Structural and Genetic Analyses Reveal a Key Role in Prophage Excision for the TorI Response Regulator Inhibitor

Latifa Elantak, Mireille Ansaldi, Francoise Guerlesquin, Vincent Méjean, Xavier Morelli

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 280, NO. 44, pp. 36802–36808, November 4, 2005

Structural and Genetic Analyses Reveal a Key Role in Prophage Excision for the TorI Response Regulator Inhibitor*

Received for publication, July 8, 2005, and in revised form, August 1, 2005 Published, JBC Papers in Press, August 2, 2005, DOI 10.1074/jbc.M507409200

Latifa ElAntak[‡1], Mireille Ansaldi[§1], Françoise Guerlesquin[‡], Vincent Méjean[§], and Xavier Morelli[‡2]

From the [‡]Unité de Bioénergétique et Ingénierie des Protéines and the [§]Laboratoire de Chimie Bactérienne, IBSM-CNRS, 31 chemin Joseph Aiguier, 13402 Marseille Cedex 20, France

TorI (Tor inhibition protein) has been identified in *Escherichia coli* as a protein inhibitor acting through protein-protein interaction with the TorR response regulator. This interaction, which does not interfere with TorR DNA binding activity, probably prevents the recruitment of RNA polymerase to the *tor*C promoter. In this study we have solved the solution structure of TorI, which adopts a prokaryotic winged-helix arrangement. Despite no primary sequence similarity, the three-dimensional structure of TorI is highly homologous to the $^{\lambda}$Xis, Mu bacteriophage repressor (MuR-DBD), and transposase (MuA-DBD) structures. We propose that the TorI protein is the structural missing link between the $^{\lambda}$Xis and MuR proteins. Moreover, *in vivo* assays demonstrated that TorI plays an essential role in prophage excision. Heteronuclear NMR experiments and site-directed mutagenesis studies have pinpointed out key residues involved in the DNA binding activity of TorI. Our findings suggest that TorI-related proteins identified in various pathogenic bacterial genomes define a new family of atypical excisionases.

Lambdoid phages, exemplified by λ itself, constitute a wide family of temperate phages, which genomes are found either as a circular double-stranded DNA, or as a prophage integrated into the host chromosome. The phage-encoded integrase ($^{\lambda}$Int) catalyzes both integration and excision reactions helped in these functions by several accessory proteins (1). Among them two are absolutely required, the host-encoded integration host factor (IHF)[3] is required for integration and excision, whereas the phage-encoded excisionase ($^{\lambda}$Xis) is necessary for excision only. Excisionase proteins are also called recombination directionality factors (RDFs), (2) since their role is to control the activity of the integrase and to direct the reaction toward excision. $^{\lambda}$Xis plays a critical role during excision by allowing the formation of a specific complex called intasome together with the integrase, IHF, and a third accessory protein Fis (3). Simultaneously to the excision of the prophage genome, $^{\lambda}$Xis also inhibits reintegration by converting the phage attachment site (*attP*) into a catalytically inactive structure (4). $^{\lambda}$Xis functions require binding to the integrase as well as to DNA in the *attR* region at two tandemly arranged binding sites (X1 and X2). Binding to these sites promotes sharp bending of DNA and assist the integrase DNA binding to allow intasome formation (5, 6).

The TorI protein was first identified as a TorR response regulator inhibitor (7). TorR is part of the two-component system TorS/TorR required for *torCAD* operon expression in *Escherichia coli* (8, 9). In response to the presence of trimethylamine-*N*-oxide (TMAO) in the environment, the TorS sensor kinase autophosphorylates and transfers a phosphoryl group to TorR through a four-step phosphorelay leading to the induction of the *torCAD* operon, which encodes the TMAO reductase respiratory system (10). The complex phosphorelay occurring between TorS and TorR led us to suspect the presence of intermediate checkpoints in the TMAO signal transduction pathway. Indeed, the *torI* gene has been identified as a negative regulator of the *torCAD* operon using a genetic multicopy approach. The negative effect was due to a previously unidentified small open reading frame (66 amino acids) that we called *torI* for *Tor i*nhibition. Further studies showed that TorI does not interfere with the phosphorelay but rather acts at the level of the TorR response regulator. Interestingly, we showed that TorI binds to the C-terminal domain of TorR without affecting its DNA binding capacity, and we proposed that TorI prevents the recruitment of RNA polymerase to the *torC* promoter (7). So far, TorI is a unique case of response regulator inhibitor acting through protein-protein interaction with the DNA binding domain of a response regulator without interfering with its DNA binding activity.

A look for TorI homologues on finished and unfinished bacterial genomes led us to find two categories of homologous proteins. The first one contained proteins that show 100% identity with TorI. These proteins are the products of gene *hkaC* in the coliphage HK620 genome and gene 18 in the genome of the *Shigella flexneri* phage Sf6 (11, 12). To date, no biological function has been assigned to these predicted proteins. In the second category of homologous proteins were found several proteins with 25–35% identity to TorI, present in various pathogenic bacteria such as uropathogenic *E. coli* O157:H7, *Yersinia pseudotuberculosis*, *Vibrio cholerae*, and *S. flexneri.* Analysis of the DNA sequences surrounding the genes of the TorI homologues revealed that most of these genes are located near a phage integrase encoding gene, in a pathogenic island, or in a characterized prophage region. The analysis of the genetic context of *torI* indicated that it also belonged to the defective prophage KplE1 genome sequence (7, 13). All of these proteins share a small size (less than 80 residues) and a high proportion of basic residues, which are typical characteristics of RDF proteins (2). Indeed, two homologues of TorI have been recently described as pathogenic island excisionases, namely Hef and Rox in *Y. pseudotuberculosis*, and *S. flexneri*, respectively (14, 15). All of these data led us to suspect a role for TorI in the excision of the KplE1 cryptic prophage in *E. coli*. In this study, we show that TorI is capable of excisionase activity *in vivo* and presents a three-dimensional fold highly similar to that of both $^{\lambda}$Xis and Mu repressor proteins. We also define the structural features of a new family of atypical excisionases.

* This work was supported by the CNRS and the Université de la Méditerranée. 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.

The atomic coordinates and structure factors (code 1Z4H) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed. Tel.: 33-491-164-647; Fax: 33-491-164-578; E-mail: morelli@ibsm.cnrs-mrs.fr.

[3] The abbreviations used are: IHF, integration host factor; RDF, recombination directionality factor; TMAO, trimethylamine-*N*-oxide; IPTG, isopropyl β-D-thiogalactopyranoside; HSQC, heteronuclear single quantum correlation; NOE, nuclear Overhauser effect; CSI, chemical shift index; r.m.s.d., root mean square deviation.

