# Bacterial pathogen deploys the iminosugar glycosyrin to manipulate plant glycobiology

Nattapong Sanguankiattichai[1,2], Balakumaran Chandrasekar[1†], Yuewen Sheng[3], Nathan Hardenbrook[4], Werner W. A. Tabak[5], Margit Drapal[6], Farnusch Kaschani[7], Clemens Grünwald-Gruber[8], Daniel Krahn[9], Pierre Buscaill[1], Suzuka Yamamoto[10], Atsushi Kato[10], Robert Nash[11], George Fleet[12], Richard Strasser[13], Paul D. Fraser[6], Markus Kaiser[5], Peijun Zhang[3,4*], Gail M. Preston[1*] and Renier A. L. van der Hoorn[1*]

[1] Department of Biology, University of Oxford; Oxford, United Kingdom.

[2] Department of Microbiology, Faculty of Science, Mahidol University; Bangkok, Thailand.

[3] Diamond Light Source, Harwell Science and Innovation Campus; Didcot, United Kingdom.

[4] Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford; Oxford, United Kingdom.

[5] ZMB Chemical Biology, Faculty of Biology, University of Duisburg-Essen; Essen, Germany.

[6] Department of Biological Sciences, Royal Holloway University of London; Egham, United Kingdom.

[7] Analytics Core Facility Essen (ACE), Chemical Biology, Faculty of Biology, Universität Duisburg-Essen, ZMB; Essen, Germany.

[8] Core Facility Mass Spectrometry, BOKU University; Vienna, Austria

[9] Leibniz Institut für analytische Wissenschaften ISAS e.V.; Dortmund, Germany.

[10] Department of Hospital Pharmacy, University of Toyama; Toyama, Japan.

[11] Institute of Biological, Environmental and Rural Sciences/ Phytoquest Limited; Aberystwyth, United Kingdom.

[12] Chemistry Research Laboratory, Department of Chemistry, University of Oxford; Oxford, United Kingdom.

[13] Institute for Plant Biotechnology and Cell Biology, Department of Biotechnology and Food Sciences, BOKU University; Vienna, Austria

[†] Present address: Department of Biological Sciences, Birla Institute of Technology and Science, Pilani; Pilani, India

* Corresponding authors: peijun.zhang@diamond.ac.uk, gail.preston@biology.ox.ac.uk, and renier.vanderhoorn@biology.ox.ac.uk

## ABSTRACT

**The extracellular space (apoplast) in plants is a key battleground during microbial infections. To avoid recognition, bacterial model phytopathogen *Pseudomonas syringae* pv. *tomato* DC3000 produces glycosyrin to inhibit the plant-secreted β-galactosidase BGAL1, which otherwise initiates the release of immunogenic peptides from bacterial flagellin. Here, we report the structure, biosynthesis, and multifunctional roles of glycosyrin. High-resolution Cryo-EM and chemical synthesis revealed that glycosyrin is an iminosugar with a five-membered pyrrolidine ring and a hydrated aldehyde that enables monosaccharide mimicry. Glycosyrin biosynthesis is controlled by virulence regulators and branches from purine metabolism via reductase GsnB. Glycosyrin-production is common in bacteria and not only prevents flagellin recognition but also alters the extracellular glycoproteome and metabolome of infected plants, highlighting glycobiology manipulation across the plant kingdom.**

The extracellular space in plant tissues (the apoplast) is an important molecular battleground during plant-pathogen interactions (*1*). This microenvironment is colonized by bacteria, fungi and oomycetes that must have evolved various strategies to avoid recognition, suppress immune responses and manipulate host physiology. Yet most of these apoplastic plant-pathogen interactions remain to be elucidated. Our previous work on the interaction between *Nicotiana benthamiana* plants and the model bacterial pathogen *Pseudomonas syringae* revealed the role of plant apoplastic β-galactosidase BGAL1 in plant immunity (*2*). BGAL1 contributes to the hydrolytic release of immunogenic peptides from glycosylated flagella of *P. syringae* that activate plant defences (*2*). Interestingly, we also found that during infection, *P. syringae* pv. *tomato* DC3000 (*Pto*DC3000) produces a small molecule inhibitor of BGAL1 (*2*), which we named here glycosyrin. In this work, we report the molecular structure of glycosyrin, its biosynthesis and regulation, and its impact on plant glycobiology.

**Glycosyrin biosynthesis gene cluster is activated by virulence regulators**

To identify genes required for glycosyrin biosynthesis, we transformed *Pto*DC3000 Δ*hopQ1-1* (*3*) (called wild-type (WT) for glycosyrin production in this work) with *lacZ* encoding the β-galactosidase from *Escherichia coli*, which is routinely used for blue staining with X-gal (5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside). We then performed Tn5-transposon mutagenesis and selection on virulence-inducing medium containing X-gal and identified darker blue colonies of glycosyrin-deficient mutants (**Fig. 1A**). The loss of glycosyrin was confirmed in activity assays with purified LacZ and a fluorogenic substrate (**Fig. S1**) and transposon insertion sites were identified for 140 glycosyrin-deficient mutants (**Data S1**). These Tn5 insertion sites concentrated in four virulence gene regulators (*hrpR, hrpS, hrpL* and *rhpS*) and one putative glycosyrin biosynthesis gene cluster (*gsn*, locus tags PSPTO_0834-0838, new NCBI locus tags

PSPTO_RS04425-RS04445, **Fig. 1B, C**). The deletion mutant lacking the *gsn* cluster (*Δgsn*) was unable to produce the inhibitor, and transformation of this mutant with a plasmid carrying the *gsn* cluster restored inhibitor production (**Fig. 1D**). Glycosyrin production was also established in *E. coli* upon transformation with the plasmid carrying the *gsn* cluster (**Fig. 1D**). These results confirm that the *gsn* gene cluster is necessary and sufficient for glycosyrin production in bacteria.

The *gsn* cluster contains five genes (*gsnABCDE,* **Fig. 1C**). *gsnA* encodes an alcohol dehydrogenase (Pfam PF00107, PF08240) and *gsnB* encodes a reductase (PF01872). The *ΔgsnA* and *ΔgsnB* mutants were unable to produce the inhibitor, indicating that these two genes are indispensable (**Fig. S2A, B**). *gsnC* encodes a HAD family phosphatase (PF13419) and *gsnE* encodes an MFS family transporter (PF07690), possibly involved in the export of glycosyrin or transport of other metabolites related to biosynthesis. The *ΔgsnC* or *ΔgsnE* mutants had partially reduced glycosyrin production (**Fig. S2C**), likely because the genome encodes many HAD and MFS members that tend to act promiscuously and redundantly (*4*, *5*). Finally, *gsnD* encodes a protein with a domain of unknown function (DUF1349, PF07081) and has a concanavalin A-like fold (**Fig. S3A**), which is present in some lectins, glycosidases and REE1, a protein implicated in regulating galactose metabolism (*6*). However, the *ΔgsnD* mutant produced similar inhibitor levels compared to WT bacteria (**Fig. S2C**).

The promoter of the *gsn* gene cluster contains the *hrp* box, a conserved binding site for transcription factor HrpL (*7*), which is transcriptionally regulated by HrpR/S and RhpS (*8*) (**Fig. 1C**), RhpS, HrpR/S and HrpL are master regulators of virulence genes induced during plant infection, including type-III effectors such as *avrPtoB* (*8*). Indeed, expression of the *gsn* cluster was impaired in *rhpS*, *hrpR/S* and *hrpL* mutants, like *avrPtoB* (**Fig. 1E**). Consequently, as demonstrated with a *gsn:lux* reporter strain, the *gsn* cluster was transcribed from the initial to late

stages of infection (**Fig. 1F**), consistent with inhibitor production during infection (*2*). When compared to WT bacteria, the *Δgsn* mutant had reduced growth in *N. benthamiana* (**Fig. 1G**) but not in vitro (**Fig. S4A**), indicating that the *gsn* cluster contributes to virulence during infection. This is also consistent with reduced virulence for a *gsnA* mutant on *Arabidopsis thaliana* described previously (*7*). The virulence role of the *gsn* cluster, however, was not observed in the *bgal1-1* mutant of *N. benthamiana* (**Fig. S4B**), suggesting that BGAL1 is the major virulence target of glycosyrin at the tested conditions.

