Structural basis of a novel histidine-DNA nicking/joining mechanism for gene transfer and promiscuous spread of antibiotic resistance

Radoslaw Pluta[a,b,1,2], D. Roeland Boer[a,b,1,3], Fabián Lorenzo-Díaz[c,4], Silvia Russi[a,b,5], Hansel Gómez[a,d], Cris Fernández-López[c], Rosa Pérez-Luque[a,b], Modesto Orozco[a,d,e], Manuel Espinosa[c] and Miquel Coll[a,b,5]

[a]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028 Barcelona, Spain; [b]Institute of Molecular Biology of Barcelona (IBMB-CSIC), Baldiri Reixac 10-12, 08028 Barcelona, Spain; [c]Centro de Investigaciones Biológicas (CIB-CSIC), Ramiro de Maeztu, 9, 28040 Madrid, Spain; [d]Joint BSC-IRB Research Program in Computational Biology, Institute for Research in Biomedicine. Barcelona, Spain; [e]Department of Biochemistry and Molecular Biology, University of Barcelona, 08028 Barcelona, Spain

[1]R.P. and D.R.B contributed equally to this work.
[2]Present address: International Institute of Molecular and Cell Biology in Warsaw, 02-109 Warsaw, Poland.
[3]Present address: CELLS-ALBA Synchrotron Light Source, 08290 Cerdanyola del Vallès, Spain.
[4]Present address: Dept. Biochemistry, Microbiology, Cell Biology, and Genetics Universidad de La Laguna, and Unidad de Investigación, Hospital Universitario Nuestra Señora de Candelaria, Santa Cruz de Tenerife, Spain.
[5]Present address: Stanford Synchrotron Radiation Lightsource / SLAC National Accelerator Laboratory, 2575 Sand Hill Road, MS 99, Menlo Park, California 94025, USA.
[5]To whom correspondence may be addressed. Email: miquel.coll@irbbarcelona.org

Significance Statement

Nearly 90% of lethal antibiotic-resistant infections in the US are caused by Gram-positive pathogens—*Staphylococcus aureus* accounting for over half of them (11,300 deaths/year). Antibiotic resistance is often encoded by plasmids and integrative elements that are exchanged freely between bacteria through conjugative DNA transfer. During conjugation, a relaxase protein binds, nicks, and covalently attaches to the 5'-end of the DNA molecule, guiding it to the recipient cell, where it restores its circular closed form. We show that relaxase MobM from the promiscuous plasmid pMV158 uses a hitherto unseen mechanism for DNA nicking/closing that is based on the formation of a protein-DNA phosphoramidate adduct. Moreover, our analysis reveals that MobM-like histidine relaxases account for 85% of all relaxases in *S. aureus* isolates.

Abstract

Relaxases are metal-dependent nucleases that break and join DNA for the initiation and completion of conjugative bacterial gene transfer. Conjugation is the main process through which antibiotic resistance spreads among bacteria, multidrug-resistant staphylococci and streptococci infections being a major threat to human health. The MOB_V family of relaxases accounts for about 85% of all relaxases found in *Staphylococcus aureus* isolates. Here we present six structures of the MOB_V relaxase MobM from the promiscuous plasmid pMV158 in complex with several origin of transfer DNA fragments. A combined structural, biochemical and computational approach reveals that MobM follows a previously uncharacterized histidine/metal-dependent DNA processing mechanism, which involves the formation of a covalent phosphoramidate histidine-DNA adduct for cell-to-cell transfer. We discuss how the chemical features of the high-energy phosphorus-nitrogen bond shape the dominant position of MOB_V histidine relaxases among small promiscuous plasmids and their preference towards Gram-positive bacteria.

\body

Introduction

Acquisition of exogenous genetic material by bacteria is achieved via conjugative DNA transfer of mobile genetic elements, such as plasmids and especially Integrative and Conjugative Elements (ICEs), and Integrative and Mobilizable Elements (IMEs) (1). Such processes of horizontal gene transfer (HGT) are considered a strong driving force in bacterial evolution and in the ability of bacteria to colonize different niches (2). In addition to permitting the rapid evolution of the bacterial pangenome, HGT is involved in the acquisition of genetic traits that may confer selective advantages to the recipient bacteria, among them antibiotic resistance (3). This is particularly important when resistance genes encoded by mobile elements are explosively spread among bacteria in hospitals, posing a serious threat to the public health systems (www.cdc.gov/drugresistance/threat-report-2013; www.who.int/drugresistance/documents/surveillancereport/en/). Thus, the so-called mobilome (4) participates in the spread of antibiotic resistance, that is expected to cause 10 million casualties per year by 2050 and the consequent huge economic burden (http://amr-review.org/sites/default/files/Report-52.15.pdf). This has aroused a unanimous claim for novel approaches to deal with infectious diseases caused by pathogenic bacteria (5).

A main performer in HGT is a protein termed relaxase, which is a topoisomerase-like enzyme that cleaves supercoiled plasmid DNA in a strand and sequence-specific manner and ligates it after cell-to-cell transfer. Relaxases start DNA transfer by conjugation upon recognition of their target DNA, the origin of transfer (*oriT*), upon which they mediate generation of a hairpin-loop structure that leaves the dinucleotide to be cleaved (the *nic* site) in a single-stranded (ss) DNA configuration (Fig. 1B) (6, 7). On the *oriT*, the relaxase assembles with other proteins participating in HGT, namely the coupling protein (CP) and the proteins that constitute the Type-IV Secretion System (T4SS). To date, relaxase-mediated nucleophilic attack at the *nic* site has been shown to generate a covalent linkage between a tyrosine residue and the 5'-phosphate DNA of the cleaved dinucleotide (8). This reaction leaves a free 3'-hydroxyl end that serves as primer for DNA

replication by conjugative rolling-circle replication (RCR) (9, 10). The current model for conjugation hypothesizes that the covalent phosphotyrosine DNA-relaxase complex is pumped to the recipient cell by means of the CP and the T4SS (11, 12). In the recipient cell, as shown for small mobilizable plasmids that replicate by the RCR mechanism, the transferred ssDNA is converted into double-stranded (ds) DNA molecules by replication from a lagging strand origin (13), followed by a second relaxase-mediated reaction to close the newly synthesized strand and supercoiling of the dsDNA by the recipient gyrase (9).

On the basis of homologies along the first 300 residues, relaxases were classified into six families: MOB_F, MOB_Q, MOB_P, MOB_V, MOB_H, MOB_C (14). The first four families include 'classical' plasmid-encoded conjugative relaxases that belong to the superfamily of HUH endonucleases (8, 14). These endonucleases are characterized by the presence of the HUH motif (two His residues used for metal coordination, separated by a hydrophobic residue) and the Y motif containing either one (Y1) or two (Y2) active site Tyr residues which deliver the nucleophilic hydroxyl group(s) for endonuclease-recombinase reactions (8). HUH endonucleases depend on a single metal ion for the activation of the scissile phosphate for DNA cleavage and formation of a transient protein-5'-DNA covalent adduct, and finally for the joining of the nicked DNA ends. Crystal structures of members of the Y1 and Y2 relaxase families are known; from the Y1-MOB_Q relaxases: MobA_pR1162 (15) and NES_pLW043 (16), and from the Y2-MOB_F relaxases: TrwC_pR388 (17, 18), TraI_pF (19), and TraI_pCU1 (20).

Biochemical and biophysical, but not structural, studies have been performed on relaxases encoded by small plasmids, most of them belonging to the MOB_V family (also termed MOB_{Pre}) (14, 21). To date, the best characterized relaxase encoded by these plasmids is MobM from the promiscuous plasmid pMV158, which has become the prototype for the MOB_{V1} subfamily of relaxases (14, 21). To advance in our knowledge of these proteins, we undertook the structural characterization of the relaxase domain of MobM (MobMN199; a construct that retains DNA-binding and relaxase activities (7, 22). The different crystal structures solved in complex with *oriT* DNA sequences reveal that, unexpectedly, MOB_V relaxases use a histidine (H22 of MobM), and not a tyrosine,

as catalytic residue. Prompted by this finding, which implies a hitherto undescribed metal-dependent mechanism for DNA cleavage, protein-DNA adduct formation, and DNA ligation, we dissected the active site mechanism by mutating all the residues involved and tested the functional effect. In addition, we performed a computational study to describe the catalytic mechanism at the atomistic scale. Interestingly, the theoretical calculations show that a conserved glutamate (E129 of MobM) serves as: i) a general acid by protonating the 3' oxygen-leaving group in the nicking reaction (while the phosphodiester bond is broken by H22), and ii) a general base abstracting a proton from the 3'-hydroxyl nucleophilic end in the DNA-joining reaction (phosphoramidate bond breaking). Our results unravel a singular catalytic mechanism for the metal ion-dependent role of the conserved histidine residue present in all MOB_V relaxases, thereby stressing the role of these proteins in the spread of the thousands of small mobile genetic elements found in many pathogenic Gram-positive bacteria, especially streptococci, and staphylococci.

