence and CD Spectra of Purified gpNu1ΔK. (A) Fluorescence spectrum of gpNu1ΔK. The full emission spectrum is shown for excitation frequencies (λ_{ex}) of 285 (solid line) and 295 nm (dashed line). (B) Far UV−CD spectrum of gpNu1ΔK. (C) Near UV−CD spectrum of gpNu1ΔK. All spectra were recorded as described in Experimental Procedures.

C-terminus of the protein (Catalano, C. E., and Hanagan, A., unpublished).

Virtually all of the expressed gpNu1ΔK protein was present in the soluble fraction of the crude cell lysate (Figure 2B). This is in stark contrast to full-length gpNu1, which partitions exclusively into the insoluble lysis pellet (*34*, *35*). Figure 2B shows gpNu1ΔK at each stage of the purification procedure and demonstrates that the protein was purified to homogeneity. Analysis of the gel data yielded an apparent molecular weight of ≈11 500, consistent with the M_r = 11 478 predicted from the gene sequence. Consistently, MOLDI-TOF mass spectral analysis of the purified protein yielded a molecular weight of 11 478.9. Western blot analysis shows a single immunoreactive band with no evidence for proteolysis (Figure 2C), suggesting that the truncated protein is folded and stable in the crude cell lysate and throughout the purification. The protocol described here yields 6.5 mg of highly purified protein per liter of cell growth.

Preliminary Characterization of gpNu1ΔK. Unlike full-length gpNu1, purified gpNu1ΔK was fully soluble under a variety of buffer conditions and at elevated (>15 mg/mL) protein concentrations (see NMR experiments below). The protein possesses a UV spectrum (not shown) typical of a globular protein that is devoid of contaminating DNA (A_{280}: A_{260} = 1.8) (*48*−*50*). An extinction coefficient (ϵ_{280}) of 15.2 $mM^{-1}\cdot cm^{-1}$ for gpNu1ΔK was determined by the method of Gill and von Hipple (*51*, *52*). The fluorescence spectrum of the protein exhibits an emission maximum of 336 nm that remains unchanged using excitation frequencies between 260 and 285 nm ($\lambda_{ex,max}$ = 273 nm) (Figure 3A); however, selective excitation of tryptophan residues (λ_{ex} = 295 nm) results in a red shift of the emission maximum to from 336 to 342 nm. This suggests that, unlike most tyrosines in proteins, the three tyrosines in gpNu1ΔK contribute

significantly to the fluorescence spectrum when excited at shorter wavelengths (*49*, *53*). Furthermore, this red-shifted emission maximum is close to that of tryptophan free in solution (348 nm), suggesting that at least one of the two tryptophans in the protein is partially solvent exposed (*49*, *53*).

The far-UV circular dichroism (CD) spectrum of gpNu1ΔK shows strong negative maxima at 222 and 208 nm, suggesting that this protein possesses a significant amount of α-helical structure (Figure 3B) (*49*, *54*). Deconvolution analysis of the spectrum is consistent with a protein containing primarily (50%) α-helical structure as well as 27% of the residues being in a *β*-sheet conformation. These values agree well with the Chou and Fasman secondary structure predictions for the protein (53% α-helix, 26% *β*-sheet). The near-UV CD spectrum of gpNu1ΔK is shown in Figure 3C. Unlike the full-length protein, which exhibits little signal in this region of the spectrum (*35*), the deletion mutant displays a signal rich in vibronic fine structure consistent with the existence of significant tertiary structure. The intense bands observed at 296 and 289 nm likely represent the two vibronic components of the tryptophan ^1L_b band transition (*54*, *55*). The intensity of these bands suggests that at least one of the two tryptophans in the protein has restricted mobility. Furthermore, these bands are significantly red-shifted from their canonical positions of 292 and 285 nm, suggesting that there are specific interactions with nearby side chains. This may result from proximity to a nearby carboxylate residue (*56*) and/or through space electronic coupling to the lowest energy states of a nearby aromatic group (*55*).[2]

NMR Spectra of gpNu1ΔK. The fold and structural integrity of gpNu1ΔK were further investigated using NMR spectroscopy, and a typical 1D ^{1}H NMR spectrum is displayed in Figure 4A. This spectrum did not change subsequent to prolonged storage (>100 days) at 4 °C, demonstrating the structural stability of the protein. Moreover, spectra collected between 7.5 and 35 °C exhibited no significant differences, other than the expected temperature-related changes in resonance line widths and amide proton chemical shifts (not shown). The line widths of ^{1}H resonances in NMR spectra are related to the molecular mass and the shape of the molecule. Analysis of the line widths in the 1D ^{1}H NMR spectrum of gpNu1ΔK suggests that this protein is at least a dimer at these protein concentrations. Furthermore, the presence of only two tryptophan $H^{\epsilon 1}$ resonances in the spectrum ("Trp-H $^{\epsilon 1}$s", Figure 4A) suggest that the dimer is symmetric, or that neither of the two tryptophans in the protein resides within the subunit interface.

Amino acid H^α chemical shifts have been correlated with the participation of the respective amino acids in secondary structural elements (*57*). It is difficult to resolve α-helical content in a 1D spectrum due to spectral overlap; however, the NMR spectrum for gpNu1ΔK possesses several resonances in the chemical shift region associated with amino acids involved in *β*-sheet structures ("H^αs", Figure 4A). These data are consistent with the α/*β*-fold predicted by CD

