PLOS GENETICS

RESEARCH ARTICLE

ArdC, a ssDNA-binding protein with a metalloprotease domain, overpasses the recipient *hsdRMS* restriction system broadening conjugation host range

Lorena González-Montes, Irene del Campo, M. Pilar Garcillán-Barcia, Fernando de la Cruz*, Gabriel Moncalián*

Departamento de Biología Molecular, Universidad de Cantabria and Instituto de Biomedicina y Biotecnología de Cantabria (IBBTEC), CSIC-Universidad de Cantabria, Santander, Cantabria, Spain

* delacruz@unican.es (FdlC); moncalig@unican.es (GM)

Abstract

Plasmids, when transferred by conjugation in natural environments, must overpass restriction-modification systems of the recipient cell. We demonstrate that protein ArdC, encoded by broad host range plasmid R388, was required for conjugation from *Escherichia coli* to *Pseudomonas putida*. Expression of *ardC* was required in the recipient cells, but not in the donor cells. Besides, *ardC* was not required for conjugation if the *hsdRMS* system was deleted in *P. putida* recipient cells. *ardC* was also required if the *hsdRMS* system was present in *E. coli* recipient cells. Thus, ArdC has antirestriction activity against the HsdRMS system and consequently broadens R388 plasmid host range. The crystal structure of ArdC was solved both in the absence and presence of Mn^{2+}. ArdC is composed of a non-specific ssDNA binding N-terminal domain and a C-terminal metalloprotease domain, although the metalloprotease activity was not needed for the antirestriction function. We also observed by RNA-seq that ArdC-dependent conjugation triggered an SOS response in the *P. putida* recipient cells. Our findings give new insights, and open new questions, into the antirestriction strategies developed by plasmids to counteract bacterial restriction strategies and settle into new hosts.

Author summary

Horizontal gene transfer is the main mechanism by which bacteria acquire and disseminate new traits, such as antibiotic resistance genes, that allow adaptation and evolution. Here we identified a gene, *ardC*, that enables a plasmid to increase its conjugative host range, and thus positively contributes to plasmid fitness. The crystal structure of the antirestriction protein ArdC revealed a fold different from other antirestriction proteins. Our results have wide implications for understanding how a gene enlarges the environments a plasmid can colonize and point to new targets to harness the bacterial DNA uptake control.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Horizontal gene transfer (HGT) is the transmission of genetic material between organisms that are not in a parent–progeny relationship [1]. The clinical relevance of the HGT process lies in the acquisition and dissemination of genes involved in conferring bacterial resistance to antibiotics (Ab^R) between unrelated pathogens. When bacteria face selective pressures, as those exerted by antibiotics, horizontal acquisition of Ab^R allows the diversification of the genomes, increasing survival opportunities [2]. Conjugation is the main HGT process that allows the transfer of genes encoded in autonomous plasmids. This process requires the machinery to build a direct contact between a donor and a recipient cell [1]. Conjugation can be modulated by environmental factors or bacterial strategies based on genetic approaches that are coded in the chromosome (host barriers) or plasmid DNA (plasmid barriers). Plasmid barriers include entry exclusion [3] or fertility inhibition [4], which reduce conjugative transfer. Host barriers can be mediated through SOS response modulation [5,6], CRISPR-Cas systems [7] or restriction and modification (R-M) systems.

R-M systems allow bacteria to discern between self and foreign DNA invading the cell, leading to its destruction. They require two enzymatic activities: a methyltransferase that provides protection to its own DNA and an endonuclease that cleaves the unmethylated invading DNA [8]. There are four main groups of R-M systems. Type I R-M, the most sophisticated R-M system, requires three genes: *hsdR*, *hsdM*, and *hsdS* and their products associate in R$_2$M$_2$S complexes. The S subunit recognizes 13–15 bp sequences, usually asymmetric and bipartite. DNA cleavage is at a location away from the specificity site [9–11]. There is a coevolutionary arms race between bacteria to avoid entrance of foreign DNA molecules and parasitic DNA molecules such as plasmids or bacteriophages to enter a putative host avoiding the restriction by bacterial R-M systems. The antirestriction mechanisms to counteract R-M systems can be divided into four main types based on its mode of action: DNA modification, transient occlusion of restriction sites, sabotage of host R-M activities, and inhibition of restriction enzymes [10].

R388 plasmid is the prototype of the IncW incompatibility group of plasmids. IncW plasmids have a low copy number, a wide range of Ab^R, and a broad host range (BHR) [12]. R388 has 35 genes assorted in functional groups or modules, among them, a gene coding for an antirestriction protein called ArdC [12]. Here, we present ArdC crystal structures and the *ardC* role in interspecies conjugation. We have also identified transcriptional changes associated with *ardC*-mediated conjugation. These results show that ArdC is involved in broadening the R388 plasmid host range.

Results

ardC is required for R388 conjugation from *E. coli* to *P. putida*

R388 plasmid (GenBank Accession Number BR000038.1) is composed of three functional sectors (S1 Fig): one for general maintenance (modules of replication, stable inheritance and establishment) located in the leading region, a sector for Ab^R and integration, and a third one for conjugation (modules of DNA transfer replication and mating pore formation) [12]. We expected the stable inheritance and establishment region to be required in interspecies conjugation. pSU2007, a Kn^R R388 derivative, was transferred with different efficiencies from *E. coli* BW27783-Nx^R to other bacteria (Fig 1A). The transfer of pIC10 (R388Δ*kfrA-orf14)*, an R388 derivative without the stability and maintenance region, was more dissimilar to that of pSU2007 from *E. coli* to *P. putida* KT2440 where the conjugation frequency dropped around 1000-fold. In the stability and maintenance gene region deleted in pIC10 there are 13 genes
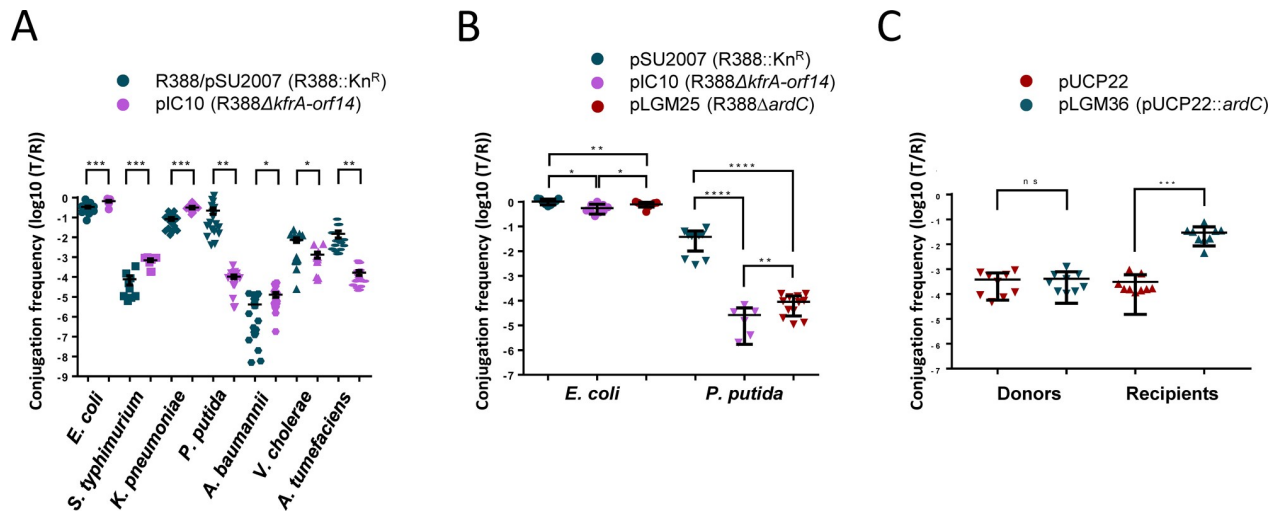
Fig 1. ArdC effect in conjugation. A) Effect of the *kfrA-orf14* region on R388 plasmid conjugative transfer from *E. coli* to different bacteria. Conjugations were performed as described in Materials and Methods at 37 °C except for *P. putida* and *A. tumefaciens* (done at 30 °C) for 1 h except for *A. baumannii* and *V. cholerae* (done for 4 h). R388 was used in conjugations towards *E. coli*, *S. typhimurium* and *K. pneumoniae*. The pSU2007 plasmid was used in conjugations towards the rest of the strains. Donor *E. coli* BW27783-RifR cells were used as donors in mating experiments with *E. coli*, *S. typhimurium*, and *K. pneumoniae*. Donor *E. coli* BW27783-NxR cells were employed in matings with the rest of the strains. Conjugation frequencies are shown as transconjugants per recipient (T/R). Horizontal bars represent the mean ± SD of N = 9–20 (Student's t-test: * p $<$ 0.1, ** p $<$ 0.01, *** p $<$ 0.001, **** p $<$ 0.0001). B) Effect of *ardC* and *kfrA-orf14* deletions on plasmid conjugative transfer (1 h at 37°C) from *E. coli* BW27783-Nx^{R} to *E. coli* BW27783-Rif^{R} or *P. putida* KT2440 (N = 6–12). C) Effect in the conjugation frequency of pLGM25 when expressing *ardC* in donors or recipients. The effect of the presence of plasmid pUCP22 or pUCP22::*ardC* in donors or recipients is shown. Conjugation was done for 1 h at 37°C with 0.1 mM IPTG in the mating mixture (N = 9).

that code for proteins homologous to some with predicted function of fertility inhibition (*osa*) [13], proteins of unknown function (*klcB*, *nuc1*, *nuc2*, *orf7*, *orf8*, *orf9*, *orf12*, and *orf14*), transcriptional regulators (*kfrA*, and *ardK*), ssDNA binding protein (*ssb*), and antirestriction (*ardC*). ArdC protein (297 amino acids and 33.2 KDa, GenBank Acc. No. FAA00054.1) exhibited an *in vitro* antirestriction function towards Type I and II R-M systems [14]. Thus, we constructed plasmid pLGM25 (R388Δ*ardC*) to check if the effect observed in conjugation with pIC10 could be due to the lack of the *ardC* gene. This plasmid was introduced into *E. coli* BW27783-Nx^{R} and then conjugated to *E. coli* BW27783-Rif^{R} or *P. putida* KT2440 (Fig 1B). We observed that the absence of *ardC* in the conjugative plasmid pLGM25 reduced the conjugation frequency to *P. putida* from 3.8E-02 to 9.0E-05, but not to *E. coli*. Thus, the results observed for pIC10 could be explained to a large extent by *ardC* absence.

ardC is needed in recipient cells

To check if ArdC was needed in donor or recipient cells, transfer of Δ*ardC* pLGM25 plasmid was measured when complemented by the overexpression of *ardC* in donor *E. coli* cells, or in recipient *P. putida* cells. As shown in Fig 1C, *ardC* did not improve the conjugation frequency when overexpressed in donors. On the other hand, overexpression of *ardC* in recipient cells increased the conjugation frequencies, reaching pSU2007 conjugation levels. Thus, it seems that the expression of ArdC is specifically required in the recipient, and not in donor cells.