The *gsn* cluster is present in many strains across the major phylogroups of *P. syringae* (**Fig. S5A**), but the phylogeny of *gsnA* is incongruent with that of *P. syringae* (**Fig. S5B**). The *gsn* cluster is flanked by transposable elements and located downstream of $tRNA^{Lys}$ loci (**Fig. S5A, C**), which are typical for integrase sites (*9*). The *gsn* cluster also has a lower GC content than its neighboring regions and the genomic average (**Fig. S5D**). These findings indicate that the *gsn* cluster has been dispersed in *P. syringae* through horizontal gene transfer.

Strains carrying the *gsn* cluster produced the inhibitor, whereas strains lacking the cluster did not (**Fig. S5E**). The *hrp* box in the *gsn* cluster promoter is also conserved across *P. syringae* (**Fig. S5F**). These data suggest that different strains produce glycosyrin during infection of various plants. Indeed, the bean pathogen *P. syringae* pv. *phaseolicola* 1448A carries the *gsn* cluster and also inhibited apoplastic β-galactosidase activity during infection of *Phaseolus vulgaris* (common bean), its native host plant (**Fig. S22E**).

GsnA homologs (alcohol dehydrogenases, ADHs) are present in diverse bacterial species in different gene clusters with similar gene functions (**Fig. S6**), some of which are known to produce distinct iminosugars, potent glycosidase inhibitors with sugar-like structures containing a nitrogen instead of oxygen in the ring (**Fig. S6**) (*10–12*). However, unlike previously

characterised gene clusters, the *gsn* cluster also encodes GsnB (homolog of reductase RibD) (**Fig. S6**), suggesting that glycosyrin is produced by a different metabolic pathway.

**Glycosyrin is a hydrated iminosugar mimicking galactose**

To elucidate the molecular structure of glycosyrin, we immobilised His-tagged LacZ on a metal affinity resin to capture glycosyrin from the crude secretome of the glycosyrin-producing WT strain until LacZ saturation. Subsequent washing and elution with imidazole yielded a LacZ-glycosyrin complex with a high degree of inhibitor saturation (**Fig 2A, B**). Using cryo-electron microscopy (cryo-EM), we resolved the structure of the LacZ-glycosyrin complex at 1.9 Å resolution and detected an electron density in the active site that was absent in the negative control generated using the secretome of the *Δgsn* mutant (**Fig. 2C, D, S7**). This density revealed that glycosyrin consists of a five-membered ring with three defined chiral centers: two with putative hydroxyls and one with a putative branching geminal diol group, which likely forms upon hydration of an aldehyde group (**Fig. 2D**). We next chemically synthesised this molecule and obtained a 1.4 Å resolution structure of its complex with LacZ, which is identical to the native glycosyrin (**Fig. 2D, S7**), confirming the structure of glycosyrin.

Since glycosyrin binding to LacZ is non-covalent (**Fig. S8A**), we analysed soluble metabolites extracted from the captured LacZ-glycosyrin complex to further validate the glycosyrin structure using gas chromatography-mass spectrometry (GC-MS, **Fig. S8B**). The peaks of the synthetic glycosyrin standard were identical to those detected in native glycosyrin and absent in the *Δgsn*-derived sample (**Fig. S8C**) and these spectra were consistent with the identified structure (**Fig. S8C, S14**). We detected the same glycosyrin signals in apoplastic fluid extracted from *N. benthamiana* leaves infected with WT but not *Δgsn* mutant *P. syringae* (**Fig. S8D**).

Glycosyrin binds to the enzyme active site and closely mimics the orientation of hydroxyl groups of galactose (**Fig. 2E**)(*13*). Remarkably, the hydrated aldehyde allows the five-membered glycosyrin ring to mimic the conformation of the six-membered galactose ring. Additionally, the protonated nitrogen of glycosyrin electrostatically interacts with the conserved catalytic glutamic acid (E538), and establishes a cation-pi interaction with the aromatic tryptophan (W569) (**Fig. 2E**). Similar interactions were predicted for active site of BGAL1 (**Fig. S9**). Indeed, inhibition asays showed that synthetic glycosyrin is a potent inhibitor of both LacZ and BGAL1, with an $IC_{50}$ below that of 1-deoxy-galactonojirimycin and galactostatin, two well-known iminosugars with 6-membered rings that inhibit β-galactosidases (**Fig. S10**).

**Glycosyrin biosynthesis branches from purine pathway**

To resolve the biosynthesis pathway of glycosyrin, we first focused on GsnB as it is unique to the *gsn* cluster in our comparative genomics analysis (**Fig. S6**). GsnB is homologous to RibD reductase, which functions in the riboflavin synthesis pathway (**Fig. S11A**). The Alphafold2-predicted structure of GsnB contains conserved active site pockets similar to those in the crystal structures of RibD in complex with the substrate analog ribose-5-phosphate (R5P) and cofactor NADPH (*14*) (**Fig. S11B**), suggesting that GsnB could act on a similar substrate. This GsnB substrate likely also contains an amine group since the *gsn* cluster lacks an aminotransferase, unlike other clusters containing GsnA homologs (**Fig. S6**).

Considering that glycosyrin is a 5-carbon sugar-like molecule, we hypothesised that 5-phosphoribosyl-1-amine (PRA) might be a substrate of GsnB (**Fig. 3A, S11C**). Structural modeling also predicted NADPH and PRA binding in the active site pockets of GsnB (**Fig. S11D**). PRA is produced by PurF from 5-phosphoribosyl-1-pyrophosphate (PRPP) and is used by PurD in purine synthesis (**Fig. 3A**)(*15*). Indeed, when grown on purines to complement for

purine deficiency (**Fig. S12A**), the *ΔpurF* mutant was unable to produce glycosyrin, unlike the *ΔpurD* mutant (**Fig. 3B**). Notably, the *ΔpurD* mutant produced more glycosyrin than WT bacteria (**Fig. S12B**), likely because more PRA is available for glycosyrin production in the absence of PurD (**Fig. 3A**). These findings establish that purine intermediate PRA is the precursor for glycosyrin biosynthesis.

Given the reductase activity of RibD (**Fig. S11A**), we speculated that GsnB could similarly reduce PRA into 1-amino-1-deoxy-D-ribitol-5-phosphate (1ADRP) using cofactor NADPH (**Fig. 3C, S11C**). Subsequently, the putative phosphatase GsnC might remove the phosphate of 1ADRP to produce 1-amino-1-deoxy-D-ribitol (1ADR). Structural modeling of GsnC indeed predicts the binding of 1ADRP with its phosphate close to the conserved catalytic aspartic acid residues and cofactor $Mg^{2+}$, consistent with the expected phosphatase catalytic mechanism (**Fig. S13**) (*16*). Finally, the putative oxidase GsnA might oxidise the secondary hydroxyl group in 1ADR to produce the ketose 5-amino-5-deoxy-L-ribulose (5ADR) (**Fig. 3C**). 5ADR can spontaneously convert into the detected hydrated glycosyrin (**Fig. 3D**), as explained below.

To confirm enzymatic steps of glycosyrin biosynthesis, we produced and purified four enzymes (PurF, GsnB, GsnC and GsnA, **Fig. S15C**) and incubated them with the predicted substrate PRPP and cofactors. β-galactosidase inhibitor was produced from the mixture with all four enzymes but not when any of the four enzymes was omitted (**Fig. 3E**), demonstrating that these enzymes are required and sufficient for glycosyrin biosynthesis in vitro.

