THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 31, Issue of August 5, pp. 28683–28691, 2005
Printed in U.S.A.

The Unstructured N-terminal Tail of ParG Modulates Assembly of a Quaternary Nucleoprotein Complex in Transcription Repression*

Received for publication, February 1, 2005, and in revised form, May 23, 2005
Published, JBC Papers in Press, June 12, 2005, DOI 10.1074/jbc.M501173200

Emma Carmelo, Daniela Barillà‡, Alexander P. Golovanov, Lu-Yun Lian, Andrew Derome, and Finbarr Hayes§

From the Faculty of Life Sciences, The University of Manchester, Manchester M60 1QD, United Kingdom

ParG is the prototype of a group of small (<10 kDa) proteins involved in accurate plasmid segregation. The protein is a dimeric DNA binding factor, which consists of symmetric paired C-terminal domains that interleave into a ribbon-helix-helix fold that is crucial for the interaction with DNA, and unstructured N-terminal domains of previously unknown function. Here the ParG protein is shown to be a transcriptional repressor of the *parFG* genes. The protein assembles on its operator site initially as a tetramer (dimer of dimers) and, at elevated protein concentrations, as a pair of tetramers. Progressive deletion of the mobile N-terminal tails concomitantly decreased transcriptional repression by ParG and perturbed the DNA binding kinetics of the protein. The flexible tails are not necessary for ParG dimerization but instead modulate the organization of a higher order nucleoprotein complex that is crucial for proper transcriptional repression. This is achieved by transient associations between the flexible and folded domains in complex with the target DNA. Numerous ParG homologs encoded by plasmids of Gram-negative bacteria similarly are predicted to possess N-terminal disordered tails, suggesting that this is a common feature of partition operon autoregulation. The results provide new insights into the role of natively unfolded domains in protein function, the molecular mechanisms of transcription regulation, and the control of plasmid segregation.

Plasmids are of inherent interest because of their contribution to bacterial genome plasticity, as model systems for investigating a variety of biological processes, and for their utility in gene cloning technology. The role of plasmids in the dissemination of antibiotic resistance and other properties is also particularly significant (1). The events that contribute to stable plasmid inheritance can be unraveled by analyzing the molecular mechanisms that underlie plasmid segregation. The active segregation of low copy number plasmids typically requires two plasmid-encoded proteins, one of which is most commonly a member of the ParA superfamily of Walker-type ATPases. The second protein is a DNA binding factor that interacts directly with a *cis*-acting partition site (2). The assembly of the ParA- and DNA-binding proteins on the partition site engenders a nucleoprotein complex that is required for the directional movement of paired plasmids away from the cell median (3). Plasmid movement is most likely mediated by ATP-dependent polymerization of the ParA protein or by its depolymerization (4). Unknown host factors are also probably required during segregation, for example as tethers for plasmid pairing at the mid-cell position.

The partition locus of multidrug-resistance plasmid TP228 consists of the tandem *parF* and *parG* genes and additional essential sequences located upstream of these genes (5). The ParF protein (22.0 kDa), a member of a distinctive subgroup of the ParA family, and ParG protein (8.6 kDa) assemble on the upstream region; ParG binds to this locus, and ParF is recruited through interactions with ParG, thereby forming a nucleoprotein complex whose precise architecture is likely to be crucial for correct partitioning (6). ParG is dimeric (6), consisting of a pair of unstructured N-terminal tails and intertwined, symmetric C-terminal domains that adopt a ribbon-helix-helix (RHH)[1] fold similar to that of the Arc/MetJ family of DNA binding transcriptional repressors (7). The interaction of ParG with its binding sites in the region 5′ of *parFG* is presumed to occur via the double-stranded antiparallel β-structure of the RHH domain that is inserted into the DNA major groove in a manner analogous to Arc/MetJ-type transcriptional repressors (8–10).

Plasmid partition genes are organized in operons whose expression is precisely regulated. The control of partition gene expression is mediated, at least in part, by transcriptional autoregulation exerted by one or both of the partition proteins (11–14). The importance of this regulation is revealed by the perturbation that takes place in segregation when partition genes are overexpressed, especially when the partition site is coresident (11, 13, 15, 16). Inappropriate production of partition proteins might result in their permanent occupation of the partition site, which could interfere with the migration of replication and transcription complexes (17, 18). The experiments presented here establish that the flexible N termini of dimeric ParG are essential for the organization of a stable quaternary protein complex on the *parFG* operator site and thus play a crucial regulatory role. This is achieved by transient associations between the flexible and folded domains in complex with the target DNA.

EXPERIMENTAL PROCEDURES

Strains and Media—Escherichia coli was grown at 37 °C in Luria-Bertani medium with appropriate antibiotics when needed. MacConkey agar plates with 1% maltose were used for two-hybrid assays. Strain DH5α was used for cloning, BL21(DE3) (Novagen) for protein overproduction and reporter assays, and SP850 (19) was employed for two-hybrid analysis.

* This work was supported by grants from The Wellcome Trust and the Biotechnology and Biological Sciences Research Council (to F. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ A Medical Research Council New Investigator.

§ To whom correspondence should be addressed: Faculty of Life Sciences, The University of Manchester, P. O. Box 88, Sackville St., Manchester M60 1QD, UK. Tel.: 44-161-3068934; Fax: 44-161-2360409; E-mail: finbarr.hayes@manchester.ac.uk.

[1] The abbreviations used are: RHH, ribbon-helix-helix; CDO, catechol 2,3-dioxygenase; EMSA, electrophoretic mobility shift assay; IR, inverted repeat; HSQC, heteronuclear single quantum correlation.

Plasmids and Proteins—The *parG* gene previously was amplified by PCR and cloned in pET22b(+) for ParG overproduction and in pT18 and pT25 vectors for two-hybrid analysis (6). Deletions in which 27, 57, and 90 bp were removed from the 5′-end of *parG* were constructed by PCR amplification and also cloned in pET22b(+), pT18, and pT25 vectors. Deletion proteins were purified as described for full-length ParG (6). A 117-bp DNA fragment covering the *parFG* promoter region and the first 30 bp downstream of the *parF* translational start was amplified using oligonucleotides 5′-CGTCGAATTCCATATTAACCTTTACTC-3′ and 5′-GATAGGATCCTTTCGGATT-3′. This fragment was cloned in the BamHI site of plasmid pDM3.0 (20) to generate a transcriptional fusion to the *xylE* reporter gene (pDM-Oper).

Two-hybrid Analysis—ParG-mediated association and functional complementation of the T18 and T25 fragments of the catalytic domain of *Bordetella pertussis* adenylate cyclase leads to cAMP synthesis, which triggers transcriptional activation of the maltose operon (6). This results in red colonies on MacConkey plates with maltose at 30 °C or white colonies when this operon is not activated (21). Pentapeptide scanning mutagenesis of ParG produced by the pT25 vector was performed as described previously (22–24) with the aim of assaying the effect of the insertions on protein dimerization using two-hybrid analysis.

Chemical Cross-linking—Chemical cross-linking was performed using dimethyl pimelimidate (10 mM) (Sigma) as cross-linking agent and 20 μM protein according to the protocol detailed elsewhere (6).

Sedimentation Velocity Analysis—Proteins (~1 mg/ml) were centrifuged in a 1.2-cm path length two-sector aluminum centerpiece cell in an An-60 Ti analytical rotor at 50,000 rpm in an Optima XL-I ultracentrifuge (Beckman) at 20 °C. Solute concentration changes were detected by Rayleigh interference and by monitoring A_{275}. Results were subjected to g(s*) analysis (25) using DCDT+ Version 1.13 (26) and to whole boundary analysis using Sedfit Version 8.0 (27).

