Structural and functional characterization of proteins from the fire blight pathogen *Erwinia amylovora*. A review on the state of the art

Stefano Benini¹

Received: 18 June 2020 / Accepted: 9 October 2020 / Published online: 2 November 2020
© The Author(s) 2020

Abstract

Together with genome analysis and knock-out mutants, structural and functional characterization of proteins provide valuable hints on the biology of the organism under investigation. Structural characterization can be achieved by techniques such as X-ray crystallography, NMR, Cryo-EM. The information derived from the structure are a good starting point to comprehend the details of the proteins molecular function for a better understanding of their biological role. This review aims at describing the progress in the structural and functional characterization of proteins from the plant pathogen *Erwinia amylovora* obtained by structural biology and currently deposited in the Protein Data Bank.

Keywords *Erwinia amylovora* · Protein structure · PDB · Pathogenicity · Enterobacteriaceae

Introduction

Erwinia amylovora, the plant pathogen responsible for the fire blight disease in Rosaceae still causes huge losses in apple and pear orchards worldwide (Vrancken et al. 2013). *E. amylovora* is one of the top ten plant pathogen for its impact on global apple and pear production (Mansfield et al. 2012).

In *E. amylovora* three main molecular systems are required for a successful infection. The type three secretion system (T3SS), that is necessary to secrete effector proteins and suppress the host organism's defenses (Vrancken et al. 2013). The exopolysaccharide biosynthetic pathway, producing amylovoran and levan, required to protect the bacteria and causing wilting of shoots and blight symptoms (Langlotz et al. 2011; Gross et al. 1992). The siderophore mediated iron uptake system, which scavenges the iron required for the pathogen metabolism from the host (Smits and Duffy 2011). Besides these, a series of virulence and pathogenicity factors are necessary for the pathogen to thrive (Borruso et al. 2017).

For a better understanding of the molecular basis regulating the biology and pathogenicity of an organism, it is worth to study in vitro the proteins involved in its life cycle. Genome analysis and knock-out mutants are an excellent starting point useful in identifying potential targets for further investigation. 3D structure determination coupled to biochemical characterization provides relevant details to study structure and function relationship of proteins and enzymes. Different experimental techniques are used to elucidate the structure of proteins, Nuclear Magnetic Resonance (NMR), X-ray crystallography, and Cryo-Electron Microscopy are still the most used techniques (Cavalli et al. 2007; Muench et al. 2019). The availability of sequenced genomes has allowed the development of “structural genomics” strategies. Structural genomics consortium have been created in order to explore and elucidate unknown protein folds, infer proteins function from their structure by comparison, and deposit their results in the Protein Data Bank (PDB) using an high-throughput approach (Grabowski et al. 2016). The approach of integrative structural biology, aiming at modelling the structure of biological systems, makes use of data from different techniques (NMR, X-ray, homology modelling, molecular docking, Small Angle X-ray Scattering, etc.) providing information to understand cell biology also useful in drug discovery (Rout and Sali 2019). The availability of *E. amylovora* genome (Smits et al. 2010) has allowed to study the fire blight pathogen using different techniques. From genome analysis and comparison to identify genes (and therefore proteins) that are crucial in developing

✉ Stefano Benini
stefano.benini@unibz.it

¹ Bioorganic Chemistry and Bio-Crystallography Laboratory (B(2)CI), Faculty of Science and Technology, Libera Università di Bolzano, Piazza Università, 5, 39100 Bolzano, Italy

the disease, to the study of knock-out mutants and phenotype analysis, and to the cloning of genes with production of proteins to be studied in vitro.

This review aims at describing the progress in the structural and functional characterization of the proteins from the plant pathogen *E. amylovora* whose structure has been obtained to date in chronological order since publication and release from the PDB.

The Protein Data Bank

The Protein Data Bank (PDB <http://www.rcsb.org/>) (Berman et al. 2003; Burley et al. 2019) collects, validates and offers public access to experimental data of 3D structures of biological macromolecules (proteins, DNA, and RNA). The availability of the 3D structures enables researchers belonging to different fields, to study the function, and by comparison, the evolution of biological macromolecules, to enhance knowledge in fundamental biology, biomedicine, and biotechnology. The PDB is currently archiving 165,117 entries (06/15/2020) and Table 1 collects the entries related to *E. amylovora*.

RcsB

The gene of the regulator of exopolysaccharide synthesis RcsB from *E. amylovora* was first cloned by Bereswill et al. in 1997 to investigate its role in modulating exopolysaccharide synthesis. Its overexpression in strains containing high-copy number plasmids carrying the *rscB* gene, caused an increased amylovoran production while suppressing levansucrase activity (Bereswill and Geider 1997). It has been demonstrated that RcsB forms a heterodimer with RcsA to form a protein-DNA complex acting cooperatively to bind to the promoter of the *ams* operon responsible for amylovoran biosynthesis. More precisely the specific binding of the complex RcsA-RcsB is to a region upstream of the

translation initiation codon of *amsG*, the first gene in the *ams* operon in *E. amylovora* (Kelm et al. 1997; Wehland et al. 1999). The importance of the Rcs phosphorelay system in regulating amylovoran production and its influence on virulence and pathogenicity were proved by gene knockout experiments and site directed mutagenesis (Wang et al. 2009; Ancona et al. 2015). The 24 kDa RcsB protein is composed of two domains, the N-terminal domain contains a phosphorylation motif with three aspartic acid (Asp10, Asp11 and Asp56), while the C-terminal domain features a helix-turn-helix (HTH) DNA binding motif (residues 151–194) homologous to LuxR-type regulators (Henikoff et al. 1990). The solution structure of the C-terminal DNA-binding domain of RcsB (C-RcsB, residues 129–215) was determined by heteronuclear magnetic resonance spectroscopy (NMR) in 2003. Beside the structure determination, Pristovsek et al. carried out the characterization of the interaction between RcsB and one of its target DNA, the RcsAB box, and the identification of the residues involved by NMR chemical shift perturbation mapping (Pristovsek et al. 2003). Twenty energy-minimized NMR structures were selected to represent the ensemble of possible conformations in the native state in solution. The structure of C-RcsB features five helices (from $\alpha 6$ to $\alpha 10$) of which the central ones $\alpha 8$ and $\alpha 9$ are part of the HTH motif and interacts with $\alpha 7$ while $\alpha 10$ completes the hydrophobic core of the domain (Fig. 1). The three helices $\alpha 7$ to $\alpha 9$ are held together by a cluster of hydrophobic residues side chains (Val158, Ile171, Leu175, Ile182). The residues of C-RcsB involved in RcsAB box DNA binding were identified, showing specific but weak interactions, by NMR chemical shift perturbation mapping (Lys153, Val158, Leu168, Thr170, Arg177, Ser178, Ile179, Thr181, Ile182). The identified residues are mostly located in the HTH motif ($\alpha 8$ to $\alpha 9$) and $\alpha 7$ (Fig. 1). The overall fold of the protein was not perturbed by the interaction with the specific DNA confirming that the affinity of RcsB to the RcsAB box is lower in the absence of RcsA of one

Table 1 *E. amylovora* 3D structures currently available at the PDB (06/15/2020)

Protein	Technique used	PDB IDs	Max resolution (Å)	Reference
RcsB	NMR	1P4W	—	(Pristovsek et al. 2003)
Levansucrase	X-ray cryst.	4D47	2.77	(Wuerger et al. 2015)
AmsI	X-ray cryst.	4D74	1.57	(Salomone-Stagni et al. 2016)
AmyR	X-ray cryst.	5FRK, 5FR7	2.12, 1.95	(Bartho et al. 2017)
HsvA	X-ray cryst.	5WPI	2.30	(Shanker et al. 2017)
GalU	X-ray cryst.	4D48	2.46	(Benini et al. 2017)
DfoJ	X-ray cryst.	5O5C	2.10	(Salomone-Stagni et al. 2018b)
DfoA	X-ray cryst.	5O8P, 5O8R	2.75, 2.80	(Salomone-Stagni et al. 2018b)
DfoC	X-ray cryst.	5O7O	2.11	(Salomone-Stagni et al. 2018b)
SrlD	X-ray cryst.	5O3Z	1.84	(Salomone-Stagni et al. 2018a)
AvrRpt2	X-ray cryst.	6HQZ	1.80	(Bartho et al. 2019)

NMR Nuclear Magnetic Resonance, X-ray cryst. X-ray crystallography

order of magnitude. In conclusion heterodimerization of RcsB with RcsA is not necessary for DNA interaction although the formation of the heterodimer RcsB-RcsA substantially increases DNA binding affinity for the RcsAB box (Pristovsek et al. 2003).

