

RESEARCH ARTICLE

# The crystal structure of *Erwinia amylovora* AmyR, a member of the YbjN protein family, shows similarity to type III secretion chaperones but suggests different cellular functions

Joseph D. Bartho<sup>1</sup>, Dom Bellini<sup>2,3</sup>, Jochen Wuerges<sup>1</sup>, Nicola Demitri<sup>4</sup>, Mirco Toccafondi<sup>1</sup>, Armin O. Schmitt<sup>1,5</sup>, Youfu Zhao<sup>6</sup>, Martin A. Walsh<sup>2,3</sup>, Stefano Benini<sup>1\*</sup>

**1** Bioorganic Chemistry and Bio-Crystallography laboratory (B2CI), Faculty of Science and Technology, Free University of Bolzano, Piazza Università 5, Bolzano, Italy, **2** Diamond Light Source, Harwell Science and Innovation Campus, Didcot, United Kingdom, **3** Research Complex at Harwell, Harwell Science and Innovation Campus, Didcot, United Kingdom, **4** Elettra—Sincrotrone Trieste, S.S 14 km 163.5 in Area Science Park, Basovizza, Trieste, Italy, **5** Georg-August-Universität Göttingen, Dept. Nutztierwissenschaften, Breeding informatics, Margarethe von Wrangell-Weg 7, Göttingen, Germany, **6** Department of Crop Sciences, University of Illinois, 1201 W. Gregory Dr., Urbana, IL, United States of America

\* [stefano.benini@unibz.it](mailto:stefano.benini@unibz.it)



**OPEN ACCESS**

**Citation:** Bartho JD, Bellini D, Wuerges J, Demitri N, Toccafondi M, Schmitt AO, et al. (2017) The crystal structure of *Erwinia amylovora* AmyR, a member of the YbjN protein family, shows similarity to type III secretion chaperones but suggests different cellular functions. PLoS ONE 12 (4): e0176049. <https://doi.org/10.1371/journal.pone.0176049>

**Editor:** Eric Cascales, Centre National de la Recherche Scientifique, Aix-Marseille Université, FRANCE

**Received:** January 12, 2017

**Accepted:** April 4, 2017

**Published:** April 20, 2017

**Copyright:** © 2017 Bartho et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All coordinates and reflections files are available from the PDB database (accession number(s): 5FRK, 5FR7).

**Funding:** JDB was supported by a predoctoral fellowship from the Free University of Bolzano MT was supported by a postdoctoral fellowship by the Fondazione Libera Università di Bolzano. The work was supported by the grant: "A structural

## Abstract

AmyR is a stress and virulence associated protein from the plant pathogenic Enterobacteriaceae species *Erwinia amylovora*, and is a functionally conserved ortholog of YbjN from *Escherichia coli*. The crystal structure of *E. amylovora* AmyR reveals a class I type III secretion chaperone-like fold, despite the lack of sequence similarity between these two classes of protein and lacking any evidence of a secretion-associated role. The results indicate that AmyR, and YbjN proteins in general, function through protein-protein interactions without any enzymatic action. The YbjN proteins of Enterobacteriaceae show remarkably low sequence similarity with other members of the YbjN protein family in Eubacteria, yet a high level of structural conservation is observed. Across the YbjN protein family sequence conservation is limited to residues stabilising the protein core and dimerization interface, while interacting regions are only conserved between closely related species. This study presents the first structure of a YbjN protein from Enterobacteriaceae, the most highly divergent and well-studied subgroup of YbjN proteins, and an in-depth sequence and structural analysis of this important but poorly understood protein family.

## Introduction

*Erwinia amylovora* is a Gram-negative plant pathogen that belongs to the Enterobacteriaceae family, which is closely related to other pathogenic Enterobacteria such as *Escherichia coli*, *Yersinia pestis*, *Salmonella enterica* and *Shigella flexneri* [1]. *E. amylovora* is the causal agent of Fire Blight, a severe necrotic disease affecting species of the Rosaceae family, and of particular

genomics approach for the study of the virulence and pathogenesis of *Erwinia amylovora*” from the Autonomous Province of Bolzano and the grant: “GenSelEa” from the Faculty of Science and Technology, Free University of Bolzano. The above mentioned funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.” The funder Diamond Light Source Ltd. provided beamtime at the beamlines I04 and I04-1 for data collection, and support in the form of salaries for the authors DB and MAW but did not have any additional role in the study design, data analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the ‘author contributions’ section.

**Competing interests:** The authors declare that no competing interests exist and the presence of Diamond Light Source Ltd as funder does not alter our adherence to PLOS ONE policies on sharing data and materials.

importance to the apple and pear fruit industries [2]. The infection strategy of *E. amylovora* hinges on the expression of the hrp (hypersensitive response and pathogenicity) type III secretion system (T3SS) and the production of an exopolysaccharide (EPS) matrix [3–5]. The matrix is mainly composed of the two EPSs, levan and amylovoran [4]. While levansucrase is the only enzyme designated to produce levan [6–9], amylovoran biosynthesis requires twelve proteins encoded by the genes of the *ams* operon and proteins involved in the galactose metabolism [10–12]. Amylovoran is one of the most important *E. amylovora* pathogenicity factors and strains with defective amylovoran biosynthesis are non-pathogenic [13]. Regulation of amylovoran production involves a complex signalling network, which is still poorly understood [14].

The amylovoran repressor gene *amyR* of *E. amylovora* is a member of the *ybjN* gene family, which is highly conserved among Enterobacteria [15,16], although distantly related members of the *ybjN* family appear in a wide spectrum of bacteria. The *ybjN* gene from *E. coli* has been defined as a stress-related gene, whose expression depends on the temperature and the growth stage [15]. Microarray analysis demonstrated that the over-expression of *ybjN* in *E. coli* induces several pleiotropic effects, promoting the transcription of stress-related genes such as the toxin-antitoxin module, SOS stress response, cold-shock, and phage-shock genes [17]. At the same time, *ybjN* down-regulates metabolic pathways such as energy metabolism and biosynthetic processes. *E. coli* strains over-expressing *ybjN* were observed to slow down their growth rate and decrease their swarming motility [17]. Moreover, there is a marked reduction in the biosynthesis of the capsule, biofilm formation and increased acid susceptibility. Expression of *ybjN* is induced under several stress conditions [18–20]. Therefore, *ybjN* in *E. coli* has an important role in stress-response and is able to regulate some virulence factors shared among other Enterobacteriaceae [17,21]. Recent studies have demonstrated that *E. amylovora* AmyR exerts a similar function on the metabolism of this organism in stress conditions [16]. Importantly, the over-expression of *amyR* negatively regulates the production of the two *E. amylovora* exopolysaccharides, amylovoran and levan. In fact, an *E. amylovora amyR* knockout strain demonstrated an overproduction of amylovoran 8-fold higher than the control in MBMA medium and colonies on LB plates exhibited a mucoid phenotype [16]. Therefore, *amyR* in *E. amylovora*, as with *ybjN* in *E. coli*, is involved in regulating exopolysaccharide production.

AmyR regulates the *ams* operon and *E. amylovora* strains with mutated *amyR* exhibited expression of *amsG*, *amsC* and *amsD* up to five-fold higher than wild-type levels [16]. Ams proteins are required for amylovoran production and their overexpression correlates with an increased EPS biosynthesis. Furthermore, the *amyR* mutation also significantly increased the transcription level of the genes *galF* and *galE* involved in galactose metabolism [16]. The encoded protein GalF could be considered an isoform of GalU, an uridylyltransferase responsible for UDP-glucose production that the second protein, GalE, converts to UDP-galactose [22–24]. These two enzymes together have a pivotal role in amylovoran production as the providers of galactose, the essential building block for this biosynthetic pathway [11]. In contrast, the over-expression of *amyR* led to a marked inhibition of amylovoran production, including in several amylovoran-overproducing mutants, and resulted in a loss of pathogenicity on apple shoots and leaves [16].

The Rcs phosphorelay system is an atypical two-component signal transduction (TCST) system present only in Enterobacteriaceae and positively regulates amylovoran biosynthesis by activating the *ams* operon transcription [25–27]. AmyR may regulate the Rcs phosphorelay system, potentially through interactions with RcsA or RcsB, the two response regulator proteins of the Rcs TCST, by preventing RcsB phosphorylation or interfering with RcsA DNA binding, or acting on the upstream transducers RcsC and RcsD [17,26,27]. However, while

*amyR* and *rscC* mutants show very similar changes in gene expression, *in vivo* infection assays show that *rscC* mutants are non-pathogenic while *amyR* mutants have similar virulence to wild-type strains, indicating that the activity of *amyR* is more complex than simply inhibiting the Rcs TCST system [16,26].

Recently the protein structure of a distantly related ybjN protein, DR1245 from *Deinococcus radiodurans*, was published [28]. The crystal structure of DR1245 revealed a fold similar to that of T3SS chaperones (T3C), despite not being implicated in any T3SS role. A structural similarity search also identified another distantly related YbjN protein from the cyanobacteria *Synechococcus elongatus* (unpublished, PDB ID: 2PLG). DR1245 does not have significant sequence similarity to the YbjN proteins of Enterobacteriaceae, with only 10% protein sequence identity to AmyR. DR1245 is an interaction partner of DdrB, a novel single stranded DNA binding protein upregulated in response to radiation stress [28]. The *E. coli* *ybjN* gene has since been shown to improve radiation resistance [29], indicating a shared stress response role of the *ybjN* family throughout Eubacteria.

This study reports the crystal structure of *E. amylovora* AmyR, the first representative structure of a YbjN protein from an Enterobacteriaceae species, which is the most highly divergent and intensely studied YbjN sub-family. The results provide specific implications for the mechanism of AmyR function, as well as relevance to the evolution and conserved functions of YbjN proteins across Eubacteria.

## Materials and methods

### Cloning

The *amyR* gene (Genbank FN434113.1, 1374785–1375270) was PCR amplified from genomic DNA of *E. amylovora* strain CFBP1430, using the forward primer 5' – ATTATTCCATGGGTATGGTTTCACTGGTTCG–3' and the reverse primer 5' – ATTATTCTCGAGTTAATGCAGCAGGGAATGGCTG –3'. Primers were designed to introduce an NcoI restriction site at the 5' end and an XhoI restriction site at the 3' end of the gene sequence. The PCR product was purified using the QIAquick PCR Purification Kit (Qiagen, Germany). The PCR product and the plasmid pETM-11 were double digested with the restriction enzymes NcoI and XhoI (NEB, USA). The digested PCR product and plasmid were run on a 1% agarose gel, and the appropriate bands were excised and gel purified (Qiagen, Germany). The *amyR* sequence was ligated into the pETM-11 expression vector [30], which encodes for a six histidine tag sequence and the Tobacco Etch Virus protease (TEV) cleavage site before the *amyR* gene insertion. The resultant protein after TEV protease cleavage has two non-native amino acids (GA) at the N-terminus. The construct was propagated in *E. coli* NovaBlue cells (EMD4 Biosciences, Germany) and was purified using a DNA miniprep kit (Sigma, USA). The plasmid molecular weight was checked by agarose gel electrophoresis after single and double digestion and the correctness of the inserted region was verified by sequencing at Eurofins facilities (Eurofins Genomics, Italy). *E. coli* BL21 (DE3) chemically competent cells (EMD4 Biosciences, Germany) were transformed with the pETM-11-*amyR* construct to express the recombinant protein.

