Functional Properties and Structural Requirements of the Plasmid

pMV158-Encoded MobM Relaxase Domain

Cris Fernández-Lópeza, Radoslaw Pluta,b,c, Rosa Pérez-Luque,b,c, Lorena Rodríguez-González,a, Manuel Espinosaa, Miquel Collb,c, Fabián Lorenzo-Díaz#d, D. Roeland Boer#b,c

a Centro de Investigaciones Biológicas, CSIC, Madrid, Spain.
b Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain
c Institute of Molecular Biology of Barcelona (IBMB-CSIC) Barcelona, Spain
d Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias, Tenerife, Spain

# To whom correspondence should be addressed:
D. Roeland Boer, Institute for Research in Biomedicine, Barcelona Science Park, Baldiri Reixac 10, 08028 Barcelona, Spain; and Institut de Biologia Molecular de Barcelona (CSIC), Barcelona Science Park, Baldiri Reixac 10, 08028 Barcelona, Spain. E-mail: rbocri@ibmb.csic.es. Phone: +34 93 4034953. Fax: +34 934 034 979.

Fabián Lorenzo-Díaz, Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias, Astrofísico Francisco Sánchez, s/n, 38203 La Laguna, Tenerife, Spain. E-mail: florenzo@ull.edu.es. Phone: +34 922 318 484. Fax: +34 922 318 490.

Running title: MobM endonuclease properties
A crucial element in the horizontal transfer of mobilizable and conjugative plasmids is the relaxase, a single-stranded endonuclease that nicks the origin of transfer (oriT) of the plasmid DNA. The relaxase of the pMV158 mobilizable plasmid is MobM (494 residues). In solution, MobM forms a dimer through its C-terminal domain, which is proposed to anchor the protein to the cell membrane and to participate in T4SS protein-protein interactions. In order to gain a deeper insight into the structural MobM requirements for efficient DNA catalysis, we studied two endonuclease domain variants that include the first 199 or 243 amino acid residues (MobMN199 and MobMN243, respectively). Our results confirmed that the two proteins behaved as monomers in solution. Interestingly, MobMN243 relaxed supercoiled DNA and cleaved single-stranded oligonucleotides harboring the oriT_{pMV158}, whereas MobMN199 was active only on supercoiled DNA. Protein stability studies using gel electrophoresis and mass spectrometry showed increased susceptibility to degradation at the domain boundary between the N- and C-terminal domains, suggesting that the domains change their relative orientation upon DNA binding. Overall, these results demonstrate that MobMN243 is capable of nicking the DNA substrate independently of its topology and that the amino acids 200-243 modulate substrate specificity, but not the nicking activity per se. These findings suggest that these amino acids are involved in positioning the DNA for the nuclease reaction rather than in the nicking mechanism itself.

**Keywords:** relaxase; bacterial conjugation; DNA transfer; DNA–protein complex
Conjugation is one of three fundamentally different ways, in addition to transformation and transduction, by which the genetic content of a cell can be altered (1, 2). The conjugative process involves active DNA transfer from one living organism to another through the Type 4 secretion system (T4SS), which is encoded on the donor cell side (3, 4). The shuttling of plasmids and Integrative and Conjugative Elements (ICEs) between organisms has proven to be a major contributor to evolutionary diversity as many genes have been exchanged rather than evolved by mutation during evolution (5). Although conjugation is generally associated with the exchange of genetic material, it most probably originated as a protein transport mechanism, since DNA transfer is always preceded by protein transfer and T4SS transfer is frequently used for translocation of proteinaceous virulence factors (6). The protein that plays a central role in the conjugational transfer of DNA is termed relaxase (7, 8). It associates with other factors to form a complex called the relaxosome (8), which provides the context for the controlled cleavage and processing of transferable DNA. The relaxase recognizes a specific DNA sequence, the origin of transfer (oriT), which generally contains one or more inverted repeats (9). The relaxase then nicks the DNA at a specific position and forms a stable complex with the oriT-harboring strand in the 5’ terminus, leaving a free 3’-OH end. The DNA-protein adduct is then transferred to the recipient cell, while the opposite strand is replicated in the donor cell using the available 3’-oxygen, generally through a rolling circle replication mechanism (10). Finally, the single-stranded (ss) DNA in the recipient cell is replicated, possibly through lagging-strand synthesis (11).