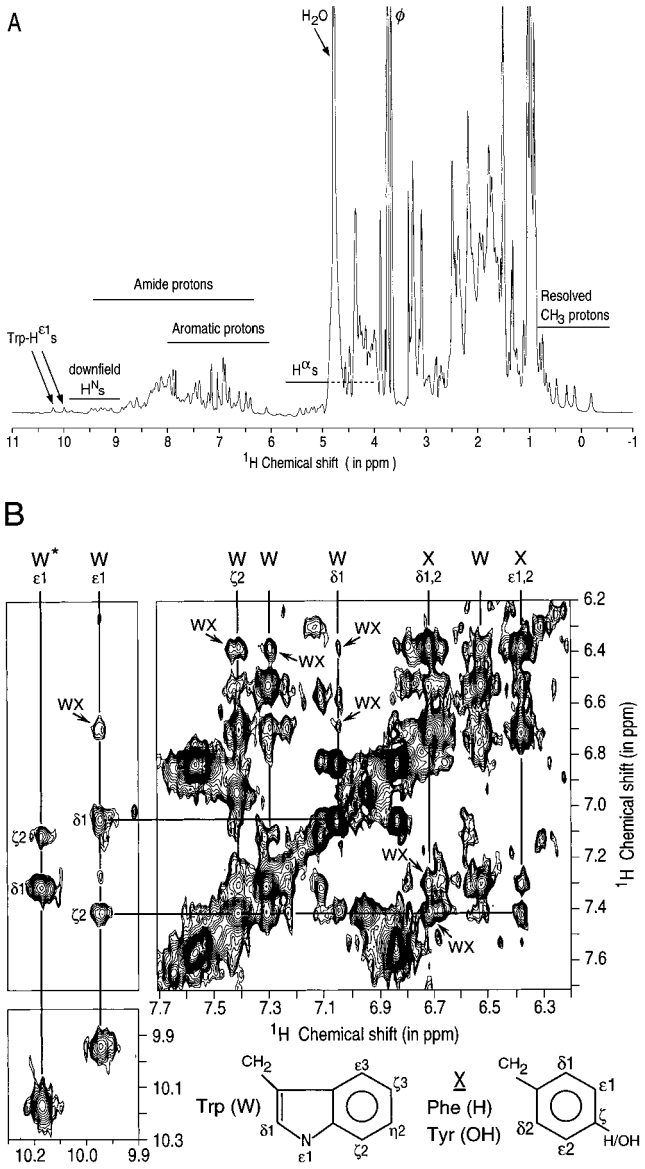

FIGURE 4: NMR Spectra of gpNu1ΔK. (A) 1D ^{1}H 600 MHz NMR spectrum of gpNu1ΔK (1 mM) in 10 mM sodium phosphate buffer, pH 7.2, recorded at 35 °C. Regions of resolved ^{1}H resonances stemming from the two tryptophan indole protons (Trp-$H^{\epsilon 1}$s), amide protons (downfield H^Ns), aromatic protons, α protons (H^α), and methyl (CH_3) protons are indicated in the figure. H_2O points to the position of the water resonance, while *φ* indicates the resonances of contaminating glycerol. (B) Fragment of the 2D ^{1}H 600 MHz NOE spectrum recorded at 25 °C. WX indicates putative through space NOE correlations between the side chains of a tryptophan residue and another aromatic residue, probably a tyrosine. The nomenclature of tryptophan, phenylalanine, and tyrosine side chains has been included in the figure for clarity. Tentative assignments of Trp and Tyr/Phe resonances were made on the basis of NOE and DIPSI (not shown) spectra.

spectroscopy (*vide supra*). The 1D spectrum further shows a broad dispersion of amide (H^N), aromatic, H^α, and methyl ^{1}H chemical shifts. Such a dispersion results from equivalent protons that reside in different local environments, which is the direct consequence of the folded nature of a protein. For example, the upfield shifts of several methyl protons ("resolved CH_3 protons", Figure 4A) reflects the shielding of these protons from the solvent by nearby amino acids. In addition, the downfield shift of several amide protons with respect to their random coil values ("downfield H^Ns", Figure

[2] Through space coupling to the lowest energy states of a nearby disulfide may also result in red-shifted tryptophan bands; however, there are no cysteine residues in gpNu1ΔK. A final, though unlikely, possibility is that the red-shifted bands in the near-UV CD spectrum result from a deprotonated phenolate, where the 0−0 band is know to occur at 295 nm (*55*).

4A) indicates their involvement in strong hydrogen bonds or their proximity to aromatic rings (*58*, *59*).

The 2D ^{1}H NOE spectrum of gpNu1ΔK shows the presence of numerous methyl−aromatic, H$^\alpha$−aromatic, and aromatic−aromatic ^{1}H−^{1}H NOE cross-peaks (Figure 4B). ^{1}H NOEs are observed only for proton pairs positioned within 5 Å of each other, and this 2D spectrum provides additional evidence that the deletion mutant is folded in solution. Preliminary analysis of this NOE spectrum confirms and extends some of the observations made with the other spectroscopic techniques. The H$^{\epsilon 1}$ chemical shifts of the two tryptophans in gpNu1ΔK (also depicted in the 1D spectrum, Figure 4A) and the NOE patterns involving these protons are clearly different, indicating that the two tryptophans reside in different chemical environments. This suggests that the solvent-exposed Trp observed in the fluorescence spectrum (Figure 3A) and the mobility-restricted Trp observed in the near-UV CD spectrum (Figure 3C) represent different residues. The NMR data thus confirm that one of the tryptophans in gpNu1ΔK is burried and mobility-restricted within the protein while the other is partially solvent-exposed, presumably closer to the protein surface. An attractive candidate for the latter is Trp22, which lies within the helix-turn-helix DNA binding motif and must be located near the surface in the DNA binding cleft. The second tryptophan residue observed in the 2D spectrum shows ^{1}H−^{1}H NOEs consistent with proximity to an aromatic residue in the protein, possibly a tyrosine (see Figure 4B). This tryptophan is likely the same residue observed in the near-UV CD spectrum of the protein and suggests that the red-shifted CD bands result from this aromatic−aromatic interaction. These data suggest, but do not prove, that the second tryptophan, Trp49, represents the buried tryptophan observed in both the CD and 2D NMR spectra.

gpNu1ΔK Is a Dimer. The above data clearly demonstrate that deletion of the C-terminal 81 amino acids of gpNu1 obviate self-association behavior, leading to aggregate formation and insolubility of the protein. The 1D ^{1}H NMR spectrum (above) suggests that gpNu1ΔK may be a dimer, and gel permeation chromatographic analysis of the protein confirms this suggestion. Purified gpNu1ΔK elutes from a Superose 12 column as a symmetric peak with an apparent molecular weight of ≈30 000 (Figure 5), consistent with a complex composed of ≈2.6 monomers. The purified gpNu1ΔK complex was further analyzed by sedimentation equilibrium experiments. Global analysis of the data at three protein concentrations and at four rotor speeds are well-described by a model that assumes a single ideal species with a subunit stoichiometry of approximately two monomers per complex (see Table 1). Similar experiments utilizing protein concentrations up to 500 μM were conducted at several temperatures (4, 27, and 37 °C) and in the absence or presence of 150 mM NaCl.[3] In each case, global analysis of the data was consistently well-described by model for a gpNu1ΔK homodimer. Including the nonideal coefficient in the analysis and/or fitting the data to higher-order models (monomer ↔ dimer, monomer ↔ trimer, etc.) did not improve the quality of the fits and in many cases yielded inferior results (not shown). All of our data are thus consistent with a gpNu1ΔK homodimer complex under all experimental conditions examined.

[3] Details of the sedimentation equilibrium experiments will be published under separate cover.

