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, Carretera de la Llum 2-26, Cerdanyola del Vallés, 08290 Barcelona, Spain
b Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), C. Nicolás Cabrera 1, Universidad Autónoma, Canto Blanco, 28049 Madrid, Spain

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 machin-ery 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...
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 occurrence of a -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 interleaded, 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. Aux1 pLS20 binds to a 25 bp region, which contained an inverted repeat with sequence 5′-TGTGAC-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 hundred bases. To explain this, it was proposed that Aux2 pLS20 binds the TGTGACAT sequence, which was the only sequence to be present in the three oriT regions. The oligomerization state of these proteins was studied by an 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 homolog in G+ conjugative systems. Unlike other TraM homologs, 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, Aux2 pLS20, and Rel pLS20 mixtures.

To determine the elution volumes of the separate proteins, 1.2 nmol of each full length Rel pLS20, RelC-ter pLS20 and RelN-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 Aux1pLS20:RelpLS20, Aux1pLS20:RelN-terpLS20, Aux1pLS20:RelC-terpLS20 and Aux2pLS20:Aux1pLS20. All samples were incubated for 30 min on ice before injection. Thus, the maximum amount of Aux1pLS20 and Aux2pLS20 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_e) and M_w was V_e = -0.6815 · 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 RelpLS20 solution at 87.2 μM (4.25 mg/ml) were titrated with 300 μL of Aux1pLS20 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 RelpLS20 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 Aux1pLS20 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 RelpLS20 with Aux1pLS20. 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–Aux1pLS20 complexes. Binding constant (K_b) and number of binding sites (N) were calculated using the Origin 7.0 software. The binding constant (K_b = 3.40 × 10^4 ± 2.64 × 10^3 mol⁻¹) and number of binding sites (N = 2.55 ± 0.0375) resulted from nonlinear least square data fitting. Dissociation constant (K_d = 29.41 μM) was calculated as the inverse of K_b. The standard Gibbs free energy (ΔG° = -6180.92 kJ mol⁻¹) was calculated using the equation: ΔG° = ΔH° - TΔS°

where T = 298 K and ΔH° and ΔS° 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_0 + S_0, indicating a good shape of the binding isotherm to calculate accurate thermodynamic parameters from the calorimetry data [28].

2.4. SAXS experiments on Aux1pLS20

SAXS experiments have been performed on the NCD-SWEET beamline at the synchrotron ALBA (Cerdanyola del Vallés, Barcelona, Spain) at 12.4 keV on Aux1pLS20 with C-terminal His tag. The data were collected on a Pilatus 1 M detector (with a pixel size of 172.0 × 172.0 μm²). 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 envelope has been restored with DAMMIF [33].

2.5. Aux2pLS20 crystallization and structure solution

Purified Aux2pLS20 was concentrated to a final protein concentration of 10 mg/ml in 20 mM Tris (pH 8.0), 300 mM NaCl. Crystallization experiments of Aux2pLS20 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_2_2_2 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-L. Crystals of the other system belonged to space group P4_1_2_2_2 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% w/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-S. Crystals of the other system belonged to space group P4_1_2_2_2 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% w/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 Aux2pLS20, 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, RelpLS20 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
individual purified proteins (Suppl. Fig. 1) were used to calculate the M_W based on a calibration of the columns using standard proteins. Both constructs of these three proteins (Fig. 1C and Suppl. Fig. 2) were found between the different constructs of these three proteins (Fig. 1C and Suppl. Fig. 2). To confirm the Aux1plLS20/RelpLS20 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 ΔG^0 of ~6180.92 kJ mol⁻¹, mainly driven by the exothermic component of the binding, with a ΔH^0 of ~5722 ± 120.4 kJ mol⁻¹ and a ΔS^0 of 1.54 kJ mol⁻¹. Both parameters suggest that the interaction between Aux1plLS20 and RelpLS20 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 Aux1plLS20

Aux1plLS20 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 Aux1plLS20, 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 Å³. 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 tetrameric 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 Aux1plLS20 forms a triangle instead of a tubular structure. The tip of this triangle is likely occupied by the additional helical residues present in the Aux1plLS20 sequence but not in Arc [51].

3.2. Crystal structures of the tetramerization domain of Aux2plLS20

The function of Aux2plLS20 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 Aux2plLS20 (Fig. 3), lacking the RHH domains. Since we attempted the crystallization of the full-length protein, these fragments occurred through in situ degragation, 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 cons-

Table 1
Summary of the data processing and refinement statistics of the crystallographic analysis of the Aux2plLS20 structures.