The nuclease functionality of relaxases is provided by a domain that is conserved among relaxases and replication initiators of rolling-circle replication from various genetic entities (7, 12, 13). The nuclease domains representing the MOBc (TraI_pUC1, TrwC_R388 and TraI_F) and MOBa (MobA_R1162 and NES_pLW1043) families have been structurally characterized (14-19). A comparison of these structures shows a similar fold of the nuclease domain and suggests a like mechanism of action. All relaxases studied up to now were found to use a metal-dependent nucleophilic attack of a catalytic residue to the phosphodiester bond. The metal ion is believed to polarize the phosphodiester group, facilitating the attack by the nucleophile. The nucleophilic amino acid of the relaxases that have been characterized to date was found to be a tyrosine. Despite the general similarity between the nuclease domains of the different relaxases, they can be classified into families on the basis of their sequences and...
properties (7). The classification is likely to be related to details in the DNA processing reactions, e.g. to the type of metal used for bond polarization. Structural information on the Mob families other than MOBz or MOBz is scarce, especially for relaxases of plasmids isolated from Gram-positive (G+) bacteria (20). Recently, the structure of NES, the MOBz relaxase from the staphylococcal plasmid pLW1043 has been determined (19). This represents the only structure of a relaxase isolated from G+ bacteria to date. The MobM relaxase from the mobilizable streptococcal plasmid pMV158 (20, 21) is the prototype of the MOBV family of relaxases, which holds more than 100 members (7). The full-length MobM protein (Ser2-Phe494, Met1 is removed during production) has a predicted molecular weight of 57,692 Da and behaves as a dimer in solution (21). Its nuclease activity resides in the N-terminal domain, in which three conserved motifs can be identified: i) motif I (HxR), of unknown function; ii) motif II (NYD/EL), which contains the proposed catalytic tyrosine (22); and iii) motif III (HxDExxPHuH), also known as the 3H motif and probably involved in the coordination of a divalent metal (7, 22). The conjugative transfer is initiated by nicking of the oriT of pMV158 at the phosphodiester linkage between guanidine 3,595 and thymine 3,596 (Fig. 1 and Table 1; (23). The oriT of pMV158 has an intricate pattern of inverted repeats (20), allowing the formation of at least three mutually exclusive hairpin-loop structures (see Fig. 1). Using ssDNAs and a truncated MobM protein containing the first 199 residues, we have recently defined a minimal oriT binding sequence of 26 nucleotides (20).

Here we report the DNA nicking properties, oligomeric state and degradation behavior of the full-length MobM protein and two N-terminal constructs that included the first 199 (MobMN199) or 243 (MobMN243) amino acid residues. The substrate specificities, physiochemical properties and stabilities differed depending on the construct length. A comparative functional assay demonstrated that the extra 44 amino acid residues of MobMN243 protein (compared to MobMN199) are essential for efficient ssDNA processing. To the best of our knowledge, this study provides the first in-depth characterization of the functional properties and structural requirements of a DNA relaxase domain belonging to the MOBV family.

MATERIALS AND METHODS
Bacterial strains, plasmids and culture conditions

*Streptococcus pneumoniae* 708 (end-1, exo-2, trt-1, hex-4, malM594) harboring pMV158 was used to purify plasmid DNA. *Escherichia coli* BL21 (DE3) (r*rocB* m*trc*, gal, ompT, int::*PlacUV5-T7* gene 1 imm 21 nin5; a gift of F.W. Studier) was used for the purification of MobM, MobMN199 and MobMN243. For overexpression of the *mobM* gene, plasmid pLGM2 was used; it contains two copies of the *mobM* gene under the control of the *ϕ10* promoter of phage T7 (23). In the case of MobMN199 and MobMN243 proteins, plasmids pMobMN199 and pMobMN243 (pET24b based-vectors, see below) were used; they contain one copy of the *mobMN199* or *mobMN243* genes under the control of the *ϕ10* promoter of phage T7 (23). *E. coli* was grown on tryptone-yeast extract (TY) medium (Pronadisa) at 37°C, whereas *S. pneumoniae* was grown in AGCH medium supplemented with 0.2% yeast extract and 0.3% sucrose (24, 25). In the case of plasmid-harboring cells, the medium was supplemented with tetracycline (Tet, 1 µg/ml) for *S. pneumoniae* and with ampicillin (Amp, 100 µg/ml) for *E. coli*/pLGM2, pMobMN199 and pMobMN243.

DNA purification, amplification and cloning

Plasmid DNA from pMV158 (26) was purified from *S. pneumoniae* by two consecutive CsCl gradients as described (27). Plasmids pET24b, pMobMN199, pMobMN243 and pLGM2 were purified by minipreparations using the High Pure Plasmid Isolation Kit (Roche Applied Science). The oligonucleotides used for nicking and stability studies were synthesized and HPLC-purified by Integrated DNA Technologies (Coralville, IA, USA) and are listed in Table 1.

Plasmids pLGM2 and pMobMN199 were constructed as described (23, 20). To obtain plasmid pMobMN243, a 768-bp fragment that encodes the first 243 residues of MobM protein was amplified by PCR on pMV158 DNA (accession number X15669) with the primers Nde-F and P244-Stop-Xho (Table 1). To facilitate cloning, forward and reverse primers included changes to generate recognition sites for *NdeI* and *XhoI* restriction enzymes, respectively. Furthermore, a double-base change was also introduced in the reverse primer to introduce an amber stop codon instead of Pro244 (P244Stop). PCR amplifications were done in 50-µl reaction mixtures containing reaction buffer 1X [16 mM (NH₄)₂SO₄, 6 mM Tris–HCl (pH 8.8), 1.5 mM MgCl₂, 0.2 mM of each dNTP (Roche), 0.4 mM of each primer, 0.65 U of Phusion DNA
polymerase (Finnzymes) and 1 ng of DNA template. The conditions were as follows: an initial denaturing step at 98°C (30 s), 30 cycles of denaturation at 98°C (10 s), annealing at 55°C (20 s) and extension at 72°C (30 s), followed by a final extension at 72°C (10 min). PCR products were purified with the QIAquick gel extraction kit (Qiagen). After digestion with \texttt{NdeI} and \texttt{XhoI} enzymes, the fragments were ligated into pET24b previously digested with the same enzymes. All the constructs were confirmed by Sanger DNA sequencing (Secugen S.L., Madrid, Spain).