MATERIALS AND METHODS

DNA Manipulations—Small-scale plasmid extractions were carried out by using the Miniprep plasmid kit (Promega), and DNA fragments were purified with the Qiagen PCR purification kit (Qiagen Inc.). DNA sequencing was performed on purified plasmids and PCR products at MWG Biotech.

Strain Construction—Strain LCB970 is a derivative of strain MC4100 (Casadaban) and was constructed by insertion of the chloramphenicol acetyl transferase (*cat*) gene in the KplE1 prophage between *yfdO* and *yfdP* according to the method of Datsenko and Wanner (16). Briefly, the *cat* gene was PCR-amplified using pKD3 as a template with the following primers: KplE1-Cm1 (5′-CAGGCGAATTTCGTTTGCCCAGGCTGTCCAGTTCGGTTCTGTGTAGGCTGGAGCTGCTTC) and KplE1-Cm2 (5′-AGCAGGCCGCCGAATGTGACGGCGAGGTGGTTCGTCCCAACATATGAATATCCTCCTTAG), where the underlined sequences are homologous to the DNA sequence within the KplE1 prophage to allow site-specific recombination by the λ-Red recombination system. The resulting PCR product was then transformed into a strain containing pKD20, which encodes the Red recombinase under an arabinose inducible promoter. The presence of the *cat* gene was then verified by PCR amplification of the chromosomal region (the sequence of the primers is available upon request to the authors).

Excision Test—Strain LCB970 carrying *torI* encoding plasmids pJFi (7) was grown in LB medium until the OD_{600} reached 0.5 units, and IPTG (1 mM) was added for 2 h at 37 °C under agitation. Culture dilutions were prepared and plated onto rich medium containing either 50 μg/ml ampicillin or 5 μg/ml chloramphenicol. Numeration of the colonies plated on both antibiotics was performed and the ratio of ampicillin-resistant/chloramphenicol-resistant colonies was calculated. Values represent the average of at least three independent determinations. To confirm prophage DNA excision, a PCR test was performed on randomly chosen colonies plated onto ampicillin. A control colony containing the empty vector was included in the PCR assay. A basic PCR amplification was performed (30 s denaturation at 94 °C, 30 s annealing at 55 °C, 30 s elongation at 72 °C) using the GoTaq® DNA polymerase (Promega) and the following primers: ptorI1 (5′-GAGCCATACAGCCTCACACTCGATGAGG) inside KplE1 prophage and Ext3′KplE1 (5′-CTTATTCGGCCTGCTAGTTCG) outside of the prophage.

Site-directed Mutagenesis—Mutagenesis of residues Tyr^{28} and Arg^{45} was performed as described previously (17). Briefly, the entire pJFi plasmid (pJF119EH*torI*) was PCR-amplified with divergent overlapping primer pairs, one primer of each pair carrying the desired mutation, and a high fidelity thermostable DNA polymerase (Expand High Fidelity DNA, Roche Diagnostic). Mutations of Tyr^{28} into Phe or Ser and Arg^{45} into Gln or Lys were performed by introducing degenerate codons T*Y*C and MAA (where *Y* indicates A or C, and M indicates C or T), respectively, at the desired positions in one primer per pair. PCR reaction was conducted with 200 ng of template plasmid, 200 μM concentration of each dNTPs, and 5 units of DNA polymerase, in the presence of 5% Me_2SO, and only 10 amplification cycles were performed to avoid nondesired mutation. The template plasmid was then hydrolyzed by addition of 20 units of the DpnI enzyme (Biolabs), and the purified PCR products were directly transformed into a *recA* strain (JM109) to allow homologous recombination on both sides of the PCR product to generate a mutated circular plasmid. The resultant plasmids were then purified and sequenced (MWG Biotech) to check for the presence of the desired mutations and the absence of additional mutation in the rest of the *torI* gene and transformed into the test strain LCB970. To check that TorI mutants were produced and stable in strain LCB970, crude extracts of LCB970 overproducing TorI wild type and mutants were prepared and submitted to Western blot analysis using a TorI antiserum.

TABLE ONE

Structural statistics based on the 17 best structures calculated with ARIA

r.m.s.d. (Å) with respect to the mean	
Backbone, 2° structures	0.30 (±5.85E-02)
Heavy atoms, 2° structures	0.67 (±8.26E-02)
Backbone, all residues (6–63)	0.69 (±0.25)
Heavy atoms, all residues (6–63)	1.18 (±0.24)
No. of experimental restraints	
Total number of meaningful distance restraints	1173
Intraresidual ($i = j$)	505
Sequential ($\|i - j\| = 1$)	296
Medium range ($1 < \|i - j\| \leq 4$)	184
Long range ($\|i - j\| > 4$)	188
Average number of restraints per residues	17.8
Dihedral angle restraints from TALOS	68
ϕ angles from $^3J_{HN\text{-}H\alpha}$	26
CSI-derived angles	74
Restraint violations	
NOE distances with violations >0.5	0
Dihedral with violations >5°	0
r.m.s.d for experimental restraints	
All distances restraints (1173)	0.027 (±1E-03))
Dihedral restraints (TALOS)	0.470 (±0.22)
Dihedral restraints (CSI)	0.645 (±0.102)
r.m.s.d (Å) from idealized geometry	
Bonds	0.00400 (±1E-04)
Angles	0.557406 (±1.08E-02)
Improper	1.39144 (±0.08416)
Ramachandran analysis (residues 1–66)	
Residues in favored regions (%)	81.9
Residues in additional allowed regions (%)	14.5
Residues in generously allowed regions (%)	2.6
Residues in disallowed regions (%)	1.0

TorI Labeling and Purification—TorI protein was produced in BL21(DE3) cells harboring plasmid pETsI (7). To obtain the double-labeled protein, cells were grown in M9 minimal medium supplemented with 2 g/liter [^{13}C]glucose (Eurisotop) and/or 1 g/liter [^{15}N]NH_4Cl until the OD_{600} reached 0.8 units, and IPTG (1 mM) was added for 2 h at 37 °C. French-pressed cell lysate extract was equilibrated with 40 mM Tris buffer (pH 7.4) and loaded onto a HiTrap SP column (Amersham Biosciences). The protein was eluted with a step gradient of KCl and was found in the 0.3 M KCl-containing fraction.

NMR Measurements, Assignment, and Interaction with DNA—The HSQC spectra were first compared at 278, 283, 288, and 298 K to pick up the best stability/intensity ratio. The NMR experiments were then carried out at 278K on a 500 MHz Bruker DRX spectrometer and on a 800 MHz Varian Inova spectrometer equipped with a triple resonance (^{1}H, ^{15}N, ^{13}C) probe including shielded z-gradients. For the backbone and side chain resonance assignments, three-dimensional HNCO, CBCA(CO)NH, HNCA, HN(CO)CA, HN(CA)CO, (H)CCH-TOCSY, NOESY-(^{15}N,^{1}H)-HSQC, and TOCSY-(^{15}N,^{1}H)-HSQC spectra were recorded (18–21). $^3J_{HN\alpha}$ coupling constants were measured using three-dimensional HNHA experiments (22). NOE distance constraints were obtained from a two-dimensinoal NOESY, three-dimensional NOESY (^{15}N,^{1}H)-HSQC spectra with mixing times of 100 and 120 ms.