Results

Abundance and features of relaxases of the MOB_V family. An in-depth search of the bacterial mobilome (Mob_Pre search in NCBI CDD; www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) retrieved more than 5,830 relaxases of the MOB_V family (Fig. S1, Table S1). Most of them belonged to Firmicutes phylum (5,765 hits), and the top-listed species were encountered in life-threatening bacteria like *Staphylococcus aureus* (3,432 hits), *Streptococcus agalactiae* (453 hits), *Enterococcus faecalis* (200 hits), and *E. faecium* (109 hits). Genomic data for some pathogenic bacteria like *S. aureus* are over-represented in the data base, and thus a high number of hits is obtained. However, strikingly, MOB_V members cover about 85% of all conjugative relaxases reported in the NCBI Conserved Domains Database for *S. aureus* (Table S1). Moreover, when comparing the number of MOB_V relaxases found in *S. aureus* (the NCBI CDD data) with the number of *S. aureus* genomes and plasmids available in the NCBI Genome resource, it can be appreciated that up to 72% of all sequenced isolates of *S. aureus* contain MOBv relaxases (vs 13% with other types of relaxase and

~15% without a relaxase discovered to date, relaxases; Table S2). Similar analysis for other bacteria revealed that MOBv relaxases are frequent in the genomes of *Enterococcus faecium* (68%), *E. faecalis* (45%) and *S. agalactiae* (56%) and very rare in *S. pneumoniae* (3%) and *S. pyogenes* (3%). Furthermore, a NCBI Nucleotide BLAST search on the sequence of plasmid pMV158 *oriT* (*oriT*$_{pMV158}$, 41 nucleotides) revealed over 100 almost identical sequences, found mainly in *S. aureus genomes* (Table S3). This figure could be much higher if we consider that MobM can cleave sequences that share ~70% identity with the *oriT*$_{pMV158}$ (23). With the advent of inexpensive genomic sequencing most MOB$_V$ relaxases are being found to be encoded in bacterial genomes within various integrative elements; however they were traditionally found mainly in RCR plasmids, MobM being representative of the entire family (21). Depending on homologies at their N-200 residues, these relaxases are grouped into two main subfamilies represented by plasmids pMV158 (MOB$_{V1}$) and pBBR1 (MOB$_{V2}$) (14). A further search for plasmids harboring relaxases of the MOB$_{V1}$ subfamily retrieved a high number of mobile elements ranging from 2.7 to 30 kb; half of them encoding antibiotic resistance genes or other virulence traits (21). The genetic organization of the mobilization region was similar in all members: i) the *oriT* was located upstream, and close to the *mob* gene; and ii) the *nic* site was found to be in the same strand as the *oriT* (21). In the case of the MOB$_V$ relaxases, the following three conserved motifs, from the N- to C-terminus, were described (21): i) motif I, H(N/D)(Q/E)R, of no assigned function; ii) motif II, NY(D/E)L, in which the Y residue (Y44 of MobM) was proposed to be the catalytic one (24), and iii) motif III, HxDExTPHMH which contains the HUH metal-coordination signature of all relaxases (8). Curiously, residue Y44 of MobM is not conserved in all members of the MOB$_V$ family (14). Furthermore, single mutants of all seven tyrosine residues of the Mob$_{pBBR1}$ relaxase (a member of the MOB$_V$ family) do not substantially alter the plasmid mobilization frequencies. This observation does therefore not support a ‘classical’ single tyrosine-mediated initiation of conjugation of this particular plasmid (25). These findings prompted us to tackle the structural features of the MobM relaxase in complex with its target DNA.

General MobM structure. Six crystal structures of the MobMN199 relaxase (residues 2-196; (7) in complex with various DNA oligonucleotides were determined to 1.9-3.1Å resolution (Table S4). The experimental phases were obtained by the SAD method, using a selenomethonine protein derivative data set. The overall structure of MobMN199 shows that it has the relaxase α/β plate fold, which is reminiscent of the palm domain of the DNA polymerase Klenow fragment (Fig. 1). It consists of a central β sheet formed by 5 anti-parallel β-strands (with topology β1, β3, β5, β4, β2) that are flanked on each side by a pair of α-helices, α1/α2 and α4/α6, respectively. The active site is located on one face of the β sheet. The histidine triad (H126 and the HUH motif H133 and H135) that coordinates the manganese ion required by the protein (7) is located on strands β4 and β5, whereas the catalytic histidine (H22; see below) is located at helix α1 (Fig. 1). The extended α6-α7 loop and the last α-helix (helix α7) create a thumb (Region 5) that wraps around the ssDNA substrate and directs it into the active site (Fig. 1 and 2). On the opposite side, the active site is flanked by a β-turn located on the α1-3_{10} loop (Region 2) and, from the bottom, by the β4-β5 loop, which is partially ordered in a β-turn motif. Finally, the α1 helix (Region 1), which includes the catalytic H22, closes the active site from the top. Regarding the MobM elements that bind the dsDNA, the α2-β3 loop (Region 3) interacts with the minor groove of the DNA hairpin, and the fragment composed of the β6-β7 hairpin, the α5 helix and the α5-α6 loop (Region 4) interact with major groove. Details of these interactions are described in the *MobM-DNA interactions* section below.

DNA structure. The $oriT_{pMV158}$ is relatively complex, since it contains three inverted repeats (IR1, IR2 and IR3) that can generate mutually exclusive DNA hairpins (7, 21) (Fig. S3C). The IR3 sequence spans 31 nucleotides upstream of the *nic* site and is formed by extension of the IR1 hairpin (18 nucleotides) by a 3-base ATA bulge (IR3 right arm) and a 5-bp IR3 base. The formation of the IR1 or IR3 hairpin excludes the formation of the IR2 hairpin, the function of which remains unknown.

The DNA oligonucleotides used for crystallization trials were chosen based on previous work that determined the minimal substrate for MobM binding, namely the IR3 sequence without the first five nucleotides (7). We crystallized two types of DNA substrates mimicking the IR1/IR3 hairpin in complex with MobM (Fig. 1C). One type consisted of a continuous 26-base oligonucleotide upstream from the *nic* site which covers a sequence of the IR1 hairpin extended by 8-bases of the IR3 right arm (DNA26 oligonucleotide; bases 6-31 of $oriT_{pMV158}$, herein numbered 1-26; Fig. 1C). The second type comprised two annealed oligonucleotides of the IR1 left arm and the IR3 right arm, resulting in a DNA molecule that lacks the 4-base loop of the IR1/IR3 hairpin (the nic series of oligonucleotides; bases 6-12 and 17-32 of $oriT_{pMV158}$; numbered 1-7 and 12-27). The DNA molecules used for protein-DNA complex crystallization either i) terminate at the 3'-oxygen before the cleavage site (dna26 and nic0 oligonucleotides), which mimics the 5'-side DNA cleavage product (and later serves as substrate in the DNA re-joining reaction), or ii) include the scissile phosphate (nic0+P, nic0+SP and nic0+1 oligonucleotides), which represents the cleavage substrate. In both types of MobM-DNA structure the DNA helix of the hairpin stem adopts the B-form. However, the sequence includes a small A-tract (A14-A16) that differs from the canonical B-DNA (Fig. 2). A-tracts have a compressed minor groove, high-propeller twisted A:T base pairs (bps), and bifurcated H-bonds at the major groove (26). This conformation is related to the DNA recognition at the stem by the β-turn RxD/N motif (see below).

The IR1 downstream sequence forms an extended single-stranded structure with the exception of a Thy23-Gua26 wobble base pair that holds the ends of a U-like turn and directs the scissile phosphate into the active site of the protein (the Nic series of structures). In the DNA26 structure, bases of the hairpin loop of one protein-DNA complex enter the active site of a neighboring complex (Fig. S2C) and form non-canonical base pairs with DNA bases of that neighboring complex (wobble Thy23-Gua8' and sheared Gua24-Ade9' base pairs); this configuration displaces the metal ion and the DNA substrate from the active site (Fig. S2D). Intriguingly, in both types of MobM-DNA structures the guanine that forms the wobble Thy-Gua bp superimposes perfectly (Gua26 in the Nic series

of structures, and Gua8' in the DNA26 structure), thereby suggesting a preferred binding site for guanine in this region of the MobM-DNA complex.

MobM-DNA interactions. The specific recognition of *oriT*$_{pMV158}$ by MobM is achieved through formation of nine protein side chain-nucleobase hydrogen bonds between seven protein residues (R7, K10, R71, R74, K149, Q161 and H186) and eight DNA bases (Thy4, Ade6, Ade16, Gua17, Ade19, Ade21, Gua22 and Thy23) (Fig. 2). Interactions with nucleobases also include seven hydrogen bonds between the protein backbone (M8, F152, S182, A184 and K187) and four bases (Thy20, Ade21, Gua22 and Gua24). The rest of the protein-DNA interactions include 14 hydrogen bonds between 11 side chains of MobM and seven sugar-phosphate moieties of the DNA backbone, five hydrogen bonds between protein and DNA backbones and a number of van der Waals contacts (Fig. 2).

In all the structures solved herein, MobM binds the DNA hairpin through a long track of positively charged residues that stack out from the protein surface (Fig. S2F). The specific interactions within the hairpin stem region are limited to four bases interacting with the conserved R71, R74 and K149 residues (Fig. 2 and S3A; note that the R71-Ade6 interaction is visible only in the DNA26 structure, which has the full DNA hairpin - Fig. S3A). In the R71 to R74 stretch, which is part of the α2-β3 loop, the protein backbone runs extended along the minor groove, with main-chain amide to phosphate backbone interactions. This conformation allows the deep penetration of the two arginines in the minor groove, which narrows down to 6.8Å (O4'-O4' distance), sandwiching the R74 guanidinium moiety between opposite sugars across the groove. The side chain of R74 is stabilized inside the minor groove by the side chain of D76, with residues 74 to 77 forming a β-turn. This feature (β-turn RxD/N motif) appears to be a characteristic DNA binding motif that is crucial for the specific recognition of the *oriT* hairpin stem (see below and Fig. S4).

Whereas specific recognition of the IR1 hairpin stem is mediated by only three residues, (R71, R74 and K149), which contact only a few bases, a different situation arises for the ssDNA bases downstream IR1, namely Ade19-Thy20-

Ade21 of the ATA bulge and Gua22-Thy23-Gua24-Thy25-Gua26 of the right arm of the IR3 stem (IR3-R). Two kinks are observed in the DNA backbone between Thy18-Ade19 and Thy20-Ade21 nucleotides of the ATA bulge, which disrupts the stacking of respective bases (Fig. 2D and S3B). The C-terminal thumb folds over the DNA backbone at the bases Gua22 and Thy23, interacting with every phosphate or nucleobase until reaching the scissile phosphate (Fig. 2A, 2D and S3B). With the exception of Thy25, which is solvent-exposed, all bases turn towards the protein and are, in fact, embedded into the protein surface, whereas the phosphate backbone turns outward towards the solvent (Fig. 2A and 2D-E).