| Metric                         | Aux2-S          | Aux2-L          |
|--------------------------------|-----------------|-----------------|
| Data collection                |                 |                 |
| Beamline                      | XALOC (ALBA)    | XALOC (ALBA)    |
| PDB code                      | 7MUV            | 7QNO            |
| λ (Å)                         | 0.9793          | 0.9793          |
| Space group                    |                  |                 |
| P 4_2_2                        | P 4_2_2         |
| Unit cell parameters           |                 |                 |
| a = 44.840, b = 44.840, c = 141.153 | a = 37.854, b = 137.854, c = 121.709 |
| Resolution range (Å)           | 47.05–1.759     | 47.05–1.759     |
| # of reflections: total        | 147,161 (5383)  | 2,316,344 (16690) |
| # of reflections: unique      | 12,790 (641)    | 87,560 (4373)   |
| Completeness                   | 90.3 (92.8)     | 96.1 (57.4)     |
| ellipsoid (%)                  | 10.0            | 14.3 (1.5)      |
| Average multiplicity           | 11.5            | 26.5 (2.65)     |
| Rmerge (%)                     | 36.5            | 0.19 (3.15)     |
| Rfree (%)                      | 38.3            | 0.20 (3.1)      |
| CC(1/2) (%)                    | 96.8            | 0.998 (0.522)   |
| Structure Refinement           |                 |                 |
| Rmerge (%)                     | 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            | 0.00            |
| Ramachandran                   | 99.21           | 100             |
| Favoured (%)                   | 0.96            | 0.75            |
| Overall score (Å)              | 27.0            | 44.16           |
| Isotropic B factor analysis    |                 |                 |
| Average model B-factors (Å²)   | 18.0            | 48.00           |
| B-factor from Wilson plot (Å²) |                 |                 |

a Throughout the table, the values in parentheses are for the outermost resolution shell.
b Rmerge = Σ h k | Fobs|/ Σ h k | Fcalc|, where h k is the number of times a reflection is measured.
c Rfree = Σ [ h k | Fobs|/ Σ h k | Fcalc|] when h k is the number of times a reflection is measured.
d Rmerge = Σ h k | Fobs|/ | Fcalc|, where T represents a test set comprising ~5% of all reflections excluded during refinement.
e Rfree = Σ h k | Fobs|/ | Fcalc|, where T represents a test set comprising ~5% of all reflections excluded during refinement.

mixture and with a V_el of 2.00 ml for the Aux1plLS20/RelC-terplLS20 mixture. In contrast, no additional peak appeared in mixtures of Aux1plLS20 and RelN-terplLS20 (Suppl. Fig. 2). No other interactions were found between the different constructs of these three proteins (Fig. 1C and Suppl. Fig. 2).
connected by a loop of 16 amino acids (A98-E113). The first three N-terminal turns of the long alpha helices are kinked with respect to the rest of the \( \alpha \)-helix. The long \( \alpha \)-helices form a coiled coil structure, interacting with the long \( \alpha \)-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

Table 2
Estimated molecular weights (Mw) of the relaxosome proteins of pLS20.

| Protein | Vel (ml) | \( M_w \), monomer (kDa) | Estimated \( M_w \) based on Vel (kDa) | Calculated oligomerization state | Predicted oligomerization state |
|---------|----------|--------------------------|-------------------------------------|---------------------------------|---------------------------------|
| Aux1pLS20 | 2.05     | 9.02                     | 46.94                               | Pentamer (5.20)                 | Tetramer                        |
| Aux2pLS20 | 1.82     | 17.03                    | 85.78                               | Pentamer (5.03)                 | Tetramer                        |
| RelpLS20,N-ter | 2.26 | 26.92 | 27.06 | Monomer (1.01) | Monomer                        |
| RelpLS20,C-ter | 2.21 | 28.27 | 30.85 | Monomer (1.07) | Monomer                        |
| RelpLS20 | 2.10     | 48.75                    | 41.17                               | Monomer (0.84)                 | Monomer                        |

![Fig. 1. In vitro interactions between relaxosome proteins. A) The SEC elution profile of RelpLS20:Aux1pLS20 mixtures at different stoichiometries. B) the SEC elution profile of RelpLS20,N-ter:Aux1pLS20 mixtures at different stoichiometries. C) The SEC elution profile of RelpLS20:Aux2pLS20 mixtures at different stoichiometries.](https://example.com/f1.png)

(I. Crespo, N. Bernardo, A. Cuppari et al. Computational and Structural Biotechnology Journal 20 (2022) 757–765)
long helices, but interact with the long alpha helices of chains proto-
mers $n + 2$ 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$\text{pf}$ protein from plasmid F (PDB codes 3D8A and 2G7O)
and that of pED208 (PDB code 3ON0) as shown in Fig. 3D [52–
It should be noted that the positions of the shorter helices are layout to the structure of the TraD pF/TraMpF complex, with only traverses the region corresponding to the binding site of TraD pF, [54]. TraM also consists of two I. Crespo, N. Bernardo, A. Cuppari et al. Computational and Structural Biotechnology Journal 20 (2022) 757–765.

minor adjustments of the residues in the loop between the helices. The latter structure reveals the complex of the full-length protein bound to DNA. Comparison of the Aux2pLS20 four-helix bundle with that of the 3D8A and 3ON0 structures shows that the tetramerization domain in Aux2pLS20 is considerably longer than in the other analyzed structures, suggesting that the N-terminal turns of the long alpha helices in Aux2pLS20 may form part of the DNA binding domain. This region is not included in the crystal structures of the TraMpF 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 TraMpF 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 Aux2pLS20 with that of TraMpF (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 TraMpF. 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 Aux2pLS20, 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 TraMpF structure despite the swap in position.