### Expression and purification of the native protein

An overnight pre-culture was grown in 20 ml 2xYT medium containing kanamycin (30 mg L<sup>-1</sup>) at 310 K with shaking at 220 rpm. Two 2.5 L baffled flasks containing 1 L of LB medium were inoculated with 5 ml of the pre-culture and grown at 310 K with shaking at 220 rpm until they reached an OD600 of 0.6. The cultures were then cooled in an ice water bath for 10 min. Expression was then induced with 0.2 mM IPTG, and the cultures were incubated overnight at 293 K with shaking at 220 rpm. After 16 hours the cells were pelleted by centrifugation at

4500g for 10 min at 277 K, the supernatant removed, and the cells were re-suspended in ice-cold buffer A (50 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole, 2 mM DTT, 3% (w/v) glycerol) at 4 ml per gram wet weight of cell pellet with 0.2 mg ml<sup>-1</sup> lysozyme (Sigma Aldrich, USA), 20 µg ml<sup>-1</sup> DNase (Sigma-Aldrich, USA) and a protease inhibitor cocktail (Sigma-Aldrich, USA). The cells were disrupted by sonication on ice for 90 s using 30 s cycles (15.6 MHz) at 2 min intervals. The lysate was cleared by centrifugation at 25,000g for 30 min at 277 K and the supernatant filtered with a 0.45 µm cellulose acetate filter. The lysate solution was loaded onto a 5 ml HisTrap HP column (GE Healthcare, Sweden) pre-equilibrated with buffer A at a flow rate of 3 ml min<sup>-1</sup>. The column was washed with 20 ml of buffer A and the bound protein was eluted with a gradient of buffer B (50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole) over 90 ml. The eluted AmyR peak was pooled in cellulose dialysis tubing (12.5 kDa MWCO) with 6His-tagged TEV protease at a 1:20 (TEV:AmyR) ratio, and dialysed overnight into buffer C (20 mM HEPES pH 7.5, 200 mM NaCl, 20 mM imidazole, 1 mM DTT) at 277 K to cleave the N-terminal 6His tag and remove the excess imidazole. The protein solution was then passed through a 5 ml HisTrap HP column pre-equilibrated with buffer C at 3 ml min<sup>-1</sup> to remove the TEV protease, uncleaved AmyR, cleaved 6His tag, and other contaminating proteins. The flow through was collected and concentrated using a Vivaspin 20 ultrafiltration unit (Sartorius, Germany) with a 10,000 Da MWCO membrane to a volume of about 2 ml. The sample was run through a Superdex S75 16/60 column (GE Healthcare, Sweden) equilibrated with buffer D (20 mM HEPES pH 7.5, 200 mM NaCl, 1 mM DTT) at a flow rate of 1 ml min<sup>-1</sup>, and the AmyR peak was collected. All purification steps were carried out at 277 K. Purity of recombinant AmyR was confirmed by SDS-PAGE electrophoresis.

### Expression and purification of the selenomethionine substituted protein

An overnight pre-culture was grown in 20 ml of minimal media (1x M9 salts, 2 mM MgSO<sub>4</sub>, 0.4% (w/v) glucose, 0.1 mM CaCl<sub>2</sub>, 1x Kao & Michayluk vitamins, 1x trace metals (0.1 mM MnCl<sub>2</sub>·4H<sub>2</sub>O, 0.1 mM FeCl<sub>3</sub>, 5 µM ZnSO<sub>4</sub>·7H<sub>2</sub>O, 5 µM CoCl<sub>2</sub>·6H<sub>2</sub>O, 5 µM CuCl<sub>2</sub>·2H<sub>2</sub>O, 5 µM NiCl<sub>2</sub>·6H<sub>2</sub>O, 5 µM H<sub>3</sub>BO<sub>3</sub>, 5 µM NaMoO<sub>4</sub>·2H<sub>2</sub>O)) with 30 mg L<sup>-1</sup> kanamycin at 310 K with shaking at 220 rpm. Two 2.5 L baffled flasks containing 1 L of minimal medium with 30 mg L<sup>-1</sup> kanamycin were inoculated with 5 ml of the pre-culture and grown at 310 K with shaking at 220 rpm until reaching an OD<sub>600</sub> of 0.5. The cultures were then supplemented with 100 mg of L-lysine, 100 mg of L-phenylalanine, 100 mg of L-threonine, 50 mg of L-isoleucine, 50 mg of L-leucine, 50 mg of L-valine, and 50 mg of L-selenomethionine to inhibit methionine synthesis and incorporate the selenomethionine [31]. The cultures were incubated at 293 K with shaking at 220 rpm for 20 minutes, then protein expression was induced with 0.2 mM IPTG and the cultures were grown for a further 16 h.

The cells were harvested and the protein extracted and purified following the method described for the native protein, with the addition of 10 mM DTT to all buffers. Selenomethionine incorporation was confirmed by ESI mass spectrometry.

### Crystallisation

The purified native AmyR was concentrated to 10 mg ml<sup>-1</sup> in buffer D. Protein concentration was determined by direct UV measurement at 280 nm on a NanoVue spectrophotometer (GE Healthcare, Sweden) using an extinction coefficient of 12,490 M<sup>-1</sup> cm<sup>-1</sup>, calculated by ProtParam (ExPASy) [32]. Crystallisation trials were performed using the microbatch-under-oil method in 96-wells MRC plates (Molecular Dimensions, UK) using volatile oil (MD2-06, Molecular Dimensions, UK) at 293 K and 277 K. Drops of 1 µl of protein solution were equilibrated against the same volume of commercially available crystallization kits PACT, CSS1,

CSS2, JCSG+, MIDAS, and MORPHEUS (Molecular Dimensions, UK). Crystals grew in the condition JCSG+ B10, which consists of 0.2 M magnesium chloride, 0.1 M sodium cacodylate pH 6.5, 50% (v/v) PEG 200. This crystallisation condition was optimised by sitting-drop vapour-diffusion, using 0.5  $\mu$ l of protein solution with 0.5  $\mu$ l of precipitant and 100  $\mu$ l reservoirs in MRC plates (Molecular Dimensions, UK), with varying pH, temperature, and protein, salt and PEG concentrations. Optimum crystals grew at 293 K 4–8 days with a protein concentration of 8 mg ml<sup>-1</sup> using 0.1 M sodium cacodylate pH 6.9, 0.2 M magnesium chloride, 46% (v/v) PEG 200 at 293 K. Crystals were harvested from the plate using cryoloops (Hampton Research, USA) and flash cooled in liquid nitrogen with cryoprotection provided by the crystallisation condition.

Crystallisation of the selenomethionine substituted protein was attempted under the same conditions as the native protein, but only produced showers of small crystals. The conditions were optimised by the same method described for the native protein. Optimum crystals grew at 293 K with a protein concentration of 8 mg ml<sup>-1</sup> and the precipitant condition 0.1 M sodium cacodylate-Tris pH 7.3, 0.3 M magnesium chloride, 42% (v/v) PEG 200 at 293 K. The sodium cacodylate-Tris buffer was prepared by titrating 1 M unbuffered Tris into 1 M sodium cacodylate pH 6.9 until reaching the desired pH of 7.3. This condition produced crystals of a similar quality to those observed for the native protein, and were collected and flash cooled in the same manner.

## Diffraction data collection and processing

X-ray diffraction data of the native protein were collected to a resolution of 1.95 Å on beamline I04-1 at Diamond Light Source (DLS), Didcot, UK. Data for MAD phasing were collected from the selenomethionine substituted protein to 2.12 Å resolution on beamline I04 at DLS. Data were processed using XDS [33] within XIA2 [34]. Data collection statistics are provided in Table 1. The Se substructure was determined using SHELX [35]. Programs of the CCP4 suite [36] were used for structure solution and model refinement. Building of the structure was aided by automated procedures in Buccaneer [37] and manual building was performed with COOT [38]. The selenomethionine substituted protein was used as a search model for molecular replacement with the native data using PHASER [39]. The structure was refined with REFMAC5 [40], and validated with MolProbity [41]. Baverage was used to extract mean B factors values [36]. Figures of the structures were prepared with CCP4MG [42]. Refinement statistics are provided in Table 2.

## Bioinformatics

Similar protein structures to AmyR were identified using the protein structure comparison service PDBeFold from the European Bioinformatics Institute (<http://www.ebi.ac.uk/msd-srv/ssm>) [43] and proteins structure alignments performed in COOT [38]. Similar protein sequences to AmyR were identified by NCBI protein BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) [44] and the NCBI Conserved Domain Database (CDD) [45]. The CDD was used to generate a list of 100 aligned sequences from the protein family “YbjN putative bacterial sensory transducer regulator” using the type selection “the most diverse members”, and the sequences were exported in mFasta format. Sequences of larger proteins containing YbjN domains were removed, and the sequences of *E. amylovora* AmyR, *E. coli* YbjN, *D. radiodurans* DR1245 and *S. elongates* T110839 were added for comparison. The protein sequences were aligned and visualised using M-Coffee from the Centre for Genomic Regulation of Barcelona (<http://tcoffee.crg.cat/apps/tcoffee/index.html>) [46,47] and ESPript 3.0 (<http://esprict.ibcp.fr>) [48]. A phylogenetic tree was built using Clustal Omega [49].

**Table 1. Data collection statistics for native and selenomethionine substituted AmyR.** Values in parentheses are for the highest resolution shell.

	Native	Selenomethionine		
		Peak	Inflection Point	High Energy Remote
Diffraction source	DLS I04-1	DLS I04	DLS I04	DLS I04
Wavelength (Å)	0.91730	0.9794	0.9796	0.9740
Temperature (K)	100	100	100	100
Detector	Dectris Pilatus 6M	Dectris Pilatus 6M	Dectris Pilatus 6M	Dectris Pilatus 6M
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2
<i>a</i> , <i>b</i> , <i>c</i> (Å)	137.16, 49.82, 57.79	138.03, 49.62, 57.81	138.03, 49.62, 57.81	138.03, 49.62, 57.81
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution limit (Å)	53.26–1.95 (2.14–1.95)	53.34–2.13 (2.19–2.13)	53.34–2.12 (2.18–2.12)	49.66–2.22 (2.28–2.22)
Completeness	99.7 (99.6)	97.8 (79.0)	97.4 (76.2)	99.8 (98.5)
Multiplicity	5.0 (5.0)	24.7 (17.8)	24.6 (17.5)	25.4 (20.0)
<i>I</i> / $\sigma$	9.2 (1.8)	23.2 (1.4)	24.0 (1.4)	17.7 (1.3)
R <sub>meas</sub> ( <i>I</i> )	0.080 (1.230)	0.124 (2.282)	0.110 (2.238)	0.136 (2.474)
R <sub>meas</sub> ( <i>I</i> +/-)		0.091 (2.242)	0.089 (2.203)	0.125 (2.411)
CC half		0.999 (0.512)	1.000 (0.582)	0.999 (0.512)
Wilson B factor	37.15	47.36	46.34	48.48
Anomalous completeness		97.6 (78.3)	97.2 (75.4)	99.8 (98.1)
Anomalous multiplicity		13.0 (9.0)	12.9 (8.9)	13.4 (10.3)
Anomalous correlation		0.957 (0.032)	0.927 (0.024)	0.802 (0.030)
Total observations	147243 (34524)	553915 (23088)	557484 (22227)	516612 (28673)
Total unique	29587 (6947)	22462 (1298)	22681 (1272)	20357 (1434)

<https://doi.org/10.1371/journal.pone.0176049.t001>

## Results and discussion

### Crystallisation and structure solution

Two different crystal forms developed in the conditions used for crystallisation; rectangular crystals and rhombic plates (S1 Fig), with the thicker rectangular crystals regularly producing better diffraction quality crystals. No structures of similar sequence were identified in the protein data bank, and all attempts at solving the native protein structure by molecular replacement failed to produce a solution, including using a type III secretion chaperone model, CesT (PDB: 1K3E), identified by homology modelling in ROSETTA [50].