ArdC is a ssDNA-binding protein with a metalloprotease domain

R388 ArdC crystal structure was solved at 2.6 Å resolution using a selenomethionine-derivative protein structure solved by single anomalous dispersion as described in Materials and Methods. Using this preliminary structure, the apo ArdC structure was solved at 2.0 Å resolution by

Table 1. Data collection and refinement statistics for ArdC structures [a].

	ArdC SeMet	ArdC native	ArdC-Mn
Wavelength	(peak)	0.9792	0.9793
Resolution range	47–2.6 (2.69–2.6)	39.5–2.0 (2.07–2.0)	54.8–2.7 (2.8–2.7)
Space group	R 3: H	R 3: H	P 32
Unit cell	a = b = 136.9 c = 51.3 $\alpha = \beta = 90$ $\gamma = 120$	a = b = 136.8 c = 51.7 $\alpha = \beta = 90$ $\gamma = 120$	a = b = 116.5 c = 162.1 $\alpha = \beta = 90$ $\gamma = 120$
Total reflections	443798 (44597)	537291 (33696)	1158240 (114678)
Unique reflections	10970 (1098)	24366 (2408)	66345 (6492)
Multiplicity	40.5 (40.6)	22.1 (13.9)	17.5 (17.4)
Completeness (%)	99.82 (99.64)	99.8 (98.5)	96.6 (96.9)
Mean I/sigma(I)	66.78 (6.54)	37.1 (3.1)	25.7 (2.6)
Wilson B-factor	46.69	31.87	47.81
R-merge	0.5962 (1.055)	0.7636 (1.274)	0.6992 (1.47)
R-meas	0.6046 (1.07)	0.7814 (1.326)	0.7198 (1.514)
CC1/2	0.937 (0.901)	0.773 (0.474)	0.691 (0.572)
CC[*]	0.984 (0.974)	0.934 (0.802)	0.904 (0.853)
Reflections used in refinement		24339 (2406)	65278 (6491)
Reflections used for R-free		1230 (111)	3008 (320)
R-work		0.1726 (0.2154)	0.2207 (0.2823)
R-free		0.1976 (0.2410)	0.2943 (0.3499)
CC(work)		0.864 (0.814)	0.855 (0.672)
CC(free)		0.817 (0.837)	0.845 (0.581)
Number of non-hydrogen atoms		2412	16374
Protein residues		276	2008
RMS(bonds)		0.008	0.010
RMS(angles)		1.16	1.17
Ramachandran favored (%)		97.76	92.79
Ramachandran allowed (%)		1.49	5.71
Ramachandran outliers (%)		0.75	1.5
Rotamer outliers (%)		0.00	0.00
Clashscore		3.46	16.63
Average B-factor		36.39	46.75

[a] Statistics for the highest-resolution shell are shown in parentheses.

https://doi.org/10.1371/journal.pgen.1008750.t001

molecular replacement (MR). Apo ArdC crystallized in the H3 space group containing one molecule per asymmetric unit. Data collection and refinement statistics are given in Table 1.

ArdC is composed of two structural domains: An N-terminal domain (residues 1–134) and a C-terminal domain (residues 151–297) joined by a long and flexible loop (135–150) (Fig 2A and 2B). Electron density was not observed for the N-terminal residues 1–6, the flexible small loop residues 33–39, residues 136–141 in the region connecting both domains and C-terminal residues 294–297. The N-terminal domain is composed of three α-helices (α1-α3), a three-stranded β-sheet (β1, β3, and β4) that supports a long and protuberant β-hairpin (β3-β4), a smaller two-stranded antiparallel β-sheet formed by β2 and β5, as well as three 3_{10} helices labeled from η1 to η3 (Fig 2A and 2B). The C-terminal domain is composed of six α-helices (α4-α9) and short three-stranded antiparallel β-sheets (β6-β8) as shown in Fig 2A and 2B.

The ArdC structure was compared to those deposited in the PDB using the Dali server [16]. ArdC N-terminal domain closest structural homolog is present in a nucleotide excision repair
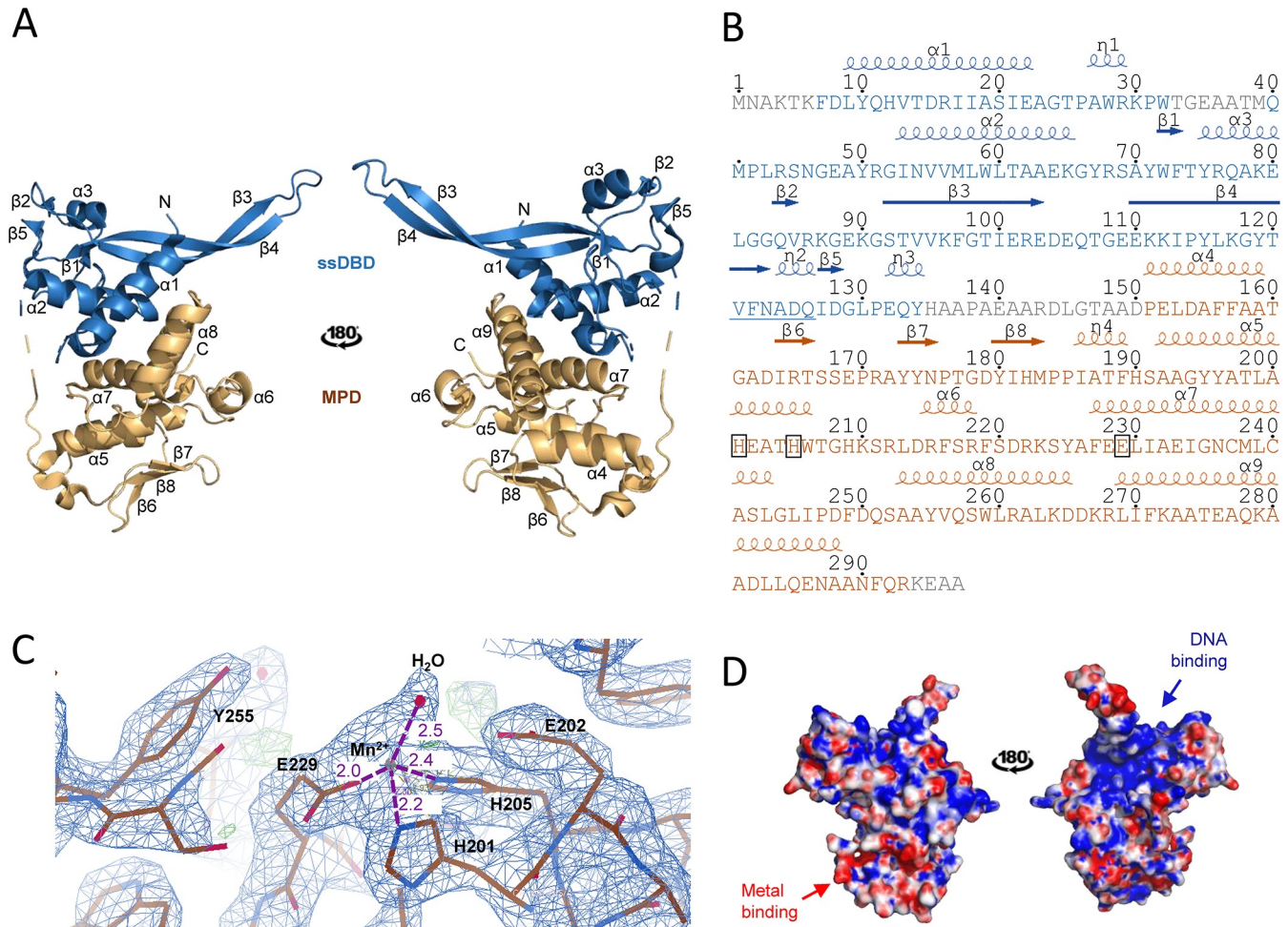
Fig 2. ArdC structure. A) Cartoon representation of two views of the ArdC structure. N-terminal ssDNA-binding domain (ssDBD) is shown in blue and C-terminal metalloprotease domain (MPD) in orange. α-helices are labeled from α1 to α9 and β-strands are labeled from β1 to β8. A dashed line schematizes the disordered loop joining both domains. B) ArdC sequence with secondary structure information. ArdC sequence is colored by domains and α-helices and β-strands are labeled as in A). 3_{10} helices are labeled from η1 to η3. The residues involved in metal coordination are framed. The "squiggle" signature proposed by [15] for Rad4 is underlined in blue. C) Electron density of the metal-binding site in the ArdC-Mn crystal structure solved at 2.7 Å resolution. Residues and molecules involved in metal coordination (H201, H205, E229, and H_2O) or activity are labeled. Distance in Å to the metal is shown in purple. D) Electrostatic potential surface. The negative surface is colored in red and the positive surface in blue (calculated by APBS tool). The expected binding areas for DNA and metal cofactor are indicated.