The role of conserved H186 needs to be highlighted since its nitrogen ND1 atom forms a hydrogen bond with oxygen O4 of Thy23, thereby stabilizing the wobble base pair Thy23:Gua26. Moreover, the imidazole ring of H186 stacks with the guanidinium group of R7, which itself forms hydrogen bonds with Gua22 (Fig. 2E and S3B). The last protein-nucleobase specific interactions are found between the Gua24 base and K187 peptide bond; however, a number of hydrogen bonds between protein residues and the DNA backbone are also found in this region. Additionally, the Gua24 base is involved in a stacking interaction with F192, which appears to be relevant given the fully conserved ring character of the residues (Phe, Tyr or His) at that position (Fig. S3D). Finally, the formation of a Thy23-Gua26 wobble bp leads to positioning of the scissile phosphate in the center of the MobM active site (Fig. 2A and 2E).

Structural comparison of MobM and other relaxases. In spite of very low sequence homology (9 to 13% sequence identity for structurally superimposable residues), structural superimposition of the structure of MobM on that of other relaxases, i.e. TrwC_pR388, TraI_F and TraI_pCU1 from the MOB_F family and MobA_pR1162 and NES_pLW043 from the MOB_Q family, demonstrate general structural similarity among these proteins (2.9 to 4.1Å C^{α} rmsd for structurally superimposable residues) (Fig. S4A). Nonetheless, we observe significant differences that distinguish MobM from members of the other families. In general, MobM appears to be a simplified version of the fold with smaller loops and less additional secondary structure decoration. In this respect, it is closer to the

structure of the MOB_Q family members MobA_pR1162 and NES_pLW043, which are also shorter than the MOB_F family representatives. Notably, in MobM, the C-terminal thumb directing ssDNA to the active site (residues 175-196) adopts an extended conformation plus one α-helix (α7) while in MOB_F relaxases, like TrwC, it is formed by a much longer stretch (residues 210-265) that adopts a mainly α-helical conformation. In MOB_Q relaxases, this element lacks structural description, since the MobA structure is based on a shorter protein construct (MobA_1-184), and for the NES relaxase construct (NES_2-220) the C-terminal residues that would form a thumb (i.e. residues 196-220) are untraceable in the crystal structure. The second evident difference is that MobM (MOB_V) and NES (MOB_Q) lack the extensive secondary structures that interact with the tip of the DNA hairpin, in contrast to TrwC (MOB_F). As a result, the tip region of the hairpin is not covered by the protein and protrudes from the complex more prominently than for TrwC.

Regarding similarities, all three MobM, NES, and TrwC relaxase-DNA hairpin crystal structures have a common β-turn RxD/N motif (MobM_74-76:RKD, NES_78-80:RKN and TrwC_75-77:RQD), which in MobM is located at the Region 3 of the α2-β3 loop (Fig. 2A-B and S4C-D). Within the motif, as described above, the Arg and Asp/Asn residues enter the hairpin minor groove, while the middle residue positions itself between the phosphates of the DNA backbone. The superimpositions of MobM structure on that of TraI_F, TraI_pCU1, and MobA suggest that these relaxases also use this β-turn RxD/N motif to bind DNA (TraI_F_68-70: RMD, and MobA_66-68: RAN) (Fig. S4C-D). However, in these cases the comparison is impaired due to lack of DNA hairpins in these crystal structures. An exception is the TraI_pCU1 relaxase, which does not have the RxD/N sequence. Interestingly, TraI_pCU1 binds its cognate DNA weakly and in a sequence-independent manner and it was suggested to rely upon a second DNA binding protein to selectively bind the $pCU1_{oriT}$ (20).

At the major groove of the hairpin, interactions unique to each of the relaxases are found. MobM (MOB_{V1} subfamily) enters the major groove using just one residue (K149) from the short 4-residue α5 helix of Region 4, which defines the minimal major groove-interacting element among all structurally described

relaxases (Fig. S4). The MOB_{Qu} relaxase NES places R151, N154, and Y156 from an 8-residue β hairpin at the major groove (Fig. S4C). In the MOB_{Q1} relaxase MobA, K161 superposes to MobM K149 and presumably would interact in major groove. K161 belongs to an 11-residue insertion that includes R143 which structurally superposes to MobM R71 of the minor groove-interacting Region 3 (Fig. S4C). The $MOB_{F1.1}$ relaxase TrwC and the $MOB_{F1.2}$ relaxase TraI_F also have a single Lys residue (K181 and K179, respectively) equivalent to MobM K149. However, the dsDNA-containing crystal structure of TrwC shows that MOB_F relaxases use additional residues (TrwC R128 and K130; Fig. S4D). These derive from an extensive β-hairpin structure that enters the major groove at the opposite face of the DNA from where the RxD/N motif enters the DNA minor groove. This area of the major groove is solvent-exposed in the MobM complex. Thus, MobM seems to have the minimal structural determinants for hairpin stem recognition.

Active site architecture. In all Nic0 structures of MobM, residues H126, H133, H135 and E129 coordinate the Mn^{2+} ion with octahedral geometry (Fig. 1D and 3) which identity was unequivocally confirmed by anomalous diffraction (Fig. S2E). In Nic0_A, Nic0_B, Nic0+P and Nic0+SP structures the fifth Mn^{2+} ion ligand is invariably the O3' oxygen of Gua26 and the sixth ligand is either the scissile phosphate oxygen (Nic0+P and Nic0+SP), the catalytic H22 (partial occupancy in Nic0_A and full occupancy in Nic0_B), or a water molecule (partial occupancy in Nic0_B).

The asymmetric unit of the Nic0+1 crystal contains two protein-DNA complexes that represent two distinct states: i) a DNA-free active site (complex B in the crystal asymmetric unit - Nic0+1/molB; Fig. 3A), and ii) an active site in which the Gua26 O3' atom and the scissile phosphate OP1 atom are located at a considerable distance from the metal ion (3.6Å and 4.2Å, respectively; note that corresponding distances equal 2.7Å and 2.1Å in the Nic0+P structure and 2.2Å and 2.3Å in the Nic0+SP structure; complex A in the crystal asymmetric unit - Nic0+1/molA; Fig. 3D). In both complexes of the Nic0+1 structure, the metal ion attracts H22 for the interaction instead, thereby showing that in the absence of

the DNA-scissile phosphate ligand, the metal ion ligates the catalytic H22 (Fig. 3A and 3D). In complex B the DNA electron density map ends at phosphate 25, indicating that the last three bases are disorder and most probably do not enter the active site. In this case, the Thy23-Gua26 wobble-bp is not formed and the position of the Gua26 3'-oxygen (O_3') is occupied by a water molecule. In complex A, despite the wobble bp formation and the placement of the DNA substrate in the active site, the Gua26 3'-oxygen and the scissile phosphate are shifted away from the center of the active site.

In all Nic structures, except the Nic0+1/molA structure, a group of fully conserved residues is engaged in the formation of a hydrogen bond network (Fig. 1D and 3), in which: i) D128 contacts N43 and R25; ii) R25 interacts with the H22 peptide bond, the scissile phosphate and the metal-ligating E129; and iii) E129 - depending on the structure- interacts with either the scissile phosphate OP1 oxygen (active site with the cleavage substrate; Nic0+P and Nic0+SP structures; Fig. 3B and 3E) or the Gua26 O3' oxygen (active site with the cleavage product/ligation substrate: the Nic0_A and Nic0_B structures; Fig. 3C and 3F), or the metal-bound water molecule, which superimposes with the Gua26 O3' oxygen (DNA-free active site; Nic0+1/molB structure; Fig. 3D). The Nic0+SP structure accommodates an additional contact for E129, a hydrogen bond between E129 and N32 (Fig. 1D). In the Nic0+1/molA structure this extensive array of contacts is limited to only one, namely the R25 – D128 interaction; this is due to displacement of the R25 side chain towards a location that, in other structures, would create clashes with N43 (Fig. 3D).

Mobility of the catalytic histidine. Depending on the crystal structure the catalytic H22 side chain occupies three distinct positions: the IN position (metal-bound; the NE2 nitrogen atom is 2.2-2.8Å away from the Mn^{2+} ion; Fig. 3A, 3C, 3D and 3F), the INTERMEDIATE position (NE2 nitrogen is 6.2Å away from the Mn^{2+} ion and 4.0Å away from the scissile phosphate phosphorous atom; Fig. 3B) and the OUT position (H22 side chain is far from the Mn^{2+} ion and the scissile phosphate; Fig. 3B, 3E and 3F). For structures that were determined at pH>4.6, the catalytic histidine side chain is placed in the metal-bound IN position in two

settings: i) in the DNA-free and the DNA-misplaced active site structures (Nic0+1/molB and Nic0+1/molA, respectively; Fig. 3A and 3D), and ii) in the presence of the 5' cleavage product (Gua26 3'OH) in the active site (Nic0_A; Fig. 3C). The above two cases show that in the absence of the interaction of the scissile phosphate with the metal ion, H22 ligates this ion. Conversely, the ligation of the scissile phosphate to the active site metal ion causes the catalytic H22 to move away from the metal-bound IN position to the INTERMEDIATE and OUT positions as can be seen in the Nic0+SP structure, which has the H22 side chain in two alternative conformations: the INTERMEDIATE position with major occupancy and the OUT position with minor occupancy (Fig. 3B). Similarly, protonation of the H22 imidazole ring at pH 4.6 induces the histidine side chain to move away from the metal ion, as can be observed when comparing higher and lower pH structures containing either the 5' cleavage product (the IN position at pH5.5 *vs.* the IN/OUT positions at pH4.6 for Nic0_A and Nic0_B structures, respectively; Fig. 3C and 3F) or the cleavage substrate (the INTERMEDIATE/OUT positions at pH6.8 *vs.* the OUT position at pH4.6 for Nic0+SP and Nic0+P structures, respectively; Fig. 3B and 3E). In fact, the relaxation activity of the MobM plasmid is highest at pH 6.5, which is slightly above the histidine imidazole p*Ka* (6.0) and it decreases steeply below pH 6.0, dropping by 30% at pH 4.6 (23). All together, these findings show that the H22 side chain has freedom of movement and that its interaction with the metal ion depends on the stage of the reaction.