To verify the order of these reactions, we first produced intermediates from each enzymatic step in vitro and then heat-inactivated the enzymes. We were then able to produce the inhibitor from these intermediates by adding the subsequent enzymes and cofactors in the expected order (**Fig. 3F**). Furthermore, using GC-MS, the intermediates (PRA, 1ADRP and

1ADR) were detected after each enzymatic step and these intermediates were depleted upon the addition of subsequent enzymes (**Fig. S15**). To confirm the final enzymatic step, we also detected glycosyrin formation from synthetic 1ADR by GsnA (**Fig. S16, S17B**), an $NAD^+$-dependent oxidase (**Fig. S17C**). Notably, the pink color of purified GsnA indicated that cobalt ion is a preferred cofactor, which was confirmed in vitro (**Fig. S17D**), unlike the annotation of this alcohol dehydrogenase family as being zinc-dependent (*17*). Structural modeling predicted the binding of 1ADR to GsnC such that the catalytic threonine residue and cofactors $NAD^+$ and $Co^{2+}$ are in close proximity to the hydroxyl group to be oxidised on 1ADR (**Fig. S17F**) (*18*). Taken together, these results confirmed the biosynthesis pathway of glycosyrin (**Fig. 3A, C**).

Finally, we propose a chemical conversion pathway for the final steps of glycosyrin formation (**Fig. 3D**). First, the amine and ketone groups in 5ADR will react to produce the cyclic imine form of glycosyrin. This imine substructure will undergo a spontaneous Heyns rearrangement (*19*), resulting in the aldehyde form of glycosyrin. This aldehyde is then hydrated to yield the hydrate form (**Fig. 3D**). To confirm that this pathway occurs spontaneously, we chemically synthesised both the imine and aldehyde derivatives of glycosyrin (**Fig. S18**) and found that, upon removing the protecting group in water, they both spontaneously converted into the hydrate form, detected by both LC-MS and NMR as the major product (**Fig. 3G, S19**). In addition, GC-MS analysis of glycosyrin produced in vitro by GsnA from 1ADR detected both imine and aldehyde forms of glycosyrin because GC-MS was performed in anhydrous conditions (**Fig. S16**). These results demonstrate that spontaneous chemical conversions generate the hydrate form of glycosyrin.

**Glycosyrin manipulates plant glycobiology**

The originally identified target of glycosyrin is BGAL1, which acts in plant defence by facilitating the release of immunogenic peptides from glycosylated flagellin protein (*2*). BGAL1 also removes the terminal β-D-galactose from *N*- and *O*-glycans of transiently expressed recombinant proteins (*20*). Taking advantage of a terminal β-D-galactose-specific lectin called RCAI (*21*), we found that apoplastic fluids isolated from the *bgal1-1* mutant contain RCAI-positive proteins (**Fig. 4A**), indicating that BGAL1 also processes endogenous plant glycoproteins. Importantly, RCAI-positive glycoproteins also accumulated in WT plants upon infection with WT *P. syringae*, more than with the *Δgsn* mutant, while no differential RCAI-positive glycoproteins were observed in *bgal1-1* mutant plants (**Fig. 4A**), indicating that this modification of the host glycoproteome occurs through BGAL1 inhibition by glycosyrin. *N*-glycan analysis of the apoplastic proteome revealed that glycosyrin production during infection suppressed the accumulation of *N*-glycans that are likely produced by BGAL1 (**Fig. S20G**). We also purified RCAI-positive proteins from the apoplast of infected plants and found that FLA17 (Fasciclin-like arabinogalactan protein17, NbL16g25050) was differentially glycosylated upon infection without showing differential protein accumulation (**Fig. S20, Data S2**).

Moreover, untargeted metabolomics of apoplastic fluid from plants infected with WT and *Δgsn P. syringae* revealed that glycosyrin also triggered the accumulation of galactosylglycerol (873 ± 150 µM) and trehalose (46 ± 12 µM) in the apoplast of leaves infected with WT *P. syringae* (**Fig. 4B, S21, Data S3, S4**). These metabolites were not consumed by *P. syringae* in vitro (**Fig. S21D**). Their accumulation was independent of BGAL1 (**Fig. S21A**), suggesting that glycosyrin may target other glycosidases involved in glycoside processing.

Indeed, we found that glycosyrin inhibited several other glycosidases, including other β-galactosidases and β-glucosidases during infection of *N. benthamiana* leaves (**Fig. S22A-C**) and in apoplastic fluids (**Fig. S22D**), as well as other plant α- and β-glucosidases (**Fig. S22E, S23**).

Taken together, these findings demonstrate that glycosyrin is a multifunctional iminosugar produced by *P. syringae* that manipulates different host glycosidases and influences various aspects of plant glycobiology during infection (**Fig. 4C**).

## Discussion

We resolved the molecular structure of glycosyrin and its biosynthesis and regulation, and discovered its additional impact on plant glycobiology. Amongst many identified natural iminosugars and derived synthetic analogues (*10*), glycosyrin is unique by carrying an aldehyde substituent attached to the heterocyclic ring. The hydrated form of the aldehyde is stable at physiological conditions and allows glycosyrin to efficiently mimic galactose. Glycosyrin and its derivatives may have medicinal applications since iminosugars are used to treat type-II diabetes and Fabry disease (*22*, *23*). The glycosyrin biosynthesis pathway is distinct among iminosugars because it uses reductase GsnB to convert an intermediate from the purine biosynthesis pathway. In contrast, biosynthesis of other imunosugars start with an aminotransferase acting on a sugar-phosphate precursor (*11*, *12*, *24*, *25*). The presence of GsnA homologs in biosynthetic gene clusters in diverse bacterial genomes indicates the immunosugar biosynthesis is common in bacteria.

In plants, glycosyrin inhibits BGAL1, which acts in plant immunity against bacteria that carry modified viosamine (mVio) in the O-glycan that decorates flagellin (*2*). The fact that the virulence role of glycosyrin in *Pto*DC3000 is dependent on BGAL1 (**Figs. 1G** and **S4B**) and that BGAL1 suppresses bacterial growth only of strains producing mVio (*2*), indicates that at the tested conditions, glycosyrin promotes bacterial growth by avoiding the recognition of the flagellin elicitor. However, there are many strains that lack the mVio transferase gene (*vioT*) and

can still produce glycosyrin (**Fig. S5**), indicating that these strains may produce glycosyrin for a different purpose.

Glycosyrin also changes the host glycoproteome during infection. *N*-glycan analysis indicated that glycosyrin increases the accumulation of putative BGAL1 products and high-mannose, the latter is consistent with previous findings with *Pto*DC3000-infected Arabidopsis (*26*, *27*). Differential *N*-glycosylation might disturb the function of proteins such as receptor kinase MIK2, which carries an *N*-glycan that is essential for complex formation with coreceptor BAK1 (*28*). Glycosyrin also causes differential glycosylation of FLA17, which carries *O*-linked arabinogalactans (*29*), an obvious substrate for apoplastic β-galactosidases (*30*). FLA proteins are implicated in cell wall development and remodeling, as well as cell-to-cell communication and adhesion (*31*, *32*). It is also likely that glycosyrin changes cell wall properties directly by inhibiting β-galactosidases (*33–35*).

Glycosyrin production also triggers an accumulation of galactosylglycerol and trehalose in the apoplast, which is independent of BGAL1 and may result from inhibition of other glycosidases. Galactosylglycerol may originate from galactolipids in the thylakoid membrane, which are also precursors of the immune signaling molecule azelaic acid (*36*, *37*). Although pathogens often induce accumulation of host-derived sugars and metabolites in the apoplast as nutrient sources (*38*, *39*), trehalose and galactosylglycerol are not consumed by *P. syringae* in vitro (**Fig. S21D**). However, these glycosides are well-known osmolytes (*40*, *41*) and may therefore contribute to the aqueous apoplast that promotes virulence (*42*, *43*). Elevated trehalose levels may also dampen plant defence responses to promote *P. syringae* infection (*44*). In addition, by inhibiting β-glucosidases, glycosyrin may also prevent the activation of defence-related glucosides, such as salicylic acid (SA)-glucoside and cyanogenic glucosides (*45*). Thus, there are many additional mechanisms through which glycosyrin may promote virulence.