**Protein purification and secondary structure prediction**

MobM, MobMN199 and MobMN243 proteins were purified as described (20, 21, 23). After cell disruption of the cell pellet of the induced culture (4 L), the sample was centrifuged to remove cell debris. The nucleic acids were precipitated by 0.2% (v/v) polyethylenimine (Sigma) and proteins of the supernatant were precipitated at 70% (w/v) ammonium sulfate saturation. Proteins in the precipitate were dissolved in buffer A [25 mM Tris-HCl, pH 7.6, 1 mM dithiothreitol, 1% glycerol (v/v), 1 mM EDTA] with 300 mM NaCl and after equilibration in same buffer by dialysis, the sample was loaded in a 100 ml heparin-agarose (BioRad) column. The protein was eluted with 400 ml of a 0.3–0.8 M NaCl gradient and fractions were analyzed by 12% Sodium dodecyl sulfate-Tris-glycine polyacrilamide gel electrophoresis (SDS-PAGE), using Bio-safe Coomassie (Bio-Rad Laboratories) as staining agent. Fractions containing the peaks of MobM, MobMN243 or MobMN199 were pooled and concentrated by filtering them through 3 kDa cut-off membranes (Pall). The sample was subjected to a gel-filtration column (HiLoad Superdex-200 from GE-Healthcare Life Science) connected to a fast-pressure liquid chromatography system (FPLC; Biologic DuoFlow from Bio-Rad). Fractions were analyzed and concentrated as above until reaching a concentration of 2 mg/ml and then stored at -80°C. The concentration of MobM, MobMN243 and MobMN199 was estimated by spectrophotometric measurements considering theoretical molar extinction coefficients of 37,485, 24,410 and 17,420 M⁻¹ cm⁻¹, respectively. After running a purified MobMN243 sample in a 12% SDS-PAGE gel, the protein bands were transferred to a polyvinylidene difluoride (PVDF) blotting membrane. Edman’s sequential degradation was performed directly from the PVDF membrane in order to analyze N-terminal sequences in a Procise 494 protein sequencer (Perkin Elmer) as reported (23).
Secondary-structure predictions of MobMN243 were carried out with the programs PSIPRED (28), Jpred (29), NPS@ (30) and SABLE (31) using the website http://www.expasy.ch/. The structure prediction was manually aligned to the secondary structure elements of the published structure of NES_plW1043 (19).

Analytical ultracentrifugation

Sedimentation equilibrium experiments were performed at 20°C in an Optima XL-A (Beckman-Coulter) analytical ultracentrifuge (CIB, Madrid). MobMN243 (10 and 20 µM) was examined in buffer UA [20 mM Tris–HCl pH 7.6, 500 mM NaCl, 1 mM EDTA, 1% (v/v) glycerol]. Samples were centrifuged at two successive speeds (20,000 and 35,000 rpm) and absorbance was measured at 3h intervals to determine that samples had reached equilibrium. Apparent average molecular masses (M$_{w,app}$) of MobMN243 were obtained using the program Heteroanalysis v1.1.44 (Univ. of Connecticut, Storrs, CT). Sedimentation velocity experiments were performed at 48,000 rpm and 20°C with samples of MobMN243 (10 and 20 µM, in buffer UA). The sedimentation coefficient distribution for this protein was calculated by least-squares boundary modeling of sedimentation velocity data using the sedimentation coefficient distribution method, as implemented in the SEDFIT 12.1b program (32). The coefficient was corrected to standard conditions to obtain the corresponding S$_{20,w}$ value using the SEDNTERP program (33). The translational frictional coefficient of MobMN243 ($f$) was determined from the molecular mass and sedimentation coefficient of the protein (34), whereas the frictional coefficient of the equivalent hydrated sphere ($f_0$) was estimated using a hydration of 0.4506 g H$_2$O per g of protein (35).

Supercoiled (sc) and single-stranded (ss) DNA nicking

Nicking of scDNA by purified MobMN199 and MobMN243 proteins were performed as reported (20, 23).