One metal-ion catalysis with a histidine as a nucleophile. Nucleases cleave phosphodiester bonds by a general acid–base catalysis, where the general base activates the nucleophile by deprotonation and the general acid facilitates product formation by protonation of the leaving group (27, 28). This reaction goes typically through the following three stages: i) nucleophilic attack; ii) formation of a highly negative pentacovalent transition state, and iii) breakage of the scissile bond. The mechanism requires that the nucleophile, the phosphorus atom and the leaving group be in-line. In the case of metal ion-dependent DNA nucleases, metal ion B, which resides on the leaving group side and is common for one- and two-metal

ion nucleases, stabilizes the pentacovalent transition state and, in certain coordination geometries, can promote the nucleophilic attack by destabilization of the enzyme-substrate complex (29). The enzymes that cleave by a two-metal ion mechanism, use the second metal ion (metal ion A) to activate a catalytic water molecule that acts as a nucleophile, as well as to stabilize the transition state. In the one-metal ion enzymes, the activation role of the second metal ion is performed by a protein residue and the catalytic water molecule can be replaced by a serine or a tyrosine hydroxyl group (28, 30). The later was the case of the first relaxase described (17, 18), and other relaxases analyzed thereafter.

In the case of the relaxase MobM, the protein-DNA structures show that a constellation of amino acids is required for the formation of the cleavage-competent active site that is best reflected in the Nic0+SP structure. In this structure the Mn^{2+} ion interacts with both, the scissile phosphate OP1 oxygen atom (2.2Å distance), and the Gua26 O3' oxygen atom (2.3Å distance). This ion is therefore well positioned to contribute to the stabilization of the electron-rich pentacovalent transition state. Importantly, the catalytic H22 is in the proper location for the nucleophilic in-line attack on the phosphorus atom (H22 NE2 nitrogen being 4.0Å away) and superposes with the position of the catalytic tyrosines in the crystal structures of other HUH endonucleases (MOB_Q relaxases, MOB_F relaxases, RepB replication initiator and TnpA transposase; Fig 4). In the case of superposition to MobA, which is the most similar to MobM among structurally described relaxases, the MobM H22 NE2 nitrogen atom superposes perfectly with the MobA catalytic Y25 hydroxyl oxygen atom (Fig. 4A).

The above findings reveal MobM as the first example of a metal-dependent nuclease that uses a histidine nitrogen atom for the nucleophilic attack on the DNA substrate. Moreover, the use of a histidine nitrogen atom, instead of an oxygen atom of a tyrosine/serine residue, makes MobM the first example of a DNA breaking-joining enzyme that operates through a phosphohistidine adduct. Known examples of nucleases belonging to the phospholipase D (PLD) superfamily of phosphodiesterases hold catalytic histidines and operate through a transient phosphohistidine intermediate adduct. However, these nucleases bear no sequence or structure relationship with relaxases, have separate

domains for substrate recognition and for cleavage, they are metal ion-independent enzymes and they do not support DNA ligation (28). Importantly, the acid dissociation constant of histidine is the closest to the cytoplasmic pH among known nuclease nucleophiles (His pKa=6; Tyr pKa=10; ribose 2'OH p*Ka*=12-14; Ser p*Ka*=13; H_2O p*Ka*=16) (28). Thus, activation of the histidine imidazole group by deprotonation is readily achieved, and no auxiliary residue appears to be required for its activation. In contrast, other relaxases that harbor a catalytic tyrosine, such as TrwC, appear to have a residue that closely interacts with it (an Asp in TrwC) to facilitate proton abstraction.

In the MobM Nic0-SP structure, the metal-bound phosphate OP1 oxygen atom interacts with the R25 NH1 nitrogen atom, taking a position not far from K265 of TraI (19) (Fig. 4B), which was proposed to act as the equivalent of the second metal ion (30). Similarly, in yeast type II topoisomerase the conserved R781 residue contacts and stabilizes the covalent phosphotyrosine moiety of the protein-DNA adduct (31) and mutation of this arginine to a residue other than lysine dramatically reduces DNA cleavage and relaxation activity (32). In MobM R25 is fixed in its position at the active site by a double H-bond interaction with D128.

In the MobM Nic0+SP structure the second oxygen atom of the scissile phosphate, OP2, shares a hydrogen bond with the E129 OE1 oxygen, which further interacts with the N32 NE2 nitrogen, while the E129 OE2 oxygen ligates the metal ion and contacts the R25 NH2 nitrogen (Figs. 1D and 3B). Use of a carboxylic residue as fourth protein ligand for metal ion ligation has not been described for any relaxase to date. Intriguingly, the MobM genetic companion on pMV158, the RepB replication initiation protein (which is not a relaxase although it belongs to the HUH superfamily), in addition to the histidine triad, uses D42 for the Mn ion ligation (Fig. 4C). However, its role other than metal ion coordination was not tested (33). Interestingly, a hydrogen bond between E129 and the DNA 3'OH end in the Nic0_A and Nic0_B structures suggests that E129 acts as a general base in Gua26 3'OH activation for the nucleophilic attack on the H22-DNA adduct during the DNA ligation reaction (see below *Theoretical study of the catalytic mechanism*).

Mutational analysis. To confirm the information derived from the crystal structures, several mutations designed to change key residues of MobM were constructed, and the resulting mutated proteins were purified. The following mutations were done: i) H22 was changed to either Ala (H22A) or Tyr (H22Y); ii) R25 was changed to Ala (R25A); iii) E129 was changed to either Ala (E129A) or to Gln (E129Q), and iv) Y44 was changed to phenylalanine (Y44F), the latter change because Y44 was proposed in a previous study to be the catalytic residue (24). The purified mutant proteins were tested for their ability to relax supercoiled plasmid DNA (nicking reaction) or to generate relaxosomes *in vitro*. This latter assay is based on the selective precipitation of DNA-protein covalent complexes by SDS and KCl (34). In both cases, pMV158 DNA was treated with the purified proteins and the reaction products were analyzed on 1% agarose gels. The assays performed with the H22A, H22Y, and R25A mutants confirmed the essential role of these residues: the wild-type MobM protein generated ~60% open circle (oc) relaxation products, whereas no significant oc-forms were observed for any of these mutant proteins (Fig. 5A). Equivalent results were found when generation of relaxosomes was analyzed: protein-DNA adducts were readily detected in the wild-type, but not in the H22 or R25 mutants. This observation demonstrates the crucial role of amino acids of Motif I in the cleavage and generation of stable DNA-protein adducts (Fig. 5B). Relaxation assays done with the Y44F mutant protein showed that it behaved like the wild-type MobM, thus also confirming the results of crystal structures. Finally, the E129A and E129Q mutants showed a slight (if any) residual relaxation activity (Fig. 5A). Neither of the two E129 mutants were able to generate relaxosomes (Fig. 5B), thus emphasizing the significance of E129 for the formation of the covalent phosphohistidine adduct (see below *Theoretical study of the catalytic mechanism*). Indeed, besides metal coordination, E129 interacts with the scissile phosphate and R25, the latter in turn contacting D128. We conclude that R25 of MobM contributes to anchoring the scissile phosphate and to the stabilization of the transition state and the phosphohistidine adduct, as indicated by the calculation performed (Fig. S5D).

The *in vivo* assays were performed by constructing pMV158-derivative plasmids harboring either the H22A or the Y44F mutations. These plasmids have a single mutation in the *mobM* gene (that is, a single amino acid change, H22A or Y44F, in the entire 495-residue MobM protein). Mobilization assays were done in the conditions previously described (35) using *S. pneumoniae* cells harboring the wild type or the mutated plasmids as donors and either pneumococcal or *Enterococcus faecalis* cells as recipients. The results showed that whereas mobilization frequencies of 2.5(±0.2) x 10^{-4} (average of five experiments) were found for pMV158 and pMV158Y44F plasmids, no transconjugant (frequency < 1.5x 10^{-10}, i.e. the experimental detection limit) was rescued in the five independent experiments performed with pMV158H22A (Table S5). Taken together, the results from the crystal structure analysis and the functional assays demonstrate that the H22 residue is essential for the activity of the protein, thereby supporting its role as the nucleophilic residue in the DNA cleavage reaction.

Theoretical study of the catalytic mechanism. We performed theoretical calculations including classical Molecular Dynamics (MD) simulations and Quantum Mechanics/Molecular Mechanics (QM/MM) free energy calculations to describe the catalytic mechanism performed by MobM. Two MobM-DNA crystal structures were used for MD simulations. One, in which the H22 NE2 nitrogen is positioned in-line for the nucleophilic attack on the reaction center (i.e. phosphorus atom of the scissile P-O bond) from the opposite site of the O3' leaving group, i.e. the Nic0+SP structure, and the other one, in which H22 ligates the Mn ion and the above mentioned atoms are not in line (the Nic0+1/molA structure). We were able to model the covalent complex between H22 and the DNA substrate showing the expected inverted configuration of the phosphorus atom for the former (Fig. S5B) but not for the latter, for which modeling of the nucleolytic reaction failed (Fig. S5C). Moreover, we observed that the formation of the covalent intermediate is possible only when the O3' leaving group atom is assumed to be protonated.