**Table 2. Refinement statistics for native and selenomethionine substituted AmyR.**

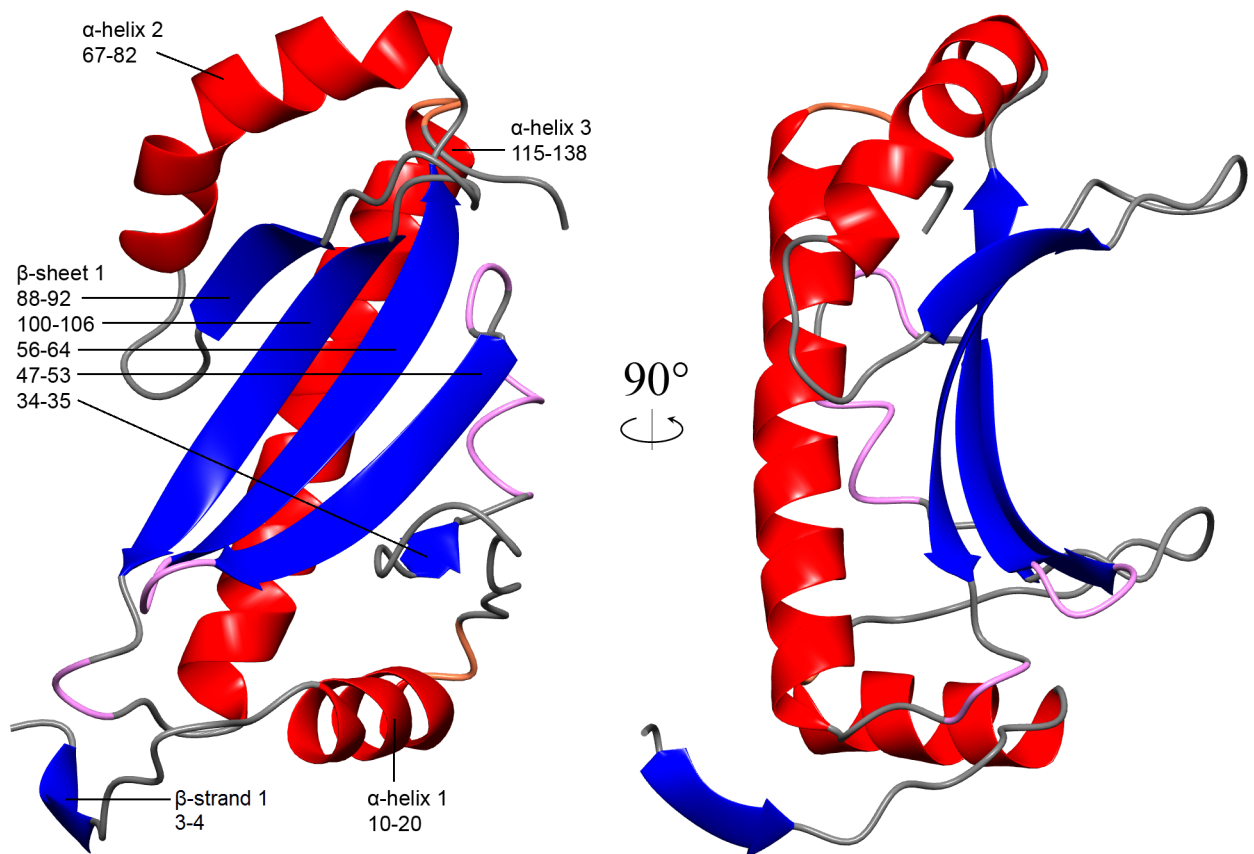
	Native	Selenomethionine
R <sub>work</sub>	0.224	0.215
R <sub>free</sub>	0.269	0.274
RMS deviation from ideal values		
Bond lengths (Å)	0.016	0.016
Bond angles (°)	1.797	1.833
Average B-factor (Å <sup>2</sup> )		
Protein atoms	42.82	50.88
Water molecules	43.49	49.71
Ramachandran plot		
Favoured (%)	284 (99.0)	283 (98.6)
Allowed (%)	1 (0.3)	4 (1.4)
Outliers (%)	2 (0.7)	0 (0)

<https://doi.org/10.1371/journal.pone.0176049.t002>

The crystal structure of AmyR substituted with selenomethionine was solved at 2.12 Å resolution by multiple wavelength anomalous diffraction (PDB: 5FRK), and native AmyR was determined at 1.95 Å by molecular replacement (PDB: 5FR7) using the selenomethionine structure as a search model. Both crystals were orthorhombic and belonged to space group  $P 2_12_12$ .

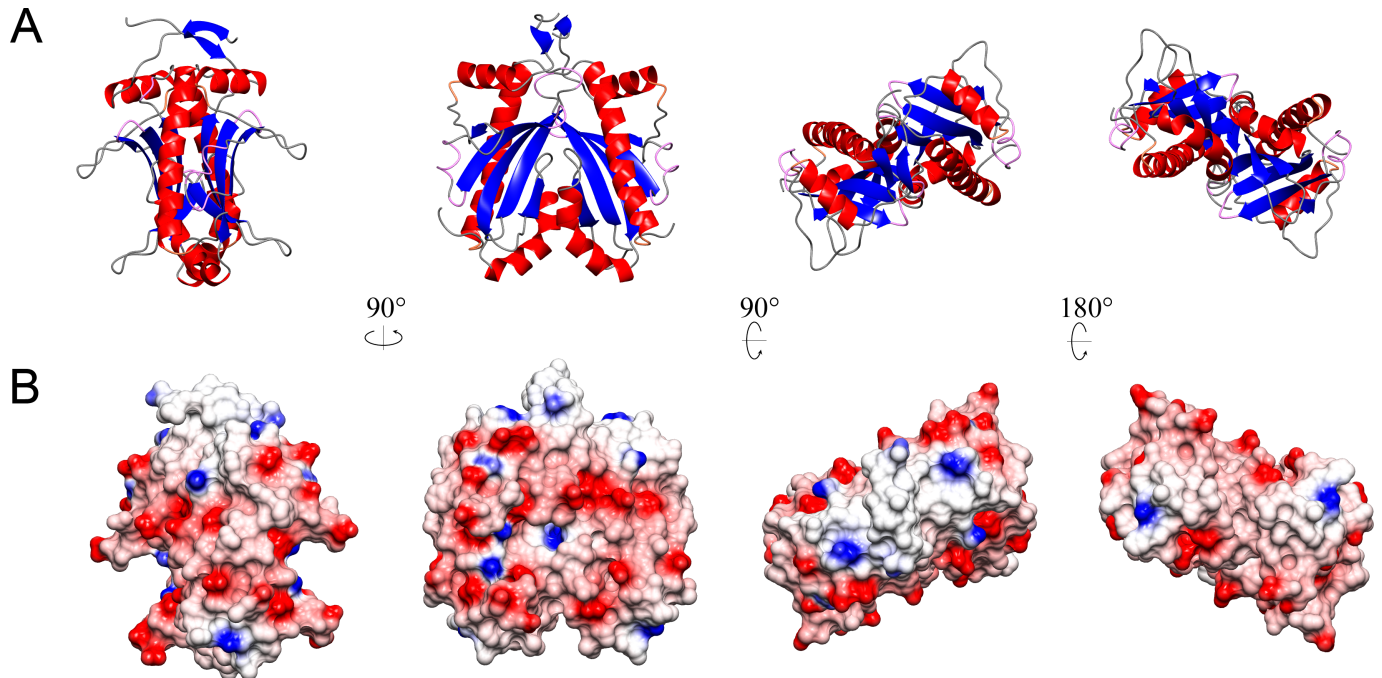
### Overall structure of AmyR

AmyR forms a homodimer in the asymmetric unit with a solvent content of 60.12% ( $V_m = 3.08 \text{ \AA}^3/\text{Da}$ ). Residues 1 to 144 of chain A and 1 to 145 of chain B were modelled and proved an excellent fit to the density. The remaining C-terminal regions were mostly unstructured, lacked defined density, and could not be modelled. The protein contains 3  $\alpha$ -helices (residues 10–20; 67–82; 115–138) surrounding a 5 stranded antiparallel  $\beta$ -sheet (residues 34–35; 47–53; 56–64; 88–92; 100–106) with the concave surface facing the solvent and flanked by two protruding loops (residues 27–33; 93–99) to create a channel (Figs 1 & 2). Flanking loop<sub>27-33</sub> is stabilised by a disulphide bridge between C28 and C31 which is coordinated by the side-chain of H35 from the  $\beta$ -strand (Fig 3). This stabilises the loop, which otherwise lacks coordinating interactions, resulting in a relatively rigid structure excluding residue S30 at the tip of the loop. The density for the disulphide bridge does not fit perfectly and it is likely some cysteines in the crystal were reduced, but the density favours the majority forming the bridge.



**Fig 1. Ribbon representation of the overall structure of the AmyR monomer.** The AmyR monomer is coloured by secondary structure. Secondary structure elements are labelled with residue numbers. The right panel is rotated 90°.  $\beta$ -strand 1 interacts with the equivalent residues in the dimer, shown in Fig 2.

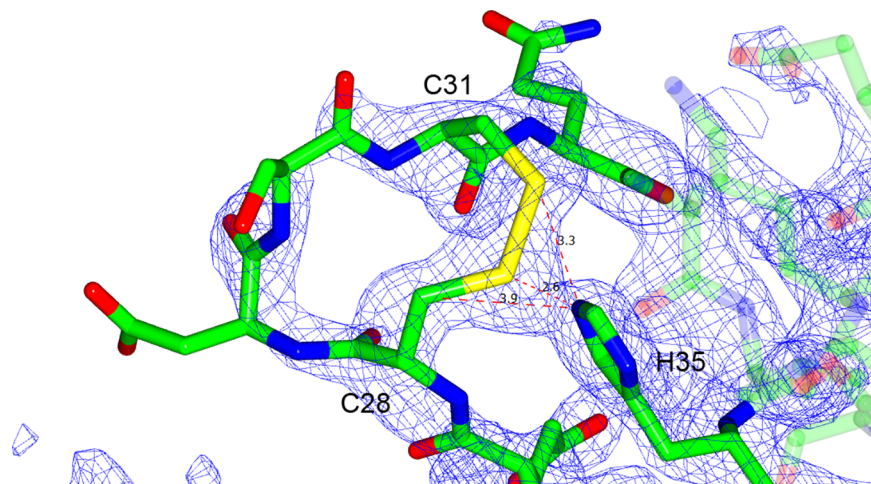
<https://doi.org/10.1371/journal.pone.0176049.g001>



**Fig 2. Rotated views of the *E. amylovora* AmyR dimer.** (A) Ribbon diagram, coloured by secondary structure. (B) Surface topography, coloured by surface charge at pH 7.4, red = negative, blue = positive.