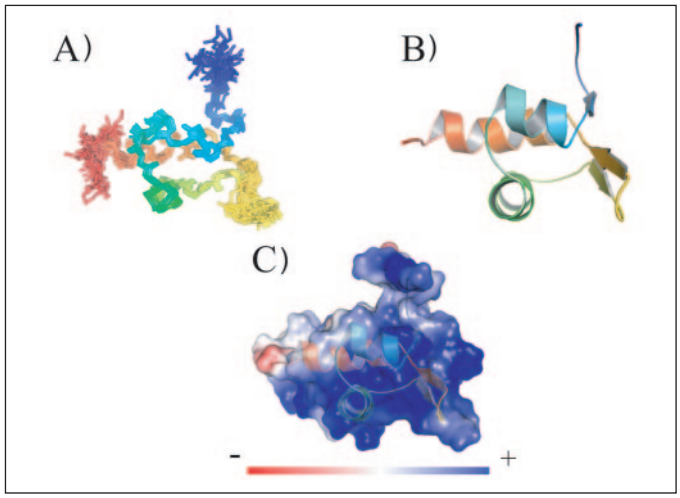

FIGURE 1. **NMR solution structure of TorI.** *A*, backbone view of the NMR ensemble (17 structures). The protein is colored from N-terminal (*blue*) to C-terminal (*red*); *B*, ribbon view of a representative TorI structure (closest to average) for residues 1–66; *C*, electrostatic surface potential of the TorI protein overlaid with the closest to average structure in ribbon view. Electronegative (acidic) regions are colored *red*, electropositive (basic) regions are colored *blue*. The structures are displayed using the molecular graphics program PyMOL (39).

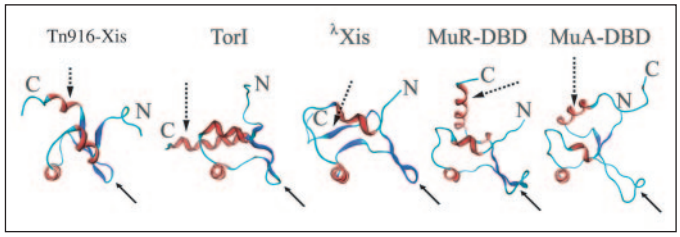

FIGURE 2. **Structural relationships between TorI and Tn916-Xis, λXis, MuR, and MuA.** From *left* to *right*, ribbon diagrams of Tn916-Xis, of TorI, of the excisionase from bacteriophage λ (λXis) and of the internal activation sequence binding domains from the Mu repressor (MuR-DBD) and Mu transposase (MuA-DBD) proteins. All proteins contain structurally conserved α-helices (colored *red*), which are packed against a β-sheet (colored *blue*). TorI possesses features in common with each protein family. *Dashed arrows* show the secondary structure present in the C terminus of each protein (α-helix for TorI and Tn916, MuR-DBD, and MuA-DBD, while λXis possesses a β-sheet). *Plain arrows* pinpoint the size of the wing (small for TorI, Tn916, and λXis, large for MuR-DBD and MuA-DBD). The structures are displayed using the molecular graphics program MOE (www.chemcomp.com).

All NMR spectra were processed using XWin-NMR (Bruker) and analyzed using Felix (Accelrys). The DNA sequence used for the titration experiments corresponds to 5′-GGGTAAAATA (Fig. 3). Two 10-base complementary oligonucleotides (MWG Biotech) were annealed in 10 mM Tris-HCl, 300 mM NaCl, and ethanol-precipitated. The double strand DNA was then resuspended in 40 mM phosphate (pH 5.9), 100 mM NaCl at a concentration of 0.1 M. The TorI-DNA complex formation was monitored by recording a series of two-dimensional ^{1}H-^{15}N HSQC spectra of a 125 mM ^{15}N-labeled TorI solution (40 mM phosphate (pH 5.9), 50 mM NaCl) with final DNA concentrations of 15 and 30 mM (0.1 and 0.2 equivalents).

Chemical Shift-derived Restraints—The Cα, Cβ, Co, Hα, and N chemical shifts of 36 residues served as input for the TALOS program (23). TALOS derives information on the ϕ and ψ backbone dihedral angles from a comparison of secondary chemical shift corresponding to known conformations. A conservative approach was chosen requiring that all 10 best matches agree for a prediction to be accepted. The TALOS predictions were converted into dihedral angle restraints as the average ϕ and ψ angles ± 2 S.D. or a minimum of ± 10°.

Structure Calculations—Distance restraints were obtained from two-dimensional NOESY and three-dimensional ^{15}N-edited NOESY experiments (with mixing times of 100 and 120 ms). ϕ angle dihedral restraints were estimated from $^{3}J_{HN\text{-}H\alpha}$ coupling constants and use of the Karplus equation. TALOS-derived dihedral angle constraints were used as described above. Structure calculations were performed with CNS/ARIA (default parameter with the Amersham Bioscience steps parameter *8 to increase convergence). The best 17 structures were selected based on the lowest total restraint energy and analyzed using Procheck and Procheck_NMR (24).

Coordinates—The Protein Data Bank accession number for the coordinates is 1Z4H.

RESULTS AND DISCUSSION

The TorI Response Regulator Inhibitor Is a Winged Helix Protein—To solve the three-dimensional structure of TorI we produced the recombinant protein in the presence of either [^{13}C]glucose or [^{15}N]NH_4Cl or both labeled substrates. After recording a series of HSQC spectra at different temperatures, we finally picked up 278 K as the best stability/