The proposed mechanism of the histidine-mediated attack to the

phosphodiester bond, as well as the corresponding free energy profile, is depicted in Fig. 7A and 7B, respectively. The reaction takes place through an associative S_N2-like mechanism involving a pentacoordinated phosphorus transition state (TS, see Fig. S5B), just as described in the theoretical study of the CheA histidine kinase (36). Moreover, E129 would protonate the O3' atom thus playing a role of general acid in the catalytic mechanism (Fig. 6A). This glutamate is hydrogen-bonded to the OP2 of the target phosphate group in the pre-catalytic structure (Fig. 6A-B), as also suggested by visual inspection of the MobM Nic0+SP structure. During catalysis the HE2 proton from E129 reorients to form a hydrogen bond with its final acceptor atom (i.e. O3') to finally be transferred (Fig. 6A-B). It should be noted that in our modeling study this proton-transfer process takes place spontaneously (i.e. it was not explicitly considered in the reaction coordinates; see Fig. 6B), thereby supporting the hypothesis of E129 acting as the general acid in the nucleolytic reaction. Moreover, the proton transfer process lags behind the nucleophilic attack and is concomitant with the drop in energy observed right after the maximum in the free energy profile (Fig. 6B).

The free energy barrier and reaction energies computed were 14.2 and 5.3 kcal/mol, respectively (Fig. 6B). Moreover, consistent with the very low (if any) residual activity observed for the E129A and E129Q mutants (Fig. 5A) we were not able to find a catalytic pathway for these mutants.

Finally, our calculations also support the electrostatic stabilizing role of R25 and, accordingly, the observed inactivity of the mutant R25A (Fig. S5D).

Phosphoramidate bond chemistry and consequences for pMV158 lifestyle. The phosphoramidate bond in phosphohistidine is thermodynamically less stable than the phosphoester bond in phosphotyrosine (37). Free phosphohistidine has a half-life of seconds in acidic solution, and of a few minutes at pH=7. Such lability would impair the transfer of the protein-DNA adduct across the cell membranes, to the receiving cell. However, in proteins the stability of the P-N bond is influenced by the protein environment, in particular by the side chains surrounding the phosphohistidine at the active site. For instance, a hydrogen

bond to the non-phosphorylated imidazole nitrogen would make the phosphoramidate bond significantly less labile. There are two partially conserved residues, namely N23 and H30, that could establish a hydrogen bond with the H22 ND1 nitrogen, once the adduct is formed. N23 is the residue adjacent to the catalytic histidine, and already makes a hydrogen bond to the H22 ND1 in the Nic0_B structure. H30 is located in a flexible loop after helix α1 where the catalytic histidine is located. This loop is disordered and could not be traced in most of our structures, except in the Nic0+SP structure where H30 ND1 is located 4.4Å away from the H22 ND1 nitrogen, while a water molecule bridges the interaction. A structure of the MobM-His22-DNA adduct could clarify if these residues (or other) are critical in regulating the stability of the P-N bond. Unfortunately, so far we have not been able to crystallize such adduct.

The physiological consequences of the use of a highly transferable phosphohistidine versus highly stable phosphotyrosine bond for the DNA transfer process in MOB_V relaxases remain to be unveiled. However, we propose that MOB_{V1} relaxases lost the need for a tyrosine hydroxyl and evolved to employ a different group for substrate attack because of the lowest p*Ka* (6.0) of histidine among known nucleophiles, thus making histidine independent from deprotonating/activating auxiliary residues which are, however, required for other nucleophiles (tyrosine p*Ka* ~10-11). From the physiological point of view, we have to consider that most MOB_V relaxases are encoded by Firmicutes with low G+C% content and that many of these bacteria tend to strongly acidify the media (below to pH 5.5), perhaps as a mechanism through which to out-compete other bacterial species. Furthermore, since these small plasmids (average size ~5 kb; (21)) lack partitioning systems, they need to achieve the average copy number (~30 copies per genome equivalent) before cell division so that newborn cells should receive enough plasmid copies to guarantee stable inheritance (38). Having an easy to break relaxase-DNA covalent link would confer a selective advantage, allowing rapid reconstitution of dsDNA-replicating molecules. In this regard it is worth mentioning that the RCR-replicase protein of plasmid pMV158 (RepB), although being a tyrosine HUH endonuclease, generates unusually transient and difficult to capture protein-DNA intermediates (39).

An interesting difference between tyrosine and histidine relaxases is that the former are found in both Gram-positive and Gram-negative bacteria, whereas the newly described histidine relaxase like MobM, appear to be found almost exclusively in the Gram-positive Firmicutes phylum (Fig. S1). It has been suggested that relaxases of the MOB_V family have been evolutionary selected by small/mobilizable plasmids (40). We could relate this notion to the high energy of the P-N bond between the scissile phosphate and the catalytic histidine, which makes the phosphate group readily transferable to other groups (37). In such scenario, the use of a catalytic histidine could facilitate the DNA closing reaction; on the other hand it would also limit the size of the DNA cargo, since the P-N bond would be prone to breakage if the DNA transfer takes too long. MOB_V/Mob_Pre relaxases are also found in long plasmids as in the case of closely related multi-drug resistance plasmids pSK41 (46.4 kb) and pGO1 plasmid (54 kb), the latter being the prototype for class III staphylococcal plasmids (41). However, in these multi-drug resistance/multi-relaxase plasmids, which use MOB_Q-type relaxases for conjugative transfer, the MOB_V relaxase-encoding gene is located (together with aminoglycoside and bleomycin resistance genes) within a transposable 5 kb cassette-like module flanked by the IS257-mediated recombination elements. Whether they support alternative conjugative transfer events of the 5kb integrant or the full plasmids remains to be explored. Another characteristic chemical feature of the phosphoramidate bond is its acidic liability (37). The existence of MOB_V/Mob_Pre nucleases almost exclusively in Gram-positive bacteria could be related to the fact that there is no maintenance of the pH homeostasis in the periplasm of Gram-negative bacteria, i.e. the periplasm pH matches that of the external environment (42). Therefore, conjugal DNA transfer mediated by a His relaxase in Gram-negative bacteria in an acidic environment could lead to the P-N bond exposure to low pH and breakage of the relaxase-DNA complex during the transfer through type IV secretion system (T4SS) spanning the periplasm. On the contrary, in the case of Gram-positive bacteria, conjugation is mediated by a cell-wall hydrolase (43), which allows intimate cell-to-cell contacts without the need for the transferred DNA to be exposed to extracellular low pH.

Conclusion

The work presented here provides the first structure-function analysis of the representative member of the MOB_{V1} family of relaxases. Our combined approach of X-ray crystallography, protein-DNA biochemistry, *in vivo* and computational has allowed us to describe a novel metal-dependent histidine nucleolytic catalysis, represented by the MobM relaxase, which is encoded by a promiscuous plasmid actively involved in the spread of antibiotic-resistance. Homologs of MobM are found in many plasmids and other mobile genetic elements of pathogenic bacteria such as *S. aureus* (Table S1 and S2). MobM is the first example of a metal-dependent nuclease that uses histidine nitrogen for the nucleophilic attack on the scissile phosphate. Furthermore, it is also the first example of a histidine relaxase, a DNA-breaking and -joining enzyme, that operates through a phosphorus-nitrogen protein-DNA adduct for cell-to-cell DNA transfer.

Materials and Methods

A detailed description of protein purification, relaxation and, mobilization assays, crystallography and theoretical calculations can be found in SI Materials and Methods.

Acknowledgements

We thank W.T.Chan for construction of plasmid pMV158H22A, L. Rodríguez for help in protein purification, the staff of the Platform for Automated Crystallization (IBMB, Barcelona, Spain), and D. Aparicio and L. Martinelli for collecting the Nic0+1 diffraction data. IRB Barcelona is the recipient of a Severo Ochoa Award of Excellence from the Spanish Ministry of Economy and Competitiveness (MINECO). IBMB-CSIC Structural Biology Unit is the recipient of a Maria de Maeztu Award of Excellence from MINECO. This work was supported grants from MINECO (BFU2008-02372/BMC, CSD-2006-23, BFU2011-22588, CSD-2008/00013, BIO2013-49148-C2-2-R, BFU2014-53550-P, BIO2015-69085-REDC and BIO2015-64802), the Generalitat de Catalunya (2014-SGR134, 2014-

SGR1530) and the European Commission (GA No. 260644, ERC_SimDNA and H2020 program (MuG and BioExcel projects)). RP and HG received a La Caixa/IRB Barcelona International PhD and Juan de La Cierva fellowship, respectively. MO is an ICREA Academia researcher.

