RESEARCH ARTICLE
Abdian *et al.*, *Microbiology* 2022;168:001284
DOI 10.1099/mic.0.001284

Fusion of a bacterial cadherin-like domain and green fluorescent protein as a specific probe to study biofilm matrix formation in *Rhizobium* spp

Patricia L. Abdian[1,*], María Soledad Malori[2], Julio J. Caramelo[2], Abi Maglio Checchi[1], Daniela M. Russo[2], Angeles Zorreguieta[2], Marcelo F. Berretta[1] and Graciela Benintende[1]

Abstract

Rhizobium adhering proteins or 'Raps' are secreted proteins identified in a very restricted group of rhizobial strains, specifically those belonging to *R. leguminosarum* and *R. etli*. The distinctive feature of members of the Rap family is the presence of one or two cadherin-like domains or CHDLs that are also present in numerous extracellular bacterial and archaeal proteins and were proposed to confer carbohydrate binding ability. We have previously made an in-depth characterization of RapA2, a calcium-binding lectin, composed by two CHDLs, involved in biofilm matrix remodelling in *R. leguminosarum* bv. *viciae* 3841. In this study, CHDLs derived from RapA2 were analysed in detail, finding significant structural and functional differences despite their considerable sequence similarity. Only the carboxy-terminal CHDL retained properties similar to those displayed by RapA2. Our findings were used to obtain a novel fluorescent probe to study biofilm matrix development by confocal laser scanning microscopy, and also to shed some light on the role of the ubiquitous CHDL domains in bacterial secreted proteins.

INTRODUCTION

Bacteria belonging to the genus *Rhizobium* are important soil microorganisms involved in the conversion of atmospheric nitrogen into ammonia, a suitable compound for plant nutrition [1]. Rhizobia go through a complex life cycle, being able to survive the harsh conditions in bulk soil and outcompete the rhizosphere microbiota before establishing symbiosis with the host plant [2]. Rhizobial communities inhabit diverse microenvironments forming biofilms [3–5], now accepted as the common lifestyle of microorganisms in natural environments [6]. In biofilms, cells are immersed in a self-produced polymeric matrix consisting of exopolysaccharides, proteins, nucleic acids, lipids and other biopolymers [7]. This extracellular matrix preserves the physical integrity of the community, but also represents the context in which all the interactions among resident microorganisms take place [8]. Matrix formation is fundamental to biofilm development [9], and has therefore become a growing field of microbiological research. However there is still fragmentary information regarding matrix composition, assembly and dynamics [10, 11].

Using the pea symbiont *Rhizobium leguminosarum* bv. *viciae* as a model organism we have previously studied the biochemical and biophysical properties of the calcium-binding lectin RapA2, a matrix protein that has a profound impact on biofilm development [12]. The major component of the biofilm matrix is an acidic exopolysaccharide [13], a hydrogel of great mechanical strength [14] to which RapA2 binds with high specificity. Later, we also showed that interaction of RapA2 with the acidic exopolysaccharides influences bacterial adhesion and biofilm matrix assembly [15].

RapA2 belongs to a family of proteins termed 'Raps' for *Rhizobial adhering proteins* that harbour one or two cadherin-like (CHDL) domains and are secreted by the type I PrsDE secretion system [16, 17]. RapA2 is the only member of the family characterized to date [12]. The lectin is a 24 kDa calcium-binding protein, composed solely by two CHDL domains, with a

Received 05 July 2022; Accepted 28 November 2022; Published 13 December 2022
Author affiliations: [1]Instituto de Microbiología y Zoología Agrícola (IMyZA), G.V. al IABiMo, INTA-CONICET, Nicolás Repetto y de los Reseros s/n, 1686, Hurlingham, Buenos Aires, Argentina; [2]Fundación Instituto Leloir, IIBBA CONICET, Patricias Argentinas 435, 1405, Buenos Aires, Argentina.
***Correspondence:** Patricia L. Abdian, abdian.patricia@inta.gob.ar
Keywords: biofilm matrix; cadherin-like domain; calcium; exopolysaccharide; GFP-fusion protein; *Rhizobium*; lectin.
Abbreviations: BIA, binding inhibition assay; C-CHDL, C-terminal cadherin-like domain; CD, circular dichroism; CHDL, cadherin-like domain; CLSM, confocal laser scanning microscopy; GFP, green fluorescent protein; ITC, isothermal titration calorimetry; N-CHDL, N-terminal cadherin-like domain; PhoA, alkaline phosphatase; Rap, *Rhizobium* adhering protein; SEC, size exclusion chromatography.
Four supplementary figures are available with the online version of this article.

predominance of beta-strand structural elements. It recognizes a structure containing glucuronic acid (GlcA) residues in the acidic exopolysaccharide. RapA2 calcium binding was shown by isothermal titration calorimetry (ITC) [12]. The location of the unique calcium-binding site derived from ITC experiments, first suspected to reside in a loop within both CHDLs, remains elusive despite extensive site directed mutagenesis analysis ([12]; P. Abdian unpublished results). Both folding and activity of RapA2 are dependent on calcium binding, which contributes greatly to protein stability. These properties may have implications for its secretion and function in the extracellular milieu [12, 18].

A considerable number of bacterial extracellular proteins contain CHDL domains, but to date there are few studies on the properties they confer to their cognate proteins [12, 19]. In order to generate a tool to study biofilm matrix, in this study, the CHDLs of RapA2 were produced as isolated proteins and their functional and structural properties evaluated and compared with the complete lectin. Interestingly, we found structural and functional differences among CHDLs, suggesting these protein domains fulfil distinct roles. The results obtained herein lead us to develop a sensitive probe that specifically detects the acidic exopolysaccharides produced by *R. leguminosarum*. The implications of these findings are discussed, and application of the fluorescent probe in combination with CLSM is exemplified with a study of biofilm matrix assembly.

METHODS

Bacterial strains and growth conditions

R. leguminosarum bv. *viciae* strain 3841 [20, 21] used in this study was grown at 28 °C in Tryptone Yeast (TY) medium [22] or in Y minimal medium [23] supplemented with 0.2% w/v mannitol and streptomycin (400 µg ml^{-1}). *Escherichia coli* strains used in this study, DH5α and BL21(DE3), were grown at 37 °C in Lysogeny Broth (LB) containing the appropriate antibiotics. Antibiotics were supplemented as required at the following concentrations: ampicillin, 200 µg ml^{-1}; chloramphenicol, 34 µg ml^{-1}; and tetracycline, 5 µg ml^{-1}.