Levansucrase

E. amylovora produces amylovoran and levan, two exopolysaccharides which are components of its protective biofilm (Koczan et al. 2009). The proteins involved in the biosynthesis of amylovoran, considered one of the major pathogenicity factors, are encoded in the *ams* operon (Langlotz et al. 2011), while the synthesis of levan, proposed to be a virulence factors of the bacterium, is carried out by the enzyme levansucrase (Lsc, EC 2.4.1.10) (Zhao et al. 2005; Geier and Geider 1993). In low sucrose concentrations, Lsc catalyzes the hydrolysis of sucrose giving fructose and glucose. In high sucrose concentrations, as the one encountered during flower infection into the nectaries, a transfructosylation reaction occurs, in which a fructosyl moiety is transferred from an enzyme-fructosyl intermediate to an acceptor, in a ping-pong mechanism (Chambert et al. 1974). The acceptors can either be sucrose or fructose thus giving rise to short chain fructo-oligosaccharides (FOS), or a long chain levan.

E. amylovora Lsc (EaLsc) was recombinantly produced and crystallized and its crystal structure solved (Caputi et al. 2013a; Wuerges et al. 2015). EaLsc predominantly produces short-chain FOS of three to six units (Caputi et al. 2013b) while the Gram-positive *Bacillus megaterium* SacB produces long chain levan (Strube et al. 2011). EaLsc may have evolved towards shorter FOS production for two main reasons: to help the pathogen survive the high sucrose concentration in the

nectar during flower infection, and to provide glucose as a building block for the biosynthesis of the pathogenicity factor amylovoran. EaLsc shows the typical fold of glycoside hydrolase families 32 and 68 (Lammens et al. 2009) featuring a five-bladed β -propeller. Four twisted strands make up each β -sheet, with the outer β -strand almost perpendicularly oriented to the inner strand. A forty residue long ‘clamp’, begins at the N-terminus and stretches up to the first blade, wrapping around half the circumference of the propeller (Fig. 2).

EaLsc crystal structure was compared to those of fructosyltransferases from Gram-positive and Gram-negative bacteria with the aim of contributing to the identification of determinants of the diverse product spectra. Together with a structural alignment of the available Lscs, Wuerges et al. performed a multiple sequence alignment of levansucrases from different *Erwinia* species and strains to identify differences between pathogenic and non-pathogenic species (Wuerges et al. 2015). While in *E. amylovora* deletion of the levansucrase gene *lsc* reduces the virulence of the bacterium (Geier and Geider 1993), in *Erwinia tasmaniensis* the *lsc* gene produces Lsc but the bacterium is epiphytic and non-pathogenic (Kube et al. 2010). Potentially significant deviations in EaLsc is Phe198 that is changed to Tyr198 in *E. tasmaniensis* levansucrase (EtLsc), which is the same amino acid found in the long chain levan producing BmSacB (Strube et al. 2011). The comparison between EaLsc and EtLsc recently published support the hypothesis that the role of Lsc in both bacteria is mostly related to survival and metabolism rather than strictly connected to the virulence or pathogenicity of *E. amylovora*, as both bacteria share the same 3D structure and very similar reaction products profile (Polsinelli et al. 2019b).

Fig. 1 Cartoon representation of the NMR structure of RcsB DNA binding motif. Helices are in red and loops in green. Highlighted in Cyan are the residues identified by NMR chemical shift perturbation mapping (Lys153, Val158, Leu168, Thr170, Arg177, Ser178, Ile179, Thr181, Ile182, identified residues are mostly located in the HTH motif, $\alpha 8$ to $\alpha 9$ and $\alpha 7$)

Fig. 2 Cartoon representation of the crystal structure of *E. amylovora* levansucrase. Helices are depicted in red, loops and β -strands in yellow. In sticks are represented the products of sucrose hydrolysis, glucose and fructose, in the active site cavity (carbon in grey and oxygen in red)

AmsI

One of the most important pathogenicity factor of *E. amylovora* is amylovoran, an exopolysaccharide characterized by the repetition of units containing galactose, glucose, glucuronic acid and pyruvate (Nimtz et al. 1996). Amylovoran biosynthesis requires the concerted action of 12 genes (from *amsI* to *amsL*) clustered in the 16 kb *ams* region in *E. amylovora* genome (Bugert and Geider 1995). The AmsI protein, coded by the *amsI* gene, is a key enzyme of amylovoran metabolism. AmsI is a protein tyrosine phosphatase (PTP) involved in signaling (Bugert and Geider 1997). The kinase-phosphatase dual system has a critical role in bacterial virulence and cell signaling, PTPs catalyse the dephosphorylation of cognate kinase phospho-tyrosine substrates which in *E. amylovora* is represented by AmsA. AmsI overproduction leads to inhibition of amylovoran synthesis due to increased dephosphorylation of the cognate kinase (Bugert and Geider 1997). AmsI is a member of the low molecular weight protein tyrosine phosphatases family (LMW-PTP with a M_r of about 18 kDa), which are found both in Prokarya and Eukarya. LMW-PTPs share a consensus sequence CXGNXCRSP, where X could be leucine, isoleucine, threonine or phenylalanine, part of a conserved P-loop motif (Caselli et al. 2016). LMW-PTPs share the overall fold, the P-loop at the N-terminal region and the catalytic residues. Variability is observed among the residues forming the protein surface around the active site, important for the interaction with the cognate kinase. The interaction between *Escherichia coli* Wzb (analog of AmsI) and Wzc (analog of AmsA, AmsI cognate kinase) has been studied in detail by Temel et al. (Temel et al. 2013). In LMW-PTPs the active site residues form a cradle that provides hydrogen-bonding interactions with the phosphate group of the substrate (Evans et al. 1996). The catalytic mechanism of LMW-PTPs starts with the formation of a covalent phospho-enzyme intermediate involving as a nucleophile the S_γ -atom of a cysteine. This first step is assisted by the protonation of the leaving group by a conserved aspartate residue. Substrate binding and stabilization of the reaction intermediate are favored by a conserved arginine. In the second step of the reaction the phospho-cysteine intermediate is attacked by a water molecule regenerating the free enzyme and releasing inorganic phosphate (Denu and Dixon 1995; Denu et al. 1996).

In order to investigate the properties of AmsI, Salomone-Stagni et al. solved its crystal structure to a resolution of 1.57 Å and studied its activity by a steady state kinetic analysis (Salomone-Stagni et al. 2016). AmsI active site contains a water molecule that bridges the catalytic arginine and a sulfate molecule derived from the crystallization condition (Fig. 3). The sulfate in AmsI is mimicking the phosphate group of the substrate while the water molecule, nestled within the P-loop in close proximity of both the sulfate and the catalytic

cysteine, was proposed to attack the phospho-cysteine intermediate during the second step of the reaction.

The kinetic study on AmsI showed that citrate pH 5.5 is the optimal buffer with a decreased activity towards basic pH. To complete the investigation, homology modelling and molecular docking were used to propose a model of interaction between AmsI and AmsA. Potential putative interaction patches on the surface of the two proteins, to be validated by mutational studies, were identified and reported in the supplementary information of the published paper (Salomone-Stagni et al. 2016).