CDO Reporter Assays—E. coli BL21(DE3) was transformed with pDM3.0 or pDM-Oper or with pDM-Oper and pET22b derivatives carrying full-length *parG* or 5′-*parG* deletions. A colony was inoculated in Luria-Bertani broth under selective conditions and incubated at 37 °C until A_{600} ~0.5. Cells were incubated for an additional 1 h before pelleting and resuspending in 100 mM potassium phosphate buffer, pH 7.4, with 10% acetone. After sonication, the cell lysate was cleared by centrifugation, and protein concentrations were determined using the Bio-Rad protein assay. CDO activity was measured by monitoring the A_{375} change for 1 min at 24 °C with 0.2 mM catechol. Measurements were performed at least in triplicate, each from an independent culture containing the plasmids. One CDO unit is the amount of enzyme that oxidizes 1 μmol catechol/min at 24 °C (28).

Electrophoretic Mobility Shift Assays (EMSA)—Biotinylated DNA substrates were a 48-bp oligonucleotide corresponding to the *parFG* operator (FS-48), a replica of this oligonucleotide in which one half of the palindrome was randomized (HS-48), and a shorter fragment (HS-23) covering one half-site (see Fig. 3*E*). Oligonucleotides (1 nM) were incubated at 25 °C for 20 min in binding buffer (10 mM Tris-HCl, pH 7.5, 50 mM KCl, 1 mM dithiothreitol, 5 mM $MgCl_2$, 2.5% glycerol, 0.05 μg/ml poly(dI-dC) with different amounts of ParG proteins. Reactions were electrophoresed on 6% polyacrylamide native gels in TBE (Tris borate-EDTA) buffer and blotted onto positively charged nylon membranes (Roche Applied Science). DNA was immobilized by UV cross-linking and detected using the LightShift chemiluminescent EMSA kit (Pierce).

Surface Plasmon Resonance—Surface plasmon resonance measurements used a Biacore 3000 instrument (Biacore AB). The surface of the CM5 sensor chip (Biacore) was sensitized using the Amine Coupling kit (Biacore AB) and coated with extravidin (Sigma). The same oligonucleotides used in EMSA were bound to the chip surface with the same resonance units to facilitate comparison of results between different substrates. An oligonucleotide with an identical base composition to FS-48 but with an unrelated sequence was bound to one flow cell as a reference. Proteins were injected in running buffer (10 mM HEPES-KOH, pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 1 mM dithiothreitol, 0.005% Tween 20) at a flow rate of 5 μl/min. Data were reference-subtracted using the oligonucleotide of unrelated sequence and analyzed using Biaevaluation 3.1 software (Biacore AB).

Ferguson Analysis—Δ9ParG-FS-48 and Δ9ParG-HS-48 complexes and accurate protein standards were electrophoresed in polyacrylamide gels ranging from 7 to 12% and stained with Coomassie Blue. Logarithms of the relative migrations of each species were plotted against acrylamide concentration. The $\log_{10}$ of the negative slope of each line (retardation coefficient) was plotted against the $\log_{10}$ of the molecular weight, and molecular weights of protein-DNA complexes were determined from the resulting linear plot (29).

Circular Dichroism (CD) Measurements—CD spectra in the near UV region (250–320 nm) were measured in a JASCO J-810 spectropolarimeter using a path length of 0.1 cm at 20 °C. ParG (10 μM) and FS-48 (1.25–20 μM) were prepared in CD buffer (20 mM potassium phosphate, pH 7.4, 50 mM KCl) and incubated at 20 °C for 20 min before recording the CD spectra. The stoichiometry of the ParG-FS-48 complex was determined by normalizing the change of ellipticity at 273 nm of the DNA in solution and in complex with ParG.

NMR Spectroscopy—NMR measurements were performed on a Bruker DRX600 spectrometer with a triple resonance CryoProbe. Sequence-specific resonance assignment of ParG used three-dimensional ^{15}N-edited TOCSY-HSQC and NOESY-HSQC spectra with uniformly ^{15}N-labeled protein (1 mM). Spectra were acquired at T = 293 K. Proton chemical shifts were referenced to the methyl resonance of sodium 2,2-dimethyl-2-silapentane-5-sulfonate at 0 ppm. Chemical shift mapping analysis used proteins (50–100 μM) in buffer containing 90% H_2O, 10% ^2H_2O, 100 mM NaCl, 50 mM NaH_2PO_4/Na_2HPO_4, 1 mM dithiothreitol, 40 mM arginine, 40 mM glutamate, pH 5.5. Arginine and glutamate were added to reduce protein aggregation and to increase the concentration of soluble protein (30). Time domain spectral data were processed by NMRPipe (31), and spectra were visualized and assigned with NmrView (32). Weighted amide chemical shift differences δ in spectra 1 and 2 resulting from deletion mutations or DNA binding were measured as shown in Equation 1

$$\delta_{1-2} = \sqrt{(\delta_1^{H} - \delta_2^{H})^2 + ((\delta_1^{N} - \delta_2^{N})/10)^2} \qquad \text{(Eq. 1)}$$

where δ^H and δ^N are proton and nitrogen chemical shifts, respectively. If signals disappeared from the spectra, arbitrary values of $\delta > 0.1$ppm were assigned solely to represent this.

NMR Analysis of Protein Complexes with FS-48—FS-48 oligonucleotide (80 μM) in 2 ml of NMR buffer was split in four equal aliquots. Concentrated protein solutions (50 μl) were added to 0.5 ml of DNA solution so that the concentration of dimeric protein in each case was <75 μM, allowing only one binding site to be occupied on each oligonucleotide.

RESULTS

ParG Dimerization Does Not Require the Unstructured N-terminal Tail—ParG self-association can be monitored *in vivo* using a two-hybrid system based on reconstitution of adenylate cyclase activity in *E. coli* (6, 21). A set of ParG derivatives with insertions at 13 different positions was constructed by pentapeptide scanning mutagenesis (22–24). None of nine insertions in the ParG N-terminal tail detectably affected the association with wild-type ParG in two-hybrid analysis (Fig. 1*A*). In contrast, all four pentapeptide insertions in the folded domain abolished heterodimerization with wild-type ParG (Fig. 1*A*).

The preceding pentapeptide insertions influence the sequence and composition of the ParG flexible tail but alter tail length less dramatically. More radical modifications of the N-terminal tail involved deleting 9, 19, and 30 residues (Δ9ParG, Δ19ParG, and Δ30ParG, respectively) (Fig. 1*A*). These deletions were chosen on the basis of NMR data that showed variations in the dynamic characteristics of the N-terminal domain, suggesting that this domain is not wholly a random coil (7). The deletion variants both interacted with wild-type ParG (Fig. 1*B*) and self-associated (Fig. 1*C*) in the two-hybrid system as readily as full-length ParG.

In chemical cross-linking, ParG principally forms dimeric species (6). Δ9ParG, Δ19ParG, and Δ30ParG displayed similar patterns over a range of incubation temperatures (Fig. 1*D*). Furthermore, sedimentation velocity measurements revealed the predominant presence of a single dimeric species for each of the three proteins (Fig. 1*E*), similar to results obtained from analytical ultracentrifugation studies of full-length ParG (6). Taken collectively, the two-hybrid, cross-linking, and sedimentation velocity data conclusively demonstrate that the N-terminal mobile tail of ParG plays no detectable role in dimerization.