DNA manipulation and plasmid construction

RapA2 (UniProt Q1M755, 236 amino acids) contains two bacterial cadherin-like domains (CHDLs) listed in the Pfam Database [24] as cadherin_4 family members under the ID PF17803. Domain delimitation was based on Pfam prediction and our previous studies [12]. The amino-terminal end of RapA2 containing the N-terminal cadherin-like domain (N-CHDL) comprises amino acids 1–118, while the carboxy-terminal end of RapA2 containing the C-terminal cadherin-like domain (C-CHDL) includes residues 119–236. The DNA encoding these domains were PCR-amplified with specific primers containing sites for *Nde*I and *Xho*I restriction endonucleases, using genomic DNA of *R. leguminosarum* bv. *viciae* strain 3841 as template. The amplified fragments were cloned into the pET22b vector. Primer pairs were as follows: 5′-GAGGAGAGTCATATGGCTTCGCC-3′ and 5′-GACCTCGAGTACGCCTTGGATGTTCAGG-3′ for N-CHDL, and 5′-ATTCATATGACCCAGGTCAAGCCGATCG-3′ and 5′-GCTGATCACTCGAGAGCGTCGGTGC-3′ for C-CHDL (restriction sites underlined).

To obtain the GFP-C-CHDL fusion protein a sequential cloning strategy was used to insert a spacer sequence that ensures correct folding of both proteins. First, the *gfp* gene (enhanced GFP or EGFP) from plasmid pHC60 [25] was amplified with a primer pair containing sites for *Not*I endonuclease (5′-AAGCGGCCGCAACGGTTTCCCTCTAGAAATAATTTTG-3′ and 5′-AAGCGGCCGCTGCGCAGCCGGATCCTTTGTATAGTTC-3′). The *gfp* amplicon was cloned into the *Not*I site of plasmid pTYB11 (NEB) that provides a spacer sequence of 24 nucleotides at the 3′-end of *gfp*. Then the DNA coding for C-CHDL was PCR-amplified from genomic DNA of *R. leguminosarum* 3841 with a pair of primers containing *Xho*I sites (5′-ATTCTCGAG AACCCAGGTCAAGCCGATC-3′ and 5′-GCTGATCACTCGAGAGCGTCGGTGC-3′). The PCR product was ligated with *Xho*I-digested pTYB11-GFP, linking the GFP- and C-CHDL-encoding fragments. Finally, the resulting construct, pTYB11-GFP-C-CHDL, was used as template in a second PCR reaction to amplify the fusion with primers 5′-AAAGGTACCATGGCTAG-CAAAGGAGAAG-3′ and 5′-AAAAGCTTCCTCGAGAGCGTCGGTGCCTG-3′, followed by cloning of the amplicon into the *Kpn*I and *Hind*III sites of pRSETB (Invitrogen), rendering pRSETB-FUS from which expression of the fusion protein with an N-terminal 6xHis-tag was achieved. Moreover, pRSETB-FUS was digested with *Eco*RI and *Hind*III, blunted and re-ligated to obtain pRSETB-GFP, used as a control. The constructs were verified by sequencing of the complete open reading frames. Plasmids were transferred to *E. coli* by electroporation using a Gene Pulser Transfection System (Bio Rad).

Protein expression and purification

Protein production was induced in *E. coli* BL21(DE3) cells harbouring plasmids pET22b-N-CHDL or pET22b-C-CHDL by addition of 0.5 mM isopropylthiogalactoside (IPTG) to bacterial cultures with an optical density at 600 nm (OD_{600}) of 0.6, and incubated for an additional 3 h at 37 °C.

Constructs pRSETB-FUS and pRSETB-GFP (control) were transferred to *E. coli* BL21(DE3)/pLysS by electroporation, and protein expression was achieved by addition of 0.5 mM IPTG when OD_{600} reached 0.8, followed by incubation for 16 h at 20 °C .

Cells were harvested by centrifugation at 3000 ***g*** for 15 min at 4 °C, washed and suspended in lysis buffer containing 10 mM imidazole. Subcellular fractionation and localization were performed by successive centrifugation at 2000 ***g***, 15 000 ***g***, and 100000 ***g*** during 1 h. Aliquots of each fraction were then analysed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Cell disruption, buffer composition and purification of N-CHDL and C-CHDL were performed as described before for RapA2 [12].

Purification of the fusion protein GFP-C-CHDL and control GFP were accomplished from 15000 ***g*** supernatants of *E. coli* BL21(DE3)/pLysS cells disrupted by sonication (Omni Sonic Ruptor 4000, 30 s, power 50 W, pulse 50%). Proteins were purified in the batch mode using HisPur Ni-NTA resin (Thermo Scientific) according to the manufacturer's instructions. Protein concentration was determined by Bradford assay using bovine serum albumin as a standard. Proteins were separated by 10% SDS-PAGE and visualized by Coomasie blue staining, or transferred to polyvinylidene difluoride (PVDF) membranes and revealed with a monoclonal alkaline phosphatase (PhoA)-conjugated anti-6xHis-tag antibody (SIGMA).

Structural and functional analysis of proteins

Secondary structure analysis of CHDL domains was performed by Far-UV circular dichroism (CD). Decalcified protein preparations were obtained by incubation with 10 mM ethylene glycol-bis(2-aminoethylether)-N,N,N′,N′-tetraaceticacid (EGTA) and re-purification by size exclusion chromatography (SEC). Protein samples were diluted to 25 µM (0.34 mg ml^{-1}) in 20 mM Tris-HCl and 150 mM NaCl, pH 8.0, with or without the addition of 1 mM $CaCl_2$, and placed in a 1 mm path length cell. Measurements were carried out at 25 °C on a Jasco J-815 spectropolarimeter. Spectra were acquired over the wavelength range of 195–250 nm. Each spectrum was the average of at least seven scans to reduce noise. Calcium titration was performed on decalcified protein samples by increasing $CaCl_2$ additions and measuring the CD signal at a fixed wavelength.

Functional analysis of the CHDL proteins was performed by means of the binding inhibition assay developed for RapA2 as previously described [12]. Proteins were used at a concentration of 50 ng ml^{-1}. Binding inhibition was evaluated by preincubation of the proteins with serial tenfold dilutions of the exopolysaccharides (*R. leguminosarum* acidic exopolysaccharide or xanthan gum, as control) before addition to the wells. Protein bound to the immobilized exopolysaccharide was detected by incubation with a monoclonal anti-6xHis-tag, followed by a PhoA-conjugated secondary antibody (SIGMA), addition of the 4-nitrophenyl phosphate substrate and colorimetric determination at 405 nm. Data represent the mean values of a representative experiment of three independent assays done in triplicate.