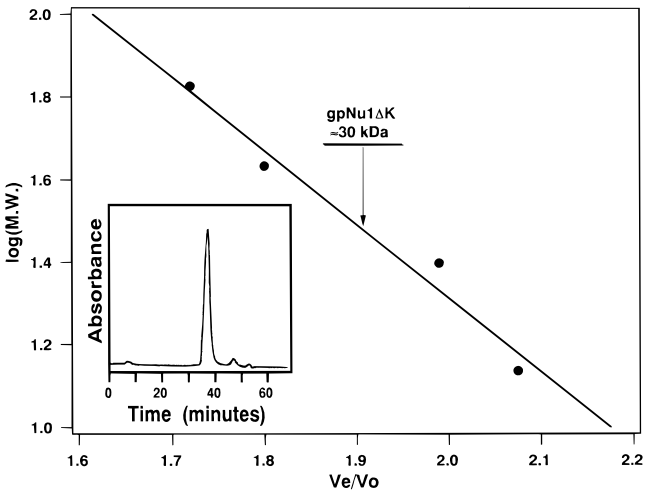

FIGURE 5: Analytical Gel Permeation Chromatographic Analysis of gpNu1ΔK. The experiment was performed as described in Experimental Procedures. The position of gpNu1ΔK (elution time, 38.1 min) is indicated with an arrow yielding an apparent molecular weight of ≈30 000, as indicated. Inset: Chromatographic profile for gpNu1ΔK. Note that the small peaks eluting at 48 and 52 min were also observed in the buffer blank.

Table 1: Sedimentation Equilibrium Analysis of gpNu1ΔK[a]

temp (°C)	[Nu1ΔK] (μM)	s^b	σ (95% conf)	N_{app}
4	5			
4	25	0.0153	1.55 (1.45−1.65)	2.1
4	150			
27	5			
27	25	0.0125	1.24 (1.15−1.32)	1.8
27	150			

[a] Sedimentation equilibrium experiments were performed at the indicated temperatures and protein concentrations. The data were analyzed by assuming an ideal monomer model as described in Experimental Procedures. Global analysis included data collected at rotor speeds of 22.5K, 27K, 32.5K, and 40K rpm and all three protein concentrations indicated in the table (12 data sets per analysis). [b] Square root of the variance of the fit. The 95% confidence intervals for σ are indicated in the table. N_{app} represents the apparent oligomerization state of the protein based upon the molecular weight calculated from the experimentally-derived σ values in the table and the molecular weight predicted from the DNA sequence of the protein. Similar N_{app} values were obtained from every data set examined.

No Interaction of gpNu1ΔK with gpA. Terminase is purified as a gpA$_1$−gpNu1$_2$ holoenzyme complex that is catalytically active *in vitro* (*17*, *20*, *47*, *60*, *61*). This subunit stoichiometry is retained throughout the purification procedures, suggesting strong protein−protein interactions between the subunits. Genetic studies have shown that the C-terminal ≈90 amino acids of gpNu1 are important for interactions with the larger gpA subunit (*27*), and the gpNu1ΔK deletion mutant was thus predicted to be deficient in holoenzyme complex formation. In the experiment presented in Figure 6, terminase holoenzyme was "reconstituted" from a C-terminal hexaHistidine-tagged gpA subunit (gpA-H6)[4] and either full-length gpNu1 or the gpNu1ΔK mutant protein. The protein mixtures were passed through a nickel chelate column which retained the gpA-H6 subunit. After the column was washed to remove any unbound protein, gpA-H6 was eluted and all of the fractions were analyzed by SDS−PAGE. The data presented in Figure 6 clearly demonstrate that while

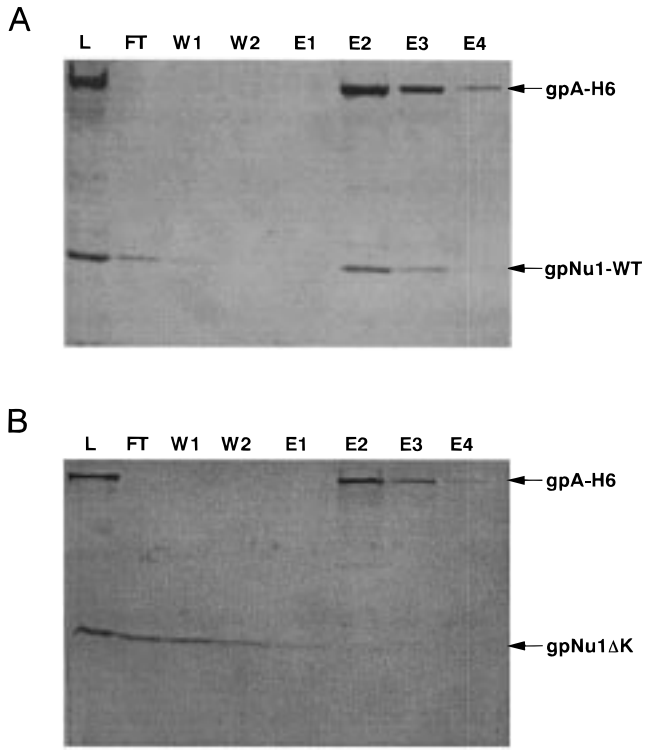

FIGURE 6: gpNu1ΔK Not Interacting with the Terminase gpA Subunit. gpA-H6 was preincubated with either wild-type gpNu1 (A) or gpNu1ΔK (B) as described in Experimental Procedures and loaded onto a nickel−chelate column. The column was washed and protein eluted as described. L, load fraction; FT, column flow through; W1−W2, column wash fractions; E1−E4, column elution fractions. The position of gpA-H6 and the relevant gpNu1 protein bands are indicated with arrows.

full-length gpNu1 readily forms a holoenzyme complex with the gpA subunit, gpNu1ΔK elutes from the column in the flow through and wash fractions with no evidence for interaction with the large terminase subunit.

gpNu1ΔK Binding to DNA. Currently accepted models for terminase assembly at *cos* suggest that gpNu1 cooperatively binds at the three R-elements found within *cosB* (Figure 1) (*6*, *9*, *19*). gpNu1ΔK possesses the putative helix-turn-helix DNA binding motif found in the full-length protein, and it was anticipated that the deletion mutant might retain the capacity to bind DNA. Initial gel retardation experiments confirmed that gpNu1ΔK indeed bound to *cos*-containing DNA but with only modest specificity (Figure 7A). Analysis of the data yields $C_{1/2}$ values of 50.5 ± 0.4 and 75.9 ± 0.7 μM for *cos*-containing and nonspecific DNA substrates, respectively. The following experiments were thus performed to ensure that gpNu1ΔK effectively discriminates between *cos*-containing and nonspecific DNA substrates. 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 (*62*). We thus examined the effect of NaCl on the stability of the gel retarded gpNu1ΔK−DNA complexes. Figure 7B shows that salt disrupts the observed protein−DNA complexes with a linear concentration dependence. Analysis of the data reveals that complexes formed with nonspecific DNA (slope = −0.387 ± 0.015) are more severely affected than are those formed with *cos*-containing DNA substrates (slope = −0.235 ± 0.060).