Glycosyrin production is widely distributed in *P. syringae* infecting diverse host plants, including almond, olive, leek, tomato and bean. Homologous iminosugar biosynthesis gene clusters are also present in other plant pathogens such as *Acidovorax* and *Erwinia*, and plant-associated bacteria such as *Kosakonia*, *Bacillus* and *Paenibacillus*. Taken together with the facts that glycosyrin can inhibit multiple glycosidases and that this inhibition is difficult to avoid because glycosyrin closely mimics monosaccharide substrates, it is likely that glycosidase inhibition is a common strategy used by these bacteria to manipulate the glycobiology of host plants.

## REFERENCES


1. G. Doehlemann, C. Hemetsberger, Apoplastic immunity and its suppression by filamentous plant pathogens. *New Phytol.* **198**, 1001–1016 (2013).
2. P. Buscaill, B. Chandrasekar, N. Sanguankiattichai, J. Kourelis, F. Kaschani, E. L. Thomas, K. Morimoto, M. Kaiser, G. M. Preston, Y. Ichinose, R. A. L. van der Hoorn, Glycosidase and glycan polymorphism control hydrolytic release of immunogenic flagellin peptides. *Science* **364** (2019).
3. C. Wei, B. H. Kvitko, R. Shimizu, E. Crabill, J. R. Alfano, N. Lin, G. B. Martin, H. Huang, A. Collmer, A Pseudomonas syringae pv.tomato DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. *Plant J.* **51**, 32–46 (2007).
4. D. Drew, R. A. North, K. Nagarathinam, M. Tanabe, Structures and general transport mechanisms by the major facilitator superfamily (MFS). *Chem. Rev.* **121**, 5289–5335 (2021).
5. E. Kuznetsova, M. Proudfoot, C. F. Gonzalez, G. Brown, M. V. Omelchenko, I. Borozan,

L. Carmel, Y. I. Wolf, H. Mori, A. V. Savchenko, C. H. Arrowsmith, E. V. Koonin, A. M. Edwards, A. F. Yakunin, Genome-wide analysis of substrate specificities of the *Escherichia coli* haloacid dehalogenase-like phosphatase family. *J. Biol. Chem.* **281**, 36149–36161 (2006).

6. I.-D. Choi, M.-Y. Jeong, M.-S. Ham, H.-C. Sung, C.-W. Yun, Novel Ree1 regulates the expression of ENO1 via the Snf1 complex pathway in *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* **377**, 395–399 (2008).
7. M. Vencato, F. Tian, J. R. Alfano, C. R. Buell, S. Cartinhour, G. A. DeClerck, D. S. Guttman, J. Stavrinides, V. Joardar, M. Lindeberg, P. A. Bronstein, J. W. Mansfield, C. R. Myers, A. Collmer, D. J. Schneider, Bioinformatics-enabled identification of the HrpL regulon and type III secretion system effector proteins of *Pseudomonas syringae* pv. *phaseolicola* 1448A. *Mol. Plant-Microbe Interact.* **19**, 1193–1206 (2006).
8. Y. Xie, X. Shao, X. Deng, Regulation of type III secretion system in *Pseudomonas syringae*. *Environ. Microbiol.* **21**, 4465–4477 (2019).
9. K. P. Williams, Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: Sublocation preference of integrase subfamilies. *Nucleic Acids Res.* **30**, 866–875 (2002).
10. A. A. Watson, G. W. J. Fleet, N. Asano, R. J. Molyneux, R. J. Nash, Polyhydroxylated alkaloids - Natural occurrence and therapeutic applications. *Phytochemistry* **56**, 265–295 (2001).
11. C. Nuñez, N. A. Horenstein, Functional analysis of a gene cluster from *Chitinophaga pinensis* involved in biosynthesis of the pyrrolidine azasugar DAB-1. *J. Nat. Prod.* **82**, 3401–3409 (2019).
12. L. F. Clark, J. V. Johnson, N. A. Horenstein, Identification of a gene cluster that initiates

azasugar biosynthesis in *Bacillus amyloliquefaciens*. *ChemBioChem* **12**, 2147–2150 (2011).

13. D. H. Juers, T. D. Heightman, A. Vasella, J. D. McCarter, L. Mackenzie, S. G. Withers, B. W. Matthews, A structural view of the action of *Escherichia coli* (lacZ) β-galactosidase. *Biochemistry* **40**, 14781–14794 (2001).

14. P. Stenmark, M. Moche, D. Gurmu, P. Nordlund, The crystal structure of the bifunctional deaminase/reductase RibD of the riboflavin biosynthetic pathway in *Escherichia coli* : implications for the reductive mechanism. *J. Mol. Biol.* **373**, 48–64 (2007).

15. Y. Zhang, M. Morar, S. E. Ealick, Structural biology of the purine biosynthetic pathway. *Cell. Mol. Life Sci.* **65**, 3699–3724 (2008).

16. A. Gohla, Do metabolic HAD phosphatases moonlight as protein phosphatases? *Biochim. Biophys. Acta - Mol. Cell Res.* **1866**, 153–166 (2019).

17. J. Mistry, S. Chuguransky, L. Williams, M. Qureshi, G. A. Salazar, E. L. L. Sonnhammer, S. C. E. Tosatto, L. Paladin, S. Raj, L. J. Richardson, R. D. Finn, A. Bateman, Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).

18. S. B. Raj, S. Ramaswamy, B. V. Plapp, Yeast alcohol dehydrogenase structure and catalysis. *Biochemistry* **53**, 5791–5803 (2014).

19. Z. Wang, “Heyns Rearrangement” in *Comprehensive Organic Name Reactions and Reagents* (2010).

20. R. Kriechbaum, E. Ziaee, C. Grünwald-Gruber, P. Buscaill, R. A. L. van der Hoorn, A. Castilho, BGAL1 depletion boosts the level of β-galactosylation of N- and O-glycans in *N. benthamiana*. *Plant Biotechnol. J.* **18**, 1537–1549 (2020).

21. Y. Itakura, S. Nakamura-Tsuruta, J. Kominami, N. Sharon, K. -i. Kasai, J. Hirabayashi, Systematic comparison of oligosaccharide specificity of *Ricinus communis* agglutinin I

and Erythrina lectins: a search by frontal affinity chromatography. *J. Biochem.* **142**, 459–469 (2007).

22. R. J. Nash, A. Kato, C. Y. Yu, G. W. Fleet, Iminosugars as therapeutic agents: Recent advances and promising trends. *Future Med. Chem.* **3**, 1513–1521 (2011).

23. B. G. Winchester, Iminosugars: from botanical curiosities to licensed drugs. *Tetrahedron Asymmetry* **20**, 645–651 (2009).

24. R. Miyauchi, C. Ono, T. Ohnuki, Y. Shiba, Nectrisine biosynthesis genes in *Thelonectria discophora* SANK 18292: Identification and functional analysis. *Appl. Environ. Microbiol.* **82**, 6414–6422 (2016).

25. H. E. Beal, N. A. Horenstein, Comparative genomic analysis of azasugar biosynthesis. *AMB Express* **11** (2021).

26. S. J. Kim, D. D. Bhandari, R. Sokoloski, F. Brandizzi, Immune activation during *Pseudomonas* infection causes local cell wall remodeling and alters AGP accumulation. *Plant J.* **116**, 541–557 (2023).

27. G. Beihammer, A. Romero-Pérez, D. Maresch, R. Figl, R. Mócsai, C. Grünwald-Gruber, F. Altmann, E. J. M. Van Damme, R. Strasser, *Pseudomonas syringae* DC3000 infection increases glucosylated N-glycans in Arabidopsis thaliana. *Glycoconj. J.* **40**, 97–108 (2023).

