Computational and Structural Biotechnology Journal 20 (2022) 757–765

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

# Structural and biochemical characterization of the relaxosome auxiliary proteins encoded on the *Bacillus subtilis* plasmid pLS20

Isidro Crespo [a,1], Nerea Bernardo [a,1], Anna Cuppari [a], Barbara M. Calisto [a], Jorge Val-Calvo [b], Andrés Miguel-Arribas [b], Wilfried J.J. Meijer [b], Xavi Carpena [a], Fernando Gil-Ortiz [a], Marc Malfois [a], D. Roeland Boer [a,*]

[a] ALBA Synchrotron Light Source, Carrer de la Llum 2-26, Cerdanyola del Vallès, 08290 Barcelona, Spain
[b] Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), C. Nicolás Cabrera 1, Universidad Autónoma, Canto Blanco, 28049 Madrid, Spain

ARTICLE INFO

*Article history:*
Received 14 March 2021
Received in revised form 22 December 2021
Accepted 30 December 2021
Available online 05 January 2022

*Keywords:*
Structural biology
Bacterial conjugation
Relaxosome
Auxiliary protein
DNA binding protein
Ribbon-Helix-Helix
Antibiotic resistance
*Firmicutes*
Horizontal gene transfer

ABSTRACT

Bacterial conjugation is an important route for horizontal gene transfer. The initial step in this process involves a macromolecular protein-DNA complex called the relaxosome, which in plasmids consists of the origin of transfer (*oriT*) and several proteins that prepare the transfer. The relaxosome protein named relaxase introduces a nick in one of the strands of the *oriT* to initiate the process. Additional relaxosome proteins can exist. Recently, several relaxosome proteins encoded on the *Bacillus subtilis* plasmid pLS20 were identified, including the relaxase, named $Rel_{pLS20}$, and two auxiliary DNA-binding factors, named $Aux1_{pLS20}$ and $Aux2_{pLS20}$. Here, we extend this characterization in order to define their function. We present the low-resolution SAXS envelope of the $Aux1_{pLS20}$ and the atomic X-ray structure of the C-terminal domain of $Aux2_{pLS20}$. We also study the interactions between the auxiliary proteins and the full-length $Rel_{pLS20}$, as well as its separate domains. The results show that the quaternary structure of the auxiliary protein $Aux1_{pLS20}$ involves a tetramer, as previously determined. The crystal structure of the C-terminal domain of $Aux2_{pLS20}$ shows that it forms a tetramer and suggests that it is an analog of $TraM_{pF}$ of plasmid F. This is the first evidence of the existence of a $TraM_{pF}$ analog in gram positive conjugative systems, although, unlike other $TraM_{pF}$ analogs, $Aux2_{pLS20}$ does not interact with the relaxase. $Aux1_{pLS20}$ interacts with the C-terminal domain, but not the N-terminal domain, of the relaxase $Rel_{pLS20}$. Thus, the pLS20 relaxosome exhibits some unique features despite the apparent similarity to some well-studied G- conjugation systems.

© 2022 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

## 1. Introduction

Horizontal gene transfer (HGT) is the exchange of genes between organisms not related to transmission of genes between parents and offspring and has important implications on evolution. Conjugation is one of the routes that allow HGT and is common in the bacterial realm [1–6]. A membrane-embedded cellular machinery named the Type 4 Secretion System (T4SS) is used in conjugation for the transfer of a conjugative element from a donor to a recipient cell [7–11]. The relaxosome is another important component of T4SS-mediated conjugation and prepares the DNA for transfer [5,12]. Conjugation also plays an important role in bacterial virulence, as the T4SS system is also used for the extrusion of virulence factor into the host environment (*eg* [13,14]). Bacterial T4SS systems occur both in Gram positive (G+) and Gram negative (G-) bacteria and may be encoded in Mobile Genetic Elements (MGEs) such as plasmids [6].

Conjugation requires the initial preparation of the DNA that is to be transferred [5,10,11,15]. This step involves an initial cut or nick at a specific site of one of the DNA strands, which then allows for the unwinding of this strand and subsequent transfer [5]. This critical step is performed by the relaxase [16–18], which have in common that they consist of an N-terminal endonuclease domain, followed by additional domains with a variety of functions [18–21]. The endonuclease domain of the relaxase remains covalently attached to the ssDNA strand and this complex is transferred to

*Abbreviations:* HGT, Horizontal Gene Transfer; MGE, Mobile Genetic Element; T4SS, Type IV secretion system; EM, Electron Micoscopy; RHH, Ribbon-Helix-Helix; EDTA, Ethylenediaminetetraacetic acid; IPTG, Isopropyl β-D-1-thiogalactopyranoside; PMSF, phenylmethylsulfonyl fluoride; SEC, Size Exclusion Chromatography; SDS-PAGE, sodium dodecyl sulphate–polyacrylamide gel electrophoresis; SAXS, Small-angle X-ray scattering; *oriT*, Origin of Transfer; AUC, Analytical Ultracentrifugation; ITC, Isothermal titration calorimetry.

* Corresponding author.
*E-mail address:* rboer@cells.es (D.R. Boer).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.csbj.2021.12.041
2001-0370/© 2022 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

the recipient cell. The process is strictly regulated at various levels, which includes the implication of auxiliary proteins that form part of the relaxosome complex [5,22]. The auxiliary factors can be plasmid encoded or provided by the donor. In most cases, the auxiliary proteins are either indispensable or enhance conjugation frequency. The functions of the auxiliary proteins include specific DNA binding on the relaxosome, connecting the relaxosome with the T4SS channel, and interacting with the relaxase.

Recently, the relaxase of the *Bacillus subtilis* plasmid pLS20 was identified [23], which was named $Rel_{pLS20}$. Sequence comparison showed that the N-terminal domain of $Rel_{pLS20}$ contains sequence motifs corresponding to relaxase domains and was marked as the founding member of a new class of relaxases, named $MOB_L$ [23]. However, other authors suggest that this relaxase could also be classified as belonging to the $MOB_P$ subfamily, $MOB_{MG}$ [19,24,25]. In addition to the relaxase, two auxiliary proteins of plasmid pLS20 have been identified, which were named $Aux1_{pLS20}$ and $Aux2_{pLS20}$ [26]. They are encoded immediately upstream of $Rel_{pLS20}$ and were predicted to be Ribbon-Helix-Helix (RHH) proteins based on the occurrence of a β-strand followed by two α-helices at their N termini. RHH proteins are transcription factors that have a range of regulatory functions in prokaryotes and bacteriophages and have an interlaced, inseparable dimeric arrangement [27]. The well-studied *Salmonella* phage protein Arc is the founding member of this family.

It was shown that the pLS20 auxiliary RHH proteins are essential for pLS20 conjugation and that they bind to specific sites of the $oriT_{pLS20}$. $Aux1_{pLS20}$ binds to a 25 bp region, which contained an inverted repeat with sequence 5′-TGGTACCA-3′, which was proposed to be its binding site. $Aux2_{pLS20}$ is able to bind to a much larger region of the *oriT* spanning several hundreds of bases. To explain this, it was proposed that $Aux2_{pLS20}$ binds the TGTGCAT sequence which was the only sequence to be present three times in the *oriT* [26]. The oligomerization state of these proteins was studied by analytical ultracentrifugation (AUC) and was proposed to be tetrameric for $Aux1_{pLS20}$ and hexameric for $Aux2_{pLS20}$, although tetrameric forms were compatible with the sedimentation profile as well [26]. This study also used AUC to probe for interactions between the different relaxosome proteins, but none could be identified.

The exact function of the auxiliary proteins in the conjugative process of pLS20 is unknown. To gain knowledge on the function of these proteins, we characterize $Aux1_{pLS20}$ and $Aux2_{pLS20}$ using various biophysical techniques and have investigated the structures of these proteins. We present the low-resolution, SAXS envelope of the $Aux1_{pLS20}$ and the crystal structure of the C-terminal domain of $Aux2_{pLS20}$ at atomic resolution. In addition, we also study the interactions between the auxiliary proteins and the relaxase $Rel_{pLS20}$. Our results gave new insights into the relaxosome of pLS20, which has implications on similar G+ conjugative systems. Thus, we show that the C-terminal domain of $Aux2_{pLS20}$ forms a four-helix bundle and with structural homology to TraM of plasmid F, thereby providing the first evidence of the existence of a $TraM_{pF}$ analog in G+ conjugative systems. Unlike other $TraM_{pF}$ analogs, $Aux2_{pLS20}$ does not interact with the relaxase. Surprisingly, $Aux1_{pLS20}$ did show an interaction with the C-terminal domain of $Rel_{pLS20}$. We discuss the implications of the results for the mechanism of relaxosome preparation and presentation to the T4SS channel.

