**Concerted action of NIC relaxase and auxiliary protein MobC in RA3 plasmid conjugation**

Jolanta Godziszewska1,3, Gabriel Moncalián2, Matilde Cabezas2, Aneta A. Bartosik1, Fernando de la Cruz2 and Grazyna Jagura-Burdzy1 \*

1Institute of Biochemistry and Biophysics PAS, Department of Microbial Biochemistry, 02-106 Warsaw, Poland

2Departamento de Biología Molecular and Instituto de Biomedicina y Biotecnología de Cantabria (IBBTEC), Universidad de Cantabria-CSIC, C/ Albert Einstein 22, 39011 Santander, Spain

3Present address: Warsaw University of Life Sciences (WULS-SGGW), Faculty of Human Nutrition and Consumer Sciences, Laboratory of Food Chemistry, 02-776 Warsaw, Poland

\*Author for correspondence:

Institute of Biochemistry and Biophysics, PAS, 02-106 Warsaw, Pawinskiego 5A, Poland, tel: +48 22 592 1212, fax: +48 22 592 21 90, email: [gjburdzy@ibb.waw.pl](mailto:gjburdzy@ibb.waw.pl%20%20)

Running title: NIC relaxase from plasmid RA3 of IncU incompatibility group

Keywords: broad-host-range plasmid, IncU, conjugative transfer, MOBP4 relaxase, *oriT*

**Summary**

Conjugative transfer of the broad-host-range RA3 plasmid, the archetype of the IncU group, relies on the relaxase NIC that belongs to the as yet uncharacterized MOBP4 subfamily. NIC contains the signature motifs of HUH relaxases involved in Tyr nucleophilic attack. However, it differs in the residue involved in His activation for cation coordination and was shown here to have altered divalent cation requirements.

NIC is encoded in the *mobC-nic* operon preceded directly by *oriT*, where *mobC* encodes an auxiliary transfer protein with a dual function: autorepressor and stimulator of conjugative transfer. Here an interplay between MobC and NIC was demonstrated. MobC is required for efficient NIC cleavage of *oriT* in supercoiled DNA whereas NIC assists MobC in repression of the *mobC-nic* operon. A 7-bp arm of IR3 (IR3a) was identified as the binding site for NIC and the crucial nucleotides in IR3a for NIC recognition were defined. Fully active *oriT*RA3 was delineated to a 47-bp DNA segment encompassing a conserved cleavage site sequence, the NIC binding site IR3a and the MobC binding site OM. This highly efficient RA3 conjugative system with defined requirements for minimal *oriT* could find ample applications in biotechnology and computational biology where simple conjugative systems are needed.

**Introduction**

Horizontal gene transfer (HGT) has a major effect on the diversity of bacterial genomes. Conjugation, specified by mobile genetic elements (conjugative plasmids and integrative and conjugative elements), seems to be the most efficient way for horizontal exchange of genetic material (Halary *et al*., 2010).

The general scheme for conjugative ssDNA transfer comprises three important steps: i/ formation of the relaxosome at the origin of transfer (*oriT*), and preparation of ssDNA for transfer, ii/ transferosome assembly by a type IV secretion system (T4SS) involved in donor-recipient recognition and secretion of the nucleoprotein complex, and iii/ delivery of the relaxosome to the transferosome by a type IV coupling protein (de la Cruz *et al*., 2010; Zechner *et al*., 2000). The specificity of the conjugation machinery is mainly provided by interactions of the relaxase with *oriT* (Caryl and Thomas, 2006; Guasch *et al*., 2003; Hekman *et al*., 2008; Larkin *et al*., 2005; Lucas *et al*., 2010; Luo *et al*., 1995; Pansegrau *et al*., 1993; Williams and Schildbach, 2006). Conjugative relaxases recognize and bind a specific sequence in the *oriT* and cleave it in a strand specific manner, covalently binding the 5’ end of the nicked strand through a tyrosine residue in a transesterification reaction (Carballeira *et al.*, 2014; Grandoso *et al*., 2000; Llosa *et al*., 1996). They can function on their own (e.g., MobM of pMV158) or are assisted in their functions by one e.g., MobA of pC221 (Caryl *et al.*, 2004; Caryl and Thomas, 2006) or more auxiliary proteins e.g., TrwC of R388, MbeA of ColE1, TraI of F or TraI of RK2 (Bowie and Sauer, 1990; Furste *et al*., 1989; Lum *et al.*, 2002; Lovett and Helinski, 1975; Moncalian and de la Cruz, 2004; Moncalian *et al.*, 1999a; 1999b; Nelson *et al.*, 1995; Ragonese *et al*., 2007; Varsaki *et al.*, 2009; 2012; Ziegelin *et al*., 1989; 1991; 1992). The relaxases accompany the ssDNA during its transfer to the recipient, re-join the ends and may also participate in the replication of the complementary strand (de la Cruz *et al.*, 2010; Draper *et al*., 2005; Pansegrau and Lanka, 1996).

The conjugative relaxases vary in the number of active tyrosine residues in the catalytic center, the mechanism of cleavage and re-joining of plasmid DNA during conjugative transfer, type of divalent cation involved in the cleavage reaction, the ability to interact with various proteins, and domain structure (Pansegrau *et al*., 1993; 1994; Llosa *et al*., 1996; Grandoso *et al*., 1994; 2000; Dash *et al.*, 1992; Byrd *et al.*, 2002; Matson and Ragonese, 2005). They have been assigned to six MOB families (Francia *et al.*, 2004; Garcillán-Barcia *et al.*, 2009). Four of the relaxase families (MOBF, MOBQ, MOBP and MOBV) belong to the HUH endonuclease superfamily carrying the HUH motif composed of two His residues separated by a hydrophobic residue and the Y motif containing one (1Y) or two (2Y) catalytic Tyr residues (Chandler *et al*., 2013). The HUH enzymes require for activity a divalent metal ion to facilitate positioning of the phosphodiester bond for cleavage. It is presumed that Mg2+ or Mn2+ are the physiological co-factors although various divalent ions may be used *in vitro* (Boer *et al*., 2006; Datta *et al.*, 2003; Edwards *et al.*, 2013; Larkin *et al*., 2007; Lorenzo-Diaz *et al.*, 2011). The His pair of the HUH motif provides two ligands for the metal ion coordination, and the third ligand seems to be invariably a polar residue of Glu, Asp, His or Gln (Chandler *et al*., 2013). Although representatives of the three largest families, MOBP, MOBF and MOBQ, have been studied for years, structural data is only available for three relaxases of the MOBF family, TrwC of R388, TraI of pCU1 (MOBF11) and TraI of F plasmid (MOBF12), and two relaxases of the MOBQ family, MobA of R1162 plasmid and NES of pLW1043 (Datta *et al*., 2003; Edwards *et al.*, 2013; Guasch *et al*., 2003, Monzingo *et al.*, 2007; Nash *et al.*, 2010). Understanding the mechanisms of relaxase functioning will help combat plasmid spreading and is especially important for broad-host-range conjugative plasmids capable of replication in and transmission between evolutionarily distant species (Lujan *et al*., 2007). Moreover, the specific activities of the relaxases can find biotechnological applications where precise processing of DNA is required (Gonzales-Prieto *et al*., 2013).

The broad-host-range conjugative plasmid RA3, the IncU archetype, transfers with efficiency of one transconjugant per donor, replicates and is stably maintained in species from alpha-, beta- and gammaproteobacteria (Kulinska *et al*., 2008; 2011; Ludwiczak *et al.*, 2013; Godziszewska *et al*., 2014). The horizontal spreading of RA3 is determined by a contiguous region of 23 kb designated conjugative transfer module and comprising three operons (Fig. 1A). The bicistronic *mobC-nic* operon adjacent to the stability module (Kulinska *et al*., 2008; 2011) encodes the auxiliary transfer protein MobC (Godziszewska *et al*., 2014) and the relaxase NIC. The divergently oriented (back-to-back) long operon from *orf33* to *traC3* encodes fourteen Mpf (Mating pair formation) proteins, four proteins putatively involved in the Dtr (DNA transfer) functions, and a coupling protein VirD4. The tri-cistronic operon *orf34-orf35-orf36* has also been shown to be a part of the conjugative transfer module (J. Godziszewska, unpublished).

The relaxase NIC of the IncU plasmids RA3 (Kulinska *et al*., 2008), pFBAOT6 (Rhodes *et al.*, 2004), pMBU13 and pMBU17 (Brown *et al*., 2013), is a protein of 331 amino acids with an N-terminal relaxase/mobilisation nuclease domain with three conserved motifs (Fig. S1). Together with its closest homologs encoded by plasmids from the PromA group (Marques *et al.*,2001; Schneiker *et al*., 2001; Tauch *et al*., 2002; van der Auwera *et al.*, 2009) NIC\_RA3 was classified to an as yet unstudied group of MOBP4 relaxases (Garcillán-Barcia *et al*., 2009). Structural predictions classify the RA3 DNA binding protein MobC to the RHH family together with other auxiliary transfer proteins like TraY and TraM of F, TrwA of R388 and MbeC of ColE1 (Bowie and Sauer, 1990; Moncalian and de la Cruz, 2004; Ragonese *et al*., 2007; Varsaki *et al*, 2009). MobC is an auxiliary transfer protein of dual function: it binds to DNA at *mobC*p and thereby represses the *mobC-nic* operon, and is required (without DNA binding) for an efficient conjugative transfer (Godziszewska *et al.*, 2014).

*oriT* with the nick site motif characteristic for MOBP4 and MOBP11 relaxases (Tauch *et al.*, 2002; Garcillán-Barcia *et al*., 2009) has been mapped to a 417-bp intergenic region between the partitioning operon *korA-incC-korB-orf11* and the *mobC-nic* operon (Kulinska *et al*., 2008). This region is highly structured with various direct and inverted repeats and functions as *oriT* and *parS*, the *cis*-acting sites for conjugation and stable maintenance functions, respectively, and the *mobC* promoter (Fig. 1B). The two direct repeats (DR1 and DR2) and the inverted repeat IR-SnaBI in this region have been assigned the function of partitioning the *cis*-acting *parS* motifs (Kulinska *et al.*, 2011). Recently we have demonstrated that KorB of RA3 (ParB homolog) bound to *parS* acts as an additional repressor for *mobC*p (Kulinska *et al*., 2016) tuning down the transfer functions when active partition proceeds.

Further analysis (Godziszewska *et al*., 2014) has limited the fully functional *oriT* to a 116-bp region with three IR sequences (IR2, IR3 and IR4) and confirmed the role of the cleavage site in the bottom strand of the motif AATCCTG↑C, as shown in Fig. 1C. Whereas the degenerated palindromic sequence IR4 located between the -35 and -10 motifs of *mobC*pwas identified as an operator (OM) for the MobC autorepressor (Godziszewska *et al.*, 2014), neither the NIC\_RA3 binding site nor the significance of the inverted repeat motifs IR2 and IR3 with the arms moved apart (Fig. 1C) and flanking *mobC*p have been established yet.

In this report we mapped the minimal *oriT* and defined the sequence requirements for efficient NIC binding. We also analyzed the role of MobC in NIC\_RA3 relaxase processing of minimal *oriT*s. Special metal requirements of NIC are also described. The study highlights a concerted action of MobC and NIC in RA3 plasmid processing and transcriptional repression.

**Results**

***Minimal* oriT *of RA3***

RA3 *oriT* was previously mapped to a 417-bp intergenic region at the junction between the stability and conjugative transfer modules (Kulinska *et al*., 2008), and then limited to 116 bp (Fig. 1C; Godziszewska *et al.*, 2014). In this work, the minimal *oriT*RA3 was delineated by using pUC18 vector with different length *oriT* fragments inserted to establish which specific sequences are sufficient to be recognized and processed by the cognate relaxase and which are required to maximize the productivity of mobilization. First, a 61 bp long fragment (*oriT61*) encompassing the conserved motif with the nick site (*nic* sequence), single arms of the palindromic sequences IR2a and IR3a, and the MobC binding site OM (IR4), or a similar size but without IR2a and with both arms of IR3 (*oriT60hairpin*) were cloned into pUC18 and tested for mobilization by the RA3 plasmid. Both fragments inserted into pUC18 seemed to contain all the *cis-*acting sites required for mobilization by RA3 with an efficiency comparable to that of pUC18 carrying the full 417-bp intergenic fragment (almost one transconjugant/ donor cell) (Fig. 2A and the first panel in 2B). This results suggesting that neither IR2a nor IR3b are strictly required for conjugative DNA processing was confirmed by cloning a 47-bp fragment (*oriT47*, RA3 coordinates 9736-9782 in Fig. 1C), which rendered pUC18 fully mobilizable by RA3 (Fig. 2A). Two next inserts deprived of OM, a 45-bp fragment containing IR2a (*oriT45*), and a 31-bp fragment (*oriT31*) without IR2a, were also active in mobilization, although with a frequency two orders of magnitude lower than *oriT47* (Fig. 2A). The difference between the transfer frequency for *oriT45* versus *oriT31* (a few fold higher for pUC18 *oriT45*) was statistically significant by Student’s t-test (p=0.05). The role of the IR2a arm needs to be verified. The minimal *oriT25* (25 bp) that still enabled transfer of pUC18 (although with a frequency 10-fold lower than that observed for *oriT31*) contained the conserved motif with the nick site and IR3a without the two distal nucleotides. Further deletion of the remaining part of IR3a (*oriT20*) made pUC18 non-mobilizable at a detectable frequency. In conclusion, although the shortest *oriT* was limited to the *nic* sequence and an adjacent palindromic arm IR3a shortened by 2 nt (*oriT25*), an intact IR3a and the presence of the MobC binding site OM in *oriT47* provides mobilization efficiency by RA3 more than three orders of magnitude higher.

In the majority of *oriT* regions analyzed, the nick site (*nic*) is embedded in or lies within short distance from an inverted repeat recognized by thecognate relaxase or by some auxiliary relaxosomal protein (Zechner *et al*., 2000). Whereas the *nic*-proximal arm of such palindrome is sufficient to initiate the nicking reaction and the covalent binding of ssDNA by the relaxase the distal arm is important in recipient cells for hairpin formation (Carballeira *et al.*, 2014). In *oriT*RA3 next to the *nic* sequence lies an atypical IR3 with its 7-nt arms separated by 26 nt, overlapping the *mobC* promoter motifsand surrounding OM (Fig. 1C). The deletion mapping of minimal *oriT* strongly suggested an important role of IR3a but not IR3b in efficient transfer (Fig. 2) raising the question about the role of IR3 arms in NIC binding.

To analyze the role of IR3, first a 31-bp insert containing *nic*-sequence and the whole IR3 palindrome with arms brought close together (*14+8hairpin*) was cloned into pUC18*.* Such plasmid was mobilized by RA3 with a frequency <10-7 indicating that the *14+8hairpin* insert was inactive as *oriT* (Fig. 2B bottom panel). This strongly suggested that the NIC\_RA3 processes *oriT* when the palindromic arms of IR3 are separated by 26 bp, as in the natural context of RA3, but not when they lie next to each other. *In vitro* studies on the NIC relaxase activity with the use of *14+8hairpin* ss oligonucleotide (as shown in Fig. 1D) confirmed that this sequence was indeed not susceptible to be cleaved by NIC relaxase (see below).

To understand the significance of the 26-bp spacer between IR3 arms, a 60-bp version of *oriT* (*oriT60hairpin*) and its various derivatives were analyzed (Fig. 2B.) Since this DNA fragment also determines a MobC binding site and *mobC* promoter activity (RNAP binding motifs -35 and -10), the constructed derivatives had either randomized OM or RNAP binding sites or both, but preserved the distance between intact IR3 arms. The modification of *mobC*p -35/ -10 motifs and OM in pUC18 *oriT60hairpin* OM*xmobC*px derivative abolished its ability to be mobilized by RA3 (transfer efficiency <10-7). This was in contrast to pUC18 *oriT31* that lacked OM and *mobC*p but also the IR3b arm and was transferred with an efficiency 10-2. Thus, somehow the presence of intact IR3 is detrimental for NIC activity when the binding sites for MobC and/ or RNAP are inactivated. Mobilization was fully restored in pUC18 *oriT60hairpin mobC*p*x* in which OM was left intact (Fig. 2B) and partially restored in pUC18 *oriT60hairpin* OMx (transfer efficiency 6x10-4), in which RNAP binding motifs were preserved. This data suggests that the negative impact of the IR3b arm on NIC binding may be overcome by the binding of MobC in the spacer region between the IR3 arms or, to lesser extent, by RNAP interactions.

Moving IR3b 5 bp closer to IR3a in pUC18 *oriT55hairpin* had no effect on the frequency of transfer as long as OM was present (compare *oriT55hairpin* and *oriT55hairpin* OMx in Fig. 2B).