AmyR

AmyR is the protein product of the amylovoran repressor gene *amyR*, a member of the *ybjN* gene family (Wang et al. 2012). In *E. coli* *ybjN* has been described as a stress-related gene, as it depends on the temperature and the bacterial growth stage to be expressed (Chen et al. 2006). Studies have demonstrated that the over-expression of AmyR in *E. amylovora* negatively regulates the production of amylovoran and levan and that the *amyR* knockout produces 8-fold higher amylovoran than the control (Wang et al. 2012).

Bartho et al. reported the crystal structure of *E. amylovora* AmyR, the first representative structure of a YbjN protein from an Enterobacteriaceae species (Bartho et al. 2017). AmyR biologically functional assembly is a homodimer. The highest structural similarity of AmyR is with proteins belonging to class I T3C (Type 3 secretion chaperones) but showing some deviations in length and orientation in peripheric loops and secondary structure elements. Despite such structural similarity the role of AmyR is not as a secretion chaperone. The genomic location of *amyR* does not conform

Fig. 3 Cartoon representation of AmsI. In cyan the helices, in magenta the strands and in salmon the connecting loops. Represented as sticks is the sulfate molecule bound in the active site, with the sulfur in yellow and the oxygen in red

with the position of genes of secretion chaperones class I T3C which are located near the genes of the corresponding type 3 effectors (Lohou et al. 2013). The sequences alignment of 100 YbjN family proteins from a selection of species showed that the most conserved region was the stretch of residues 75–88 that is part of a hydrophobic dimerization interface. The second conserved region, comprising residues 13–24 could represent a site for protein-protein interaction similarly to class I T3Cs beside a plausible role of stabilizing protein fold (Fig. 4).

The gene regulation effect of AmyR was shown as a direct result of the presence of the AmyR protein, and not just co-regulation of the AmyR promoter since overexpression of AmyR under a non-native promoter showed changes in transcriptional levels of other genes (Wang et al. 2012). Bartho et al. proposed that AmyR could influence transcriptional regulation through an inhibitory binding, a stabilising binding, or a combination of the two. In the case of AmyR, one hypothesis is that the observed variations between the wild-type, the *amyR* knockout and the over expression mutant is because proteins of regulatory pathways could be bound and sequestered in a dose dependent manner (Bartho et al. 2017).

HsvA

Most of *E. amylovora* pathogenicity related genes are located in a Hrp pathogenicity island of its genome. Within the pathogenicity island, two operons are predicted to contain genes for the expression of proteins with enzymatic properties. They are controlled by the alternative σ factor HrpL during infection suggesting a role in bacterial virulence (Oh et al. 2005). The amidinotransferase HsvA is the product of the gene *hsvA* that together with *hsvB* and *hsvC* is found in the region of Hrp-associated enzymes required for full virulence (Oh et al. 2005). The reaction mechanism of amidinotransferases involves the action of three key residues in the active site, cysteine, histidine and aspartic acid, in a double displacement

(ping-pong) mechanism as the access to the active site does not allow simultaneous access to both amidino donor and acceptor during catalysis (Humm et al. 1997a, b). A biochemical characterization and protein structure determination was carried out in 2017 with the aim of understanding the mechanism of action, the substrate specificity and suggest a possible role in virulence (Shanker et al. 2017). The amidinotransferase activity was investigated on a panel of 16 candidate acceptors using arginine as an amidino donor and detecting the amount of ornithine released during a 30 min course reaction. The most efficient acceptor was putrescine (with a chain of four carbons between primary amines) while other polyamines from two to six carbons length varied in activity. Considering that *E. amylovora* genome contains a gene for the expression of a spermidine synthase (SpeE) the authors found that HsvA was far more efficient when using spermidine as an acceptor rather than putrescine. The X-ray crystal structure was determined showing that HsvA is a homodimer, with an overall fold of each monomer arranged as an α/β propeller domain. The α -helices and β -strands are organized into five $\beta\beta\alpha\beta$ modules around a pseudo 5-fold symmetry axis, resembling a basket. The active site contains the conserved amino acids Cys351, His249, and Asp200, forming a catalytic arrangement similar to that of cysteine proteases. These amino acids are at the base of a long narrow channel running from the protein surface to the active site cavity. Three glutamate residues are located at the entrance of the tunnel on the protein surface conferring a negative charge that could help positioning positively charged substrates (i.e., polyamines) into the channel. The arginine binding site is characterized by six amino acids in the same arrangement as in the human L-arginine:glycine amidinotransferase (Humm et al. 1997b). The preference for putrescine and spermidine is explained by the presence of Glu244 located just below the opening of the channel leading to the active site and about 11 Å from the catalytic cysteine, at a distance compatible with the length of putrescine and spermidine (Fig. 5).

Fig. 4 Left panel: cartoon representation of the biologically relevant dimer from the crystal structure of AmyR. One subunit is represented in cyan while the other one in green. Right panel: AmyR conserved regions are depicted in red and the surface in transparent grey

The role of HsvA in *E. amylovora* pathogenicity was proposed to be that of regulating the level of polyamines, possibly enhancing the fitness of the bacterium during infection, or by supporting the production of virulence factors (Shanker et al. 2017).

GalU

The biosynthesis, regulation and secretion of *E. amylovora* exopolysaccharide amylovoran requires the concerted action of the proteins encoded by the twelve genes in the *ams* operon (Bugert and Geider 1995). The main amylovoran building block galactose is provided by enzymes of the glucose and galactose metabolism. Glucose-1-phosphate uridylyltransferase (GalU; UDP-glucose pyrophosphorylase, EC 2.7.7.9) catalyses the reaction between α -D-glucose 1-phosphate (Glc-1P) and UTP to give UDP-glucose and pyrophosphate. UDP-glucose is then converted by the epimerase GalE into UDP-galactose, the main building block of amylovoran (Metzger et al. 1994; Wagstaff et al. 2015).

The reaction mechanism of GalUs to obtain UDP-glucose starts with the binding of UTP to the enzyme in the presence of a Mg^{2+} ion followed by the substrate Glc-1P. The phosphate group of Glc-1P is positioned to favor the nucleophilic attack to the α -phosphate of UTP. Kim et al. proposed a mechanism for *Helicobacter pylori* HpGalU, in which the substrates bind in sequential order, first Mg^{2+} and UTP followed by Glc-1P to form the reaction complex in a sequential Bi Bi mechanism, giving as final products UDP-glucose and PPi (Kim et al. 2010). *E. amylovora* GalU (EaGalU) may represent a good target to block amylovoran biosynthesis as its gene is located upstream of the *ams* gene cluster therefore crucial for amylovoran production. Besides the importance of this class of enzymes for bacterial survival the potential use of EaGalU in biotechnology and in chemo-enzymatic

synthesis makes it a good candidate for further studies. The structure, the enzymatic activity of EaGalU and its substrate specificity have been determined for a better understanding of its structure and function relationship (Benini et al. 2017). EaGalU was recombinantly expressed in *E. coli* and crystallized (Toccafondi et al. 2014). EaGalU structure determined by protein X-ray crystallography shows two EaGalU molecules in the asymmetric unit that by crystallographic symmetry form the dimer of dimers to give the biologically functional homotetramer observed in bacterial GalUs (Fig. 6). EaGalU tertiary structure is characterized by a central open β -sheet surrounded by two small outlying β -sheets and α -helices, similar to the Rossmann fold characteristic of nucleotide-binding proteins (Unligil and Rini 2000).