## 2. Methods

### 2.1. Protein expression and purification

All protein constructs were expressed and purified using standard protocols. The cDNAs expressing the protein constructs were $Aux1_{pLS20}$ without tags ($M_w$ 9.02 kDa), $Aux1_{pLS20}$ with C-terminal His tag ($M_w$ 10.6 kDa), $Aux2_{pLS20}$ ($M_w$ 17.03 kDa), full length construct ($M_w$ 48.75 kDa) and the C-terminal (residues 180–410, $M_w$ 28.27 kDa) and N-terminal domain (residues 1–232, $M_w$ 26.92 kDa) of $Rel_{pLS20}$. The pET28-based clone of $Aux1_{pLS20}$ with C-terminal His tag was obtained as described in [26]. All the other cDNAs were chemically synthesized, adding an N-terminal HRV-3C protease cleavage site before the methionine codon of each protein construct. The cDNA synthesis and subcloning into pHTP1 vectors were purchased from NZYTech (Campus do Lumiar, Lisboa, Portugal). *Escherichia coli* BL21 (DE3) were transformed with pHTP1 carrying the insert of the different constructs of $aux1_{pLS20}$, $aux2_{pLS20}$, $rel_{pLS20}$, $relC\text{-}ter_{pLS20}$ and $relN\text{-}ter_{pLS20}$ and were inoculated in fresh Luria-Bertani broth (LB) media complemented with 50 μg/ml kanamycin at 37 °C overnight. Then, the cells from the overnight culture were collected by centrifugation (4000 × g for 10 min) and suspended in expression media (typically 1 L of Terrific Broth (TB) with 50 μg/ml kanamycin), at a ratio of 15 ml of preculture per liter of medium. Cells in expression media were grown at 37 °C until an $OD_{600}$ = 0.8–1 was reached. After that, protein expression was induced overnight at 20 °C by addition of 1 mM isopropyl-β-D-1-thio galactopyranoside (IPTG, Omnipur). After overnight induction, cells were centrifuged at 4000×g for 30 min, and pellets were frozen and stored at −80 °C. Before purification, the pellets were thawed and resuspended in lysis buffer at a ratio of 5 ml/g of cells. The lysis buffer contained 0.5 M NaCl, 50 mM Tris pH 7.5, 5 mM imidazole, 1 mM EDTA, 5% glycerol, and 0.2 mM of PMSF (phenylmethylsulfonyl fluoride) as protease inhibitor. The cell suspension was lysed by sonication, adding DNase I to a final concentration of 200 μg/ml and lysozyme to a final concentration of 100 μg/ml during sonication. Insoluble matter was precipitated by centrifugation (18 000×g, 30 min), and the supernatant was filtered through a 0.22 μm filter and applied to a nickel-charged His-Trap™ HP chelating column 5 ml (GE Healthcare Life Sciences). The column was washed with 10 column volumes (50 ml) of binding buffer (20 mM Tris pH 8.0, 500 mM NaCl, 5 mM Imidazole) to elute unspecific bound proteins. Bound proteins were eluted using sequential steps of 4, 10 and 25% of elution buffer (0.5 M NaCl, 50 mM Tris pH 7.5, 0.5 M imidazole). Fractions were analyzed on SDS-PAGE and fractions containing pure protein were combined and concentrated over an Amicon ultra 10 kDa MWCO (Millipore). The buffer was exchanged using three HiTrap 5 ml Desalting columns (Cytiva) connected in series to 20 mM Tris pH 8.0, 500 mM NaCl. After elution, proteins were incubated overnight at 4 °C with 1 mg of HRV-3C protease (in-house preparation). After incubation, HRV-3C, tags and uncut protein were removed by application of the incubation solution through a His-Trap™ HP chelating column. The flow-through was concentrated and further purified by Size Exclusion Chromatography to remove aggregates and other contaminants. The SEC column (Generon ProSEC 16/60 3–70 HR) was equilibrated in 20 mM Tris pH 8.0, 500 mM NaCl. Fractions containing the target protein were concentrated over an Amicon ultra 10 kDa MWCO (Millipore). Typically, a yield of ∼20 mg of $Aux1_{pLS20}$, ∼10 mg for $Aux2_{pLS20}$, ∼100 mg for full length $Rel_{pLS20}$, ∼20 mg for $RelN\text{-}ter_{pLS20}$ and $RelC\text{-}ter_{pLS20}$ were obtained from 10 g of pellet. Purity was assessed to be >95% by SDS–PAGE, followed by BlueSafe staining (NZYTech, Campus do Lumiar, Lisboa, Portugal). The protein concentration was determined by nanodrop, using the theoretical extinction coefficient to calculate the concentrations, immediately prior to usage where possible. Surplus aliquots were stored at −80 °C until usage.

### 2.2. Analytical size exclusion chromatography on $Aux1_{pLS20}$, $Aux2_{pLS20}$ and $Rel_{pLS20}$ mixtures

To determine the elution volumes of the separate proteins, 1.2 nmol of either full length $Rel_{pLS20}$, $RelC\text{-}ter_{pLS20}$ and $RelN\text{-}ter_{pLS20}$ were used for all the interaction assays. For complex-

binding stoichiometry assays, molar ratios of 1:1, 5:1 and 10:1 were prepared for $Aux1_{pLS20}$:$Rel_{pLS20}$, $Aux1_{pLS20}$:RelN-$ter_{pLS20}$, $Aux1_{pLS20}$:RelC-$ter_{pLS20}$ and $Aux2_{pLS20}$:$Aux1_{pLS20}$. All samples were incubated for 30 min on ice before injection. Thus, the maximum amount of $Aux1_{pLS20}$ and $Aux2_{pLS20}$ injected was 12 nmol in the 1:10 M ratio. Samples of the proteins alone and in complex were prepared in a final volume of 30 μL in 20 mM Tris pH 8, 300 mM NaCl and 25 μL were injected on a Superdex 200 increase 5/150 GL equilibrated with 20 mM Tris pH 8, 300 mM NaCl. The elution was run at a flow rate of 0.3 ml/min and the absorption of the elute was monitored at 260 and 280 nm. To estimate the molecular weights ($M_W$) of the homo- and heterocomplexes, a calibration of the Superdex 200 increase 5/150 GL column was performed using proteins with known $M_W$ (*i.e.* BLC (239.64 KDa), RCO (81.28 KDa), BSA (69.29 KDa), MpARF3 (49.52 KDa), MpARF2 (44.63 KDa), AtARF5 (44.29 KDa), AtARF1 (41.13 KDa) and p69 (24.1 KDa), which were eluted using the same elution buffer used above. The derived relation between the elution volume ($V_{el}$) and $M_W$ was $V_{el} = -0.6815 \cdot \log(M_w) + 5.1906$, with an $R^2 = 0.933$.

### 2.3. Isothermal titration calorimetry (ITC)

Calorimetric measurements were carried out using a VP-ITC instrument from MicroCal Inc. (Northampton, USA), in the Polymorphism and Calorimetry Unit of the Scientific and Technological Centers of the University of Barcelona. In the experiment, 1.4 ml of $Rel_{pLS20}$ solution at 87.2 μM (4.25 mg/ml) were titrated with 300 μL of $Aux1_{pLS20}$ at 2770 μM (25 mg/ml) at 298 K. The reference cell was filled with double deionized water ($ddH_2O$). All measurements were carried out in 20 mM Tris-HCl pH 8, 300 mM NaCl. The buffer solution and $ddH_2O$ were degassed at room temperature with stirring under vacuum for ≥30 min. Upon experimental setup, the $Rel_{pLS20}$ solution present in the sample cell was stirred at 300 rpm. The titration was initiated at 298 K after a stable baseline was achieved, with an initial injection of 2 μL of $Aux1_{pLS20}$ during 4 s. The initial injection was followed by 29 injections of 10 μL during 20 s, each spaced by 300 s.