ParG Is the Autorepressor of the parFG Genes and Interacts with an Operator Site That Overlaps the Putative parFG Promoter—A 259-bp fragment upstream of the *parF* translational

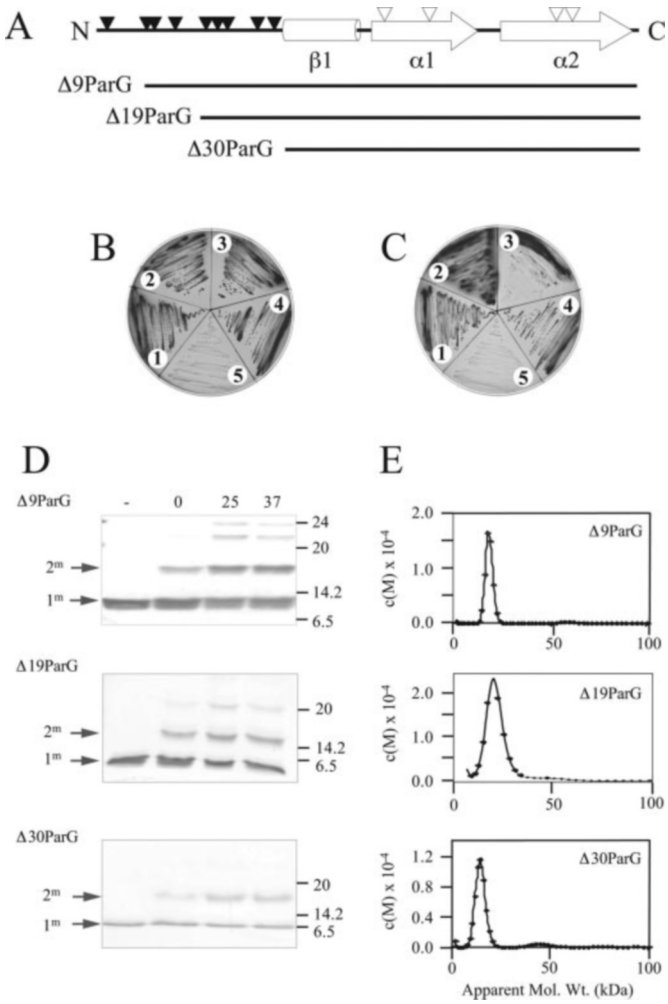

FIG. 1. **Mutational analysis of the flexible N-terminal tail of ParG.** *A*, secondary structure elements are indicated by *cylinders* (β-sheets) and *arrows* (α-helices) (7). *Filled* and *open* triangles indicate pentapeptide insertions that have no effect or abolish ParG dimerization in a two-hybrid assay, respectively. The insertions follow residues 4, 9, 10, 15, 22, 23, 24, 27, 29, 44, 53, 66, and 67. *B*, two-hybrid analysis showing the effect of N-terminal deletions in ParG on interaction with full-length ParG. Sectors show strains producing the following proteins: *1*, ParG + Δ9ParG; *2*, ParG + Δ19ParG; *3*, ParG + Δ30ParG; *4*, ParG + ParG; *5*, vector + ParG. In each case, *parG* was cloned in pT25 and the truncated derivatives in pT18. Colonies in which an interaction between the test proteins is evident or not are *red* and *white*, respectively, and appear as *dark* and *light* colonies in this image. *C*, two-hybrid analysis showing the effect of ParG N-terminal deletions on homodimerization. Sectors show strains producing the following proteins: *1*, Δ9ParG + Δ9ParG; *2*, Δ19ParG + Δ19ParG; *3*, Δ30ParG + Δ30ParG; *4*, ParG + ParG; *5*, two empty vectors. In each case, *parG* genes were cloned in both pT18 and pT25. *D*, chemical cross-linking using dimethyl pimelimidate with N-terminal deletions of ParG at 0, 25, and 37 °C. Reactions without dimethyl pimelimidate are in the *leftmost lanes*. Protein oligomers and molecular mass markers (kDa) are indicated on the *left* and *right*, respectively. *E*, sedimentation velocity analysis of Δ9ParG, Δ19ParG, and Δ30ParG. Plots of radial scans at 275_{nm} are shown.

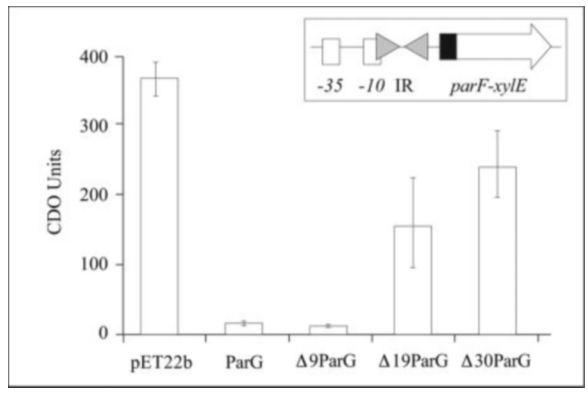

FIG. 2. **ParG is a transcriptional autorepressor of the *parFG* promoter assessed by CDO activity of the *xylE-parF* fusion in the presence and absence of ParG and its deletion derivatives.** Results are averages ± S.D. of at least three experiments performed in triplicate. *Inset*, schematic representation of the *parF-xylE* construct.

start codon is necessary for plasmid partitioning (5) and is bound by ParG (6). The 80-bp region immediately upstream of the *parF* initiation codon includes an inverted repeat (IR) that consists of imperfect 16-bp half-sites separated by a 4-bp spacer (see Fig. 3*E*). *In silico* analysis pinpointed the probable *parFG* promoter to the 5′-end of the 80-bp region. The candidate −10 and −35 elements for this promoter each contain 4/6 matches to the consensus promoter for σ^{70} of *E. coli*, with optimal separation of the boxes. The −10 element overlaps the left half-site of the 36-bp IR (Fig. 2, *inset*). This organization suggests that the IR is an operator site implicated in regulation of *parFG* expression.

To assess the role of the upstream region in *parFG* expression, a *xylE* transcriptional fusion was constructed by cloning a 117-bp fragment covering the promoter, IR, and 5′-end of *parF* in a promoterless *xylE* cassette in plasmid pDM3.0 (Fig. 2). This fusion produced 301 ± 15 catechol 2,3-dioxygenase (CDO) units, compared with the low activity of the promoterless plasmid (~0.5 CDO units), demonstrating that the *parFG* promoter is indeed located in the short region immediately upstream of the *parF* translational start.

The ParG protein was provided from a pET22b(+) expression vector, and its effect in CDO assays was examined to assess whether it was a transcriptional repressor of the *parFG* promoter. Expression of the *parF-xylE* fusion was reduced 20-fold from 360 ± 25 CDO units in the presence of pET22b(+) to 18 ± 7 units in the presence of ParG (Fig. 2). To test the role of the IR in this repression, the rightward half of the palindrome was scrambled to a similar sequence to that present in the *in vitro* substrate HS-48 (Fig. 3). CDO production from a construct containing the mutated IR was similar to that from the unmutated plasmid, demonstrating that promoter activity is unaltered by the mutations. However, ParG repressed expression from this promoter only 3-fold. These results demonstrate that ParG is a transcriptional repressor of *parFG* expression and that the IR is an operator site required for this repression.

The interaction between ParG and the IR region was explored further by EMSA with a biotinylated 48-bp oligonucleotide covering the IR (FS-48) (Fig. 3*E*). ParG formed two well defined complexes (I and II) with this substrate that eventually resolved into the more slowly migrating species (II) at high protein concentration (Fig. 3*A*). ParG did not retard a half-site of the IR located on either a 23-bp oligonucleotide (HS-23) (Fig. 3*A*) or on a 28-bp substrate bearing 6-bp extensions on both ends of the half-site (data not shown). In contrast, a substrate (HS-48) in which the right half-site of FS-48 was randomized formed a single complex with a migration similar to complex I (Fig. 3*A*). Thus, although ParG does not associate with an isolated half-site (HS-23), addition of the unrelated sequence in HS-48 stabilizes the interaction, perhaps by the association of ParG dimers bound to opposite half-sites. In this case, ParG bound to the intact half-site might be positioning additional ParG protomers non-specifically on the randomized sequences.