Kinetic analysis in the synthetic direction for EaGalU revealed values of K_m that are in the μM range ($7.7 \pm 0.9 \mu M$ for Glc-1P and $27.3 \pm 2.6 \mu M$) comparable with the one of the GalU enzyme from *E. coli* (EcGalU) (Wagstaff et al. 2015). Substrate specificity of EaGalU was investigated using a range of sugar 1-phosphates and UTP and the reaction products were analysed by 1H NMR. EaGalU showed a clear preference for its natural substrate α -D-glucose 1-phosphate (Glc-1P) with 100% conversion to UDP-glucose. The tested *gluco*-configured molecules were also substrates and converted to the corresponding UDP-sugar derivative. A conversion of 100% was observed for α -D-glucosamine 1-phosphate (GlcN-1P) and for the pentose α -D-xylose 1-phosphate (Xyl-1P) while *N*-acetyl- α -D-glucosamine 1-phosphate (GlcNAc-1P) showed a 74% conversion. The C2 epimer of the natural substrate α -D-glucose 1-phosphate, α -D-mannose 1-phosphate (Man-1P) is still converted to 70%. The enzymatic performance is much lower using *galacto*-configured substrates (C4 epimers) and α -D-galactose 1-phosphate (Gal-1P) is converted very slowly to 28% and no conversion was observed in the case of α -D-galactosamine 1-phosphate (GalN-

Fig. 5 Cartoon representation of the biologically relevant homodimer of HsvA. Chain A is in green while chain B in cyan. Highlighted in red is Glu244 located just below the opening of the channel leading to the active site and regulating substrate selectivity

1P) and for α -D-galacturonic acid 1-phosphate (GalA-1P). In silico molecular docking computations were employed to rationalize the different measured enzymatic activities and to understand the mechanism by which the different sugar 1-phosphates bind to the enzyme and initiate the reaction. The binding mode of all the sugar 1-phosphates used were simulated using the EaGalU structure. The formation of an H-bond between Lys202 N ζ and the phosphate group of the sugar 1-phosphates together with the establishment of an H-bond between Glu201 O ϵ and O2 on the sugar ring stabilizes the interaction between EaGalU and the cognate 1-phosphates sugars, and was proposed to help positioning of the substrate thus favoring phosphate nucleophilic attack to UTP. These results were proposed to serve as a guide for selecting sugar 1-phosphates with the ideal configuration to obtain UDP-sugar derivatives by EaGalU (Benini et al. 2017).

Desferrioxamine pathway, DfoJ, DfoA, DfoC

Iron is an essential nutrient in almost every living organism. The siderophore desferrioxamine E (DFO-E) plays an important role in *E. amylovora* pathogenesis as iron bioavailability in the host is very low. DfoJ, DfoA and DfoC are the enzymes responsible for desferrioxamine biosynthesis starting from lysine. (Fig. 7).

In *E. amylovora* a single gene cluster (*dfoJAC* operon) encodes for the three proteins required for performing the four enzymatic reactions necessary for the biosynthesis of DFO-E (Polsinelli et al. 2019a). Mutants of *E. amylovora* CFBP1430,

defective in the siderophore biosynthesis (disrupted *dfoA* gene) or siderophore uptake (ferrioxamine receptor FoxR defective) showed a two orders of magnitude reduced growth on apple flowers compared to the wild-type (Dellagi et al. 1998). Desferrioxamines (DFOs) consist of diamine and dicarboxylic acid building blocks linked by amide bonds. Desferrioxamine E is the major product of DFO biosynthesis in *E. amylovora* while other DFOs (D2, X1–7 and G1–2) are produced in lower quantities (Feistner et al. 1993; Feistner 1995).

Beside scavenging iron from the host, DFO-E protects the bacterium against the oxidative burst triggered by the plant against the infection, and could also enhance the oxidative stress induced by hairpins causing an increase of electrolyte leakage from the host plant (Dellagi et al. 1998).

With the complete structural characterization of the whole pathway Salomone-Stagni et al. confirmed that *E. amylovora* DFO-E biosynthesis begins with the decarboxylation of lysine by DfoJ, a lysine decarboxylase (group II pyridoxal-dependent enzymes, PFAM: PF00282) to give cadaverine. Cadaverine is immediately monooxygenated by DfoA, thus generating N-hydroxyl-cadaverine. In the last steps DfoC first catalyzes the condensation of N-hydroxyl cadaverine with a succinyl moiety, to give N-5-aminopentyl-N-(hydroxyl)-succinamic acid, which is then promptly processed by DfoC synthetase domain, for trimerization and cyclization, to give desferrioxamine E. Beside the structural characterization, in vitro activity of DfoA and DfoJ was also tested in the same work, to confirm the role of these enzymes in the pathway (Salomone-Stagni et al. 2018b).

DfoJ is composed of three domains: the N-terminal domain (NTD), the PLP-binding domain and the C-terminal domain (CTD). The NTD is composed of three α -helices and a long loop connecting the NTD to the large PLP binding domain which is formed by a seven-stranded β -sheet surrounded by α -helices. The CTD is small and formed by a four stranded anti-parallel β -sheet and three α -helices located opposite to the large PLP binding domain. DfoJ forms stable biologically relevant dimers that are related by crystallographic symmetry, as observed in the crystal, and confirmed by PDBePISA (Krissinel 2010). The PLP binding site is located at DfoJ dimer interface and residues from both subunits are making up the active site and are involved in the reaction mechanism.

DfoA is a FAD and NADPH dependent N-hydroxylating monooxygenase/hydroxylase (NMO) belonging to the LucD family of enzymes (PFAM PF13434) (Olucha and Lamb 2011). The structure of DfoA, solved in its holoform to 2.8 Å resolution contains the ternary complexes of DfoA with both FAD and NADP⁺. DfoA belongs to class B flavoprotein monooxygenases (van Berkel et al. 2006), contains two Rossmann-fold dinucleotide-binding domains for binding FAD and NADP(H) and forms biologically functional homotetramers. Homotetrameric oligomerization is also observed in other bacterial and fungal flavoprotein

Fig. 6 Cartoon representation of the biologically relevant EaGalU homotetramer, a dimer of dimers, and colored according to secondary structure elements. In one dimer the helices are colored in red, the strands in yellow and the turns in green, in the other dimer the helices in cyan, the strands in magenta and the turns in salmon

monooxygenases belonging to the NMOs subgroup such as the ornithine hydroxylase from *Kutzneria* sp. 744 (Setser et al. 2014), and from *Aspergillus fumigatus* (Franceschini et al. 2012).

KtzI from *Kutzneria* sp. 744 (Setser et al. 2014) is the most similar structure to DfoA as shown by a PDB search carried out with PDBeFOLD (Krissinel and Henrick 2004). To discover the determinant of substrate selectivity, a cadaverine molecule was manually placed in DfoA active site using as a guide the position of ornithine in KtzI. In DfoA Ala60 and Leu238 allow hydrophobic interactions with the aliphatic portion of the cadaverine, the backbone oxygen of Tyr268 may bind to cadaverine N1, and Gln55 contributes to the formation of the cadaverine binding pocket. DfoA Asp392 sidechain is in a perfect position to bind cadaverine N7 amino group. Furthermore, Thr241(O γ) may form a hydrogen bond with cadaverine N7 amino group (Salomone-Stagni et al. 2018b).

Among siderophore biosynthesis enzymes DfoC is an unusual example as the succinyl transferase and siderophore synthetase domains are part of a single protein. DfoC N-terminal domain is responsible for the transfer of a succinyl moiety from a succinyl-CoA donor to the primary amine of N-hydroxyl-cadaverine and is therefore a succinyl transferase domain. The C-terminal domain belongs to the type-C siderophore synthetases (Non-Ribosomal Peptide Synthetase (NRPS)-Independent Siderophore (NIS) protein family (Challis and Naismith 2004)). Within the synthetase domain the nucleotide triphosphate (NTP) required for the reaction, binds at the base of the substrate binding pocket. Access to the NTP binding site would be blocked when the N-5-

aminopentyl-N-(hydroxyl)-succinamic acid substrate is loaded. The authors identified a “channel” that would allow NTP recycling without requiring the major substrates to be released (Salomone-Stagni et al. 2018b). DfoC structure features a large cavity bridging across the DfoC dimerization interface that could allow the ‘head’ of N-5-aminopentyl-N-(hydroxyl)-succinamic acid to enter into the binding pocket of one subunit, and the ‘tail’ to enter into the binding pocket of the other subunit, therefore suggesting a cooperative synthesis of the DFO-E ring structure (Salomone-Stagni et al. 2018a, b). The expression of the N-terminal succinyl transferase and the C-terminal siderophore synthetase domains in a single protein might suggest cooperativity or coordinated substrate transfer but the active site locations of the N- and C-terminal domains do not. The expression of these two enzymes as a single protein may still improve substrate trafficking between their active site even if the substrate is not actively transported. The N-terminal domains may move as ‘balls on a string’ thanks to the linker loops connecting them to the dimerized C-terminal domains. This movement would bring the active sites of each domain closer to each other, enabling faster and more efficient substrates delivery from the N-terminal to the C-terminal domains Salomone-Stagni et al. 2018b).