TraMpF interacts with the plasmid F coupling protein, TraD, and the structure of the complex has been determined (PDB code 3D8A) [53]. The TraDpF peptide binds to the loop region that connects the N- and C-terminal helices (Fig. 3D). The helices in the 3ON0 structure TraMpF, not bound to the peptide, show a similar layout to the structure of the TraDpF/TraMpF complex, with only minor adjustments of the residues in the loop between the helices. In contrast, in Aux2pLS20, the loop between the helices of Aux2pLS20 traverses the region corresponding to the binding site of TraMpF, 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 Aux1pLS20 and two high-resolution structures of the tetramerization domain of Aux2pLS20. In addition, we analyze the possible interactions between the auxiliary proteins and the relaxase using 3D8A and pre-globular shape of Aux1pLS20. Given that RHH domains are intrinsically dimeric in nature, it follows that the oligomerization state of Aux1pLS20 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 Aux2pLS20 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 Aux2pLS20 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 Aux2pLS20 is highly unlikely given that the protein contains dimeric RHH domain. Interestingly, the oligomerization of Aux2pLS20 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 Aux2pLS20 are reminiscent of the TraM protein of plasmid F and pED208 [52–54,58]. Thus, Aux2pLS20 is a structural analog of the TraMpF 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 Aux2pLS20 C-terminal domain with that of TraMpF suggests that one of the functions of Aux2pLS20 is to interact with the coupling protein of pLS20.

Unlike TraMpF, which was proposed to also interact with the relaxase TraFpF [60], no interactions could be detected between Aux2pLS20 and RelpLS20. 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 Aux2pLS20 and the relaxase and it is likely that Aux2pLS20 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 Aux2pLS20 with that of TraMpF also provides clues on how different connector proteins select their substrates. The swapped position of the two helices with respect to TraMpF 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 Aux2pLS20, if it exists, is likely different from that observed for plasmid F-like connector proteins. Future work should elucidate if Aux2pLS20 interacts directly with the pLS20 coupling protein, and if the TraMpF binding pocket is preserved or is located in a different region of the protein.

We found that Aux1pLS20 is able to interact with the C-terminal domain of RelpLS20 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 RelpLS20 as well as its C-terminal domain, but not the N-terminal.

763
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 Aux1pLS20 and Aux2pLS20 likely bridge the various components of the relaxosome and mediate the interaction with the T4SS channel. We identify Aux2pLS20 as a structural analog of TraMpsp, which suggests that Aux2pLS20 links the relaxosome to the T4SS channel through interactions with the oriT. Furthermore, we show that Aux1pLS20 interacts with RelpLS20, which suggests that it stabilizes the binding of the relaxase to the DNA sequence, as both that RelpLS20 and Aux1pLS20 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 Aux1pLS20 and two X-ray structures of Aux2pLS20. The combined data presented here show that Aux1pLS20 and Aux2pLS20 are likely tetramers in solution. We identify Aux2pLS20 as a TraMpsp analog and show that Aux1pLS20 interacts with the C-terminal domain of RelpLS20.

Funding

This work was supported by Ministry of Economy and Competitiveness of the Spanish Government grants BIO2016-77883-C2-1-P and PID2019-108778BGB-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-Bourgert N, Guédon G. Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. FEMS Microbiol Rev 2014;38(4):720–60. https://doi.org/10.1111/1574-6976.12056.
[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, Zechnel 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] Similic V, Carcillén-Barcia MP, Francia MV, Rocha EP, de la Cruz F. Mobility of Plasmids. Microbiol Mol Biol Rev 2010;74(4):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 2016;1843(8):1578–91. https://doi.org/10.1016/j.bbamcr.2015.12.019.
[8] Low HH, Cubellini F, Rivera-Cañada A, Braun N, Connery S, Dujeancourt A, et al. Structure of a type IV secretion system. Nature 2014;508(7497):350–3. https://doi.org/10.1038/nature13083.
[9] Li YC, Hu Bo, Christie PJ, Sandvik M, Cascales E, Christie PJ. Biological and structural diversity of type IV secretion systems. Microbiol Spectr 2019;7(2).
[10] Zechnel EL, Lang S, Schüttelkopf A. 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] Wongs JWW, Lu J, Glover JM. 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, Kubon 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, Flucienickacz 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.2017.02153.
[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] Carcillén-Barcia MP, Redondo-Salvo S, Vilá E, de la Cruz F. MOBscan: automated annotation of MOB relaxases. Methods Mol Biol 2020:2075. https://doi.org/10.1007/978-1-4939-9878-7_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] Cormican R, Ruth 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.