Polysaccharide-binding assay of the GFP-C-CHDL fusion protein

A dot blot test was set up to examine the fluorescence and the ability of the fusion protein to bind the acidic exopolysaccharide. The polyanionic uronic acid polysaccharides were reported to bind strongly to nylon membranes and could lately be detected by staining with cationic dyes [26]. Stock solutions (10 µg µl^{-1}) of *R. leguminosarum* acidic exopolysaccharide and xanthan gum (control) were diluted in water and 2 µl droplets were spotted on two identical strips of Hybond *N*+ membrane (Amersham) in the range of 16 µg to 5 ng. After air-drying, one of the strips was stained with a 1 mg ml^{-1} solution of methylene blue (SIGMA). The other strip was blocked for 1 h with 0.5% bovine serum albumin (BSA) in phosphate buffer saline (PBS) at room temperature. After washing, the membrane strip was incubated for 2 h with a 25 µg ml^{-1} solution of the fusion protein in PBS plus 1 mM $CaCl_2$. Washing steps were performed using PBS-Ca^{2+}, added with 0.005% Tween-20. The membrane strip was allowed to air-dry and finally observed under UV light with a portable lamp (UVP UVGL-58 254/365 nm) to check for fluorescence.

CLSM imaging of rhizobial biofilm matrix

Biofilms were formed in chambered cover glass slides (Lab-Tek Nunc) at 28 °C in a 4 day-experiment [13] to observe matrix development at different stages. The chambers contained *R. leguminosarum* 3841 in Y-mannitol minimal medium at an initial OD_{600} of 0.001. At different time points, liquid was carefully removed and the chambers were filled with 200 µl of the fluorescent lectin solution obtained herein (0.5 mg ml^{-1} of GFP-C-CHDL fusion protein in PBS, 1 mM Ca^{2+}). In a control experiment, the fluorescent lectin was replaced by partially purified GFP. Samples were incubated for 30 min in the dark at room temperature. After careful removal of the solution, the samples were washed three times with PBS-Ca^{2+} to remove unbound protein and finally covered with the same buffer to avoid drying. Images were acquired using a Zeiss LSM710 AxioObserver microscope with the 40× or the 63× oil-immersion objective, setting the filter for GFP (488/595) and operated with the Zen software (Zeiss, Germany). Optical sections were taken at various focal planes to create Z-stacks with 0.4 µm intervals in the regions of interest.

RESULTS

Heterologous expression and purification of CHDLs

To understand the biological role of CHDL domains and determine their functional and structural properties, DNA fragments encoding the amino- and carboxy-terminal domains of RapA2 (N-CHDL and C-CHDL, respectively), were amplified with specific primers and subsequently cloned into an expression vector. Fusions of the CHDL proteins with a C-terminal 6xHis-tag