Cleavage activity of ssDNA by purified MobMN199 or MobMN243 proteins was tested in standard reaction mixtures (20 µl) containing fluorophore-labeled oligonucleotide (60 nM) in buffer [25 mM Tris-HCl (pH 7.6), 0.1 mM EDTA, 10% glycerol (v/v), 1 mM dithiothreitol] to which different concentrations of purified MobMN199 or MobMN243 were added. Standard reactions were performed at 30°C for 20 min in...
the presence of 8 mM MnCl₂ and were stopped by addition of SDS (0.5%) and proteinase K (1 mg/ml), followed by incubation for 30 min at 37°C. DNA was precipitated by adding 0.3 M Na acetate and 100% ethanol and incubation for 30 min at -20°C. After centrifugation at 13,000 rpm for 10 min at 4°C the pellet was dried (10 min, RT) and dissolved in distilled water with formamide. Samples were heated to 90°C for 3 min, after which they were flash-cooled to 4°C. Nicking of oligonucleotides was monitored by electrophoresis on 20% polyacrylamide (PAA) plus 7M urea and visualized by Phosphorimager (Fuji). Band intensities were measured using Bio-Rad Laboratories software.

To test the capacity of MobMN243 to form stable complexes with ssDNA, 10 µM of protein was incubated in a mix (20 µl) containing buffer [25 mM Tris-HCl (pH 7.6), 0.1 mM EDTA, 10% glycerol (v/v), 1 mM dithiothreitol] and the oligonucleotide (10 or 20 µM). After 20 min at 30°C, samples were loaded in a 12% SDS-PAGE gel and stained with SYPRO Ruby (Bio-Rad) following the manufacturer's recommendations.

**Protein stability assays**

The time-dependent stability of MobMN243 was assayed by SDS-PAGE as follows: purified protein (180 µM) from one single batch in buffer A plus 100 mM NaCl was supplemented with either: i) MnCl₂ (15 mM final concentration); ii) oligonucleotide IR1+8 (in 10% excess to MobMN243), or iii) MnCl₂ (15 mM final concentration) plus oligonucleotide IR1+8 (in 10% excess to MobMN243). The three samples were incubated at 4°C for 2 months and then analyzed by SDS-PAGE (12%). The IR1+8 oligonucleotide (Table 1) represents the minimal oriT recognized with a high binding affinity by MobMN199 (20). For comparison purposes, freshly purified full-length MobM and MobMN199 were also included in the gels.

DNA oligonucleotide IR1+12, encompassing the minimal oriT binding site (IR1+8) and 4 nucleotides downstream of the cleavage site (see Table 1), was used for mass spectrometry (MS) stability experiments. MobMN243 (40 µM) in buffer A plus 100 mM NaCl was supplemented with either phosphoramidate or with a mixture of 15 mM MnCl₂ or MgCl₂ and the IR1+12 oligonucleotide (50 µM). Phosphoramidate was synthesized as described (36). The samples were then incubated at 30°C for 20 min and subsequently supplemented with 40 mM EDTA (final concentration) and placed on ice. The resulting samples were then either loaded on a Superdex-200 column or HiTrap Desalting 5 ml column.
Size-exclusion chromatography was performed at 4°C. Fractions containing protein-DNA complex were concentrated using centrifuge concentrators with a 10 kDa cutoff and subjected to the MS analysis.

Electrospray mass spectrometry (ESI) was performed on a LTQ-FT Ultra instrument (Thermo Scientific), which was equipped with the nanoflow ESI source. The NanoMate (Advion BioSciences, Ithaca, NY, USA) aspirated the samples from a 384-well plate (protein Lobind) with disposable, conductive pipette tips, and infused the samples through the ESI Chip, which consists of 400 nozzles in a 20x20 array. The spray voltage was 1.7-1.85 kV and delivery pressure 0.3-0.35 psi. Samples, provided in 100 mM ammonium acetate, pH 6.8 were diluted either in 0.1% acetonitrile/formic acid or 1% acetonitrile/water/formic acid to a final concentration of 1-5 µM in order to perform MS analysis (final formic acid concentration of 0.1-0.5%).

RESULTS & DISCUSSION

Purification of MobMN243. To gain a deeper insight into the structure-function properties of the MobM endonuclease domain and its binding to the pMV158 oriT (Fig. 1), we designed a strategy to obtain a protein containing the first 243 N-terminal residues. A truncated MobMN243 protein was obtained by replacing the P244 amino acid residue by a stop codon. To purify MobMN243, the protocol used to obtain the MobMN199 protein was followed (20). The fractions recovered after the various purification stages are shown in Figure 2A. After the ammonium sulfate precipitation (lanes 6 and 7), the sample was subjected to heparin-affinity chromatography and the pure fractions were collected and subjected to a final gel-filtration step. The purest fractions were concentrated to 2 mg/ml (lane 8). The MobMN243 gel-filtration profile was then compared to the chromatograms of MobMN199 and full-length MobM (Fig. 2B). MobM eluted in a single peak at 61.5 ml while MobMN243 and MobMN199 eluted at 80.5 ml and 87.1 ml, respectively. Since MobM and MobMN199 behave as dimer and monomer, respectively (20, 21), the elution behavior suggests that MobMN243 most probably is a monomer in solution. Under denaturing conditions, MobMN243 migrated with a relative molecular mass (Mr) between 30-45 kDa, which differs slightly from the theoretical mass 28,393.6 Da (Table 2). An increase in the amount of SDS in the gels...
abolished this effect, resulting in band migration corresponding to the molecular weight (between 20-30 kDa).