It was noticed previously that introduction of a three-nucleotide substitution into the *nic*-proximal part of IR3a (GCAATTT→*tag*ATTT; IR3aX) within the 417-bp *oriT* fragment (Godziszewska *et al*., 2014) decreased mobilization of pJSB7.18 plasmid by RA3 more than 100-fold in comparison to the plasmid with the wt insert. Now we have observed that truncation of IR3a by two *nic*-distal nucleotides in the *oriT25* insert led only to a 10-fold decrease in mobilization in comparison to *oriT31* with the intact IR3a (Fig. 2A).

To define the role of particular nucleotides in IR3a, a series of o*riT31*s with single substitutions in IR3a at positions C1G2T3T4A5A6A7 (numbering indicates distance from the cleavage site in the bottom strand in Fig. 1C) were cloned into pUC18 and tested for mobilization by RA3 (Fig. 3). The modified *oriT31*X with three substitutions in IR3a (C1-T3, variant IR3aX tested previously in *oriT417*, Godziszewska *et al*., 2014) was included in the analysis as well as the wt *oriT31* insert as a control. pUC18 *oriT31*X and three derivatives with a single change at positions C1, G2 or T3 were not mobilized by RA3 at a detectable level (frequency <10-7), while substitutions in other positions only decreased the transfer frequency 5- to 12-fold in comparison to pUC18 *oriT31*. These results confirmed that the three IR3a nucleotides closest to the nick site are essential for RA3 mobilization.

***Regulation of* mobC*p by MobC and NIC***

The conserved nicksite and the IR3a sequence recognized by NIC\_RA3 relaxase are adjacent to the *mobC* promoter (50 bp upstream of the predicted transcription start site for *mobC*, Fig. 1C), which suggested that NICcould affect expression from *mobC*p. To check whether NIC regulates *mobC*pon its own or possibly assists MobC in the repression of *mobC*p,the following three-plasmid system was used*.* The promoter-probe vector pPT01 (Macartney *et al.*, 1997) was used to transcriptionally fuse the *mobC*p fragment (417 bp) with a promoter-less *xylE* cassette to obtain pJSB7.9 (Godziszewska *et al.*, 2014). To deliver MobC and NIC *in trans*, pJSB4.1(*tac*p*-mobC*RA3), a derivative of the medium-copy-number expression plasmid pAMB39 based on pBBR1MCS-1, and pJSB5.3 (*tac*p*-nic*RA3), a derivative of the high-copy-number expression vector pGBT30, were used. The expression of *mobC*p (monitored by spectrophotometric assay of XylE activity) was analyzed in the triple transformants grown without IPTG induction. It had previously been shown (Godziszewska *et al.*, 2014) that MobC is a potent repressor and even under such conditions the leaky character of *tac*pprovides a sufficient amount of MobC to cause a 10-fold repression of *mobC*p.

*E. coli* DH5(pJSB7.9.*mobC*p*-xylE*) was transformed with four pairs of plasmids: pGBT30/ pAMB39 (empty expression vectors as a control); pJSB4.1 *tac*p*-mobC/* pGBT30, pAMB39/ pJSB5.3 *tac*p*-nic*, or pJSB4.1 *tac*p*-mobC/* pJSB5.3 *tac*p*-nic.* The extracts from logarithmically growing cultures of these four strains were assayed for XylE activity. These experiments indicated that MobC but not NIC is the repressor of *mobC*p(Fig. 4A, top panel). Although on its own NIC had no effect on the transcriptional activity of *mobC*p, it potentiated up to 10-fold the repression exerted by MobC. The differences in XylE activity in the presence of MobC *versus* MobC and NIC were significant according to Student’s paired two tailed t-test (p<0.01).

From a collection of mutant derivatives of pJSB7.9 plasmid (Godziszewska *et al.*, 2014) we chose two variants, pJSB7.17 having a two-nucleotide substitution in the MobC binding site OM (further described as OMxx) and pJSB7.18 with a three-nucleotide substitution in the IR3a sequence tested in *oriT31*C1-T3 (further described as IR3ax). It has previously been demonstrated that the 417-bp fragment carrying OMxx is bound by MobC with a 10-fold lower affinity *in vitro* and shows a 15-fold lower MobC repression index for *mobC*p *in vivo*, but exhibits an unaffected mobilization frequency of pJSB7.17 in comparison to pJSB7.9 (Godziszewska *et al.*, 2014). In contrast, the IR3ax presence in the 417-bp fragment that decreased the mobilization frequency of pJSB7.18 in comparison to pJSB7.9, affected neither the MobC binding affinity *in vitro* nor *in vivo*. Expression of the mutant versions of *mobC*p-*xylE* transcriptional fusions was analyzed in the presence of combinations of expression vectors as for wt *mobC*p-*xylE* fusion.

The *mobC*pOMxx variant produced only a 20% decrease of activity (measured by XylE level) in the presence of MobC (Fig. 4A middle panel) and a 40% decrease in the presence of both MobC and NIC proteins (not statistically significant differences according to Student’s paired two-tailed t-test, p-value>0.01). This result confirmed that NIC repression of *mobC*p relied on interactions with MobC bound at its operator site (OM). The mutated IR3ax in pJSB7.18 had no impact on the extent of repression by MobC on its own (Fig. 4A), and no statistically significant enhancement of the MobC repression by NIC was observed for this mutant. So it has been concluded that a cooperative repression of *mobC*p by MobC and NIC is observed only when the both proteins effectively interact with DNA at their specific binding sites.

To verify these observations triple transformants’ cultures were grown as for the XylE assays and used for RNA isolation. RT-qPCR with pairs of primers for *xylE* and the *repA* gene of the promoter-probe vector, was then performed. In general the relative level of *xylE* transcript versus *repA* transcript confirmed the previous enzymatic assays (Fig. 4B). A significant enhancement of MobC repression by NIC was observed only in wt *mobC*p (pJSB7.9) when both proteins had their binding sites intact. It was noticeable that the presence of a high-copy-number plasmid with *tac*p-*nic* even when uninduced had a strong effect on the cells when MobC was not delivered *in trans* and led to the variations in relative transcript levels. The changes observed in expression of *mobC*p in the presence of NIC alone seemed to be non-significant according to Student’s t-test (p>0.01).

***Characteristics of NIC\_RA3 activities* in vitro**

MOBP represents the largest relaxase family (Garcillán-Barcia *et al*., 2009) with several distinguishable clades, among them MOBP4 which includes the IncU plasmids. Studies on the RA3 relaxosome thus help to characterize the MOBP4 group. In order to better understand the role of the auxiliary protein MobC in NIC\_RA3 activity, we reproduced DNA processing by the NIC\_RA3relaxase *in vitro*.

***a/ DNA recognition and cleavage activity of NIC\_RA3***

The NIC\_RA3 relaxase was overproduced in *E. coli* BL21 (DE3)(pJSB6.3 T7p-*nic*) and purified as a His6-NIC hybrid protein. Since even low concentrations of IPTG inhibited culture growth, the yield of purified NIC protein was low. The purified His6-NIC was tested for DNA binding and nicking activity with a Cy5-labeled 61-nt oligonucleotide (Fig. 1D) used as ss *oriT* or ds *oriT* after annealing to the complementary unlabeled strand.

In the EMSA experiments (Fig. 5A, B & C) His6-NIC bound both substrates, ss *oriT* and ds *oriT*, with a similar affinity, *K*app about 450 nM(*K*apparent – concentration of the protein at which 50% of the DNA fragment is bound). A single NIC- ds *oriT* complex was formed (Fig. 5B) whereas two complexes of different mobility were seen with ss *oriT*RA3 probably due to the binding of NIC to the original 61-nt oligonucleotide and a 36-nt fragment after cleavage (Fig. 5A). No interactions of His6-NIC with non-specific ssDNA were detected (Fig. 5C).

Cy5-labeled ssDNA containing *oriT*RA3 (61 nt) was also used to directly test the cleavage activity of His6-NIC. Under conditions used (Experimental Procedures), the 61-nt oligonucleotide was cut with low efficiency into two fragments, of which only the 36-nt one retained the 5’ label and was detected on the gel (Fig. 5D). No cleavage of the corresponding ds fragment was detected (data not shown). Since the N-terminal His-tag could have been responsible for the observed low enzymatic activity of His6-NIC, the *nic* gene was cloned into pET29c to produce NIC-His6. The C-terminally tagged NIC-His6 was efficiently over-produced in *E coli* C41 without a noticeable toxic effect. No difference in the cleavage activity of NIC-His6 on ssDNA *oriT* was observed in comparison to His6-NIC (Fig. 5E). To check if MobC would enhance the NIC activity, His6-MobC was over-produced, purified as described previously (Godziszewska *et al.*, 2014) and added to the NIC reaction mixture. No significant enhancement of the NIC cleavage activity *in vitro* was detected in the presence of His6-MobC (Fig. 5E).

***b/covalent complexes of NIC with cleaved oligonucleotides***

Conjugative relaxases cleave DNA in a site- and strand-specific manner by a transesterification reaction mediated by a tyrosine residue in the presence of divalent metal ions (mainly Mg2+). After cleavage the relaxase remains covalently bound to the 5’-end of the cleaved DNA strand via a phosphotyrosyl linkage. For better visualization of the NIC cleavage reaction we decided to analyse the intermediate products - complexes of NIC covalently bound to 5’-end of the cleaved oligonucleotide.

NIC-His6 tagged and untagged NIC were efficiently over-produced in *E coli* C41 and purified by sequential column chromatography and assayed forcleavage/ binding activity. Three different ss oligonucleotides similar to those successfully used for TrwC analysis were designed (Lucas *et al*., 2010). They all contained *nic* sequence (Fig. 1D) and IR3a, the *nic*-proximal arm of IR3 palindrome. The oligonucleotides designated *14+8* and *14+18* had additional sequences of different lengths downstream of the cleavage site. The third oligonucleotide *14+18hairpin* (also tested *in vivo* in mobilization experiments, Fig. 2B) contained the IR3b arm added next to IR3a (Fig. 1D). SDS-PAGE analysis of reaction mixtures comprising the NIC or NIC-His6 relaxase, oligonucleotides and Mg2+ revealed slower-migrating bands corresponding to NIC with covalently bound DNA substrate after nicking reaction revealing the same substrate specificity of the two NIC forms. They poorly cleaved oligonucleotides 14+8 and 14+18 and did not process the 14*+*18*hairpin* oligonucleotide (data not shown).

When different divalent cations were tested in the binding-cleavage reactions the activities of both NIC and NIC-His6 were significantly enhanced by Mn2+ or Ni2+ presence in the reaction mixture in comparison to Mg2+ (Fig. 6A).

To find the structural differences between NIC and other HUH relaxases that would explain their differential cation requirements, we performed structural prediction and modeling of NIC on the Protein Homology/analogY Recognition Engine (PHYRE) Web Server. As expected, NIC was found to be a homolog of RSF1010/ R1162 MobA (Monzingo *et al*., 2007) and other HUH relaxases such as TrwC (Guasch *et al.*, 2003) and TraI (Datta *et al*., 2003). A 3D model of NIC was generated by homology modeling using R1162\_miniMobA (PDB access 2SN6) as a template. Residues 42-186 of NIC could be modelled at >90% accuracy (Fig. 6C), although the sequence identity between NIC and miniMobA is only 25% (Fig. S4). The residues involved in the Tyr nucleophilic attack and the HUH motif shown to participate in metal coordination in members of this family of proteins are conserved (Boer *et al*., 2006; Larkin *et al*., 2005; Monzingo *et al*., 2007). However, in NIC\_RA3, the residue potentially involved in the His activation for cation coordination is different than in the other studied homologs. Instead of an acidic residue (e.g., E74 in MobA of R1162, Fig. S2), NIC contains the hydrophobic residue V108 unable to stabilize the His coordination conformation and thus supposedly Mg2+ cannot be coordinated efficiently by NIC\_RA3.

Two mutant *nic* alleles were constructed to produce relaxases with amino acids substitutions V108A or V108E, respectively. Purified C-terminally His-tagged NIC variants were tested for cleavage and covalent binding of oligonucleotide 14+18 in the presence of Mn2+ or Mg2+ (Fig. 6D and E). Both variants demonstrated similar cleavage/ covalent binding activity in the presence of 10 mM Mn2+ as WT NIC-His6 but differed in the activity in the presence of 10 mM Mg2+ (Fig. 6D). Quantification of ssDNA-NIC complex formation showed that the WT NIC-His6 activity in the presence of Mg2+equaled 45% of that with Mn2+. The V108A substitution increased the activity with Mg2+ to 85%, while the V108E one to 115% of the activity observed of WT NIC-His6 with Mn2+. The different activities of the two NIC mutant variants are even more pronounced at a lower Mg2+ concentration of 1 mM (Fig. 6E).

The *nicV108A* and *nicV108E* alleles were also introduced into pBGS18 carrying a complete conjugative module (pJSB1.83 and pJSB1.84, respectively) and tested for conjugative transfer between *E. coli* strains. No difference was noticed in the transfer frequency (1 transconjugant per donor) between a plasmid producing WT NIC\_RA3 and its mutated versions NICV108A and NICV108E confirming that the modified proteins were fully active *in vivo*.

***c/ Nicking of supercoiled plasmid DNA* in vitro**

Previous studies on the role of MobC in conjugation (Godziszewska *et al.*, 2014) have demonstrated that deletion of *mobC* leads to a 1000-fold decrease in the frequency of transfer of a plasmid with the RA3 conjugative module. However, no effect of MobC on the cleavage of *oriT* oligonucleotide (ss or ds) by NIC\_RA3 was detected in the present study using an *in vitro* assay (Fig. 5D). To clarify this apparent discrepancy, the *in vitro* activity of the NIC-His6 relaxase in the presence or absence of His6-MobC was analyzed using supercoiled (SC) pUC18 with 417-bp *oriT*RA3 (pJSB2.9). As a control, supercoiled pUC18 without the insert was used (Fig. 7A). Neither NIC-His6 nor His6-MobC alone affected the balance between the different topological forms of pJSB2.9. However, NIC-His6 and His6-MobC together produced a shift from the supercoiled (SC) towards the open-circle (OC) form of pJSB2.9 whereas pUC18 superhelicity was unaffected (Fig. 7A). These results show that MobC is required for the *oriT-*specific nicking activity of NIC\_RA3 towards supercoiled DNA.

The pUC18 derivatives with different *oriT* variants previously analyzed for mobilization ability were then used in the nicking assay (Fig. 7B). pJSB2.43 *oriT61*, pJSB2.44 *oriT45*, and pJSB2.50 *oriT31* were also nicked efficiently by NIC-His6 only in the presence of His6-MobC. The relaxation of pJSB2.44 and pJSB2.50, both with OM deleted, confirmed the earlier observation that MobC binding to DNA is not required to enhance nicking by NIC. The 31-bp insert with three nucleotides in the IR3a arm modified (*oriT31*C1-T3, pJSB2.54) did not allow nicking of the plasmid under conditions used, as expected from the mobilization assay.

**Discussion**

The conjugative transfer system of the IncU plasmid RA3 represents one of the most efficient, broad-host-range HGT processes. The frequency of the RA3 transfer between *E. coli* and representatives of alpha-, beta- and other gammaproteobacteria is as high as one transconjugant per donor cell in 2-h mating. On the basis of the amino acid sequence of the NIC relaxase, the RA3 conjugative transfer system has been classified into the MOBP family (Garcillán-Barcia *et al*., 2009). The widely occurring MOBP relaxases belong to the HUH endonuclease superfamily with a single Tyr residue in the active center (Pansegrau *et al*., 1993) and are clearly split into several clades designated MOBP1 to MOBP12 (Garcillán-Barcia *et al.*, 2009). Representatives of the IncU group, e.g., RA3 (Kulinska *et al.*, 2008), pFBAOT6 (Rhodes *et al*., 2004), pMBU13 and pMBU17 (Brown *et al*., 2013) constitute the MOBP4 clade together with the PromA group of plasmids (van der Auwera *et al*., 2009). Besides sequencing data, no information about relaxases of the MOBP4 group is available. Thus, in this work we characterized a range of NIC activities and the role of the auxiliary protein MobC in their stimulation.

The fully active minimal *oriT* of RA3 was delineated to 47 bp containing the *nic* sequence, a 7-bp arm of the IR3 palindrome (IR3a) and the MobC binding site OM. Constructs with *oriT* deprived of OM were mobilized by RA3 with a 100-fold lower frequency. These results suggested that by binding at OM the transfer auxilliary MobC protein could enhance the NIC\_RA3 activity at *oriT*. However, *in vitro* experiments showed that MobC was required not only to induce NIC cleavage of supercoiled plasmid DNA with *oriT417* or *oriT61* but also with *oriT45* and *oriT31* with OM deleted. In conclusion, MobC mainly stimulates the NIC activity through direct protein-protein interactions but, when bound to OM, it may further assist the NIC\_RA3 positioning at *oriT*. This concerted action could be particularly important under natural conditions of a limited amount of the relaxase produced by RA3, as exemplified by the mobilization experiments of the pUC18 derivatives. The interactions between NIC and MobC will be studied further.