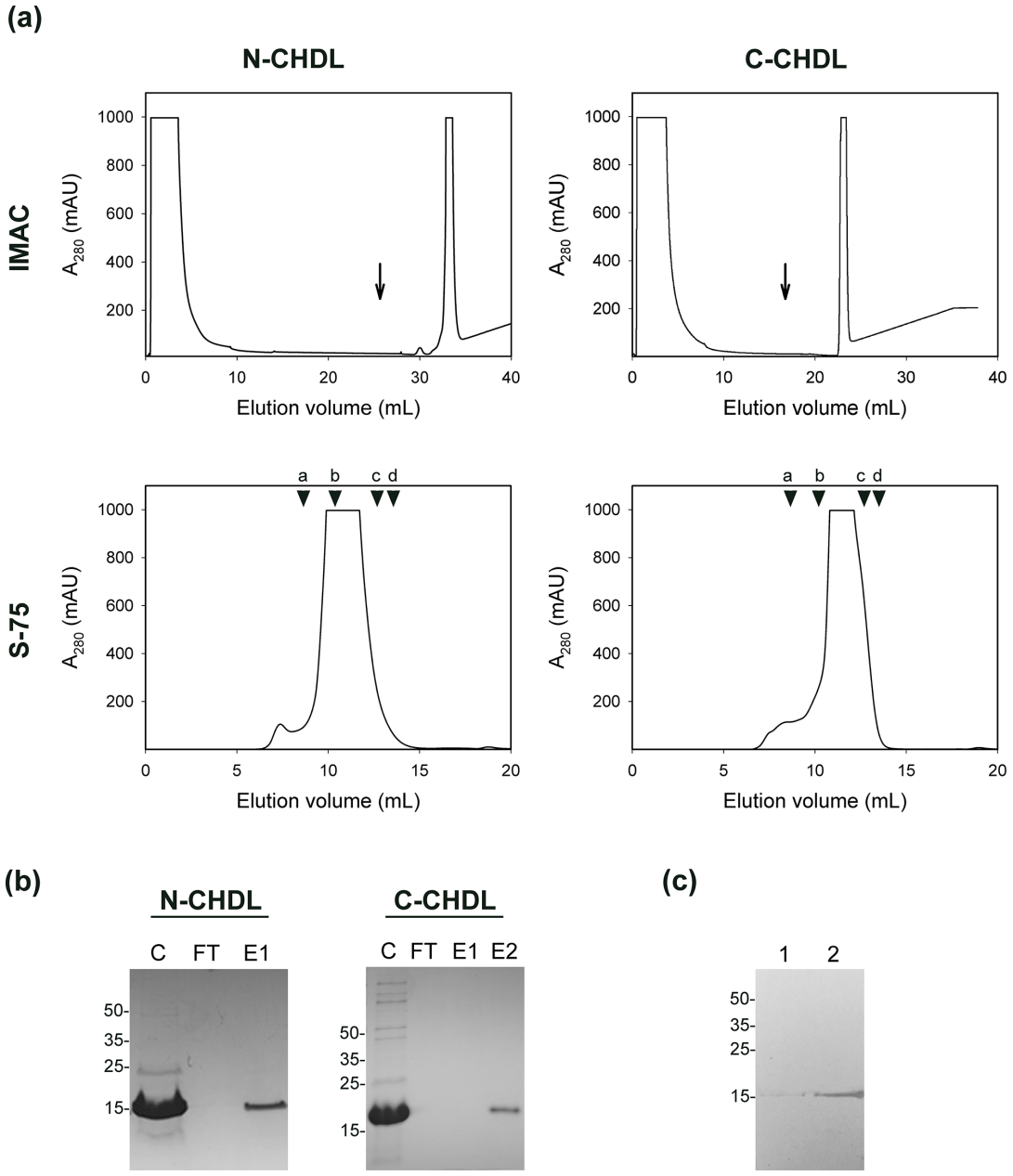

Fig. 1. Two step purification scheme of CHDLs. (a) Protein elution profiles from affinity chromatography (IMAC) and size exclusion chromatography (S-75). MW standards: a, 158 kDa; b, 43 kDa; c, 17 kDa; d, 13 kDa. The arrow in IMAC profiles points the initiation of elution with an imidazole gradient. (b) SDS-PAGE of the IMAC protein samples after ultrafiltration (C, concentrated) and fractions eluting from the S-75 column (FT, flow through; E, elution). (c) Western blot analysis of purified N-CHDL and C-CHDL after EGTA treatment (lanes 1 and 2, respectively) using monoclonal anti-6xHis-tag antibody.

(~13 kDa) achieved high levels of heterologous expression in *E. coli*. After cell disruption, analysis by SDS-PAGE showed both proteins localized to the soluble fraction (Fig. S1, available in the online version of this article).

A two-step protocol, consisting of Immobilized Metal Affinity Chromatography (IMAC) and SEC, was followed to purify the CHDLs from soluble fractions of *E. coli* (Fig. 1a). As judged by SDS-PAGE, both N-CHDL and C-CHDL were purified to homogeneity (Fig. 1b). To note, their theoretical MW (13 kDa) did not agree with that observed in SEC (43 kDa for N-CHDL and 39 kDa for C-CHDL). These differences denote an increase in the hydrodynamic radius of the CHDLs, distant from that expected for globular proteins. A similar observation was previously made for RapA2 [12] and for several calcium-binding proteins, for which overestimation of MW in SEC is due to highly extended conformation [27]. Decalcification of the purified CHDL protein samples with EGTA reduced their apparent MW in SEC (39 and 34 kDa for N- and C-CHDL, respectively) (Fig.

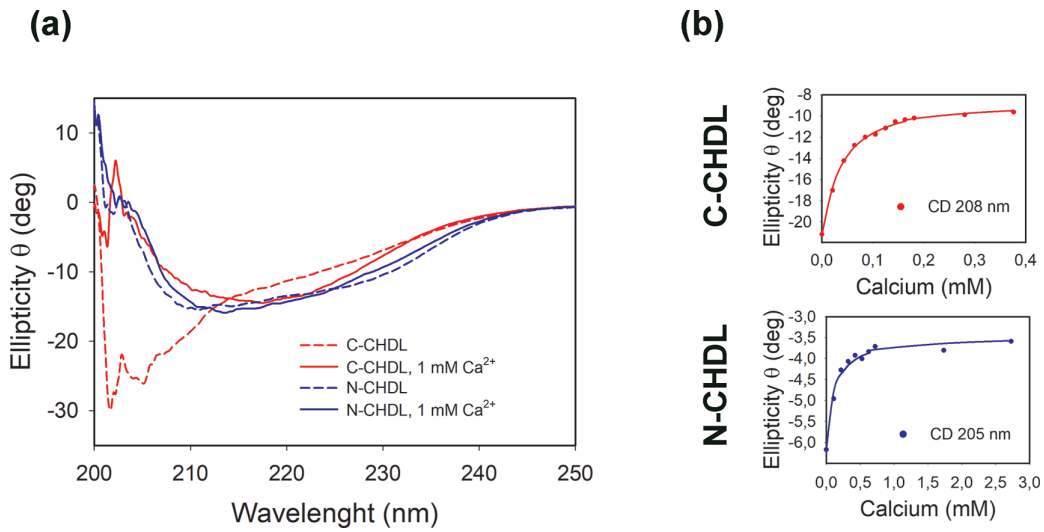

Fig. 2. Effect of calcium on the structures of N-CHDL and C-CHDL. (a) Far UV-CD spectra of CHDLs with or without 1 mM $CaCl_2$ addition. (b) Variation of ellipticity at a fixed wavelength in response to increasing calcium concentrations.

S2). Furthermore, analysis of decalcified protein samples by Western blot, confirmed N-CHDL and C-CHDL equal size (near 13 kDa) and the presence of the 6xHis-tag (Fig. 1c).

CD-Spectroscopic analysis of CHDLs

It was previously shown that calcium ions are important for the folding of RapA2 to its native structure [12]. With the ability to produce the isolated CHDL proteins in purified soluble state, their structural analysis was approached by CD spectroscopy. Spectra in the far-UV of C-CHDL and N-CHDL were recorded in the absence/presence of calcium to determine whether the CHDLs retained the ability to bind the ion. The folding properties of C-CHDL closely resemble those of the parent protein, RapA2. As shown in Fig. 2, C-CHDL displays a spectrum similar to that of denatured proteins in the absence of calcium, with a minimum near 201 nm. Upon calcium addition, a significant change in the CD spectrum was produced, displaying a minimum at 216 nm, a clear sign of beta-sheet conformation. On the other hand, N-CHDL seems not fully unfolded in the absence of calcium and the spectrum may represent a molten globule-like state [28]. When calcium was added, a small shift in the spectrum of N-CHDL was observed, without reaching the typical native CD-spectra displayed by RapA2 and C-CHDL. Moreover, titration of CHDLs with calcium followed at a fixed wavelength, allowed estimation of the equilibrium dissociation constants (K_D) (Fig. 2b). These were 67.4 µM and 19.6 µM for N-CHDL and C-CHDL, respectively.

Functional analysis of CHDLs

To compare the lectin activity of the isolated CHDLs with RapA2, a modification of the ELISA technique was used. The binding inhibition assay (BIA) measures binding of residual lectin to the acidic exopolysaccharide immobilized in multiwell plates, after pre-incubation of the lectin with serial dilutions of the polysaccharide solution [12]. Detection of bound lectin was done by incubation with anti-6xHis-tag antibodies conjugated to PhoA, followed by addition of the phosphatase substrate and measurement of optical density at 405 nm. As shown in Fig. 3, a concentration-dependent inhibition of binding to the immobilized polysaccharide was observed, both for RapA2 and the C-CHDL protein. However, a clear change in the affinity for the polysaccharide was measured, with C-CHDL showing a decrease in affinity of one order of magnitude compared to that of RapA2. On the other hand, the N-CHDL protein showed no ability to recognize or bind the acidic exopolysaccharide. In addition, none of the proteins examined were able to bind xanthan, an acidic heteropolysaccharide of different structure used as control. These results clearly show that the specific activity of RapA2 is confined in its carboxy-terminus and is retained, although with less affinity, by the C-CHDL protein.