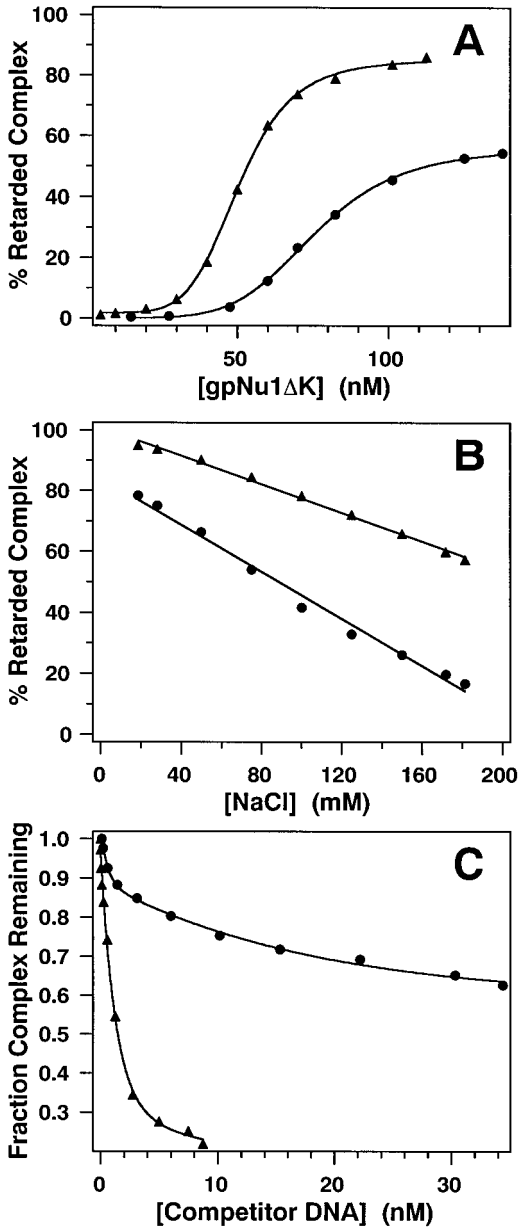

FIGURE 7: gpNu1ΔK Binding to DNA. (A) Binding of *cos*-containing vs nonspecific DNA by gpNu1ΔK. The gel retardation experiments were conducted as described and included 100 mM NaCl and either *cos*-containing (▲) or nonspecific (●) DNA. Analysis of the data yields $C_{1/2}$ values of 50.5 ± 0.4 and 75.9 ± 0.7 μM for specific and nonspecific DNA substrates, respectively. (B) Salt disrupts gpNu1ΔK−DNA complexes. The gel retardation experiments were conducted as described in Experimental Procedures using 100 μM gpNu1ΔK and 20 pM *cos*-containing (▲) or nonspecific (●) DNA substrates. NaCl was added to the binding buffer as indicated in the figure. Linear regression analysis of the data yielded salt-dependent slopes of −0.235 ± 0.060 and −0.387 ± 0.015 for specific and nonspecific DNA substrates, respectively. (C) Competitive discrimination between specific and nonspecific DNA. The experiments were conducted as described in Experimental Procedures and included 50 μM gpNu1ΔK, 20 pM radiolabeled *cos*-containing DNA, and 50 mM NaCl. Cold specific (▲) or nonspecific (●) competitor DNA was added as indicated in the figure.

[4] The cloning, expression, and characterization of the C-terminal hexaHIS gpA protein will be published separately. We note here, however, that no differences in either the endonuclease, ATPase, or strand-separation activities of terminase reconstituted from hexaHIS-tagged and wild-type gpA have been observed.

The experiment presented in Figure 7A suggests that gpNu1ΔK 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 (*63*), and DNA binding specificity is better demonstrated by direct competition between the substrates (*64*−*66*). In the experiment presented in Figure 7C, gpNu1ΔK 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 specific DNA substrate effectively competes with complex formation and that ≈1.5 nM cold competitor is required to decrease the radiolabeled complex by 50%. Similar titrations with a nonspecific DNA competitor demonstrate that significantly higher concentrations (≈50−100 nM, estimated from the curve) are required to disrupt the radiolabeled complex to a similar extent and that gpNu1ΔK binds *cos*-containing DNA with 35−70-fold greater affinity than nonspecific DNA sequences.

Catalytic Activity of gpNu1ΔK. 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 gpNu1ΔK does not interact with the terminase gpA subunit in solution (see Figure 6). The protein retains DNA binding activity, however, and it was feasible that a DNA-bound holoenzyme complex might form and possess catalytic activity. Figure 8A demonstrates that while neither the isolated gpA subunit nor the full-length gpNu1 subunit alone possess any detectable nuclease activity under these experimental conditions, terminase enzyme reconstituted from the individual subunits (gpA_1−$gpNu1_2$) is fully active. Reconstitution of terminase holoenzyme from gpA and the gpNu1ΔK deletion mutant does not, however, yield any detectable nuclease activity, even at significantly elevated protein concentrations. It is important to note that no *cos*-cleavage activity is detected even at 50 μM gpNu1ΔK, a concentration where strong binding to *cos*-containing DNA is observed (see Figure 7).