SrID

Sorbitol is used as a transport carbohydrate in plants belonging to the Amygdaloideae subfamily including apple, pear and quince (Zimmermann and Ziegler 1975) that are hosts for *E. amylovora* (Vanneste 1995). Sorbitol is used by *E. amylovora* as a carbon source thanks to the sorbitol operon

Fig. 7 Schematic view of the desferrioxamine E biosynthetic pathway

(*srlA*, *E*, *B*, *D*, *M*, *R* genes) which encodes the six proteins required for sorbitol uptake. *SrlD* is a member of the short-chain dehydrogenases/reductases (SDRs), that catalyze NAD(P)(H)-dependent redox reactions. It is a sorbitol-6-phosphate 2-dehydrogenase (S6PDH; EC 1.1.1.140) that catalyzes the interconversion of D-sorbitol 6-phosphate (glucitol 6-phosphate; S6P) to D-fructose 6-phosphate (F6P). Mutations in the sorbitol operon of *E. amylovora* impair sorbitol utilization in apple shoots, thus preventing efficient colonization of host plant tissue (Aldridge et al. 1997). It has been proposed that, differences in host specificity between the Spiraeoideae infecting strains and the *Rubus* infecting strains might be due to the lack of a complete sorbitol operon in the *Rubus* infecting strains (Borruso et al. 2017). For a better understanding of *SrlD* mechanism of action its activity and 3D structure were determined together with structural and sequence comparisons (Salomone-Stagni et al. 2018a). The structure of *SrlD* was solved to a maximum resolution of 1.84 Å. *SrlD* structure features a seven strands (β A– β G) central β -sheet, surrounded by 10 α -helices, a typical Rossmann fold characteristic of the SDR superfamily (Oppermann et al. 2003; Persson and Kallberg 2013). *SrlD* quaternary structure is a homotetramer. Each monomer has two interaction surfaces which are involved in establishing contacts with equivalent surfaces of two neighboring monomers with about 25% of the monomer surface buried in the contact areas. The cofactor NAD⁺/NADH and substrate S6P binding site is evident on the protein surface, as a large pocket coming from the protein core and ending with a small channel. *SrlD* active site features the conserved catalytic tetrad (Asn112, Ser141, Tyr154, Lys158) as previously determined for bacterial SDRs (Philippsen et al. 2005). In the active site, and relevant for selectivity towards S6P rather than unphosphorylated sorbitol (towards which the protein is inactive) is Lys142. Moreover the authors propose that the substrate S6P would enter into the active site keeping the phosphate group near the protein surface where it would interact with four lysines (Lys142, Lys149, Lys203 and Lys218) that would form H-bonds with it (Salomone-Stagni et al. 2018a) (Fig. 8).

AvrRpt2

In order to suppress the host organism's defenses, *E. amylovora* secretes several proteins into the host cytoplasm via the Type 3 Secretion System (T3SS) (Vrancken et al. 2013; Borruso et al. 2017). The cysteine protease AvrRpt2 is one of the type III effector proteins which are secreted by *E. amylovora* (AvrRpt2_{EA}) (Zhao et al. 2006; Khan et al. 2012; Schropfer et al. 2018) AvrRpt2 is activated by peptidyl-prolyl *cis/trans* isomerization after secretion inside the host cell by T3SS (Coaker et al. 2006). Prolyl *cis/trans* isomerization of *Pseudomonas syringae* AvrRpt2 (AvrRpt2_{PS}) depends on the eukaryotic cyclophilin family of peptidyl prolyl *cis/trans* isomerases (PPIases) (Aumuller et al. 2010), and

Fig. 8 The entrance to *SrlD* active site showing the four lysines (Lys142, Lys149, Lys203 and Lys218) proposed by Salomone-Stagni et al. to be involved in substrate selectivity and orientation and located at the entrance of the active site channel

occurs at 4 conserved GPxL motifs (Coaker et al. 2006). AvrRpt2_{PS} and AvrRpt2_{EA} are both activated in a cyclophilin-dependent manner as they share high sequence similarity and conservation of the GPxL motifs. Upon secretion into the host cell cyclophilin-dependent autocatalytic cleavage of an N-terminal fragment occurs, between the conserved residues G45 and G46 (AvrRpt2_{EA}) or G71 and G72 (AvrRpt2_{PS}) (Jin et al. 2003; Zhao et al. 2006). Cleavage of the N-terminal fragment has no effect on proteolytic activity (Chisholm et al. 2005; Aumuller et al. 2010).

The secretion of AvrRpt2_{EA} into the host cell can increase virulence, but is not essential for *E. amylovora* infection (Zhao et al. 2006). In *Malus × robusta* 5 the presence of AvrRpt2_{EA} is recognised by FB_MR5, the only identified resistance protein in a *Malus* species, thus preventing *E. amylovora* infection (Vogt et al. 2013). *E. amylovora* strains carrying an AvrRpt2_{EA} mutation of cysteine 156 (Cys156) to a serine (Ser156) are able to overcome resistance. This naturally occurring mutation was found in several *E. amylovora* strains in North America and reproduced in AvrRpt2_{EA} deletion strains by Vogt et al. (Vogt et al. 2013). The induced expression of AvrRpt2_{EA} in transgenic plants of the fire blight-susceptible cultivar Pinova caused shoots necrosis, browning of older leaves and increased expression of salicylic acid (Schropfer et al. 2018).

The structure of AvrRpt2_{EA70–222} (from residue number 70 to the C-terminal) was solved by X-ray crystallography providing insights into the cyclophilin-mediated maturation of AvrRpt2, and the mechanism of recognition by FB_MR5 (Bartho et al. 2019). AvrRpt2_{EA70–222} consists of a six stranded anti-parallel β -sheet flanked on both sides by 5 α -helices. AvrRpt2_{EA70–222} retains a conserved fold with other protease-

like protein domains despite highly divergent sequences and different biological functions. The active site is located in a cleft formed by the connecting loops between beta-strands. Besides the residues responsible for the proteolytic cleavage (Cys88, His173, and Asp191), the cleft contains a negatively charged internal area and, at its edge, a positively charged arginine sidechain (Arg194) possibly contributing to substrate binding and specificity. The four conserved cyclophilin-binding motifs (CBM1 - CBM4) are present in the AvrRpt2_{EA70–222} structure in the trans configuration, as the crystallized protein was never exposed to a eukaryotic cyclophilin. The four CBMs are located in loops (CBM1, CBM3) or next to secondary structure elements (CBM2, CBM4). It was hypothesized that the structural changes induced by proline isomerization at CBM4 (residues 195–198, GPDL) would reposition the catalytic residue Asp191 within the active site, thus altering the surrounding substrate binding cleft and possibly modifying the substrate binding area and substrate specificity (Bartho et al. 2019). The structure of AvrRpt2_{EA70–222} shows that Cys156, implicated in recognition by the FB_MR5 resistance protein of *Malus × robusta* 5, is on the opposite side of the protein in respect to the active site catalytic triad (Cys88, His173 and Asp191). Cys156 is thus not involved in substrate binding at the catalytic site, or in determining substrate specificity of AvrRpt2_{EA}. In the structure of AvrRpt2_{EA70–222}, Cys156 is co-located with CBM3 (residues 159–163, GPIMF, Fig. 9), a post-translational modification, or even a disulfide bridge formation may interfere with cyclophilin binding, blocking cis-trans isomerisation of the Gly-Pro site and preventing the AvrRpt2_{EA} from activating or from interacting with plant host factor. These types of post-translational modification would not be possible on a serine therefore explaining the different behaviour of the two genetic variants (Cys156 vs Ser156) (Bartho et al. 2019). The resistance of *Malus × robusta* 5 to *E. amylovora* strains featuring AvrRpt2_{EA}-Cys156 can be due to either FB_MR5 direct interaction with the Cys156 of AvrRpt2_{EA}, or to a post-translational modification of Cys156 which would not be possible with a serine instead of a cysteine. Another hypothesis proposed by Bartho et al. is that the FB_MR5 resistance protein may recognize both the AvrRpt2_{EA}-Cys156 and AvrRpt2_{EA}-Ser156 forms, but only the latter would be correctly processed and able to suppress the resistance-signaling pathway triggered by FB_MR5 (Bartho et al. 2019).