28. F. Jia, Y. Xiao, Y. Feng, J. Yan, M. Fan, Y. Sun, S. Huang, W. Li, T. Zhao, Z. Han, S. Hou, J. Chai, N-glycosylation facilitates the activation of a plant cell-surface receptor. *Nat. Plants*, doi: 10.1038/s41477-024-01841-6 (2024).

29. A. Leszczuk, P. Kalaitzis, J. Kulik, A. Zdunek, Review: structure and modifications of arabinogalactan proteins (AGPs). *BMC Plant Biol.* **23**, 45 (2023).

30. E. Nguema-Ona, M. Vicré-Gibouin, M. Gotté, B. Plancot, P. Lerouge, M. Bardor, A.

Driouich, Cell wall O-glycoproteins and N-glycoproteins: Aspects of biosynthesis and function. *Front. Plant Sci.* **5**, 1–12 (2014).

31. Y. Ma, T. Shafee, A. M. Mudiyanselage, J. Ratcliffe, C. P. MacMillan, S. D. Mansfield, A. Bacic, K. L. Johnson, Distinct functions of FASCILIN-LIKE ARABINOGALACTAN PROTEINS relate to domain structure. *Plant Physiol.* **192**, 119–132 (2023).
32. X. Wu, Y. Lai, L. Lv, M. Ji, K. Han, D. Yan, Y. Lu, J. Peng, S. Rao, F. Yan, H. Zheng, J. Chen, Fasciclin-like arabinogalactan gene family in *Nicotiana benthamiana*: genome-wide identification, classification and expression in response to pathogens. *BMC Plant Biol.* **20**, 305 (2020).
33. J. Sampedro, C. Gianzo, N. Iglesias, E. Guitián, G. Revilla, I. Zarra, AtBGAL10 is the main xyloglucan β-galactosidase in arabidopsis, and its absence results in unusual xyloglucan subunits and growth defects. *Plant Physiol.* **158**, 1146–1157 (2012).
34. T. Kotake, S. Dina, T. Konishi, Molecular cloning of a β-galactosidase from radish that specifically hydrolyzes β-(1-> 3)-and β-(1->6)-galactosyl residues of arabinogalactan protein. *Plant Physiol.* **138**, 1563–1576 (2005).
35. G. H. Dean, H. Zheng, J. Tewari, J. Huang, D. S. Young, T. H. Yeen, T. L. Western, N. C. Carpita, M. C. McCann, S. D. Mansfield, G. W. Haughn, The Arabidopsis MUM2 gene encodes a β-galactosidase required for the production of seed coat mucilage with correct hydration properties. *Plant Cell* **19**, 4007–4021 (2007).
36. C. Wang, R. Liu, G.-H. Lim, L. de Lorenzo, K. Yu, K. Zhang, A. G. Hunt, A. Kachroo, P. Kachroo, Pipecolic acid confers systemic immunity by regulating free radicals. *Sci. Adv.* **4**, 1–11 (2018).
37. Q. Gao, K. Yu, Y. Xia, M. B. Shine, C. Wang, D. Navarre, A. Kachroo, P. Kachroo, Mono- and digalactosyldiacylglycerol lipids function nonredundantly to regulate systemic

acquired resistance in plants. *Cell Rep.* **9**, 1681–1691 (2014).

38. F. El Kasmi, D. Horvath, T. Lahaye, Microbial effectors and the role of water and sugar in the infection battle ground. *Curr. Opin. Plant Biol.* **44**, 98–107 (2018).

39. X. Zhu, D. Fang, D. Li, J. Zhang, H. Jiang, L. Guo, Q. He, T. Zhang, A. P. Macho, E. Wang, Q. H. Shen, Y. Wang, J. M. Zhou, W. Ma, Y. Qiao, *Phytophthora sojae* boosts host trehalose accumulation to acquire carbon and initiate infection. *Nat. Microbiol.* **8**, 1561–1573 (2023).

40. B. C. Freeman, C. Chen, G. A. Beattie, Identification of the trehalose biosynthetic loci of *Pseudomonas syringae* and their contribution to fitness in the phyllosphere. *Environ. Microbiol.* **12**, 1486–1497 (2010).

41. N. Pade, N. Linka, W. Ruth, A. P. M. Weber, M. Hagemann, Floridoside and isofloridoside are synthesized by trehalose 6-phosphate synthase-like enzymes in the red alga *Galdieria sulphuraria*. *New Phytol.* **205**, 1227–1238 (2015).

42. C. Roussin-Léveillée, D. Mackey, G. Ekanayake, R. Gohmann, P. Moffett, Extracellular niche establishment by plant pathogens. *Nat. Rev. Microbiol.*, doi: 10.1038/s41579-023-00999-8 (2024).

43. X. F. Xin, B. Kvitko, S. Y. He, *Pseudomonas syringae* : What it takes to be a pathogen. *Nat. Rev. Microbiol.* **16**, 316–328 (2018).

44. X. Wang, Y. Du, D. Yu, Trehalose phosphate synthase 5-dependent trehalose metabolism modulates basal defense responses in *Arabidopsis thaliana*. *J. Integr. Plant Biol.* **61**, 509–527 (2019).

45. A. V. Morant, K. Jørgensen, C. Jørgensen, S. M. Paquette, R. Sánchez-Pérez, B. L. Møller, S. Bak, β-Glucosidases as detonators of plant chemical defense. *Phytochemistry* **69**, 1795–1813 (2008).

46. J. A. Vizcaíno, A. Csordas, N. Del-Toro, J. A. Dianes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-Riverol, F. Reisinger, T. Ternent, Q.-W. Xu, R. Wang, H. Hermjakob, 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–D456 (2016).

47. P. A. Bronstein, M. J. Filiatrault, C. R. Myers, M. Rutzke, D. J. Schneider, S. W. Cartinhour, Global transcriptional responses of *Pseudomonas syringae* DC3000 to changes in iron bioavailability in vitro. *BMC Microbiol.* **8**, 1–15 (2008).

48. G. L. Winsor, E. J. Griffiths, R. Lo, B. K. Dhillon, J. A. Shay, F. S. L. Brinkman, Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res.* **44**, D646–D653 (2016).

49. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

50. C. Ritz, F. Baty, J. C. Streibig, D. Gerhard, Dose-response analysis using R. *PLoS One* **10**, e0146021 (2015).

51. B. H. Kvitko, A. Collmer, "Construction of *Pseudomonas syringae* pv. *tomato* DC3000 mutant and polymutant strains" in *Methods in Molecular Biology (Clifton, N.J.)* (2011; http://link.springer.com/10.1007/978-1-61737-998-7_10), pp. 109–128.

52. A. Schäfer, A. Tauch, W. Jäger, J. Kalinowski, G. Thierbach, A. Pühler, Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145**, 69–73 (1994).

53. S. Obranić, F. Babić, G. Maravić-Vlahoviček, Improvement of pBBR1MCS plasmids, a very useful series of broad-host-range cloning vectors. *Plasmid* **70**, 263–267 (2013).

54. R. Soldan, N. Sanguankiattichai, M. Bach-Pages, I. Bervoets, W. E. Huang, G. M. Preston, From macro to micro: a combined bioluminescence-fluorescence approach to monitor bacterial localization. *Environ. Microbiol.* **23**, 2070–2085 (2021).

55. M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).

56. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J. M. Jumper, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).

57. J. Boitreaud, J. Dent, M. McPartlon, J. Meier, V. Reis, A. Rogozhnikov, K. Wu, Chai-1: decoding the molecular interactions of life. (2024). https://doi.org/10.1101/2024.10.10.615955.

58. J. Wohlwend, G. Corso, S. Passaro, M. Reveiz, K. Leidal, W. Swiderski, T. Portnoi, I. Chinn, J. Silterra, T. Jaakkola, R. Barzilay, Boltz-1 democratizing biomolecular interaction modeling. (2024). https://doi.org/10.1101/2024.11.19.624167.