https://doi.org/10.1371/journal.pgen.1008750.g002

protein called Rad4 (PDB: 2QSG; Z-score: 4.5), a component of the eukaryotic nucleotide excision repair (NER) pathway. Rad4 is composed of an inactive transglutaminase fold domain and three different β-hairpin domains (BHD). The three tandem BHD domains form a large DNA binding surface [17]. ArdC N-terminal domain is more similar at a sequence level to the second BHD domain of Rad4 (BHD2), although BHD2 is considerable smaller (about 50 amino acids long compared to the 134 residues of ArdC-N). It lacks some ArdC-N structural features, such as the starter ArdC α1, α2, and final 3_{10} motifs. Moreover, the protuberant β-hairpin formed by β3 and β4 is larger in ArdC (Fig 3A). ArdC N-terminal domain possesses the $V^{121}FNADQ^{126}$ sequence located within a 3_{10} helix (η2) between β4 and β5 (Fig 2B). This region forms a crossover with the β2 to β3 region and creates a sharp twist of the chain known as the "squiggle" motif [15]. This motif in Rad4 is proposed to be responsible for a highly flexible region that could facilitate recognition of DNA sequences.
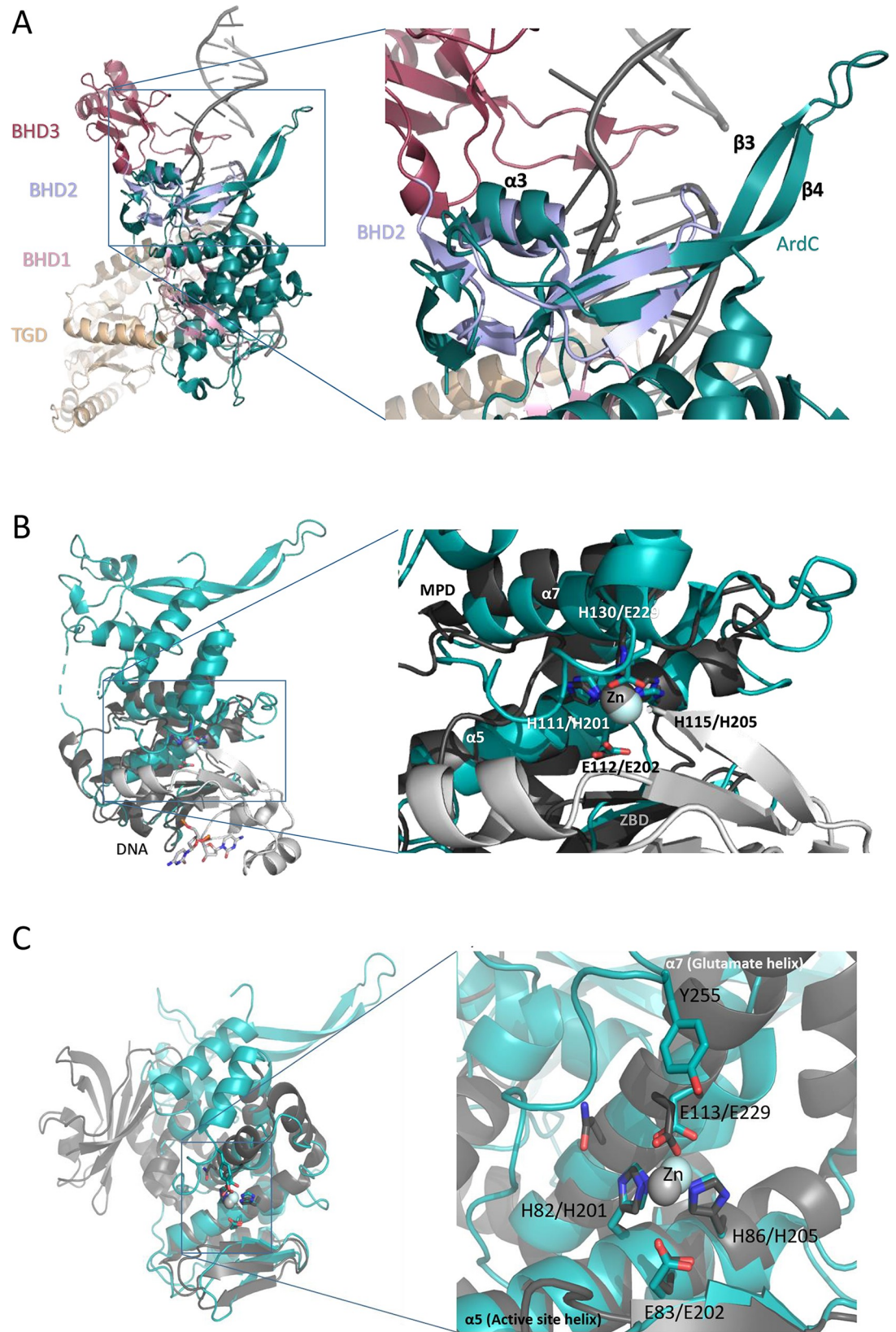
Fig 3. ArdC structural homologs. A) Superposition of ArdC (green) with the Rad4 protein bound to UV-damaged DNA (2QSG). The Rad4 TGD domain is shown in beige, BHD1 in pink, BHD2 in purple, and BHD3 in red. For clarity, Rad23 (present in 2QSG structure) has been removed. Detailed view of ArdC ssDBD domain superposed to the Rad4 BHD2 domain. B) Superposition of ArdC (green) with the Spartan SprT domain (6MDX; grey). The SprT Zn2+-binding sub-domain (ZBD) is shown in light grey and the metalloprotease sub-domain (MPD) is shown in dark grey. Detailed view of the metalloprotease active center with the residues involved in catalysis in sticks numbered as (MPD/ArdC). C) Superposition of the ArdC structure (green) with the IrrE-Zn protein from *Deinococcus radiodurans* (3DTI, grey). Detailed view of the active center with the residues involved in catalysis in sticks numbered as (IrrE/ArdC).

The surface electrostatic map (Fig 2D) reveals a positively charged groove in the region of the N-terminal domain adjacent to the C-terminal domain, suggesting a DNA binding site between both structural domains. By electrophoretic mobility shift assays (EMSA), we determined that ArdC preferentially binds ssDNA oligonucleotides over dsDNA molecules (Fig 4) in accordance with previous results [14]. Moreover, binding to partial dsDNA with 5' or 3' terminal ssDNA overhangs is preferred over binding to perfectly paired complementary dsDNA duplex (Fig 4). We will name ArdC N-terminal domain hereafter ssDNA-binding domain (ssDBD).

Regarding the C-terminal domain (amino acids 151–297), according to the Dali server [16] the closest structural homologs were the DNA binding metalloproteases Spartan (PDB: 6MDW; Z-score: 6.4) and IrrE (PDB: 3DTE; Z-score: 5.5) (Fig 3C and 3D). Spartan is a protein involved in the cleavage of proteins irreversibly cross-linked to DNA to preserve this way genome stability [18]. IrrE protects *D. radiodurans* from UV radiation DNA damage by proteolysis of a negative transcriptional regulator, DdrO, which represses the expression of DNA damage response genes involved in SOS response [19]. These proteins belong to the gluzinzin metalloprotease family characterized by the presence of the conserved residues HExxH located on the "active site helix" (α5 in ArdC) and an additional conserved motif (E,H)xx(A,F,T,S,G) located in the contiguous α-helix or "glutamate helix" (α7 in ArdC) [20] (Fig 2A and 2B). The surface electrostatic map revealed a negatively charged catalytic pocket (Fig 2D). We will name ArdC C-terminal metalloprotease domain MPD.
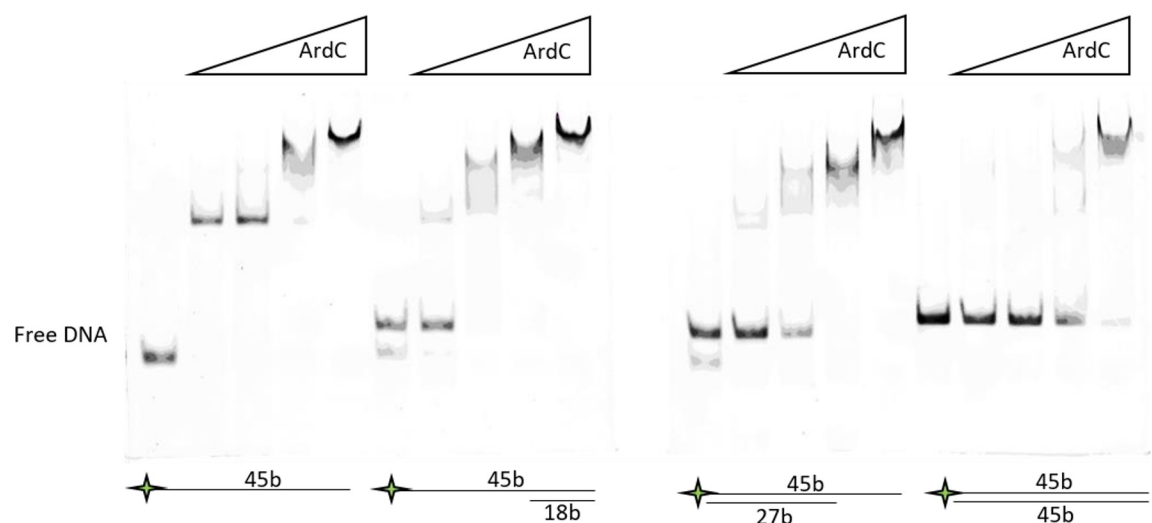
Fig 4. ArdC DNA-binding preferences assessed by EMSA. ArdC binding of a 6FAM-labeled 45 bases ssDNA oligonucleotide (Fluor-T87I2), a perfectly paired complementary 45bp dsDNA duplex (Fluor-T87I2 + T87I1), and two partial dsDNA with 5' or 3' terminal ssDNA overhangs (Fluor-T87I2 + Mid1) and (Fluor-T87I2 + Mid2) was performed at increasing concentrations of ArdC (0, 125 nM, 250 nM, 500 nM, and 1 μM), as described in Materials and Methods. Protein-DNA complexes were resolved by native 10% polyacrylamide gels and visualized by a fluorescent image analyzer.

The two histidines of the active site helix and the glutamic acid of the glutamate helix have been shown to coordinate a catalytic divalent metal ion, usually zinc. However, gluzinzin metalloproteases maintain the catalytic activity with Co^{2+}, Mn^{2+} or Ni^{2+} too due to the flexibility of these three metal coordination geometries [20,21]. To assess the metal used by ArdC the stability of ArdC was assayed in the presence of different metal cofactors. ArdC showed increased thermal stability (assayed by ThermoFluor) in the presence of Ni^{2+}, Mn^{2+} or Co^{2+} ($\Delta T_M > 4°C$), but not in the presence of Zn^{2+}, Ca^{2+}, Mg^{2+}, Cu^{2+} or Fe^{3+} (S1 Table). Moreover, to know the conformation of the active site when bound to metals, ArdC was crystallized in the presence of $MnCl_2$ (Materials and Methods). ArdC-Mn crystallized in the P32 space group containing eight molecules per asymmetric unit and the structure was solved at 2.7 Å (Table 1). Mn^{2+} is tetrahedrally coordinated by H201, H205, E229 and an H_2O molecule (Fig 2C). H205 is oriented towards the metal by interaction with the conserved E228 through the non-coordinating nitrogen atom. The E202 residue of the $HE^{202}xxH$ motif orients and acts as a catalytic base for the activation of a water molecule that coordinates the metal. The H_2O molecule could act as a Lewis acid to allow the nucleophilic attack [20]. By analogy with other gluzinzin metalloproteases, the conserved ArdC residue Y255 could stabilize by a hydrogen bond the polypeptide chain to be cleaved [22]. The metalloprotease sub-domain (MPsD) in Spartan shares the active center structure with ArdC MPD except that MPsD uses a third histidine instead of a glutamic acid for metal coordination (Fig 3B).

It had been proposed that ArdC could avoid ssDNA degradation by HhaI, a type II restriction enzyme able to cleave both ssDNA and dsDNA [14]. According to our structural results, this ArdC DNA protection could be due to ArdC MPD activity targeting the restriction enzyme. To test this hypothesis, the inhibition of HhaI by ArdC was assayed in the presence of ssDNA M13mp18 (7.2 kb) and Mg^{2+}. As observed in S2 Fig, ArdC was able to avoid ssDNA cleavage by HhaI but we did not observe HhaI degradation by ArdC.

Since its 3D structure defined ArdC as a protease, we tried to find a specific protein target. ArdC mutant E229A (supposed to be inactive) was purified and used as prey for co-purification of potential targets in *P. putida* KT2440 cell lysate by the pull-down technique. The only protein that co-eluted with ArdC was PP_0941, a protein of unknown function similar to the 50S ribosome subunit associated protein YjgA (S3 Fig).