Conclusions

The availability of sequenced genomes and the current technology allows to study the structure and function of proteins and enzymes reaching a high level of details. The structural data deposited and available in the PDB enable, researcher from different fields and background, to use the results

accomplished by the structural biologists and get a better understanding of the structure and function relationships of macromolecules. Structural and functional characterization of proteins could provide, not only a deeper understanding of the biology of the organism under investigation, but also starting points for developing molecules to modulate or inhibit enzymatic activity and to control diseases.

This review reports examples of the wealth of information provided by structural biology for a better understanding of *E. amylovora* biology. The study of the transcriptional regulator RcsB allowed the identification and location of the residues involved in DNA binding. Ealse structure in complex with glucose and fructose represents a snapshot of the reaction mechanism showing the product of sucrose hydrolysis trapped in the active site. In the crystal structure of AmsI it was possible to identify the water molecule supposed to attack the phospho-cysteine intermediate in the second step of AmsI reaction. In the case of AmyR the structural comparison of its crystal structure against the PDB database was crucial to discover the 3D similarity with T3Cs and to locate the conserved residues, as bioinformatics on its aminoacidic sequence alone could not give any clue about it. HsvA structure determination, beside confirming its role as a polyamine amidinotransferase, allowed not only the characterization of the active site, but also the identification of the residue (Glu244) that determines the enzyme substrate specificity. The structure of EaGalU was used as a model for molecular docking calculation to explain the different enzymatic efficiency towards a series of phosphorylated sugars. The structural characterization of the whole desferrioxamine E biosynthetic pathway confirmed the role of the enzymes involved, as well as, providing unprecedented structural information for DfoC. The structure of DfoC is the first of a siderophore synthetase coupling an acyltransferase domain at its N-term and a Non-Ribosomal Peptide Synthetase (NRPS)-Independent

Fig. 9 Cartoon representation showing that in AvrRpt2_{EA70–222} Cys156 is co-located with CBM3 (residues 159–163, GPIMF)

Siderophore (NIS) domain at its C-term. With the identification of four lysine (Lys142, Lys149, Lys203 and Lys218) around the entrance of the active site, the structure of SrlD provided an explanation for its enzymatic selectivity. SrlD is indeed active only towards phosphorylated sorbitol while completely inactive on sorbitol. The four lysines would favor binding of phosphorylated sorbitol into the active site for the catalysis thus providing the required selectivity. Finally, the structure of AvrRpt2^{EA70–222} provided crucial information about the location of the CBM motifs and of Cys156. The latter is exposed on the surface of the protein and therefore prone to protein-protein interaction (e.g., by disulfide bridge formation) or post-translational modifications. The obtained information could not be determined by sequence comparison alone or in vivo experiments.

In conclusion structural biology is an important tool to provide information towards the knowledge of the molecular mechanisms by which biomolecules operate and interact with each other, thus pushing forward the frontier of biochemical science.

Acknowledgements Open access funding provided by Libera Università di Bolzano within the CRUI-CARE Agreement. Coworkers and collaborators are acknowledged for their contribution to the projects. Fabio Rezzonico is acknowledged for critically reading the manuscript. Figures were made using “The PyMOL Molecular Graphics System, Schrödinger, LLC”.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aldridge P, Metzger M, Geider K (1997) Genetics of sorbitol metabolism in *Erwinia amylovora* and its influence on bacterial virulence. *Mol Gen Genet* 256:611–619
- Ancona V, Chatnaparat T, Zhao Y (2015) Conserved aspartate and lysine residues of RcsB are required for amylovan biosynthesis, virulence, and DNA binding in *Erwinia amylovora*. *Mol Gen Genomics* 290:1265–1276
- Aumuller T, Jahreis G, Fischer G, Schiene-Fischer C (2010) Role of prolyl cis/trans isomers in cyclophilin-assisted *Pseudomonas syringae* AvrRpt2 protease activation. *Biochemistry* 49:1042–1052
- Bartho JD, Bellini D, Wuerges J, Demitri N, Toccafondi M, Schmitt AO, Zhao Y, Walsh MA, Benini S (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:e0176049
- Bartho JD, Demitri N, Bellini D, Flachowsky H, Peil A, Walsh MA, Benini S (2019) The structure of *Erwinia amylovora* AvrRpt2 provides insight into protein maturation and induced resistance to fire blight by *Malus x robusta* 5. *J Struct Biol* 206:233–242
- Benini S, Toccafondi M, Rejzek M, Musiani F, Wagstaff BA, Wuerges J, Cianci M, Field RA (2017) Glucose-1-phosphate uridylyltransferase from *Erwinia amylovora*: activity, structure and substrate specificity. *Biochim Biophys Acta* 1865:1348–1357
- Bereswill S, Geider K (1997) Characterization of the *rscB* gene from *Erwinia amylovora* and its influence on exopolysaccharide synthesis and virulence of the fire blight pathogen. *J Bacteriol* 179:1354–1361
- Berman H, Henrick K, Nakamura H (2003) Announcing the worldwide protein data Bank. *Nat Struct Biol* 10:980
- Borruso L, Salomone-Stagni M, Polsinelli I, Schmitt AO, Benini S (2017) Conservation of *Erwinia amylovora* pathogenicity-relevant genes among *Erwinia* genomes. *Arch Microbiol* 199:1335–1344
- Bugert P, Geider K (1995) Molecular analysis of the *ams* operon required for exopolysaccharide synthesis of *Erwinia amylovora*. *Mol Microbiol* 15:917–933
- Bugert P, Geider K (1997) Characterization of the *amsI* gene product as a low molecular weight acid phosphatase controlling exopolysaccharide synthesis of *Erwinia amylovora*. *FEBS Lett* 400:252–256
- Burley SK, Berman HM, Bhikadiya C, Bi C, Chen L, Di Costanzo L, Christie C, Dalenberg K, Duarte JM, Dutta S, Feng Z, Ghosh S, Goodsell DS, Green RK, Guranovic V, Guzenko D, Hudson BP, Kalro T, Liang Y, Lowe R, Namkoong H, Peisach E, Periskova I, Prlc A, Randle C, Rose A, Rose P, Sala R, Sekharan M, Shao C, Tan L, Tao YP, Valasatava Y, Voigt M, Westbrook J, Woo J, Yang H, Young J, Zhuravleva M, Zardecki C (2019) RCSB protein data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res* 47:D464–D474
- Caputi L, Cianci M, Benini S (2013a) Cloning, expression, purification, crystallization and preliminary X-ray analysis of EaLsc, a levansucrase from *Erwinia amylovora*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 69:570–573
- Caputi L, Nepogodiev SA, Malnoy M, Rejzek M, Field RA, Benini S (2013b) Biomolecular characterization of the levansucrase of *Erwinia amylovora*, a promising biocatalyst for the synthesis of fructooligosaccharides. *J Agric Food Chem* 61:12265–12273
- Caselli A, Paoli P, Santi A, Mugnaioni C, Toti A, Camici G, Cirri P (2016) Low molecular weight protein tyrosine phosphatase: multifaceted functions of an evolutionarily conserved enzyme. *Biochim Biophys Acta* 1864:1339–1355
- Cavalli A, Salvatella X, Dobson CM, Vendruscolo M (2007) Protein structure determination from NMR chemical shifts. *Proc Natl Acad Sci U S A* 104:9615–9620
- Challis GL, Naismith JH (2004) Structural aspects of non-ribosomal peptide biosynthesis. *Curr Opin Struct Biol* 14:748–756
- Chambert R, Treboul G, Dedonder R (1974) Kinetic studies of levansucrase of *Bacillus subtilis*. *Eur J Biochem* 41:285–300
- Chen X, Shen D, Zhou B (2006) Analysis of the temperature-sensitive mutation of *Escherichia coli* pantothenate kinase reveals YbjN as a possible protein stabilizer. *Biochem Biophys Res Commun* 345: 834–842
- Chisholm ST, Dahlbeck D, Krishnamurthy N, Day B, Sjolander K, Staskawicz BJ (2005) Molecular characterization of proteolytic cleavage sites of the *Pseudomonas syringae* effector AvrRpt2. *Proc Natl Acad Sci U S A* 102(6):2087–2092
- Coaker G, Zhu G, Ding Z, Van Doren SR, Staskawicz B (2006) Eukaryotic cyclophilin as a molecular switch for effector activation. *Mol Microbiol* 61:1485–1496