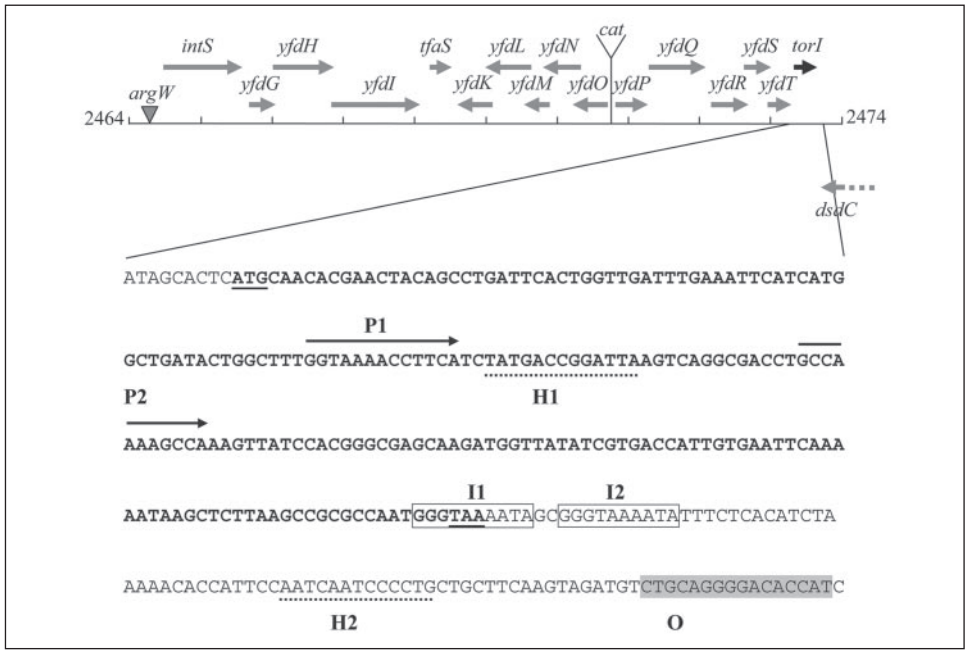

FIGURE 3. **Representation of the KplE1 prophage.** The 16 Open reading frames of the KplE1 prophage are mentioned, with the *arrows* indicating the direction of transcription. The insertion site of the *cat* gene in strain LCB970 lies between *yfdO* and *yfdP*. The nucleotide sequence of the attR site (3′ end) of the prophage is indicated below: the *torI* gene is in boldface characters, with the ATG and stop codons *underlined*. The characteristic sites for the intasome formation are indicated as follows: core sequence (*O*), gray background; side-arm integrase binding sites (*P1*, *P2*), *plain arrows*; IHF binding sites (*H1*, *H2*), *dashed line*; TorI binding sites (*I1*, *I2*), *boxed sequence*.

intensity ratio. The structure of TorI was then solved at 278 K (40 mM $NaPO_4$ (pH 5.9)) using conventional homonuclear two- and multidimensional heteronuclear NMR spectroscopy on a 500 MHz Bruker DRX spectrometer and at high resolution field (800 MHz Varian Inova spectrometer), making use of uniformly ^{15}N- and ^{13}C-labeled protein. The NMR assignment of ^{1}H, ^{15}N, and ^{13}C resonances was accomplished using a combination of ^{1}H-^{1}H TOCSY, ^{1}H-^{1}H NOESY, ^{1}H-^{15}N TOCSY-HSQC, ^{1}H-^{15}N NOESY-HSQC, and ^{1}H-^{15}N HNHA experiments and the standard multidimensional heteronuclear experiments HNCO, HNCA, HN(CO)CA, HN(CA)CO, and CBCA(CO)NH (18–21). The global fold was established using unambiguous assigned long range NOEs. Several long range NOEs between the helices and the β-strands were obtained from a three-dimensional NOESY (^{1}H-^{15}N)-HSQC spectrum. The unassigned NOEs with multiple possible assignments were used in ARIA (25) as ambiguous restraints. The final ensemble of structures was calculated using non-bonded interaction for simulated annealing and refinement of the final structures in an explicit water box. A total of 1341 restraints have been used to calculate the structures. 1173 distance restraints were identified from the two-dimensional and three-dimensional NOESY; 68 dihedral angle constraints were obtained from TALOS (23) on the basis of backbone chemical shift values; 100 dihedral angle constraints were estimated from the chemical shift index (CSI), $^{3}J_{HN\text{-}H\alpha}$ coupling constant, and from the use of the Karplus equation (26). TABLE ONE summarizes the experimental restraints and the structural statistics of the 17 best structures. A superimposition of the final ensemble of 17 simulated annealing structures is shown in Fig. 1*A*; a view of the representative (closest to average) structure is shown in Fig. 1*B*, and the electrostatic surface potential of the TorI protein overlaid with the ribbon diagram is presented in Fig. 1*C*.

The conformers within the ensemble do not exhibit any NOE, dihedral angle, or scalar coupling constant violations greater than 0.4 Å, 5°, or 2 Hz, respectively. Residues Gln^{6}–Arg^{63} are well structured in solution and the coordinates of their backbone and heavy atoms can be superimposed to the average structure with a root mean square deviation (r.m.s.d.) of 0.69 (±0.25) Å and 1.18 (±0.24) Å, respectively.

TABLE TWO

Excision activity of TorI and mutants

Excision test have been performed in strain LCB970 containing pJF119EH derivatives as described under "Materials and Methods." pJFi encodes the wild-type TorI protein, and the mutations of the variants are indicated. All activities were measured in the presence of 1 mM IPTG as the inductor for torI gene expression. Excision percentages of the mutants are calculated using the TorI wild-type activity as the reference (100%).

	Ap^{R}/Cm^{R}	Excision %
pJFi (ptac-torI)	351	100
pJF119EH (empty vector)	1	0
pJFi Y28F	1.2	0.06
pJFi Y28S	1.4	0.11
pJFi R45Q	25	6.8
pJFi R45K	35	9.7

The TorI protein adopts an overall structure classified as an unusual "winged" helix structure that is formed by three α-helices (α1, residues Lys^{14}–Asp^{19}; α2, Lys^{24}–Lys^{32}; α3, Arg^{51}–Arg^{63}), which are packed against three extended strands (Fig. 1*B*). The first two helices are separated by an ordered 5-residue turn (T1, Asp^{19}–Gly^{23}) and are positioned approximately orthogonal to one another to generate a L-shaped structure. Immediately following helix α2, the peptide chain adopts an extended strand conformation (S1, residues Ser^{33}–Lys^{38}), before leading into a two-stranded anti-parallel β-sheet (β2, residues Ala^{39}–Val^{41}; β3, residues Ala^{46}–$Tryp^{48}$) whose strands are connected by a 4-residue reverse turn ("wing") (W, residues Ile^{42}–Arg^{45}). Strand S2 (S2, residues Leu^{49}–Asp^{52}) then follows strand β2 and is situated directly adjacent to strand S1. The structure is completed by the first β-strand (β1, residues Asp^{8}–Val^{11}), which packs against β2, limiting the potential mobility of the wing. Finally, the well defined third helix is parallel to α-helix 1 (opposite orientation) and positioned approximately orthogonal to α-helix 2 generating a final C-shaped structure between the three α-helices.

The winged helix structural motif is a derivative of the usual helix-turn-helix DNA binding motif (27–30). It is known for some winged helix proteins that DNA binding occurs through two kinds of interactions: one α-helix interacts with the major groove of the DNA target, while the wing penetrates the minor groove of the DNA target. This wing can be seen as an anchor that increases the specificity of the interaction. In the TorI structure the winged helix is composed of helix α2 and of the wing in between β2 and β3. This motif thus constitutes a potential DNA binding site on the TorI protein.