The stoichiometry of ParG bound to FS-48 was assessed by a variation of the Ferguson method (33) adapted for protein-DNA complexes (29). Gel filtration, analytical ultracentrifugation, and dynamic light scattering showed that the mobile tail of

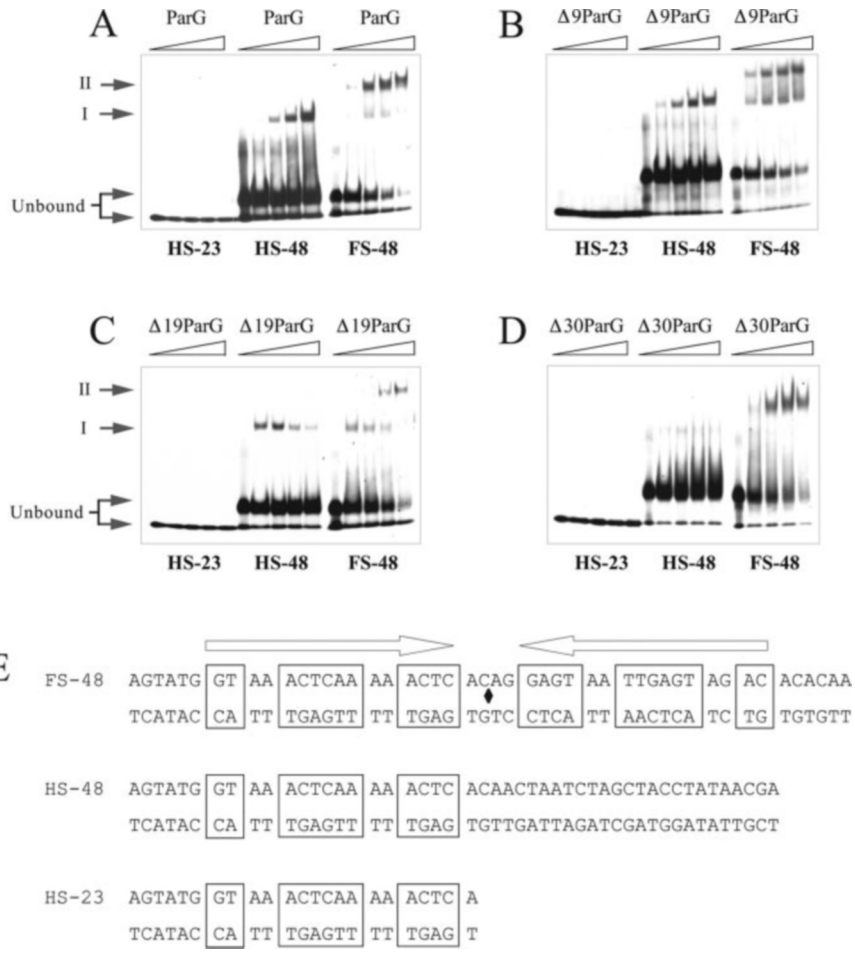

FIG. 3. **DNA binding by purified His-tagged ParG and deletion derivatives.** EMSA of ParG (*A*), Δ9ParG (*B*), Δ19ParG (*C*), and Δ30ParG (*D*) to the FS-48 (*right*), HS-48 (*center*), and HS-23 (*left*) oligonucleotides. Protein concentrations assayed were (*left to right* in each set): 0, 0.1, 0.25, 0.5, and 1 μM (monomers). Positions of complexes I and II and unbound substrates are shown. Note that the fastest migrating species in lanes containing HS-48 or FS-48 are likely to be either the same fragments with an atypical structure or single-stranded oligonucleotides. *E*, sequences of oligonucleotides employed in surface plasmon resonance and EMSA. A *black diamond* indicates the axis of symmetry, with half-sites indicated with *open arrows*. Conserved positions are *boxed*.

ParG greatly increases its hydrodynamic radius.[2] To reduce this effect, Δ9ParG was selected for Ferguson analysis as it retains the activities of full-length ParG in EMSA (see following section) but lacks approximately one-third of the tail. A set of protein standards and complexes I and II formed between Δ9ParG and FS-48 or HS-48 were analyzed in gels containing 7–12% acrylamide. The relative migrations of the main bands were plotted against the polyacrylamide concentrations of the gels (Fig. 4*A*). The retardation coefficients were represented against the molecular masses, and the sizes of the nucleoprotein complexes were estimated from the resulting plot (Fig. 4*B*). This yielded molecular masses of 72 and 101 kDa for complexes I and II, respectively. These figures are in close agreement with the 64.5 kDa calculated for complexes formed by FS-48 and two Δ9ParG dimers (29.5 kDa DNA + 2 × 17.5 kDa protein) and 99.5 kDa for FS-48 and four Δ9ParG dimers (29.5 kDa DNA + 4 × 17.5 kDa protein).

The ParG:FS-48 binding stoichiometry was also analyzed by CD measurements, showing an increase of ellipticity at 273 nm caused by the titration of ParG with FS-48 compared with the free DNA. Fig. 4*C* shows the quantitative monitoring of that change, plotted as a percentage of the maximum change. The plot indicates that the ellipticity variations attain a plateau at a 4:1 molar ratio of ParG dimers:FS-48, establishing a complex similar to the previously described complex II.

The N-terminal Tail Is Required for Repression and Modifies the Interaction of ParG with Its Operator Site—The Δ9ParG, Δ19ParG, and Δ30ParG derivatives were examined for repression of the *parF-xylE* reporter to assess whether the unstructured tail influences transcriptional repression by ParG (Fig. 2). Expression of the fusion was strongly repressed by Δ9ParG (13 ± 6.6 CDO units), producing a repression ratio very similar to that determined for full-length ParG (Fig. 2). In contrast, repression of the *parFG* promoter was dramatically reduced for Δ19ParG and Δ30ParG, which had repression ratios of only 2.3- and 1.5-fold, respectively. Thus the unstructured tail of ParG plays a crucial role in transcriptional repression of the *parFG* promoter, and increasingly longer deletions of the tail are accompanied by progressively weaker repression.

The involvement of the N-terminal tail in the interaction of ParG with the operator was investigated further by EMSA (Fig. 3, *B–D*). Δ9ParG and Δ19ParG displayed patterns of binding to FS-48 that were very similar to those observed for ParG: complex I was most evident at low protein concentrations, with the appearance of complex II at higher protein:DNA ratios. For Δ9ParG and Δ19ParG, a large proportion of the substrate was bound into complex II at the highest protein concentration (Fig. 3, *B* and *C*). In contrast, interaction of Δ30ParG with FS-48 produced a smeared band with a different migration from either complexes I or II, suggesting that the association of Δ30ParG with the substrate was particularly perturbed (Fig. 3*D*). The singular behavior of Δ30ParG was emphasized by experiments with HS-48, with which the protein failed to form a defined complex, only smearing it from the unbound position. In contrast, Δ9ParG and Δ19ParG both associated with HS-48 with similar patterns as full-length ParG, producing complex I in each case. Like ParG, none of the deletion proteins interacted detectably with HS-23.

Surface plasmon resonance experiments were performed to quantitate the ParG-operator interactions. Using a Biacore

[2] F. Hayes, D. Barillà, and E. Carmelo, unpublished data.