59. G. Corso, A. Deng, B. Fry, N. Polizzi, R. Barzilay, T. Jaakkola, Deep confident steps to new pockets: strategies for docking generalization. (2024). https://doi.org/10.48550/arXiv.2402.18396.

60. E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H.

Morris, T. E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

61. M. van Kempen, S. S. Kim, C. Tumescheit, M. Mirdita, J. Lee, C. L. M. Gilchrist, J. Söding, M. Steinegger, Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* **42**, 243–246 (2024).

62. J. Rappsilber, M. Mann, Y. Ishihama, Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).

63. M. C. Chambers, B. Maclean, R. Burke, D. Amodei, D. L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt, J. Egertson, K. Hoff, D. Kessner, N. Tasman, N. Shulman, B. Frewen, T. A. Baker, M.-Y. Brusniak, C. Paulse, D. Creasy, L. Flashner, K. Kani, C. Moulding, S. L. Seymour, L. M. Nuwaysir, B. Lefebvre, F. Kuhlmann, J. Roark, P. Rainer, S. Detlev, T. Hemenway, A. Huhmer, J. Langridge, B. Connolly, T. Chadick, K. Holly, J. Eckels, E. W. Deutsch, R. L. Moritz, J. E. Katz, D. B. Agus, M. MacCoss, D. L. Tabb, P. Mallick, A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920 (2012).

64. V. Demichev, C. B. Messner, S. I. Vernardis, K. S. Lilley, M. Ralser, DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2020).

65. J. Cox, M. Y. Hein, C. A. Luber, I. Paron, N. Nagaraj, M. Mann, Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics*, doi: 10.1074/mcp.M113.031591 (2014).

66. S. Tyanova, T. Temu, P. Sinitcyn, A. Carlson, M. Y. Hein, T. Geiger, M. Mann, J. Cox, The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat.*

*Methods* **13**, 731–740 (2016).

67. J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, M. Mann, Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).

68. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.*, doi: 10.1038/nbt.1511 (2008).

69. J. J. Almagro Armenteros, K. D. Tsirigos, C. K. Sønderby, T. N. Petersen, O. Winther, S. Brunak, G. von Heijne, H. Nielsen, SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.*, doi: 10.1038/s41587-019-0036-z (2019).

70. M. Kessler, O. Acuto, C. Storelli, H. Murer, M. Müller, G. Semenza, A modified procedure for the rapid preparation of efficiently transporting vesicles from small intestinal brush border membranes. Their use in investigating some properties of D-glucose and choline transport systems. *Biochim. Biophys. Acta - Biomembr.* **506**, 136–154 (1978).

71. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

72. B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, R. Lanfear, IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).

73. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).

74. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic

tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).

75. C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T. L. Madden, BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).

76. R. Agarwala, T. Barrett, J. Beck, D. A. Benson, C. Bollin, E. Bolton, D. Bourexis, J. R. Brister, S. H. Bryant, K. Canese, C. Charowhas, K. Clark, M. Dicuccio, I. Dondoshansky, S. Federhen, M. Feolo, K. Funk, L. Y. Geer, V. Gorelenkov, M. Hoeppner, B. Holmes, M. Johnson, V. Khotomlianski, A. Kimchi, M. Kimelman, P. Kitts, W. Klimke, S. Krasnov, A. Kuznetsov, M. J. Landrum, D. Landsman, J. M. Lee, D. J. Lipman, Z. Lu, T. L. Madden, T. Madej, A. Marchler-Bauer, I. Karsch-Mizrachi, T. Murphy, R. Orris, J. Ostell, C. O'sullivan, A. Panchenko, L. Phan, D. Preuss, K. D. Pruitt, K. Rodarmer, W. Rubinstein, E. Sayers, V. Schneider, G. D. Schuler, S. T. Sherry, K. Sirotkin, K. Siyan, D. Slotta, A. Soboleva, V. Soussov, G. Starchenko, T. A. Tatusova, K. Todorov, B. W. Trawick, D. Vakatov, Y. Wang, M. Ward, W. J. Wilbur, E. Yaschenko, K. Zbicz, Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **44**, D7–D19 (2016).

77. S. Thakur, B. S. Weir, D. Guttman, Phytopathogen genome announcement: Draft genome sequences of 62 *Pseudomonas syringae* type and pathotype strains. *Mol. Plant-Microbe Interact.* **29**, MPMI-01-16-0013-TA (2016).

78. D. A. Baltrus, H. C. McCann, D. S. Guttman, Evolution, genomics and epidemiology of *Pseudomonas syringae*: Challenges in Bacterial Molecular Plant Pathology. *Mol. Plant Pathol.* **18**, 152–168 (2017).

79. N. A. O'Leary, M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B. Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y. Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M. Farrell, T.

Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K. Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Neill, S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz, H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb, W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy, K. D. Pruitt, Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745 (2016).

80. S. Lu, J. Wang, F. Chitsaz, M. K. Derbyshire, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, G. H. Marchler, J. S. Song, N. Thanki, R. A. Yamashita, M. Yang, D. Zhang, C. Zheng, C. J. Lanczycki, A. Marchler-Bauer, CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* **48**, D265–D268 (2020).

81. M. Yamamoto, M. Ohnishi-Kameyama, C. L. Nguyen, F. Taguchi, K. Chiku, T. Ishii, H. Ono, M. Yoshida, Y. Ichinose, Identification of genes involved in the glycosylation of modified viosamine of flagellins in pseudomonas syringae by mass spectrometry. *Genes (Basel).* **2**, 788–803 (2011).

82. S. Altschul, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).

83. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

84. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

85. T. Wagner, F. Merino, M. Stabrin, T. Moriya, C. Antoni, A. Apelbaum, P. Hagel, O. Sitsel, T. Raisch, D. Prumbaum, D. Quentin, D. Roderer, S. Tacke, B. Siebolds, E.

Schubert, T. R. Shaikh, P. Lill, C. Gatsogiannis, S. Raunser, SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* **2**, 218 (2019).

86. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. Hagen, E. Lindahl, S. H. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7** (2018).

87. K. Jamali, L. Käll, R. Zhang, A. Brown, D. Kimanius, S. H. W. Scheres, Automated model building and protein identification in cryo-EM maps. (2023). https://doi.org/10.1101/2023.05.16.541002.

88. P. Emsley, K. Cowtan, Coot : model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

89. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).

90. B. M. O'Leary, A. Rico, S. McCraw, H. N. Fones, G. M. Preston, The infiltration-centrifugation technique for extraction of apoplastic fluid from plant leaves using *Phaseolus vulgaris* as an example. *J. Vis. Exp.*, doi: 10.3791/52113 (2014).

91. S. E. Stein, P. Linstrom, Y. Mirokhin, D. Tchekhovskoi, X. Yang, W. Gary, M. O. D. Sparkman, "NIST Standard Reference Database 1A (NIST 08)" (2011).

92. J. Kopka, N. Schauer, S. Krueger, C. Birkemeyer, B. Usadel, E. Bergmuller, P. Dormann, W. Weckwerth, Y. Gibon, M. Stitt, L. Willmitzer, A. R. Fernie, D. Steinhauser, GMD@CSB.DB: the Golm Metabolome Database. *Bioinformatics* **21**, 1635–1638 (2005).

93. G. Paglia, S. Hrafnsdóttir, M. Magnúsdóttir, R. M. T. Fleming, S. Thorlacius, B. Palsson,

I. Thiele, Monitoring metabolites consumption and secretion in cultured cells using ultra-performance liquid chromatography quadrupole-time of flight mass spectrometry (UPLC-Q-ToF-MS). *Anal. Bioanal. Chem.*, doi: 10.1007/s00216-011-5556-4 (2012).

94. V. P. Kamath, J. Xue, J. J. Juarez-Brambila, C. B. Morris, R. Ganorkar, P. E. Morris, Synthesis of analogs of forodesine HCl, a human purine nucleoside phosphorylase inhibitor—Part I. *Bioorg. Med. Chem. Lett.* **19**, 2624–2626 (2009).