Structural Similarity of TorI with Other Proteins—Interestingly a search for homologous structure using the programs DALI (www.ebi.ac.uk/dali/) and SSM (www.ebi.ac.uk/msd-srv/ssm) allowed us to identify prophage excisionase-type molecules of the $^{\lambda}$Xis family (r.m.s.d. of 2.63 Å with a Qscore of 0.26 with the structure of full-length excisionase from bacteriophage HK022, Protein Data Bank code 1pm6 and r.m.s.d. of 2.7 Å with a Z-score of 2.5 with the structure of the $^{\lambda}$Xis protein, Protein Data Bank code 1lx8) (27, 28) (Fig. 2). These proteins belong to the RDF family of proteins that comprises a diverse group of proteins involved in controlling the directionality of integrase-mediated site-specific recombination (2). In lambdoid phages the $^{\lambda}$Xis proteins are the essential partners of the tyrosine family of site-specific recombinases (Int) for the excision of the prophage DNA during the lytic cycle. Indeed, the integrase plays a key role in both integration and excision reactions together with the host factor IHF, but the directionality of recombination is driven by the excisionase that interacts with the integrase and DNA during prophage excision (3, 31). The three-dimensional conformation of $^{\lambda}$Xis has been described to be similar to the DNA binding domain of the Mu bacteriophage repressor (MuR-DBD), to the DNA binding domain of the Mu transposase (MuA-DBD) protein and recently to the excisionase protein from the conjugative transposon Tn916 (32, 33) (see below for discussion).

Identification of TorI DNA Target and in Vivo Excisionase Activity—We previously found that the *torI* gene was actually part of a cryptic prophage genome, the prophage KplE1 (or CPS53) that extends

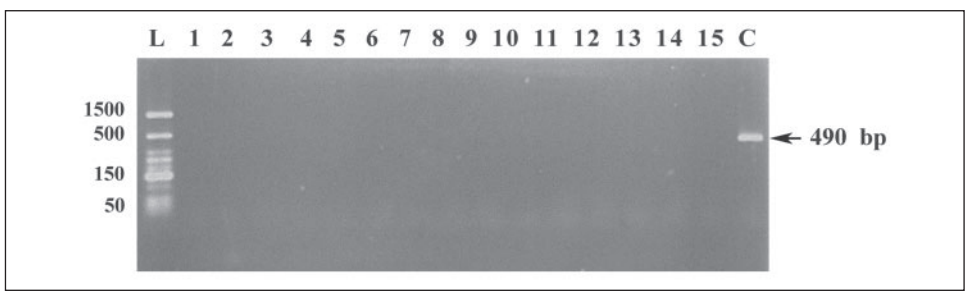

FIGURE 4. **Effect of TorI overexpression on the excision of the KplE1 prophage DNA.** PCR amplifications were performed on randomly chosen colonies grown on rich medium overproducing (*lanes 1–15*) or not (*lane C*) the TorI protein. *L*, ladder.