References

1. Bellanger X, Payot S, Leblond-Bourget N, & Guédon G (2014) Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. *FEMS Microbiology Reviews* 38(4):720-760.
2. Wozniak RAF & Waldor MK (2010) Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat. Rev. Microbiol.* 8:552-563.
3. Lanza VF, Tedim AP, Martinez JL, Baquero F, & Coque TM (2015) The plasmidome of Firmicutes: Impact on the emergence and the spread of resistance to antimicrobials. *Microbiol Spectr* 3(2):PLAS-0039-2014.
4. Siefert JL (2009) Defining the mobilome. *Methods Mol. Biol.* 532:13 - 27.
5. Chan WT, Balsa D, & Espinosa M (2015) One cannot rule them all: Are bacterial toxins-antitoxins druggable? *FEMS Microbiol. Rev.* 39:522-540.
6. Cabezón E, Ripoll-Rozada J, Peña A, de la Cruz F, & Arechaga I (2015) Towards an integrated model of bacterial conjugation. *FEMS Microbiology Reviews* 39:81-95.
7. Lorenzo-Díaz F*, et al.* (2011) The MobM-relaxase domain of plasmid pMV158: thermal stability and activity upon Mn^{2+}-and DNA specific-binding. *Nucleic Acids Res.* 39:4315-4329.
8. Chandler M*, et al.* (2013) Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nature Rev. Microbiol.* 11:625-538.
9. Espinosa M (2013) Plasmids as models to study macromolecular interactions: the pMV158 paradigm. *Res. Microbiol.* 164:199-204.
10. Gomis-Rüth FX & Coll M (2006) Cut and move: protein machinery for DNA processing in bacterial conjugation. *Curr. Opin. Struct. Biol.* 16:744-752.
11. Gomis-Ruth FX, Sola M, de la Cruz F, & Coll M (2004) Coupling factors in macromolecular type-IV secretion machineries. *Current pharmaceutical design* 10(13):1551-1565.
12. Llosa M, Gomis-Rüth FX, Coll M, & de la Cruz F (2002) Bacterial conjugation: a two-step mechanism for DNA transport. *Mol. Microbiol.* 45:1-8.
13. Kramer MG, Khan SA, & Espinosa M (1997) Plasmid rolling circle replication: identification of the RNA polymerase-directed primer RNA and requirement of DNA polymerase I for lagging strand initiation. *EMBO J.* 16:5784-5795.
14. Garcillan-Barcia MP, Francia MV, & de la Cruz F (2009) The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev* 33(3):657 - 687.
15. Monzingo AF, Ozburn A, Xia S, Meyer RJ, & Robertus JD (2007) The structure of the minimal relaxase domain of MobA at 2.1 Å resolution. *J. Mol. Biol* 366:165-178.
16. Edwards JS*, et al.* (2013) Molecular basis of antibiotic multiresistance transfer in *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. USA* 110(8):2804-2809.
17. Boer 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(3):857-869.
18. Guasch A*, et al.* (2003) Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nat. Struct. Biol.* 10:1002-1010.
19. Larkin C*, et al.* (2005) Inter- and Intramolecular Determinants of the Specificity of Single-Stranded DNA Binding and Cleavage by the F Factor Relaxase. *Structure* 13(10):1533-1544.
20. Nash RP, Habibi S, Cheng Y, Lujan SA, & Redinbo MR (2010) The mechanism and control of DNA transfer by the conjugative relaxase of resistance plasmid pCU1. *Nucl.*

Acids Res. doi:10.1093/nar/gkq303:1-15.
21. Lorenzo-Díaz F, Fernández-López C, Garcillán-Barcia MP, & Espinosa M (2014) Bringing them together: Plasmid pMV158 rolling circle replication and conjugation under an evolutionary perspective. *Plasmid* 74:15-31.
22. Fernández-López C*, et al.* (2013) Functional properties and structural requirements of the plasmid pMV158-encoded MobM relaxase domain *J. Bacteriol.* 195:3000-3008.
23. Fernández-López C*, et al.* (2013) Nicking activity of the pMV158 MobM relaxase on cognate and heterologous origins of transfer. *Plasmid* 70:120-130.
24. de Antonio C, Farias ME, de Lacoba MG, & Espinosa M (2004) Features of the plasmid pMV158-encoded MobM, a protein involved in its mobilization. *J. Mol. Biol.* 335(3):733-743.
25. Szpirer CY, Faelen M, & Couturier M (2001) Mobilization function of the pBHR1 plasmid, a derivative of the broad-host-range plasmid pBBR1. *J. Bacteriol.* 183:2101-2110.
26. Coll M, Frederick CA, Wang AH, & Rich A (1987) A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc. Natl. Acad. Sci. USA* 84(23):8385-8389.
27. Lassila JK, Zalatan JG, & Herschlag D (2011) Biological Phosphoryl-Transfer Reactions: Understanding Mechanism and Catalysis. *Annual Review of Biochemistry* 80(1):669-702.
28. Yang W (2011) Nucleases: diversity of structure, function and mechanism. *Quarterly reviews of biophysics* 44(1):1-93.
29. Yang W, Lee JY, & Nowotny M (2006) Making and Breaking Nucleic Acids: Two-Mg^{2+}-Ion Catalysis and Substrate Specificity. *Molecular Cell* 22(1):5-13.
30. Yang W (2008) An equivalent metal ion in one- and two-metal-ion catalysis. *Nature structural & molecular biology* 15(11):1228-1231.
31. Schmidt BH, Burgin AB, Deweese JE, Osheroff N, & Berger JM (2010) A novel and unified two-metal mechanism for DNA cleavage by type II and IA topoisomerases. *Nature* 465(7298):641-644.
32. Okada Y*, et al.* (2000) Assignment of functional amino acids around the active site of human DNA Topoisomerase IIα. *Journal of Biological Chemistry* 275(32):24630-24638.
33. Boer DR*, et al.* (2009) Plasmid replication initiator RepB forms a hexamer reminiscent of ring helicases and has mobile nuclease domains. *EMBO J.* 28:1666-1678.
34. Trask DK, DiDonato JA, & Muller MT (1984) Rapid detection and isolation of covalent DNA/protein complexes: application to topoisomerase I and II. *EMBO J* 3:671-676.
35. Lorenzo-Díaz F & Espinosa M (2009) Large-scale filter mating assay for intra- and inter-specific conjugal transfer of the promiscuous plasmid pMV158 in Gram-positive bacteria. *Plasmid* 61(1):65-70.
36. Shi T*, et al.* (2011) Mechanism for the autophosphorylation of CheA Histidine Kinase: QM/MM Calculations. *The Journal of Physical Chemistry B* 115(41):11895-11901.
37. Attwood PV, Piggott MJ, Zu XL, & Besant PG (2007) Focus on phosphohistidine. *Amino Acids* 32:145-156.
38. Novick RP (1998) Contrasting lifestyles of rolling-circle phages and plasmids. *Trends Biochem. Sci.* 23:434-438.
39. Moscoso M, Eritja R, & Espinosa M (1997) Initiation of replication of plasmid pMV158: mechanisms of DNA strand transfer reactions mediated by the initiator RepB protein. *J. Mol. Biol.* 268:840-856.
40. Smillie C, Garcillan-Barcia MP, Francia MV, Rocha EPC, & de la Cruz F (2010) Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* 74(3):434-452.
41. Caryl JA & O'Neill AJ (2009) Complete nucleotide sequence of pGO1, the prototype conjugative plasmid from the staphylococci. *Plasmid* 62(1):35-38.
42. Slonczewski JL, Fujisawa M, Dopson M, & Krulwich TA (2009) Cytoplasmic pH measurement and homeostasis in Bacteria and Archaea. *Advances in Microbial Physiology*, ed Robert KP (Academic Press), Vol Volume 55, pp 1-317.
43. Arends K*, et al.* (2013) TraG encoded by the pIP501 Type IV Secretion System is a two-domain peptidoglycan-degrading enzyme essential for conjugative transfer. *J. Bacteriol.* 195(19):4436-4444.
44. Studier FW, Rosenberg AH, Dunn JJ, & Dubendorff JW (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Meth. Enzymol.* 185:60-89.
45. del Solar G, Díaz R, & Espinosa M (1987) Replication of the streptococcal plasmid

pMV158 and derivatives in cell-free extracts of *Escherichia coli*. *Mol. Gen. Genet.* 206:428-435.
46. Guzmán LM & Espinosa M (1997) The mobilization protein, MobM, of the streptococcal plasmid pMV158 specifically cleaves supercoiled DNA at the plasmid *oriT*. *J Mol Biol* 266(4):688-702.
47. Leslie AGW & Powell HR (2007) Processing Diffraction Data with Mosflm. *Evolving Methods for Macromolecular Crystallography*, eds Read RJ & Susman JL), Vol 245, pp 41-51.
48. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallographica Section D Biological Crystallography* 62(Pt 1):72-82.
49. Schneider TR & Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallographica Section D Biological Crystallography* 58(Pt 10 Pt 2):1772-1779.
50. McCoy AJ*, et al.* (2007) Phaser crystallographic software. *Journal of applied crystallography* 40(Pt 4):658-674.
51. Cowtan K (2001) Fast Fourier feature recognition. *Acta Crystallographica Section D Biological Crystallography* 57(Pt 10):1435-1444.
52. Terwilliger TC (2003) SOLVE and RESOLVE: automated structure solution and density modification. *Methods in Enzymology* 374:22-37.
53. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography* 60(Pt 12 Pt 1):2126-2132.
54. Murshudov GN, Vagin AA, & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D Biological Crystallography* 53(Pt 3):240-255.
55. Davis IW*, et al.* (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35(Web Server issue):W375-383.
56. DeLano WL (2002) The Pymol Molecular Graphics System. *on World Wide Web http:/*www.pymol.org.
57. Case DA*, et al.* (2014) AMBER 14 (University of San Francisco).
58. Maier JA*, et al.* (2015) ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of Chemical Theory and Computation* 11(8):3696-3713.
59. Ivani I*, et al.* (2016) Parmbsc1: a refined force field for DNA simulations. *Nat Meth* 13(1):55-58.
60. Foster JP & Weinhold F (1980) Natural hybrid orbitals. *Journal of the American Chemical Society* 102(24):7211-7218.