95. O. Fiehn, Metabolomics by gas chromatography–mass spectrometry: combined targeted and untargeted profiling. *Curr. Protoc. Mol. Biol.* **114** (2016).

96. Z. Lai, O. Fiehn, Mass spectral fragmentation of trimethylsilylated small molecules. *Mass Spectrom. Rev.* **37**, 245–257 (2018).

97. D. J. Harvey, P. Vouros, Mass spectrometric fragmentation of trimethylsilyl and related alkylsilyl derivatives. *Mass Spectrom. Rev.* **39**, 105–211 (2020).

98. K. Zheng, J. C. Lyu, E. L. Thomas, M. Schuster, N. Sanguankiattichai, S. Ninck, F. Kaschani, M. Kaiser, R. A. L. van der Hoorn, The proteome of *Nicotiana benthamiana* is shaped by extensive protein processing. *New Phytol.* **243**, 1034–1049 (2024).

99. F. Altmann, J. Helm, M. Pabst, J. Stadlmann, Introduction of a human- and keyboard-friendly N-glycan nomenclature. *Beilstein J. Org. Chem.* **20**, 607–620 (2024).

100. R. E. Parales, C. S. Harwood, Construction and use of a new broad-host-range lacZ transcriptional fusion vector, pHRP309, for Gram - bacteria. *Gene* **133**, 23–30 (1993).

101. H. Fones, C. A. R. Davis, A. Rico, F. Fang, J. A. C. Smith, G. M. Preston, Metal hyperaccumulation armors plants against disease. *PLOS Pathog.* **6**, e1001093 (2010).

102. D. H. Figurski, D. R. Helinski, Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc. Nati. Acad. Sc* **76**, 1648–1652 (1979).

103. Y. Bao, D. P. Lies, H. Fu, G. P. Roberts, An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of gram-negative bacteria. *Gene* **109**, 167–168 (1991).

**ACKNOWLEDGEMENTS**

We thank Pedro Bota for GC-MS maintenance and training; Urszula Pyzio for plant care; Sarah Rodgers, Caroline O'Brien and Patricia Bowman for technical support; Brian Mooney and Jie Huang for feedback on the manuscript. We thank Diamond Light source for access and support of the cryoEM facilities at the UK National Electron Bio-Imaging Centre (eBIC), proposal NT21004, NT29812, and BI28713. Computation was performed at the Diamond Light Source and the Oxford Biomedical Research Computing (BMRC) facility, a joint development between the Wellcome Centre for Human Genetics and the Big Data Institute (BDI) supported by Health Data Research UK and the NIHR Oxford Biomedical Research Centre.

**Funding:**

BBSRC grant BB/T015128/1 (NS, GMP, RH), BB/R017913/1 (PB, RH) and BB/W013932/1 (RH, and PDF)

ERC Advanced Grant 101019324 (RH) and 101021133 (PZ)

National Institutes of Health U54AI170791-7522 (PZ)

UK Wellcome Trust Investigator Award 206422/Z/17/Z (PZ)

Wellcome Trust Core Award Grant 203141/Z/16/Z (PZ)

Oxford Interdisciplinary Bioscience DTP BB/M011224/1 (NS, GMP)

Royal Thai Government Scholarship (NS)

European Union via HORIZON MSCA GLYCO-N training network (MK)

Japanese Society for the Promotion of Science (JSPS KAKENHI) JP24K09706 (AK)

**Author contributions:**

Conceptualization: NS, BC, GMP, RH

Methodology: NS, BC, DK, GMP, RH, YS, NH, PZ, WWAT, MK, MD, PDF, SY, AK, FK, CG, RS

Investigation: NS, BC, DK, PB, YS, NH, WWAT, MD, SY, AK, RN, GF, FK, CG, RS

Funding acquisition: AK, PDF, MK, PZ, GMP, RH

Supervision: AK, PDF, MK, PZ, GMP, RH

Writing – original draft: NS, RH

Writing – review & editing: all authors

**Competing interests:** The authors declare no competing interests.

**Data and materials availability:** All data are available in the manuscript, supplementary materials and cited references. The cryoEM density maps and corresponding atomic models have been deposited in the EMDB and PDB, respectively. The accession codes are EMDB-19182 and PDB 8RI7 for LacZ with native inhibitor (WT); EMDB-19181 and PDB 8RI6 for LacZ with *Δgsn* negative control; EMDB-19183 and PDB 8RI8 for LacZ with synthetic glycosyrin. The High-resolution HILIC-MS data have been deposited in Mendeley data (DOI: 10.17632/kbmbn6hvn7.1). The mass spectrometry proteomics data for the on-bead digestions have been deposited to the ProteomeXchange Consortium via the PRIDE (*46*) partner repository (https://www.ebi.ac.uk/pride/archive/) with the dataset identifier PXD059601. Materials and code are available upon request.

**Supplementary Information**

Materials and Methods

Fig. S1 to S23

Tables S1 to S6

Data S1 to S4

References (46-103)