<https://doi.org/10.1371/journal.pone.0176049.g002>

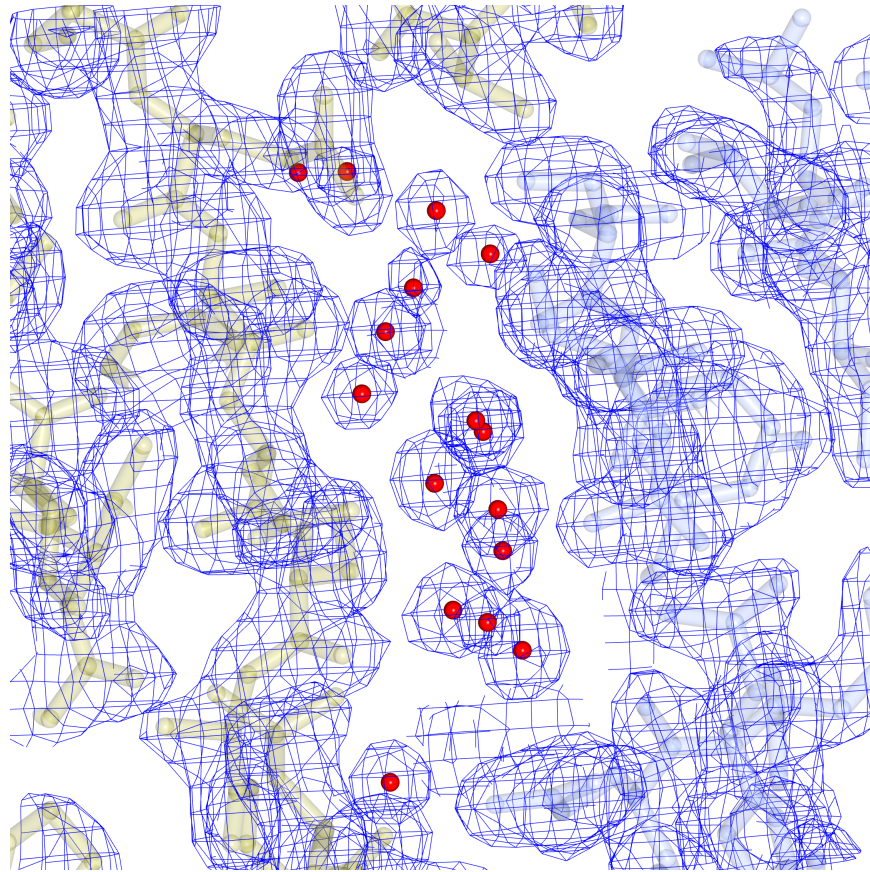
Flanking loop<sub>93-99</sub> on the other side of the  $\beta$ -sheet forms a crystal contact at its tip between D95 and T68 of the adjacent molecule, resulting in it being highly coordinated in the described structure. However, loop<sub>93-99</sub> otherwise lacks coordinating interactions, so in solution could be quite flexible and capable of folding over to partially encase the channel formed by the  $\beta$ -sheet and flanking loops. Half of the dimer interface is characterized by hydrophobic interactions between the bent  $\alpha$ -helix<sub>67-82</sub> and  $\beta$ -strand<sub>88-92</sub> of each chain, while the other half of the



**Fig 3. Coordination of residues in the  $\beta$ -sheet flanking loop.** The loop of residues 27 to 33, showing the disulphide bond between C28 and C31 coordinated with H35. The electron density shows the coordinated region of the loop is well ordered, while the tip remains flexible. Atomic distances between H35NE2 and the disulphide bridge are indicated in angstroms.

<https://doi.org/10.1371/journal.pone.0176049.g003>





**Fig 4. The water molecules between the dimer interface.** Cross-section of the hydrophilic pocket at the dimer interface, showing the A and B chains coloured in gold and silver, with the highly ordered water molecules contained within.

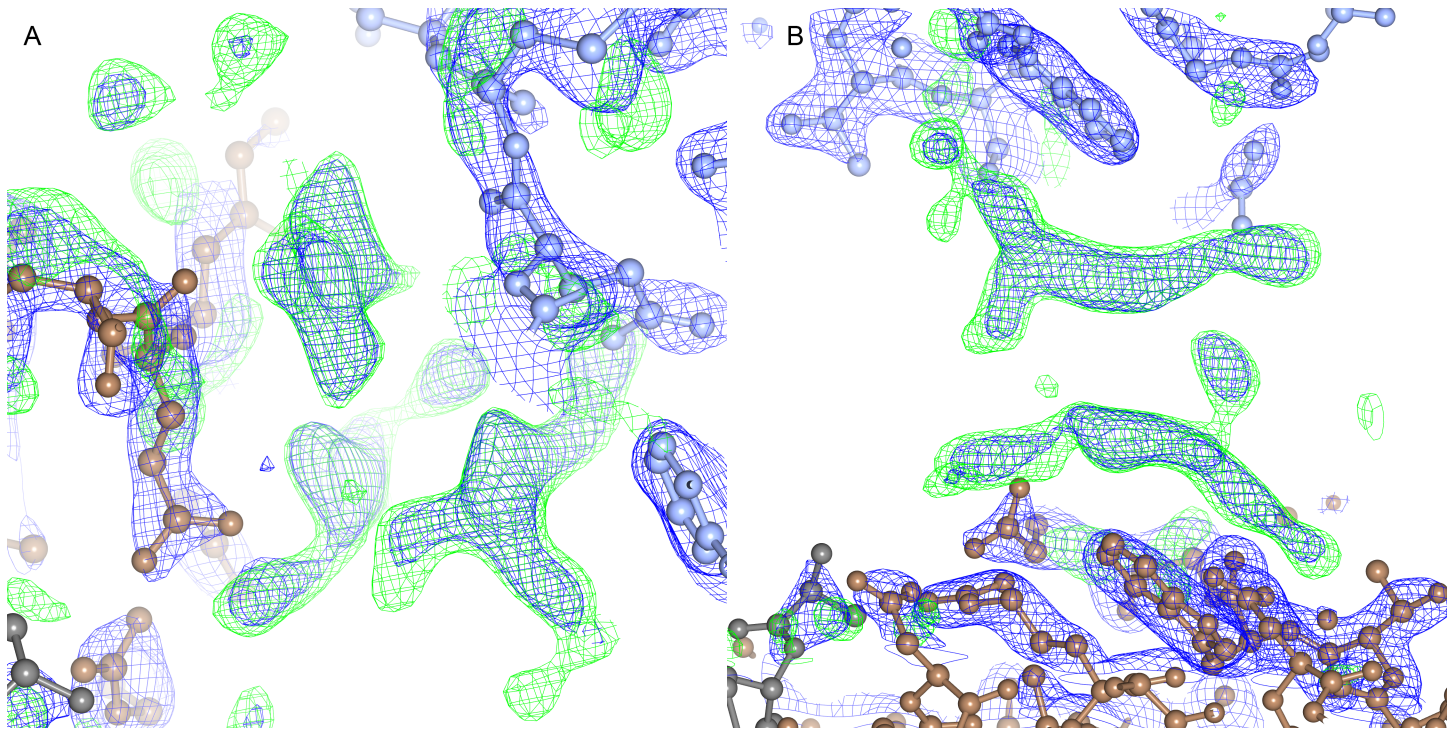
<https://doi.org/10.1371/journal.pone.0176049.g004>

interaction face forms a hydrophilic pocket containing 18 highly ordered water molecules (Fig 4), and is capped by hydrogen bonding between the two N-terminal tails.

The carboxy-terminal 17 amino acids of AmyR form an unstructured chain, accounting for 10% of the protein sequence. The electron density is observed protruding into the solvent channel from both A and B chains of the dimer, and can be followed with decreasing accuracy for roughly 4 residues before the density becomes completely ambiguous. The chain appears to reconnect with the main structure to form a crystal contact between residues 26 to 28 of chain A with the equivalent residues of chain B on the symmetry related molecule, but density could not be unambiguously modelled because of the short length of the observed density without any clearly defined structural features (Fig 5).

### YbjN proteins share a fold with T3SS chaperone proteins but appear to function independently of the T3SS

A structural homology search with PDBeFold [43] showed the closest similar structures to AmyR to be *D. radiodurans* DR1245, *S. elongatus* T110839, and class I T3C proteins. Structural alignment shows how highly conserved the domain architecture is (Fig 6). There is some variation in the length and orientation of the periphery loops and secondary structural elements, but the core domain structures are almost identical. The main variation is the orientation of the dimer interaction, rather than the protein fold itself.

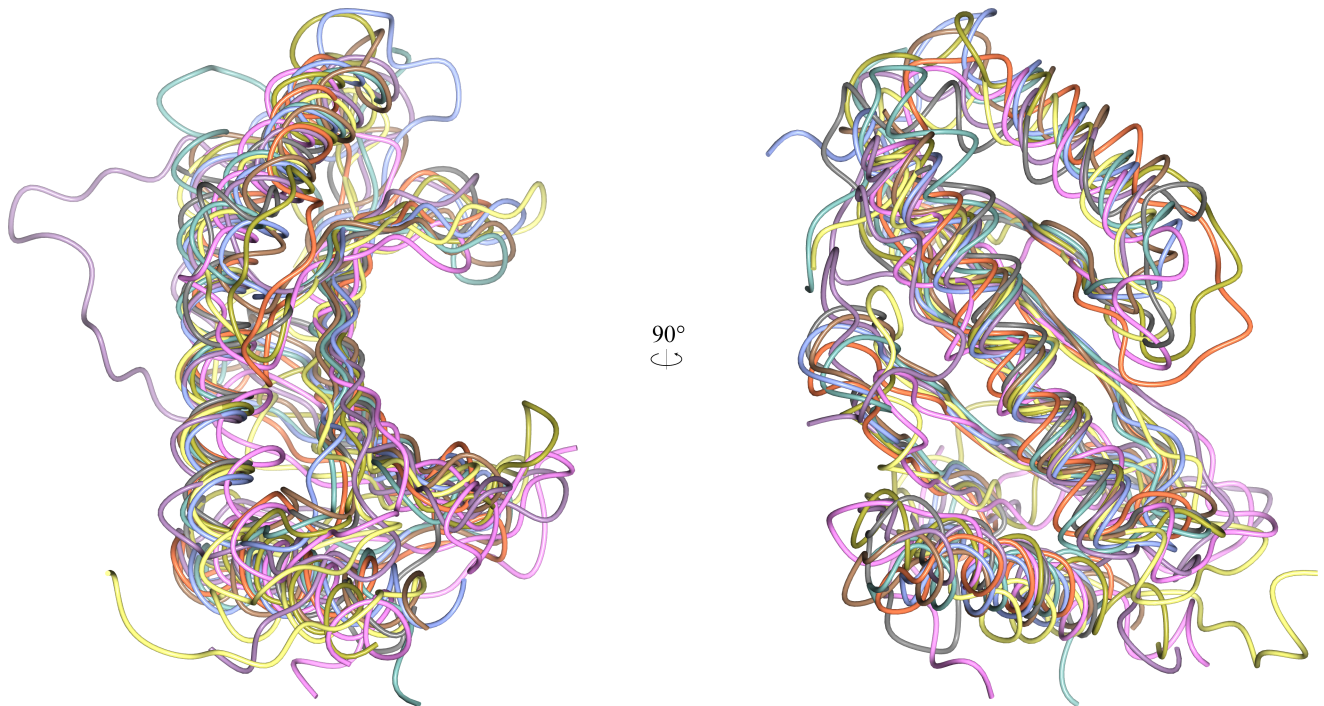


**Fig 5. Unmodelled density at crystal contact between molecules.** (A) Electron density and difference map of the unmodelled density at the crystal contact. The main crystal contact forming density is central, while the two other chains are seen below. (B) An alternative view of the two chains.

<https://doi.org/10.1371/journal.pone.0176049.g005>

Despite the structural similarity with the class I T3C proteins, the genomic location of *amyR* does not conform with a class I T3C. The genes for class IA T3Cs are located adjacent to the genes of the type III effector which they partner with [51]. In contrast, the next downstream gene from *amyR* in the *E. amylovora* genome is *potF* (genbank CBA20255.1) [52,53]. This gene is annotated to code a periplasmic receptor component of the PotFGHI putrescine ABC transporter, with high similarity to other such proteins in Enterobacteriaceae species, including the *E. coli* homologue with 78% protein sequence identity, which has been structurally and functionally characterised [54]. Upstream of *amyR* is the *mdaA* gene, coding for an oxygen-insensitive nitroreductase [52,53]. Neither PotF or MdaA are T3SS associated proteins. PotF is targeted to the periplasm, which is bypassed by the T3SS, while MdaA is cytoplasmic. Therefore, AmyR cannot act as a class IA T3C for PotF or MdaA. Class IB T3Cs are located in T3SS related operons [55], which again does not apply to *amyR*.

To confirm this is a common feature of *ybjN* genes, the genomic locations of *ybjN* genes from 13 additional species were investigated. Species were selected on the criteria of containing both a *ybjN* gene and a T3SS, and that the whole genome sequence was available, including virulence plasmid sequences where applicable. Assuming that the genomic organisation of closely related species would be similar, the selection was limited to one member of any given Genus. The species *Escherichia coli*, *Yersinia pestis*, *Salmonella enterica*, *Shigella flexneri*, *Vibrio parahaemolyticus*, *Serratia plymuthica*, *Burkholderia glumae*, *Chlamydia trachomatis*, *Sodalis praecaptivus*, *Ralstonia insidiosus*, *Rhizobium phaseoli*, *Photorhabdus asymbiotica*, and *Desulfovibrio vulgaris* were included. In no case were any *ybjN* genes located in T3SS gene clusters, or adjacent to T3SS associated proteins (as some effector proteins are not located in T3SS operons, e.g. *E. amylovora* AvrRpt2). These results show that *ybjN* genes occupy distinct genomic locations from T3Cs.