Figures

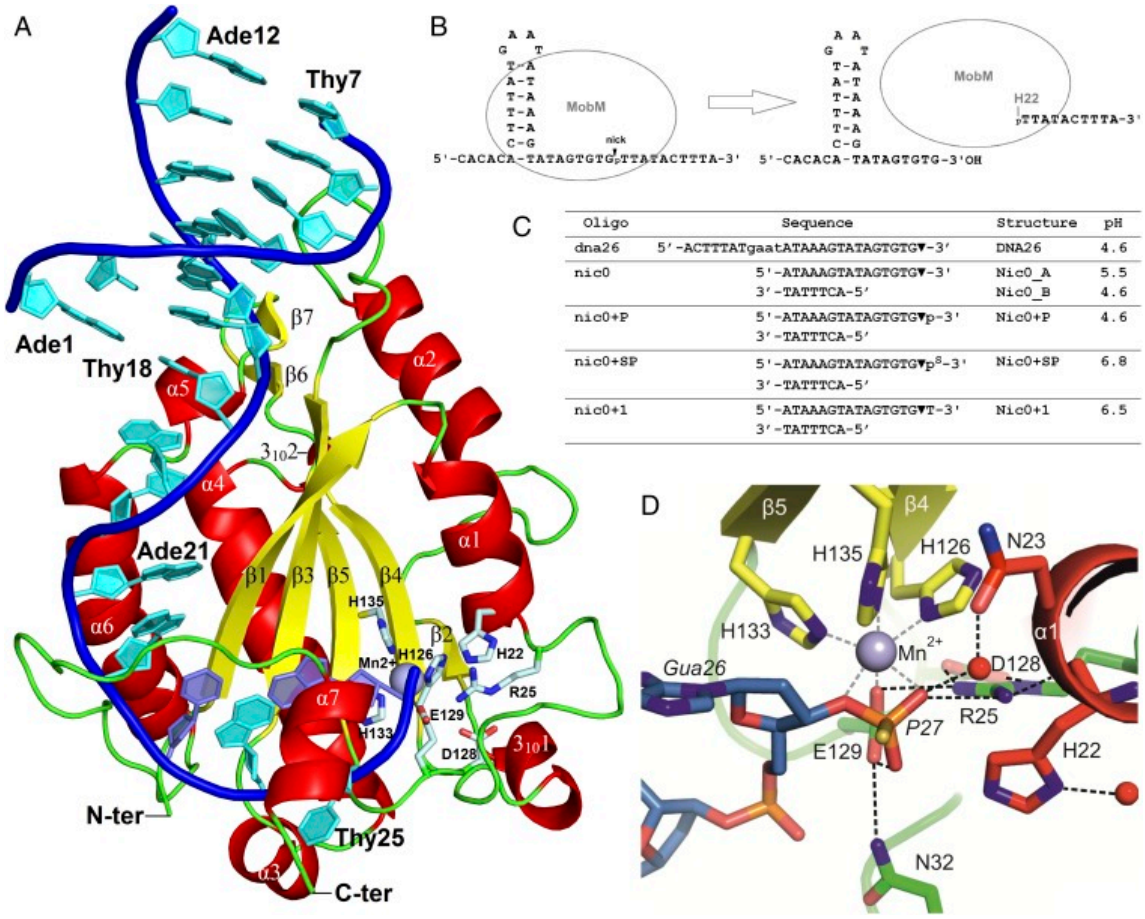

Fig. 1. Structure of the MobM-DNA complex. **(A)** Crystal structure of the N-terminal MobM relaxase domain bound to the nic0+SP oligonucleotide containing the scissile phosphate (Nic0+SP structure). (B) Scheme of oriT processing by MobM. (C) Oligonucleotides used for crystallization, structure names and pH. (D) Active site details.

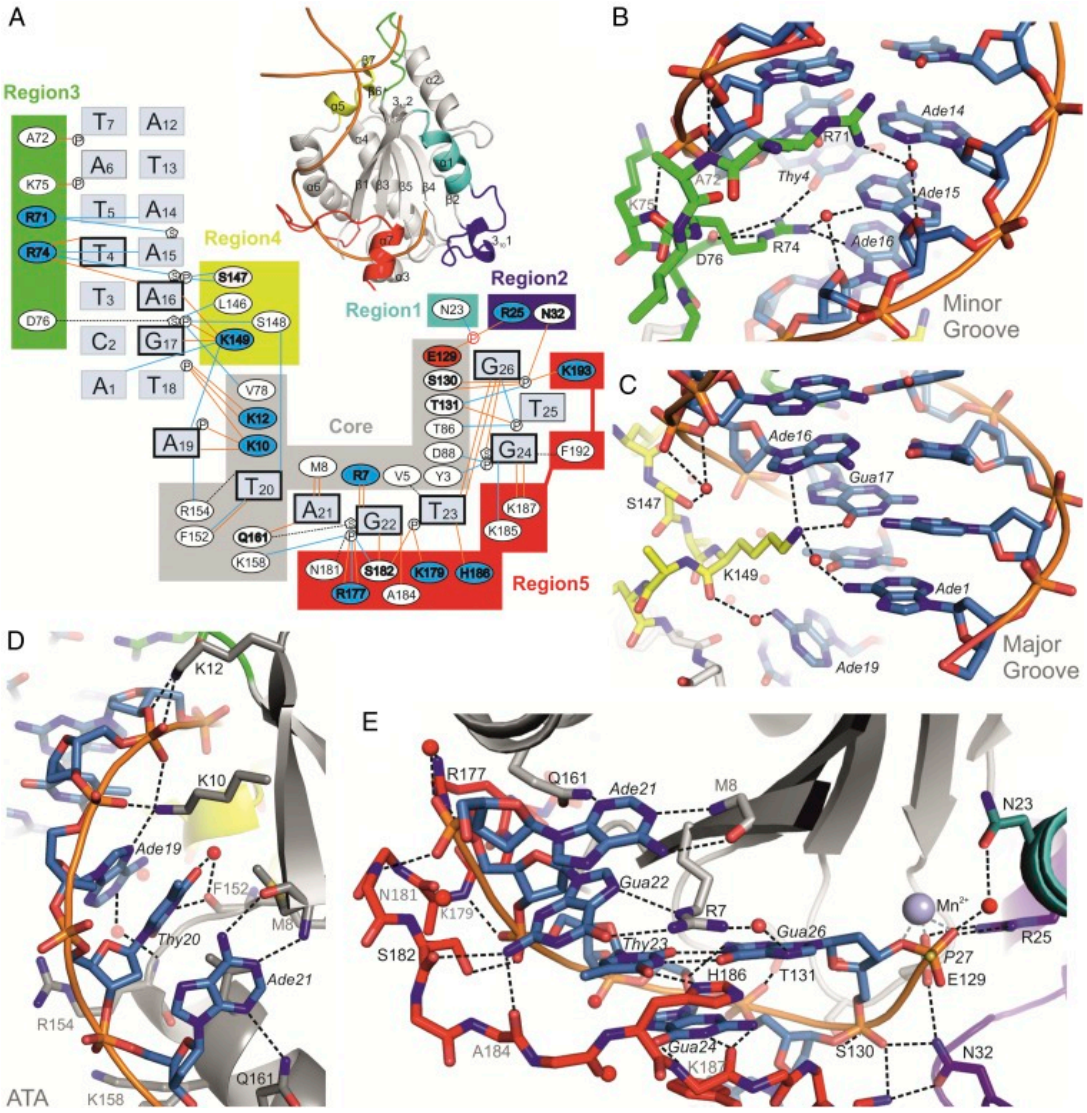

Fig. 2. MobM-DNA interactions. (A) Scheme representing the MobM–DNA interactions in the Nic0+SP complex structure. Brown lines indicate hydrogen bonds, blue lines water-mediated hydrogen bonds, and black dashed lines van der Waals contacts. Residues that form side chain-DNA hydrogen bonds are indicated in bold (positively and negatively charged residues are shown on blue and red background, respectively). Nucleotides that form nucleobase-protein hydrogen bonds and the Thy23-Gua26 wobble bp are indicated in boldface letters. **(B)** Major groove binding by MobM Region 3 (loop α2-β3). **(C)** Minor groove binding by MobM Region4 (helix α5). **(C)** ATA bulge binding by MobM Core Region (strand β1 and helix α6). **(D)** MobM-ssDNA binding by the Core Region, Region5 (loop α6-α7 and helix α7), Region2 (loop α1-3_{10}1) and Region1 (helix α1). Residues that form side-chain – DNA hydrogen bonds are labeled in black. Residues that form main-chain – DNA hydrogen bonds are labeled in grey.