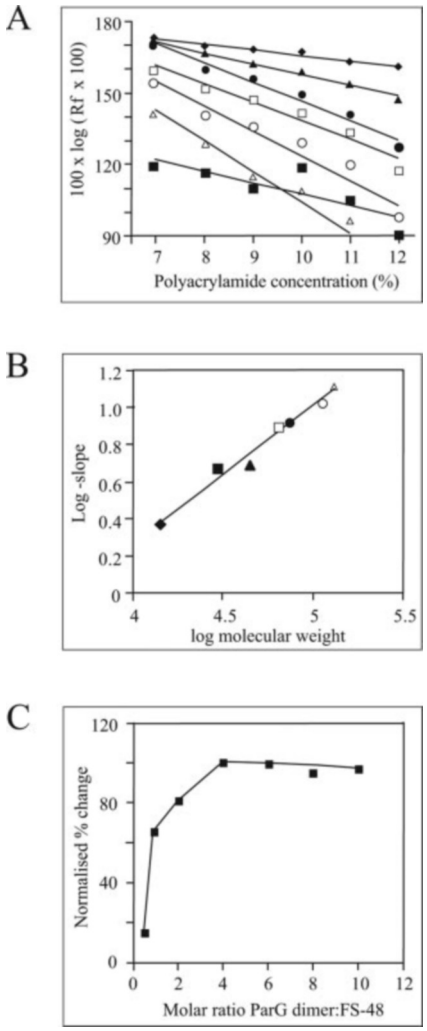

FIG. 4. **Stoichiometry of ParG binding to the operator site.** *A*, Ferguson analysis of Δ9ParG binding to FS-48 and HS-48. Relative migrations of protein standards and protein-DNA complexes were plotted against the acrylamide concentration of the gel. ■, carbonic anhydrase; ▲, chicken egg albumin; □, BSA monomer; △, BSA dimer; ◆, α-lactalbumin; ●, HS-48-Δ9ParG; ○, FS-48-Δ9ParG. *B*, logarithms of negative slopes of best fits of plots in *panel A* are represented against logarithms of molecular weights of each standard. *C*, representation of the % change in ellipticity at 273 nm induced upon titration of ParG with FS-48.

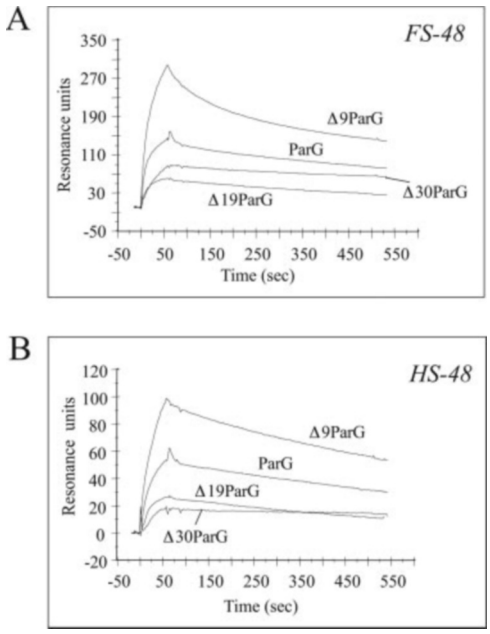

FIG. 5. **Surface plasmon resonance studies of the interaction of ParG and deletion derivatives with FS-48 and HS-48 substrates.** Sensorgrams of ParG, Δ9ParG, Δ19ParG, and Δ30ParG (0.5 μM) binding to FS-48 (*A*) and HS-48 (*B*).

3000 instrument, a CM Sensor Chip (Biacore) was extravidin coated and derivatized with FS-48 in Flow-Cell (FC) 1, HS-48 in FC2, and 48 bp of unrelated DNA as a reference in FC3. Proteins were passed over the chip surfaces at 0.05–5 μM (monomer equivalents) for 60 s (association) and allowed to wash off subsequently for 450 s (dissociation). ParG and the N-terminal deletion variants bound strongly to FS-48 at a concentration of 0.5 μM (Fig. 5*A*), and a similar effect was observed at every protein concentration tested (data not shown). Nevertheless, the interaction with this target was different between the proteins. Δ9ParG reproducibly gave the strongest response with FS-48, as opposed to Δ19ParG whose response was much weaker. Sensorgrams of ParG, Δ9ParG, and Δ19ParG had noticeably curved association and dissociation phases, strikingly different from those of Δ30ParG, which presented a much flatter pattern particularly in the dissociation phase when the slope was almost horizontal. The observation in EMSA analysis that ParG established two different complexes with FS-48 suggested that ParG and the N-terminal deletion derivatives follow a complex model in operator site binding. None of the kinetic models proposed by BIAevaluation 3.1 software gave a satisfactory close curve fitting for this complex interaction, therefore excluding them from derivation of binding constants for the FS-48 binding.

Binding of full-length ParG and the N-terminal deletion derivatives to HS-48 presented some characteristics unnoticed in previous sensorgrams (Fig. 5*B*). First, the profiles are almost straight lines in both association and dissociation phases, as opposed to the arched curves produced by protein binding to FS-48. This suggests that very different kinetics describe the binding of ParG and its N-terminal deletion derivatives to the FS-48 and HS-48 fragments. Second, in apparent contrast to the absence of binding to HS-48 in EMSA (Fig. 3*D*), Δ30ParG exhibits some binding to HS-48 in the surface plasmon resonance experiments. This effect might be because of nonspecific interactions.

Intramolecular Interactions between the N- and C-terminal Domains of ParG—The role of the N-terminal domain of ParG in complex formation with its operator site was analyzed further by NMR spectroscopy. It was first assessed whether the N- and C-terminal domains of ParG interact using chemical shift changes. HSQC spectra were recorded for uniformly ^{15}N-labeled ParG and its derivatives Δ9ParG, Δ19ParG, and Δ30ParG under identical conditions (Fig. 6). Comparison of spectra pairs revealed that deletions Δ9 and Δ19 induced alterations in the positions of signals originating only from residues as far as 7 positions immediately after the deletion site, suggesting that region 1–19 does not interact with the folded domain (Fig. 7*A*). The largest changes were induced by the Δ30 mutation, which affected not only the N terminus residues 31–33 immediately following the deletion site but also a significant number of residues (34–37, 40, 41, 43, 44, 46–48) within the β-strand and the first α-helix of the folded domain. These results suggest that part or all of region 20–30 interacts transiently with the folded domain of ParG. To analyze whether this region adopts any weak secondary structure, chemical shift deviations of C^{α}, C^{β}, CO, and H^{α} atoms from values typical for the random coil state were examined (34). In addition to the three stable regions of secondary structure (β-strand 34–41, α-helices 42–55 and 60–74) for each ParG monomer (7), the

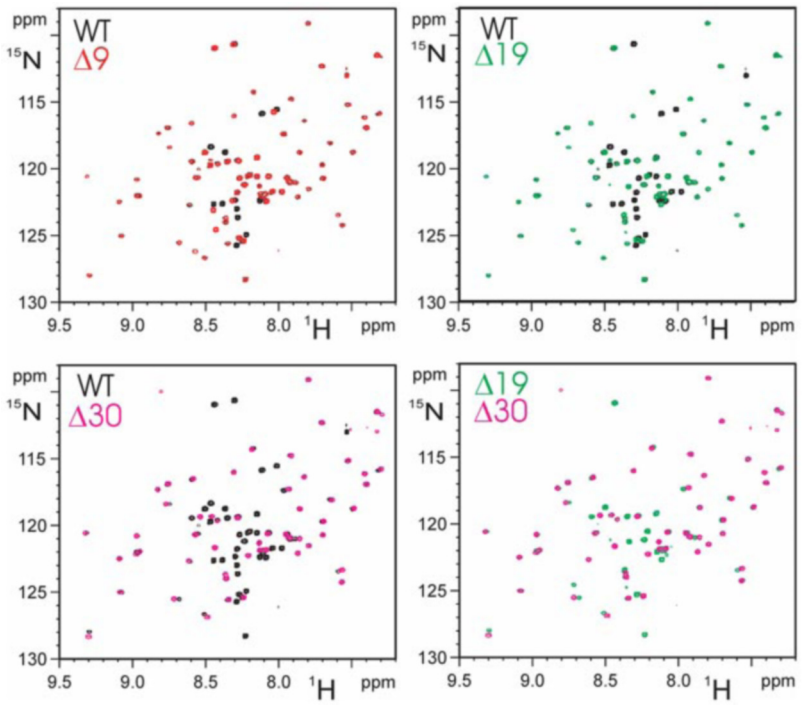

FIG. 6. **Pairwise overlay of two-dimensional ^{1}H-^{15}N-HSQC spectra of full-length ParG (*black*) and its N-terminal deletion derivatives, Δ9ParG (*red*), Δ19ParG (*green*), and Δ30ParG (*magenta*).**