**Fig 6. Structural alignment of AmyR monomer with other YbjN and T3C proteins.** The two views are rotated horizontally 90°. The alignment includes the PDB entries 1JYO:D (*Salmonella enterica* SicP, ice blue), 1S28:A (*Pseudomonas savastanoi* AvrPphF ORF1, gold), 1TTW:A (*Yersinia pestis* SycH, coral), 1XKP:B (*Yersinia pestis* SycN, grey), 2PLG:B (*Synechococcus elongates* T110839, pink), 3EPU:A (*Salmonella enterica* STM2138, sea green), 3KXY:J (*Escherichia coli* ExsC, brown), and 4H5B:A (*Deinococcus radiodurans* DR\_1245, lilac), and 5FR7:A (*Erwinia amylovora* AmyR, yellow).

<https://doi.org/10.1371/journal.pone.0176049.g006>

The phenotypes observed by varying AmyR expression cannot be explained purely by the function of a T3C. Secretion chaperones facilitate secretion and/or stabilise cytoplasmic T3SS effector and gate proteins. The chaperoning effect should be the same when expressed sufficiently or to excess. While an *amyR* knockout mutation would be expected to have an observable effect, the overexpression of AmyR results in regulatory changes in *E. amylovora*, which cannot be explained by AmyR functioning solely as a T3C [16]. None of the proteins secreted via the T3SS have been shown to regulate *E. amylovora* exopolysaccharide levels, and T3SS mutants can still produce wild-type levels of exopolysaccharide [56]. Furthermore, *amyR* knockouts show increased virulence, rather than decreased, as would be expected for a T3C [16].

Some class I T3Cs have shown transcriptional regulation activity along with chaperoning of T3SS proteins, such as *Shigella flexneri* Spa15 [57], *Pseudomonas aeruginosa* ExsC [58], and *Pseudomonas syringae* HrpG [59]. Unlike AmyR and other YbjN proteins, the currently identified class I T3Cs with transcriptional regulation activity are located in T3SS operons, their interactions are with T3SS associated proteins, and their regulatory activity is restricted to T3SS gene clusters [57–60]. One notable exception is *Chlamydia trachomatis* Scc4 (CT663), which can trigger global changes in gene expression by directly interacting with RNA polymerase to inhibit  $\sigma^{66}$ -dependent transcription [61,62]. However, Scc4 is classified in the CesT superfamily of T3Cs [45], genetically distinct from YbjN proteins, and is located in a cluster of T3SS associated genes in the bacterial genome [63].

The functions of related proteins also support YbjN proteins not performing a T3SS role, as YbjN proteins have been generally characterised as stress-response sensory transduction

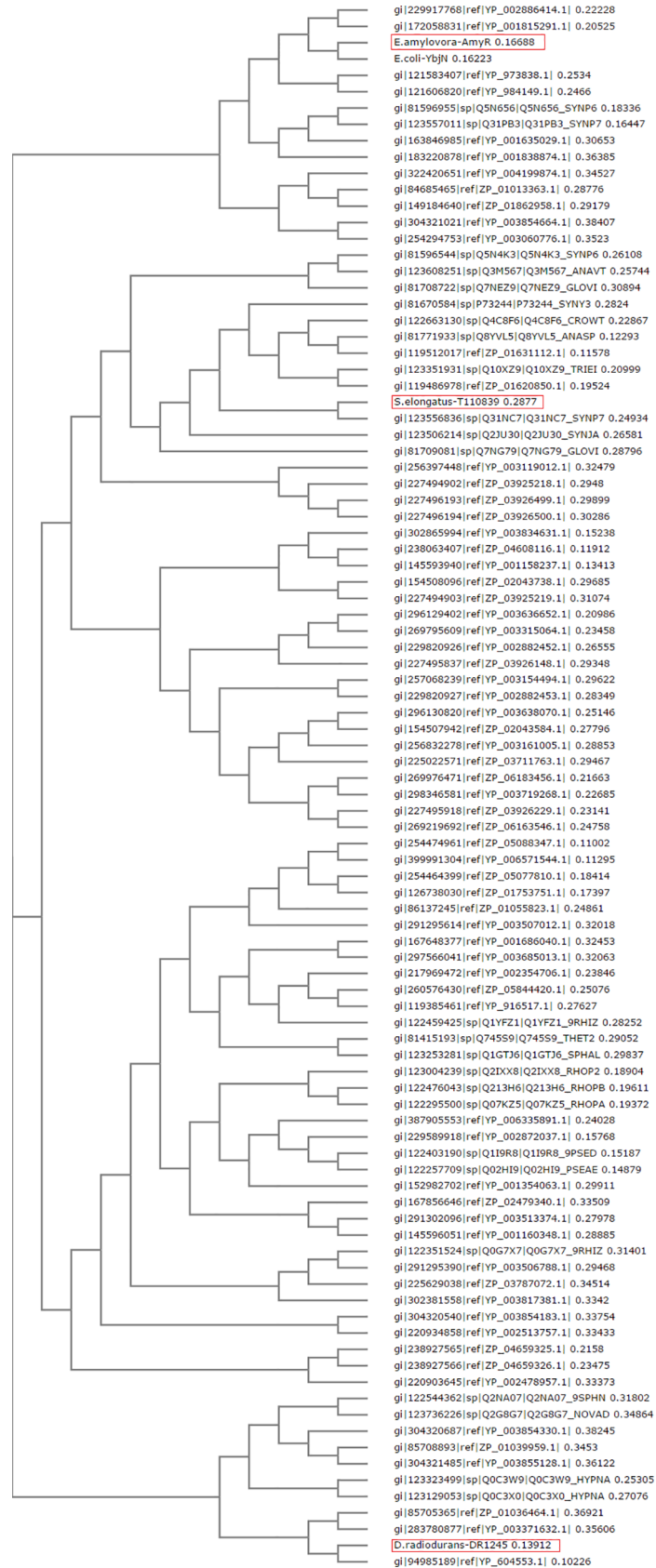
regulators. The reported functional analysis of the *E. coli* homolog YbjN has no relation to the T3SS [15,17]. The structural homolog DR1245 from *D. radiodurans* also performs functions and interactions entirely unrelated to infection or the T3SS [28]. Despite the clear structural homology, none of these proteins show sequence homology to T3SS chaperones, while *E. amylovora* AmyR and *E. coli* YbjN have 67% sequence identity and are functionally conserved [16,17]. Furthermore, the CDD shows that YbjN family proteins are present throughout Eubacteria, including in species which are non-pathogenic and lack a T3SS [45]. Even though pairwise alignments of distantly related YbjN proteins might lack sequence similarity, the CDD shows through multiple alignment that YbjN proteins classify into a distinct protein family, which excludes T3Cs. While a T3SS related role cannot be completely ruled out for all YbjN proteins, at this time there is no evidence for YbjN proteins, including AmyR, to function in conjunction with the T3SS. Therefore, the shared fold between YbjN proteins and T3Cs suggests a shared mechanism and ancestral origin while performing different cellular functions.

### Identification of conserved residues and regions in YbjN proteins

The sequences of 100 YbjN family proteins were aligned from a diverse selection of species to identify conserved regions or specific residues. The alignment showed consensus regions among species that were more closely evolutionarily related, but no overall consensus sequence for YbjN proteins. The Enterobacteriaceae species were the most divergent of the sequences, with *E. amylovora* and *E. coli* sharing the lowest identity score against the average of all sequences, while the majority of other sequences rated between 40 and 52. The sequences from Cyanobacteria made up the remainder of sequences with low identity to the consensus, and also lacking similarity to the Enterobacteriaceae sequences represent an independent variation from the YbjN consensus. A phylogenetic tree of the protein sequences (Fig 7) shows that the three presently determined YbjN protein structures (AmyR, DR1245, and T110839) are placed in each of the three major divisions of the tree. Therefore, these three structures represent three of the most divergent members of the YbjN protein family, with almost no sequence similarity between them, yet the alignment of their structures shows how extremely conserved the protein fold is within the YbjN family (Fig 6).

The M-Coffee alignment of the YbjN protein sequences (S2 Fig) revealed two regions which were the most highly conserved, which correlated to residues 13–24 and 75–88 of the AmyR sequence, while ESPript identified 21 highly conserved positions (S3 Fig) correlating to residues L14, L18, H38, L77, S78, I80, N81, T86, K88, F90, L91, L101, L103, L107, A109, L113, T114, F118, M122, I129, and A136. Residue N97 on the  $\beta$ -sheet flanking loop<sub>93-99</sub> was identified but shown to be incorrectly assigned, because this loop is significantly longer in AmyR than for most YbjN proteins, and the correct alignment was with L101.

Almost all the conserved residues had internalised side chains mediating the hydrophobic interactions between secondary structure elements in the core of the protein and hydrophobic dimerization interface (Fig 8A). There will only be limited amino acids that could adopt these positions with side chains capable of fitting in the hydrophobic spaces between secondary structure elements without significantly altering their orientation, given the high structural conservation of the fold adopted by YbjN and T3C proteins. The few conserved hydrophilic residues also mediate structural interactions with minimal solvent exposure. The most highly conserved region was residues 75–88, which also contained around one quarter of the specific residues identified. This mapped to  $\alpha$ -helix<sub>67-82</sub> and the subsequent loop, which mediate the hydrophobic dimerization interface (Fig 8B). The other conserved region in the alignment, between residues 13–24, seemed to be poorly conserved in Enterobacteriaceae species, and the



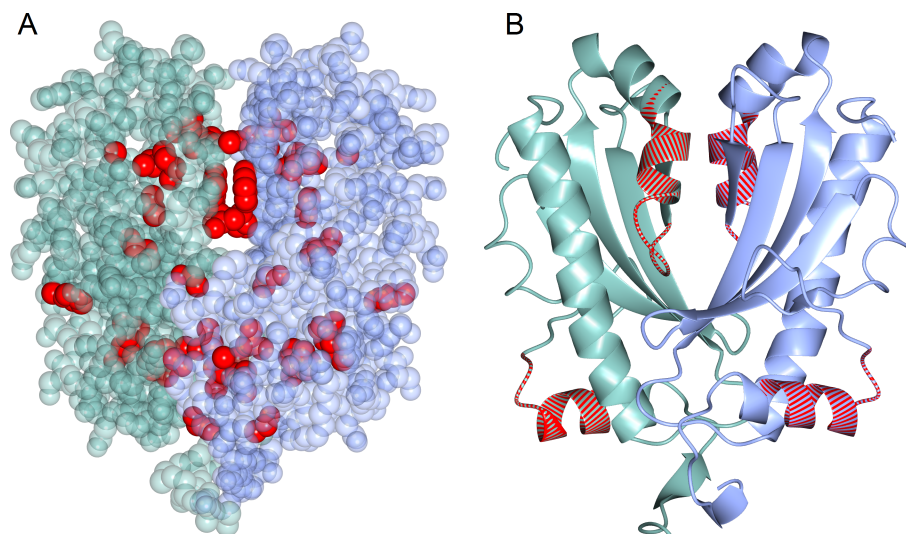
**Fig 7. Phylogenetic tree of 100 YbjN protein sequences selected as described in methods.** The three sequences with protein structures are boxed in red.

<https://doi.org/10.1371/journal.pone.0176049.g007>

AmyR sequence was a poor match in this region. This region mapped to the C-terminal end of the first  $\alpha$ -helix of the AmyR structure, which was not predicted to be involved in interactions. The two most highly conserved amino acids in this region were internally facing hydrophobic residues, with side chains facing the internal side of the  $\beta$ -sheet.