It was shown here that IR3a is a NIC\_RA3 recognition site. It is a part of IR3, an atypical inverted repeat with the two arms separated by 26 bp and in this way enclosing the OM and *mobC* promoter motifs. Single-substitution mutant studies clearly demonstrated that not all IR3a nucleotides are equally important for the NIC binding, indicating a critical role of the three nucleotides CGT proximal to the nick site.

In the *in vitro* cleavage/ covalent binding tests only oligonucleotides containing the single nick-proximal arm of IR3, IR3a, were processed by NIC\_RA3 whereas the 14+8*hairpin* containing the both arms of the palindrome without an intervening sequence was completely refractory to the cleavage/ binding reaction. Also *in vivo* this palindromic structure apparently precluded NIC binding and *oriT* processing since the 14+8*hairpin* insert failed to direct plasmid mobilization by RA3.

Those results were unexpected since in other *oriT*s the presence of an inverted repeat in the same position relative to the *nic* sequence as in the 14+8*hairpin* is critical for their activity, the *nic*-proximal arm of the palindrome being required for transfer initiation and the nick-distal arm for hairpin formation in the recipient and completion of the transfer (Carballeira *et al*., 2014; Luo *et al*., 1995; Williams and Schildbach, 2006). The inhibitory effect of IR3b on NIC binding at IR3a was absent in pUC18*oriT60hairpin* when the arms of the palindrome were separated by the native 26-nt spacer bearing OM and RNAP binding motifs (*mobC*p). Randomization of the 26 nt between IR3a and IR3b to eliminate binding of MobC and RNAP, abolished mobilization while the presence of intact -35 and -10 motifs in the 26-nt spacer partially restored the mobilization ability. Full mobilization activity was exhibited by all constructs where OM was left intact. This suggests steric hindrance by proteins bound in the spacer region (with a major role of MobC) as a means of protecting IR3a from negative interference by IR3b and facilitating NIC cleavage of *oriT*. Such interference may have an important regulatory function during establishment of RA3 plasmid in the new host when repressors like MobC are not yet synthesized. Further studies are needed to verify this hypothesis and to elucidate how NIC\_RA3 overcomes the requirement for hairpin formation in the recipient to complete the processing of the transferred DNA in all constructs deprived of IR3b.

The NIC\_RA3 binding site, IR3a, overlaps -35 motif of *mobC*p. The regulatory studies presented here have revealed an important role of NIC in potentiating the repression of *mobC*p by MobC, making it a tightly autoregulated circuit. A similar autoregulatory role of a relaxase MobA in conjunction with an auxiliary protein has been demonstrated for p1-p3 promoters in the *oriT* of RSF1010/R1162 (Frey *et al*., 1992).

Both NIC relaxase and auxiliary MobC protein participate in two seemingly self-excluding processes: initiation of transfer and limiting their own synthesis. To fully succeed in any of these processes they both should be bound to specific sites in DNA. Further studies are required to establish the signals for the fine tuning of the switching on/off of the RA3 conjugation process.

Our *in vivo* studies were complemented by *in vitro* analysis. It showed that NIC binds linear dsDNA as well as ssDNA and does not need MobC for the cleavage of such oligonucleotides.

Experiments with purified NIC have demonstrated that DNA cleavage and covalent binding of the cleaved oligonucleotide proceeds more efficiently in the presence of Ni2+ or Mn2+ rather than Mg2+. This is in stark contrast to, e.g., TraI relaxase of F plasmid which preferentially uses Mg2+ as cofactor (Datta *et al*., 2003; Larkin *et al*., 2005; Lujan *et al*., 2007). A 3D homology model of NIC with R1162\_miniMobA (PDB access 2SN6) as a template confirmed the conservation of residues involved in the Tyr nucleophilic attack on DNA and of the HUH motif participating in metal coordination, but the residue involved in His activation for the cation coordination was different in NIC\_RA3 than those in MobA and other analyzed relaxases (Boer *et al*., 2006; Larkin *et al*., 2005; 2007; Monzingo *et al*., 2007). Whereas the anionic residues D85 in R388 TrwC, D81 in F plasmid TraI and E74 in R1162 MobA (Fig. S2) are involved not only in the deprotonation of the active Tyr, but also in stabilization of the His for metal ion coordination, the NIC\_RA3 equivalent V108 should not be able to help Mg2+ to be efficiently coordinated. After changing this hydrophobic residue for Glu, NICV108E variant showed a restored cleavage activity in the presence of Mg2+. The V108A mutant also showed better Mg2+-cleavage activity than WT NIC although it was less active than V108E. Thus, the presence of valine avoids Mg2+-dependent cleavage not only because of the lack of H-bond with H152, but also by stereo chemical impediment. Notably, NES of pLW1043 (Edwards *et al.*, 2013) and MobM of pMV158 (Lorenzo-Diaz *et al.*, 2011) reported to be more active in the presence of Mn2+ or Ni2+ than Mg2+ also contain a hydrophobic residue in the position equivalent to V108 of NIC\_RA3 (Fig. S3). Moreover, a hydrophobic residue is found in all other MOBP4 relaxases instead of an acidic residue (Fig. S1). If our model is correct this suggests similar requirements for metal cofactors of these proteins.

This study defined the minimal requirements for RA3 conjugation demonstrating that RA3 is one of the genetically simplest conjugative systems found in nature. It also specified the biochemical properties of NIC\_RA3, the representative of MOBP4 subgroup. The interplay between NIC relaxase and its auxilliary protein MobC may provide the fine tuning of conjugative process by controlling the level of relaxase and its activity.

**Experimental Procedures**

***Bacterial strains and growth conditions***

*Escherichia coli* strains used were DH5 [F*-*(*80*d*lacZM15*) *recA1 endA1 gyrA96 thi-1 hsdR17*(*rk-mk+*) *supE44 relA1 deoR (lacZYA-argF)U196*] and its RifR derivative, BL21(DE3) F- *ompT* *hsdS*B (rB-mB-) *gal dcm* (DE3) (Novagen Inc) and C41(DE3) derivative of BL21(DE3) suitable for overproduction of toxic proteins (Novagen Inc). Bacteria generally were grown in L broth (Kahn *et al*., 1979) at 37ºC or on L agar (L broth with 1.5% w/v agar) supplemented with appropriate antibiotics: benzyl penicillin, sodium salt (150 g ml-1 in liquid media and 300 g ml-1 in agar plates) for penicillin resistance, kanamycin 50 g ml-1 for kanamycin resistance and chloramphenicol 10 g ml-1 for chloramphenicol resistance. L agar used for blue/white screening contained 0.1 mM IPTG and X-gal (40 μg ml-1).

***Plasmid DNA isolation, analysis, cloning and manipulation***

Plasmid manipulations were carried out by standard procedures (Sambrook *et al.*,1989). All plasmids used in this study are listed in Table 1. Standard PCR (Mullis *et al*., 1986) was performed with pairs of primers listed in Table S1. PCR derived clones were verified by sequencing at the DNA Sequencing and Oligonucleotide Synthesis Laboratory, Institute of Biochemistry and Biophysics, using dye terminator sequencing and an ABI 377 Perkin Elmer automated sequencer.

Complementary oligonucleotides encoding variants of *oriT* region were annealed and cloned into SmaI cut pUC18 vector. Two expression vectors, medium-copy number pAMB9.37 based on pBBR1MCS-1 of IncA/C group (Kovach *et al.*, 1994) and the high-copy number vector pGBT30 (Jagura-Burdzy *et al*.,1991), based on the pMB1 replicon were used for allele cloning in the regulatory studies, whereas for overproduction and purification the *nic* orf was re-cloned into pET28a, pET29c or pET3a vectors (Novagen Inc).

***Site-directed mutagenesis in vitro***

To introduce mutations in the *nic* gene an enzymatic assembly of DNA molecules was used (Gibson *et al.*, 2009). Complementary mutagenic primers for *nicV108A* introduced KpnI restriction site whereas mutagenic primers for *nicV108E* introduced SacI restriction site for screaning. Each mutagenic primer was paired with a primer annealing to the vector sequence and used in standard PCR reactions to amplify fragments with ends overlapping required vector sequence (pET29c or pBGS18) and *nic* gene. Linearized vector (pET29c or pBGS18) and two appropriate PCR fragments introducing *nic* mutation were assembled by use of NEBuilder HiFi DNA Assembly Master Mix (NEW ENGLAND BioLabs). Plasmid DNA from transformants was screened for the presence of the restriction site introduced by the mutagenic primers and then introduced modification was verified by DNA sequencing.

***Bacterial transformation***

Competent cells of *E. coli* were prepared by the standard CaCl2 method (Sambrook *et al*., 1989).

***Conjugation procedure***

*E. coli* DH5R was transformed with pUC18 derivatives carrying different parts of *oriT*RA3 and such double transformants were used as donors in conjugation with DH5 RifR as a recipient. Overnight cultures of the donor and recipient strains (100 l each) were mixed on L agar plates and incubated for 2 hours at 37ºC. Cells were scrapped, re-suspended in L broth and aliquots of serial 10-fold dilutions were plated onto L agar plates with 100 μg ml-1 rifampicin and 300 μg ml-1 penicillin to estimate the number of transconjugants. In parallel 100 l of donor strain culture was incubated on L agar plate for 2 hours at 37ºC and treated the same way to establish the titer of donor cells in the conjugation mixture. The transfer frequency was calculated as the number of transconjugants per donor cell.

***Determination of catechol 2,3-dioxygenase activity (XylE)***

XylE activity (the product of *xylE*) was assayed spectrophotometrically in the cleared extracts from logarithmically growing cultures (Zukowski *et al*., 1983). One unit of catechol 2,3-dioxygenase activity is defined as the amount of enzyme needed to convert 1 μmol of catechol to 2-hydroxymuconic semialdehyde in 1 minute per mg of protein. Protein concentration was determined using the Bradford method (Bradford, 1976).

***RNA isolation and RT-qPCR analysis***

Three independent cultures of each analyzed strain (transformants of DH5 used in the regulatory *xylE* assays) were inoculated and grown overnight in L broth at 37oC supplemented with appropriate antibiotics. Cultures were diluted 1:100 into fresh L broth and propagated with shaking at 37oC. 2 ml aliquots of each culture in logarithmic phase of growth (OD600=0.5-0.6) were subjected to RNA extraction.

Total RNA was isolated from three cultures of each strain (biological replicates) using an RNeasy mini-kit (Qiagen) according to the manufacturer’s instructions. The DNase treatment of total RNA using TURBO DNase kit (Ambion) was applied to eliminate DNA contamination. The RNA concentration was estimated using Nano Drop ND-1000 Spectrophotometer.

The RNA samples were used for RT-qPCR analysis. Three biological replicates of total RNA (6 μg) from each strain served as templates for cDNA synthesis with SuperScipt VILO Master Mix reverse transcriptase (Invitrogen). The cDNAs were purified using QiaQuick PCR purification Kit (Qiagen) and then used as templates in qPCR performed with 5x HOT FIREPol EvaGreen qPCR Mix Plus (Solis BioDyne). The specifically designed qPCR primers (Table S1) were used to amplify reference *repA* gene of pSC101 vector and target *xylE* gene. Before use, primers were tested for equal efficiency of the qPCR reactions. Only efficiency values of about 0.95 or more were accepted.

Each PCR reaction contained 4 l 5x HOT FIREPol EvaGreen qPCR Mix Plus (Solis BioDyne), 2 l of diluted cDNA, and each of the forward and reverse primers at 0.6 M; nuclease-free water was added to obtain a final volume of 20 l. qPCR was performed using the Light Cycler 480 (Roche). PCR products were amplified according to the following protocol: one cycle at 95°C for 15 min, followed by 40 cycles at 95°C for 15 s, 60oC for 20 s and 72°C for 20 s. The melting curve was 65 to 95°C with increments of 0.5°C/s. In each run, a negative control (no cDNA) and a standard curve to calculate the efficiency of the quantitative PCR reaction for each primer set were included.

30 ng of cDNA from each biological replicate were analyzed in three technical replicates. Changes in *xylE* reporter gene expression between control strain with empty expression vectors and strains producing either NIC, MobC or NIC/ MobC were calculated with normalization of Ct values to mean Ct value for reference plasmidic gene *repA*. Relative quantification was performed by the ΔΔCt method (Livak and Schmittgen, 2001; Pfaffl, 2001).

***Over-expression and purification of MobC and NIC***

His6-MobC was purified as described before (Godziszewska *et al*., 2014). Since over-production of His6-NIC in the *E. coli* strain BL21(DE3)(pJSB6.3) was detrimental for the growth the standard IPTG induction was modified (Supplementary Methods). Alternatively two other forms of NIC: untagged and C-terminally tagged NIC-His6, or mutated NICV108A-His6 and NICV108E-His6 were efficiently over-produced in *E coli* C41 cells and purified as described in the Supplementary methods. Plasmid pET3a:*nic* carries the *nic* gene cloned between the NdeI and BamHI sites of vector pET3a whereas pET29c:*nic-*His6, pET29a:*nicV108A-*His6,pET29a:*nicV108E-*His6 carry the *nic* alleles without stop codon and cloned between the NdeI and XhoI sites of the vector.

***Analysis of protein-DNA interactions by electrophoretic mobility shift assay (EMSA)***

The substrates used in the reactions were either Cy5 labeled oligonucleotide (#31, #32) or ds oligonucleotides (annealed primers #31 and #11; Table S1). The protein-DNA binding reactions were performed in the final volume of 20 μl of the binding buffer (50 mM Tris-HCl pH 8.0; 10 mM MgCl2; 50 mM NaCl; 0.1 mg ml-1 BSA), and increasing amounts of His6-NIC or NIC-His6 added. Complexes were analyzed on a native 10% polyacrylamide gel in 0.5 TBE buffer (45 mM Tris-borate, 1 mM EDTA). The gels were run at 100 V for 1.5 h and DNA was visualized using FluorchemQ MultiImage III camera (Alpha Innotech) and Alpha Innotech – Alpha View program.

***Assays of NIC relaxase nicking activity***

***A/ Nicking activity on supercoiled plasmid DNA***

Reaction mixtures (final volume 20 μl) contained 100 ng of supercoiled DNA of pUC18 or its derivatives, 10 or 20 pmole of purified NIC-His6 and/or 17.5 pmole of His6-MobC in the nicking buffer (20 mM Tris-acetate, pH 7.9, 50 mM potassium acetate, 10 mM magnesium acetate, 10 mg ml-1 BSA). After 1.5 h at 37°C a proteinase K (5 mg ml-1) was added together with SDS (0.5% w/v) and incubated for 30 min at 37°C. Products of nicking reactions were separated on 1% agarose gels, stained with the ethidium bromide and visualized under UV light.

***B/ Analysis of ssDNA cleavage products***

The 2.5 pmoles of Cy5-labeled oligonucleotide #31 (Table S1) were used in the cleavage reaction conducted with different amounts of His6-NIC protein (0, 4.5, 9, 19 pmole, respectively) in the nicking buffer for 1.5 h at 37°C. To some reaction mixtures 17.5 pmole of His6-MobC was added. The proteins were removed by incubation with proteinase K (final concentration 5 mg ml-1) and 0.5% (w/v) SDS for 30 min at 37°C. The products were analyzed on a native 10% polyacrylamide gel in 0.5 TBE buffer (45 mM Tris-borate, 1 mM EDTA). The conditions of DNA separation and visualization were as for EMSA experiments.

***C/ ssDNA-NIC complex formation***

DNA substrates (Fig. 1D) were commercially synthesized by Sigma-Aldrich (Spain). Each substrate was resuspended in miliQ water, heated to 95ºC for 10 min, and then either allowed to cool passively to room temperature or snap cooled on ice. Cleavage was carried out by incubating 1 mg ml-1 NIC (or NIC-His6) with 15 M of each oligonucleotide at 37ºC for one hour in 10 mM Tris-HCl, pH 7.6, 0.01-10 mM MgCl2 (as indicated), 375 mM NaCl and 15 M EDTA. MgCl2 was substituted by MnCl2 or NiCl2 when indicated. The reaction was stopped by adding SDS and heating. The cleavage activity was checked by the lower mobility of the Protein-DNA covalent complexes on 10% or 12% acrylamide gels analyzed by SDS-PAGE. There was always a molar excess of oligonucleotides to guarantee that all the protein is in complex with some oligonucleotide.

**Acknowledgements**

Work in GJB laboratory was financed by the National Science Centre in Poland (grant 2011/03/B/NZ1/06540) and bygrant MNiSW 7793/B/P01/2011/40. Work in the FdlC and GM laboratories was financed by the Spanish Ministry of Economy and Competitivity (BFU2014-55534-C2) and by the European Seventh Framework Program (project 612146/FP7-ICT-2013-10 (PLASWIRES)).