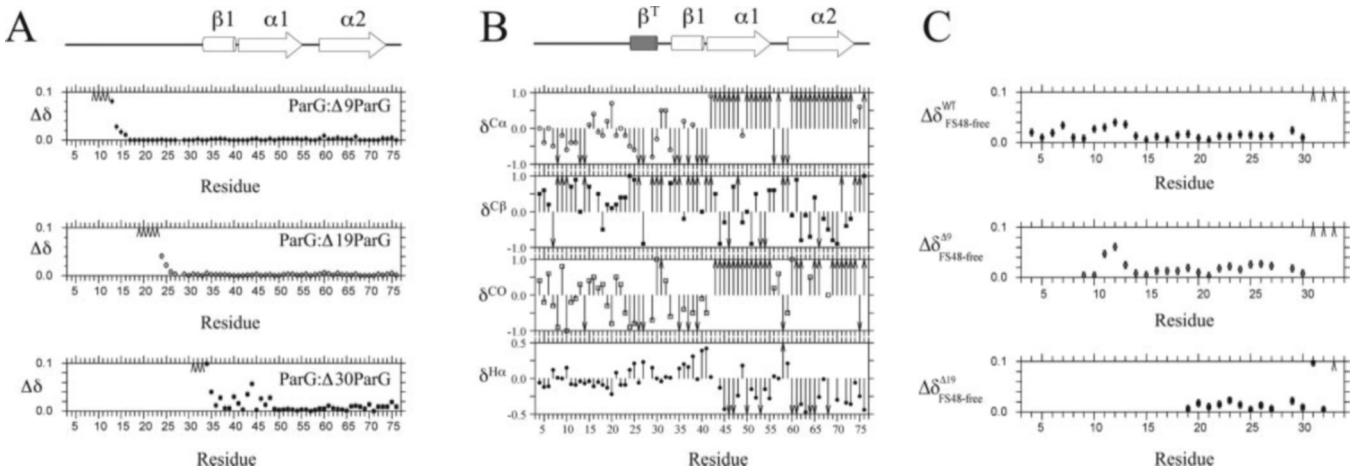

FIG. 7. **Interactions of the mobile N-terminal domain of ParG determined by NMR.** *A*, weighted distance between amide signal positions for pairs of two-dimensional ^{1}H-^{15}N-HSQC spectra of ParG-Δ9ParG, ParG-Δ19ParG, and ParG-Δ30ParG. Positions of secondary structure elements are shown schematically at the top. *B*, difference between chemical shifts of C^{α}, C^{β}, CO, and H^{α} atoms in ParG and corresponding chemical shift values observed for the unfolded state of the protein (36). The *shaded cylinder* indicates the region (β^T) that, according to consensus differences, transiently adopts a β-structure conformation. *C*, weighted distance Δδ (in ppm) between cross-peaks in pairs of ^{1}H-^{15}N-HSQC spectra shown on profiles for N-terminal residues. From *top to bottom*: ParG free and bound to FS-48, Δ9ParG free and bound to FS-48, and Δ19ParG free and bound to FS-48.

flexible region 23–29 showed chemical shift indices reminiscent of a β-structure (Fig. 7*B*, β^T). This part of the polypeptide chain might fold back to transiently form an additional strand of antiparallel β-structure per monomer, thereby forming an extended four-stranded β-sheet per ParG dimer with two stable antiparallel strands situated in the middle and two transient β^T strands on either side.

Chemical Shift Mapping of ParG-DNA Interactions—To establish which residues of ParG participate in interactions with DNA and whether these interactions are affected by the N-terminal deletions, mixtures of ^{15}N-labeled ParG, Δ9ParG, Δ19ParG, and Δ30ParG were incubated with the FS-48 substrate, and two-dimensional HSQC spectra were analyzed (Fig. 8). In a second experiment the same proteins were mixed with nonspecific genomic DNA (data not shown). In all cases, signals from the folded domain of ParG disappeared, due to the slow tumbling of the complex, confirming that this domain bound both to specific and nonspecific DNA as established previously for the full-length protein (7). Interestingly, the N-terminal domain behaved differently in the specific and nonspecific complexes. In complexes with nonspecific DNA, no significant changes in chemical shifts of signals from the flexible N terminus parts of ^{15}N-labeled ParG, Δ9ParG, and Δ19ParG were observed (data not shown). However, the magnitude and extent of changes of the N terminus residues with FS-48 are Δ9ParG >ParG >Δ19ParG, with residues His-7, Lys-11, Lys-12, Met-13, Asn-18, Arg-19, Arg-23, Thr-26, Ala-27, and Val-29 most affected (Fig. 7*C*). As the magnitude and extent of chemical shift changes can be indicators of protein-DNA complex stability, the Δ9ParG-FS-48 complex apparently is more stable than the complex involving the full-length protein. Positively charged residues Lys-11 and Lys-12 demonstrated the largest change in chemical shifts between bound and unbound form. Furthermore, signals from Ser-31, Gly-32, and Lys-33 were not

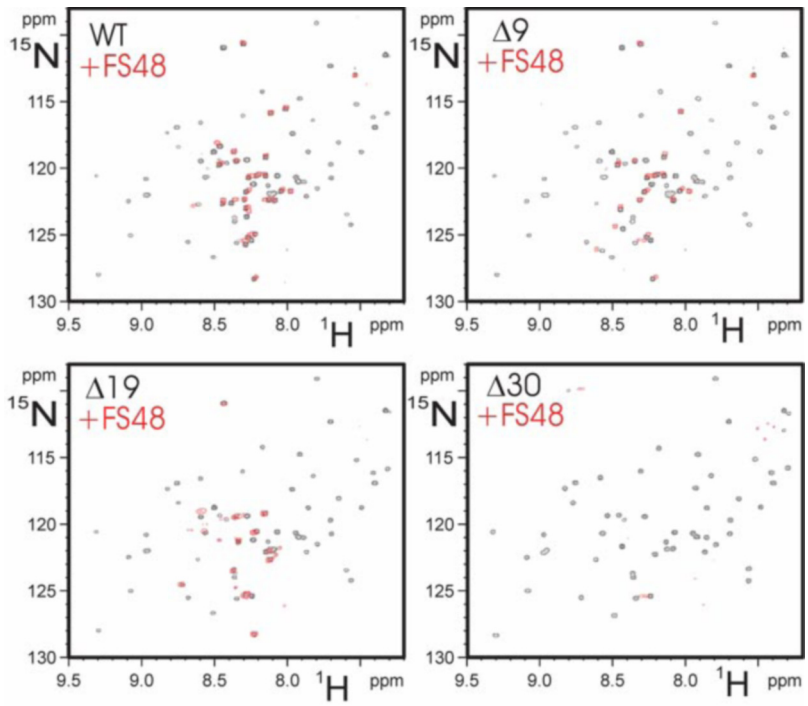

FIG. 8. **Pairwise overlay of two-dimensional ^{1}H-^{15}N-HSQC spectra of full-length ParG and its N-terminal deletion derivatives, Δ9ParG, Δ19ParG, and Δ30ParG in the absence (*black*) and presence (*red*) of oligonucleotide FS-48.**

visible in the spectrum of the Δ9ParG-FS-48 complex, but signals from residues Ser-31 and Gly-32 reappeared with Δ19ParG, probably because of a loss in cooperativity (Fig. 7*C*). Only signals from Gly-32 and the C-terminal His tag were visible in the Δ30ParG-FS-48 spectrum (Fig. 8), showing that deletion of the entire N-terminal domain did not inhibit complex formation.