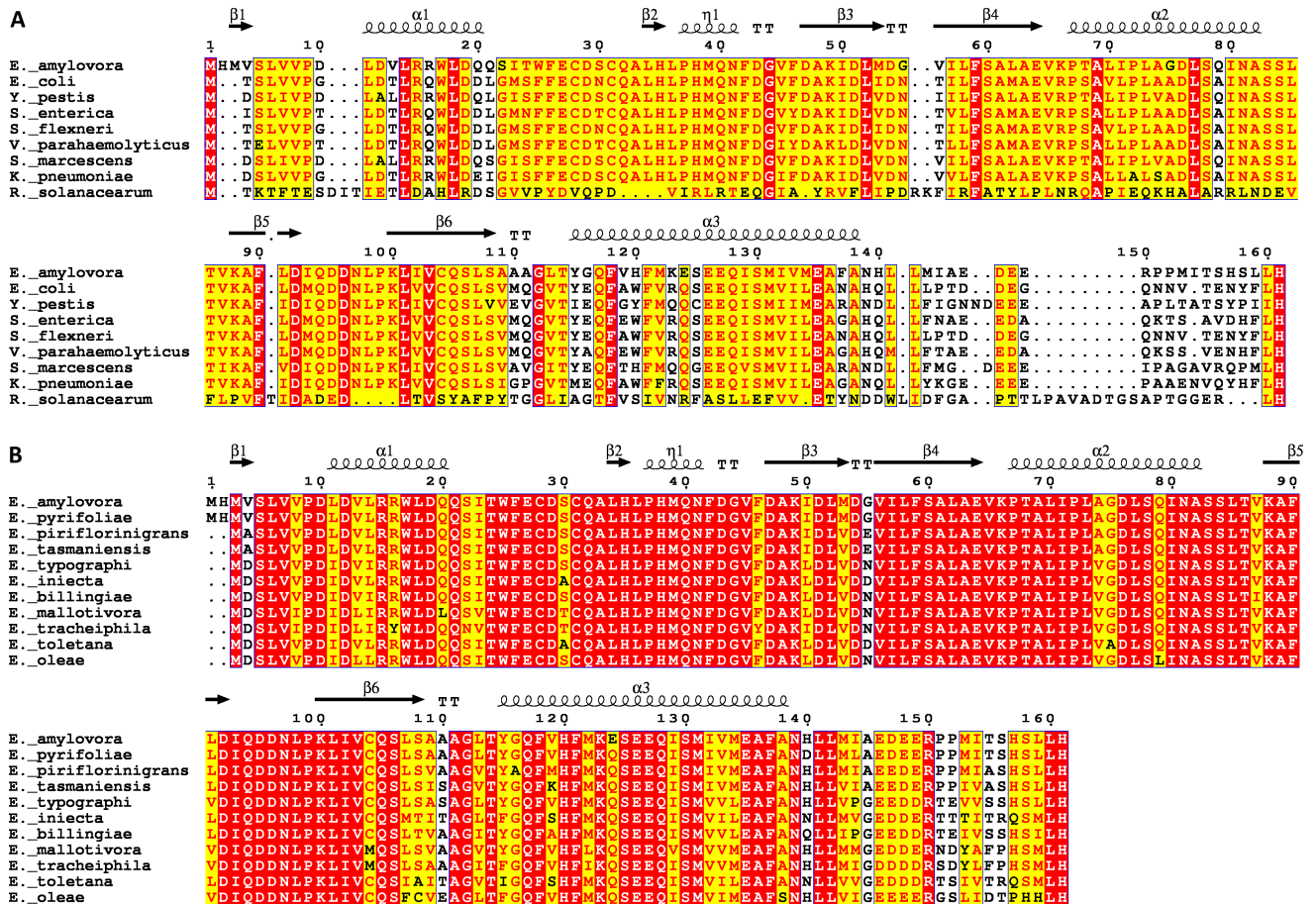
This analysis shows that there are no conserved surface areas in YbjN proteins across the entire protein family. This corresponds well with the idea that YbjN proteins, like T3Cs, share a common fold and common mechanism of action, but interact with different target proteins which are determined by the solvent exposed surface, particularly on and around the  $\beta$ -sheet. Therefore, closely related proteins like *E. amylovora* AmyR and *E. coli* YbjN are functionally conserved, while other distantly related YbjN proteins are active through different specific interactions.

To better understand the important regions of Enterobacteriaceae specific YbjN proteins, an alignment was performed with protein sequences from 7 other Enterobacteriaceae species selected for their importance to infectious disease (*Escherichia coli*, *Yersinia pestis*, *Salmonella enterica*, *Shigella flexneri*, *Vibrio parahaemolyticus*, *Serratia marcescens*, *Klebsiella pneumoniae*), and one Beta Proteobacteria as an outlier (*Ralstonia solanacearum*) (Fig 9A). This highlighted the high sequence conservation within Enterobacteriaceae, with all sequences showing between 64% to 74% identity with *E. amylovora* AmyR. The main area of divergence was the C-terminal 24 amino acids, accounting for the tip of the last  $\alpha$ -helix, proceeding loop, and unstructured tail in the AmyR structure. However, a conserved Leu-His motif was observed at the terminal two amino acids. The presence of disordered C-termini in similar structures indicates this is a conserved feature in YbjN proteins. There were also minor divergences in



**Fig 8. Location of conserved regions and residues of YbjN proteins superimposed on *E. amylovora* AmyR structure.** (A) Sphere representation of AmyR displaying the two chains of the dimer in grey and green, with the atoms of conserved sidechains shown in red Van der Waals spheres. The main chains are partially transparent to show that the conserved residue side chains are mostly internalised or at the dimerization interface. (B) The same view as a ribbon diagram, with the YbjN protein conserved regions in red.

<https://doi.org/10.1371/journal.pone.0176049.g008>



**Fig 9. Sequence alignments of YbjN proteins.** (A) Alignments of YbjN protein sequences from Enterobacteriaceae species, *Erwinia amylovora* (NCBI accession CBA20254); *Escherichia coli* (AJE55158); *Yersinia pestis* (ABP40742); *Salmonella enterica* (WP\_000624814); *Shigella flexneri* (ABF03066); *Vibrio parahaemolyticus* (KKY40576); *Serratia marcescens* (OKP20636); and Beta Proteobacteria *Ralstonia solanacearum* (WP\_042591288). (B) Alignment of YbjN protein sequences from *Erwinia* species, *E. amylovora* (CBA20254); *E. pyrifoliae* (CAX56090); *E. piriflorinigrans* (WP\_023654581); *E. tasmaniensis* (WP\_012441893); *E. typographi* (WP\_034891635); *E. iniecta* (WP\_052896595); *E. billingiae* (WP\_013201489); *E. mallotivora* (WP\_034941268); *E. tracheiphila* (WP\_016189698); *E. toletana* (WP\_017802422); *E. oleae* (WP\_034945603).

<https://doi.org/10.1371/journal.pone.0176049.g009>

residues 10–21 and 110–125, corresponding to  $\alpha$ -helix 1, and the preceding loop and start of  $\alpha$ -helix 3, respectively.

Alignment of the YbjN protein sequences from 11 *Erwinia* species also only showed significant variance in the unstructured region of the C-terminus, with nearly identical folded domain sequences (Fig 9B). Since there is not significant variation in the protein interaction regions of YbjN proteins in Enterobacteriaceae, their interaction partners are also expected to be homologous. Differences in activity between Enterobacteriaceae species will be a result of variation in the downstream transcriptional regulation pathways, rather than the direct activity of the YbjN proteins. However, in more divergent bacteria this may not be the case.

### Protein interactions

Given their shared architecture, YbjN proteins will most likely interact with partner proteins in the same manner as Class I T3Cs. These function through protein-protein interactions by binding unfolded chains and secondary structure elements of partially or fully folded peptides

in an ATP-independent manner [64,65]. Binding is strong and specific, and effector release is coupled to the activity of an independent ATPase at the T3SS gate [66].

The only reports of the observed fold in complex with interaction partners have been for T3Cs, co-crystallised with either effector proteins [67,68] or T3SS gate proteins [69]. The interaction is usually observed as the wrapping of unstructured or partially structured amino acid chains around the chaperone dimer perpendicular to the axis of the dimerization interface (PDB examples: 1TTW [67]; 1JYO [70], but interactions with folded domains are also possible (PDB examples: 2FM8 [71]; 1XKP [69]). Analysis of the surface charges of AmyR show many negatively charged groups surrounding the dimer in the plane of chain wrapping observed in the T3C structures (Fig 2B), but hydrophobic interactions are also likely to be important for a chaperone-like protein. Surface grooves are oriented to accommodate interactions with peptide chains in the direction similar to T3Cs, including the internal groove of the beta-sheet and flanking loops (Fig 2A), which is capable of fitting  $\alpha$ -helical assemblies or unfolded chains.

Class I T3Cs are divided into two subclasses, class IA which associates with a single effector protein, and class IB which are able to associate with several effectors [55]. Both subclasses share the same fold and cannot be distinguished by comparing structural features. There was some indication that class IB chaperones adopt a wider dimerization conformation with a reduced interface [72]. However, there is significant variation at the dimer interface within class IA, class IB and YbjN proteins, from the increasing number of structures now available, and insufficient evidence to classify YbjN proteins as targeting single or multiple interaction partner proteins based on quaternary structure [65,73]. If AmyR has more than one interaction partner it could interact with the Rcs phosphorelay system, since matching changes in gene regulation are observed between *amyR* and *rcsC* mutants, but also interact with one or more other pathways, resulting in the observed phenotypic differences between *amyR* and *rcsC* mutants [16,26,27].

The functional significance of the disordered C-terminus is unknown. The sequence is not well conserved even in closely related species, yet the presence and relative length is conserved, along with the C-terminal LH motif (Fig 9). Its function may be similar to the P23 chaperone where specific binding interactions are determined by the folded domain, while the unstructured C-terminus only assists in non-specific interactions of proteome stabilization [74]. Alternatively, the conserved terminal LH residues could be a binding motif on a flexible arm, similar to the disordered C-terminus of the DnaK chaperone [75].

## Gene regulation by AmyR through specific interactions

The overexpression of AmyR under a non-native promoter showed that the changes in transcriptional levels of other genes are a direct result of the presence of the AmyR protein, and not just co-regulation of the AmyR promoter [16]. Transcriptional regulation may occur through an inhibitory binding, a stabilising binding, or a combination of the two, with one or more target proteins. Cases of some class I T3Cs provide examples of how the observed chaperone fold can regulate transcription through interactions with partner proteins, either by direct interaction with a transcriptional regulator (eg. *P. syringae* chaperone HrpG suppresses the negative regulator HrpV [59]), indirect interaction with a transcriptional regulator via a co-activator/antiactivator (e.g. *S. flexneri* Spa15 chaperone acts as a co-antiactivator in complex with OspD1 to sequester the transcriptional activator MxiE [57]), or to sequester an inhibitor of a transcriptional regulator (e.g., *P. aeruginosa* ExsC chaperone sequesters the antiactivator ExsD from the ExsD-ExsA complex, releasing and activating the transcriptional regulator ExsA [58]). In the case of AmyR, regulatory pathway proteins could be bound and sequestered in a dose dependent manner, explaining the observed variations between the wild-type, knock-out and over expression *amyR* mutants [16].



Chaperone-like binding can also stabilise the target proteins, increasing their accumulation and subsequent activity. In mutation studies of the Enterobacteria *Yersinia*, *Salmonella* and *Shigella*, non-secreting mutants are shown to accumulate effector proteins in the cytoplasm. However, when the class I T3C associated with an effector protein is also knocked out the effector accumulation is reduced, despite equivalent transcript levels [55,76–82]. The inherent instability of the effectors in the cytoplasm appears to be caused by their chaperone binding region, the removal of which can stabilise the effector protein in the absence of the chaperone [83]. In the same manner AmyR could stabilise transcriptional regulatory pathway proteins or inhibitors, determining their abundance through association, and explaining the dose dependent transcriptional regulation observed for AmyR. Furthermore, the currently studied YbjN proteins are expressed in response to stress, and specific stabilization could have the dual effect of increasing the protein accumulation and protecting the target protein from stresses such as heat shock and radiation [15,28,29].