The calorimetric signal was integrated to obtain the enthalpy changes caused by complexation of $Rel_{pLS20}$ with $Aux1_{pLS20}$. Data were analyzed using the software Origin 7.0 (January 2004) and fitted to one single-site binding model, subtracting the average heat of the last three measurements after saturation of the $Rel_{LS20}$ binding sites to correct experimental heats of the $Rel_{pLS20}$-$Aux1_{pLS20}$ dilution. The enthalpy ($\Delta H = -5722 \pm 120.4\ kJ\ mol^{-1}$), entropy ($\Delta S = 1.54\ kJ\ mol^{-1}$), binding constant ($K_b = 3.40 * 10^4 \pm 2.64 * 10^3\ mol^{-1}$) and number of binding sites ($N = 2.55 \pm 0.0375$) resulted from nonlinear least square data fitting. Dissociation constant ($K_d = 29.41\ \mu M$) was calculated as the inverse of $K_b$ and the standard Gibbs free energy value ($\Delta G^0 = -6180.92\ kJ\ mol^{-1}$) was calculated using the equation:

$$\Delta G^0 = \Delta H^0 - T\Delta S^0$$

where T = 298 K and $\Delta H^0$ and $\Delta S^0$ are the thermodynamic values resulting from experimental data fitting at the same temperature. The *c* value of the assay (*c* = 7.56) was calculated as $c = n * P_t * K_b$, indicating a good shape of the binding isotherm to calculate accurate thermodynamic parameters from the calorimetry data [28].

### 2.4. SAXS experiments on $Aux1_{pLS20}$

SAXS experiments have been performed on the NCD-SWEET beamline at the synchrotron ALBA (Cerdanyola del Vallès, Barcelona, Spain) at 12.4 keV on $Aux1_{pLS20}$ with C-terminal His tag. The data were collected on a Pilatus 1 M detector (with a pixel size of 172.0 × 172.0 $\mu m^2$). The distance sample/detector was 2556.00 mm. 40 images were collected for each concentration (1.0 mg/ml, 2.5 mg/ml, 5.0 mg/ml, 7.5 mg/ml and 10 mg/ml) with an exposure time of 0.1 s. The q-axis calibration was obtained by measuring silver behenate [29]. The program pyFAI [30] was used to integrate the 2D data into 1D data. The 1D data has been averaged, subtracted, normalized by the concentration, extrapolated to zero concentration and merged with Primus [31] from the ATSAS package. The radius of gyration $R_g$ and the maximum distance $D_{max}$ has been determined with GNOM [32]. The low-resolution envelop has been restored with DAMMIF [33].

### 2.5. $Aux2_{pLS20}$ crystallization and structure solution

Purified $Aux2_{pLS20}$ was concentrated to a final protein concentration of 10 mg/ml in 20 mM Tris (pH 8.0), 300 mM NaCl. Crystallization experiments of $Aux2_{pLS20}$ were performed using the sitting-drop vapor-diffusion method at 18 °C, by equilibration of drops of 1 μL protein + 1 μL crystallization buffer against 100 μL of the crystallization buffer. Crystals were harvested from two different crystallization conditions. Cryo-cooling in liquid nitrogen was performed by soaking crystals on a cryo-protecting solution consisting in reservoir buffer complemented with 10% glycerol, followed by direct plunge-freezing in liquid nitrogen. Data was collected on two crystal systems on 21/09/2021 at the BL13-XALOC beamline of the ALBA synchrotron Light Source [34,35], using MXCuBE software for data collection [36]. Crystals of one system belonged to the space group $P4_12$ and were grown using 0.2 M Magnesium chloride hexahydrate and 20% w/v Polyethylene glycol 3,350 as crystallization buffer. Data for this system was processed to a resolution of 1.76 Å, with two protein monomers in the asymmetric unit. We will refer to this crystal structure as Aux2-S. Crystals of the other system belonged to space group $P4_32_12$ and were obtained using 0.06 M D-Glucose, 0.06 M D-Mannose, 0.06 M D-Galactose, 0.06 M L-Fucose, 0.06 M D-Xylose, 0.06 M N-Acetyl-D-Glucosamine, 0.06 M Tris-Bicine pH 8.5, 20% v/v Ethylene glycol, 10 % w/v PEG 8000 as crystallization buffer. The data of this system was processed to a resolution of 1.89 Å with two protein tetramers in the asymmetric unit. We will refer to this crystal structure as Aux2-L. Data were processed with Autoproc v1.0.5 [37–42]. See Table 1 for further statistics.

To solve the structure of $Aux2_{pLS20}$, a search for four alpha helices was performed using the program Arcimboldo lite [43]. The size of the helices was adjusted according to the secondary structure predictions computed with PSIPRED [44]. The initial model was built by auto-tracing in the output map with Buccaneer [45]. The structure was completed through alternate manual model building with Coot v0.8.9 [46] and refinement with PHENIX v1.9.2–4158 [47]. The model was validated and further adjusted and refined using MolProbity [48]. The crystallographic and refinement parameters are given in Table 1. Figures were prepared using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.)

## 3. Results

The relaxosome of pLS20 was shown to involve two RHH proteins [26], in addition to the recently identified relaxase, $Rel_{pLS20}$ and the *oriT* region [23]. To gain more insight into the role of these proteins in the relaxosome formation and their structures, we have performed biophysical and structural studies of the relaxosome proteins.

As a first step, we estimated the molecular weight ($M_W$) of the pLS20 relaxosome proteins in solution using Size Exclusion Chromatography (SEC). The elution volumes of the SEC profiles of the

**Table 1**
Summary of the data processing and refinement statistics of the crystallographic analysis of the $Aux2_{pLS20}$ structures.

| | Aux2-S | Aux2-L |
|---|---|---|
| Data collection | | |
| Beamline | XALOC (ALBA) | XALOC (ALBA) |
| PDB code | 7NUV | 7QNQ |
| λ (Å) | 0.9793 | 0.9793 |
| Space group | *P* 4 $2_1$ 2 | *P* $4_3$ $2_1$ 2 |
| Unit cell parameters (Å) | *a* = 44.840, *b* = 44.840, *c* = 141.153 | *a* = 37.854, *b* = 137.854, *c* = 121.709 |
| Resolution range (Å) [a] | 47.05–1.759 (1.874 – 1.759) | 97.48–1.892 (1.964 – 1.892) |
| # of reflections: | | |
| total | 147,161 (5833) | 2,316,344 (116090) |
| unique | 12,790 (641) | 87,560 (4373) |
| Completeness ellipsoidal (%) | 90.2 (39.8) | 96.1 (57.4) |
| <I / σ(I)> | 10.0 | 14.3 (1.5) |
| Average multiplicity | 11.5 | 26.5 (26.5) |
| $R_{merge}$ (%) [b] | 36.5 | 0.19 (3.15) |
| $R_{meas}$ (%) [c] | 38.3 | 0.20 (3.1) |
| CC(1/2) (%) | 96.8 | 0.998 (0.522) |
| Structure Refinement | | |
| $R_{cryst}$ [d] /$R_{free}$ [e] (%) | 20.61 / 24.17 | 18.0 / 20.7 |
| r.m.s.deviation from target values: | | |
| Bond lengths (Å) | 0.007 | 0.005 |
| Bond angle distances (Å) | 0.912 | 0.662 |
| Molprobity scores: | | |
| Clashscore (‰) | 1.92 | 0.82 |
| Poor rotamers (%) | 0.85 | 0.36 |
| Ramachandran Outliers (%) | 0.00 | 0 |
| Ramachandran Favoured (%) | 99.21 | 100 |
| Overall score (Å) | 0.96 | 0.75 |
| Isotropic B factor analysis | | |
| Average model B-factors (Å$^2$) | 27.0 | 44.16 |
| B-factor from Wilson plot (Å$^2$) | 18.0 | 48.00 |

[a] Throughout the table, the values in parentheses are for the outermost resolution shell.
[b] $R_{merge} = \Sigma_h | \hat{I}_h - I_{h,i} | / \Sigma_h \Sigma_i I_{h,i}$, where $\hat{I}_h = (1/n_h) \Sigma_i I_{h,i}$ and $n_h$ is the number of times a reflection is measured.
[c] $R_{meas} = [\Sigma_h (n_h/[n_h-1])^{½} \Sigma_i | \hat{I}_h - I_{h,i} |] / \Sigma_h \Sigma_i I_{h,i}$, where $\hat{I}_h = (1/n_h) \Sigma_i I_{h,i}$ and $n_h$ is the number of times a reflection is measured.
[d] $R_{cryst} = \Sigma_{hkl} | |F_{obs}| - k |F_{calc}| | / \Sigma_{hkl} |F_{obs}|$
[e] $R_{free} = \Sigma_{hkl \subset T} | |F_{obs}| - k |F_{calc}| | / \Sigma_{hkl \subset T} |F_{obs}|$ where T represents a test set comprising ~ 5% of all reflections excluded during refinement.

individual purified proteins (Suppl. Fig. 1) were used to calculate the $M_W$ based on a calibration of the columns using standard proteins with known $M_W$. The results of the $M_W$ estimation are shown in Table 2. The $M_W$ of the full length relaxase ($Rel_{pLS20}$), its N-terminal domain (RelN-$ter_{pLS20}$) and its C-terminal domain (RelC-$ter_{pLS20}$), respectively, correspond well to the theoretical values of a monomer, indicating that the protein is globular and monomeric in solution in the absence of other interaction partners. However, the estimated $M_W$ of $Aux1_{pLS20}$ and $Aux2_{pLS20}$ SEC both correspond to a pentamer and therefore deviate from the expected values based on the fact they are RHH proteins.