SOS response is activated in *P. putida* recipient cells by the transfer of an *ardC*-containing plasmid

IrrE, the bacterial closest structural homolog to ArdC, triggers SOS response by cleaving the transcriptional regulator DdrO in an analogous way to the RecA-LexA system [19]. To check if ArdC could have similar activity on plasmid conjugation, we analyzed by RNA-seq changes in gene expression when an *ardC*-containing plasmid was transferred from *E. coli* to *P. putida*. As described in Materials and Methods, *P. putida* KT2440 was mixed in a conjugation filter with either *E. coli* BW27783-Nx^R bearing no plasmid (NP), *E. coli* BW27783-Nx^R bearing pSU2007 (*ardC*$^+$) or *E. coli* BW27783-Nx^R bearing pLGM25 (*ardC*$^-$).

As expected according to the results shown in Fig 1B, significant conjugation frequency differences between *ardC*$^+$ and *ardC*$^-$ conditions were observed (S2 Table). RNA-seq results (S3–S5 Tables and S4 Fig) showed that: (a) R388 genes involved in conjugation are highly upregulated in the *ardC*$^+$ condition regarding the *ardC*$^-$ condition (S5 and S6 Tables and S4A Fig). This is consistent with the zygotic induction observed in the recipient cells after conjugation [23]; (b) several donor *E. coli* genes and pathways involved in flagellar motility, SOS and stress responses and different metabolic pathways are downregulated in the *ardC*$^-$ condition regarding the NP or *ardC*$^+$ conditions (S5 and S7 Tables, and S4B Fig); (c) SOS genes are upregulated

in recipient *P. putida* cells when receiving *ardC*[+] containing plasmid with respect to NP or *ardC*[−] conditions. Differential expression of *P. putida* genes in recipient cells when the ardC-containing plasmid is transferred is shown in S8 Table and S4C Fig.

ArdC metalloprotease activity is not required for conjugation to *P. putida*

Being ArdC a ssDNA binding protein with a metalloprotease domain, we checked if this proteolytic activity was needed for ArdC activity in conjugation. The *ardC* gene was mutated to *ardC_E229A* in pSU2007 to generate plasmid pLGM33. Mutation of the glutamic acid of the active site involved in metal coordination to alanine, E229A, is expected to deactivate the proteolytic center of ArdC, as it occurs in other family members. Plasmids pSU2007 or pLGM33 were conjugated from *E. coli* to *P. putida*. Surprisingly, pLGM33 conjugation frequency resulted to be like that of pSU2007, around 0.01 transconjugants per recipient (T/R) (Fig 5A). Besides, the *ardC_E229A* mutant gene was tested for its ability to complement pLGM25 (R388 *ΔardC*) plasmid in mating experiments from *E. coli* to *P. putida* KT2440. The *in trans* expression of *ardC_E229A* in *P. putida* KT2440 recipient cells was able to increase the conjugation frequency of pLGM25 at the same levels as the expression of the wt *ardC* gene (Fig 5B). Thus, ArdC metalloprotease activity is not required for the host range broadening activity at least in *P. putida*.

ArdC counteracts the HsdRMS system in both *P. putida* and *E. coli*

To identify the functional target of ArdC, different *P. putida* mutant strains were assessed as recipients in mating experiments. RecA dependent SOS response is activated in *ardC*[+] conjugation recipient cells as shown by the RNA-seq experiments (S8 Table). To check the role of this response in conjugation, firstly pSU2007 and pLGM25 were conjugated from *E. coli* to *P. putida* KT2440*ΔrecA*. The frequency of conjugation of pSU2007 to *P. putida* KT2440*ΔrecA*
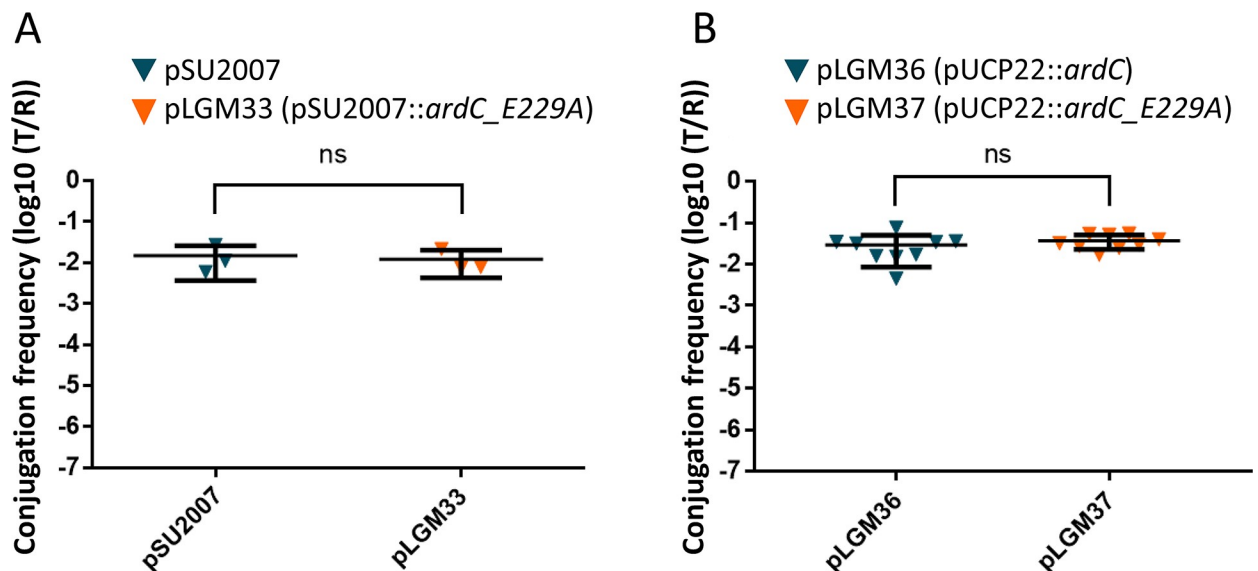
Fig 5. Effect of ArdC E229A mutant on plasmid conjugative transfer. A) Conjugation of *E. coli* BW27783 bearing pSU2007 or pLGM33 (pSU2007_ardC_E229A) to *P. putida* KT2440. Conjugation was performed for 1 h at 37˚C. Horizontal bars represent the mean ± SD of N = 3 observations. B) Effect in the conjugation frequency of pLGM25 when expressing *ardC_E229A* in recipient cells. Conjugation of pLGM25 in *E. coli* BW27783 donor cells to *P. putida* KT2440 recipient cells bearing pUCP22::*ardC* or pUCP22:: *ardC_E229A*. Conjugation was performed for 1 h at 37˚C with 0.1 mM IPTG added to the mating mixture (N = 9).

https://doi.org/10.1371/journal.pgen.1008750.g005

was about 10^{-2} T/R (Fig 6A). Thus, RecA dependent SOS response in *P. putida* recipient cells is not essential for conjugation. Moreover, the frequency of conjugation of pLGM25 to *P. putida* KT2440*ΔrecA* was around 10^{-4} T/R, meaning that the absence of *recA* does not enhance conjugation in the absence of *ardC*. In addition, pSU2007 and pLGM25 were conjugated from *E. coli* to *P. putida* EM42, which carries deletions of several prophages and other accessory genes (*Δprophage1 Δprophage4 Δprophage3 Δprophage2 Δtn7 ΔendA-1 ΔendA-2 ΔhsdRMS Δflagellum Δtn4652*) that could harass the heterologous gene expression (because their association to genetic instability or attributed to the unfruitful usage of metabolic resources). Any of these genes removed from *P. putida* EM42 could affect the establishment of the plasmids acquired by conjugation. The conjugation frequency towards EM42 strain was not affected by the *ardC* deletion (Fig 6A) indicating that ArdC could be counteracting the action of the products of one or more of the deleted genes in the EM42 strain. To identify the gene(s) responsible for the observed phenotype, pSU2007 and pLGM25 were conjugated from *E. coli* to *P. putida* mutants with deletions in single or several genes. pLGM25 only reached pSU2007 conjugation levels (around 0.1 T/R) in the *P. putida* KT2449*ΔhsdRMS* strain EM422 (Fig 6A). Similar results were observed at 30˚C (S5 Fig). The *hsdRMS* operon was thus the main responsible for the effect observed in EM42. pLGM25 was efficiently transferred between *P. putida* KT2440, being its conjugation frequency around 0.1 T/R (Fig 6B). However, pLGM25 transfer drastically dropped from *P. putida* EM422 (*ΔhsdRMS*) to *P. putida* KT2440 unless pLGM36 was present in the *P. putida* KT2440 recipient cell. Thus, *ardC* is counteracting the effect of the HsdRMS R-M system in the incoming DNA.

The *E. coli* strain BW27783 used as donor and recipient in the mating experiments shown in Fig 1B was a non-restricting and modifying strain ($r_K^- m_K^+$), and thus not suitable to check for the ArdC effect in R388 conjugation between *E. coli* cells. So, to evaluate if the *E. coli* HsdRMS R-M system was also targeted by ArdC, a non-restricting and non-modifying ($r_B^- m_B^-$) *E. coli* strain, BL21(DE3), was used as a donor. When pLGM25 was transferred to *E. coli* MG1655 ($r_K^+ m_K^+$) its conjugation frequency significantly decreased in comparison to pSU2007 transfer, but not to *E. coli* BW27783 ($r_K^- m_K^+$) (Fig 6C). Moreover, the conjugation frequency was again rescued when ArdC was expressed in the recipient cell.