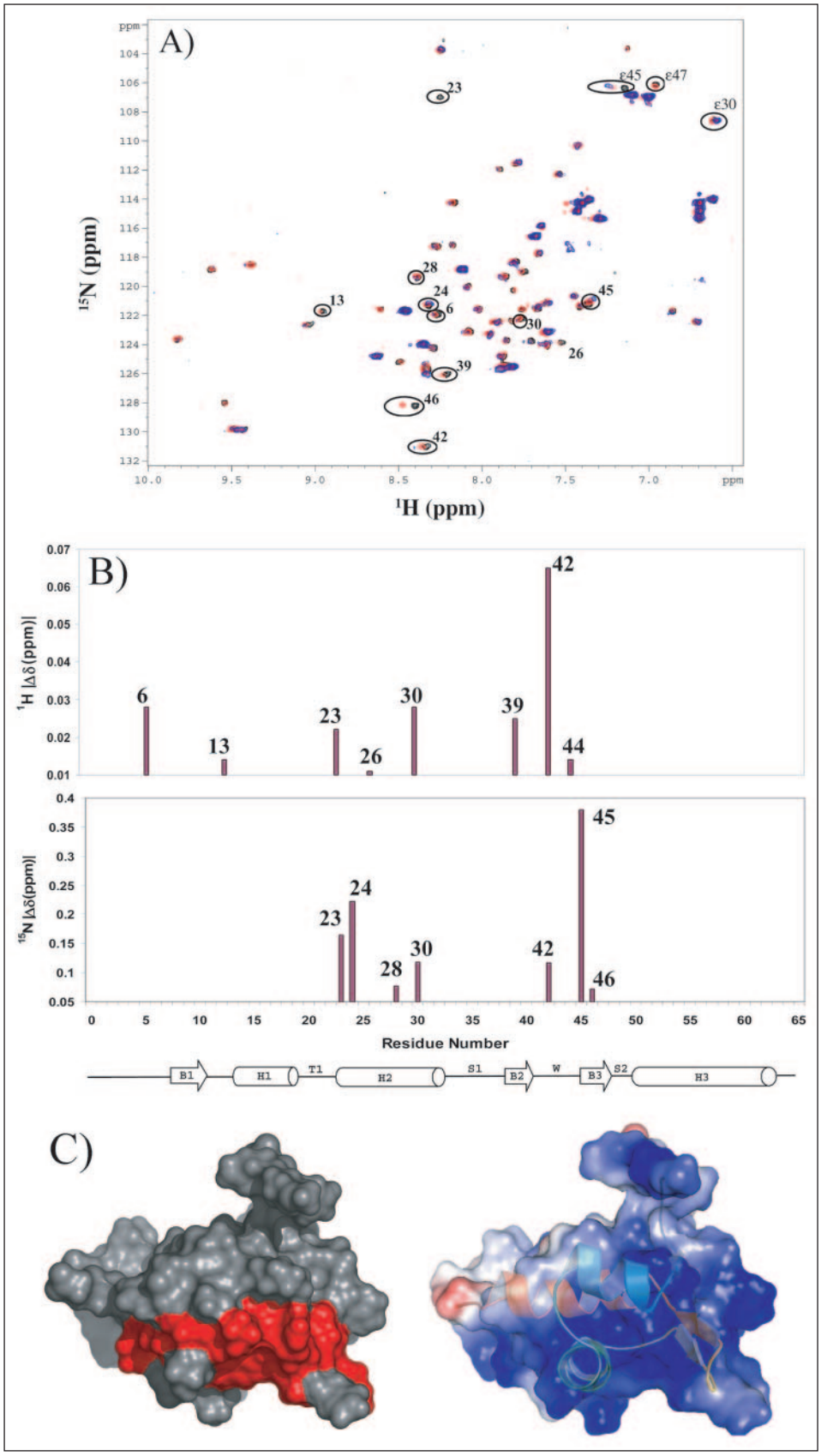

FIGURE 5. **Mapping TorI-DNA interaction site by heteronuclear NMR.** *A*, overlaid contour plots of the ^{1}H-^{15}N HSQC spectra of the free protein (*black*) and in presence of 0.1 equivalent of DNA (*red*) and 0.2 equivalent of DNA (*blue*). The cross-peaks showing chemical shift variations upon addition of DNA are *circled* and labeled with their amino acid numbers. *B*, plot of the TorI backbone amide proton (NH) chemical shift variations upon titration of DNA. The cross-peaks in the presence of 0.1 equivalent of DNA were chosen when the corresponding cross-peaks in the presence of 0.2 equivalent of DNA were not available. The secondary structure is showed at the *bottom* of the figure to pinpoint that mainly α-helix 2 and the wing are affected by the presence of the DNA. *C*, *left*: surface representation of the unaffected residues (*gray*) and affected residues (6, 13, 23, 24, 26, 28, 30, 39, 42, 44, 45, 46) upon DNA binding. *Right*, electrostatic surface potential of TorI overlaid with the backbone structure in ribbon view.

from the *argW* tRNA gene to the *dsdC* gene (around 2475 kb on the *E. coli* MG1655 chromosome) (7, 13). The *torI* gene is located at the 3′ end of this prophage, whereas the *intS* gene, encoding a putative tyrosine integrase, is located at the other extremity (Fig. 3). The KplE1 cryptic prophage was identified by DNA sequence homology to *S. flexneri* bacteriophages Sf6 and V and coliphage HK620, all of them being λ-type temper-

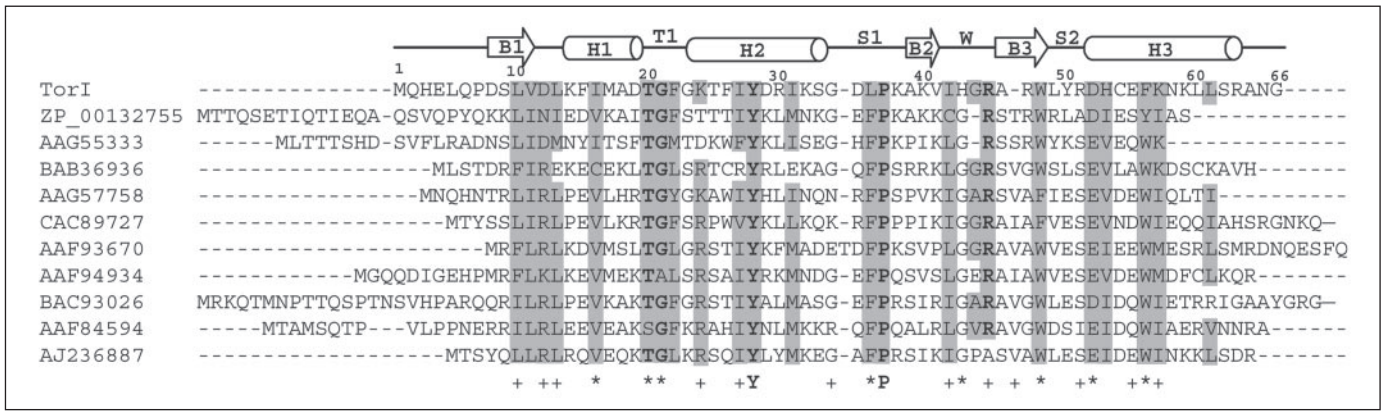

FIGURE 6. **TorI homologue alignment related to the TorI structure.** The secondary structure elements of TorI are presented on *top* of the alignment. TorI homologues retrieved by tBLASTn search and showing 25–35% identity with TorI using a Fasta local alignment (7) are presented. Conserved residues are indicated according to the percentage of representation as follow: >60%, +; >80%, *; 100%, the actual residue letter. Important conserved residues are shown in *bold*. Similar residues are highlighted in *gray*. The proteins used for the alignment are: TorI, protein of *E. coli* K-12; ZP_00132755 of *Haemophilus somnus*; AAG55333, AAG57758, and BAB36936 of *E. coli* O157:H7; CAC89727 of *Yersinia pestis*; AAF93670 and AAF94934 of *Vibrio cholerae*; BAC93026 of *Vibrio vulnificus*, AAF84594 of *Xylella fastidiosa*; AJ236887 (Rox) of *Y. pseudotuberculosis*.

ate phages. We thus looked for the characteristic sequences involved in prophage excision at the 3′ end of the KplE1 prophage (*attR*). As expected, we could predict all of the sequences necessary for the intasome structure formation (Fig. 3). This region indeed comprises the integrase binding sites, composed of a core sequence and two arm-type sequences (O, P1 and P2, respectively), two IHF binding sites (H1, H2), and two perfect repeats of a 10-bp motif that are good candidates for being TorI binding sites (I1, I2).

To demonstrate the *in vivo* activity of TorI as a prophage excisionase we constructed a test strain in which the chloramphenicol acetyltransferase encoding gene was inserted in a non-coding region of the prophage KplE1. We then checked for the effect of TorI overexpression on KplE1 excision by estimating the number of bacteria (colony-forming units) that lost the ability to grow on chloramphenicol. Upon IPTG induction of the *torI* gene, most of the colonies proved to be chloramphenicol sensitive compared with the control condition in the presence of the vector alone (TABLE TWO). PCR amplifications of the 3′ end of the KplE1 DNA were then performed on a sample of randomly chosen colonies plated on rich medium to confirm the disappearance of the KplE1 prophage DNA from the bacterial chromosome upon TorI overexpression. As expected, no amplification product could be obtained in cells overproducing TorI (Fig. 4, *lanes 1–15*), whereas a PCR product at the expected size (490 bp) was obtained in control cells containing the vector alone (Fig. 4, *lane C*). These results show (i) that the KplE1-defective prophage can be excised and (ii) that TorI overexpression promotes KplE1 prophage excision and allows the *in vivo* validation of the structural homology of TorI with λ-type excisionases. Based on similarity with the $^{\lambda}$Xis protein and identification of the winged helix DNA binding motif identified in the three-dimensional structure, we anticipate that TorI should also interact with a specific DNA target in the *attR* region of the prophage.