## Conclusions

AmyR from *E. amylovora* is a member of the YbjN family, which have been implicated in regulation of the stress-response in bacteria. The structure of AmyR was solved at 1.95 Å resolution, revealing a class I type III secretion (T3C) chaperone-like fold. Functional studies have previously shown AmyR as an essential regulator of virulence factors in *E. amylovora*. Class I T3C-type proteins such as AmyR are capable of forming very specific interactions with one or more partner proteins. Therefore, the crystal structure presented in this study points to a mechanism where AmyR could accomplish virulence regulation by either binding and blocking partner protein activity, or through the stabilisation of partner proteins promoting cytoplasmic accumulation. Knowledge of the fold, and consequently possible modes of action, of AmyR will guide future investigation into the virulence regulation mechanisms of *E. amylovora*.

The structure of AmyR is the first of a YbjN protein from Enterobacteriaceae, the most evolutionarily divergent and highly studied subgroup of YbjN proteins, and the first from the phylum Proteobacteria. The structures of YbjN proteins show exceptionally high conservation of tertiary structure, despite highly divergent primary structure. The conserved sequence features from YbjN proteins were identified and mapped to the AmyR structure. These revealed that only residues involved in the stabilisation of both the protein core and the dimerization interface were conserved, while surface exposed residues completely lacked similarity across the YbjN family. However, within closely related bacterial groupings there was very high sequence similarity even at the putative surface-exposed interaction sites, indicating that related YbjN proteins, such as those within the Enterobacteriaceae family, interact with homologous proteins.

Despite the shared fold with class I T3Cs, the current evidence suggests that YbjN proteins are not implicated in a T3SS associated role. This includes their genomic location not matching that of T3Cs, their presence in species lacking a T3SS, their classification as an independent protein family from T3Cs, the regulation of their expression being associated with stress rather than infection, their transcriptional regulation activity not being associated with T3SS genes, and specifically for AmyR, knockout mutants showing increased virulence rather than decreased. While a T3SS associated role cannot be ruled out for all YbjN proteins, at this time there is no evidence of YbjN proteins performing a T3SS associated role, despite their shared architecture with class I T3Cs. The presented structural and phylogenetic analyses of YbjN proteins provides a foundation for future studies into this poorly understood but increasingly important protein family.

## Supporting information

**S1 Fig. The two observed crystal forms of AmyR.**

(PDF)

**S2 Fig. MCoffee alignment of 100 YbjN protein sequences.**

(PDF)

**S3 Fig. ESPript alignment of 100 YbjN protein sequences.**

(PDF)

## Acknowledgments

Plasmid pETM-11 was obtained from the European Molecular Biology Laboratory (EMBL) under a signed Material Transfer Agreement. The authors would like to thank Diamond Light Source for beamtime, and the staff of beamlines I04 and I04-1 for assistance with data collection. J.B thanks DLS and the Research complex at Harwell for hosting him as a visiting scientist whilst performing part of this work.

## Author Contributions

**Conceptualization:** JDB SB.

**Funding acquisition:** SB.

**Investigation:** JDB DB JW ND MT.

**Project administration:** SB.

**Supervision:** SB DB MAW.

**Validation:** JDB DB ND SB.

**Writing – original draft:** JDB.

**Writing – review & editing:** JDB DB JW ND MT AOS YZ MAW SB.

## References

1. Vanneste JL. Fire Blight: The Disease and its Causative Agent, *Erwinia amylovora*. Wallingford: CABI; 2000.
2. Vrancken K, Holtappels M, Schoofs H, Deckers T, Valcke R. Pathogenicity and infection strategies of the fire blight pathogen *Erwinia amylovora* in Rosaceae: State of the art. *Microbiology* 2013 May 01; 159(Pt 5):823–832. <https://doi.org/10.1099/mic.0.064881-0> PMID: 23493063
3. Wei Z, Kim JF, Beer SV. Regulation of hrp genes and type III protein secretion in *Erwinia amylovora* by HrpX/HrpY, a novel two-component system, and HrpS. *Mol Plant-Microbe Interact* 2000; 13(11):1251–1262. <https://doi.org/10.1094/MPMI.2000.13.11.1251> PMID: 11059492
4. Koczan JM, McGrath MJ, Zhao Y, Sundin GW. Contribution of *Erwinia amylovora* exopolysaccharides amyovorin and levan to biofilm formation: Implications in pathogenicity. *Phytopathology* 2009; 99(11):1237–1244. <https://doi.org/10.1094/PHYTO-99-11-1237> PMID: 19821727
5. Koczan JM, Lenneman BR, McGrath MJ, Sundin GW. Cell surface attachment structures contribute to biofilm formation and xylem colonization by *Erwinia amylovora*. *Appl Environ Microbiol* 2011; 77(19):7031–7039. <https://doi.org/10.1128/AEM.05138-11> PMID: 21821744
6. Gross M, Geier G, Rudolph K, Geider K. Levan and levansucrase synthesized by the fireblight pathogen *Erwinia amylovora*. *Physiol Mol Plant Pathol* 1992; 40(6):371–381.
7. Caputi L, Nepogodiev SA, Malnoy M, Rejzek M, Field RA, Benini S. Biomolecular characterization of the levansucrase of *Erwinia amylovora*, a promising biocatalyst for the synthesis of fructooligosaccharides. *J Agric Food Chem* 2013; 61(50):12265–12273. <https://doi.org/10.1021/jf4023178> PMID: 24274651

8. Caputi L, Cianci M, Benini S. Cloning, expression, purification, crystallization and preliminary X-ray analysis of EaLsc, a levansucrase from *Erwinia amylovora*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2013; 69(5):570–573.
9. Wuerges J, Caputi L, Cianci M, Boivin S, Meijers R, Benini S. The crystal structure of *Erwinia amylovora* levansucrase provides a snapshot of the products of sucrose hydrolysis trapped into the active site. *J Struct Biol* 2015; 191(3):290–298. <https://doi.org/10.1016/j.jsb.2015.07.010> PMID: 26208466
10. Bernhard F, Coplin DL, Geider K. A gene cluster for amylovoran synthesis in *Erwinia amylovora*: characterization and relationship to cps genes in *Erwinia stewartii*. *Molec Gen Genet* 1993; 239(1–2):158–168. PMID: 8389975
11. Metzger M, Bellemann P, Bugert P, Geider K. Genetics of galactose metabolism of *Erwinia amylovora* and its influence on polysaccharide synthesis and virulence of the fire blight pathogen. *J Bacteriol* 1994; 176(2):450–459. PMID: 7507102
12. Bugert P, Gelder K. Molecular analysis of the ams operon required for exopolysaccharide synthesis of *Erwinia amylovora*. *Mol Microbiol* 1995; 15(5):917–933. PMID: 7596293
13. Bellemann P, Geider K. Localization of transposon insertions in pathogenicity mutants of *Erwinia amylovora* and their biochemical characterization. *J Gen Microbiol* 1992; 138(5):931–940. <https://doi.org/10.1099/00221287-138-5-931> PMID: 1322951
14. Zhao Y, Wang D, Nakka S, Sundin GW, Korban SS. Systems level analysis of two-component signal transduction systems in *Erwinia amylovora*: Role in virulence, regulation of amylovoran biosynthesis and swarming motility. *BMC Genomics* 2009; 10.
15. Chen X, Shen D, Zhou B. Analysis of the temperature-sensitive mutation of *Escherichia coli* pantothenate kinase reveals YbjN as a possible protein stabilizer. *Biochem Biophys Res Commun* 2006; 345(2):834–842. <https://doi.org/10.1016/j.bbrc.2006.04.101> PMID: 16701556
16. Wang D, Korban SS, Pusey PL, Zhao Y. AmyR Is a Novel Negative Regulator of Amylovoran Production in *Erwinia amylovora*. *PLoS ONE* 2012; 7(9).
17. Wang D, Calla B, Vimolmangkang S, Wu X, Korban SS, Huber SC, et al. The orphan gene ybjN conveys pleiotropic effects on multicellular behavior and survival of *Escherichia coli*. *PLoS ONE* 2011; 6(9).
18. Hayashi S, Aono R, Hanai T, Mori H, Kobayashi T, Honda H. Analysis of organic solvent tolerance in *Escherichia coli* using gene expression profiles from DNA microarrays. *J Biosci Bioeng* 2003; 95(4):379–383. PMID: 16233423
19. Bore E, Hébraud M, Chafsey I, Chambon C, Skjæret C, Moen B, et al. Adapted tolerance to benzalkonium chloride in *Escherichia coli* K-12 studied by transcriptome and proteome analyses. *Microbiology* 2007; 153(4):935–946.
20. Keeney D, Ruzin A, Mcaleese F, Murphy E, Bradford PA. MarA-mediated overexpression of the AcrAB efflux pump results in decreased susceptibility to tigecycline in *Escherichia coli*. *J Antimicrob Chemother* 2008; 61(1):46–53. <https://doi.org/10.1093/jac/dkm397> PMID: 17967850
21. Piqué N, Miñana-Galbis D, Merino S, Tomás JM. Virulence factors of *Erwinia amylovora*: A review. *Int J Mol Sci* 2015; 16(6):12836–12854. <https://doi.org/10.3390/ijms160612836> PMID: 26057748
22. Chang H-, Lee J-, Deng W-, Fu T-, Peng H-. Virulence and outer membrane properties of a galU mutant of *Klebsiella pneumoniae* CG43. *Microb Pathog* 1996; 20(5):255–261. <https://doi.org/10.1006/mpat.1996.0024> PMID: 8861391
23. Frey PA. The Leloir pathway: A mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. *FASEB J* 1996; 10(4):461–470. PMID: 8647345
24. Toccafondi M, Cianci M, Benini S. Expression, purification, crystallization and preliminary X-ray analysis of glucose-1-phosphate uridylyltransferase (GalU) from *Erwinia amylovora*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2014; 70(9):1249–1251.
25. Bernhard F, Poetter K, Geider K, Coplin DL. The rcsA gene from *Erwinia amylovora*: identification, nucleotide sequence, and regulation of exopolysaccharide biosynthesis. *Mol Plant Microbe Interact* 1990; 3(6):429–437. PMID: 2131100
26. Wang D, Korban SS, Zhao Y. The Rcs phosphorelay system is essential for pathogenicity in *Erwinia amylovora*. *Mol Plant Pathol* 2009; 10(2):277–290. <https://doi.org/10.1111/j.1364-3703.2008.00531.x> PMID: 19236575
27. Wang D, Qi M, Calla B, Korban SS, Clough SJ, Cock PJA, et al. Genome-wide identification of genes regulated by the rcs phosphorelay system in *Erwinia amylovora*. *Mol Plant-Microbe Interact* 2012; 25(1):6–17. <https://doi.org/10.1094/MPMI-08-11-0207> PMID: 21936662
28. Norais C, Servant P, Bouthier-de-la-Tour C, Coureux P-, Ithurbe S, Vannier F, et al. The *Deinococcus radiodurans* DR1245 Protein, a DdrB Partner Homologous to YbjN Proteins and Reminiscent of Type III Secretion System Chaperones. *PLoS ONE* 2013; 8(2).