**Author contributions**

GJB and FC contributed in the conception and design of the study. JG, GM, MC and AAB participated in data acquisition, analysis and presentation. GJB, FC and GM wrote the manuscript. All authors contributed in editing and revising the manuscript. All authors approved the content of the article.

All authors declare no conflict of interest.

**References**

Boer, R., Russi, S., Guasch, A., Lucas, M., Blanco, A.G., Pérez-Luque, R., *et al.* (2006) Unveiling the molecular mechanism of a conjugative relaxase: The structure of TrwC complexed with a 27-mer DNA comprising the recognition hairpin and the cleavage site. *J Mol Biol* **358**: 857–869.

Bowie, J.U., and Sauer, R.T. (1990) TraY proteins of F and related episomes are members of the Arc and Mnt repressor family. *J Mol Biol* **211**: 5–6.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram 486 quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–254.

Brown, C.J., Sen, D., Yano, H., Bauer, M.L., Rogers, L.M., Auwera, G.A. Van der, and Top, E.M. (2013) Diverse broad-host-range plasmids from freshwater carry few accessory genes. *Appl Environ Microbiol* **79**: 7684–7695.

Byrd, D.R. (2002) Structure-function analysis of *Escherichia coli* DNA helicase I reveals non-overlapping transesterase and helicase domains. *J Biol Chem* **277**: 42645–42653.

Caryl, J.A., Smith, M.C.A. and Thomas, C.D. (2004) Reconstitution of a staphylococcal plasmid-protein relaxation complex *in vitro*. *J Bacteriol* **186**: 3374–3383.

Caryl, J.A., and Thomas, C.D. (2006) Investigating the basis of substrate recognition in the pC221 relaxosome. *Mol Microbiol* **60**: 1302–1318.

Carballeira, J.D., Gonzalez-Perez, B., Moncalian, G., and de la Cruz, F. (2014) A high security double lock and key mechanism in HUH relaxases controls *oriT*-processing for plasmid conjugation. *Nucleic Acids Res* **42**: 10632–10643.

Chandler, M., Cruz, F. de la, Dyda, F., Hickman, A.B., Moncalian, G., and Ton-Hoang, B. (2013) Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nat Rev Micro* **11**: 525–538.

Dash, D.K., Traxler, B.A., Panicker, M.M., Hackney, D.D., and Minkley, E.G. Jr. (1992) Biochemical characterization of *Escherichia coli* DNA helicase I. *Mol Microbiol* **6**: 1163–1172.

Datta, S., Larkin, C., and Schildbach, J.F. (2003) Structural insights into single-stranded DNA binding and cleavage by F factor TraI. *Structure* **11**: 1369–1379.

de la Cruz, F., Frost, L.S., Meyer, R.J., and Zechner, E.L. (2010) Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol Rev* **34**: 18–40.

Draper, O., Cesar, C., Elvira, Machon, C., de la Cruz, F., and Llosa, M. (2005) Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. *PNAS* **102**: 16385–16390.

Edwards, J.S., Betts, L., Frazier, M.L., Pollet, R.M., Kwong, S.M., Walton, W.G., *et al.* (2013) Molecular basis of antibiotic multiresistance transfer in *Staphylococcus aureus*. *Proc Natl Acad Sci* **110**: 2804–2809.

Francia, M.V., Varsaki, A., Garcillán-Barcia, M.P., Latorre, A., Drainas, C., and de la Cruz, F. (2004) A classification scheme for mobilization regions of bacterial plasmids. *FEMS Microbiol Rev* **28**: 79–100.

Frey, J., Bagdasarian, M.M., and Bagdasarian, M. (1992) Replication and copy number control of the broad host-range plasmid RSF1010. *Gene* **113**: 101–106.

Fürste, J.P., Pansegrau, W., Ziegelin, G., Kröger, M., and Lanka, E. (1989) Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. *Proc Natl Acad Sci USA* **86**: 1771–1775.

Garcillán-Barcia, M.P., Francia, M.V., and de la Cruz, F. (2009) The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev* **33**: 657–687.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A. 3rd and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6:** 343-345.

Godziszewska, J., Kulińska, A., and Jagura-Burdzy, G. (2014) MobC of conjugative RA3 plasmid from IncU group autoregulates the expression of bicistronic *mobC-nic* operon and stimulates conjugative transfer. *BMC Microbiol* **14**: 235.

González-Prieto, C., Agúndez, L., Linden, R.M., and Llosa, M. (2013) HUH site-specific recombinases for targeted modification of the human genome. *Trends Biotechnol* **31**: 305–312.

Grandoso, G., Avila, P., Cayon, A., Hernando, M.A., Llosa, M., and de la Cruz, F. (2000) Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. *J Mol Biol* **295**: 1163–1172.

Grandoso, G., Llosa, M., Zabala, J.C., and de la Cruz, F. (1994) Purification and biochemical characterization of TrwC, the helicase involved in plasmid R388 conjugal DNA transfer. *Eur J Biochem* **226**: 403–412.

Guasch, A., Lucas, M., Moncalian, G., Cabezas, M., Perez-Luque, R., Gomis-Ruth, F.X., *et al.* (2003) Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nat Struct Mol Biol* **10**: 1002–1010.

Halary, S., Leigh, J.W., Cheaib, B., Lopez, P., and Bapteste, E. (2010) Network analyses structure genetic diversity in independent genetic worlds. *Proc Natl Acad Sci USA* **107**: 127–132.

Hekman, K., Guja, K., Larkin, C., and Schildbach, J.F. (2008) An intrastrand three-DNA-base interaction is a key specificity determinant of F transfer initiation and of F TraI relaxase DNA recognition and cleavage. *Nucleic Acids Res* **36**: 4565–4572.

Jagura-Burdzy, G., Ibbotson, J.P., and Thomas, C.M. (1991) The korF region of broad-host-range plasmid RK2 encodes two polypeptides with transcriptional repressor activity. *J Bacteriol* **173**: 826–833.

Kahn M, Kolter R, Thomas C, Figurski D, Meyer R, Remaut E, and Helinski DR (1979) Plasmid cloning vehicles derived from plasmids ColE1, F, R6K, and RK2. *Methods Enzymol* **68**: 268–280.

Kovach, M.E., Phillips, R.W., Elzer, P.H., Roop, R.M., and Peterson, K.M. (1994) pBBR1MCS: a broad-host-range cloning vector. *Biotechniques* **16**: 800–802.

Kulinska, A., Cao, Y., Macioszek, M., Hayes, F., and Jagura-Burdzy, G. (2011) The centromere site of the segregation cassette of broad-host-range plasmid RA3 is located at the border of the maintenance and conjugative transfer modules. *Appl Environ Microbiol* **77**: 2414–2427.

Kulinska, A., Czeredys, M., Hayes, F., and Jagura-Burdzy, G. (2008) Genomic and functional characterization of the modular broad-host-range RA3 plasmid, the archetype of the IncU group. *Appl Environ Microbiol* **74**: 4119–4132.

Kulinska, A., Godziszewska, J., Wojciechowska, A., Ludwiczak, M., Jagura-Burdzy, G. (2016) Global transcriptional regulation of backbone genes in broad-host-range plasmid RA3 from the IncU group involves segregation protein KorB (ParB Family). *Appl Environ Microbiol* **82**: 2320–2335.

Larkin, C., Datta, S., Harley, M.J., Anderson, B.J., Ebie, A., Hargreaves, V., and Schildbach, J.F. (2005) Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase. *Structure* **13**: 1533–1544.

Larkin, C., Haft, R.J., Harley, M.J., Traxler, B., Schildbach, J.F. (2007) Roles of active site residue and the HUH motif of the F plasmid TraI relaxase. *J Biol Chem* **282**:33707-33713.

Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real time quantitative PCR and the 2-ΔΔCt method. *Methods* **25:** 402–408.

Llosa, M., Grandoso, G., Hernando, M.A., and de la Cruz, F. (1996) Functional domains in protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. *J Mol Biol* **264**: 56–67.

Lorenzo-Diaz, F., Dostal, L., Coll, M., Schildbach, J.F., Menendez, M., and Espinosa, M. (2011) The MobM relaxase domain of plasmid pMV158: thermal stability and activity upon Mn2+ and specific DNA binding. *Nucleic Acids Res* **39**: 4315–4329

Lovett, M.A., and Helinski, D.R. (1975) Relaxation complexes of plasmid DNA and proteins II. Characterization of the proteins associated with the unrelaxed and relaxed complexes of plasmid ColE1. *J Biol Chem* **250**: 8790– 8795.

Lucas, M., Gonzalez-Perez, B., Cabezas, M., Moncalian, G., Rivas, G., and de la Cruz, F. (2010) Relaxase DNA binding and cleavage are two distinguishable steps in conjugative DNA processing that involve different sequence elements of the *nic* site. *J Biol Chem* **285**: 8918–8926.

Ludwiczak, M., Dolowy, P., Markowska, A., Szarlak, J., Kulinska, A., and Jagura-Burdzy, G. (2013) Global transcriptional regulator KorC coordinates expression of three backbone modules of the broad-host-range RA3 plasmid from IncU incompatibility group. *Plasmid* **70**: 131–145.

Lujan, S. A., Guogas, L. M., Ragonese, H., Matson, S. W., and Redinbo, M. R. (2007) Disrupting antibiotic resistance propagation by inhibiting the conjugative DNA relaxase. *Proc Natl Acad Sci U. S. A.* **104**: 12282–12287.

[Lum, P.L](http://www.ncbi.nlm.nih.gov/pubmed/?term=Lum%20PL%5BAuthor%5D&cauthor=true&cauthor_uid=12206773)., [Rodgers, M.E](http://www.ncbi.nlm.nih.gov/pubmed/?term=Rodgers%20ME%5BAuthor%5D&cauthor=true&cauthor_uid=12206773)., and [Schildbach, J.F](http://www.ncbi.nlm.nih.gov/pubmed/?term=Schildbach%20JF%5BAuthor%5D&cauthor=true&cauthor_uid=12206773). (2002) TraY DNA recognition of its two F factor binding sites. [*J Mol Biol*.](http://www.ncbi.nlm.nih.gov/pubmed/12206773) **321**: 563-578.

[Luo, Y](http://www.ncbi.nlm.nih.gov/pubmed/?term=Luo%20Y%5BAuthor%5D&cauthor=true&cauthor_uid=7596285), [Gao, Q](http://www.ncbi.nlm.nih.gov/pubmed/?term=Gao%20Q%5BAuthor%5D&cauthor=true&cauthor_uid=7596285)., and [Deonier, R.C](http://www.ncbi.nlm.nih.gov/pubmed/?term=Deonier%20RC%5BAuthor%5D&cauthor=true&cauthor_uid=7596285). (1995)Boundaries of the nicking region for the F plasmid transfer origin, *oriT*. [*Mol Microbiol* **15**: 829-837*.*](http://www.ncbi.nlm.nih.gov/pubmed/7596285)

Macartney, D.P., Williams, D.R., Stafford, T., and Thomas, C.M. (1997) Divergence and conservation of the partitioning and global regulation functions in the central control region of the IncP plasmids RK2 and R751. *Microbiology* **143**: 2167–2177.

Marques, M.V., Silva, A.M. da, and Gomes, S.L. (2001) Genetic organization of plasmid pXF51 from the plant pathogen *Xylella fastidiosa*. *Plasmid* **45**: 184–199.

Matson, S.W., and Ragonese, H. (2005) The F-Plasmid TraI protein contains three functional domains required for conjugative DNA strand transfer. *J Bacteriol* 187: 697–706.

[Moncalián, G](http://www.ncbi.nlm.nih.gov/pubmed/?term=Moncali%C3%A1n%20G%5BAuthor%5D&cauthor=true&cauthor_uid=10593894)., [Cabezón, E](http://www.ncbi.nlm.nih.gov/pubmed/?term=Cabez%C3%B3n%20E%5BAuthor%5D&cauthor=true&cauthor_uid=10593894)., [Alkorta, I](http://www.ncbi.nlm.nih.gov/pubmed/?term=Alkorta%20I%5BAuthor%5D&cauthor=true&cauthor_uid=10593894)., [Valle, M](http://www.ncbi.nlm.nih.gov/pubmed/?term=Valle%20M%5BAuthor%5D&cauthor=true&cauthor_uid=10593894)., [Moro, F](http://www.ncbi.nlm.nih.gov/pubmed/?term=Moro%20F%5BAuthor%5D&cauthor=true&cauthor_uid=10593894)., [Valpuesta, J.M](http://www.ncbi.nlm.nih.gov/pubmed/?term=Valpuesta%20JM%5BAuthor%5D&cauthor=true&cauthor_uid=10593894)., [Goñi, F.M](http://www.ncbi.nlm.nih.gov/pubmed/?term=Go%C3%B1i%20FM%5BAuthor%5D&cauthor=true&cauthor_uid=10593894)., and [de la Cruz, F](http://www.ncbi.nlm.nih.gov/pubmed/?term=de%20La%20Cruz%20F%5BAuthor%5D&cauthor=true&cauthor_uid=10593894). (1999a) Characterization of ATP and DNA binding activities of TrwB, the coupling protein essential in plasmid R388 conjugation. [*J Biol Chem*](http://www.ncbi.nlm.nih.gov/pubmed/10593894) **274**: 36117-24.

Moncalián, G., and de la Cruz, F. (2004) DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. *Biochim Biophys Acta* **1701**: 15–23.

[Moncalián, G](http://www.ncbi.nlm.nih.gov/pubmed/?term=Moncali%C3%A1n%20G%5BAuthor%5D&cauthor=true&cauthor_uid=10209739)., [Valle, M](http://www.ncbi.nlm.nih.gov/pubmed/?term=Valle%20M%5BAuthor%5D&cauthor=true&cauthor_uid=10209739)., [Valpuesta, J.M](http://www.ncbi.nlm.nih.gov/pubmed/?term=Valpuesta%20JM%5BAuthor%5D&cauthor=true&cauthor_uid=10209739)., and [de la Cruz, F](http://www.ncbi.nlm.nih.gov/pubmed/?term=de%20la%20Cruz%20F%5BAuthor%5D&cauthor=true&cauthor_uid=10209739). (1999b) IHF protein inhibits cleavage but not assembly of plasmid R388 relaxosomes. [Mol Microbiol](http://www.ncbi.nlm.nih.gov/pubmed/10209739?dopt=Abstract) **31**: 1643-52.

Monzingo, A.F., Ozburn, A., Xia, S., Meyer, R.J., and Robertus, J.D. (2007) The structure of the minimal relaxase domain of MobA at 2.1 \AA resolution. *J Mol Biol* **366**: 165–178.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986) Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* **1**: 263–273.

Nash, R.P., Habibi, S., Cheng, Y., Lujan, S.A., Redinbo, M.R. (2010) The mechanism and control of DNA transfer by the conjugative relaxase of resistance plasmid pCU1. *Nucleic Acids Res* **38**: 5929–5943.

Nelson, W.C., Howard, M.T/, Sherman, J.A., and Matson, S.W. (1995) [The *traY* gene product and integration host factor stimulate *Escherichia coli* DNA helicase I-catalyzed nicking at the F plasmid *oriT*.](http://www.ncbi.nlm.nih.gov/pubmed/7499339) *J Biol Chem* **270**: 28374-80.

Pansegrau, W. and Lanka, E. (1996) Mechanisms of initiation and termination reactions in conjugative DNA processing. Independence of tight substrate binding and catalytic activity of relaxase (TraI) of IncPα plasmid RP4. *J Biol Chem* **271**:13068–13076.

Pansegrau, W., Lanka, E., Barth, P.T., Figurski, D.H., Guiney, D.G., Haas, D., *et al.* (1994) Complete nucleotide sequence of Birmingham IncPα plasmids: compilation and comparative analysis. *J Mol Biol* **239**: 623–663.

Pansegrau, W., Schröder, W., and Lanka, E. (1993) Relaxase (TraI) of IncP alpha plasmid RP4 catalyzes a site-specific cleaving-joining reaction of single-stranded DNA. *Proc Natl Acad Sci USA* **90**: 2925–2929.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real time RT–PCR. *Nucleic Acids Res* **29**: 2002–2007.

Ragonese, H., Haisch, D., Villareal, E., Choi, J.-H., and Matson, S.W. (2007) The F plasmid-encoded TraM protein stimulates relaxosome-mediated cleavage at *oriT* through an interaction with TraI. *Mol Microbiol* **63**: 1173–1184.

Rhodes, G., Parkhill, J., Bird, C., Ambrose, K., Jones, M.C., Huys, G., *et al.* (2004) Complete nucleotide sequence of the conjugative tetracycline resistance plasmid pFBAOT6, a member of a group of IncU plasmids with global ubiquity. *Appl and Environ Microbiol* **70**: 7497–7510.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular cloning, a laboratory manual*. 2 nd ed., Cold Spring Harbor, NY: Cold Spring Harbor Press.