We then set out to map the interactions between respective pairs of the three pLS20 encoded proteins that have been shown to be essential for conjugation: $Aux1_{pLS20}$, $Aux2_{pLS20}$, and $Rel_{pLS20}$ [23,26], performing SEC analysis with mixtures of $Aux1_{pLS20}$, $Aux2_{pLS20}$, full length $Rel_{pLS20}$, and the N- and C- terminal domains of $Rel_{pLS20}$. We found that $Aux1_{pLS20}$ interacts with the full-length $Rel_{pLS20}$ and RelC-$ter_{pLS20}$ (Fig. 1 A-C), as evidenced by the appearance of new peaks with a $V_{el}$ of 1.97 ml for the $Aux1_{pLS20}$/$Rel_{pLS20}$ mixture and with a $V_{el}$ of 2.00 ml for the $Aux1_{pLS20}$/RelC-$ter_{pLS20}$ mixture. In contrast, no additional peak appeared in mixtures of $Aux1_{pLS20}$ and RelN-$ter_{pLS20}$ (Suppl. Fig. 2). No other interactions were found between the different constructs of these three proteins (Fig. 1C and Suppl. Fig. 2). To confirm the $Aux1_{pLS20}$/$Rel_{pLS20}$ interaction and to determine its strength, we measured the thermodynamic parameters of the interaction in solution using isothermal titration calorimetry (ITC, Fig. 1D). The $K_d$ determined by the ITC measurements was 29.4 μM. The ITC results indicate that the interaction was favorable in terms of enthalpy and entropy, with a calculated $\Delta G^0$ of −6180.92 kJ mol$^{-1}$, mainly driven by the exothermic component of the binding, with a $\Delta H^0$ of −5722 ± 120.4 kJ mol$^{-1}$ and a $\Delta S^0$ of 1.54 kJ mol$^{-1}$. Both parameters suggest that the interaction between $Aux1_{pLS20}$ and $Rel_{pLS20}$ is strongly affected by hydrogen bond formation and van der Waals interactions [49]. The entropy component, although small, may be related to the burial of water-accessible surface area upon binding, resulting in the release of interfacial water molecules to the solvent, contributing favorably to the total entropy of interaction [50].

### 3.1. SAXS on $Aux1_{pLS20}$

$Aux1_{pLS20}$ is an RHH protein and is predicted to consist of 1 β-strand followed by three α-helices (Fig. 2A). The occurrence of the first sheet and two consecutive α-helices is consistent with the RHH motif. However, the third α-helix is an addition to RHH proteins, as it is predicted to span 50 amino acids. To probe the effect of this helix on the overall structure of $Aux1_{pLS20}$, we performed SAXS measurements at different concentrations (Suppl Fig. 3). The data are consistent with a $M_W$ of 38.9 kDa, which corresponds well with the expected $M_W$ of 44.0 kDa for a tetramer, and is in accordance with AUC data [26]. The radius of gyration ($R_g$) was determined to be 34 Å, the maximum dimension ($D_{max}$) as 130 Å and the Porod volume ($V_{Porod}$) as 89 Å$^3$. Reconstruction of the envelop using DAMMIN resulted in an elongated overall shape (Fig. 2B). Although the overall shape is overestimated by the SAXS analysis, it is consistent with the tetramer structure of the Arc RHH structure bound to DNA (PDB code 1PAR, see Fig. 2C) and the tetrameric configuration determined previously using analytical ultracentrifugation [26]. The $R_g$, $R_{max}$ and $V_{Porod}$ are also slightly overestimated, which is likely related with the formation of higher oligomerization states, which has also been observed in the SEC analysis, where a second species with lower elution volume appears (Suppl. Fig. 1). There is additional density in the SAXS volume compared to Arc, indicating that $Aux1_{pLS20}$ forms a triangle instead of a tubular structure. The tip of this triangle is likely occupied by the additional helical residues present in the $Aux1_{pLS20}$ sequence but not in Arc [51].

### 3.2. Crystal structures of the tetramerization domain of $Aux2_{pLS20}$

The function of $Aux2_{pLS20}$ has so far not been determined and cannot be easily inferred from the sequence. A high-resolution structure could provide insights into its function by structural homology to known proteins and we therefore set out to determine its structure. We were able to determine two crystal structures, named Aux2-L and Aux2-S, both containing essentially the same structure of a C-terminal fragment of $Aux2_{pLS20}$ (Fig. 3), lacking the RHH domains. Since we attempted the crystallization of the full-length protein, these fragments occurred through *in situ* degradation, indicating flexibility between the RHH and C-terminal domains. Both structures show a tetrameric oligomerization state, where each protein chain folds into two consecutive alpha helices. The N-terminal helix consists of 12 helical turns and the C-terminal helix consists of 3 turns. The short and long alpha helices are con-

**Table 2**
Estimated molecular weights ($M_W$) of the relaxosome proteins of pLS20.

| Protein | $V_{el}$ (ml) | $M_{w,\ monomer}$ (kDa) | Estimated $M_w$ based on $V_{el}$ (kDa) | Calculated oligomerization state | Predicted oligomerization state |
|---|---|---|---|---|---|
| $Aux1_{pLS20}$ | 2.05 | 9.02 | 46.94 | Pentamer (5.20) | Tetramer |
| $Aux2_{pLS20}$ | 1.82 | 17.03 | 85.78 | Pentamer (5.03) | Tetramer |
| $Rel_{pLS20}$N-ter | 2.26 | 26.92 | 27.06 | Monomer (1.01) | Monomer |
| $Rel_{pLS20}$C-ter | 2.21 | 28.27 | 30.85 | Monomer (1.07) | Monomer |
| $Rel_{pLS20}$ | 2.10 | 48.75 | 41.17 | Monomer (0.84) | Monomer |