TorI Residues Involved in DNA Binding Activity—To investigate the molecular basis of TorI interaction with DNA and to locate the residues important for DNA binding, we performed a titration assay of TorI by adding increasing amounts of the putative TorI DNA target, using NMR chemical shift perturbations. This method consists of recording the two-dimensional ^{1}H-^{15}N HSQC spectrum of TorI upon successive additions of the 10-bp putative target DNA (5′-GGGTAAAATA) (Fig. 3). This method detects residues that interact directly with DNA or that are indirectly affected by its binding (Fig. 5). Upon DNA binding, which leads to modifications of the chemical and/or magnetic environments, the resonances of the unbound TorI underwent chemical shift changes (Fig. 5*B*). The free and DNA-bound proteins are in fast exchange on the NMR time scale. The largest effects in the ^{1}H and ^{15}N resonances occurred in the N terminus, the first turn, the helix 2, and the wing (reverse turn connecting strands β2 and β3). More precisely, the largest chemical shifts changes were detected for Gly^{23}/Lys^{24} (end of the first turn) and Arg^{45} (in the wing) whereas none of the C-terminal residues of TorI were affected by DNA binding (Fig. 5, *B* and *C*). Remarkably, these residues are located in the HTH motif and in the wing, which can thus be defined as the TorI DNA binding motif.

To confirm this model of interaction we decided to change by site-directed mutagenesis two highly conserved residues (Tyr^{28} and Arg^{45}) that underwent chemical shifts upon DNA binding (Figs. 5 and 6). We first checked that all TorI mutants were produced and stable in strain LCB970 by Western blot analysis using a whole TorI antiserum (data not shown). As indicated in TABLE TWO, both substitutions of Tyr^{28} with either Ser or Phe completely abolished the excisionase activity of TorI. On the other hand, the TorI proteins where Arg^{45} was mutated to either Gln or Lys, were still able to promote some excision of the prophage, although the activity was strongly impaired by these mutations (about 7 to 10% of the wild-type activity). These results thus indicate that Tyr^{28} in helix α2 and Arg^{45} in the wing are involved in TorI excisionase activity. However, Tyr^{28} seems absolutely required as suggested by the phenotype observed with both conservative mutations, whereas Arg^{45} when changed to Gln or Lys retains some excisionase activity, suggesting that DNA binding through the wing is less stringent than through helix α2.

TorI, the Structural Missing Link between $^{\lambda}$Xis and MuR Proteins?—According to its three-dimensional structure and to the mutagenesis study, TorI is a prokaryotic winged helix protein, which contains a HTH motif (helix α1-turn1-helix α2) and a loop (wing) contacting the DNA. Moreover, TorI is closely related to the structure of $^{\lambda}$Xis and to the excisionase of the conjugative transposon Tn916 (32, 34) but also to the DNA binding domains of the Mu bacteriophage repressor (MuR-DBD) and transposase (MuA-DBD) proteins (30, 33) (Fig. 2). The MuA and Tn916 transposases are similar to the recombination directionality factors in functioning to modulate the efficiency of the transposition of Mu bacteriophage and Tn916, respectively, whereas the MuR protein establishes lysogeny by shutting down phage transposition functions and by competing with the transposase in the operator region (35).

The structures of the excisionase-DNA complex of $^{\lambda}$Xis and of the Mu-repressor-DNA complex have been determined (28, 36). These two complexes both reveal interactions between the helix α2 and the wing of the proteins and the bound target DNA. In the $^{\lambda}$Xis complex structure the helix α2 (equivalent to the TorI helix 2) is inserted into the major groove, while the wing contacts the adjacent minor groove. This model is consistent with several mutagenesis studies where mutants with

amino acid substitutions within the helix α2 show a decreased excisionase activity *in vivo* and are defective in DNA binding (28, 37). Moreover, NMR chemical shift mapping data on MuR with DNA (30) showed that the largest deviations in the ^{1}H and ^{15}N shifts occur in the helix-turn-helix unit (H1-T-H2) and the flexible loop between strands β2 and β3 (wing). Interestingly, the most drastic ^{1}H and ^{15}N chemical shift changes occur in the amide proton of Gly^{30} (in turn T) and the amide nitrogen of Ala^{57} (in the wing connecting strands β2 and β3), respectively. These two residues correspond to the Gly^{23} (turn T1) and Arg^{45} (wing) of TorI that were strongly affected by the binding of DNA (Fig. 5*B*). Therefore the similar fold between these proteins (TorI, $^{\lambda}$Xis, and MuR/MuA) seems to be intricately correlated to their DNA binding properties.

Despite these convergent structural similarities, the TorI protein presents new structural features compared with the excisionase $^{\lambda}$Xis. The major difference is located at the C-terminal end of TorI where 12 residues form a well structured alpha helix (residues Tyr^{51}–Ser^{63}) (Fig. 2), whereas the $^{\lambda}$Xis protein contains a β-sheet at its C terminus that was shown to interact with the integrase (3, 38). In contrast to $^{\lambda}$Xis, the Mu repressor protein harbors a C-terminal helix similar to TorI, which is also comprised of 12 residues (residues Thr^{67}–Gly^{79}) (30). These data suggest that TorI is more related to the Mu repressor than to the excisionase $^{\lambda}$Xis. However, the Mu repressor possesses a large wing (residues Glu^{50}–Lys^{56}) compared with TorI and $^{\lambda}$Xis (residues Ile^{42}–Arg^{45} and Asp^{37}–Glu^{40}, respectively). The size of the wing has been described to be intrinsically correlated to its dynamic properties. Indeed, the authors have shown that the large wing of MuR undergoes a major structural rearrangement upon DNA binding (36). In fact, in the absence of DNA, the MuR wing is mostly disordered, but upon DNA binding, insertion into the minor groove severely dampens rapid motions within the wing. In the case of $^{\lambda}$Xis the short wing is ordered in the absence of DNA (34). By this way the $^{\lambda}$Xis protein avoids the necessary entropic cost to block the wing to save the energy needed for DNA distortion (28). Through the structural similarities of $^{\lambda}$Xis and MuR, the authors suggested that the unrelated enteric lambda and Mu bacteriophages RDF proteins could be derived from a common ancestor (28) that has yet to be defined. Since TorI possess a short wing similar to that of $^{\lambda}$Xis and a well defined C-terminal helix as found in MuR, we propose that TorI represents the structural missing link between the $^{\lambda}$Xis and MuR proteins.