To further probe the stability of the ParG-FS-48 complex an "NMR-footprinting" approach was used in which DNaseI was added to the nucleoprotein complex and DNA degradation and its subsequent effect on the spectrum were observed *in situ* (7). DNaseI causes dissociation of ParG oligomers assembled on nonspecific DNA (7). However, upon addition of DNaseI to the ParG-FS-48 complex, the FS-48 fragment was largely degraded as was apparent from the presence of numerous sharp lines in one-dimensional ^{1}H NMR arising from degraded nucleotides and the disappearance of characteristic resonances for intact DNA signals close to 13 ppm (data not shown). However, signals from the folded domain of ParG remained broadened and undetectable, suggesting that it was still bound to DNA and that ParG was protecting part of the DNA from DNaseI cleavage. In contrast, the majority of signals from the N-terminal domain that changed their chemical shifts due to binding with FS-48 reverted to their chemical shift positions in the free state (Fig. 7*C*). Concomitantly, the signals from residues Ser-31, Gly-32, and Lys-33 reappeared in the spectrum, strengthening the hypothesis that the N-terminal domain participates in specific DNA binding in a cooperative manner with the folded domain. As the N-terminal domain interaction does not confer protection of the DNA against DNaseI digestion, the interactions formed are transient and with those regions of DNA that are accessible to the enzyme, presumably at the periphery of the nucleoprotein complex. In summary, ParG protects FS-48 from complete degradation by nuclease probably because this is a much tighter complex than between ParG and nonspecific DNA. Thus the N-terminal tail of ParG contributes to specific protein-DNA interactions.

DISCUSSION

The RHH fold defines a class of proteins that exert transcriptional repression by insertion of a double-stranded antiparallel β-structure into the DNA major groove of their operator sites (35). In addition to their ordered RHH C-terminal domains, the prototypical Arc, Mnt, and MetJ proteins contain mobile N-terminal tails that make further DNA contacts and undergo conformational changes when the repressors bind to their target sites (8). These tails provide additional binding specificity in the case of Arc and Mnt (36). In contrast, the flexible C-terminal region of the ParD RHH protein is responsible for binding with a partner protein, ParE (37), and the disordered C-terminal extension of MetJ is required for co-repressor binding (38). The RHH domain of the nickel-responsive NikR protein is flanked by folded C-terminal regions that mediate tetramerization and high affinity nickel binding (39). Thus, major differences among RHH proteins reside in their N- or C-terminal extensions, which are generally crucial for their biological function. In many RHH proteins, however, the roles of these extensions remain to be established. The current work provides the first evidence that the RHH protein ParG is a transcriptional repressor of *parFG* expression and identifies the relevant operator site. The flexible N terminus of ParG modulates this repressor activity, with progressive deletions alleviating repression, thereby dramatically disrupting the activity of the protein. The disordered tails of ParG apparently are not required for dimerization as neither peptide insertions nor deletion of the tails detectably moderated the intermonomer interaction *in vivo* or *in vitro*.

Inherently unstructured proteins and protein domains are estimated to constitute a large fraction of the proteome (40). The coordinated DNA binding and local folding of unstructured proteins or protein regions have been described for a variety of systems (41). For example, heterodimerization of the MAT**a**1 and MATα2 homeodomain proteins in yeast is mediated by the disordered C-terminal tail of MATα2, part of which folds to form a helix that packs against MAT**a**1. The MAT**a**1-MATα2 complex exhibits a far higher degree of specificity and affinity for its target operator site than do the individual proteins (42,

43). Similarly, the lymphoid enhancer binding factor-1 exists in a partially disordered state but undergoes a cooperative folding transition upon DNA binding. This transition is accompanied by DNA bending that is crucial for the recruitment of other regulatory factors (44, 45). Unlike these and many other proteins in which mobile domains become structured upon DNA binding, the N-terminal domain of ParG retains significant flexibility when assembled on DNA. The changes observed in chemical shifts of amide signals from this domain between free and bound forms indicate that the interactions between the tail and DNA are transient. A number of amino acid residues in the tail are implicated in these interactions, with a significant proportion being positively charged, possibly forming electrostatic contacts with the negatively charged phosphate groups on the target. Furthermore, there is a distinct difference in the behavior of the N-terminal domain with specific and nonspecific DNA: the chemical shift changes of residues in the N-terminal domain observed in the ParG-FS-48 complex are absent in the nonspecific complex with genomic DNA, implying that the flexible N terminus of ParG particularly contributes to DNA recognition specificity. Similarly, binding to its operator site might induce alterations in the C-terminal region of ParG that are transduced to the mobile N-terminal domain. This behavior has parallels with other RHH proteins in which flexible loops that immediately precede the β-ribbons change conformation significantly when bound to the target sites, looping away from the protein surfaces to wrap around phosphate groups on the DNA (8, 10, 46).

Truncations of the mobile tail of ParG induced extensive structural transitions of nearby residues, most notably in Δ30ParG in which perturbations of the polypeptide chain were transmitted through the β-strand and into the first α-helix of the folded domain (Fig. 7*A*). The structural disruptions in Δ30ParG include pronounced alterations in the β-strand elements that fold into an antiparallel β-sheet that is predicted to be involved in DNA binding. This disruption is most likely responsible for the significantly altered properties exhibited by Δ30ParG, such as (i) the divergent kinetics of Δ30ParG binding to the operator site, (ii) abolition of the interaction with the HS-48 substrate, and (iii) the marked reduction in transcriptional repression exerted by this protein. Interestingly, region 20–30 of the N-terminal tail of ParG has a propensity to form the β^T structure that transiently interacts with the β-sheet DNA recognition element of the folded domain to give an extended β-sheet, providing two additional short-lived antiparallel β-strands per ParG dimer. Because the double β-strand of the folded domain of ParG is predicted to insert in the major groove of DNA, it is tempting to suggest that the transient β^T strands could partially cover the DNA-binding site in ParG. In other words, the β^T region 20–30 of the flexible domain might play an autoinhibitory role. In this scenario, only the most optimal contacts between the folded and N-terminal domains of ParG and the DNA will provide enough energy to overcome the thermodynamic costs associated with the removal of the transient β^T structure from the folded domain, exposing the latter for interaction with the binding site. Thus, the deletion of β^T in Δ30ParG and the subsequent exposure of the β-sheet DNA recognition element could explain the altered behavior of this protein.

The C-terminal folded domain of Δ19ParG is entirely intact, and its transcriptional repression (Fig. 2) and DNA binding (Fig. 5) defects cannot be explained simply by perturbation of the interaction between the unfolded and folded domains. Thus, in addition to transient interactions with the DNA-binding domain via β^T, the disordered N-terminal tails of ParG might mediate protomer-protomer interactions within the nucleoprotein complex, thereby contributing directly to the formation of a stable nucleoprotein complex at the operator site. In summary, it is clear that the flexible N-terminal domain has one or more complex regulatory roles in ensuring the specificity of ParG interactions with its target site.