Discussion

Antibiotic resistance determinants and xenobiotic degradation genes are extensively disseminated in different ecological niches by conjugative BHR plasmids. BHR plasmids evolved different strategies to avoid obstacles to their entrance in new recipient cells. In this article, we determined that ArdC protein, produced by the IncW BHR plasmid R388, is required for interspecies conjugation from *E. coli* to *P. putida*. ArdC was first studied by [14], who showed an *in vitro* antirestriction function towards Type I and II R-M systems. They observed that ArdC showed a 38% identity with the N-terminal region (about 300 amino acids, DUF1738 domain (pfam08401)) of TraC1 primase from RP4 plasmid. Since TraC1 travels to the recipient cell during conjugation presumably bound to the ssDNA that is being transferred (T-strand) [14], they proposed that ArdC could be also transferred during conjugation bound to the plasmid T-strand. Besides, they proposed that ArdC protects the incoming DNA from host endonucleases through restriction site occlusion. However, they failed to detect any significant influence of ArdC on the efficiency of an IncW plasmid transfer between a non-restricting and non-modifying donor *E. coli* and an EcoK-restricting recipient *E. coli* strain. We demonstrated in this work that ArdC plays indeed an *in vivo* function in IncW plasmid conjugation. ArdC role involves its expression in recipient cells. Consequently, complementation of pLGM25 (R388*ΔardC*) with *ardC* in donor cells neither recovered wt plasmid conjugation frequencies
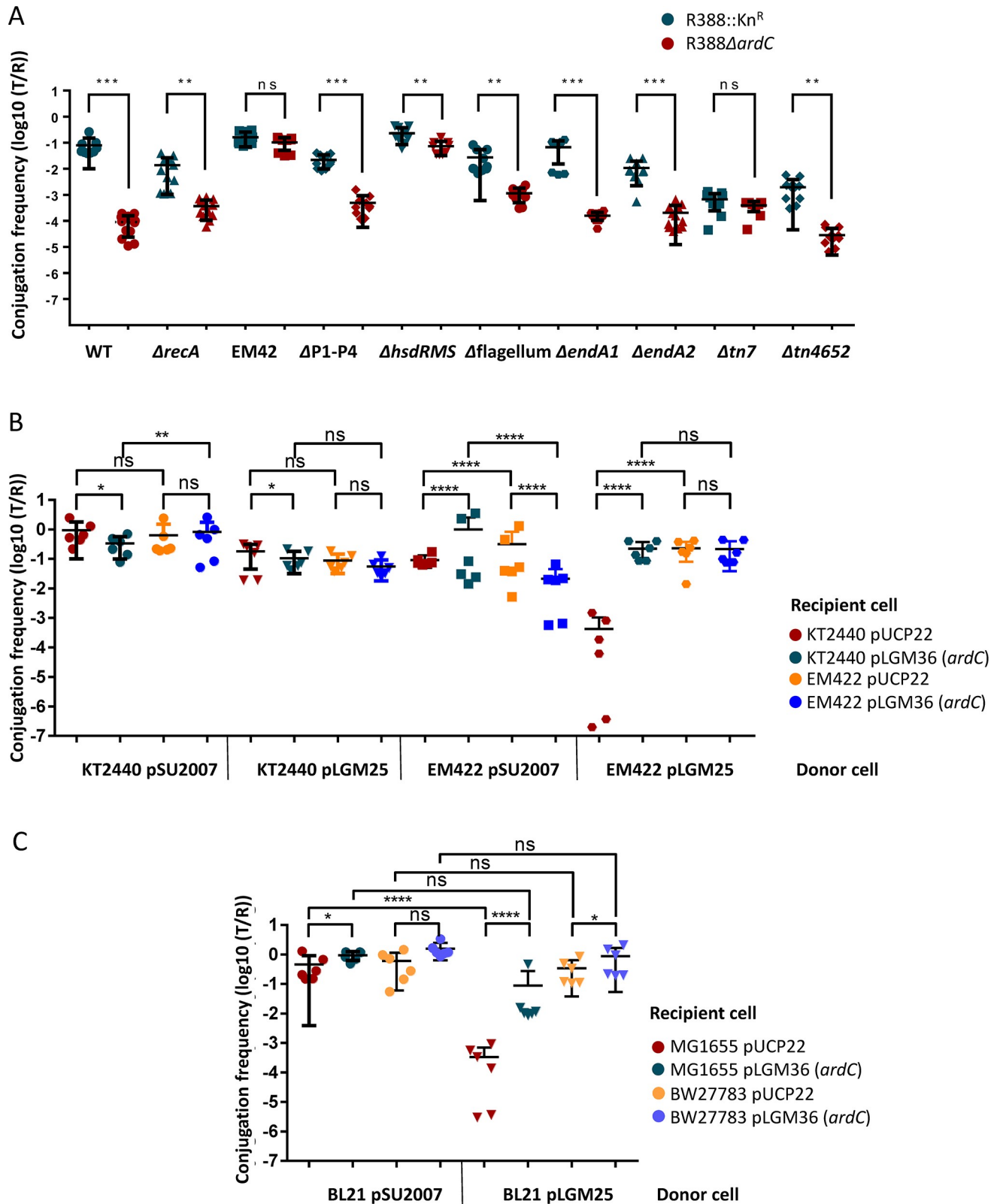
Fig 6. Effect of HsdRMS systems on R388 plasmid conjugation. A) Effect of ArdC on plasmid conjugative transfer from *E. coli* to *P. putida* KT2440 mutants. The conjugation frequencies (T/R) to *P. putida* KT2440 wt strain or different mutants were obtained after conjugation for 1h at 37°C. The deleted gene(s) in each strain is shown. EM42 is Δprophage1, Δprophage4, Δprophage3, Δprophage2, Δ*tn7*, Δ*endA-1*, Δ*endA-2*, Δ*hsdRMS*, Δflagellum, and

from *E. coli* to *P. putida* nor between *E. coli* strains, while ArdC expression in the recipient cells did. Thus, although our experiments do not rule out ArdC export during conjugation in the wild type system, it is not a requirement for its activity, contrary to what was proposed [14]. TraC1 of RP4 and Sog primases of IncI1 plasmids are transferred to the recipient cell during conjugation[24]. They have in common a topoisomerase-primase (TOPRIM) domain located either in the C-terminal part of TraC1 or the N-terminal part of Sog. Both primases differ in that TraC1 contains the DUF1738 domain, which is also present in ArdC. TOPRIM is not present in ArdC. The antirestriction activity of the TraC1 primase has not been investigated but, in the light of our results is a reasonable hypothesis. In any case, a putative TraC1 antirestriction activity seems redundant, since antirestriction against the Type I R-M system was demonstrated both *in vivo* [25] and *in vitro* [26] for KlcA, another RP4 protein.

DUF1738 and TOPRIM domain accretion in a single protein seems to be a specific adaptation. Inspection of the RefSeq plasmid database (version 90) using Hidden Markov Model searches for TOPRIM (PF13362) and DUF1738 (PF08401) domains revealed the presence of 1,334 TOPRIM-containing proteins and 877 DUF1738-containing proteins. Only 84 plasmid proteins contained both domains and they were encoded in 83 plasmids. Additionally, 203 plasmids encoded both domains, but each in a separate protein thus totalling 286 plasmids encoding both DUF1738 and TOPRIM domains, either fused or independent (S6 Fig). An overwhelming proportion of plasmids encoding either DUF1738 or TOPRIM domains also encoded a MOB relaxase and can thus be considered transmissible by conjugation [27]: 939 out of 1242 TOPRIM-encoding plasmids, 667 out of 831 DUF1738-encoding plasmids, 252 out of 286 DUF1738+TOPRIM-encoding plasmids, and 82 out of 83 plasmids encoding both domains in the same protein. While transmissible plasmids encoding TOPRIM and DUF1738 domains in separate proteins were distributed in seven different MOB classes, all but one of those encoding both domains in the same polypeptide belonged to the MOB_P class (roughly half of them IncP1 plasmids).

Further information about ArdC activity was obtained by solving its crystal structure. Interestingly, ArdC contains an N-terminal ssDNA binding domain and a C-terminal metalloprotease domain. ArdC is structurally similar to DNA-binding dependent metalloproteases involved in the maintenance of genetic stability such as Spartan or IrrE. Human Spartan protein cleaves DNA-protein crosslinks [18], while IrrE plays a central regulatory role in DNA protection and repair pathways in response to radiation [19]. ArdC structure differs from other known plasmid-encoded antirestriction proteins. ArdA from the conjugative transposon Tn*916* (2W82) is structurally similar to the B-form DNA, this way binding Type I R-M systems to avoid DNA degradation [28]. KlcA (PDB:2KMG) from the IncP-1β plasmid pBP136 and ArdB (PDB:2WJ9) are composed of a single α/β domain inhibiting the endonuclease activity of Type I R-M systems by an indirect mechanism not related to the mimic of DNA structure [29]. According to our solved structure, we expect ArdC to provide a new antirestriction mechanism.

We observed SOS response activation in *P. putida* recipient cells during *ardC*⁺-mediated conjugation from *E. coli*. ArdC could trigger SOS response in recipient cells similarly than IrrE triggers SOS response upon radiation damage (by proteolysis of DdrO, a transcriptional regulator involved in SOS response) [30]. However, it has also been described that conjugative ssDNA activates SOS response when the plasmid cannot replicate in the recipient [31]. As

ArdC_E229A mutant is still able to promote R388 conjugation to *P. putida*, we think that activation of the SOS response in the recipient is the consequence of the ongoing conjugative process and not a direct effect of the presumed proteolytic activity of ArdC. We hypothesize that the high affinity of ArdC by ssDNA could delay the complementary strand replication resulting in a transient higher amount of ssDNA and thus triggering the SOS response. This is supported by the fact that pLGM25 transfer to *P. putida* did not result in SOS activation.

Since ArdC is not required for R388 conjugation either to *P. putida* KT2440Δ*hsdRMS* or to *E. coli* Δ*hsdR*, it is expected to play a role as counteracting Type I R-M system, probably preventing degradation of the transferred DNA. Moreover, ArdC is not required for R388 conjugation between *P. putida* KT2440 strains. Consequently, once R388 is modified by the HsdRMS system, it can be properly transferred to another cell also containing the same HsdRMS system. Type I R-M systems attack dsDNA and thus are not expected to degrade ssDNA during bacterial conjugation. However, it has been reported that the EcoKI R-M system affects the uptake of DNA by conjugation [32] and, ArdC is not expressed until plasmid DNA is in dsDNA shape. Thus, Type I R-M system could be attacking late, once ArdC is generated from the dsDNA plasmid after DNA entrance and establishment in recipient cells. In this respect, the observation that mutation of the ArdC metalloprotease active center does not reduce interspecies conjugation suggests that the metalloprotease activity is not required during conjugative transfer to *P. putida*. The presumed metalloprotease activity is not expected to play a role in HsdRMS activity. Thus, it is tempting to propose that, just by ArdC binding to DNA, the protein interferes with HsdRMS binding and thereby hindering the degradation of its target DNA.

In summary, our results indicate a new mechanism of DNA antirestriction played by protein ArdC, by which plasmids increase their conjugation host range. Interfering with ArdC activity could thus provide a new tool to hinder the transmission of antibiotic resistance.

Materials and methods

Conjugation assays

Conjugation assays were performed by mixing *E. coli* or *P. putida* cells (Table 2) containing the plasmid of interest (Table 3) with recipient cells grown overnight at their optimal growing temperature. *E. coli* cells were grown at 37 ºC and *P. putida* at 30 ºC. Cells at OD600 = 0.6 were mixed in a 1: 1 donor: recipient ratio, washed on LB medium, resuspended in 30 μL of LB and deposited on a 0.22 μm pore size cellulose acetate filters (Sartorius Stedim) in LB-agar plates previously incubated at 37ºC unless otherwise indicated. After 1 h, filters were removed with sterile tweezers and introduced in 1 mL LB, where cells were resuspended by vortexing for a few seconds. 1/10 serial dilutions were done and 10 μL drops were plated in LB agar plates with the appropriate selecting antibiotics for donors, recipients and transconjugants. Conjugation frequencies were obtained by dividing transconjugants per recipients (T/R). For conjugations in the presence of pUCP22-derived plasmids, isopropyl β-D-thiogalactoside (IPTG) was added to the conjugation mixture to a 0.1 mM IPTG final concentration. Means and standard deviations, as well as statistical tests, were calculated with GraphPad Prism1 (v 7.04) biostatistics software.