We have previously characterized the ATPase activity of terminase holoenzyme and have identified catalytic sites within each subunit of the enzyme (*30*, *47*). The active site P-loop motif is retained within gpNu1ΔK, and it was possible that this mutant protein might possess ATPase activity. We thus directly examined the catalytic activity of the isolated gpNu1ΔK subunit and terminase enzyme "reconstituted" from this protein. Figure 8B demonstrates that while the isolated gpA subunit possesses modest ATPase activity, neither the full-length gpNu1 subunit nor the C-terminal deletion mutant exhibits any detectable catalytic activity under these experimental conditions.[5] Figure 8B and Table 2 demonstrate that while reconstitution of terminase holoenzyme from gpA and the full-length gpNu1 subunit significantly stimulates ATP hydrolysis, enzyme prepared from gpNu1ΔK does not exhibit activity beyond that observed for gpA alone. Interestingly, these data suggest that the ATPase activity of the gpNu1 subunit is essentially silent in the isolated protein and that interactions with gpA in the

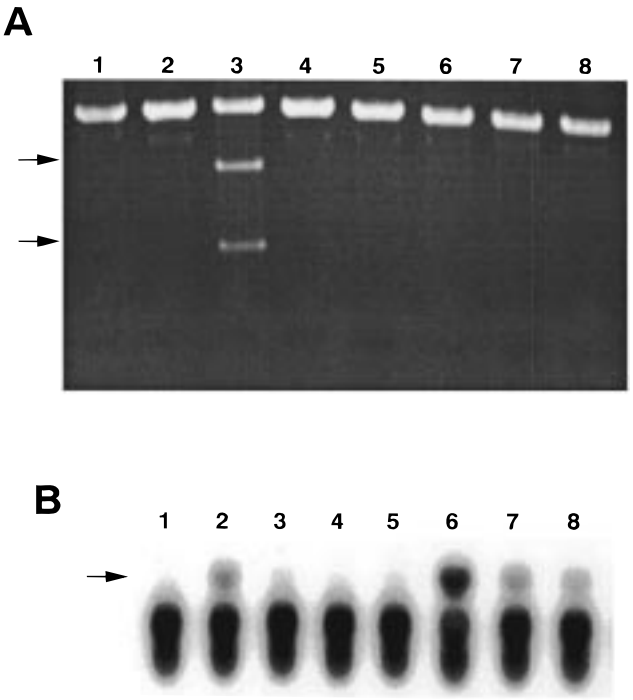

FIGURE 8: Catalytic Activity of gpNu1ΔK. (A) The *cos*-cleavage assay was performed as described in Experimental Procedures with the following additions: lane 1, wild-type gpNu1 alone (2 μM); lane 2, gpA alone (1 μM); lane 3, terminase reconstituted from gpA (1 μM) and wild-type gpNu1 (2 μM); lanes 4−7, terminase reconstituted from gpA (1 μM) and increasing concentrations of gpNu1ΔK (2, 20, 40, and 50 μM, respectively); lane 8: gpNu1ΔK alone (50 μM). The positions of the DNA product bands are indicated with arrows. (B) ATPase assay was performed as described in Experimental Procedures with the following additions: lane 1, no additions; lane 2, gpA alone (1 μM); lane 3, full-length gpNu1 alone (2 μM); lane 4, gpNu1ΔK alone (2 μM); lane 5, gpNu1ΔK alone (50 μM); lane 6, terminase reconstituted from gpA (1 μM) and full-length gpNu1 (2 μM); lanes 7 and 8, terminase reconstituted from gpA (1 μM) and gpNu1ΔK (2 μM and 50 μM, respectively). The position of the ADP product band is indicated with an arrow. Quantitation of the ATP hydrolysis data is presented in Table 2.

Table 2: ATPase Activity of the Terminase Subunits[a]

additions	ADP formed (μM)
none	ND
gpA	6
gpNu1-FL	ND
gpNu1ΔK	ND
gpNu1ΔK (50 μM)	ND
gpA + gpNu1-FL	21
gpA + gpNu1ΔK	5
gpA + gpNu1ΔK (50 μM)	4

[a] The ATPase assays were performed as described in Experimental Procedures. The terminase gpA subunit was added, as indicated, to a final concentration of 1 μM. Unless otherwise stated, the gpNu1 subunit (full length or mutant) was added, as indicated, to a final concentration of 2 μM. ND, no detectable ADP formation (<0.5 μM).

[5] Early studies suggested that the isolated gpNu1 subunit possessed an intrinsic ATPase activity with a turnover number of 38 min^{-1} and that, unlike terminase holoenzyme, it was not stimulated by DNA (*34*). More recent studies have similarly reported a gpNu1 ATPase activity, but with a much lower turnover number of 0.4−0.8 min^{-1} (*61*). Our gpNu1 preparations similarly possess an extremely weak ATPase activity, estimated to be ≈0.03 min^{-1}, that is not stimulated by DNA and that approaches wild-type levels only upon reconstitution into a holoenzyme complex. Given the ubiquitous nature of contaminating ATPases, the low ATPase activity of the isolated gpNu1 subunit, and the lack of DNA stimulation in our gpNu1 preparations, we cannot attribute the observed ATPase activity directly to gpNu1 without further experimentation.

holoenzyme complex are required for the previously characterized catalytic activity.

DISCUSSION

Terminase enzymes are found in all of the tailed, double-stranded DNA bacteriophages as well as the eukaryotic herpesvirus groups and constitute an important component of the DNA packaging "machines" (*10*−*12*). All of the known enzymes are composed of both small and large subunits, and the holoenzyme complexes utilize the energy of ATP hydrolysis to drive a single genome into the confined space within the viral capsid. Among the best characterized terminases is that of bacteriophage *λ*, and biochemical roles for each subunit have been elucidated (*6*−*9*). Several packaging intermediates have been proposed, each with specific roles in the packaging process, but the nature of these nucleoprotein complexes remains obscure.

The phage *λ* terminase gpNu1 subunit binds to three R-elements found within the *cosB* subsite of *cos* (Figure 1), and it has been suggested that protein dimers bind to these sites in a cooperative manner (*9*, *67*, *68*). Genetic studies have demonstrated that the helix-turn-helix-containing N-terminal half of the protein is required for site-specific binding to DNA (*27*). Cooperative DNA binding would require additional interactions between gpNu1 proteins bound at each of the R-elements, presumably mediated by residues found within the C-terminus of the protein. This putative domain layout is reminiscent of those found in several proteins, including the phage *λ* cI repressor and cro protein (*69*, *70*) and the *E. coli* lac repressor (*71*). Each of these proteins possesses an N-terminal, helix-turn-helix-containing domain responsible for direct DNA binding interactions and a C-terminal self-association domain important for efficient, and in some cases, cooperative DNA binding. Particularly striking similarities are observed between gpNu1 and cI as follows: (1) cI repressor dimers bind to three DNA binding elements found within the operator binding site (O_R1, O_R2, O_R3 in the P_R promoter and O_L1, O_L2, O_L3 in the P_L promoter) via a helix-turn-helix-containing N-terminal domain. (2) Self-association interactions leading to higher-order oligomers are mediated by a C-terminal domain, and protein oligomerization is required for specific and cooperative DNA binding. On the basis of these similarities, we initiated experiments directed toward identification of functional domains within gpNu1.