A Family of Atypical Excisionases—The TorI protein has been identified as a response regulator inhibitor that binds to the TorR response regulator (7). Moreover, this study shows that TorI has an excisionase activity even if no relevant primary sequence homology with $^{\lambda}$Xis was observed. Interestingly, the closest homologues of TorI are of phage origin and belong to different pathogenic species such as *V. cholerae*, *Y. pseudotuberculosis*, or *E. coli* O157:H7 species. The sequence alignment of TorI with its homologues (Fig. 6) shows the conservation of several patches of residues. Among them, Tyr^{28} (in helix 2) of TorI is strictly conserved throughout this family. Thr^{20} and Gly^{21} are also highly conserved, and these two residues constitute the turn of the helix-turn-helix motif of TorI. Pro^{37} is also strictly conserved and might be involved in the optimization of the extended structure of the polypeptide chain just before the β2-strand. The other conserved residues in this family, His^{43} and Arg^{45}, are part of the wing. As shown above both Tyr^{28} and Arg^{45} proved to be crucial for the excisionase activity of TorI. According to these sequence similarities we thus predict that all of these proteins share the same secondary structures and fold, as well as the same function, which is consistent with the *in vivo* excisionase role of at least three of them, TorI, Hef, and Rox in *E. coli*, *Y. pseudotuberculosis*, and *S. flexneri*, respectively (14, 15). Since TorI is also involved in gene regulation, we propose that this particular family of RDF proteins may be also involved in regulatory processes. The TorI structure thus provides a structural basis for the understanding of the dual role of this family of excisionases.

Acknowledgments—We thank the members of our respective laboratories for helpful discussions, Dr. Olivier Bornet for his technical assistance and advice regarding pulse sequence implementation, Bernard Chetrit for his computer assistance with the grid clustering, Laurence Théraulaz for strain construction, and finally Dr. Marianne Grant for editing the manuscript. Some of the NMR experiments were recorded on the Inova Varian 800 MHz at the National Scale Facility in Grenoble, France.

REFERENCES

1. Friedman, D. I. (1992) *Curr. Opin. Genet. Dev.* **2,** 727–738
2. Lewis, J. A., and Hatfull, G. F. (2001) *Nucleic Acids Res.* **29,** 2205–2216
3. Cho, E. H., Gumport, R. I., and Gardner, J. F. (2002) *J. Bacteriol.* **184,** 5200–5203
4. de Vargas, L. M., and Landy, A. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 588–592
5. Thompson, J. F., de Vargas, L. M., Skinner, S. E., and Landy, A. (1987) *J. Mol. Biol.* **195,** 481–493
6. Numrych, T. E., Gumport, R. I., and Gardner, J. F. (1991) *J. Bacteriol.* **173,** 5954–5963
7. Ansaldi, M., Théraulaz, L., and Méjean, V. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 9423–9428
8. Simon, G., Méjean, V., Jourlin, C., Chippaux, M., and Pascal, M. C. (1994) *J. Bacteriol.* **176,** 5601–5606
9. Simon, G., Jourlin, C., Ansaldi, M., Pascal, M. C., Chippaux, M., and Méjean, V. (1995) *Mol. Microbiol.* **17,** 971–980
10. Jourlin, C., Ansaldi, M., and Méjean, V. (1997) *J. Mol. Biol.* **267,** 770–777
11. Clark, A. J., Inwood, W., Cloutier, T., and Dhillon, T. S. (2001) *J. Mol. Biol.* **311,** 657–679
12. Casjens, S., Winn-Stapley, D. A., Gilcrease, E. B., Morona, R., Kuhlewein, C., Chua, J. E., Manning, P. A., Inwood, W., and Clark, A. J. (2004) *J. Mol. Biol.* **339,** 379–394
13. Rudd, K. E. (1999) *Res. Microbiol.* **150,** 653–664
14. Luck, S. N., Turner, S. A., Rajakumar, K., Adler, B., and Sakellaris, H. (2004) *J. Bacteriol.* **186,** 5551–5554
15. Lesic, B., Bach, S., Ghigo, J. M., Dobrindt, U., Hacker, J., and Carniel, E. (2004) *Mol. Microbiol.* **52,** 1337–1348
16. Datsenko, K. A., and Wanner, B. L. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 6640–6645
17. Ansaldi, M., Lepelletier, M., and Méjean, V. (1996) *Anal. Biochem.* **234,** 110–111
18. Bax, A., Clore, G. M., Driscoll, P. C., Gronenborn, A. M., Ikura, M., and Kay, L. E. (1990) *J. Magn. Reson.* **87,** 620–627
19. Bax, A., and Ikura, M. (1991) *J. Biomol. NMR* **1,** 99–104
20. Grzesiek, S., and Bax, A. (1992) *J. Magn. Reson.* **96,** 432–440
21. Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc.* **114,** 6291–6293
22. Vuister, G. W., and Bax, A. (1994) *J. Biomol. NMR* **4,** 193–200
23. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289–302
24. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
25. Linge, J. P., and Nilges, M. (1999) *J. Biomol. NMR* **13,** 51–59
26. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR* **4,** 171–180
27. Rogov, V. V., Lucke, C., Muresanu, L., Wienk, H., Kleinhaus, I., Werner, K., Lohr, F., Pristovsek, P., and Ruterjans, H. (2003) *Eur. J. Biochem.* **270,** 4846–4858
28. Sam, M. D., Cascio, D., Johnson, R. C., and Clubb, R. T. (2004) *J. Mol. Biol.* **338,** 229–240
29. Gajiwala, K. S., and Burley, S. K. (2000) *Curr. Opin. Struct. Biol.* **10,** 110–116
30. Ilangovan, U., Wojciak, J. M., Connolly, K. M., and Clubb, R. T. (1999) *Biochemistry* **38,** 8367–8376
31. Abremski, K., and Gottesman, S. (1982) *J. Biol. Chem.* **257,** 9658–9662
32. Abbani, M., Iwahara, M., and Clubb, R. T. (2005) *J. Mol. Biol.* **347,** 11–25
33. Clubb, R. T., Omichinski, J. G., Savilahti, H., Mizuuchi, K., Gronenborn, A. M., and Clore, G. M. (1994) *Structure* (*Camb.*) **2,** 1041–1048
34. Sam, M. D., Papagiannis, C. V., Connolly, K. M., Corselli, L., Iwahara, J., Lee, J., Phillips, M., Wojciak, J. M., Johnson, R. C., and Clubb, R. T. (2002) *J. Mol. Biol.* **324,** 791–805
35. Vogel, J. L., Li, Z. J., Howe, M. M., Toussaint, A., and Higgins, N. P. (1991) *J. Bacteriol.* **173,** 6568–6577
36. Wojciak, J. M., Iwahara, J., and Clubb, R. T. (2001) *Nat. Struct. Biol.* **8,** 84–90
37. Cho, E. H., Alcaraz, R., Jr., Gumport, R. I., and Gardner, J. F. (2000) *J. Bacteriol.* **182,** 5807–5812
38. Numrych, T. E., Gumport, R. I., and Gardner, J. F. (1992) *EMBO J.* **11,** 3797–3806
39. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA

Structural and Genetic Analyses Reveal a Key Role in Prophage Excision for the TorI Response Regulator Inhibitor

Latifa ElAntak, Mireille Ansaldi, Françoise Guerlesquin, Vincent Méjean and Xavier Morelli

J. Biol. Chem. 2005, 280:36802-36808.
doi: 10.1074/jbc.M507409200 originally published online August 2, 2005