ParG forms two EMSA complexes with its target site that resolve into a single species at high protein:DNA ratios. These nucleoprotein complexes are consistent with two- and four-ParG dimers assembled on DNA for the more quickly (I) and slowly (II) migrating complexes, respectively. ParG does not interact with an operator half-site (HS-23), but the formation of complex I with a half-site to which a randomized extension is added (HS-48) reveals that one of the two dimers within this complex is positioned, probably non-specifically, on the randomized half-site, analogously to previous observations with the Arc-operator complex (47). This supports the contention that ParG preferentially binds its cognate sites as a "dimer of dimers." Thus the minimal binding unit of ParG to the operator site is a tetramer (dimer of dimers) (complex I), each unit contacting a pair of repeats in the operator that correlates with crystal structures of other RHH proteins in complex with DNA (8–10). Higher protein:DNA ratios produce the more complex structure II in which a second tetramer is bound. The progressive deletion of 9 or 19 N-terminal residues does not drastically change this pattern of interactions in EMSA. However, deletion of the entire N terminus in Δ30ParG prevented formation of a well defined complex with FS-48. Surface plasmon resonance analysis further revealed that the binding properties of Δ30ParG differ significantly from those of longer ParG derivatives. In view of the complex interactions between ParG and FS-48, additional factors such as sequential *versus* non-sequential binding, cooperativity, whether the binding sites are independent or dependent, and whether conformational changes occur in either the target site or protein will need to be considered in future work aimed at dissecting the kinetics of ParG interactions with this substrate. Interestingly, preliminary circular dichroism analysis has already revealed that the target site undergoes conformational changes when bound to ParG.[3]

Numerous plasmids encode ParG homologs, most of which are predicted to possess N-terminal mobile extensions of ~20–40 residues (5, 7, 48). Although the primary sequences of the extensions are diverse, it is likely that these heterogeneous domains fulfil the same architectural role in transcriptional repression as does the N-terminal tail of ParG. Thus, transcriptional modulation by the action of a repressor protein bearing a flexible tail is likely to be a common theme in partition operon autoregulation by ParG-type proteins. Furthermore, in addition to its interaction with the IR, ParG binds to additional sequences that constitute the *cis*-acting partition site positioned upstream of *parFG*.[3] These sequences consist of repeats that are related to the IR but are arrayed differently (6). It will be intriguing to assess the role of the mobile N-terminal tail of ParG in the organization of the nucleoprotein complex at this alternative site.

Acknowledgments—We thank Steven West and Zorana Vujovic for contributing to the pentapeptide scanning mutagenesis studies, Andy Baron for expertise in analytical ultracentrifugation, Grazyna Jagura-Burdzy for supplying plasmid pDM3.0, John McCarthy for providing access to the Biacore instrument, and Tobias von der Haar for helpful discussions on the Biacore system.

REFERENCES

1. Hayes, F. (2003) *Methods Mol. Biol.* **235,** 1–17
2. Surtees, J. A., and Funnell, B. E. (2003) *Curr. Top. Dev. Biol.* **56,** 145–180
3. Li, Y., Dabrazhynetskaya, A., Youngren, B., and Austin, S. (2004) *Mol. Microbiol.* **53,** 93–102

[3] E. Carmelo, D. Barillà, and F. Hayes, unpublished data.

4. Barillà, D., Rosenberg, M. F., Nobbmann, U., and Hayes, F. (2005) *EMBO J.* **24,** 1453–1464
5. Hayes, F. (2000) *Mol. Microbiol.* **37,** 528–541
6. Barillà, D., and Hayes, F. (2003) *Mol. Microbiol.* **49,** 487–499
7. Golovanov, A. P., Barillà, D., Golovanova, M., Hayes, F., and Lian, L. Y. (2003) *Mol. Microbiol.* **50,** 1141–1153
8. Somers, W. S., and Phillips, S. E. (1992) *Nature* **359,** 387–393
9. Gomis-Ruth, F. X., Sola, M., Acebo, P., Parraga, A., Guasch, A., Eritja, R., Gonzalez, A., Espinosa, M., del Solar, G., and Coll, M. (1998) *EMBO J.* **17,** 7404–7415
10. Raumann, B. E., Rould, M. A., Pabo, C. O., and Sauer, R. T. (1994) *Nature* **367,** 754–757
11. Kusukawa, N., Mori, H., Kondo, A., and Hiraga, S. (1987) *Mol. Gen. Genet.* **208,** 365–372
12. Friedman, S. A., and Austin, S. J. (1988) *Plasmid* **19,** 103–112
13. Hayes, F., Radnedge, L., Davis, M. A., and Austin, S. J. (1994) *Mol. Microbiol.* **11,** 249–260
14. Kwong, S. M., Yeo, C. C., and Poh, C. L. (2001) *Mol. Microbiol.* **40,** 621–633
15. Funnell, B. E. (1988) *J. Bacteriol.* **170,** 954–960
16. Davis, M. A., Radnedge, L., Martin, K. A., Hayes, F., Youngren, B., and Austin, S. J. (1996) *Mol. Microbiol.* **21,** 1029–1036
17. Youngren, B., and Austin, S. (1997) *Mol. Microbiol.* **25,** 1023–1030
18. Sawitzke, J. A., Li, Y., Sergueev, K., Youngren, B., Brendler, T., Jones, K., and Austin, S. (2002) *J. Bacteriol.* **184,** 2447–2454
19. Shah, S., and Peterkofsky, A. (1991) *J. Bacteriol.* **173,** 3238–3242
20. Macartney, D. P., Williams, D. R., Stafford, T., and Thomas, C. M. (1997) *Microbiology* **143,** 2167–2177
21. Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 5752–5756
22. Hallet, B., Sherratt, D. J., and Hayes, F. (1997) *Nucleic Acids Res.* **25,** 1866–1867
23. Hayes, F., and Hallet, B. (2000) *Trends Microbiol.* **8,** 571–577
24. Hayes, F. (2003) *Annu. Rev. Genet.* **37,** 3–29
25. Stafford, W. F., III (1992) *Anal. Biochem.* **203,** 295–301
26. Philo, J. S. (2000) *Anal. Biochem.* **279,** 151–163
27. Schuck, P. (2000) *Biophys. J.* **78,** 1606–1619
28. Sala-Trepat, J. M., and Evans, W. C. (1971) *Eur. J. Biochem.* **20,** 400–413
29. Orchard, K., and May, G. E. (1993) *Nucleic Acids Res.* **21,** 3335–3336
30. Golovanov, A. P., Hautbergue, G. M., Wilson, S. A., and Lian, L. Y. (2004) *J. Amer. Chem. Soc.* **126,** 8933–8939
31. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6,** 277–293
32. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR* **4,** 603–614
33. Ferguson, K. A. (1964) *Metabolism* **13,** 985–1002
34. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR* **4,** 171–180
35. del Solar, G., Hernandez-Arriaga, A. M., Gomis-Ruth, F. X., Coll, M., and Espinosa, M. (2002) *J. Bacteriol.* **184,** 4943–4951
36. Knight, K. L., and Sauer, R. T. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 797–801
37. Oberer, M., Zangger, K., Prytulla, S., and Keller, W. (2002) *Biochem. J.* **361,** 41–47
38. Rafferty, J. B., Somers, W. S., Saint-Girons, I., and Phillips, S. E. V. (1989) *Nature* **341,** 705–710
39. Schreiter, E. R., Sintchak, M. D., Guo, Y., Chivers, P. T., Sauer, R. T., and Drennan, C. L. (2003) *Nature Struct. Biol.* **10,** 794–799
40. Tompa, P. (2002) *Trends Biochem. Sci.* **27,** 527–533
41. Dyson, H. J., and Wright, P. E. (2002) *Curr. Opin. Struct. Biol.* **12,** 54–60
42. Phillips, C. L., Stark, M. R., Johnson, A. D., and Dahlquist, F. W. (1994) *Biochemistry* **33,** 9294–9302
43. Li, T., Stark, M. R., Johnson, A. D., and Wolberger, C. (1995) *Science* **270,** 262–269
44. Love, J. J., Li, X. A., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) *Nature* **376,** 791–795
45. Love, J. J., Li, X., Chung, J., Dyson, H. J., and Wright, P. E. (2004) *Biochemistry* **43,** 8725–8734
46. Phillips, K., and Phillips, S. E. V. (1994) *Structure* **2,** 309–316
47. Brown, B. M., and Sauer, R. T. (1993) *Biochemistry* **32,** 1354–1363
48. Fothergill, T. J. G., Barillà, D., and Hayes, F. (2005) *J. Bacteriol.* **187,** 2651–2661