Inspection of the primary sequence of gpNu1 revealed a hydrophobic region of the protein between amino acids 100 and 140, and we reasoned that this portion of the protein might define a domain responsible for self-association interactions and efficient DNA binding. The gpNu1 mutant protein, gpNu1ΔK, was thus constructed which deleted the C-terminal 81 amino acids containing this hydrophobic "domain". Expression of this deletion mutant was efficient and yielded milligram quantities of fully soluble protein. Consistent with our prediction, deletion of the hydrophobic domain abrogates the self-association behavior of the full-length protein and gpNu1ΔK does not aggregate in solution. Interestingly, however, the protein is a dimer in solution, regardless of buffer composition or protein concentration. Gel permeation chromatography, sedimentation equilibrium, and NMR analysis of the protein clearly demonstrate a gpNu1ΔK dimer in the concentration range of 5 μM to 1 mM, with no evidence for further aggregation. Our data also place an upper limit of 1 μM on the monomomer−dimer dissociation constant. While the nature of the dimer cannot be determined with our present data, preliminary NMR analysis disfavors an asymmetric dimer.

Given the severity of the deletion, several experiments were performed to ensure that gpNu1ΔK retains a folded structure in solution. Strong evidence for both secondary and tertiary structural elements is provided by both circular dichroism and NMR spectroscopy. We note that this is in contrast to full-length gpNu1 that possesses secondary structure, but little, if any, evidence of tertiary structure in the absence of the gpA subunit (*35*). We have previously demonstrated that the thermal stability of gpNu1 is enhanced in the presence of gpA, and we postulated that protein−protein interactions in the holoenzyme complex are required for appropriate folding of gpNu1 (*35*). The data presented here further suggest that the C-terminal, gpA-interaction domain of full-length gpNu1 inhibits folding of the isolated protein and that deletion of this region allows the DNA-binding domain to adopt a fully folded conformation.

The spectroscopic data clearly demonstrate that gpNu1ΔK is folded in solution; however, functional assays are required to verify that the protein retains the native fold of the full-length protein. The extent of the deletion in gpNu1ΔK removes the gpA-interaction domain of the protein (*27*), and, as expected, the mutant fails to form a holoenzyme complex. These interactions are important to the assembly of a catalytically competent nucleoprotein complex, and the endonuclease and ATPase catalytic activities of the enzyme are similarly impaired. What activity is expected of the mutant protein? The deletion mutant retains the putative helix-turn-helix DNA binding motif, and specific DNA binding was expected. While gpNu1ΔK indeed binds to *cos*-containing DNA, the observed affinity is decreased by ≈100-fold as compared to the full-length protein.[6] Nevertheless, discrimination between specific and nonspecific DNA sequences is clearly evident in competition experiments where gpNu1ΔK binds to *cos*-containing DNA with 35−70-fold greater affinity than a nonspecific DNA competitor. Importantly, this level of discrimination is similar to that observed with the full-length protein (see footnote 6).

Our data thus demonstrate that the C-terminal 81 amino acids of gpNu1 are required for self-association interactions, for gpA interactions leading to the formation of a catalytically competent holoenzyme complex, and for high-affinity DNA binding. Deletion of this region of the protein yields a soluble protein dimer that retains the capacity to discriminate between *cos*-containing and nonspecific DNA substrates. Figure 9 presents a model for functional domains within gpNu1 and for cooperative DNA binding specificity. The

[6] We have previously reported that GpNu1 binds to *cos*-containing DNA yielding three complexes (N1−N3) in a protein concentration-dependent manner (*13*). We have reexamined the interaction between full-length gpNu1 and DNA using the reaction conditions described here (in particular, 20 pM DNA) and have shown that a single gel retarded complex (N1) is observed ($C_{1/2} \approx 0.5$ μM; Catalano, C. E., and Yang, Q. Unpublished). Nonspecific DNA complexes (N2/N3) have not been observed under these conditions using protein concentrations up to 3 μM. This places a lower limit of $C_{1/2}$ for nonspecific DNA ≈50 μM.