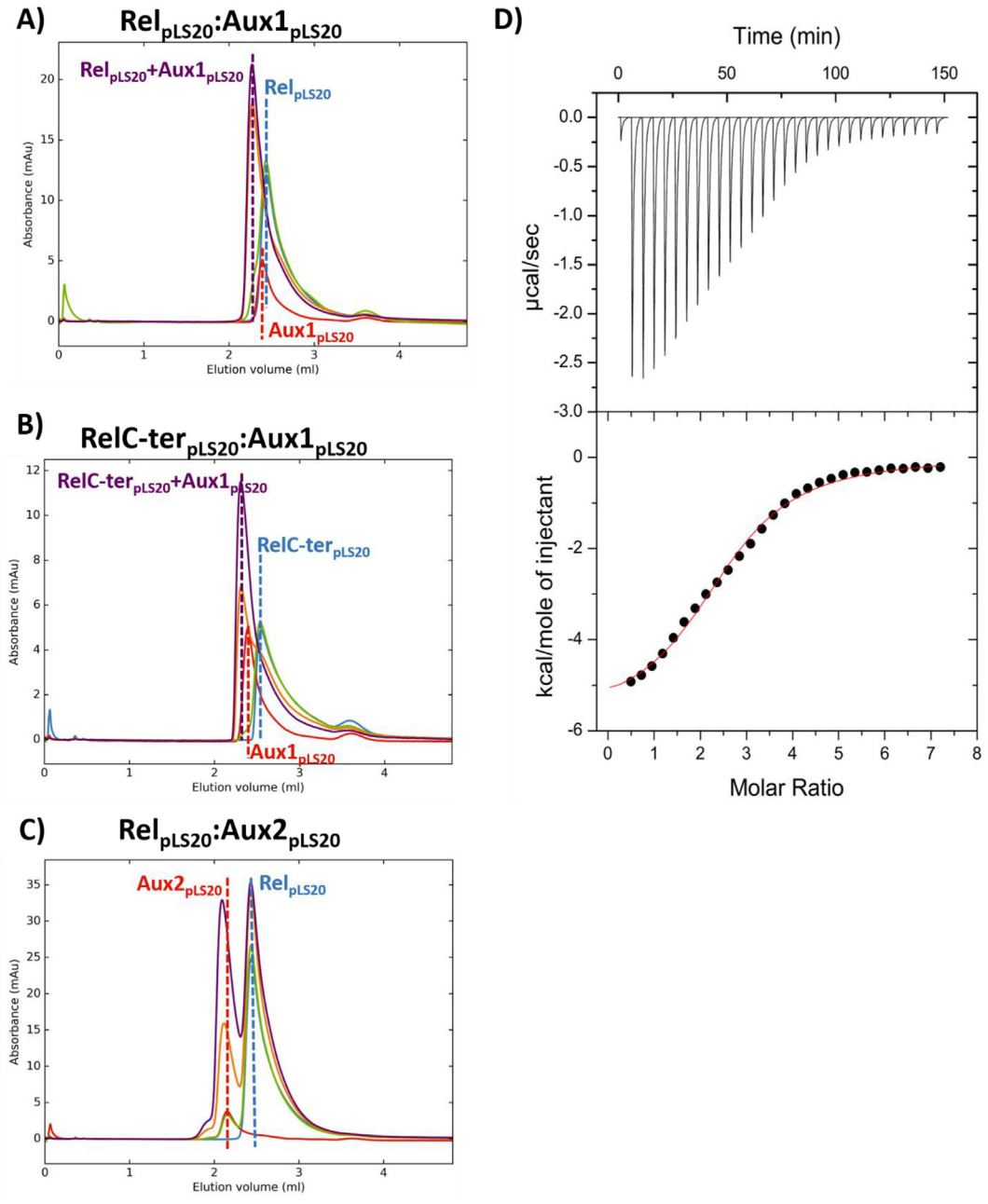


**Fig. 1.** *In vitro* interactions between relaxosome proteins. A) The SEC elution profile of $Rel_{pLS20}$:$Aux1_{pLS20}$ mixtures at different stoichiometries. B) the SEC elution profile of RelC-ter$_{pLS20}$:$Aux1_{pLS20}$ mixtures at different stoichiometries. C) The SEC elution profile of $Rel_{pLS20}$:$Aux2_{pLS20}$ mixtures at different stoichiometries. In panels A-C, the green, orange and purple lines represent the profile of mixtures at 1:1, 5:1 and 10:1 M ratio, respectively, the blue line in A-C marks the peak of the corresponding $Rel_{pLS20}$ domain, the red line in all plots represents the $Aux1_{pLS20}$ or $Aux2_{pLS20}$ and the purple line marks the peak of the complex, D) Thermogram of an $Aux1_{pLS20}$:$Rel_{pLS20}$ titration as determined by isothermal titration calorimetry. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

nected by a loop of 16 amino acids (A98-E113). The first three N-terminal turns of the long alpha helices are kinked with respect the rest of the α-helix. The long α-helices form a coiled coil structure, interacting with the long α-helices of the other chains through extensive hydrophobic packing interactions, thereby forming tetramers. This creates a tubular structure with a narrow channel at the center. The short alpha helices act as a lock to stabilize the structure. They are oriented in a parallel fashion to the

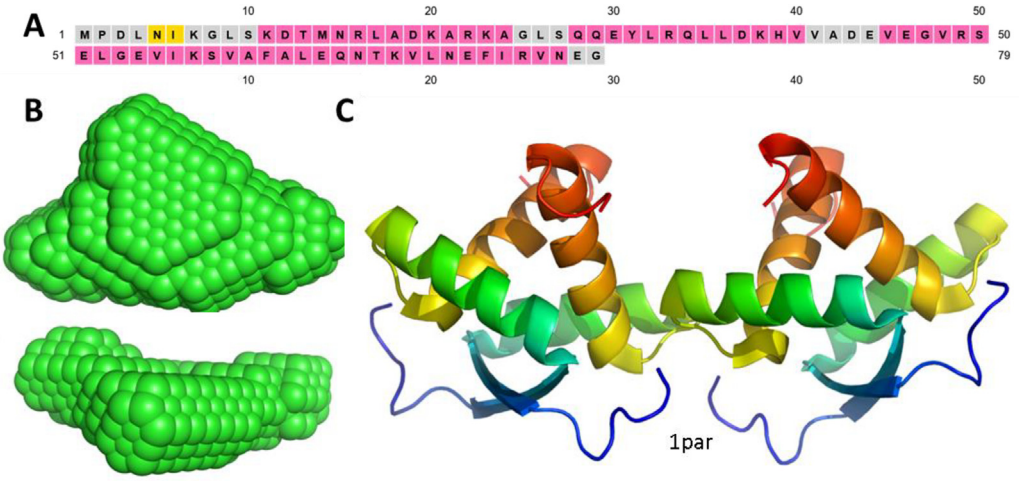


**Fig 2.** SAXS analysis of Aux1$_{pLS20}$. A) Sequence of Aux1$_{pLS20}$, showing the residues predicted as forming a β-strand in yellow and residues predicted to form α-helices in pink. B) Side (top panel) and top (bottom panel) view of the volume reconstruction based on the Aux1$_{pLS20}$ SAXS data. C) Cartoon representation of the Arc tetramer in complex with the DNA (not shown) observed in the structure of Arc (PDB code 1PAR). The residues are rainbow colored, blue for the N-terminus changing to red at the C-terminus in a gradual fashion. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

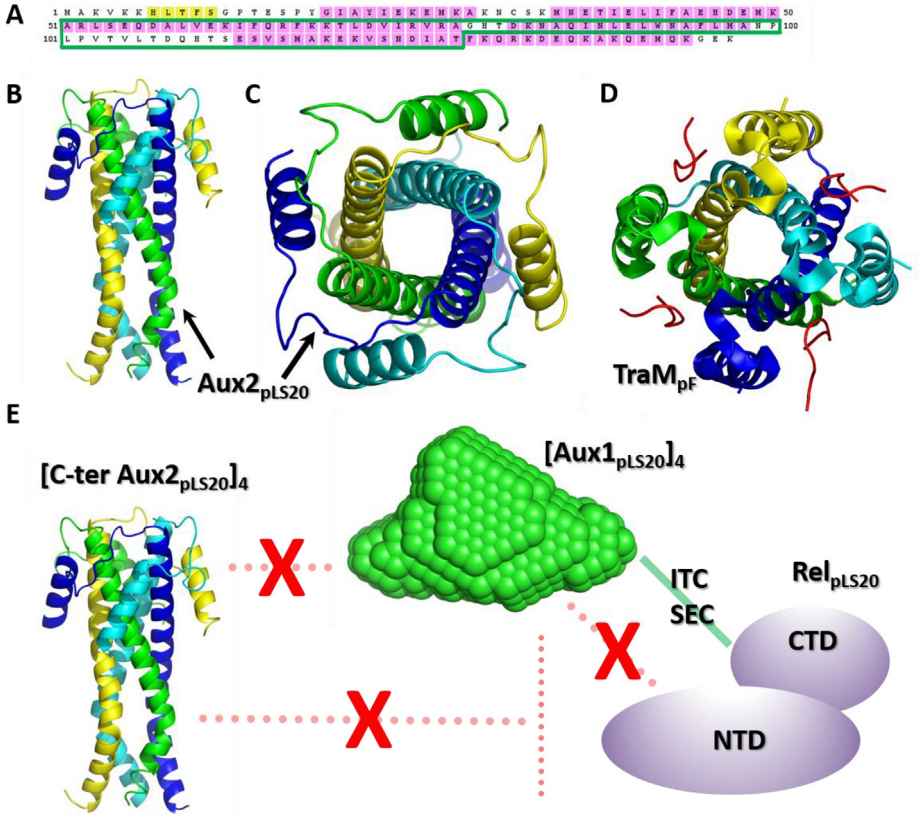


**Fig. 3.** Crystallographic structure of the tetramerization domain of Aux2$_{pLS20}$. A) Sequence of the full-length protein, predicted β-strands are colored yellow and α-helices pink. The green box indicates the fragment crystallized. B) Cartoon representation of the Aux2$_{pLS20}$ structure, showing the four monomers in different colors. The N-terminal helices are the inner and longer helices, the C-terminal helices are the external, shorter helices shown at the top. C) Top view of the Aux2$_{pLS20}$ structure along the α-helices, coloring as in B). D) Structure of the C-terminal domain of the C-terminal domain of plasmid F TraM (PDB code 3D8A). The TraM$_{pF}$ chains are colored in the same order as the Aux2$_{pLS20}$ inner helices whereas the peptide molecules of the coupling protein TraD$_{pF}$ are colored red. Note the relative shift of the C-terminal helices with respect to the inner helices compared to the Aux2$_{pLS20}$ structure shown in panel C. E) Graphical summary of the results described in this article. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

long helices, but interact with the long alpha helices of chains protomers n + 2 and n + 3, thanks to a long loop that connects both helices (Fig. 3C).