**MobMN243 oligomerization state.** Analytical ultracentrifugation analyses showed that full-length MobM behaved as an ellipsoidal dimer in solution (21), whereas MobMN199 was observed to be a monomer (20). Since the shorter version MobMN243 lacks the last 251 C-terminal amino acids, we examined its oligomeric state and hydrodynamic properties. The elution volume of the truncated protein in the size-exclusion chromatography used for purification (Fig. 2B) suggested that MobMN243 is a monomer. To corroborate this, analytical ultracentrifugation assays were performed at two protein concentrations (10 and 20 µM). The results of sedimentation equilibrium and velocity assays are shown in Figure 3. Sedimentation velocity profiles fit well with a model of single sedimenting species, with an $S_{20,w}$ of 2.5 S and an average molecular mass ($M_{w,app}$) of 27,503 ± 1,000 Da. These findings are consistent with the theoretical monomer mass. At 10 µM and 20 µM, the experimental sedimentation equilibrium data fitted best with an average $M_{w,app}$ of 30,500 Da ± 1,000 Da, which is near the value obtained for the sedimentation velocity and is also consistent with a monomeric species. The frictional ratio ($f/f_0$) calculated was 1.29, indicating that MobMN243 deviates from the behavior expected for globular particles ($f/f_0 ~1$).

Clearly, the dimerization interface resides in the C-terminal domain, perhaps through self-interaction of two protomers through a putative Leucine zipper motif (from Leu 317 to Leu338) present in this part of the protein (37).

**Substrates of MobMN243.** The oriT of pMV158 contains the MobM-mediated nick site (located between coordinates 3,591 and 3,592) and three inverted repeats (IR1, IR2 and IR3). The capacity of MobMN199 and MobMN243 to convert scDNA (form FI) into open circular species (form FII) was tested at several protein concentrations (from 8 to 256 nM) in the presence of 8 mM MnCl₂. The samples were incubated at 30°C for 20 min, as reported for the full-length protein (23). No differences in scDNA nicking efficiency between MobMN199 and MobMN243 were detected. Indeed, the results showed that both truncated proteins relaxed pMV158 DNA with the same efficiency as full-length MobM, converting up to 60% of supercoiled plasmid into the FII form (Fig. 4A). The failure to reach 100% cleavage could be explained by...
the equilibrium between the nicking and closing activities of MobMN243, which has been observed for several relaxases (38-40).

Nicking assays using oligonucleotide 19nic4 (Table 1) were performed in order to study the capacity of MobMN199 and MobMN243 to cleave ssDNA. This oligonucleotide encompasses the right arm of the IR3 (19-nt) and 4-nt after the nick site (Fig. 1). A control with no proteins and different concentrations of the two proteins (60-480 nM range) were incubated with 60 nM of the 19nic4 ssDNA (Table 1) in the presence of 8 mM MnCl₂ as described (see Materials and Methods and Fig. 4B). The expected products after a specific nick of the 19nic4 substrate are a 19-mer and a 4-mer. Contrary to the results obtained in the scDNA relaxation assays, no ssDNA cleaving activity was observed when the reaction was carried out with MobMN199 (Fig. 4B) or with an N-terminal His-tagged version of MobM243 (results not shown). However, a 19-mer band appeared in the presence of MobMN243 even at the lowest concentration tested (Fig. 4B).

Interestingly, all of the N-terminal truncations studied retained nuclease activity, indicating that all the residues required for nicking activity reside in the first 199 amino acids, which was the shortest construct tested here. As the MobM C-terminal is truncated beyond S243, the substrate range that can be cleaved becomes more restricted (see Table 2). Thus, MobMN243 behaves like full-length MobM as far as substrate nicking is concerned and both are able to cleave ssDNA as well as scDNA. Removal of the 200-243 amino acids region abolishes activity on ssDNA, although MobMN199 still cleaved the scDNA. The observation that the presence of the C-terminal 44 amino acid residues of MobMN243 are essential for efficient ssDNA cleavage suggest that the minimal endonuclease requirements for an efficient nicking depend on the DNA topology.

The minimal requirements for efficient DNA nicking have been analyzed for other conjugative relaxases belonging to the MOB₇ and MOB₂₀ families. The first 275 amino acid residues of the TrwC relaxase (R388 plasmid) catalyzed ssDNA cleavage, although a longer fragment (348 amino acid residues) was required to produce the nick on a supercoiled double-stranded DNA substrate (41). In the case of MobA (R1162 plasmid) it has been demonstrated that a 184 residue amino-terminal fragment cleaves ssDNA as well as scDNA (17). Without high-resolution structural information, we can only speculate about whether the substrate specificity of the different constructs has a structural origin or not.
whether other mechanisms are involved. However, a plausible explanation reconciles the experimental observations, and assumes that MobM residues beyond amino acid 200 participate in DNA processing, inducing a conformation amenable to ssDNA nicking. The finding that MobMN199 cuts scDNA can be attributed to the topology of the supercoiled plasmid, allowing nicking to occur in the absence of the 200-243 amino acid region. However, further processing of the DNA is expected to depend on these residues. Overall, the data provided here show that the mechanistic nuclease activity depends solely on the N-terminal domain whereas DNA processivity depends on a region closer to the C-terminal domain.