Transcriptomic analysis

Mixtures of *E. coli* BW27783-Nx^R (bearing pSU2007, pLGM25 or no plasmid) with *P. putida* KT2440 were carried out by the already described conjugation assay in a 5:1 donor: recipient ratio for 30 minutes. Harvested cells from the conjugation filter were treated with RNAprotect® Bacteria Reagent (Qiagen) and snap-frozen. Cells were lysed with lysozyme (Sigma) and

Table 2. Strains used in this study.

Strain	Phenotype	Reference
Escherichia coli		
DH5α	F- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG*, Φ80d*lacZΔM15 Δ(lacZYA-argF)*U169, *hsd*R17(rK- mK+), λ–	[33]
BL21 (DE3)	F– *ompT gal dcm lon hsdSB*(rB–mB–) λ(DE3[*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*+]K-12(λS)	[34]
BL21(DE3)-SmR	Streptomycin resistant spontaneous mutant of BL21(DE3)	This work
C41 (DE3)	F- *ompT hsdSB (rB- mB-) gal dcm* (DE3)	[35]
β834(DE3)	F- *ompT hsdSB*(rB- mB-) *gal dcm met* (DE3)	[36]
TB10	TB10 is the result of a P1 transduction from DY329 into MG1655. It has a large amount of the λ prophage genome inserted into a biotin operon. The λ red genes α, β and γ are under the control of cI857, making it temperature inducible.	[37] and [38]
DY380	Sm^R λ Cl857 (cro⁻bioA) tet (DH10B)	[39]
BW27783	F-, Δ(*araD-araB*)567, Δ*lacZ*4787(::rrnB-3), λ-, *rph-1*, Δ(*rhaD-rhaB*)568, *hsd*R514 Δ(*araH-araF*)570(::FRT), Δ*araE*p-532::FRT, φP*cp8araE*535	[40]
BW27783-NxR	Nalidixic resistant spontaneous mutant of BW27783	[41]
BW27783-RifR	Rifampicin resistant spontaneous mutant of BW27783	[41]
MG1655	K-12 F–λ–*ilvG–rfb-50 rph-1*	[42]
MG1655-RifR	Rifampicin resistant spontaneous mutant of MG1655	This work
EcMR2Δ*mutS*	MG1655, *lacI- bla, bio-, lambda-Red1, mutS*–, cmR	[43]
Pseudomonas putida		
KT2440	Wild-type *P. putida* strain; mt-2 derivative cured of its plasmid (pWW0-)	[44]
EM178	KT2440 derivative; Δprophage1 Δprophage4 Δprophage3 Δprophage2	[45]
EM42	KT2440 derivative; Δprophage1 Δprophage4 Δprophage3 Δprophage2 Δ*tn7* Δ*endA-1* Δ*endA-2* Δ*hsdRMS* Δflagellum Δ*tn4652*	[45]
EM422	KT2440 derivative; Δ*hsdRMS*	From De Lorenzo group
KT2440 Δ*recA*	KT2440 derivative; Δ*recA*	From De Lorenzo group
KT2440 Δflagellum	KT2440 derivative; Δflagellum	From De Lorenzo group
KT2440 Δ*endA1*	KT2440 derivative; Δ*endA-1*	From De Lorenzo group
KT2440 Δ*endA2*	KT2440 derivative; Δ*endA-2*	From De Lorenzo group
KT2440 Δ*tn7*	KT2440 derivative; Δ*tn7*	From De Lorenzo group
KT2440 Δ*tn4652* pSW	KT2440 derivative; Δ*tn4652 bearing pSW plasmid*	From De Lorenzo group
KT2440 Δ*tn4652*	KT2440 Δ*tn4652* derivative cured of pSW plasmid	This work
Other bacteria		
Salmonella typhimurium LT2	*Salmonella enterica* subsp. *enterica* serovar *typhimurium* str. *LT2*	ATCC 700720
Klebsiella pneumoniae K6	*K. pneumoniae* subsp. *pneumoniae*. Clinical isolate from the Medical College of Virginia, 1994. Ap^R KmR, Cm^R	ATCC 700603
Acinetobacter baumannii	Sm^R Ap^R	ATCC 19606
Vibrio cholerae CIP106855	N16961 Rif^R. Biovar Eltor, serovar O:1	CIP106855
Agrobacterium tumefaciens GMI9023	*Agrobacterium tumefaciens* C58 derivative cured of its plasmids (pTi-, pAT-)	[46]

https://doi.org/10.1371/journal.pgen.1008750.t002

Table 3. Plasmids used in this study.

Plasmid	Description	Phenotype	Size (Kb)	Reference
R388	R388 wild type plasmid	Su^R Tp^R; (IncW)	33.9	[47]
pSU2007	R388 derivative; KnR cassette insertion	Su^R Tp^R Kn^R; (IncW)	32.9	[48]
pET29c	Expression vector	Kn^R; Rep (pMB1); Overexpression controlled by T7 promoter with a 6-HisTag.	5.4	Addgene
pUA66	GFP reporter plasmid	Kn^R pSC101 replicon	4.5	[49]
pUCP22	Shuttle Vector; Escherichia-Pseudomonas broad-host-range expression vector	Ap^R Gm^R; Plac promoter.	4.7	[50]
pHERD20T	Shuttle Vector; Escherichia-Pseudomonas broad-host-range expression vector	Cb^R; PBAD promoter and araC regulator.	5.1	[51]
pLGM21	pET29c::ardC	Kn^R T7 promoter	6.1	This work
pIC10	R388ΔkfrA-orf14	Tp^R Kn^R	26.2	This work
pLGM25	R388ΔardC	Tp^R Kn^R	33.9	This work
pLGM28	pET29c::ardC_E229A	Kn^R T7 promoter	6.1	This work
pLGM33	pSU2007 (ardC_E229A)	Tp^R Kn^R	32.9	This work
pLGM36	pUCP22::ardC	Ap^R Gm^R Plac promoter	5.6	This work
pLGM37	pUCP22::ardC_E229A	Ap^R Gm^R Plac promoter	5.6	This work

https://doi.org/10.1371/journal.pgen.1008750.t003

proteinase K (Roche). Total RNA was extracted with RNeasy Mini Kit (Qiagen) and treated with RNase-free DNase (Qiagen) in column for DNA removal. Ambion TURBO DNA-free™ DNase Treatment was also applied for better DNA removal. RNA integrity and quality were validated by the Agilent RNA ScreenTape assay. The RNA integrity number equivalent (RINe) was assured to be above 8 to use the isolated RNA in the RNA-seq experiment.

Transcriptome libraries were prepared by Macrogen (Seoul, Korea) with Ribo-Zero rRNA Removal Kit and TruSeq Stranded mRNA sample preparation kit (Illumina). Libraries were sequenced by Macrogen on the Illumina HiSeq 4000 platform. The transcriptome libraries were paired-end sequenced with 100-bp reads. Raw reads in FASTQ format were quality analyzed with FastQC [52]. Reads were mapped against R388 (NCBI Accession number NC_028464.1), *E. coli* str. K-12 substr. MG1655 (U00096.3) and *P. putida* KT2440 (AE015451.2) sequences. The alignment of reads was done with Bowtie2 software [53]. Artemis program [54] was used to visualize the alignment and do the RPKM (reads per kilobase and million mapped reads) calculations. Genes with less than 10 RPKMs in all experimental conditions were removed from the analysis. DAVID online tool v6.8 [55] was used to test for gene ontology enrichment among the list of differentially expressed genes to do a functional classification.

Protein expression and purification

Plasmids pLGM21 (pET29c::*ardC*) or pLGM28 (pET29c::*ardC_E229A*) were transformed into electrocompetent *Escherichia coli* BL21 (DE3) cells (Table 2). Transformed cells were grown in 1L LB medium, in the presence of kanamycin, at 37˚C, with shaking, to an optical density of 0.5–0.6. The temperature was then reduced to 18˚C and protein expression was induced with IPTG to a final concentration of 0.5 mM. Cells were allowed to grow for 16 h. The cultures were then centrifuged at 5,000 rpm and 4˚C for 15 min and the resulting pellets were stored at −20˚C. For protein purification, pellets were resuspended in buffer A (500 mM NaCl, 20 mM imidazole, 100 mM Tris-HCl pH 7.5) supplemented with protease inhibitor phenylmethylsulfonyl fluoride (PMSF) 1% (v/v). The slurry was sonicated in a Labsonic 2000 (B. Braun) equipment at 50% of potency for 3 cycles of 1.5 min at intervals of 1 min on ice. The lysed cells were

then ultra-centrifuged at 100,000 g for 15 min at 4˚C. Supernatants were loaded onto a 5 mL HisTrap HP column (GE Healthcare) previously equilibrated with buffer A. Proteins were eluted by an imidazole concentration gradient between buffer A and buffer B (300 mM NaCl, 500 mM imidazole, 100 mM Tris-HCl pH 7.5). ArdC containing fractions were pooled and diluted to a final NaCl concentration of 200 mM. The resulting protein was then loaded onto a 5 mL HiTrap Heparin HP (GE Healthcare) equilibrated with buffer C (100 mM Tris-HCl pH 7.5, 200 mM NaCl). Proteins were eluted by a linear gradient between buffer C and D (100 mM Tris-HCl pH 7.5, 1 M NaCl). ArdC containing fractions were concentrated using Amicon Ultra 30k Centrifugal filters (Millipore, Ireland) and loaded onto a Superdex 75 GL10_30 column (GE Healthcare) previously equilibrated with buffer E (100mM Tris-HCl pH 7.5, 1 mM EDTA, 300 mM NaCl).

To crystallize ArdC with a metal cofactor, cell lysis and the two first purification steps were done as described but with 1 mM $MnCl_2$ in all buffers. Preparation of selenomethionine (SeMet)-labelled ArdC was also carried out as described above but using strain *E. coli* β834 (DE3) and minimal medium (SelenoMet Medium Base + SelenoMet Nutrient Mix) supplemented with SelenoMethionine Solution (Molecular Dimensions) as indicated by the manufacturer.

Protein crystallization and structure determination

Crystals of ArdC and ArdC-SeMet were obtained using the sitting-drop vapor-diffusion method at 22˚C by mixing 1.5 μL protein at 20 mg/mL concentration in 20 mM Tris-HCl, 50 mM NaCl, 1 mM EDTA buffer with an equal volume of the reservoir solution containing 0.1 M HEPES pH 7.5; 10% w/v polyethylene glycol 6,000 and 5% v/v (+/-)-2-methyl-2,4-pentanediol. 2-methyl-2 4-pentanediol (10–20% v/v) was added as cryoprotectant before diffraction experiments. ArdC-Mn crystallized at 12 mg/mL in 25% v/v ethylene glycol and crystals were cryoprotected with additional 15% glycerol.

For data collection, the crystals were flash-frozen in liquid nitrogen at 105 K. For single ArdC-SeMet crystals data was collected at 0.9793Å, the wavelength corresponding to the Selenium absorption maximum according to the fluorescence scan. Datasets were obtained at beamline XALOC at the ALBA Synchrotron Radiation Facility (Barcelona, Spain) with a Dectris PILATUS3 6M Pixel detector. Diffraction images were processed using iMosflm [56] and Scala [57] as part of the CCP4 package [58]. The structure was solved by single anomalous dispersion (SAD) phasing using the program AutoSol of the PHENIX package [59]. The refinement of the initial model was performed through several cycles by Phenix refine [59] until appropriate R factors were reached. Final manual modeling was done in COOT [60]. The ArdC-Mn structure was solved by MR using the ArdC structure as a template.