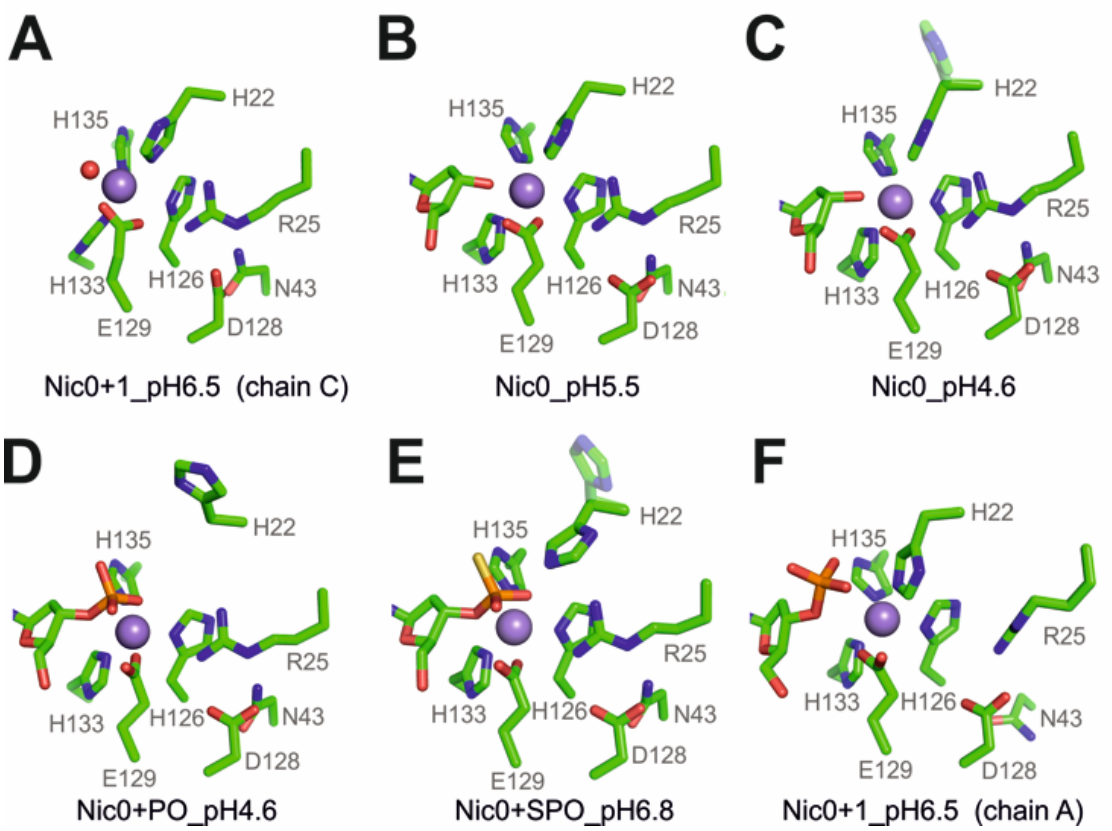

Fig. 3. Active site snapshots (A) Nic0+1/molB – DNA-free active site: H22 ligates Mn^{2+}; **(B)** Nic0+SP – active site with the DNA substrate (up to the scissile phosphate, here as a thio-phosphate): H22 is in the proper position for catalysis, in-line with the scissile bond atoms; R25 and E129 interact with the scissile phosphate. **(C)** Nic0_A – active site with the 5' DNA product/ligation substrate: H22 ligates Mn^{2+}; E129 interacts with the Mn^{2+}-bound DNA O3' of Gua26. **(D)** Nic0+1/molA – active site with the DNA substrate placed in a non-productive manner: the scissile phosphate does not interacts with the Mn^{2+} ion; H22 ligates Mn^{2+}; a number of interactions between E129, D128, R25 (partially disordered) and N43 (disordered) are broken. **(E)** Nic0+P – a low pH active site with the DNA substrate: as in (B), but H22 protonation favors the OUT position of this side chain. **(F)** Nic0_B – a low pH active site with the 5' DNA product/ligation substrate: as in (C), but H22 protonation favors the OUT position of this side chain. H22 minor occupancy rotamers in (C) and (E) are shown faded.

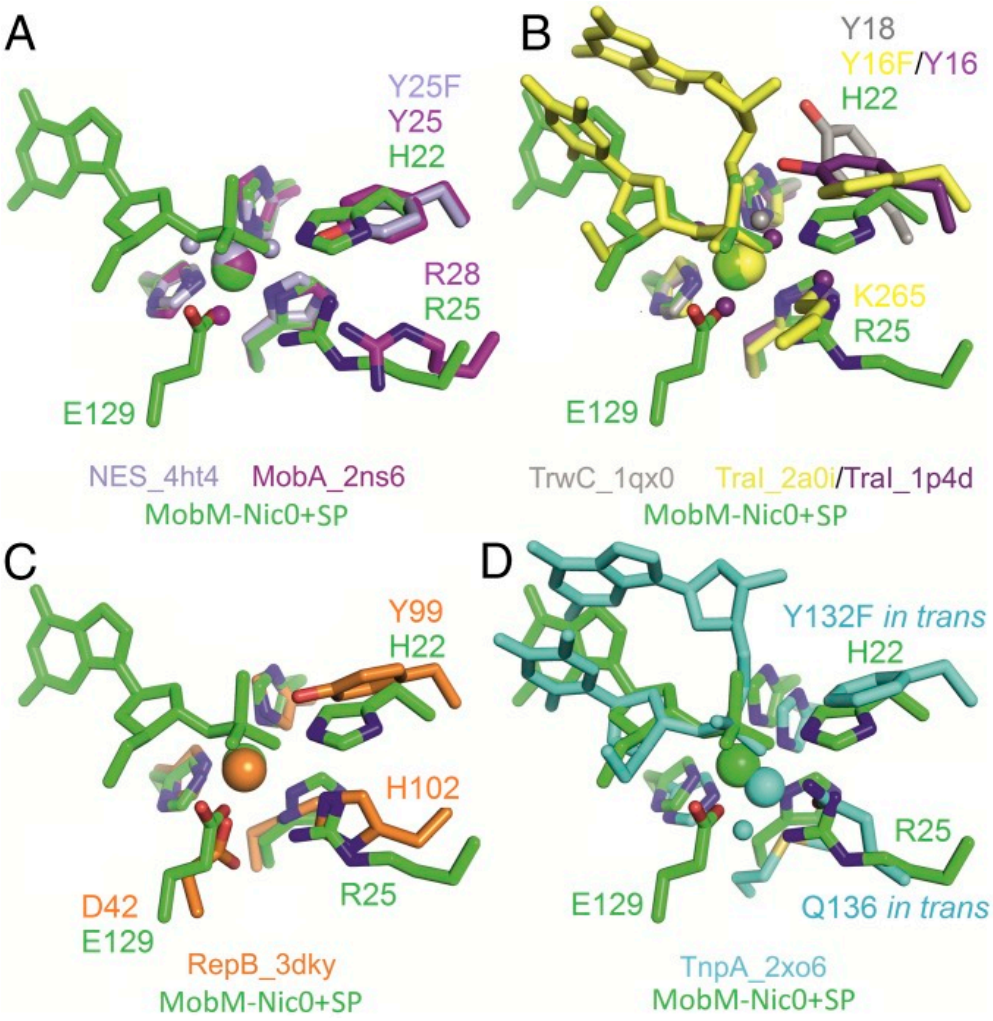

Fig. 4. MobM Nic0+SP structure active site superimposition with other HUH endonucleases. (A) MOBQ family of one-tyrosine (Y1) relaxases MobA_pR1162 and NES_pLW1043. (B) MOBF family of two-tyrosine (Y2) relaxases TraI_pF and TrwC_pR388. (C) Replication initiation protein RepB_pMV158. (D) Transposase TnpA_Dra2 of bacterial insertion sequence ISDra2.

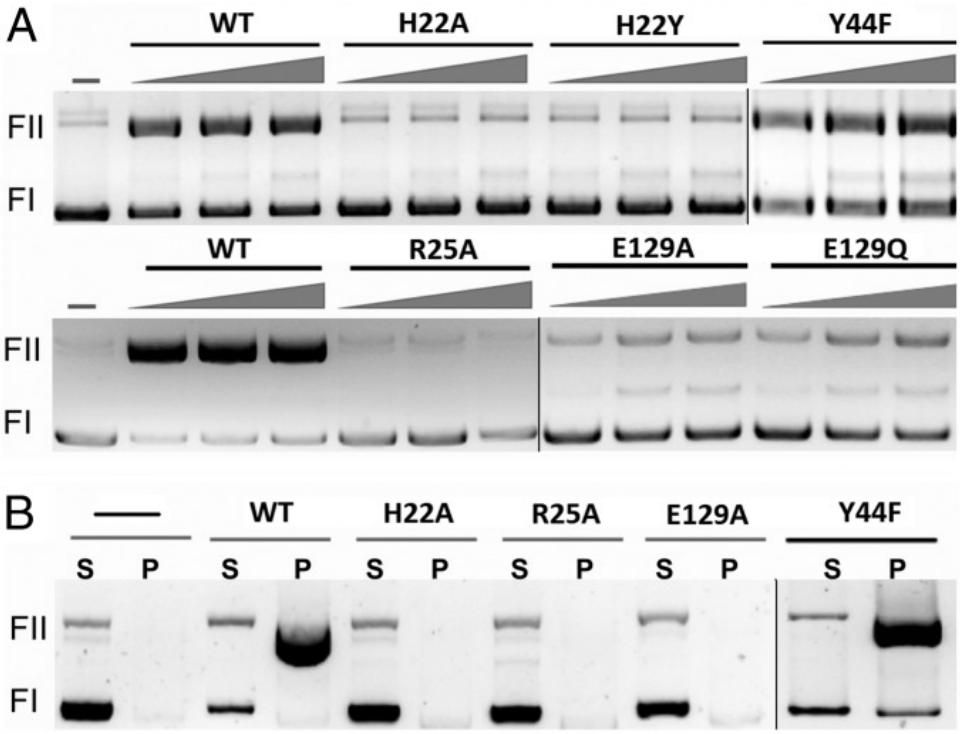

Fig. 5. *In vitro* activity of MobM mutant proteins. (A) MobM nucleolytic activity by plasmid relaxation assay and **(B)** covalent adduct formation and stabilization by protein-DNA pull-down assay. MobM activity is shown as change of plasmid supercoiled (FI) form to open circle (FII) form; S, supernatant (DNA fraction), and P, pellet (protein and protein-DNA fraction). After nicking, the relaxase remains covalently bound to plasmid DNA.

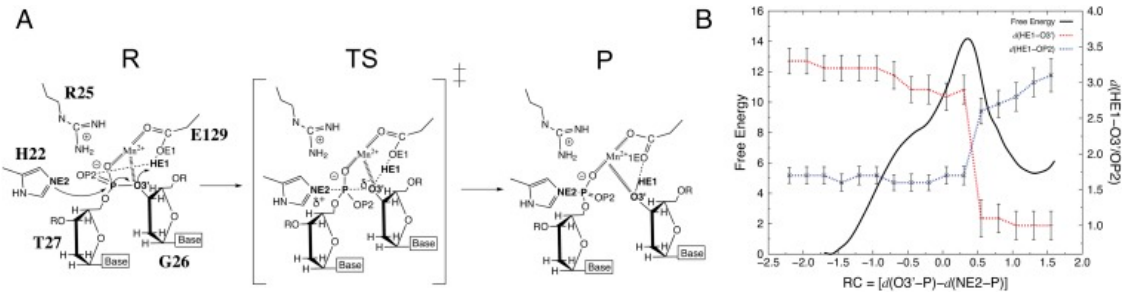

Fig. 6. Theoretical calculations. (A) Proposed mechanism of nucleolytic reaction catalyzed by MobM and **(B)** the corresponding QM/MM free energy surface. The distance relative to the protonation of the O3' leaving group atom by E129 is depicted. R (Reactants), TS (Transition State), P (Products). RC, Reaction Coordinate. Free energy is in kcal/mol and distances, including RC are in Å.