Interestingly, the structure is similar to the C-terminal domain of TraM$_{pF}$ protein from plasmid F (PDB codes 3D8A and 2G7O) and that of pED208 (PDB code 3ON0) as shown in Fig. 3D [52–

54]. TraM also consists of two α-helices that fold into a similar tetrameric stalk structure, consisting of a central α-helix followed by a second antiparallel helix that laterally packs against the central helices. The latter structure reveals the complex of the full-length protein bound to DNA. Comparison of the $Aux2_{pLS20}$ four-helix bundle with that of the 3D8A and 3ON0 structures shows that the tetramerization domain in $Aux2_{pLS20}$ is considerably longer than in the other analyzed structures, suggesting that the N-terminal turns of the long alpha helices in $Aux2_{pLS20}$ may form part of the DNA binding domain. This region is not included in the crystal structures of the $TraM_{pF}$ in absence of DNA, indicating that the density is not well defined in that region and that these residues are disordered. The structure of the DNA complex of $TraM_{pF}$ shows that this region becomes ordered upon binding of DNA. Therefore, the DNA binding may induce a conformational change in this region that would displace the N-terminal part of the long alpha helix, allowing the RHH domain to change position to adjust to the configuration of the DNA binding sites in the *oriT*.

When comparing the structure of the tetramerization domain of $Aux2_{pLS20}$ with that of $TraM_{pF}$ (PDB codes 3D8A and 3ON0), the most pronounced difference is in the position of the C-terminal helices with respect to the long helix. The loop connecting the small and long helices is shorter in $TraM_{pF}$. As a result, the interactions between the C-terminal and N-terminal helices in the four-helix bundle occur between adjacent monomers, in contrast to $Aux2_{pLS20}$, where the short helix interacts with the n + 2 monomer. It should be noted that the positions of the shorter helices are equivalent with those observed in the $TraM_{pF}$ structure despite the swap in position.

$TraM_{pF}$ interacts with the plasmid F coupling protein, TraD, and the structure of the complex has been determined (PDB code 3D8A) [53]. The $TraD_{pF}$ peptide binds to the loop region that connects the N- and C-terminal helices (Fig. 3D). The helices in the 3ON0 structure $TraM_{pF}$, not bound to the peptide, show a similar layout to the structure of the $TraD_{pF}$/$TraM_{pF}$ complex, with only minor adjustments of the residues in the loop between the helices. In contrast, in $Aux2_{pLS20}$, the loop between the helices of $Aux2_{pLS20}$ traverses the region corresponding to the binding site of $TraD_{pF}$, which is due to the swap of the position of the short C-terminal helix described above.

## 4. Discussion

In this paper, we report the low-resolution structure of $Aux1_{pLS20}$ and two high-resolution structures of the tetramerization domain of $Aux2_{pLS20}$. In addition, we analyze the possible interactions between the auxiliary proteins and the relaxase using ITC and SEC, which showed that $Aux1_{pLS20}$ and $Rel_{pLS20}$ interact, but the other proteins do not.

Our results indicate that the RHH protein $Aux1_{pLS20}$ tetramerizes and forms an elongated structure, reminiscent of the RHH proteins, of which the transcriptional repressor Arc from the *Salmonella* phage P22 is a founding member [27]. The sequential length, function and structural features of $Aux1_{pLS20}$ are similar to protein TraY of IncF plasmids: both TraY and $Aux1_{pLS20}$ are RHH proteins and are able to bind to inverted sequences on their respective *oriTs* [55,56]. Our SAXS data suggest that $Aux1_{pLS20}$ is a tetramer like $TraY_{pF}$, which is in agreement with the AUC data previously published [26]. From the SEC experiments, however, a pentameric configuration was derived. It should be noted that the elution volume of the proteins in SEC can vary according to their shape. The discrepancies of the $M_W$ derived from SEC data compared to SAXS and AUC can be explained by the elongated, non-globular shape of $Aux1_{pLS20}$. Given that RHH domains are intrinsically dimeric in nature, it follows that the oligomerization state of $Aux1_{pLS20}$ should be a multiple of dimers. Our combined data and previously published data show that the tetrameric arrangement is most likely.

The data obtained for the relaxase suggests that the protein is monomeric in solution, which is consistent with previous dynamic light scattering and AUC data [23]. Many relaxases are found in a monomeric state, both in solution and in the crystal and single particle electron microscopy structures, which is particularly true for G+ relaxases [21,57].

The structures of the C-terminal domain of $Aux2_{pLS20}$ show that it consists of two alpha helices connected by a short loop (Fig. 3). The N-terminal helix is kinked and forms a four-helix bundle, combining four monomers through extensive and tightly packed hydrophobic interactions. As a result, the structure of $Aux2_{pLS20}$ shows a tetrameric oligomerization state. However, the SEC data indicate a pentamer. As stated above, RHH domains oligomerize as multiples of dimers, and a pentameric structure for the $Aux2_{pLS20}$ is highly unlikely given that the protein contains dimeric RHH domain. Interestingly, the oligomerization of $Aux2_{pLS20}$ was previously determined as a hexamer using AUC data [26]. From the structures, it is hard to envisage how the C-terminal stalk domain could reorganize into a hexameric structure, since the four-helix bundle is held together by a tightly packed hydrophobic core. This packing would likely be lost when adding an additional two helices to form a hexamer. However, we cannot exclude the possibility of the formation of a hexamer, and additional work will be required to determine the oligomerization states of this protein at different stages of the conjugation process.

The domain architecture and the C-terminal structure of $Aux2_{pLS20}$ are reminiscent of the TraM protein of plasmid F and pED208 [52–54,58]. Thus, $Aux2_{pLS20}$ is a structural analog of the $TraM_{IncF}$ proteins. This type of protein is a common factor in gram negative relaxosomes of plasmids from different incompatibility groups [5,59]. In plasmid F and other systems, one of the functions of this protein is to interact with the coupling protein through the C-terminal domain and the relaxase, functioning as a connector protein. Thus, the structural homology of the $Aux2_{pLS20}$ C-terminal domain with that of $TraM_{pF}$ suggests that one of the functions of $Aux2_{pLS20}$ is to interact with the coupling protein of pLS20.

Unlike $TraM_{pF}$, which was proposed to also interact with the relaxase $TraI_{pF}$ [60], no interactions could be detected between $Aux2_{pLS20}$ and $Rel_{pLS20}$. This is consistent with previous AUC data which did not detect any interaction between these proteins [26]. Thus, it seems that the interaction between the T4SS and the pLS20 relaxosome is not accomplished by a direct interaction between $Aux2_{pLS20}$ and the relaxase and it is likely that $Aux2_{pLS20}$ bridges the relaxosome and T4SS channel through interactions with DNA regions of the *oriT* [26].

The comparison of the structure of the tetramerization domain of $Aux2_{pLS20}$ with that of $TraM_{pF}$ also provides clues on how different connector proteins select their substrates. The swapped position of the two helices with respect to $TraM_{pF}$ and the concomitant differences in the position of the loop between the two helices, result in a substantially altered binding pocket of the peptide fragment of the coupling protein. As a consequence, the interaction between the coupling protein of pLS20 and $Aux2_{pLS20}$, if it exists, is likely different from that observed for plasmid F-like connector proteins. Future work should elucidate if $Aux2_{pLS20}$ interacts directly with the pLS20 coupling protein, and if the $TraM_{pF}$ binding pocket is preserved or is located in a different region of the protein.