Schneiker, S., Keller, M., Dröge, M., Lanka, E., Pühler, A., and Selbitschka, W. (2001) The genetic organization and evolution of the broad host range mercury resistance plasmid pSB102 isolated from a microbial population residing in the rhizosphere of alfalfa. *Nucleic Acids Res* **29**: 5169–5181.

Spratt, B.G., Hedge, P.J., te Heesen, S., Edelman, A., and Broome-Smith, J.K. (1986). Kanamycin-resistant vectors that are analogues of plasmids pUC8, pUC9, pEMBL8 and pEMBL9. *Gene* **41**: 337–342.

Tauch, A., Schneiker, S., Selbitschka, W., Pühler, A., Overbeek, L.S. van, Smalla, K., *et al.* (2002) The complete nucleotide sequence and environmental distribution of the cryptic, conjugative, broad-host-range plasmid pIP02 isolated from bacteria of the wheat rhizosphere. *Microbiol* **148**: 1637–1653.

Williams, S.L., and Schildbach, J.F. (2006) Examination of an inverted repeat within the F factor origin of transfer: context dependence of F TraI relaxase DNA specificity. *Nucleic Acids Res* **34**: 426–435.

Van der Auwera, G., Król, J., Suzuki, H., Foster, B., Van Houdt, R., Brown, C., *et al.* (2009) Plasmids captured in *C. metallidurans* CH34: defining the PromA family of broad-host-range plasmids. *Antonie van Leeuwenhoek* **96**: 193–204.

[Varsaki, A](http://www.ncbi.nlm.nih.gov/pubmed/?term=Varsaki%20A%5BAuthor%5D&cauthor=true&cauthor_uid=22449962)., [Lamb, H.K](http://www.ncbi.nlm.nih.gov/pubmed/?term=Lamb%20HK%5BAuthor%5D&cauthor=true&cauthor_uid=22449962)., [Eleftheriadou, O](http://www.ncbi.nlm.nih.gov/pubmed/?term=Eleftheriadou%20O%5BAuthor%5D&cauthor=true&cauthor_uid=22449962)., [Vandera, E](http://www.ncbi.nlm.nih.gov/pubmed/?term=Vandera%20E%5BAuthor%5D&cauthor=true&cauthor_uid=22449962)., [Thompson, P](http://www.ncbi.nlm.nih.gov/pubmed/?term=Thompson%20P%5BAuthor%5D&cauthor=true&cauthor_uid=22449962)., [Moncalián, G](http://www.ncbi.nlm.nih.gov/pubmed/?term=Moncali%C3%A1n%20G%5BAuthor%5D&cauthor=true&cauthor_uid=22449962)., [de la Cruz, F](http://www.ncbi.nlm.nih.gov/pubmed/?term=de%20la%20Cruz%20F%5BAuthor%5D&cauthor=true&cauthor_uid=22449962)., [Hawkins, A.R](http://www.ncbi.nlm.nih.gov/pubmed/?term=Hawkins%20AR%5BAuthor%5D&cauthor=true&cauthor_uid=22449962)., and [Drainas, C](http://www.ncbi.nlm.nih.gov/pubmed/?term=Drainas%20C%5BAuthor%5D&cauthor=true&cauthor_uid=22449962). (2012) Interaction between relaxase MbeA and accessory protein MbeC of the conjugally mobilizable plasmid ColE1. [*FEBS Lett*](http://www.ncbi.nlm.nih.gov/pubmed/22449962) 586: 675-9.

Varsaki, A., Moncalián, G., Pilar Garcillán-Barcia, M. del, Drainas, C., and de la Cruz, F. (2009) Analysis of ColE1 MbeC unveils an extended ribbon-helix-helix family of nicking accessory proteins. *J Bacteriol* **191**: 1446–1455.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103–119.

Zechner, E.L., de la Cruz, F., Eisenbrandt, R., Grahn, A.M.,Koraimann, G., Lanka, E., Muth, G., Pansegrau, W., Thomas, C.M., Wilkins, B.M., Zatyka, M. (2000) Conjugative-DNA transfer processes. In: Thomas, C.M. (Ed.), *The Horizontal Gene Pool, Bacterial Plasmids Gene Spread.* Harwood Academic Publishers, Amsterdam, pp. 87–174.

Ziegelin, G., Fürste, J.P., and Lanka, E. (1989) TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer origin. *J Biol Chem* **264**: 11989–11994.

Ziegelin, G., Pansegrau, W., Lurz, R., and Lanka, E. (1992) [TraK protein of conjugative plasmid RP4 forms a specialized nucleoprotein complex with the transfer origin.](http://www.ncbi.nlm.nih.gov/pubmed/1324929) *J Biol Chem* **267**: 17279-86.

[Ziegelin, G](http://www.ncbi.nlm.nih.gov/pubmed/?term=Ziegelin%20G%5BAuthor%5D&cauthor=true&cauthor_uid=1665997)., [Pansegrau, W](http://www.ncbi.nlm.nih.gov/pubmed/?term=Pansegrau%20W%5BAuthor%5D&cauthor=true&cauthor_uid=1665997)., [Strack, B](http://www.ncbi.nlm.nih.gov/pubmed/?term=Strack%20B%5BAuthor%5D&cauthor=true&cauthor_uid=1665997)., [Balzer, D](http://www.ncbi.nlm.nih.gov/pubmed/?term=Balzer%20D%5BAuthor%5D&cauthor=true&cauthor_uid=1665997)., [Kröger, M](http://www.ncbi.nlm.nih.gov/pubmed/?term=Kr%C3%B6ger%20M%5BAuthor%5D&cauthor=true&cauthor_uid=1665997)., [Kruft, V](http://www.ncbi.nlm.nih.gov/pubmed/?term=Kruft%20V%5BAuthor%5D&cauthor=true&cauthor_uid=1665997)., and [Lanka, E](http://www.ncbi.nlm.nih.gov/pubmed/?term=Lanka%20E%5BAuthor%5D&cauthor=true&cauthor_uid=1665997). (1991) Nucleotide sequence and organization of genes flanking the transfer origin of promiscuous plasmid RP4. [*DNA Seq*](http://www.ncbi.nlm.nih.gov/pubmed/?term=TraI+plasmid+RP4+TraJ+TraK) **1**: 303-27.

Zukowski, M.M., Gaffney, D.F., Speck, D., Kauffmann, M., Findeli, A., Wisecup, A., and Lecocq, J.P. (1983) Chromogenic identification of genetic regulatory signals in *Bacillus subtilis* based on expression of a cloned *Pseudomonas* gene. *Proc Natl Acad Sci USA* **80**: 1101–1105.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tables**  **Table 1. Plasmids used in this study** | | | | |
| **Designation** | | **Relevant features** | **Copy number** | **References** | |
| pAMB9.37 | | pBBR1MCS-1 with *lacI*q, *tac*p, expression vector |  | Ludwiczak *et al*., 2013 | |
| pBBR1MCS-1 | | BHR1, IncA/C*,* CmR | medium | Kovach *et al.*, 1994 | |
| pBGS18 | | *ori*MB1, KmR | high | Spratt *et al.*,1986 | |
| pET3a | | *ori*MB1, ApR, T*7*p, T7 tag, MCS, T7 terminator | medium | Novagen | |
| pET28a | | *ori*MB1, KmR, T*7*p, *lacO*, His6-tag, T7 tag, MCS, T7 terminator | medium | Novagen | |
| pET29c | | *ori*MB1, KmR, T*7*p, *lacO*, S-tag, MCS, His6-tag, T7 terminator | medium | Novagen | |
| pGBT30 | | *ori*MB1, ApR, *lacI*q, *tac*p expression vector | high | Jagura-Burdzy *et al*., 1991 | |
| pGEM-T Easy | | *ori*MB1, linear vector for cloning of PCR products | high | Novagen | |
| pJSB2.9 | | *ori*MB1, ApR, *mobC*p (RA3 coordinates 9435-9852) | high | Godziszewska *et al.*, 2014 | |
| pJSB2.43 | | *ori*MB1, ApR, *oriT61* | high | Godziszewska *et al.*, 2014 | |
| pJSB2.44 | | *ori*MB1, ApR, *oriT45* | high | Godziszewska *et al.*, 2014 | |
| pJSB2.57 | | *ori*MB1, ApR, *oriT45*RA3, *lacI*q, *tac*p*-mobC-nic*, TraRA3, *korC*p*-korC* | high | Godziszewska *et al.*, 2014 | |
| pJSB4.1 | | pBBR1MCS-1, *lacI*q, *tac*p-*mobC* | medium | Godziszewska *et al.*, 2014 | |
| pJSB5.1 | | *ori*MB1, ApR, *lacI*q, *tac*p-*mobC* | high | Godziszewska *et al.*, 2014 | |
| pJSB6.1 | | *ori*MB1, KmR, T*7*p*-mobC*, | medium | Godziszewska *et al.*, 2014 | |
| pJSB7.9 | | *ori*SC101, KmR, *mobC*p- *xylE* | medium | Godziszewska *et al.*, 2014 | |
| pJSB7.17 | | *ori*SC101, KmR, *mobC*pmutVIII2-*xylE* | medium | Godziszewska *et al.*, 2014 | |
| pJSB7.18 | | *ori*SC101, KmR, *mobC*pmutIX3-*xylE* | medium | Godziszewska *et al.*, 2014 | |
| pPT01 | | *ori*SC101, KmR, promoter-less *xylE* | medium | Macartney *et al.*, 1997 | |
| pUC18 | | *ori*MB1, ApR | high | Yanisch-Perron *et al*., 1985 | |
| RA3 | | BHR, IncU,CmR, SmR, SuR | low | Hayes F. | |
| 1 BHR - broad-host-range  2 two-nucleotide substitution in MobC operator OM, here designated OMxx  3 three-nucleotide substitution in IR3a arm of inverted repeat, here designated IR3ax  **Plasmids constructed in this study** | | | | | |
| **Designation** | **Description** | | | | |
| pET3a-*nic* | *nic* orf cloned between NdeI and BamHI sites of pET3a | | | | |
| pET29c-*nic* | *nic* orf cloned between NdeI and XhoI sites of pET29c to produce NIC-His6 | | | | |
| pJSB1.80 | pBGS18 *mobC*p*-mobC-nic’*; EcoRI-SmaI fragment from pJSB3.4 | | | | |
| pJSB1.81 | pBGS18 *mobC*p*-mobC-nicV108A*’; an enzymatic assembly of two mutagenic fragments, PCR amplified on pJSB1.80 with the use of pairs of primers #41/ #44 and #42/ #43 | | | | |
| pJSB1.82 | pBGS18 *mobC*p*-mobC-nicV108E*’; an enzymatic assembly of two mutagenic fragments, PCR amplified on pJSB1.80 with the use of primers #39/ #44 and #40/ #43 | | | | |
| pJSB1.83 | pBGS18 *mobC*p*-mobC-nicV108A*, TraRA3, *korC*p-*korC*;SmaI-SalI fragment from pJSB2.57 carrying ‘*nic*, TraRA3, *korC*p-*korC* inserted into pJSB1.81 | | | | |
| pJSB1.84 | pBGS18 *mobC*p*-mobC-nicV108E*, TraRA3 *korC*p-*korC*; SmaI-SalI fragment from pJSB2.57 carrying ‘*nic*, TraRA3, *korC*p-*korC* inserted into pJSB1.82 | | | | |
| pJSB2.50 | pUC18 *oriT31*; annealed primers #5 and #6, inserted in the SmaI site (RA3 coordinates 9736-9766) | | | | |
| pJSB2.52 | pUC18 *oriT25*; annealed primers #7 and #8, inserted in the SmaI site of pUC18 (RA3 coordinates 9736-9760) | | | | |
| pJSB2.53 | pUC18 *oriT20*; annealed primers #3 and #4, inserted in the SmaI site (RA3 coordinates 9736-9755) | | | | |
| pJSB2.54 | pUC18 *oriT31x*; annealed primers #9 and #10; inserted in the SmaI site (RA3 coordinates 9736-9766; GCA→TAG at positions 9756-9758) | | | | |
| pJSB2.60 | pUC18 *oriT31*-1; annealed primers #13 and #14; inserted in the SmaI site (RA3 coordinates 9736-9766; G→C at position 9756) | | | | |
| pJSB2.61 | pUC18 *oriT31*-2; annealed primers #15 and #16, inserted in the SmaI site (RA3 coordinates 9736-9766; C→G at position 9757) | | | | |
| pJSB2.62 | pUC18 *oriT31*-3; annealed primers #17 and #18, inserted in the SmaI site (RA3 coordinates 9736-9766; A→T at position 9758) | | | | |
| pJSB2.63 | pUC18 *oriT31*-4; annealed primers #19 and #20, inserted in the SmaI site (RA3 coordinates 9736-9766; A→T at position 9759) | | | | |
| pJSB2.64 | pUC18 *oriT31*-5; annealed primers #21 and #22, inserted in the SmaI site (RA3 coordinates 9736-9766; T→A at position 9760) | | | | |
| pJSB2.65 | pUC18 *oriT31*-6; annealed primers #23 and #24, inserted in the SmaI site (RA3 coordinates 9736-9766; T→A at position 9761) | | | | |
| pJSB2.66 | pUC18 *oriT31*-7; annealed primers #25 and #26, inserted in the SmaI site (RA3 coordinates 9736-9766; T→A at position 9762) | | | | |
| pJSB2.68 | pUC18 *oriT60hairpin*; annealed primers #27 and #28, inserted in the SmaI site (RA3 coordinates 9736-9795) | | | | |
| pJSB2.69 | pUC18 *14+8hairpin*; annealed primers #29 and #30, inserted in the SmaI site (RA3 coordinates 9740-9761; with additional arm of repeat - IR3b) | | | | |
| pJSB2.70 | pUC18 *oriT47*; annealed primers #47 and #48, inserted in the SmaI site (RA3 coordinates 9736-9782) | | | | |
| pJSB2.71 | pUC18 *oriT60hairpin*OM*xmobC*p*x*; annealed primers #49 and #50, inserted in the SmaI site (RA3 coordinates 9736-9795, with randomized OM and *mobC*p ) | | | | |
| pJSB2.72 | pUC18 *oriT60hairpinmobC*p*x*; annealed primers #51 and #52 inserted in the SmaI site (RA3 coordinates 9736-9795, with randomized *mobC*p ) | | | | |
| pJSB2.73 | pUC18 *oriT*55*hairpin*; annealed primers #53 and #54, inserted in the SmaI site (RA3 coordinates 9736-9795, with IR3b 5 nt closer to IR3a ) | | | | |
| pJSB2.76 | pUC18 *oriT60hairpin*OM*x*; annealed primers #55 and #56, inserted in the SmaI site (RA3 coordinates 9736-9795, with randomized OM ) | | | | |
| pJSB2.77 | pUC18 *oriT55hairpin*OM*x*; annealed primers #57 and #58, inserted in the SmaI site (RA3 coordinates 9736-9795, with randomized OM and 5-nt deletion between IR3a and IR3b) | | | | |
| pJSB3.3 | pGEM-T easy *nic*, PCR fragment amplified with the use of primers #1 and #2 (RA3 coordinates 10360-11355) | | | | |
| pJSB3.4 | pGEM-T easy *mobC*p-*mobC*-*nic*; PCR fragment amplified with the use of primers #59 and #2 (RA3 coordinates 9435-11355) | | | | |
| pJSB4.3 | pBBR1MCS-1 *lacIq* *tac*p*-nic*, BamHI-SalI fragment from pJSB5.3 | | | | |
| pJSB5.3 | pGBT30 *tac*p*-nic*, EcoRI-SalI fragment from pJSB3.3 | | | | |
| pJSB6.3 | pET28a T7p*-nic*, EcoRI-SalI fragment from pJSB3.3 | | | | |
| pJSB6.78 | pET29c *nicV108A*;an enzymatic assembly of two mutagenic fragments, PCR amplified on pET29c *nic* with the use of pairs of primers #41/ #46 and #42/ #45 | | | | |
| pJSB6.79 | pET29c *nicV108E*;an enzymatic assembly of two mutagenic fragments, PCR amplified on pET29c *nic* with the use of pairs of primers #39/ #46 and #40/ # 45 | | | | |

**Figure legends**

**Fig. 1. Organization of RA3 region between stability and transfer modules**

**A. RA3 conjugative transfer module.** Orfs are coloured according to their predicted function in conjugation: light grey - homologs of Mpf (mating pair formation) system, dark grey - Dtr proteins (DNA replication and transfer), black indicates a homolog of coupling protein VirD4. Transcripts are symbolized by thin arrows above the graph. **B. Intermodular *par-tra* region containing *parS* and *oriT* sequences.** Direct repeats (DR) and inverted repeats (IR) identified in this region are indicated. The *parS* region involved in partitioning is shown according to Kulinska *et al*. (2011), where the IR-SnaBI palindrome is the primary binding site for KorB (ParB). The *nic* sequence is marked. *mobC* promoter motifs are boxed. **C. Nucleotide sequence of *mobC*p*/oriT* region**. Region between nucleotides 9722 and 9839 of RA3 is shown [GenBank: DQ401103]. Putative *mobC* promoter motifs are boxed and putative regulatory sequences are marked by arrows. Black triangle indicates the NIC cleavage site in the conserved sequence. Ribosome binding site and initiation codon for MobC are in bold. **D. Oligonucleotides used in analysis of NIC activity *in vitro*.** Three ss oligonucleotides 14+8; 14+18; and 14+8*hairpin* encompassing cleavage site and surrounding sequences (bottom strand in panel C) were tested as substrates for NIC cleavage-joining activity *in vitro*. 14+8*hairpin* has IR3b arm (in italics) moved next to IR3a. Oligonucletide 61 was a substrate in EMSA and NIC cleavage reaction. Important features are indicated.