Electrophoretic mobility shift assay (EMSA)

The binding ability of the ArdC protein to ssDNA, dsDNA, and dsDNA with ssDNA overhangs was tested by electrophoresis mobility shift assay (EMSA). 6FAM-labeled oligonucleotide Fluor-T87I2 (45 bases) was incubated alone or with T87I1 oligonucleotide (45 complementary bases), Mid1 oligonucleotide (13 5' terminal complementary bases) or Mid2 oligonucleotide (27 3' terminal complementary bases) (see S9 Table) in buffer containing 50mM Tris-HCl (pH 7.5) and 1mM EDTA for 5 min at 95˚C and the mixtures were cooled down slowly to room temperature. 50 nM DNA was incubated with various concentrations of ArdC (0, 125 nM, 250 nM, 500 nM and 1 μM) in a reaction buffer [50 mM NaCl, 25 mM Tris-HCl (pH 7.5), 0.5 mM EDTA] at room temperature for 30 min. DNA-protein complexes were analyzed using non-denaturing polyacrylamide gel electrophoresis 10% (29:1) in cold Tris

Borate EDTA (TBE) buffer 1x. Gels were run at 100 V for 75 min and analyzed using a Fujifilm fluorescent image analyzer Fla-5100. Experiments were repeated three times.

Extended Materials and Methods are included in S1 Text.

Supporting information

S1 Text. Supplementary Materials and Methods.
(DOCX)

S1 Fig. Genetic map of R388 plasmid. Genetic organization of the plasmid divided into functional modules and three different sectors: conjugation shadowed in orange, general maintenance in light blue and Ab^R and integration in grey. Region deleted in pLGM25 (shown in maroon) and pIC10 (shown in purple) are also shown. Adapted from [12].
(TIF)

S2 Fig. Retardation and protection of ssDNA (M13mp18) by ArdC from degradation by HhaI. A) Agarose gel showing ssDNA retardation by ArdC under non-denaturing conditions. Lane 1: ssDNA (5.5 nM) in the presence of $MgCl_2$. The vast majority of the molecules of M13mp18 ssDNA are circular (upper band), although some of them are present in the linear form (lower band). Lane 2: ssDNA and HhaI (7 U) in the presence of $MgCl_2$. Lane 3–7: ssDNA and HhaI at increasing concentrations of ArdC (0.95 μM (3), 1.9 μM (4), 3.8 μM (5), 5.7 μM (6) and 7.8 μM (7)) in the presence of $MgCl_2$. B) Agarose gel showing ssDNA protection by ArdC from HhaI digestion. Proteinase K and SDS were added to the samples before loading the gel to remove HhaI and ArdC proteins. Lane content as in A). C) SDS-PAGE gel showing the proteolytic activity of ArdC preincubated with ssDNA (M13mp18) to HhaI. Lane 1: ArdC (8 μM) in the presence of $MgCl_2$. Lane 2: HhaI (70 U) in the presence of $MgCl_2$. Lane 3: ArdC and HhaI in the presence of $MgCl_2$. Lane 4–5: ArdC and HhaI in the presence of ssDNA (M13, 27.5 nM) with $MgCl_2$ (4) or EDTA (5).
(TIF)

S3 Fig. Mass spectrometry analysis for protein identification. A) Protein from a pull-down experiment cleaved from the gel and sent for mass spectrometry analysis is indicated by an arrow. B) MASCOT search result for the peptides obtained after digestion with trypsin (that cuts the C-term side of KR unless next residue is P) in SwissProt database for *P. putida* KT2440. Matched peptides with pp_0941 protein are shown in purple (sequence coverage of 45%). pp_0941 has 173 amino acids and a molecular weight of 20238 Da.
(TIF)

S4 Fig. Differential expression of genes in experiment *ardC*$^+$ vs. *ardC*—to test the influence of ArdC in the three reference sequences. A) Differential expression of R388 genes. B) Differential expression of *E. coli* genes. C) Differential expression of *P. putida* genes. Upregulated genes with an RPKM fold change >2 are in the green zone, and downregulated genes with RPKM fold change <2 are in the orange zone.
(TIF)

S5 Fig. Effect of ArdC on plasmid conjugative transfer from *E. coli* to *P. putida* wt and mutants at 30˚C. The conjugation frequencies per recipient (T/R) into *P. putida* KT2440 WT strain or different mutants of *P. putida* KT2440 are shown. Conjugation was done for 1 h at 30˚C. Horizontal bars represent the mean ± SD obtained for each dataset of N = 9 (t-test: ** p < 0.01).
(TIF)

S6 Fig. Plasmids encoding TOPRIM and DUF1738 domains. Venn diagrams showing A) the proteins containing either DUF1738, TOPRIM or DUF1738+TOPRIM domains and B) plasmids encoding proteins containing DUF1738 and/or TOPRIM domains either in separate or multidomain proteins. Below, a panel indicates the fraction of plasmids potentially transmissible by conjugation (MOB$^+$) for each condition. Proteins contained in RefSeq plasmid database version 90 were screened with the HMM profiles PF13362 for TOPRIM and PF08401 for DUF1378, using the hmmsearch function of HMMER 3.1b2 with the default parameters [61]. The presence of MOB relaxases was detected by using MOBscan [27]. Venn diagrams were built using the online tool http://bioinformatics.psb.ugent.be/webtools/Venn/.
(TIF)

S1 Table. T$_M$ value of ArdC in different solutions.
(DOCX)

S2 Table. Conjugation frequencies from *E. coli* to *P. putida*.
(DOCX)

S3 Table. TruSeq Stranded mRNA Illumina sequencing results and coverage for each condition.
(DOCX)

S4 Table. Percentage of reads aligned to the three reference sequences by Bowtie2.
(DOCX)

S5 Table. Distribution of the differentially upregulated or downregulated genes for the three reference sequences and conditions.
(DOCX)

S6 Table. Expression profile of R388 genes.
(DOCX)

S7 Table. Expression profile of upregulated *E. coli* genome-encoded genes in the absence of *ardC* in the plasmid.
(DOCX)

S8 Table. Expression profile of differentially upregulated *P. putida* genes in the presence/absence of *ardC*.
(DOCX)

S9 Table. Oligonucleotides used in this study.
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: M. Pilar Garcillán-Barcia, Gabriel Moncalián.

References

1. Soucy SM, Huang J, Gogarten JP. Horizontal gene transfer: Building the web of life. Nat Rev Genet. 2015; 16: 472–482. https://doi.org/10.1038/nrg3962 PMID: 26184597

2. Craig MacLean R., San Millan A. The Evolution of Antibiotic Resistance. Science. 2019; 365: 1082–1083. https://doi.org/10.1126/science.aax3879 PMID: 31515374

3. Garcillán-Barcia MP, de la Cruz F. Why is entry exclusion an essential feature of conjugative plasmids? Plasmid. 2008; 60: 1–18. https://doi.org/10.1016/j.plasmid.2008.03.002 PMID: 18440635

4. Frost LS, Koraimann G. Regulation of bacterial conjugation: Balancing opportunity with adversity. Future Microbiol. 2010; 5: 1057–1071. https://doi.org/10.2217/fmb.10.70 PMID: 20632805

5. Petrova V, Chitteni-Pattu S, Drees JC, Inman RB, Cox MM. An SOS Inhibitor that Binds to Free RecA Protein: The PsiB Protein. Mol Cell. 2009; 36: 121–130. https://doi.org/10.1016/j.molcel.2009.07.026 PMID: 19818715

6. Bagdasarian M, Bailone A, Angulo JF, Scholz P, Bagdasarian M, Devoret R. PsiB, and anti-SOS protein, is transiently expressed by the F sex factor during its transmission to an Escherichia coli K-12 recipient. Mol Microbiol. 1992; 6: 885–893. https://doi.org/10.1111/j.1365-2958.1992.tb01539.x PMID: 1318487

7. Marraffini LA. CRISPR-Cas immunity in prokaryotes. Nature. 2015; 526: 55–61. https://doi.org/10.1038/nature15386 PMID: 26432244

8. Gormley NA, Watson MA, Halford SE. Bacterial Restriction–Modification Systems. Encycl Life Sci. 2005; 1–11. https://doi.org/10.1038/npg.els.0001037

9. Wilkins BM. Plasmid promiscuity: meeting the challenge of DNA immigration control. Environ Microbiol. 2002; 4: 495–500. https://doi.org/10.1046/j.1462-2920.2002.00332.x PMID: 12220405

10. Tock MR, Dryden DTF. The biology of restriction and anti-restriction. Curr Opin Microbiol. 2005; 8: 466–472. https://doi.org/10.1016/j.mib.2005.06.003 PMID: 15979932

11. Murray NE. Type I restriction systems: Sophisticated molecular machines. Microbiol Mol Biol Rev. 2000; 64: 412–434. https://doi.org/10.1128/mmbr.64.2.412-434.2000 PMID: 10839821

12. Fernández-López R, Pilar Garcillán-Barcia M, Revilla C, Lázaro M, Vielva L, De La Cruz F. Dynamics of the IncW genetic backbone imply general trends in conjugative plasmid evolution. FEMS Microbiol Rev. 2006; 30: 942–966. https://doi.org/10.1111/j.1574-6976.2006.00042.x PMID: 17026718

13. Chen CY, Kado CI. Inhibition of Agrobacterium tumefaciens oncogenicity by the osa gene of pSa. J Bacteriol. 1994; 176: 5697–5703. https://doi.org/10.1128/jb.176.18.5697-5703.1994 PMID: 8083162

14. Belogurov AA, Delver EP, Agafonova O V, Belogurova NG, Lee LY, Kado CI. Antirestriction protein Ard (Type C) encoded by IncW plasmid pSa has a high similarity to the "protein transport" domain of TraC1 primase of promiscuous plasmid RP4. J Mol Biol. 2000; 296: 969–977. https://doi.org/10.1006/jmbi.1999.3493 PMID: 10686096

15. Krishnan A, Burroughs AM, Iyer LM, Aravind L. Unexpected Evolution of Lesion-Recognition Modules in Eukaryotic NER and Kinetoplast DNA Dynamics Proteins from Bacterial Mobile Elements. iScience. 2018; 9: 192–208. https://doi.org/10.1016/j.isci.2018.10.017 PMID: 30396152

16. Holm L, Sander C. Dali: a network tool for protein structure comparison. Trends Biochem Sci. 1995; 20: 478–480. https://doi.org/10.1016/s0968-0004(00)89105-7 PMID: 8578593

17. Min J-H, Pavletich NP. Recognition of DNA damage by the Rad4 nucleotide excision repair protein. Nature. 2007; 449: 570–575. https://doi.org/10.1038/nature06155 PMID: 17882165

18. Li F, Raczynska JE, Chen Z, Yu H. Structural Insight into DNA-Dependent Activation of Human Metallo-protease Spartan. Cell Rep. 2019; 26: 3336–3346. https://doi.org/10.1016/j.celrep.2019.02.082 PMID: 30893605

19. Wang Y, Xu Q, Lu H, Lin L, Wang L, Xu H, et al. Protease activity of PprI facilitates DNA damage response: Mn(2+)-dependence and substrate sequence-specificity of the proteolytic reaction. Huen MS-Y, editor. PLoS ONE. 2015; 10: e01220: 1–17. https://doi.org/10.1371/journal.pone.0122071 PMID: 25811789

20. Cerdà-Costa N, Gomis-Rüth FX. Architecture and function of metallopeptidase catalytic domains. Protein Soc. 2014; 23: 123–144. https://doi.org/10.1002/pro.2400 PMID: 24596965

21. Fukasawa KM, Hata T, Ono Y, Hirose J. Metal Preferences of Zinc-Binding Motif on Metalloproteases. J Amino Acids. 2011; 574816: 1–7. https://doi.org/10.4061/2011/574816 PMID: 22312463

22. Matthews BW. Structural Basis of the Action of Thermolysin and Related Zinc Peptidases Inhibitor Binding. Acc Chem Res. 1988; 21: 333–340.