- Dellagi A, Brisset MN, Paulin JP, Expert D (1998) Dual role of desferrioxamine in *Erwinia amylovora* pathogenicity. *Mol Plant-Microbe Interact* 11:734–742
- Denu JM, Dixon JE (1995) A catalytic mechanism for the dual-specific phosphatases. *Proc Natl Acad Sci U S A* 92:5910–5914
- Denu JM, Lohse DL, Vijayalakshmi J, Saper MA, Dixon JE (1996) Visualization of intermediate and transition-state structures in protein-tyrosine phosphatase catalysis. *Proc Natl Acad Sci U S A* 93:2493–2498
- Evans B, Tishmack PA, Pokalsky C, Zhang M, Van Etten RL (1996) Site-directed mutagenesis, kinetic, and spectroscopic studies of the P-loop residues in a low molecular weight protein tyrosine phosphatase. *Biochemistry* 35:13609–13617
- Feistner GJ (1995) Proferrioxamine synthesis in *Erwinia amylovora* in response to precursor or hydroxylysine feeding: metabolic profiling with liquid chromatography-electrospray mass spectrometry. *Biomaterials* 8:318–327
- Feistner GJ, Stahl DC, Gabrik AH (1993) Proferrioxamine siderophores of *Erwinia amylovora*. A capillary liquid chromatographic/electrospray tandem mass spectrometric study. *Org Mass Spectrom* 28:163–175
- Franceschini S, Fedkenheuer M, Vogelaar NJ, Robinson HH, Sobrado P, Mattevi A (2012) Structural insight into the mechanism of oxygen activation and substrate selectivity of flavin-dependent N-hydroxylating monooxygenases. *Biochemistry* 51:7043–7045
- Geier G, Geider K (1993) Characterization and influence on virulence of the levansucrase gene from the fireblight pathogen *Erwinia amylovora*. *Physiol Mol Plant Pathol* 42:387–404
- Grabowski M, Niedzialkowska E, Zimmerman MD, Minor W (2016) The impact of structural genomics: the first quinquennial. *J Struct Funct Genom* 17:1–16
- Gross M, Geier G, Rudolph K, Geider K (1992) Levan and levansucrase synthesized by the fireblight pathogen *Erwinia amylovora*. *Physiol Mol Plant Pathol* 40:371–381
- Henikoff S, Wallace JC, Brown JP (1990) Finding protein similarities with nucleotide sequence databases. *Methods Enzymol* 183:111–132
- Humm A, Fritsche E, Steinbacher S (1997a) Structure and reaction mechanism of L-arginine:glycine amidinotransferase. *Biol Chem* 378:193–197
- Humm A, Fritsche E, Steinbacher S, Huber R (1997b) Crystal structure and mechanism of human L-arginine:glycine amidinotransferase: a mitochondrial enzyme involved in creatine biosynthesis. *EMBO J* 16:3373–3385
- Jin P, Wood MD, Wu Y, Xie Z, Katagiri F (2003) Cleavage of the *Pseudomonas syringae* type III effector AvrRpt2 requires a host factor(s) common among eukaryotes and is important for AvrRpt2 localization in the host cell. *Plant Physiol* 133:1072–1082
- Kelm O, Kiecker C, Geider K, Bernhard F (1997) Interaction of the regulator proteins RcsA and RcsB with the promoter of the operon for amylovoran biosynthesis in *Erwinia amylovora*. *Mol Gen Genet* 256:72–83
- Khan MA, Zhao Y, Korban SS (2012) Molecular mechanisms of pathogenesis and resistance to the bacterial pathogen *Erwinia amylovora*, causal agent of fire blight disease in Rosaceae. *Plant Mol Biol Rep* 30:247–260
- Kim H, Choi J, Kim T, Lokanath NK, Ha SC, Suh SW, Hwang HY, Kim KK (2010) Structural basis for the reaction mechanism of UDP-glucose pyrophosphorylase. *Mol Cells* 29:397–405
- Koczan JM, McGrath MJ, Zhao Y, Sundin GW (2009) Contribution of *Erwinia amylovora* exopolysaccharides amylovoran and Levan to biofilm formation: implications in pathogenicity. *Phytopathology* 99:1237–1244
- Krissinel E (2010) Crystal contacts as nature's docking solutions. *J Comput Chem* 31:133–143
- Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 60:2256–2268
- Kube M, Migdoll AM, Gehring I, Heitmann K, Mayer Y, Kuhl H, Knaust F, Geider K, Reinhardt R (2010) Genome comparison of the epiphytic bacteria *Erwinia billingiae* and *Erwinia tasmaniensis* with the pear pathogen *Erwinia pyrifoliae*. *BMC Genomics* 11
- Lammens W, Le Roy K, Schroeven L, Van Laere A, Rabijns A, Van den Ende W (2009) Structural insights into glycoside hydrolase family 32 and 68 enzymes: functional implications. *J Exp Bot* 60(3):727–740
- Langlotz C, Schollmeyer M, Coplin DL, Nimtz M, Geider K (2011) Biosynthesis of the repeating units of the exopolysaccharides amylovoran from *Erwinia amylovora* and stewartan from *Pantoea stewartii*. *Physiol Mol Plant Pathol* 75:163–169
- Lohou D, Lonjon F, Genin S, Vailleau F (2013) Type III chaperones & co in bacterial plant pathogens: a set of specialized bodyguards mediating effector delivery. *Front Plant Sci* 4:435
- Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, Dow M, Verdier V, Beer SV, Machado MA, Toth I, Salmond G, Foster GD (2012) Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol Plant Pathol* 13:614–629
- Metzger M, Bellemann P, Bugert P, Geider K (1994) Genetics of galactose metabolism of *Erwinia amylovora* and its influence on polysaccharide synthesis and virulence of the fire blight pathogen. *J Bacteriol* 176:450–459
- Muench SP, Antonyuk SV, Hasnain SS (2019) The expanding toolkit for structural biology: synchrotrons, X-ray lasers and cryoEM. *IUCrJ* 6:167–177
- Nimtz M, Mort A, Domke T, Wray V, Zhang Y, Qiu F, Coplin D, Geider K (1996) Structure of amylovoran, the capsular exopolysaccharide from the fire blight pathogen *Erwinia amylovora*. *Carbohydr Res* 287:59–76
- Oh CS, Kim JF, Beer SV (2005) The Hrp pathogenicity island of *Erwinia amylovora* and identification of three novel genes required for systemic infection. *Mol Plant Pathol* 6:125–138
- Olucha J, Lamb AL (2011) Mechanistic and structural studies of the N-hydroxylating flavoprotein monooxygenases. *Bioorg Chem* 39:171–177
- Oppermann U, Filling C, Hult M, Shafqat N, Wu X, Lindh M, Shafqat J, Nordling E, Kallberg Y, Persson B, Jorvall H (2003) Short-chain dehydrogenases/reductases (SDR): the 2002 update. *Chem Biol Interact* 143–144:247–253
- Persson B, Kallberg Y (2013) Classification and nomenclature of the superfamily of short-chain dehydrogenases/reductases (SDRs). *Chem Biol Interact* 202:111–115
- Philippsen A, Schirmer T, Stein MA, Giffhorn F, Stetefeld J (2005) Structure of zinc-independent sorbitol dehydrogenase from *Rhodobacter sphaeroides* at 2.4 Å resolution. *Acta Crystallogr D Biol Crystallogr* 61:374–379
- Polsinelli I, Borroso L, Caliendo R, Triboli L, Esposito A, Benini S (2019a) A genome-wide analysis of desferrioxamine mediated iron uptake in *Erwinia* spp. reveals genes exclusive of the Rosaceae infecting strains. *Sci Rep* 9:2818
- Polsinelli I, Caliendo R, Salomone-Stagni M, Demitri N, Rejzek M, Field RA, Benini S (2019b) Comparison of the levansucrase from the epiphyte *Erwinia tasmaniensis* vs its homologue from the phytopathogen *Erwinia amylovora*. *Int J Biol Macromol* 127:496–501
- Pristovsek P, Sengupta K, Lohr F, Schafer B, von Trebra MW, Ruterjans H, Bernhard F (2003) Structural analysis of the DNA-binding domain of the *Erwinia amylovora* RcsB protein and its interaction with the RcsAB box. *J Biol Chem* 278:17752–17759
- Rout MP, Sali A (2019) Principles for integrative structural biology studies. *Cell* 177:1384–1403
- Salomone-Stagni M, Musiani F, Benini S (2016) Characterization and 1.57 Å resolution structure of the key fire blight phosphatase AmsI