Fusion of C-CHDL with the green fluorescent protein

As shown in this work, C-CHDL represents only half of the whole RapA2 lectin but largely retains its structural and functional properties, binding specifically the exopolysaccharide that constitutes the major component of the biofilm matrix. Therefore, an interesting possibility arose related to the study of *Rhizobium* biofilm matrix formation. A fusion of the active C-CHDL with a stable form of GFP could allow visualization of matrix assembly by high resolution microscopy.

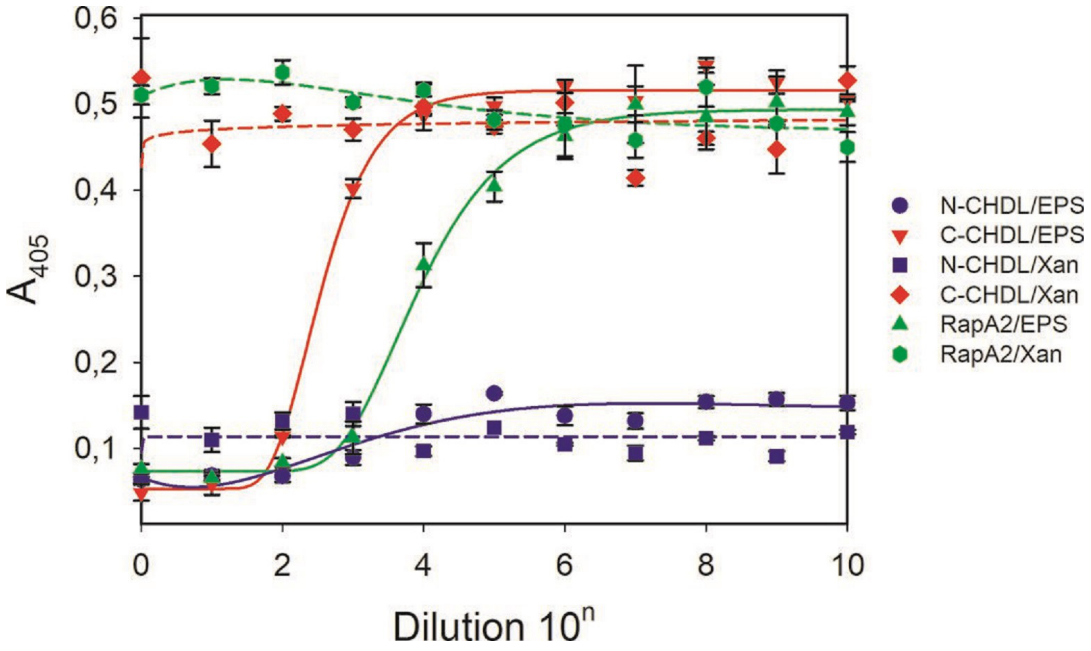

Fig. 3. Binding inhibition assay (BIA). RapA2, C-CHDL and N-CHDL were pre-incubated with serial tenfold dilutions of *R. leguminosarum* exopolysaccharide (EPS) or xanthan (Xan, used as control) before addition to the wells of a microplate. Protein bound to the immobilized exopolysaccharide was detected with anti-6xHis-tag antibody conjugated to PhoA.

The *gfp* gene was cloned upstream of the C-CHDL coding sequence, with a spacer of 24 nucleotides between them and a 6xHis-tag encoding sequence at the 5′-end to facilitate purification (Fig. 4a). The correct position and length of the spacer proved extremely important to ensure correct folding and function. The construct was introduced into *E. coli* BL21(DE3)/pLysS cells and expression conditions were optimized until soluble, fluorescent and stable fusion protein was obtained. Purification was achieved in one-step by IMAC, and protein elution was monitored by direct eye inspection since fluorescence was evident throughout the fractions that contained the protein. Analysis by SDS-PAGE showed a protein band of ~45 kDa, in agreement with the expected size for the fusion protein GFP-C-CHDL (Fig. 4b).

Validation and use of the fusion protein in the study of biofilm matrix formation

The acidic exopolysaccharide is the major structural constituent required for the three-dimensional growth of *R. leguminosarum* biofilms both *in vitro* and *in vivo* [3, 13]. The possibility to follow its appearance from very early steps until development of a

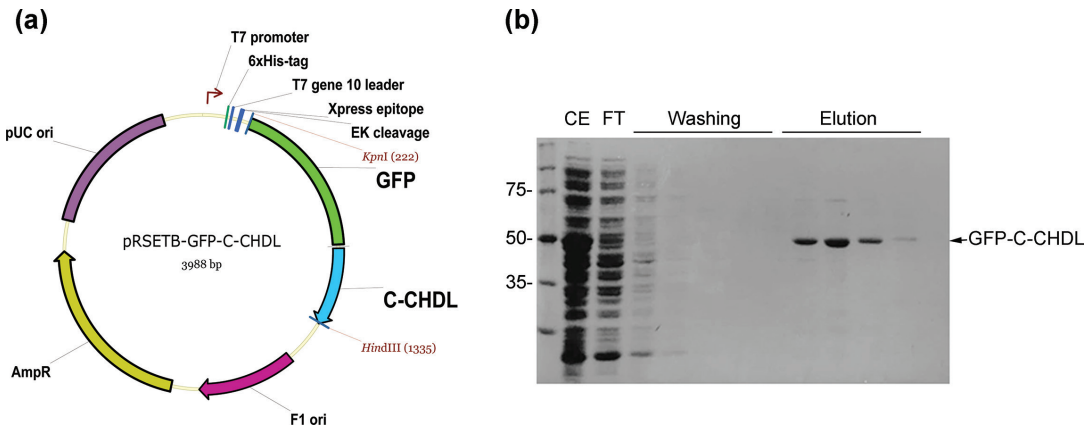

Fig. 4. Fusion of C-CHDL to GFP. (a) Physical map of the plasmid construct. AmpR, ampicillin resistance gene. (b) Examination of the IMAC fractions content during purification of the fusion protein GFP-C-CHDL by SDS-PAGE. CE, total cell extract; FT, flow through.