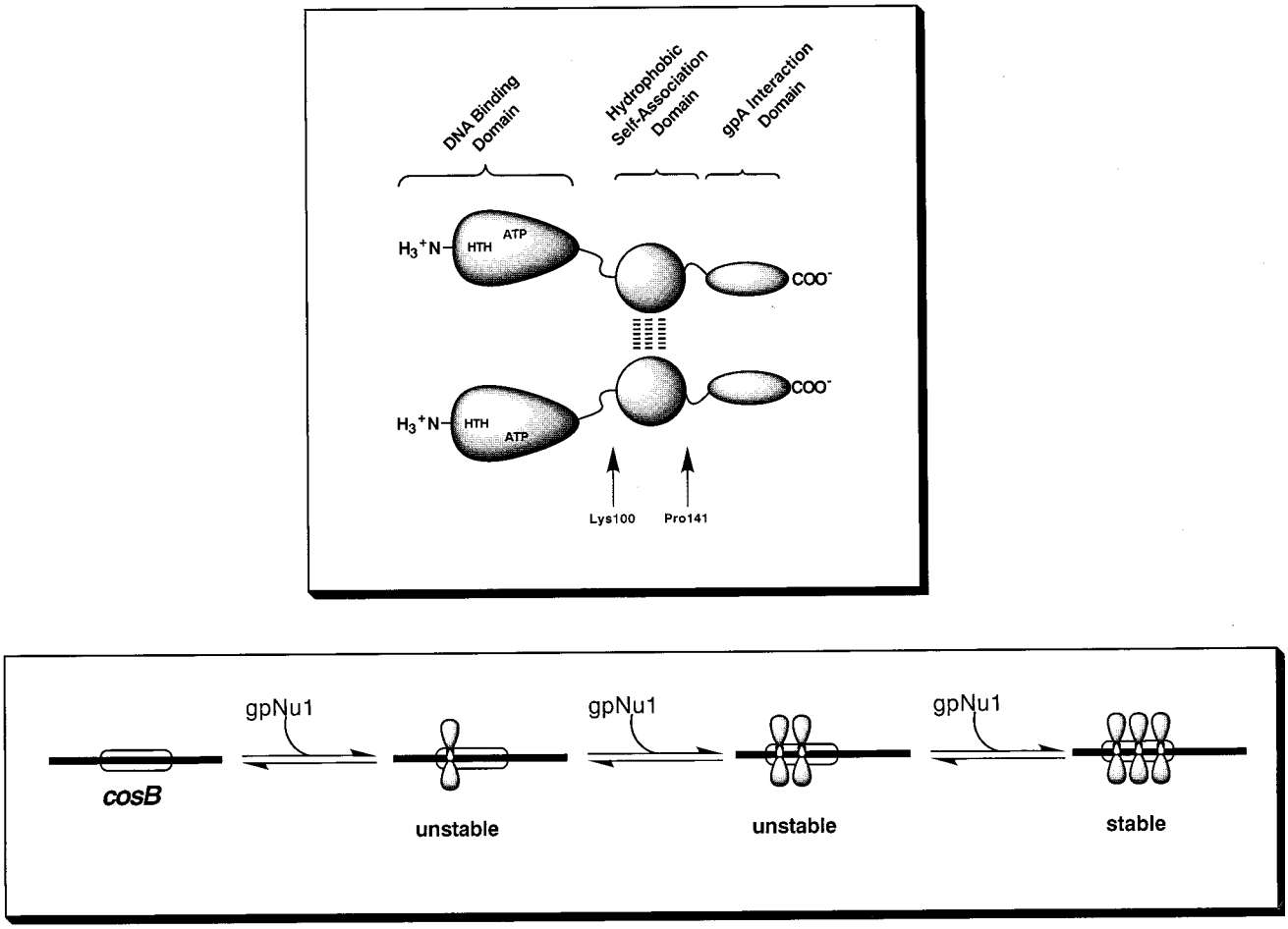

FIGURE 9: (Upper Panel) Model for the Domain Structure of gpNu1. Putative domains involved in DNA binding, self-association, and gpA interactions are indicated. HTH indicates the helix-turn-helix DNA binding motif (Lys3 to Glu22), and ATP represents the ATPase catalytic site (Val29 to Asp48) found in gpNu1ΔK. (Lower Panel) Model for the Assembly of gpNu1 at *cosB*. gpNu1 dimers are depicted as bilobed structures that bind to the three R-elements found in *cosB*. Assembly of a stable nucleoprotein complex requires occupancy of all three R-elements. Details are given in the text.

N-terminal ≈100 amino acids of the protein define a DNA binding domain required for specific DNA binding interactions (Figure 9, upper panel). Efficient assembly of gpNu1 at *cosB* requires strong interactions between proteins bound at contiguous R-elements, however, and the hydrophobic region of gpNu1 located between Lys100 and Pro141 in the primary sequence defines a domain that enhances specific and cooperative DNA binding. At this point, it is instructive to compare the nucleoprotein complexes formed by gpNu1 and the phage λ cI protein. Assembly of cI at O_R and O_L similarly involves protein dimers binding to contiguous DNA recognition elements (intrinsic binding) and in addition, protein−protein interactions (cooperative binding) in the complex (*72*). This complex has been described as a "delicate switch" that must respond to cellular conditions and either maintain a lysogenic state or allow lytic replication. This model requires reversible, facile assembly and disassembly of the nucleoprotein complex. Conversely, gpNu1 is responsible, at least in part, for the assembly of an extremely stable nucleoprotein complex whose function is to protect the newly formed "sticky end" of the viral genome from cellular nucleases. Premature disassembly of this intermediate would result in degradation of the viral chromosome and aborted viral replication. It is thus imperative that gpNu1 remain stably bound at *cosB* until a viral pro-capsid triggers terminase movement. We suggest that while intrinsic DNA binding energies are important in the specific recognition of *cosB* by gpNu1, cooperative binding energies, presumably mediated by the C-terminal domain of the protein, are equally important and play a major role in the assembly and stability of the resulting nucleoprotein complex (Figure 9, lower panel).

Finally, genetic studies that demonstrated a gpA-interactive domain within gpNu1 localized this region to the penultimate ≈90 amino acids of the protein. We suggest that this constitutes a third domain of gpNu1, separate and distinct from the self-association domain described here. We further suggest that this functional domain resides in the extreme C-terminus of the protein (Figure 9, upper panel). This predicts that deletion of the C-terminal ≈40 amino acids of gpNu1, a deletion that would retain the self-association domain of the protein, would exhibit aggregation behavior and cooperative DNA binding, but lack gpA interactions and be catalytically impaired. Construction of gpNu1ΔP, a truncation mutant deleted at Pro141, and characterization of its structural and functional properties is currently underway in our laboratory. Characterization of these, and other, mutant proteins will yield significant insight into the domain structure of gpNu1 and the role of these domains in the assembly of nucleoprotein complexes required for DNA packaging in phage λ.

ACKNOWLEDGMENT

The authors are indebted to Drs. David Bain and Michael Feiss for critical review of this manuscript. We further acknowledge the Howard Hughes Medical Institute (HHMI) for support of the University of Colorado Health Sciences Center NMR facility.