- from *Erwinia amylovora*. Acta Crystallogr F Struct Biol Commun 72:903–910
- Salomone-Stagni M, Bartho JD, Kalita E, Rejzek M, Field RA, Bellini D, Walsh MA, Benini S (2018a) Structural and functional analysis of *Erwinia amylovora* SrlD. The first crystal structure of a sorbitol-6-phosphate 2-dehydrogenase. J Struct Biol 203:109–119
- Salomone-Stagni M, Bartho JD, Polsinelli I, Bellini D, Walsh MA, Demitri N, Benini S (2018b) A complete structural characterization of the desferrioxamine E biosynthetic pathway from the fire blight pathogen *Erwinia amylovora*. J Struct Biol 202:236–249
- Schropfer S, Bottcher C, Wohner T, Richter K, Norelli J, Rikkerink EHA, Hanke MV, Flachowsky H (2018) A single effector protein, AvrRpt2_{EA}, from *Erwinia amylovora* can cause fire blight disease symptoms and induces a salicylic acid-dependent defense response. Mol Plant-Microbe Interact 31:1179–1191
- Setser JW, Heemstra JR, Walsh CT, Drennan CL (2014) Crystallographic evidence of drastic conformational changes in the active site of a flavin-dependent N-hydroxylase. Biochemistry 53:6063–6077
- Shanker S, Schaefer GK, Barnhart BK, Wallace-Kneale VL, Chang D, Coyle TJ, Metzler DA, Huang J, Lawton JA (2017) The virulence-associated protein HsvA from the fire blight pathogen *Erwinia amylovora* is a polyamine amidinotransferase. J Biol Chem 292:21366–21380
- Smits TH, Duffy B (2011) Genomics of iron acquisition in the plant pathogen *Erwinia amylovora*: insights in the biosynthetic pathway of the siderophore desferrioxamine E. Arch Microbiol 193:693–699
- Smits TH, Rezzonico F, Kamber T, Blom J, Goesmann A, Frey JE, Duffy B (2010) Complete genome sequence of the fire blight pathogen *Erwinia amylovora* CFBP 1430 and comparison to other *Erwinia* spp. Mol Plant-Microbe Interact 23:384–393
- Strube CP, Homann A, Gamer M, Jahn D, Seibel J, Heinz DW (2011) Polysaccharide synthesis of the levansucrase SacB from *Bacillus megaterium* is controlled by distinct surface motifs. J Biol Chem 286:17593–17600
- Temel DB, Dutta K, Alphonse S, Nourikyan J, Grangeasse C, Ghose R (2013) Regulatory interactions between a bacterial tyrosine kinase and its cognate phosphatase. J Biol Chem 288:15212–15228
- Toccafondi M, Cianci M, Benini S (2014) Expression, purification, crystallization, and preliminary X-ray analysis of glucose-1-phosphate uridylyltransferase (GalU) from *Erwinia amylovora*. Acta Crystallogr F Struct Biol Commun 70:1249–1251
- Unligil UM, Rini JM (2000) Glycosyltransferase structure and mechanism. Curr Opin Struct Biol 10:510–517
- van Berkel WJ, Kamerbeek NM, Fraaije MW (2006) Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. J Biotechnol 124:670–689
- Vanneste JL (1995) *Erwinia amylovora*. In: Singh US, Singh RP, Kohmoto K (eds) Pathogenesis and host specificity in plant diseases: Histopathological. Biochemical, Genetic and molecular bases, pp 21–46
- Vogt I, Wohner T, Richter K, Flachowsky H, Sundin GW, Wensing A, Savory EA, Geider K, Day B, Hanke MV, Peil A (2013) Gene-for-gene relationship in the host-pathogen system *Malus x robusta* 5-*Erwinia amylovora*. New Phytol 197:1262–1275
- Vrancken K, Holtappels M, Schoofs H, Deckers T, Valcke R (2013) Pathogenicity and infection strategies of the fire blight pathogen *Erwinia amylovora* in Rosaceae: state of the art. Microbiology 159:823–832
- Wagstaff BA, Rejzek M, Pesnot T, Tedaldi LM, Caputi L, O'Neill EC, Benini S, Wagner GK, Field RA (2015) Enzymatic synthesis of nucleobase-modified UDP-sugars: scope and limitations. Carbohydr Res 404:17–25
- Wang D, Korban SS, Zhao Y (2009) The Rcs phosphorelay system is essential for pathogenicity in *Erwinia amylovora*. Mol Plant Pathol 10:277–290
- Wang D, Korban SS, Pusey PL, Zhao Y (2012) AmyR is a novel negative regulator of amylovoran production in *Erwinia amylovora*. PLoS One 7:e45038
- Wehland M, Kiecker C, Coplin DL, Kelm O, Saenger W, Bernhard F (1999) Identification of a RcsA/RcsB recognition motif in the promoters of exopolysaccharide biosynthetic operons from *Erwinia amylovora* and *Pantoea stewartii* subspecies *stewartii*. J Biol Chem 274:3300–3307
- Wuerges J, Caputi L, Cianci M, Boivin S, Meijers R, Benini S (2015) The crystal structure of *Erwinia amylovora* levansucrase provides a snapshot of the products of sucrose hydrolysis trapped into the active site. J Struct Biol 191:290–298
- Zhao Y, Blumer SE, Sundin GW (2005) Identification of *Erwinia amylovora* genes induced during infection of immature pear tissue. J Bacteriol 187:8088–8103
- Zhao Y, He SY, Sundin GW (2006) The *Erwinia amylovora* *avrRpt2_{EA}* gene contributes to virulence on pear and AvrRpt2_{EA} is recognized by *Arabidopsis* RPS2 when expressed in *Pseudomonas syringae*. Mol Plant-Microbe Interact 19:644–654
- Zimmermann MH, Ziegler H (1975) List of all sugars and sugar alcohols in sieve-tube exudates. In: Pirson A, Zimmermann MH (eds) Encyclopedia of plant physiology. Springer Verlag, Berlin-Heidelberg-New York, pp 480–503

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.