**Fig. 2. Mapping of minimal *oriT*RA3.** Double stranded oligonucleotides of indicated length were cloned into pUC18 and introduced into *E. coli* DH5α (RA3). Sequences of oligonucletide inserts are shown with important motifs marked. pUC18 with a 417-nt *oriT* fragment (pJSB2.9) and empty pUC18 were used as controls in conjugation experiments. The DH5α (RA3)(pUC18 derivative) strains were used as donors in conjugation with DH5α RifR recipient. The frequency of plasmid mobilization is expressed as the number of ApR RifR transconjugants per donor cell with the standard deviation from at least three experiments. The level of significance in transfer frequency was tested for each individual strain against donor strain with pUC18*oriT417* using Student’s paired two-tailed t-test. p-value ≤0.05 was chosen as statistically significant. ND - not detected, frequency <10-7.

**A.** Data for all donor strains were compared individually to control strain with pUC18 *oriT417* and additionally between pairs of strains indicated by brackets. Significant differences are marked by asterisks (p-value ≤0.05). **B.** DNA sequences of different variants of *oriT60hairpin* (both arms of IR3 included) with scrambled MobC binding site (OMx) or promoter motifs (*mobC*px) or no intervening sequence between IR3 arms (*oriT14+8hairpin*) are shown. Substituted nucleotides are in lower case, deletion of 5 nt is indicated. Only unmodified OM and *mobC*p motifs are labelled as in wt sequence. Data for all donor strains were compared individually to control strain with pUC18*oriT417.* Significant differences are marked by asterisks (p value ≤0.05).

**Fig. 3. Site-directed mutational analysis of IR3a.** Thirty-one-nt long ds oligonucleotides with substitutions in IR3a arm, as indicated, were cloned in pUC18 and introduced into *E.coli* DH5α (RA3). As a control plasmid pUC18*oriT31* was applied. The double transformants were used as donors in conjugation with DH5α RifR strain. Transfer efficiency is expressed as the number of transconjugants per donor cell with standard deviation from at least three experiments. ND - below detection limit, frequency <10-7. The level of significance in transfer frequency was tested for each individual strain against donor strain with pUC18*oriT31* using Student’s paired two-tailed t-test p≤0.01. Statistically significant differences are marked by two asterisks in the diagram.

**Fig. 4. Co-repressor activity of NIC\_RA3. A. Enzymatic assay.** *E. coli* strainsDH5 (pJSB7.9 *mobC*p*-xylE*),DH5 (pJSB7.17 *mobC*pOMxx*-xylE*) and DH5 (pJSB7.18 *mobC*pIR3ax*-xylE*) were transformed with pairs of compatible expression plasmids, empty pBBR1MCS-1 and pGBT30 or their derivatives overproducing MobC (pJSB4.1 *tac*p*-mobC*) and NIC (pJSB5.3 *tac*p*-nic*). Positions of mutations in *mobC*p region are marked. XylE activity was assayed in extracts from logarithmically growing cultures of triple transformants without IPTG induction and is shown relative to the appropriate control strain with two empty expression vectors, with standard deviation from at least three independent assays. For each strain with a given *mobC*p*-xylE* transcriptional fusion the bars represent XylE activity in strains co-transformed with (top to bottom): empty vectors pBBR1MCS-1 and pGBT30, pJSB4.1 *tac*p*-mobC* and empty pGBT30, empty pBBR1MCS-1 and pJSB5.3 *tac*p*-nic*, and pJSB4.1 *tac*p*-mobC* and pJSB5.3 *tac*p*-nic*. The level of significance of difference in XylE activity was tested for each strain with combination of expression plasmids against control strain with two empty vectors using Student’s paired two-tailed t-test. Asterisks mark statistically significant differences (p<0.01). Additionally the XylE data from cultures producing MobC *versus* cultures producing MobC/ NIC in each set of strains were compared using t-test. Significant difference was found only for wt *mobC*p (p<0.01) as indicated by asterisks above the bracket. **B. Transcript analysis by RT-qPCR.** Total RNA was isolated from three cultures of each triple transformant strain used for enzymatic assays (biological replicates). After purification RNA samples were subjected to RT-qPCR analysis, cDNA from each biological replicate was analyzed in three technical replicates by qPCR with two pairs of primers for *xylE* and plasmidic *repA* gene. Changes in *xylE* reporter gene expression between control strain with empty expression vectors and strains producing either NIC, MobC or NIC/ MobC were calculated with normalization of Ct values to mean Ct value for the reference plasmidic gene *rep101*. Standard deviation for at least six replicates is indicated. The level of significance of differences in the relative *xylE* assays was tested for each strain with combination of expression plasmids against control strain with two empty vectors using Student’s paired two-tailed t-test (p<0.01). Additionally data from cultures producing MobC *versus* data from cultures producing MobC/ NIC in each set of strains were compared using t-test. Significant difference was found only for wt *mobC*p (p<0.01) as indicated by asterisks above the bracket.

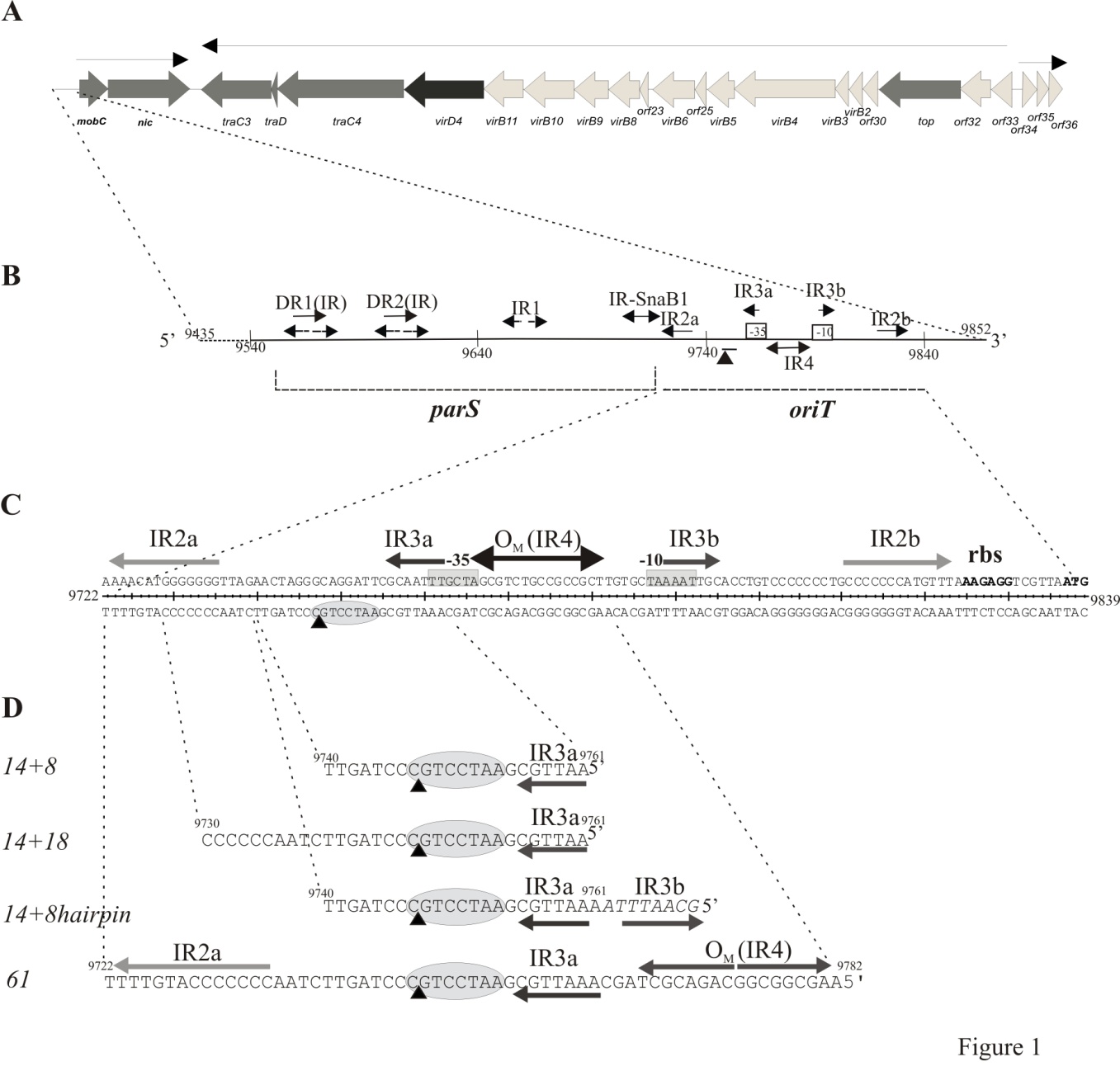
**Fig. 5. NIC binding and cleavage of *oriT61***. **A. NIC binding to ss *oriT*.** 5’ Cy5-labelled 61-nt oligonucleotide (#31 in Table S1, 2.5 pmole) was incubated with His6-NIC (0, 5, 10, or 20 pmole) in binding buffer. After separation on native 10% polyacrylamide gel, Cy5 fluorescence was visualized by FluorchemQ MultiImage III (Alpha Innotech). Arrows indicate DNA-NIC complexes. **B. NIC binding to ds *oriT*.** 5’ Cy5-labelled 61-nt oligonucleotide (#31) was annealed to unlabelled complementary strand (#11 in Table S1). The ds *oriT61* (2.5 pmole) was incubated with His6-NIC (0- 20 pmole) in binding buffer and analyzed as described above. **C. NIC binding to non-specific ssDNA.** 5’ Cy5-labelled ss oligonucleotide (#32 in Table S1, 2.5 pmole)was incubated with His6-NIC (0-20 pmole) in binding buffer and analyzed as described above. **D. Cleavage of ss *oriT61* by His6-NIC*.*** 5’ Cy5-labelled 61-nt oligonucleotide (#31 in Table S1, 2.5 pmole) was incubated with His6-NIC (0-20 pmole) in nicking buffer. After proteinase K treatment in the presence of SDS DNA fragments were separated on native 10% polyacrylamide gel and visualized as above. Arrow indicates the cleavage product of 36 nt. **E. NIC-His6 cleavage activity in the presence of MobC.** 5’ Cy5-labelled 61-nt oligonucleotide (#31 in Table S1, 2.5 pmole) was incubated with NIC-His6 (0-20 pmole) in nicking buffer. Where indicated, His6-MobC was added (15 pmole). Reaction conditions and analysis as above. Cleavage 36-nt product is indicated by arrow.

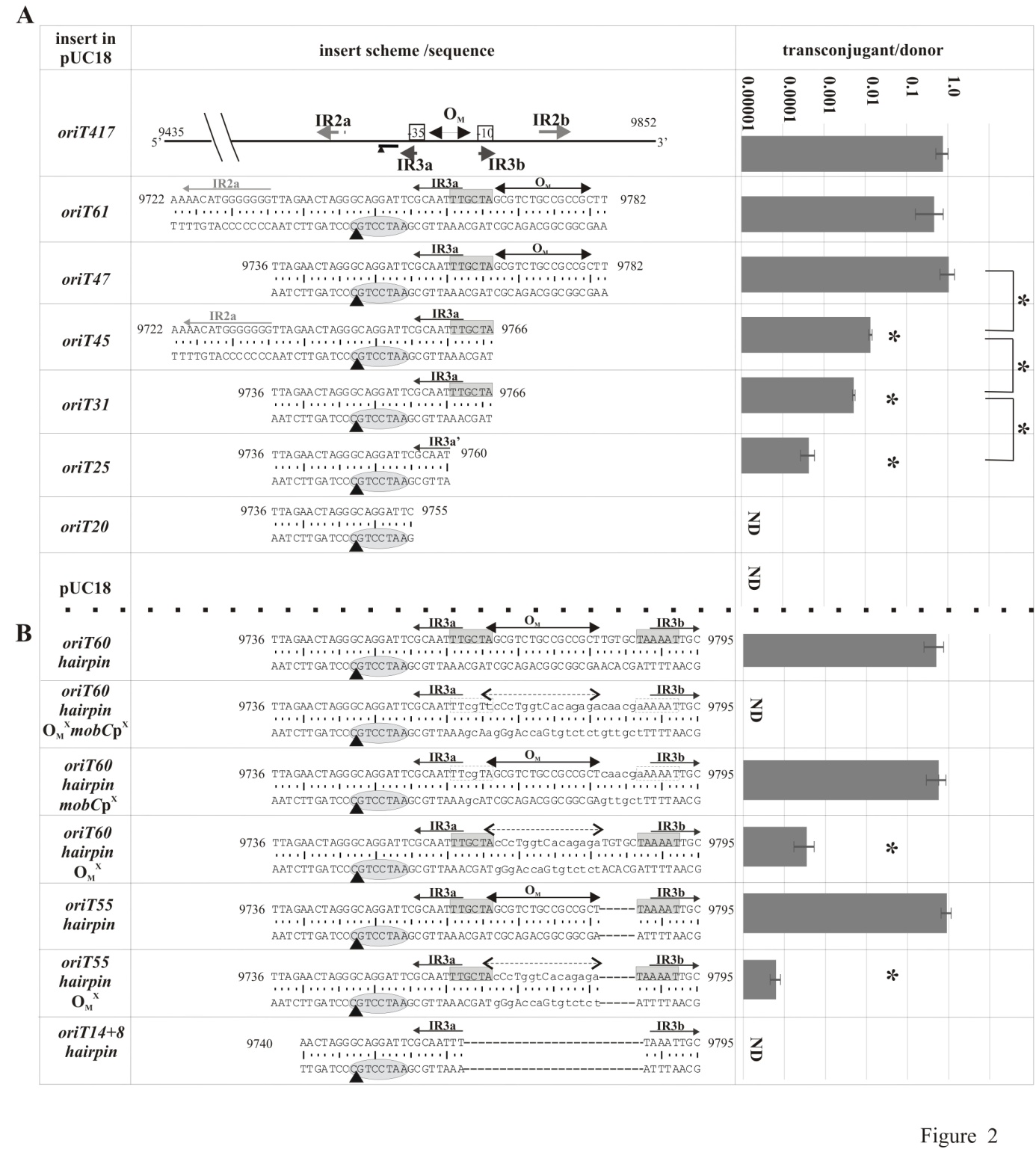
**Fig. 6. Dependence of NIC cleavage/ covalent binding activity on divalent cations.**