We found that $Aux1_{pLS20}$ is able to interact with the C-terminal domain of $Rel_{pLS20}$ in SEC and in ITC. This is to the best of our knowledge unique among the relaxosomes studied so far. The interaction is consistent in that it occurs with the full length $Rel_{pLS20}$ as well as its C-terminal domain, but not the N-terminal

domain (Fig. 1 and Suppl Fig. 2). It should be noted, however, that this interaction was not observed in AUC data previously published [26]. An explanation of this apparent discrepancy may lie in the different experimental conditions used in both techniques and that the interaction is weak as determined by ITC measurements. It is likely that it can only be detected under favorable conditions, which may explain why AUC experiments failed to detect this interaction. Finally, it should be noted that the experiments described above were performed in absence of DNA, and it is expected that the presence of DNA sequences with affinity for these proteins may induce stronger binding and stabilization of the complexes. Future experiments are needed to investigate these questions further.

The findings in this paper are summarized graphically in Fig. 3E. The combined data presented here show that $Aux1_{pLS20}$ and $Aux2_{pLS20}$ likely bridge the various components of the relaxosome and mediate the interaction with the T4SS channel. We identify $Aux2_{pLS20}$ as a structural analog of $TraM_{pF}$, which suggests that $Aux2_{pLS20}$ links the relaxosome to the T4SS channel through interactions with the *oriT*. Furthermore, we show that $Aux1_{pLS20}$ interacts with $Rel_{pLS20}$, which suggests that it stabilizes the binding of the relaxase to the DNA sequence, as both that $Rel_{pLS20}$ and $Aux1_{pLS20}$ contain DNA binding domains. In addition, the interaction provides a possible function to the C-terminal domain of the relaxase.

## 5. Conclusions

We present the envelope of $Aux1_{pLS20}$ and two X-ray structures of $Aux2_{pLS20}$. The combined data presented here show that $Aux1_{pLS20}$ and $Aux2_{pLS20}$ are likely tetramers in solution. We identify $Aux2_{pLS20}$ as a $TraM_{pF}$ analog and show that $Aux1_{pLS20}$ interacts with the C-terminal domain of $Rel_{pLS20}$.

## Funding

This work was supported by Ministry of Economy and Competitiveness of the Spanish Government grants BIO2016-77883-C2-1-P and PID2019-108778GB-C21 to W.J.J.M, PID2020-117028 GB-I00 (AEI/FEDER, EU), BIO2016-77883-C2-2-P (AEI/FEDER, EU) and FIS2015-72574-EXP (AEI/FEDER, EU), which also supported N.B., to R.B. Funding for open access charge: Ministry of Economy and Competitiveness of the Spanish Government PID2020-117028 GB-I00. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author contributions

IC, NB and AC designed and performed the experiments, BMC, XC and FGO assisted in the technical preparation of the experiments. JVC, AMA, and WJJM provided pET28-based clones and cell culture material of some of the protein samples (or subdomains) used in the initial and/or final studies. MM analyzed the SAXS data, RB designed the research and wrote the paper.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgement

We thank Rafel Prohens of the polymorphism and calorimetry unit of the CCiT at the University of Barcelona. We acknowledge the assistance of the staff of the XALOC and NCD-SWEET beamlines, as well as the Floor Coordinators at the ALBA synchrotron.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.12.041.

## References

[1] Bellanger X, Payot S, Leblond-Bourget N, Guédon G. Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. FEMS Microbiol Rev 2014;38(4):720–60. https://doi.org/10.1111/1574-6976.12058.
[2] Grohmann E, Muth G, Espinosa M. Conjugative plasmid transfer in gram-positive bacteria. Microbiol Mol Biol Rev 2003;67(2):277–301.
[3] Goessweiner-Mohr N, Arends K, Keller W, Grohmann E. Conjugative type IV secretion systems in Gram-positive bacteria. Plasmid 2013;70(3):289–302. https://doi.org/10.1016/j.plasmid.2013.09.005.
[4] Grohmann E, Christie PJ, Waksman G, Backert S. Type IV secretion in Gram-negative and Gram-positive bacteria. Mol Microbiol 2018;107(4):455–71. https://doi.org/10.1111/mmi.13896.
[5] De La Cruz F, Frost LS, Meyer RJ, Zechner EL. Conjugative DNA metabolism in Gram-negative bacteria. FEMS Microbiol Rev 2010;34(1):18–40. https://doi.org/10.1111/j.1574-6976.2009.00195.x.
[6] Smillie C, Garcillán-Barcia MP, Francia MV, Rocha EPC, de la Cruz F. Mobility of Plasmids. Microbiol Mol Biol Rev 2010;74(3):434–52.
[7] Christie PJ, Whitaker N, González-Rivera C. Mechanism and structure of the bacterial type IV secretion systems. Biochim Biophys Acta - Mol. Cell Res 2014;1843(8):1578–91. https://doi.org/10.1016/j.bbamcr.2013.12.019.
[8] Low HH, Gubellini F, Rivera-Calzada A, Braun N, Connery S, Dujeancourt A, et al. Structure of a type IV secretion system. Nature 2014;508(7497):550–3. https://doi.org/10.1038/nature13081.
[9] Li YG, Hu Bo, Christie PJ, Sandkvist M, Cascales E, Christie PJ. Biological and structural diversity of type IV secretion systems. Microbiol Spectr 2019;7(2).
[10] Zechner EL, Lang S, Schildbach JF. Assembly and mechanisms of bacterial type IV secretion machines. Philos Trans R Soc B Biol Sci 2012;367(1592):1073–87. https://doi.org/10.1098/rstb.2011.0207.
[11] Christie PJ, Valero LG, Buchrieser C. Biological diversity and evolution of type IV secretion systems. Curr Top Microbiol Immunol 2017;413. https://doi.org/10.1007/978-3-319-75241-9_1.
[12] Wong JJW, Lu J, Glover JNM. Relaxosome function and conjugation regulation in F-like plasmids – a structural biology perspective. Mol Microbiol 2012;85:602–17. https://doi.org/10.1111/j.1365-2958.2012.08131.x.
[13] Christie PJ, Vogel JP. Bacterial type IV secretion: Conjugation systems adapted to deliver effector molecules to host cells. Trends Microbiol 2000;8(8):354–60. https://doi.org/10.1016/S0966-842X(00)01792-3.
[14] Nagai H, Kubori T. Type IVB secretion systems of *Legionella* and other gram-negative bacteria. Front Microbiol 2011;2. https://doi.org/10.3389/fmicb.2011.00136.
[15] Costa TRD, Harb L, Khara P, Zeng L, Hu Bo, Christie PJ. Type IV secretion systems: advances in structure, function, and activation. Mol Microbiol 2021;115(3):436–52. https://doi.org/10.1111/mmi.14670.
[16] Kohler V, Keller W, Grohmann E. Regulation of gram-positive conjugation. Front Microbiol 2019;10:1134. https://doi.org/10.3389/fmicb.2019.01134.
[17] Wawrzyniak P, Plucienniczak G, Bartosik D. The different faces of rolling-circle replication and its multifunctional initiator proteins. Front Microbiol 2017;8. https://doi.org/10.3389/fmicb.2017.02353.
[18] Guzmán-Herrador DL, Llosa M. The secret life of conjugative relaxases. Plasmid 2019;104:102415. https://doi.org/10.1016/j.plasmid.2019.102415.
[19] Garcillán-Barcia MP, Redondo-Salvo S, Vielva L, de la Cruz F. MOBscan: automated annotation of MOB relaxases. Methods Mol Biol 2020;2075. https://doi.org/10.1007/978-1-4939-9877-7_21.
[20] Ilangovan A, Kay CWM, Roier S, El Mkami H, Salvadori E, Zechner EL, et al. Cryo-EM structure of a relaxase reveals the molecular basis of DNA unwinding during bacterial conjugation. Cell 2017;169(4):708–721.e12. https://doi.org/10.1016/j.cell.2017.04.010.
[21] Edwards JS, Betts L, Frazier ML, Pollet RM, Kwong SM, Walton WG, et al. Molecular basis of antibiotic multiresistance transfer in *Staphylococcus aureus*. Proc Natl Acad Sci U S A 2013;110(8):2804–9. https://doi.org/10.1073/pnas.1219701110.
[22] Gomis-Rüth FX, Coll M. Cut and move: protein machinery for DNA processing in bacterial conjugation. Curr Opin Struct Biol 2006;16(6):744–52. https://doi.org/10.1016/j.sbi.2006.10.004.