**MobMN243-DNA stable complexes.** The conjugative relaxases carry out cleavage reactions of a specific phosphodiester bond (nick site) in order to initiate DNA transfer. Such a reaction is based on a transesterification mechanism that produces a stable protein–DNA intermediate and a free 3′-OH terminus. To visualize the protein-ssDNA adducts expected after the nicking reaction, we incubated MobMN243 with the oligonucleotide IR1+12pA, which comprises the IR1+12 sequence and a final 15-mer long poly A tail (Table 1 and Fig. 1B) to increase the molecular weight of the protein-DNA complexes and facilitate the separation of the complex from the free-MobM protein by electrophoretic analysis. The predicted products after the IR1+12pA oligonucleotide cleavage are two shorter DNA fragments, one of them (a 21-mer) attaches to the MobMN243 protein (Fig. 5A). Highly efficient ssDNA nicking was observed for MobMN243 when the IR1+12pA oligonucleotide was used as substrate (Figure 5B). Under the same experimental conditions, the protein-DNA adducts were solved by SDS-PAGE followed by SYPRO Ruby staining. Two bands were detected, one corresponding to the free protein (MobMN243) and the other to the MobMN243-DNA21 complex (Fig. 5C). However, the ratio of covalently bound complex was quite small, which can be explained in part by the equilibrium between the nicking and closing activities of MobMN243 (42, 43). Furthermore, similar results were previously reported for the relaxase domain of TrwC, where the molecules involved in the formation of stable complexes represented less than 10% of the total protein (44).

**MobM stability.** The recent structures of NES, a relaxase found in Gram-positive bacteria (19), indicate that MobM-like nucleases are divided into two functional domains, which are separated by a boundary...
region that encompasses part of the C-terminus of MobMN243. The predicted secondary-structure of the MobMN243 protein obtained by computational methods (not shown) shows that the C-terminal domain of MobM contains a sequence of secondary structure elements similar to NES. Although the structure of the C-terminal domain may differ from that of NES (45), it is likely that MobM has a similar domain architecture to this relaxase. The long-term storage stability of MobM under a range of conditions was found to reflect the domain boundary. The electrophoretic profiles of MobMN243 showed an extra band with a smaller apparent molecular weight (not shown). In order to determine whether this band corresponds to a partner protein that copurifies with MobM derivates, an Edman-sequencing approach was followed after protein transfer and band isolation from nylon membranes. Interestingly, the sequence obtained after 8 cycles was SYMVARMQ, which corresponds to the N-terminal residues of the MobM protein (excluding the initial Met residue). This result indicates that MobMN243 presents intrinsic instability, as does the entire MobM protein. For this reason, the stability of various MobM constructs after 2 months of storage in the absence and presence of substrate was studied by gel electrophoresis under denaturing conditions (Fig. 6). Incubation of full-length MobM at 4ºC for 2 months showed that the non-degraded species was predominant for the full-length protein, although the occurrence of bands of lower molecular weight indicated some degradation (Fig. 6, lane 1). Incubation of MobMN243 without substrates at 4ºC in buffer A did not show appreciable degradation (Fig. 6, lane 3) compared to fresh sample (lane 2). Likewise, the presence of Mn²⁺ (lane 4) or IR1+8 DNA (lane 5) slightly enhanced MobMN243 proteolysis; however, the main species was still the non-degraded sample. A remarkable increase in degradation was observed in the presence of all three components, i.e. protein, metal cation and DNA for MobMN243 (lane 6), but not for MobMN199 (not shown). The main product of degradation showed a migration similar to that of MobMN199 (lane 7; fresh sample). Thus, the DNA substrate had a remarkable effect on the stability of the amino acid region beyond residue 199.

To determine where degradation occurs, the masses of the degradation products were checked with ESI-MS experiments on fresh and pure samples of MobMN199 and MobMN243. The monoisotopic masses of both intact proteins corresponded well to the theoretical values (Table 2) when taking into account that Met1 was cleaved off and that oxidation and phosphorylation of the sample occurred during production and purification (not shown). Degradation was observed for MobMN243 and occurred
exclusively at the C-terminus, at amino acids Y209 and K202 and to a smaller extent at K193, A190 and E176. Degradation in the presence of Mg$^{2+}$ was diminished under the given condition with respect to the presence of Mn$^{2+}$, but similar results were obtained for Mg$^{2+}$ when extending the incubation time twofold. The results confirm that degradation occurs predominantly around residue 199, in the region 190-210.