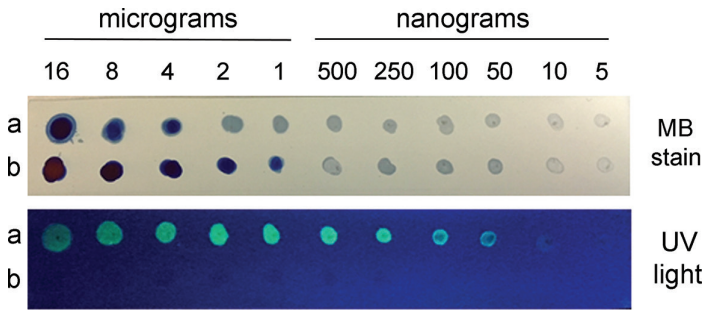

Fig. 5. Dot blots of aqueous solutions of *R. leguminosarum* exopolysaccharide (a) and xanthan gum (b) in decreasing concentrations. Upper panel, the membrane was stained with the cationic dye methylene blue (MB). Lower panel, the membrane was blocked with BSA, incubated with GFP-C-CHDL fusion protein and after drying, it was photographed under UV light.

mature biofilm, its dynamics under diverse growth conditions or even in diverse mutants derived from the parental strain, requires the combination of CLSM with a specific and sensible tool, such as the fluorescent probe described herein.

As a first step, the fusion protein GFP-C-CHDL was evaluated by means of a polysaccharide blotting method, in which serial dilutions of the acidic exopolysaccharide produced by *R. leguminosarum* or xanthan gum (negative control) in concentrations ranging from 16 µg to 5 ng were immobilized on nylon membrane strips. Blots of both exopolysaccharides showed clear staining reactions with the cationic dye methylene blue, but only *R. leguminosarum* exopolysaccharide was visible under UV light after incubation with GFP-C-CHDL (Fig. 5). In this way, the fusion protein was demonstrated to be a specific fluorescent probe of high sensitivity, with a detection limit for the exopolysaccharide in the nanogram range.

Next, the usefulness of the GFP-C-CHDL was assessed during the development of *R. leguminosarum* biofilms by CLSM. The fluorescent probe was useful to observe the localization of the acidic exopolysaccharide in early steps of biofilm formation, but also in mature 4 day biofilms (Fig. 6). Green fluorescence appeared in spots on the surface of isolated cells or surrounding them, 24 hours after inoculation of the chambered coverglasses used to observe biofilms. On day two, the exopolysaccharide was detected in the spaces between cells, apparently aiding in the formation of microcolonies. Later (day four), the fluorescent signal was stronger in the upper and most exposed surfaces of the mature biofilm. In contrast, no fluorescence was detected in biofilms incubated with partially purified GFP (Fig. S3). Therefore, the use of this specific tool made it clear that the acidic exopolysaccharide is required at all stages of biofilm formation by *R. leguminosarum*.

DISCUSSION

This study describes a useful tool derived from RapA2, a calcium-binding lectin, to analyse biofilm matrix production by *R. leguminosarum*. Construction of a fusion between GFP and the C-CHDL of RapA2 required prior analysis of the protein domains which also provide insight into the function of bacterial CHDLs.

Previous computational and experimental data [12, 19, 29] strongly suggest that CHDL domains confer carbohydrate-binding activity to the proteins bearing them. High sequence similarities among CHDLs of proteins belonging to the Rap family, was interpreted as indicative of recognition and binding to a specific glycosidic structure in the acidic exopolysaccharide produced by *R. leguminosarum*. However, our analysis of the CHDLs that constitute RapA2 (56% identity) demonstrate that the isolated protein domains do not share structural or functional properties. Circular dichroism (CD) spectroscopy showed that C-CHDL was able to bind calcium ions with higher affinity, and subsequently folded into a compact β-sheet structure. On the other hand, N-CHDL remained partially unfolded, even after EGTA incubation. Although it binds calcium, no structural transition was detected. This could be due to non-specific binding of calcium, related to a high content of acidic residues exposed in the partially unfolded N-CHDL that may act as potential ion-ligands, as described for other proteins [30, 31].

The lack of canonical calcium-binding motifs in the sequence of RapA2 precluded the prediction of the stoichiometry of calcium binding, but our previous results derived from ITC analysis clearly indicated the existence of only one calcium binding site per molecule [12]. Based on the results obtained here, we hypothesize that the unique calcium-binding site of RapA2 is located in its C-terminal domain (C-CHDL). This assumption is also supported by the observation made by Ausmees *et al.* [16] on the RapA1 protein from *R. leguminosarum* bv. *trifolii* R200. A truncated derivative of this protein lacking 34 amino acids in the C-terminus lost calcium-binding ability.

Functional BIA assays showed that only C-CHDL conserved lectin activity and binds specifically the acidic exopolysaccharide produced by *R. leguminosarum*. It was observed that preincubation of C-CHDL with EGTA prevented binding to the