REFERENCES

1. Sanger, F., Coulson, G. F., Hill, D. F., and Petersen, G. B. (1982) *J. Mol. Biol. 162*, 729−773.
2. Hershey, A. D., and Dove, W. (1971) in *Lambda II* (Hendrix, R., Roberts, J., Stahl, F., and Weisberg, R., Eds.) pp 3−11, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
3. Katsura, I. (1983) in *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W., and Weisberg, R. A., Eds.) pp 331−346, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
4. 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.
5. Feiss, M., and Becker, A. (1983) in *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W., and Weisberg, R. A., Eds.) pp 305−330, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
6. Murialdo, H. (1991) *Annu. Rev. Biochem. 60*, 125−153.
7. Becker, A., and Murialdo, H. (1990) *J. Bacteriol. 172*, 2819−2824.
8. Feiss, M. (1986) *Trends Genet. 2*, 100−104.
9. Catalano, C. E., Cue, D., and Feiss, M. (1995) *Mol. Microbiol. 16*, 1075−1086.
10. Casjens, S. R. (1985) in *Virus Structure and Assembly* (Casjens, S. R., Ed.) pp 1−28, Jones and Bartlett Publishers, Inc., Boston, MA.
11. Black, L. W. (1988) in *The Bacteriophages* (Calendar, R., Ed.) pp 321−373, Plenum Publishing Corp., New York.
12. Roizman, B., and Sears, A. E. (1991) in *Fundamental Virology* (Fields, B. N., Knipe, D. M., and Chanock, R. M., Eds.) pp 863−865, Raven Press, New York.
13. Yang, Q., Hanagan, A., and Catalano, C. E. (1997) *Biochemistry 36*, 2744−2752.
14. Rubinchik, S., Parris, W., and Gold, M. (1994) *J. Biol. Chem. 269*, 13575−13585.
15. Woods, L., Terpening, C., and Catalano, C. E. (1997) *Biochemistry 36*, 5777−5785.
16. Rubinchik, S., Parris, W., and Gold, M. (1994) *J. Biol. Chem. 269*, 13586−13593.
17. Yang, Q., and Catalano, C. E. (1997) *Biochemistry 36*, 10638−10645.
18. Higgins, R. R., Lucko, H. J., and Becker, A. (1988) *Cell 54*, 765−775.
19. Becker, A., Marko, M., and Gold, M. (1977) *Virology 78*, 291−305.
20. Tomka, M. A., and Catalano, C. E. (1993) *J. Biol. Chem. 268*, 3056−3065.
21. Rubinchik, S., Parris, W., and Gold, M. (1995) *J. Biol. Chem. 270*, 20059−20066.
22. Hwang, Y., and Feiss, M. (1995) *Virology 211*, 367−376.
23. Feiss, M., Sippy, J., and Miller, G. (1985) *J. Mol. Biol. 186*, 759−771.
24. Sippy, J., and Feiss, M. (1992) *J. Bacteriol. 174*, 850−856.
25. Frackman, S., Siegele, D. A., and Feiss, M. (1984) *J. Mol. Biol. 180*, 283−300.
26. Wu, W.-F., Christiansen, S., and Feiss, M. (1988) *Genetics 119*, 477−484.
27. Frackman, S., Siegele, D. A., and Feiss, M. (1985) *J. Mol. Biol. 183*, 225−238.
28. Kypr, J., and Mrazek, J. (1986) *J. Mol. Biol. 191*, 139−140.
29. Becker, A., and Gold, M. (1988) *J. Mol. Biol. 199*, 219−222.
30. Hwang, Y., Catalano, C. E., and Feiss, M. (1995) *Biochemistry 35*, 2796−2803.
31. Hanagan, A., Meyer, J. D., Johnson, L., Manning, M. C., and Catalano, C. E. (1998) *Int. J. Biol. Macromol. 23*, 37−48.
32. Chow, S., Daub, E., and Murialdo, H. (1987) *Gene 60*, 277−289.
33. Murialdo, H., Davidson, A., Chow, S., and Gold, M. (1987) *Nucleic Acids Res. 15*, 119−140.
34. Parris, W., Davidson, A., Keeler, C. L., and Gold, M. (1988) *J. Biol. Chem. 263*, 8413−8419.
35. Meyer, J. D., Hanagan, A., Manning, M. C., and Catalano, C. E. (1998) *Int. J. Biol. Macromol. 23*, 27−36.
36. Sreerama, N., and Woody, R. W. (1993) *Anal. Biochem. 209*, 32−44.
37. Feiss, M., Siegele, D. A., Rudolph, C. F., and Frackman, M. (1982) *Gene 17*, 123−130.
38. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.
39. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
40. Rucker, S. P., and Shaka, A. J. (1989) *J. Mol. Phys. 68*, 509−517.
41. Shaka, A. J., Lee, C. J., and Pines, A. (1988) *J. Magn. Reson. 77*, 274−293.
42. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
43. Preneta, A. Z. (1994) in *Protein Purification Applications. A Practical Approach* (Harris, E. L. V., and Angal, S., Eds.) pp 293−305, IRL Press, New York, NY.
44. Johnson, M. L., Correia, J. A., Yphantis, D. A., and Halvorson, H. R. (1981) *Biophys. J. 36*, 575−588.
45. Laue, T. M. (1995) *Methods Enzymol 259*, 427−453.
46. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.) pp 90−125, The Royal Society of Chemistry, Cambridge, U.K.
47. Tomka, M. A., and Catalano, C. E. (1993) *Biochemistry 32*, 11992−11997.
48. Dawson, R. M. C., Elliott, D. C., Elliot, W. H., and Jones, K. M. (1986) *Data for Biochemical Research*, Oxford University Press, New York.
49. Schmid, F. X. (1990) in *Protein Structure, a Practical Approach* (Creighton, T. E., Ed.) pp 251−285, IRL Press, New York.
50. Mach, H., Volkin, D. B., Burke, C. J., and Middaugh, C. R. (1995) in *Protein Stability and Folding: Theory and Practice* (Shirley, B. A., Ed.) pp 91−114, Humana Press, Inc., Totowa, NJ.
51. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
52. Gill, S. C., and von Hippel, P. H. (1990) *Anal. Biochem. 189*, 283.
53. Lakowicz, J. R. (1983) *Principles of Fluorescence Spectroscopy*, Plenum Press, New York.
54. 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.
55. 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.
56. Goux, W. J., and Hooker, T. M. (1980) *Biopolymers 19*, 2192−2208.
57. Wishart, D. S., and Sykes, B. D. (1994) *Methods Enzymol. 239*, 363−392.
58. Wagner, G., Pardi, A., and Wüthrich, K. (1983) *J. Am. Chem. Soc. 105*, 5948−5949.
59. Pardi, A., Wagner, G., and Wüthrich, K. (1983) *Eur. J. Biochem. 137*, 445−454.
60. Gold, M., and Becker, A. (1983) *J. Biol. Chem. 258*, 14619−14625.
61. Parris, W., Rubinchik, S., Yang, Y.-C., and Gold, M. (1994) *J. Biol. Chem. 269*, 13564−13574.
62. Record, T., Ha, J.-H., and Fisher, M. A. (1991) *Methods Enzymol. 208*, 291−343.
63. Letovsky, J., and Dynan, W. S. (1989) *Nucleic Acids Res. 17*, 2639−2653.

64. Singh, H., Sen, R., Baltimore, D., and Sharp, P. A. (1986) *Nature 319*, 154−158.
65. Carthew, R. W., Chodosh, L. A., and Sharp, P. A. (1985) *Cell 43*, 439−448.
66. Carey, J. (1991) *Methods Enzymol. 208*, 103−118.
67. Shinder, G., and Gold, M. (1988) *J. Virol. 62*, 387−392.
68. Higgins, R. R., and Becker, A. (1995) *J. Mol. Biol. 252*, 31−46.
69. Pabo, C. O., Sauer, R. T., Sturtevant, J. M., and Ptashne, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1608−1612.
70. Gussin, G. N., Johnson, A. D., Pabo, C. O., and Sauer, R. (1983) in *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W., and Weisberg, R. A., Eds.) pp 93−123, Cold Springs Harbor Laboratory, Cold Springs Harbor, NY.
71. Khoury, A. M., Nick, H. S., and Lu, P. (1991) *J. Mol. Biol. 219*, 623−634.
72. Ptashne, M. (1986) *The Genetic Switch*, Cell Press, Cambridge, MA.

BI981271D