Susceptibility to degradation has been linked to flexibility (46) and the increase in degradation susceptibility of the boundary region upon substrate binding therefore points to an increase in solvent exposure as a result of a substrate-induced change in conformation. The degradation behavior suggests that the N-terminal domains are flexible appendices attached to the dimerized C-terminal domains of MobM. How this localized degradation affects or modulates the conjugative process is not known. It has been proposed that the C-terminal domain of MobM is involved in the interactions with other elements such as the coupling protein (20), which mediates the interactions between the relaxosome with the secretion system (3, 4). Once this interaction takes place, it is reasonable to assume that the C-terminal part of MobM is not essential. In this regard, it would be interesting to determine whether the protein-DNA complex that enters the recipient cell (as demonstrated for TrwC_R388; (47)) contains the full-length MobM protein or a natural truncated version similar to MobMN199. Further in vivo experiments are needed to address this question and will allow us a deeper understanding of the pMV158 transfer mechanism.

In summary, we propose that the E200-S243 region where degradation occurs is important for DNA processivity and functions as a signal transducer of substrate binding to the C-terminal domain (21). Our data implies that the N- and C-terminal domains are loosely connected and that the intervening region is crucial for substrate processivity. We have thus dissected the MobM N-terminal sequence into at least two functionally distinct regions, where the first 190-200 amino acids are relevant for DNA nicking activity, while the 200-243 region acts as a nicking modulator and a hinge that allows for a change in relative orientations of the N- and C-terminal domains of MobM during substrate processing.

ACKNOWLEDGEMENTS
This work was supported by the Spanish Ministry of Science and Innovation (grants BFU2008-02372/BMC, CONSOLIDER CSD 2006-23 and BFU2011-22588 to M.C. and grants CSD2008-00013, INTERMODS and BFU2010-19597 to M.E.), the Generalitat de Catalunya (grant SGR2009-1309 to M.C.) and the European Commission (FP7 Cooperation Project SILVER - GA No. 260644 to M.C.). We thank Marina Gay and Marta Vilaseca of the Mass Spectrometry Core Facility at IRB Barcelona.
REFERENCES


### Table 1. Oligonucleotides used in this study.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5’ to 3’)</th>
<th>Coordinates (nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nde-F</td>
<td>AAGGAGGGAAAAACATAGAGTTACA</td>
<td>3718-3741</td>
</tr>
<tr>
<td>P244-Stop-Xho</td>
<td>CTTGGTTAGTTGCTCGAGTATCTAAGATTGAAC</td>
<td>4485-4452</td>
</tr>
<tr>
<td>19nic4</td>
<td>GAATATAAAGTATAGTGTG/TTAT</td>
<td>3577-3599</td>
</tr>
<tr>
<td>IR1+12pA</td>
<td>ACTTTATGAATAAAGTATAGTGTTATGG(A15)</td>
<td>3570-3599</td>
</tr>
<tr>
<td>IR1+8</td>
<td>ACTTTATGAATAAAGTATAGTG</td>
<td>3570-3595</td>
</tr>
<tr>
<td>IR1+12</td>
<td>ACTTTATGAATAAAGTATAGTGTTAT</td>
<td>3570-3599</td>
</tr>
</tbody>
</table>

a The oligonucleotide IR1+12pA contains a 15 nt-long poly(A) tail.

b Restriction sites are in boldface and base changes that generate restriction sites are underlined. The double base change introduced in the P244-Stop-Xho oligonucleotide to obtain the stop codon (P244Stop) is shown in italics. The oriT cleavage site is indicated by “/”.

c Coordinates are given with respect to the pMV158 sequence (accession number NC_010096).

d These oligonucleotides were labeled with the Cy5 fluorophore at its 5’-end.

### Table 2. Substrate processivity and masses of MobM protein and its truncated versions.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Residues *</th>
<th>Calculated M.W. (Da) b</th>
<th>Measured M.W. (Da) c</th>
<th>Activity on DNA</th>
<th>Supercoiled</th>
<th>Single stranded</th>
</tr>
</thead>
<tbody>
<tr>
<td>MobM</td>
<td>493</td>
<td>57,692.3</td>
<td>ND</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>MobMN199</td>
<td>198</td>
<td>23,115.2</td>
<td>23,195.3 *</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>MobMN243</td>
<td>242</td>
<td>28,393.6</td>
<td>28,409.7 *</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

* Number of total amino acid residues, excluding Met1, which is processed in the expression systems used in this study.

b Monoisotopic, excluding Met1 (see comment a).

c Masses as determined by ESI-MS (see text).

* Mass corresponds to MobMN199 with one phosphorylated residue.

* Mass corresponds to MobMN243 oxidized at one position.
FIGURE LEGENDS

FIG 1 Genetic map and origin of transfer (oriT) of plasmid pMV158. (A) Scheme of pMV158. Arrowheads indicate the direction of transcription of the four protein-coding genes (mobM, copG, repB and tetL). The positions of the oriT and the origins of the leading-strand (dso) and lagging-strand (ssoU and ssoA) synthesis are shown. (B) DNA sequence of the oriT of pMV158 (coordinates 3564-3605). The three identified inverted repeats (IR1, IR2 and IR3) are indicated by arrows and the nick site by a vertical arrowhead. The oligonucleotides used in this study are also depicted (see Table 1).