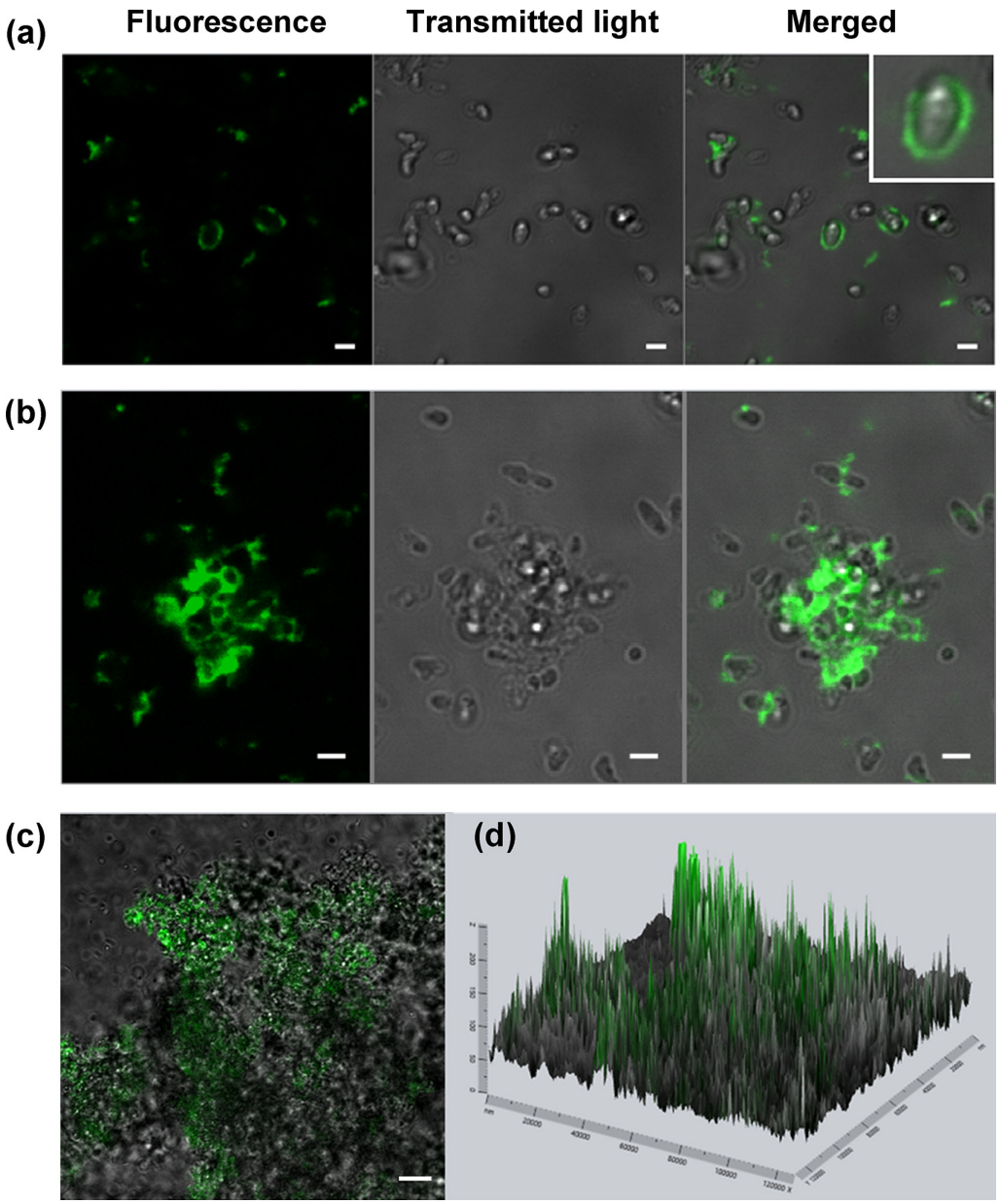

Fig. 6. Localization of exopolysaccharide by CLSM during biofilm matrix development in *R. leguminosarum*. Biofilms were grown in Y-mannitol medium using chambered glass slides. At different time points, biofilms were stained with the fluorescent probe, GFP-C-CHDL, washed and photographed. (a) Top view of isolated cells (day 1) using a 40× water immersion objective lens. Inset is 3× zoom. (b) Microcolony imaged on day 2. (c) Mature biofilm on day four and (d) surface projections of mature biofilm structure prepared using the ZEN software. Scale bars: (a) and (b), 2 µm; (c) 5 µm.

exopolysaccharide (Fig. S4), suggesting that calcium triggers folding and that calcium binding to the protein is a prerequisite for lectin activity. However, the binding activity of C-CHDL was significantly reduced compared to RapA2 then N-CHDL may somehow contribute to the lectin activity. In this regard, N-CHDL was unable to bind the exopolysaccharide in the functional assay, probably caused by failure to adopt its native structure in isolation. Indeed we have previously proposed a calcium-dependent cooperative folding mechanism for the RapA2 protein [12] and therefore N-CHDL participation in exopolysaccharide binding cannot be completely excluded.

In view of the C-CHDL properties uncovered in this work, this protein fragment was fused to GFP resulting in a novel probe that specifically detects the acidic exopolysaccharide produced by *R. leguminosarum* in the nanogram range. The fluorescent probe was tested with CLSM imaging of biofilm matrix development. As previously reported, the acidic exopolysaccharide is the main polymer in the biofilm matrix formed by *R. leguminosarum* on abiotic surfaces [13]. Use of the GFP-C-CHDL probe

to analyse *R. leguminosarum* biofilm formation on a glass surface enabled detection of the matrix polymer from the onset of cell-cell interactions until the development of a mature biofilm structure. Interestingly, on day 1, the exopolysaccharide appeared as discrete spots with polar or subpolar location on the bacterial surface. Later, it was detected as the cementing material among cells in microcolonies (day 2), and was localized mostly in the uppermost section of mature biofilms (day 4). Although reduced diffusion of the fluorescent probe in the biofilm structure could hinder access to the exopolysaccharide in the lower sections of mature biofilms, it is also possible that the polymer is specifically located in the upper zone. This particular localization of exopolysaccharides was also observed in *E. coli* K12 biofilms on agar, and was attributed to the existence of metabolically distinct cell types stratified in the structure of the biofilm, with cells in the upper zone contributing to matrix structure mainly through active production of the exopolysaccharide cellulose [32]. At present, we cannot rule out any of these possibilities, but the GFP-C-CHDL probe will allow determining the precise location of the exopolysaccharide during biofilm development. Studies are underway to clarify this issue.

According to Neu and Kuhlicke [33] lectin staining in combination with CLSM still is the fastest and most straightforward technique to evaluate carbohydrate distribution in complex biofilms. Lectins constitute powerful probes to identify biofilm subdomains and could be useful to assess biofilm sorption properties. They represent a better alternative to the use of antibodies which are difficult to obtain and depend on the recovery and purification of significant quantities of polysaccharide material from the biofilm matrix.

Undoubtedly, the binding specificity of different lectins expands the possibilities to map the structure and chemistry of polysaccharides in biofilm systems. This report intends to be a contribution in this sense, providing a detailed characterization of a bacterial lectin displaying novel specificity, and further showing a possible application in the characterization and dynamics of the extracellular matrix of a microbial biofilm.

Funding information
This research work did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements
We acknowledge Marta Bravo, Andrea Puebla and Irma Fuxan for DNA sequencing and technical assistance, as well as Dr Cora Polack for assistance with CLSM at IBIOBA. P.L.A., M.S.M., J.J.C., D.M.R., A.Z. and M.F.B. are staff members of CONICET.

Conflicts of interest
The author(s) declared that there are no conflicts of interest.