29. Byrne RT, Chen SH, Wood EA, Cabot EL, Cox MM. *Escherichia coli* genes and pathways involved in surviving extreme exposure to ionizing radiation. *J Bacteriol* 2014; 196(20):3534–3545. <https://doi.org/10.1128/JB.01589-14> PMID: 25049088
30. Dümmler A, Lawrence A-, de Marco A. Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microb Cell Fact* 2005; 4.
31. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* 1993; 229(1):105–124. <https://doi.org/10.1006/jmbi.1993.1012> PMID: 7678431
32. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, De Castro E, et al. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res* 2012; 40(W1):W597–W603.
33. Kabsch W. XDS. *Acta Crystallogr Sect D Biol Crystallogr* 2010; 66(2):125–132.
34. Winter G. Xia2: An expert system for macromolecular crystallography data reduction. *J Appl Crystallogr* 2010; 43(1):186–190.
35. Sheldrick GM. Experimental phasing with SHELXC/D/E: Combining chain tracing with density modification. *Acta Crystallogr Sect D Biol Crystallogr* 2010; 66(4):479–485.
36. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D Biol Crystallogr* 2011; 67(4):235–242.
37. Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr Sect D Biol Crystallogr* 2006; 62(9):1002–1011.
38. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr* 2010; 66(4):486–501.
39. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr* 2007; 40(4):658–674.
40. Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, et al. REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr Sect D Biol Crystallogr* 2004; 60(12 I):2184–2195.
41. Chen VB, Arendall WB III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sect D Biol Crystallogr* 2010; 66(1):12–21.
42. McNicholas S, Potterton E, Wilson KS, Noble MEM. Presenting your structures: The CCP4mg molecular-graphics software. *Acta Crystallogr Sect D Biol Crystallogr* 2011; 67(4):386–394.
43. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr Sect D Biol Crystallogr* 2004; 60(12 I):2256–2268.
44. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; 215(3):403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: 2231712
45. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res* 2015; 43(D1):D222–D226.
46. Notredame C, Higgins DG, Heringa J. T-coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 2000; 302(1):205–217. <https://doi.org/10.1006/jmbi.2000.4042> PMID: 10964570
47. Moretti S, Armougoum F, Wallace IM, Higgins DG, Jongeneel CV, Notredame C. The M-Coffee web server: A meta-method for computing multiple sequence alignments by combining alternative alignment methods. *Nucleic Acids Res* 2007; 35(SUPPL.2):W645–W648.
48. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 2014; 42(W1):W320–W324.
49. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011; 7.
50. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* 2004 Jul 1; 32(Web Server issue):W526–31. <https://doi.org/10.1093/nar/gkh468> PMID: 15215442
51. Lohou D, Lonjon F, Genin S, Vaillau F. Type III chaperones & Co in bacterial plant pathogens: A set of specialized bodyguards mediating effector delivery. *Front Plant Sci* 2013; 4(NOV).
52. Sebahia M, Bocsanczy AM, Biehl BS, Quail MA, Perna NT, Glasner JD, et al. Complete genome sequence of the plant pathogen *Erwinia amylovora* strain ATCC 49946. *J Bacteriol* 2010; 192(7):2020–2021. <https://doi.org/10.1128/JB.00022-10> PMID: 20118253
53. Smits THM, Rezzonico F, Kamber T, Blom J, Goesmann A, Frey JE, et al. Complete genome sequence of the fire blight pathogen *Erwinia amylovora* CFBP 1430 and comparison to other *Erwinia* spp. *Molecular Plant-Microbe Interactions* 2010; 23(4):384–393. <https://doi.org/10.1094/MPMI-23-4-0384> PMID: 20192826

54. Vassilyev DG, Tomitori H, Kashiwagi K, Morikawa K, Igarashi K. Crystal structure and mutational analysis of the *Escherichia coli* putrescine receptor. Structural basis for substrate specificity. *J Biol Chem* 1998; 273(28):17604–17609. PMID: [9651355](#)
55. Parsot C, Hamiaux C, Page A-. The various and varying roles of specific chaperones in type III secretion systems. *Curr Opin Microbiol* 2003; 6(1):7–14. PMID: [12615213](#)
56. Nissinen RM, Ytterberg AJ, Bogdanove AJ, Van Wijk KJ, Beer SV. Analyses of the secretomes of *Erwinia amylovora* and selected hrp mutants reveal novel type III secreted proteins and an effect of HrpJ on extracellular harpin levels. *Molecular Plant Pathology* 2007; 8(1):55–67. <https://doi.org/10.1111/j.1364-3703.2006.00370.x> PMID: [20507478](#)
57. Parsot C, Ageron E, Penno C, Mavris M, Jamoussi K, d'Hauteville H, et al. A secreted anti-activator, OspD1, and its chaperone, Spa15, are involved in the control of transcription by the type III secretion apparatus activity in *Shigella flexneri*. *Mol Microbiol* 2005; 56(6):1627–1635. <https://doi.org/10.1111/j.1365-2958.2005.04645.x> PMID: [15916611](#)
58. Dasgupta N, Lykken GL, Wolfgang MC, Yahr TL. A novel anti-anti-activator mechanism regulates expression of the *Pseudomonas aeruginosa* type III secretion system. *Mol Microbiol* 2004; 53(1):297–308. <https://doi.org/10.1111/j.1365-2958.2004.04128.x> PMID: [15225323](#)
59. Wei C, Deng W, Huang H. A chaperone-like HrpG protein acts as a suppressor of HrpV in regulation of the *Pseudomonas syringae* pv. *syringae* type III secretion system. *Mol Microbiol* 2005; 57(2):520–536. <https://doi.org/10.1111/j.1365-2958.2005.04704.x> PMID: [15978082](#)
60. Büttner D. Protein export according to schedule: Architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. *Microbiology and Molecular Biology Reviews* 2012; 76(2):262–310. <https://doi.org/10.1128/MMBR.05017-11> PMID: [22688814](#)
61. Rao X, Deighan P, Hua Z, Hu X, Wang J, Luo M, et al. A regulator from *Chlamydia trachomatis* modulates the activity of RNA polymerase through direct interaction with the beta subunit and the primary sigma subunit. *Genes Dev* 2009 Aug 1; 23(15):1818–1829. <https://doi.org/10.1101/gad.1784009> PMID: [19651989](#)
62. Shen L, Macnaughtan MA, Frohlich KM, Cong Y, Goodwin OY, Chou CW, et al. Multipart Chaperone-Effector Recognition in the Type III Secretion System of *Chlamydia trachomatis*. *J Biol Chem* 2015 Nov 20; 290(47):28141–28155. <https://doi.org/10.1074/jbc.M115.670232> PMID: [26438824](#)
63. Stephens RS, Kalman S, Lammel C, Fan J, Marathe R, Aravind L, et al. Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* 1998 Oct 23; 282(5389):754–759. PMID: [9784136](#)
64. Wattiau P, Woestyn S, Cornelis GR. Customized secretion chaperones in pathogenic bacteria. *Mol Microbiol* 1996; 20(2):255–262. PMID: [8733225](#)
65. Fattori J, Prando A, Martins AM, Dos Santos Rodrigues FH, Tasic L. Bacterial secretion chaperones. *Protein Pept Lett* 2011; 18(2):158–166. PMID: [21121896](#)
66. Akeda Y, Galán JE. Chaperone release and unfolding of substrates in type III secretion. *Nature* 2005; 437(7060):911–915. <https://doi.org/10.1038/nature03992> PMID: [16208377](#)
67. Phan J, Tropea JE, Waugh DS. Structure of the *Yersinia pestis* type III secretion chaperone SycH in complex with a stable fragment of YscM2. *Acta Crystallogr Sect D Biol Crystallogr* 2004; 60(9):1591–1599.
68. Janjusevic R, Quezada CM, Small J, Erec Stebbins C. Structure of the HopA1(21–102)-ShcA chaperone-effector complex of *Pseudomonas syringae* reveals conservation of a virulence factor binding motif from animal to plant pathogens. *J Bacteriol* 2013; 195(4):658–664. <https://doi.org/10.1128/JB.01621-12> PMID: [23204470](#)
69. Schubot FD, Jackson MW, Penrose KJ, Cherry S, Tropea JE, Plano GV, et al. Three-dimensional structure of a macromolecular assembly that regulates type III secretion in *Yersinia pestis*. *J Mol Biol* 2005; 346(4):1147–1161. <https://doi.org/10.1016/j.jmb.2004.12.036> PMID: [15701523](#)
70. Stebbins CE, Galán JE. Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion. *Nature* 2001; 414(6859):77–81. <https://doi.org/10.1038/35102073> PMID: [11689946](#)
71. Lilic M, Vujanac M, Stebbins CE. A common structural motif in the binding of virulence factors to bacterial secretion chaperones. *Mol Cell* 2006; 21(5):653–664. <https://doi.org/10.1016/j.molcel.2006.01.026> PMID: [16507363](#)
72. van Eerde A, Hamiaux C, Pérez J, Parsot C, Dijkstra BW. Structure of Spa15, a type III secretion chaperone from *Shigella flexneri* with broad specificity. *EMBO Rep* 2004; 5(5):477–483. <https://doi.org/10.1038/sj.embor.7400144> PMID: [15088068](#)
73. Cooper CA, Zhang K, Andres SN, Fang Y, Kaniuk NA, Hannemann M, et al. Structural and biochemical characterization of SrcA, a multi-cargo type III secretion chaperone in *Salmonella* required for pathogenic association with a host. *PLoS Pathog* 2010; 6(2).

74. Weikl T, Abelmann K, Buchner J. An unstructured C-terminal region of the Hsp90 co-chaperone p23 is important for its chaperone function. *J Mol Biol* 1999; 293(3):685–691. <https://doi.org/10.1006/jmbi.1999.3172> PMID: 10543959
75. Smock RG, Blackburn ME, Gierasch LM. Conserved, disordered C terminus of DnaK enhances cellular survival upon stress and DnaK in vitro chaperone activity. *J Biol Chem* 2011; 286(36):31821–31829. <https://doi.org/10.1074/jbc.M111.265835> PMID: 21768118
76. Wattiau P, Bernier B, Deslée P, Michiels T, Cornelis GR. Individual chaperones required for Yop secretion by *Yersinia*. *Proc Natl Acad Sci U S A* 1994; 91(22):10493–10497. PMID: 7937981
77. Frithz-Lindsten E, Rosqvist R, Johansson L, Forsberg Å. The chaperone-like protein YerA of *Yersinia pseudotuberculosis* stabilizes YopE in the cytoplasm but is dispensable for targeting to the secretion loci. *Mol Microbiol* 1995; 16(4):635–647. PMID: 7476159
78. Persson C, Nordfelth R, Holmstrom A, Hakansson S, Rosqvist R, Wolf-Watz H. Cell-surface-bound *Yersinia* translocate the protein tyrosine phosphatase YopH by a polarized mechanism into the target cell. *Mol Microbiol* 1995; 18(1):135–150. PMID: 8596454
79. Fu Y, Galán JE. Identification of a specific chaperone for SptP, a substrate of the centisome 63 type III secretion system of *Salmonella typhimurium*. *J Bacteriol* 1998; 180(13):3393–3399. PMID: 9642193
80. Abe A, De Grado M, Pfuetzner RA, Sánchez-SanMartín C, DeVinney R, Puente JL, et al. Enteropathogenic *Escherichia coli* translocated intimin receptor, Tir, requires a specific chaperone for stable secretion. *Mol Microbiol* 1999; 33(6):1162–1175. PMID: 10510231
81. Cambronne ED, Cheng LW, Schneewind O. LcrQ/YscM1, regulators of the *Yersinia* yop virulon, are injected into host cells by a chaperone-dependent mechanism. *Mol Microbiol* 2000; 37(2):263–273. PMID: 10931323
82. Niebuhr K, Jouihri N, Allaoui A, Gounon P, Sansonetti PJ, Parsot C. IpgD, a protein secreted by the type III secretion machinery of *Shigella flexneri*, is chaperoned by IpgE and implicated in entry focus formation. *Mol Microbiol* 2000; 38(1):8–19. PMID: 11029686
83. Woestyn S, Sory M-, Boland A, Lequenne O, Cornelis GR. The cytosolic SycE and SycH chaperones of *Yersinia* protect the region of YopE and YopH involved in translocation across eukaryotic cell membranes. *Mol Microbiol* 1996; 20(6):1261–1271. PMID: 8809777