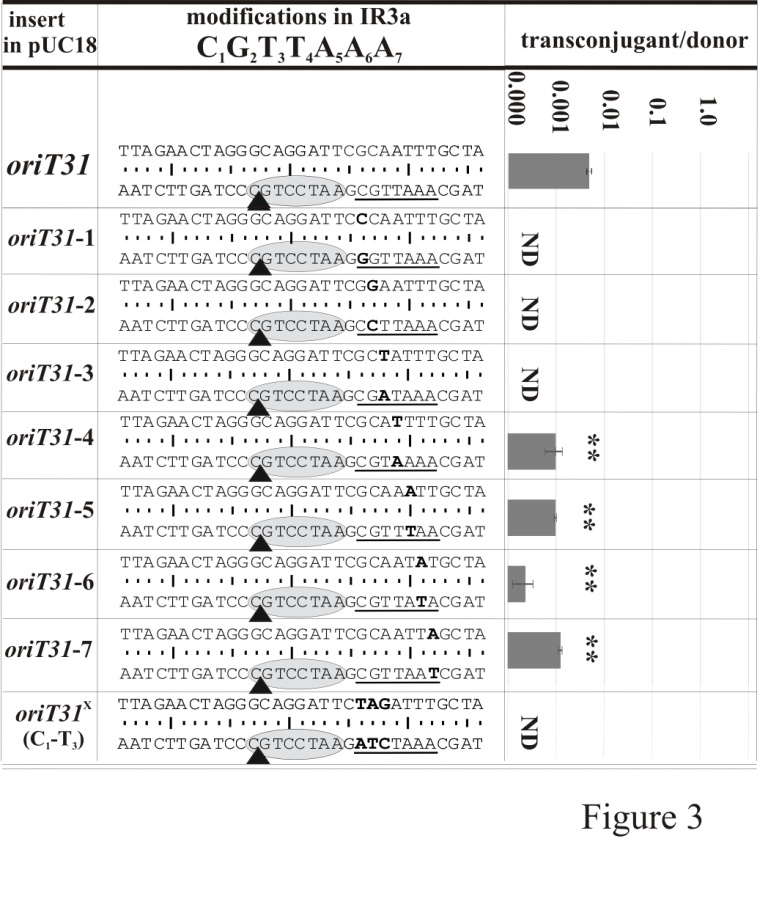
**A. Cleavage/ covalent binding of ss 14+8 oligonucleotide.** 0.1 mM of untagged NIC protein was incubated with 70 M ss 14+8 oligonucleotide (Fig. 1D) in the presence of 5 mM different divalent cations and reaction mixture was separated by SDS–PAGE on 10% gel and stained with Coomassie Brilliant Blue to visualize NIC covalently bound to 5’-end of the cleaved oligonucleotide. In the first lane NIC protein without oligonucleotide added. White arrows indicate NIC, black arrows point out the complexes between NIC-His6 and the cleaved oligonucleotide. **B. Comparison of cleavage/ covalent binding activities of NIC-His6 and untagged NIC**. 0.1 mM of NIC-His6 or untagged NIC were incubated with 70 M oligonucleotide: 14+8 or 14+18 (Fig. 1D) in the presence of 5 mM Mn2+ or Ni2+ and reaction mixtures were analyzed as in **A**. **C. 3D model of the relaxase domain of RA3 NIC.** A 3D model was generated by homology modelling using RSF1010\_miniMobA (access 2SN6) as a template. One-hundred and fifty-five residues (42-186) of NIC could be modelled at >90% accuracy, although the sequence identity between NIC and miniMobA is only 25%. Amino acid residues presumably important for activity are shown as stick models and labelled. **D. Comparison of cleavage/ covalent binding activities of NIC-His6, NICV108A-His6 and NICV108E-His6 in the presence of Mg2+ or Mn2+**. 0.1 mM of NIC variants were incubated with 70 M oligonucleotide 14+18 either in the presence of EDTA, 10 mM Mg2+ or Mn2+ cations (as indicated). The complexes were separated by SDS–PAGE on 10% gel and stained by Coomassie blue. Black arrows point out NIC-oligonucleotide complexes formed with different efficiency in the presence of Mg2+. **E. Comparison of cleavage/ covalent binding activities of NICV108A-His6 and NICV108E-His6 in the presence of different concentrations of Mg2+**. 0.1 mM of NIC variants were incubated with 70 M oligonucleotide 14+18 in the presence of 10 mM Mg2+ (lanes 1 and 6), 1 mM Mg2+ (lanes 2 and 7)¸ 0.1 mM Mg2+ (lanes 3 and 8) or 0.01 mM Mg2+ (lanes 4 and 9). Complexes were separated by SDS–PAGE on 10% gel and stained by Coomassie blue. Black arrows point out NIC-oligonucleotide complexes formed.

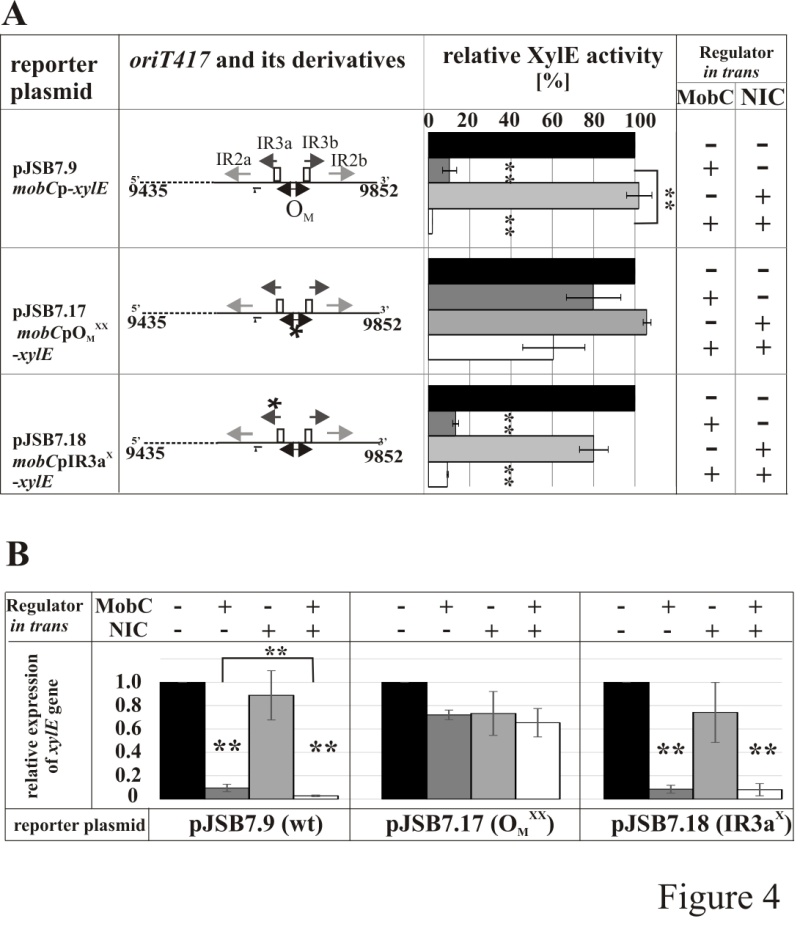
**Fig. 7. Nicking activity of NIC-His6 on supercoiled plasmid DNA. A.** Supercoiled plasmid (100 ng), control pUC18 (empty vector) or pJSB2.9*oriT417* fragment (417 bp), was incubated for 30 min at 37oC in nicking buffer with 10 or 20 pmole of NIC-His6) and, where indicated, with 7.5 pmole of His6-MobC. After proteinase K and SDS treatment the reaction mixtures were separated on 1% agarose gel, stained with ethidium bromide and visualized under UV. Arrows indicate OC forms. **B.** Supercoiled pUC18 derivatives (100 ng) with different versions of minimal *oriT* inserted were treated as above.

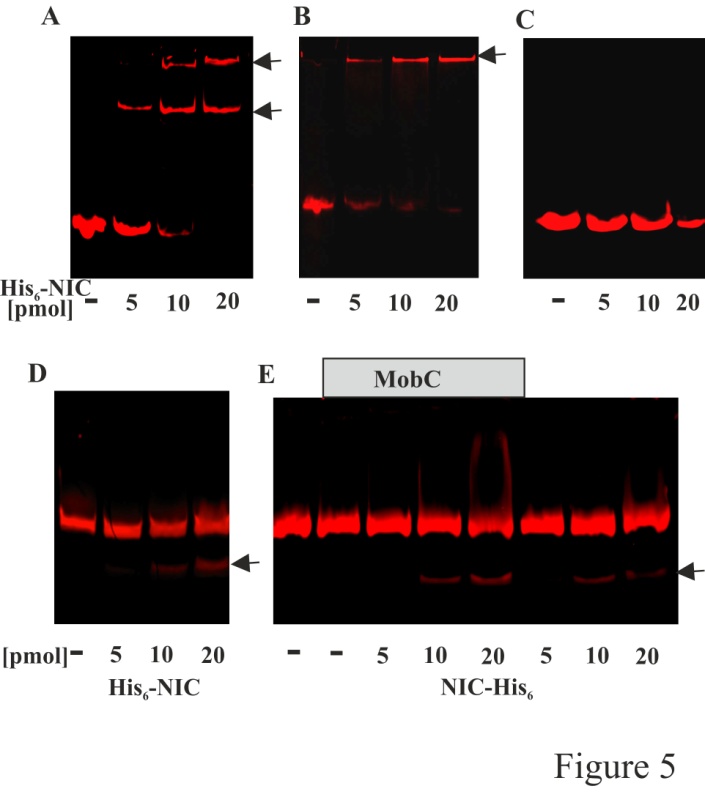
**Figures**

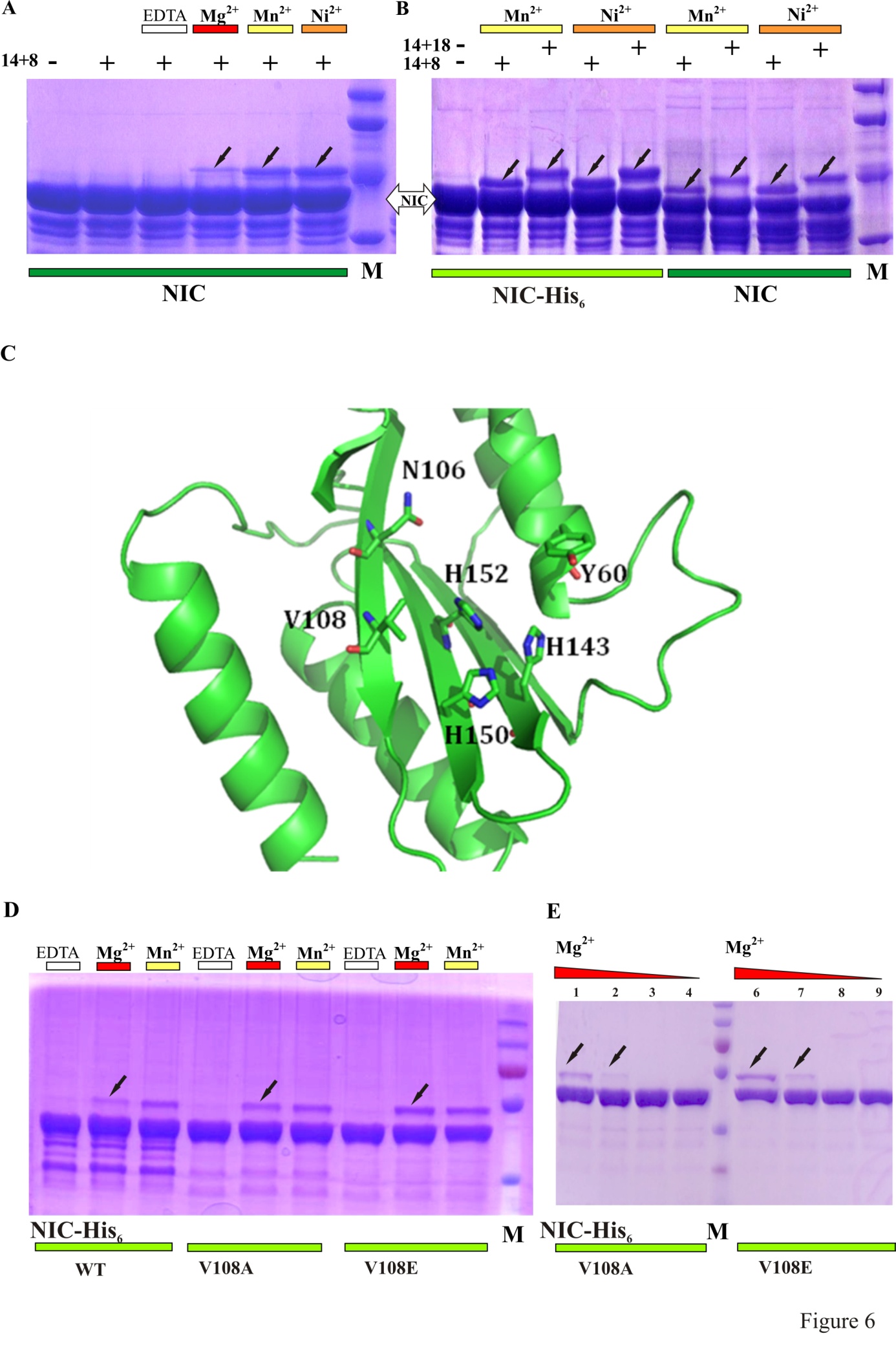


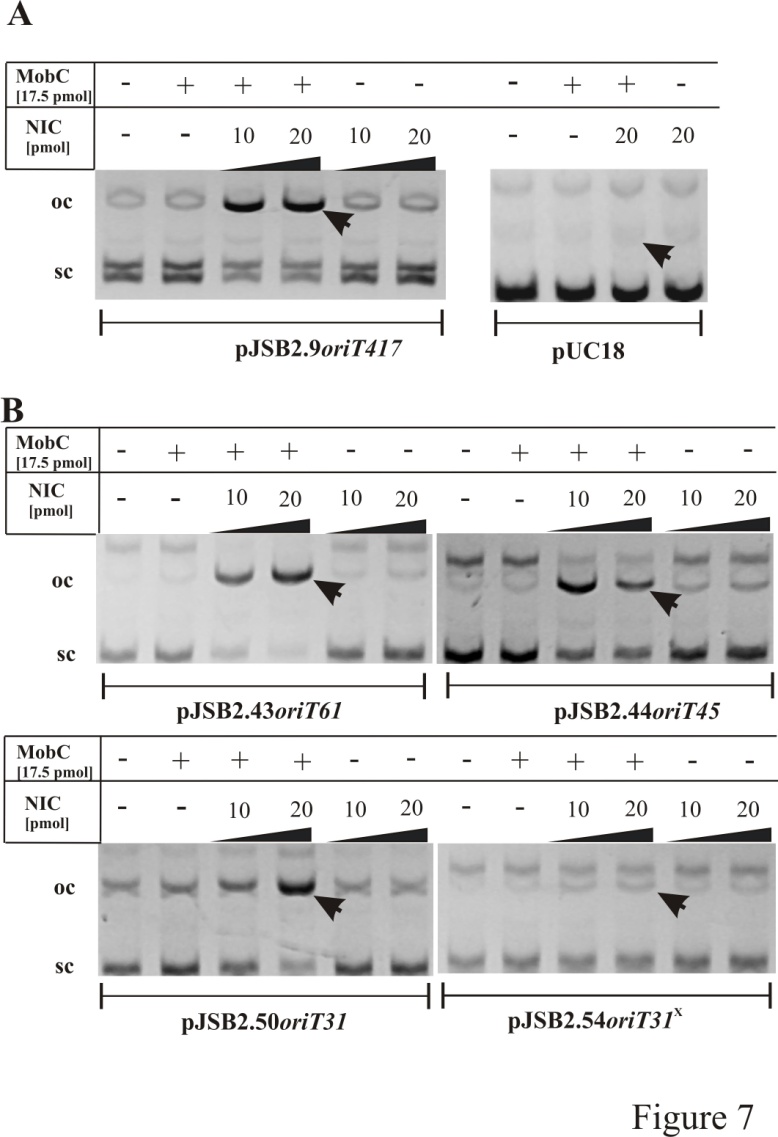




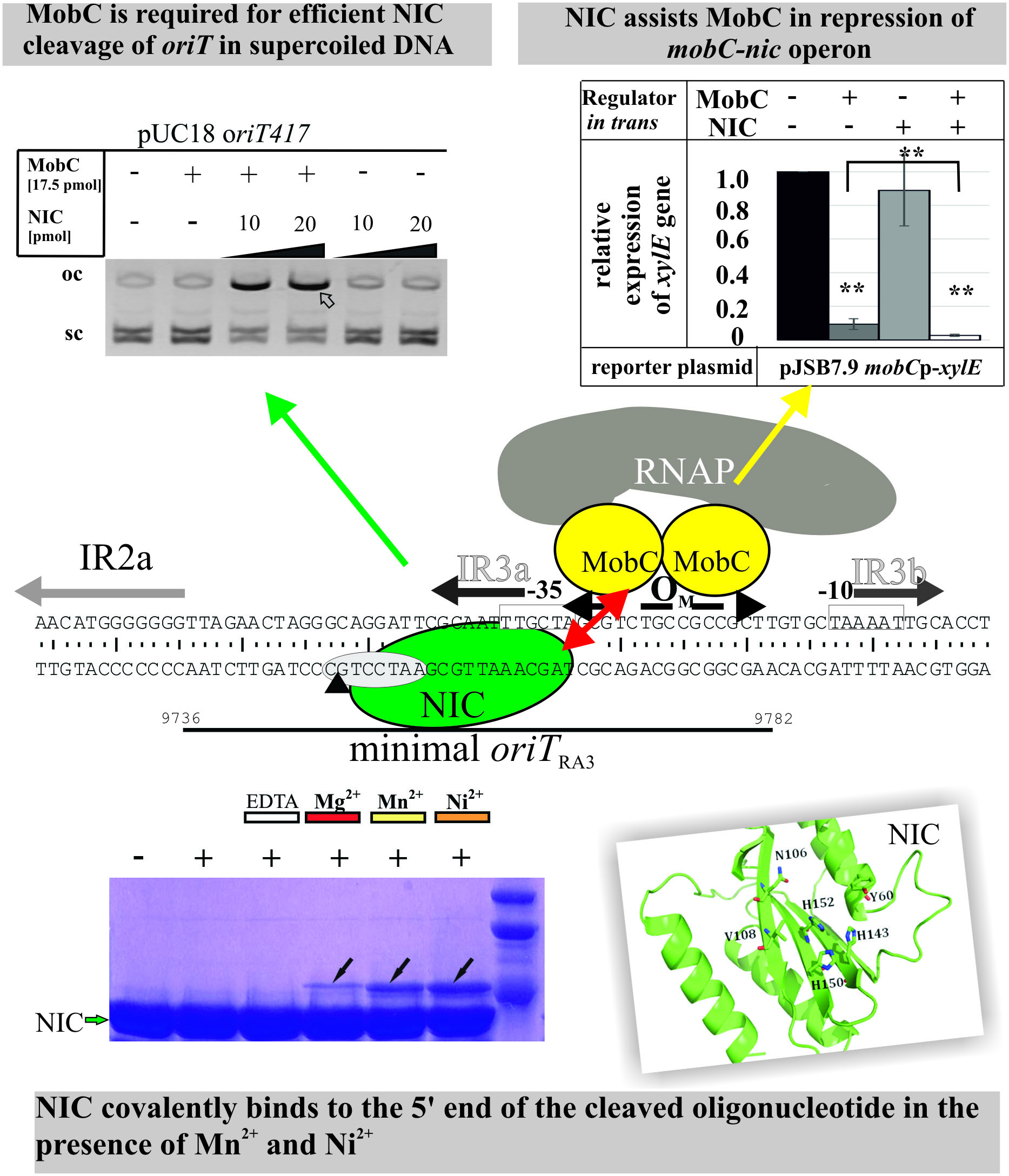








**Abbreviated Summary**



The *mobC-nic* operon of broad-host-range conjugative RA3 plasmid (IncU incompatibility group) encodes HUH relaxase NIC (MOBP4 subgroup) and a transcriptional regulator MobC. Here we delineated minimal *oriT*RA3, identified the NIC\_RA3 binding site and its metal ions preference. The interplay between these two proteins was demonstrated: the MobC participates in the efficient NIC cleavage of supercoiled *oriT*RA3 and NIC\_RA3 potentiates the MobC repression of the *mobC* promoter.