References

1. Lindström K, Mousavi SA. Effectiveness of nitrogen fixation in rhizobia. *Microb Biotechnol* 2020;13:1314–1335.
2. Poole P, Ramachandran V, Terpolilli J. Rhizobia: from saprophytes to endosymbionts. *Nat Rev Microbiol* 2018;16:291–303.
3. Downie JA. The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. *FEMS Microbiol Rev* 2010;34:150–170.
4. Oleńska E, Małek W, Kotowska U, Wydrych J, Polińska W, *et al*. Exopolysaccharide carbohydrate structure and biofilm formation by *Rhizobium leguminosarum* bv. *trifolii* strains inhabiting nodules of trifoliumrepens growing on an old Zn–Pb–Cd-polluted waste heap area. *IJMS* 2021;22:2808.
5. Velmourougane K, Prasanna R, Saxena AK. Agriculturally important microbial biofilms: present status and future prospects. *J Basic Microbiol* 2017;57:548–573.
6. Flemming HC, Wuertz S. Bacteria and archaea on Earth and their abundance in biofilms. *Nat Rev Microbiol* 2019;17:247–260.
7. Wingender J, Neu TR, Flemming HC. What are bacterial extracellular polymeric substances. In: Wingender J, Neu TR and Flemming HC (eds). *Microbial Extracellular Polymeric Substances*. Berlin, Heidelberg: Springer; 1999. pp. 1–19.
8. Flemming H-C, Wingender J, Szewzyk U, Steinberg P, Rice SA, *et al*. Biofilms: an emergent form of bacterial life. *Nat Rev Microbiol* 2016;14:563–575.
9. Flemming HC, Wingender J. The biofilm matrix. *Nat Rev Microbiol* 2010;8:623–633.
10. Abdian PL, Zorreguieta A. Extracellular factors involved in biofilm matrix formation by Rhizobia. In: Flemming HC, Neu TR and Wingender J (eds). *The Perfect Slime – Microbial Extracellular Polymeric Substances (EPS)*. London: IWA Publishing; 2016. pp. 227–247.
11. Seviour T, Derlon N, Dueholm MS, Flemming H-C, Girbal-Neuhauser E, *et al*. Extracellular polymeric substances of biofilms: suffering from an identity crisis. *Water Res* 2019;151:1–7.
12. Abdian PL, Caramelo JJ, Ausmees N, Zorreguieta A. RapA2 is a calcium-binding lectin composed of two highly conserved cadherin-like domains that specifically recognize *Rhizobium leguminosarum* acidic exopolysaccharides. *J Biol Chem* 2013;288:2893–2904.
13. Russo DM, Williams A, Edwards A, Posadas DM, Finnie C, *et al*. Proteins exported via the PrsD-PrsE type I secretion system and the acidic exopolysaccharide are involved in biofilm formation by *Rhizobium leguminosarum*. *J Bacteriol* 2006;188:4474–4486.
14. Coviello T, Burchard W, Geissler E, Maier D. Static and dynamic light scattering by a thermoreversible gel from *Rhizobium leguminosarum* 8002 exopolysaccharide. *Macromolecules* 1997;30:2008–2015.
15. Vozza NF, Abdian PL, Russo DM, Mongiardini EJ, Lodeiro AR, *et al*. A *Rhizobium leguminosarum* CHDL- (Cadherin-Like-) Lectin Participates in Assembly and Remodeling of the Biofilm Matrix. *Front Microbiol* 2016;7:1608.
16. Ausmees N, Jacobsson K, Lindberg M. A unipolarly located, cell-surface-associated agglutinin, RapA, belongs to A family of *Rhizobium*-adhering proteins (Rap) in *Rhizobium leguminosarum* bv. *trifolii*. *Microbiology* 2001;147:549–559.
17. Krehenbrink M, Downie JA. Identification of protein secretion systems and novel secreted proteins in *Rhizobium leguminosarum* bv. viciae. *BMC Genomics* 2008;9:55.
18. Sotomayor-Pérez AC, Ladant D, Chenal A. Disorder-to-order transition in the CyaA toxin RTX domain: implications for toxin secretion. *Toxins* 2014;7:1–20.

19. Cao L, Yan X, Borysenko CW, Blair HC, Wu C, *et al*. CHDL: a cadherin-like domain in Proteobacteria and Cyanobacteria. *FEMS Microbiol Lett* 2005;251:203–209.
20. Johnston AW, Beringer JE. Identification of the rhizobium strains in pea root nodules using genetic markers. *J Gen Microbiol* 1975;87:343–350.
21. Young JPW, Crossman LC, Johnston AWB, Thomson NR, Ghazoui ZF, *et al*. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol* 2006;7:R34.
22. Beringer JE. R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* 1974;84:188–198.
23. Sherwood MT. Improved synthetic medium for the growth of *Rhizobium*. *J Appl Bacteriol* 1970;33:708–713.
24. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, *et al*. Pfam: The protein families database in 2021. *Nucleic Acids Res* 2021;49:D412–D419.
25. Cheng HP, Walker GC. Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. *J Bacteriol* 1998;180:5183–5191.
26. Jeffree CE. Immobilization of pectin, xyloglucan and other soluble plant polysaccharides on blotting membranes. *New Phytol* 1993;125:695–706.
27. La Verde V, Dominici P, Astegno A. Determination of hydrodynamic radius of proteins by size exclusion chromatography. *Bio Protoc* 2017;7:e2230.
28. Uversky VN, Finkelstein AV. Life in phases: intra- and inter-molecular phase transitions in protein solutions. *Biomolecules* 2019;9:842.
29. Fraiberg M, Borovok I, Bayer EA, Weiner RM, Lamed R. Cadherin domains in the polysaccharide-degrading marine bacterium *Saccharophagus degradans* 2-40 are carbohydrate-binding modules. *J Bacteriol* 2011;193:283–285.
30. Hoyer E, Knöppel J, Liebmann M, Steppert M, Raiwa M, *et al*. Calcium binding to a disordered domain of a type III-secreted protein from a coral pathogen promotes secondary structure formation and catalytic activity. *Sci Rep* 2019;9:7115.
31. Meissner G, Conner GE, Fleischer S. Isolation of sarcoplasmic reticulum by zonal centrifugation and purification of Ca 2+ -pump and Ca 2+ -binding proteins. *Biochim Biophys Acta* 1973;298:246–269.
32. Serra DO, Richter AM, Hengge R. Cellulose as an architectural element in spatially structured *Escherichia coli* biofilms. *J Bacteriol* 2013;195:5540–5554.
33. Neu TR, Kuhlicke U. Fluorescence lectin bar-coding of glycoconjugates in the extracellular matrix of biofilm and bioaggregate forming microorganisms. *Microorganisms* 2017;5:5.

Edited by: M. Welch and N. Holden