FIG 2 Over-expression and purification of the MobMN243 protein. (A) Fractions from the different purification steps of MobMN243 were analyzed by electrophoresis on 12% SDS-PAGE gels. Samples were loaded in the following order: non-induced cultures (lane 1); cultures after induction with IPTG (lane 2) and rifampicin (lane 3); supernatant of a total cell lysate (lane 4); supernatant after PEI precipitation (lane 5); supernatant of the ammonium sulfate precipitation step dialysis against buffer A plus 300 mM NaCl (lane 6); and protein soluble fraction after dialysis (lane 7). The sample was loaded onto a heparin–agarose column and then subjected to a gel-filtration column (lane 8). M indicates the molecular weight standards (in kDa). (B) Comparative analysis of the gel-filtration elution profiles of MobM (grey dashed line), MobMN243 (continuous black line) and MobMN199 (continuous grey line) subjected to 1 ml/min isocratic flow (buffer A plus 300 mM) in a HiLoad Superdex-200 column.

FIG 3 MobMN243 oligomerization state studied by analytical ultracentrifugation profile. The sedimentation equilibrium profile (35,000 rpm and 20°C) of 20 µM MobMN243 in buffer UA is shown. The lower part shows the experimental data (circles) and the best fit (solid line) to a single species of $M_w$ (31,000 ± 1000 Da). The upper part shows residuals of the theoretical fit. The inset shows the distribution of sedimentation coefficients of the same MobMN243 protein sample in sedimentation velocity experiments (48,000 rpm at 20°C).
FIG 4 Nicking activity of MobMN199 and MobMN243. (A) Relaxation assays of pMV158 by MobMN199 and MobMN243. Supercoiled pMV158 DNA samples (8 nM) were incubated with (8, 16, 32, 64, 128, 256 nM) or without (-) MobMN199 or MobMN243 and supplemented with 8 mM MnCl₂ at 30°C, 20 min. Generation of relaxed forms (FII) from scDNA forms (FI) was analyzed by electrophoresis on 1% agarose gels stained with EtBr 1 µg/ml. The weak band above relaxed forms FII has been observed before (20) and might correspond to relaxed DNA dimers. (B) Nicking activity of MobMN199 and MobMN243 on single strand DNA harboring regions of the pMV158-oriT (19nic4). The nicking reaction yields 19-mer and 4-mer (not shown) oligonucleotides products. 19nic4 (60 nM) was incubated with a range of protein concentrations (60, 120, 240, 480 nM) at 30°C for 20 min in the presence of 8 mM MnCl₂. The reaction was stopped with 0.1% SDS and proteinase K (1 mg/ml) for 30 min at 37°C and samples were electrophoretically analyzed in a denaturing 20% PAA - 7M urea gel and visualized by Phosphorimager. Band quantification was done using the QuantityOne software (Bio-Rad) in both (A) and (B).

FIG 5 Isolation of the ssDNA-MobMN243 stable complexes after IR1+12pA oligonucleotide cleavage. (A) Scheme of the substrate and reaction products expected after IR1+12pA nicking by MobMN243. (B) DNA analysis. IR1+12pA 5’-labeled (Cy5) oligonucleotide (60 nM) was incubated with different concentrations of MobMN243 (128, 256, 512 µM) for 20 min at 30°C in the presence of 8 mM MnCl₂. The reaction was stopped by incubation with 0.1% SDS and proteinase K (100 µg/ml) for 30 min at 37°C. Samples were loaded in a denaturing 20% PAA gel (7M urea) and DNA visualized in a Phosphorimager platform. (C) Protein analysis. MobMN243 (10 µM) was incubated with DNA (10 or 20 µM) in the same conditions used for the DNA analysis. Samples were visualized by SDS-PAGE (12%) and SYPRO Ruby staining. A stable adduct corresponding to the MobMN243 protein attached to a 21-mer ssDNA fragment was identified (MobMN243-DNA21).

FIG 6 MobMN243 ligand-dependent proteolysis as analyzed by 12% SDS-PAGE. Lane 1: Full-length MobM (57.9 kDa), incubated for 2 months at 4°C. Lane 2: MobMN243 (28.4 kDa), fresh sample. Lane 3: MobMN243 incubated for 2 months at 4°C. Lane 4: MobMN243, incubated for 2 months at 4°C in the presence of MnCl₂ (15 mM final concentration). Lane 5: MobMN243 incubated for 2 months at 4°C in the
presence of IR1+8 (180 μM). Lane 6: MobMN243 incubated for 2 months at 4°C in the presence of MnCl$_2$

plus IR1+8. Lane 7: MobMN199 (23.1 kDa), non-degraded fresh sample, included for comparison with the
degraded samples in lanes 1-6. The marker lane (M) was loaded with SeeBlue Pre-stained Standards
(Invitrogen).