23. Fernandez-Lopez R, del Campo I, Revilla C, Cuevas A, de la Cruz F. Negative Feedback and Transcriptional Overshooting in a Regulatory Network for Horizontal Gene Transfer. PLoS Genet. 2014; 10: e10041: 1–15. https://doi.org/10.1371/journal.pgen.1004171 PMID: 24586200

24. Merryweather A, Rees CE, Smith NM, Wilkins BM. Role of sog polypeptides specified by plasmid ColIb-P9 and their transfer between conjugating bacteria. EMBO J. 1986; 5: 3007–3012. PMID: 3024972

25. Serfiotis-Mitsa D, Herbert AP, Roberts GA, Soares DC, White JH, Blakely GW, et al. The structure of the KlcA and ArdB proteins reveals a novel fold and antirestriction activity against Type I DNA restriction systems in vivo but not in vitro. Nucleic Acids Res. 2010; 38: 1723–1737. https://doi.org/10.1093/nar/gkp1144 PMID: 20007596

26. Goryanin II, Kudryavtseva AA, Balabanov VP, Biryukova VS, Manukhov IV, Zavilgelsky GB. Antirestriction activities of KlcA (RP4) and ArdB (R64) proteins. FEMS Microbiol Lett. 2018;365. https://doi.org/10.1093/femsle/fny227 PMID: 30239714

27. Garcillán-Barcia MP, Redondo-Salvo S, Vielva L, de la Cruz F. MOBscan: Automated Annotation of MOB Relaxases. Methods Mol Biol Clifton NJ. 2020; 2075: 295–308. https://doi.org/10.1007/978-1-4939-9877-7_21 PMID: 31584171

28. Melkina OE, Goryanin II, Zavilgelsky GB. The DNA–mimic antirestriction proteins ArdA ColIB-P9, Arn T4, and Ocr T7 as activators of H-NS-dependent gene transcription. Microbiol Res. 2016; 192: 283–291. https://doi.org/10.1016/j.micres.2016.07.008 PMID: 27664747

29. Serfiotis-Mitsa D, Herbert AP, Roberts GA, Soares DC, White JH, Blakely GW, et al. The structure of the KlcA and ArdB proteins reveals a novel fold and antirestriction activity against type I DNA restriction systems in vivo but not in vitro. Nucleic Acids Res. 2009; 38: 1723–1737. https://doi.org/10.1093/nar/gkp1144 PMID: 20007596

30. Devigne A, Ithurbide S, Bouthier de la Tour C, Passot F, Mathieu M, Sommer S, et al. DdrO is an essential protein that regulates the radiation desiccation response and the apoptotic-like cell death in the radioresistant Deinococcus radiodurans bacterium. Mol Microbiol. 2015; 96: 1069–1084. https://doi.org/10.1111/mmi.12991 PMID: 25754115

31. Baharoglu Z, Bikard D, Mazel D. Conjugative DNA transfer induces the bacterial SOS response and promotes antibiotic resistance development through integron activation. PLoS Genet. 2010; 6: e100116: 1–10. https://doi.org/10.1371/journal.pgen.1001165 PMID: 20975940

32. Roer L, Aarestrup FM, Hasman H. The EcoKI type I restriction-modification system in Escherichia coli affects but is not an absolute barrier for conjugation. J Bacteriol. 2015; 197: 337–342. https://doi.org/10.1128/JB.02418-14 PMID: 25384481

33. Grant SG, Jessee J, Bloom FR, Hanahan D. Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. Proc Natl Acad Sci. 1990; 87: 4645–4649. https://doi.org/10.1073/pnas.87.12.4645 PMID: 2162051

34. Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol. 1986; 189: 113–130. https://doi.org/10.1016/0022-2836(86)90385-2 PMID: 3537305

35. Miroux B, Walker JE. Over-production of proteins in Escherichia coli: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol. 1996; 260: 289–298. https://doi.org/10.1006/jmbi.1996.0399 PMID: 8757792

36. Budisa N, Steipe B, Demange P, Eckerskorn C, Kellermann J, Huber R. High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in Escherichia coli. Eur J Biochem. 1995; 230: 788–796. https://doi.org/10.1111/j.1432-1033.1995.tb20622.x PMID: 7607253

37. Yu D, Ellis HM, Lee E-C, Jenkins NA, Copeland NG, Court DL. An efficient recombination system for chromosome engineering in Escherichia coli. Proc Natl Acad Sci. 2000; 97: 5978–5983. https://doi.org/10.1073/pnas.100127597 PMID: 10811905

38. Silver PA, Stirling F, Bitzan L, Way J, Oliver JWK, Redfield E, et al. Rational Design of Evolutionarily Stable Microbial Kill Switches. Mol Cell. 2017; 68:.e3: 686–697. https://doi.org/10.1016/j.molcel.2017.10.033 PMID: 29149596

39. Lee EC, Yu D, Martinez De Velasco J, Tessarollo L, Swing DA, Court DL, et al. A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. Genomics. 2001; 73: 56–65. https://doi.org/10.1006/geno.2000.6451 PMID: 11352566

40. Keasling JD, Wanner BL, Skaug T, Datsenko KA, Khlebnikov A. Homogeneous expression of the PBAD promoter in Escherichia coli by constitutive expression of the low-affinity high-capacity AraE transporter. Microbiology. 2001; 147: 3241–3247. https://doi.org/10.1099/00221287-147-12-3241 PMID: 11739756

41. del Campo I, Ruiz R, Cuevas A, Revilla C, Vielva L, de la Cruz F. Determination of conjugation rates on solid surfaces. Plasmid. 2012; 67: 174–182. https://doi.org/10.1016/j.plasmid.2012.01.008 PMID: 22289895

42. Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, et al. The complete genome sequence of Escherichia coli K-12. Science. 1997; 277: 1453–1462. https://doi.org/10.1126/science.277.5331.1453 PMID: 9278503

43. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, et al. Programming cells by multiplex genome engineering and accelerated evolution. Nature. 2009; 460: 894–898. https://doi.org/10.1038/nature08187 PMID: 19633652

44. Bagdasarian M, Lurz R, Rückert B, Bagdasarian MM, Frey J, Timmis KN. Specific-purpose plasmid cloning vectors II. Broad host range, high copy number, RSF1010-derived gene cloning in Pseudomon. Gene. 1981; 16: 237–247. https://doi.org/10.1016/0378-1119(81)90080-9 PMID: 6282695

45. Martínez-García E, Nikel PI, Aparicio T, de Lorenzo V. Pseudomonas 2.0: Genetic upgrading of P. putida KT2440 as an enhanced host for heterologous gene expression. Microb Cell Factories. 2014; 13: 1–15. https://doi.org/10.1186/s12934-014-0159-3 PMID: 25384394

46. Rosenberg C, Huguet T. The pAtC58 plasmid of Agrobacterium tumefaciens is not essential for tumour induction. MGG Mol Gen Genet. 1984; 196: 533–536. https://doi.org/10.1007/BF00436205

47. Datta Naomi & Hedges R. W. Trimethoprim Resistance Conferred by W Plasmids in Enterobacteriaceae. J Gen Microbiol. 1972; 72: 349–355. https://doi.org/10.1099/00221287-72-2-349 PMID: 4562309

48. Martinez E, de la Cruz F. Transposon Tn21 encodes a RecA-independent site-specific integration system. Mol Gen Genet. 1988; 211: 320–325. https://doi.org/10.1007/bf00330610 PMID: 2832705

49. Zaslaver A, Bren A, Ronen M, Itzkovitz S, Kikoin I, Shavit S, et al. A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. Nat Methods. 2006; 3: 623–628. https://doi.org/10.1038/nmeth895 PMID: 16862137

50. West SEH, Schweizer HP, Dall C, Sample AK, Runyen-Janecky LJ. Construction of improved Escherichia-Pseudomonas shuttle vectors derived from pUC18/19 and sequence of the region required for their replication in Pseudomonas aeruginosa. Gene. 1994; 148: 81–86. https://doi.org/10.1016/0378-1119(94)90237-2 PMID: 7926843

51. Qiu D, Damron FH, Mima T, Schweizer HP, Yu HD. PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in Pseudomonas and Burkholderia spp. and other bacteria. Appl Environ Microbiol. 2008; 74: 7422–6. https://doi.org/10.1128/AEM.01369-08 PMID: 18849445

52. Andrews S. FASTQC A Quality Control tool for High Throughput Sequence Data. In: Babraham Institute [Internet]. 2015. Available: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/Help/3AnalysisModules/

53. Ben Langmead, Steven S. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2013; 9: 357–359. https://doi.org/10.1038/nmeth.1923.Fast

54. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: An integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics. 2012; 28: 464–469. https://doi.org/10.1093/bioinformatics/btr703 PMID: 22199388

55. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009; 4: 44–57. https://doi.org/10.1038/nprot.2008.211 PMID: 19131956

56. Battye TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr. 2011; 67: 271–281. https://doi.org/10.1107/S0907444910048675 PMID: 21460445

57. Evans P, IUCr. Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr. 2006; 62: 72–82. https://doi.org/10.1107/S0907444905036693 PMID: 16369096

58. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallographica Section D: Biological Crystallography. 2011. pp. 235–242. https://doi.org/10.1107/S0907444910045749 PMID: 21460441

59. Afonine P.V.; Grosse-Kunstleve R.W; Adams P.D. The Phenix refinement framework. CCP4 Newsletter on protein Crystallography 42, contribution 8. 2005.

60. Emsley P, Cowtan K. *Coot*: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60: 2126–2132. https://doi.org/10.1107/S0907444904019158 PMID: 15572765

61. Eddy SR. A new generation of homology search tools based on probabilistic inference. Genome Inform Int Conf Genome Inform. 2009; 23: 205–211.