[23] Ramachandran G, Miguel-Arribas A, Abia D, Singh PK, Crespo I, Gago-Córdoba C, et al. Discovery of a new family of relaxases in *Firmicutes* bacteria. PLOS Genet 2017;13(2):e1006586. https://doi.org/10.1371/journal.pgen.1006586.
[24] Zheng B, Tomita H, Inoue T, Ike Y. Isolation of VanB-type *Enterococcus faecalis* strains from nosocomial infections: First report of the isolation and identification of the pheromone-responsive plasmids pMG2200, encoding VanB-type vancomycin resistance and a Bac41-type bacteriocin, a. Antimicrob Agents Chemother 2009;53. https://doi.org/10.1128/AAC.00754-08.
[25] Tomita H, Ike Y. Genetic analysis of transfer-related regions of the vancomycin resistance *Enterococcus* conjugative plasmid pHTβ: Identification of oriT and a putative relaxase gene. J Bacteriol 2005;187(22):7727–37. https://doi.org/10.1128/JB.187.22.7727-7737.2005.
[26] Miguel-Arribas A, Hao J-A, Luque-Ortega JR, Ramachandran G, Val-Calvo J, Gago-Córdoba C, et al. The *Bacillus subtilis* conjugative plasmid pLS20 encodes two ribbon-helix-helix type auxiliary relaxosome proteins that are essential for conjugation. Front Microbiol 2017;8. https://doi.org/10.3389/fmicb.2017.02138.
[27] Schreiter ER, Drennan CL. Ribbon-helix-helix transcription factors: variations on a theme. Nat Rev Microbiol 2007;5(9):710–20. https://doi.org/10.1038/nrmicro1717.
[28] Srivastava VK, Yadav R. Isothermal titration calorimetry. Elsevier Inc.; 2019. doi: 10.1016/B978-0-12-816548-5.00009-5.
[29] Huang TC, Toraya H, Blanton TN, Wu Y. X-ray powder diffraction analysis of silver behenate, a possible low-angle diffraction standard. J Appl Crystallogr 1993;26(2):180–4. https://doi.org/10.1107/S0021889892009762.
[30] Kieffer J, Karkoulis D. PyFAI, a versatile library for azimuthal regrouping. J. Phys. Conf. Ser. 2013;425(20):202012. https://doi.org/10.1088/1742-6596/425/20/202012.
[31] Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J Appl Crystallogr 2003;36(5):1277–82. https://doi.org/10.1107/S0021889803012779.
[32] Svergun DI. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr 1992;25(4):495–503. https://doi.org/10.1107/S0021889892001663.
[33] Franke D, Svergun DI. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J Appl Crystallogr 2009;42(2):342–6. https://doi.org/10.1107/S0021889809000338.
[34] Juanhuix J, Gil-Ortiz F, Cuní G, Colldelram C, Nicolás J, Lidón J, et al. Developments in optics and performance at BL13-XALOC, the macromolecular crystallography beamline at the Alba Synchrotron. J Synchrotron Radiat 2014;21(4):679–89. https://doi.org/10.1107/S1600577514008825X10.1107/S160057751400825X/rv5017sup1.avi10.1107/S160057751400825X/rv5017sup2.pdf.
[35] Fauth F, Boer R, Gil-Ortiz F, Popescu C, Vallcorba O, Peral I, et al. The crystallography stations at the Alba synchrotron. Eur Phys J Plus 2015;130(8). https://doi.org/10.1140/epjp/i2015-15160-y.
[36] Oscarsson M, Beteva A, Flot D, Gordon E, Guijarro M, Leonard G, et al. MXCuBE2: the dawn of MXCuBE Collaboration. J Synchrotron Radiat 2019;26:393–405. https://doi.org/10.1107/s1600577519001267.
[37] Tickle IJ, Flensburg C, Keller P, Paciorek W, Sharff A, Vonrhein C, et al. STARANISO 2018.
[38] Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, et al. Data processing and analysis with the {\it autoPROC} toolbox. Acta Crystallogr Sect D 2011;67:293–302. https://doi.org/10.1107/S0907444911007773.
[39] Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr 2010;66(2):125–32.
[40] Evans P. Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr 2006;62(1):72–82.
[41] Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP 4 suite and current developments. Acta Crystallogr D Biol Crystallogr 2011;67(4):235–42.
[42] Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr 2013;69(7):1204–14.
[43] Sammito M, Millán C, Rodríguez DD, de Ilarduya IM, Meindl K, De Marino I, et al. Exploiting tertiary structure through local folds for crystallographic phasing. Nat Methods 2013;10(11):1099–101. https://doi.org/10.1038/nmeth.2644.
[44] Buchan DWA, Jones DT. The PSIPRED Protein Analysis Workbench: 20 years on. Nucleic Acids Res 2019;47. doi: 10.1093/nar/gkz297.
[45] Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr 2006;62(9):1002–11.
[46] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 2010;66(4):486–501. https://doi.org/10.1107/S0907444910007493.
[47] Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol 2019;75(10):861–77.
[48] Williams PA, Hughes CE, Lim GK, Kariuki BM, Harris KDM. Discovery of a new system exhibiting abundant polymorphism: M -aminobenzoic acid. Cryst Growth Des 2012;12(6):3104–13. https://doi.org/10.1021/cg3003178.
[49] Cramer J, Jiang X, Schönemann W, Silbermann M, Zihlmann P, Siegrist S, et al. Enhancing the enthalpic contribution of hydrogen bonds by solvent shielding. RSC Chem Biol 2020;1(4):281–7.
[50] Jelesarov I, Bosshard HR. Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. J Mol Recognit 1999;12:3–18. https://doi.org/10.1002/(SICI)1099-1352(199901/02)12:1<3::AID-JMR441>3.0.CO;2-6.
[51] Raumann BE, Rould MA, Pabo CO, Sauer RT. DNA recognition by β-sheets in the Arc represser-operator crystal structure. Nature 1994;367(6465):754–7. https://doi.org/10.1038/367754a0.
[52] Wong JJW, Lu J, Edwards RA, Frost LS, Glover JNM. Structural basis of cooperative DNA recognition by the plasmid conjugation factor, TraM. Nucl Acids Res 2011;39. doi: 10.1093/nar/gkr296.
[53] Lu J, Wong JJW, Edwards RA, Manchak J, Frost LS, Glover JNM. Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. Mol Microbiol 2008;70(1):89–99. https://doi.org/10.1111/j.1365-2958.2008.06391.x.
[54] Lu J, Edwards RA, Wong JJW, Manchak J, Scott PG, Frost LS, et al. Protonation-mediated structural flexibility in the F conjugation regulatory protein, TraM. EMBO J 2006;25(12):2930–9. https://doi.org/10.1038/sj.emboj.7601151.
[55] Frost LS, Ippen-Ihler K, Skurray RA. Analysis of the sequence and gene products of the transfer region of the F sex factor. Microbiol Rev 1994;58(2):162–210.
[56] Fernandez-Lopez R, de Toro M, Moncalian G, Garcillan-Barcia MP, de la Cruz F. Comparative genomics of the conjugation region of F-like plasmids: Five shades of F. Front Mol Biosci 2016;3. doi: 10.3389/fmolb.2016.00071.
[57] Pluta R, Boer DR, Lorenzo-Díaz F, Russi S, Gómez H, Fernández-López C, et al. Structural basis of a histidine-DNA nicking/joining mechanism for gene transfer and promiscuous spread of antibiotic resistance. Proc Natl Acad Sci USA 2017;114(32):E6526–35.
[58] Peng Y, Lu J, Wong JJW, Edwards RA, Frost LS, Mark Glover JN. Mechanistic basis of plasmid-specific DNA binding of the F plasmid regulatory protein, TraM. J Mol Biol 2014;426(22):3783–95. https://doi.org/10.1016/j.jmb.2014.09.018.
[59] Guglielmini J, Néron B, Abby SS, Garcillán-Barcia MP, La Cruz DF, Rocha EPC. Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. Nucl Acids Res 2014;42. doi: 10.1093/nar/gku194.
[60] Ragonese H, Haisch D, Villareal E, Choi J-H, Matson SW. The F plasmid-encoded TraM protein stimulates relaxosome-mediated cleavage at oriT through an interaction with TraI. Mol Microbiol 2007;63(4):1173–84.