SUPPLEMENTARY INFORMATION

**Concerted action of NIC relaxase and auxiliary protein MobC in RA3 plasmid conjugation**

Jolanta Godziszewska, Gabriel Moncalián, Matilde Cabezas, Aneta A. Bartosik, Fernando de la Cruz and Grazyna Jagura-Burdzy

*List of content:*

1. **Table S1.** Oligonucleotides used in this study

2. **Fig. S1.** Amino acids sequence alignment of relaxases from MOBP4 subfamily.

3. **Fig. S2.** NIC\_RA3 and MobA\_R1162 sequence alignment.

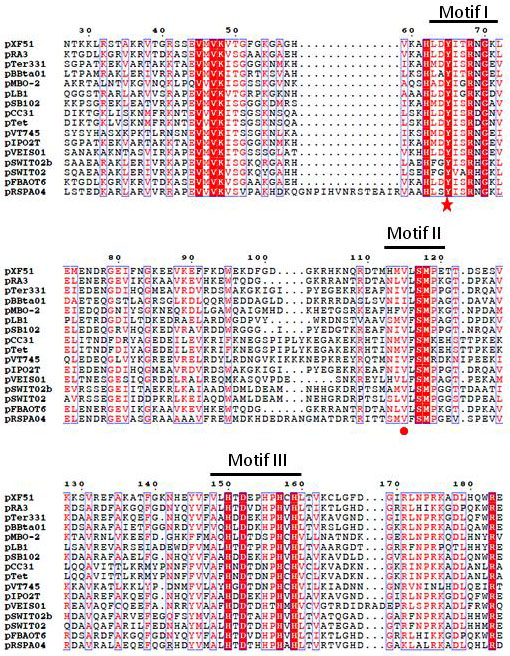
4. **Fig. S3.** Comparison of N-terminal parts of the relaxases NIC\_RA3, MobA\_R1162, NES\_pLW1043 and MobM\_pMV158.

5. **Supplementary methods**

**Table S1. Oligonucleotides used in this study**

|  |  |  |
| --- | --- | --- |
| **#1** | *nic1* | CGGAATTC**ATG**AATAAGGGCTATGACACTCTAGCCGGG |
| **#2** | *nic2* | CGGTCGACTTATCTCTCGTCTTCGTCCC |
| **#3** | *wt20D* | GAATCCTGCCCTAGTTCTAA |
| **#4** | *wt20G* | TTAGAACTAGGGCAGGATTC |
| **#5** | *wt31G* | TTAGAACTAGGGCAGGATTCGCAATTTGCTA |
| **#6** | *wt31D* | TAGCAAATTGCGAATCCTGCCCTAGTTCTAA |
| **#7** | *wt25G* | TTAGAACTAGGGCAGGATTCGCAAT |
| **#8** | *wt25D* | ATTGCGAATCCTGCCCTAGTTCTAA |
| **#9** | *31TAGg* | TTAGAACTAGGGCAGGATTCtagATTTGCTA |
| **#10** | *31TAGd* | TAGCAAATctaGAATCCTGCCCTAGTTCTAA |
| **#11** | *oriT61G* | AAACATGGGGGGGTTAGAACTAGGGCAGGATTCGCAATTTGCTAGCGTCTGCCGCCGCTTG |
| **#12** | *oriT61D* | CAAGCGGCGGCAGACGCTAGCAAATTGCGAATCCTGCCCTAGTTCTAACCCCCCCATGTTT |
| **#13** | *mut131G* | TTAGAACTAGGGCAGGATTCcCAATTTGCTA |
| **#14** | *mut131D* | TAGCAAATTGgGAATCCTGCCCTAGTTCTAA |
| **#15** | *mut231G* | TTAGAACTAGGGCAGGATTCGgAATTTGCTA |
| **#16** | *mut231D* | TAGCAAATTcCGAATCCTGCCCTAGTTCTAA |
| **#17** | *mut331G* | TTAGAACTAGGGCAGGATTCGCtATTTGCTA |
| **#18** | *mut331D* | TAGCAAATaGCGAATCCTGCCCTAGTTCTAA |
| **#19** | *mut431G* | TTAGAACTAGGGCAGGATTCGCAtTTTGCTA |
| **#20** | *mut431D* | TAGCAAAaTGCGAATCCTGCCCTAGTTCTAA |
| **#21** | *mut531G* | TTAGAACTAGGGCAGGATTCGCAAaTTGCTA |
| **#22** | *mut531D* | TAGCAAtTTGCGAATCCTGCCCTAGTTCTAA |
| **#23** | *mut631G* | TTAGAACTAGGGCAGGATTCGCAATaTGCTA |
| **#24** | *mut631D* | TAGCAtATTGCGAATCCTGCCCTAGTTCTAA |
| **#25** | *mut731G* | TTAGAACTAGGGCAGGATTCGCAATTaGCTA |
| **#26** | *mut731D* | TAGCtAATTGCGAATCCTGCCCTAGTTCTAA |
| **#27** | *oriT62G* | TTAGAACTAGGGCAGGATTCGCAATTTGCTAGCGTCTGCCGCCGCTTGTGCTAAAATTGC |
| **#28** | *oriT62D* | GCAATTTTAGCACAAGCGGCGGCAGACGCTAGCAAATTGCGAATCCTGCCCTAGTTCTAA |
| **#29** | *14+8hairpinD* | GCAATTTAAAATTGCGAATCCTGCCCTAGTT |
| **#30** | *14+8hairpinG* | AACTAGGGCAGGATTCGCAATTTTAAATTGC |
| **#31** | *Cy5oriT61D* | Cy5-CAAGCGGCGGCAGACGCTAGCAAATTGCGAATCCT  GCCCTAGTTCTAACCCCCCCATGTTT |
| **#32** | *Cy5inc230P* | Cy5-GATAGCTCTTTGCCATTAAC |
| **#33** | *14+8* | AATTGCGAATCCTGCCCTAGTT |
| **#34** | *14+18* | AATTGCGAATCCTGCCCTAGTTCTAACCCCCC |
| **#35** | *xylEG* | AACTACCCGGACCACAAACC |
| **#36** | *xylED* | GCACGGTCATGAATCGTTCG |
| **#37** | *rep101qF* | CAACCAGAGAGCTGATGAC |
| **#38** | *rep101qR* | GCTTGAGAACTTGGCATAG |
| **#39** | *NicEGSacI* | ACCGCTAACCTGGaGCTcTCCATGCCAA |
| **#40** | *NicEDSacI* | CCTTTGGCATGGAgAGCtCCAGGTTAGC |
| **#41** | *NicAGKpnI* | ACCGCTAACCTGGcGCTGTCCATGCCAAAGGGtACcGATCCGAAAG |
| **#42** | *NicADKpnI* | AGCTTTCGGATCgGTaCCCTTTGGCATGGACAGCgCCAGGTTAGCG |
| **#43** | *GpBGS* | TGTAAAACGACGGCCAGTG |
| **#44** | *DpBGS* | CACAGGAAACAGCTATGACCA |
| **#45** | *pET29G* | CCCCTCTAGAAATAATTTTGTTTAACT |
| **#46** | *pET29D* | CTTTCGGGCTTTGTTAGCAG |
| **#47** | *oriT1D* | TTAGAACTAGGGCAGGATTCGCAATTTGCTAGCGTCTGCCGCCGCTT |
| **#48** | *oriT1G* | AAGCGGCGGCAGACGCTAGCAAATTGCGAATCCTGCCCTAGTTCTAA |
| **#49** | *oriT2D* | TTAGAACTAGGGCAGGATTCGCAATTTcgttccctggtcacagaga caacgaAAAATTGC |
| **#50** | *oriT2G* | GCAATTTTtcgttgtctctgtgaccagggaacgAAATTGCGAATCCTGCCCTAGTTCTAA |
| **#51** | *oriT3D* | TTAGAACTAGGGCAGGATTCGCAATTTcgTAGCGTCTGCCGCCGCTcaacgaAAAATTGC |
| **#52** | *oriT3G* | GCAATTTTtcgttgAGCGGCGGCAGACGCTAcgAAATTGCGAATCCTGCCCTAGTTCTAA |
| **#53** | *oriT4G* | TTAGAACTAGGGCAGGATTCGCAATTTGCTAGCGTCTGCCGCCGCTTAAAATTGC |
| **#54** | *oriT4D* | GCAATTTTAAGCGGCGGCAGACGCTAGCAAATTGCGAATCCTGCCCTAGTTCTAA |
| **#55** | *oriT60 OMX* | TTAGAACTAGGGCAGGATTCGCAATTTGCTAccctggtcacagagaTGTGCTAAAATTGC |
| **#56** | *oriT60 OMX* | GCAATTTTAGCACAtctctgtgaccagggTAGCAAATTGCGAATCCTGCCCTAGTTCTAA |
| **#57** | *oriT55 OMX* | TTAGAACTAGGGCAGGATTCGCAATTTGCTAccctggtcacagagaTAAAATTGC |
| **#58** | *oriT55 OMX* | GCAATTTTAtctctgtgaccagggTAGCAAATTGCGAATCCTGCCCTAGTTCTAA |
| **#59** | *kasmob1* | CGGAATTCACATGTTTCTCGTTGGAGGGTGATCA |

Restriction enzyme recognition sites are underlined, initiation codon is in bold, lower case letters indicate nucleotide substitutions relative to native sequence. Pairs of complementary oligonucleotides used for cloning into pUC18 are marked by shadowing.

****

**Fig. S1. Amino acids sequence alignment of relaxases from MOBP4 subfamily.** Identical residues are shown in white on red background, similar residues are shown in red. A red star indicates the conserved catalytic tyrosine, red circle marks the hydrophobic residues in motif II where a polar residue is found in other relaxases. Sequences were plotted with ESPript 2.2.

**Figure S2 SI.tif**

**Fig. S2.** **NIC\_RA3 and MobA\_R1162 sequence alignment.** Identical residues are shown in white on red background, similar residues in red. Secondary structure elements of the modelled NIC\_RA3 protein are shown above the alignment. Sequences were plotted with ESPript 2.2.

Figure S3 SI.tif

**Fig. S3. Comparison of N-terminal parts of the relaxases NIC\_RA3, MobA\_R1162, NES\_pLW1043 and MobM\_pMV158.** Similar residues are shadowed, HUH motifs are labelled in green, residue at position V108 in NIC\_RA3 and residues at equivalent positions in other proteins are shown in red. Sequences MobA\_R1162 (PDB 2SN6), NES\_pLW1043 (PDB 4HT4) and MobM\_pMV158 (PDB 4LVM) were aligned with NIC\_RA3 by Clustal Omega and verified manually.

**Supplementary methods:**

***1. Purification of His6-NIC\_RA3*.**

Overnight culture of the *E. coli* strain BL21(DE3) (pJSB6.3) was diluted 100-fold into 3 l of L broth with kanamycin and grown to A600~0.4 before 0.1 mM IPTG was added for 3 h. Cells were harvested and resuspended in 8 ml of sonication buffer (150 mM NaCl, 50 mM Tris-HCl pH8.0, 10 mM imidazole, 10 mM β-mercaptoethanol, 50 mM Arg, 50 mM Glu, 0.1 mM EDTA, 0.5% Triton X-100, cocktail of protease inhibitors, lysozyme 1 mg ml-1) and incubated on ice for 1 h. After sonication the lysate was cleared by centrifugation at 16600xg for 15 min at 4°C. The supernatant was applied on Äktapurifier system with nickel agarose column (Qiagen) equilibrated with buffer I (150 mM NaCl, 20 mM Tris-HCl pH 8.0, 10 mM imidazole, 10 mM β-mercaptoethanol). His6-NIC\_RA3 was eluted by buffer II (150 mM NaCl, 20 mM Tris-HCl pH 8.0, 600 mM imidazole, 10 mM β-mercaptoethanol). Fractions with His6-NIC\_RA3 were applied on the Resource S (GE Healthcare) - ion exchange chromatography column and eluted by buffer III (10 mM NaPi pH7.5) with increasing concentration of salt (75 mM – 1 M NaCl). Fractions with His6-NIC\_RA3 (eluted at 550 mM NaCl) were concentrated to a volume 0.5 ml on column Vivaspin 2 (Sartorius Stedim Biotech). For the last stage of purification a molecular sieve, Superdex 200 10/300 GL column (GE Healthcare) was used and the protein eluted with buffer IV (10 mM NaPi pH7.5, 550 mM NaCl). The presence of His6-NIC\_RA3 in the fractions was examined by SDS-PAGE, using a PHAST system (Pharmacia) with 12.5% homogenous gel. His6-NIC\_RA3 was stored at 4°C and a concentration of the protein was measured according to standard procedure.

***2. Purification of untagged NIC\_RA3.***

The NIC\_RA3 protein was expressed and purified as follows. An overnight culture of *E. coli* C41(DE3) cells harbouring pET3a:*nic* plasmid was diluted 20-fold in 1 l of LB medium containing ampicillin (100 μg ml-1) and incubated at 37ºC with shaking until A600 = 0.6. Then, IPTG was added to a final concentration of 0.5 mM. After 3 h further incubation, cells were harvested by centrifugation and stored at -80ºC. Frozen cells were resuspended with a solution containing Tris 100 mM (pH 7.6), NaCl 500 mM, EDTA 1mM, PMSF 1%. The solution was sonicated non-continuously for 10 minutes. The lysate was centrifuged at 40,000g for 15 min at 4ºC. Supernatants were applied to a P11-phosphocellulose column equilibrated in buffer A (50 mM Tris–HCl (pH 7.6), 200 mM NaCl, 0.1 mM EDTA) and eluted at 600 mM NaCl. Fractions containing NIC\_RA3 were pooled, diluted to 200 mM NaCl, loaded to a Hi-Trap SPHP (Amersham), and eluted with a linear NaCl gradient (200 mM–1000 mM NaCl) in buffer A. Finally, a gel filtration step was carried out in a Superdex75HR column (Amersham) equilibrated in 100 mM Tris–HCl (pH 7.6), 500 mM NaCl, 0.1 mM EDTA.

***3. Purification of NIC\_RA3-His6****,* ***NIC\_RA3V108A-His6, NIC\_RA3V108E-His6***

Overproduction of NIC\_RA3-His6 was done as above from *E. coli* C41(DE3) cells harbouring pET29c:*nic* plasmid or pET29c:*nicV108A* plasmid, or pET29c:*nicV108E* plasmid. After harvesting the cells, the pellet was resuspended in 20 ml of buffer A (50 mM Tris–HCl pH 7.5, 1 M NaCl) containing 0.1 mM PMSF and then sonicated and cleared by centrifugation. Supernatant was loaded onto a 5-ml His-Trap column (GE Healthcare) equilibrated with buffer A. Bound proteins were eluted with a linear gradient of buffer B (50 mM Tris–HCl pH 7.5, 1 M NaCl, 0.5 M imidazole). After His-Trap chromatography, next purification steps were heparin column chromatography and S75 column chromatography as described above for untagged NIC\_RA3.