MDAR Reproducibility Checklist

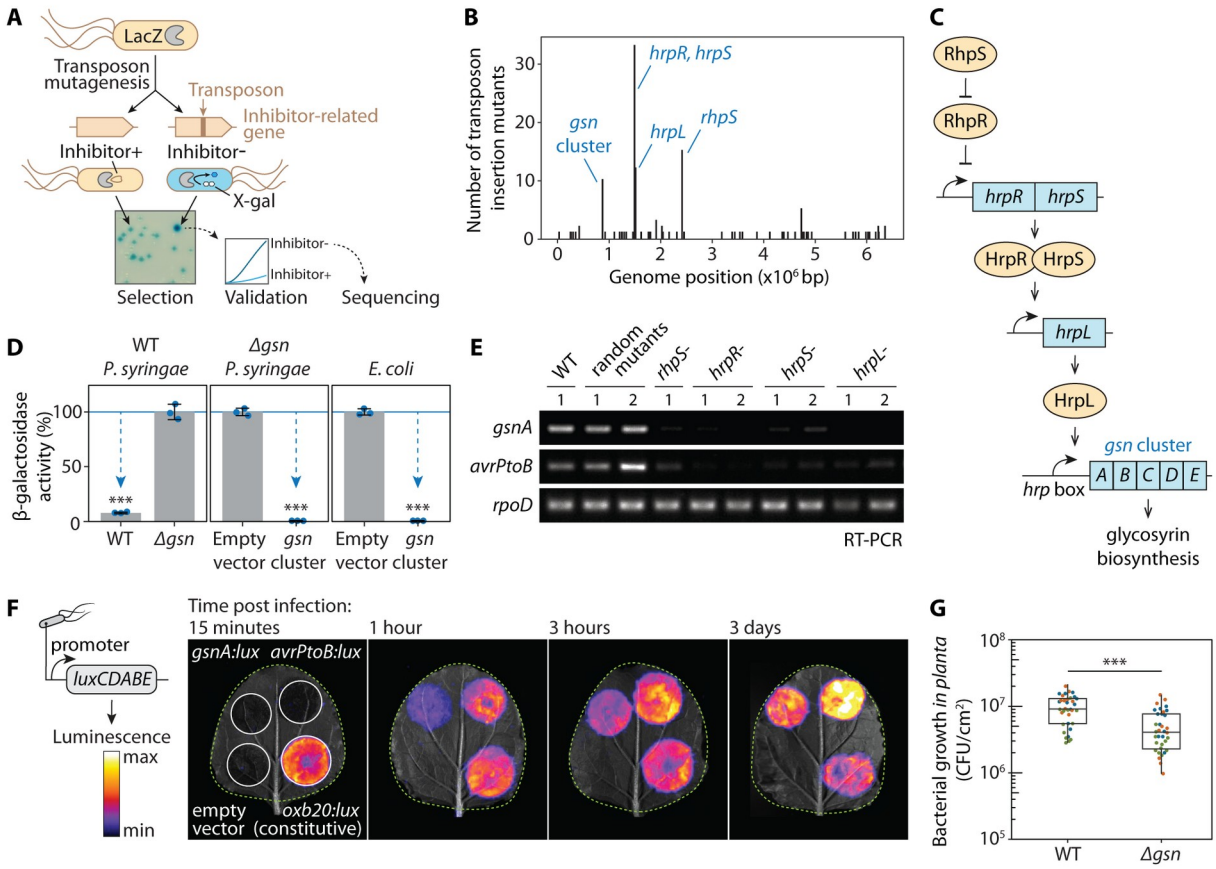


**Fig. 1.** Glycosyrin biosynthesis gene cluster and its regulators identified by forward genetics.

**(A)** Genetic screen for glycosyrin-deficient mutants. *P. syringae* expressing LacZ β-galactosidase was used to create a random transposon insertion mutant library. When plated onto minimal medium supplemented with X-gal, glycosyrin-deficient mutants cannot inhibit LacZ, resulting in a darker blue colour. These candidates were confirmed in an enzymatic assay and transposon insertion sites determined by sequencing. **(B)** Histogram of the identified transposon insertion sites along the length of the genome, highlighting genes required for glycosyrin production. **(C)** The *gsn* gene cluster confers glycosyrin biosynthesis under the control of virulence gene regulators RhpS, HrpR, HrpS and HrpL. The promotor of the *gsn* cluster contains a HrpL binding site (*hrp* box). **(D)** The *gsn* cluster confers glycosyrin biosynthesis in *P. syringae* and *E. coli*. Supernatants from bacterial cultures in minimal medium were tested for inhibitor

using the FDG hydrolysis assay with LacZ. β-galactosidase activity is reported as a percentage relative to the mean of the no-inhibitor-control (*Δgsn* or empty vector). Arrows highlight inhibition. Data are mean +/- SD (N = 3). Asterisks indicate significant difference compared to no-inhibitor-control ($P < 0.001$) from Welch's t-test. **(E)** *gsn* gene expression is dependent on *hrpR*, *hrpS*, *hrpL* and *rhpS*. RNA from bacteria grown in minimal medium was used for RT-PCR to monitor transcript levels of *gsnA*, *avrPtoB* (type III secreted effector) and *rpoD* (reference gene). **(F)** The *gsn* cluster is transcribed during infection. Bacteria carrying various promoter:*luxCDABE* reporter constructs were infiltrated into *N. benthamiana* leaves and luminescence was imaged at different time points. Signals displayed are scaled to the maximum and minimum within each image. **(G)** *gsn* cluster contributes to virulence. Bacterial strains were spray-inoculated onto *N. benthamiana* leaves and the colony forming units (CFU) per $cm^2$ were quantified at 3 days post inoculation. Results from 3 independent experiments with 12 replicates each are plotted in different colours. Asterisks indicate significant difference ($P < 0.001$) from two-way ANOVA with experiments as blocks.

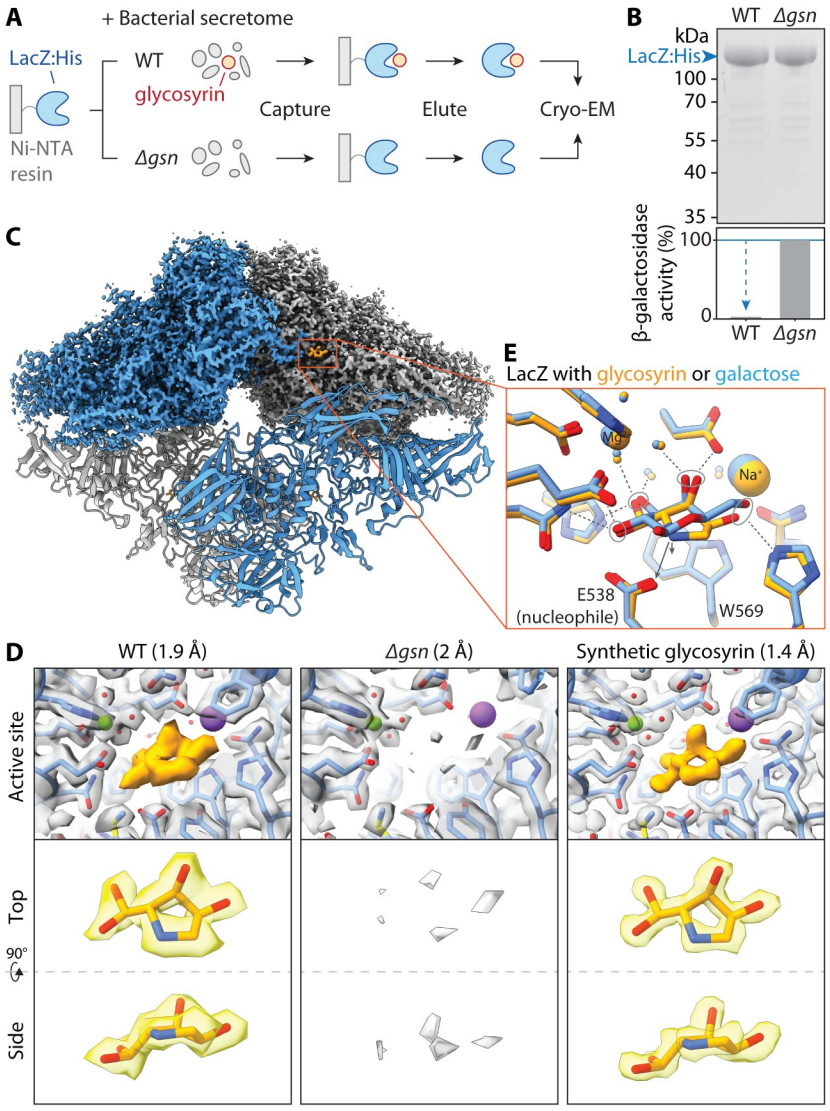


**Fig. 2.** Glycosyrin structure and mechanism revealed by high resolution cryoEM.

**(A)** LacZ-glycosyrin complex capture and downstream analyses. The Histidine-tagged β-galactosidase from *E. coli* (LacZ:His), immobilised on Ni-NTA beads was used to capture glycosyrin from the crude bacterial secretomes of glycosyrin-producing *P. syringae* (WT) or the glycosyrin-deficient mutant (*Δgsn*, negative control). After washing, the complex was eluted and used for cryo-electron microscopy (Cryo-EM). **(B)** Captured LacZ was saturated with glycosyrin. (Top) Total protein stain of eluted samples separated on SDS-PAGE. (Bottom) β-galactosidase activity of each sample measured by the FDG hydrolysis assay showing inhibition

of WT compared to *Δgsn* samples. **(C)** Structure of LacZ-glycosyrin complex from Cryo-EM. The density map is shown for the top half of the structure and a fitted model is shown for the bottom half. Each monomer of the LacZ tetramer is coloured differently. Glycosyrin is coloured orange. **(D)** Structure of glycosyrin revealed by Cryo-EM. Top: structures of the LacZ complex captures from WT or *Δgsn* strains and of LacZ incubated with synthetic glycosyrin. Density maps with fitted protein structures show the enzyme active site in the presence and absence of glycosyrin (orange). The resolution of each structure is shown in brackets. Bottom: extracted density map with fitted structure of glycosyrin from top and side view. **(E)** Glycosyrin mimics galactose binding in the active site. Overlay of structures of the LacZ active site and side chains of interacting residues in complex with glycosyrin (orange) or galactose (blue) showing similarity of overall binding pose and positioning of hydroxyl groups (circled). A positive charge on the protonated nitrogen of glycosyrin will add additional electrostatic interactions with the negatively charged catalytic glutamic acid (E538) and cation-pi interaction with the aromatic tryptophan (W569), represented by arrows. The stick representation of the molecular structure is coloured by heteroatoms (red:oxygen, blue:nitrogen) while hydrogen is not shown. Dashed lines represent hydrogen bonds. Small spheres represent water.

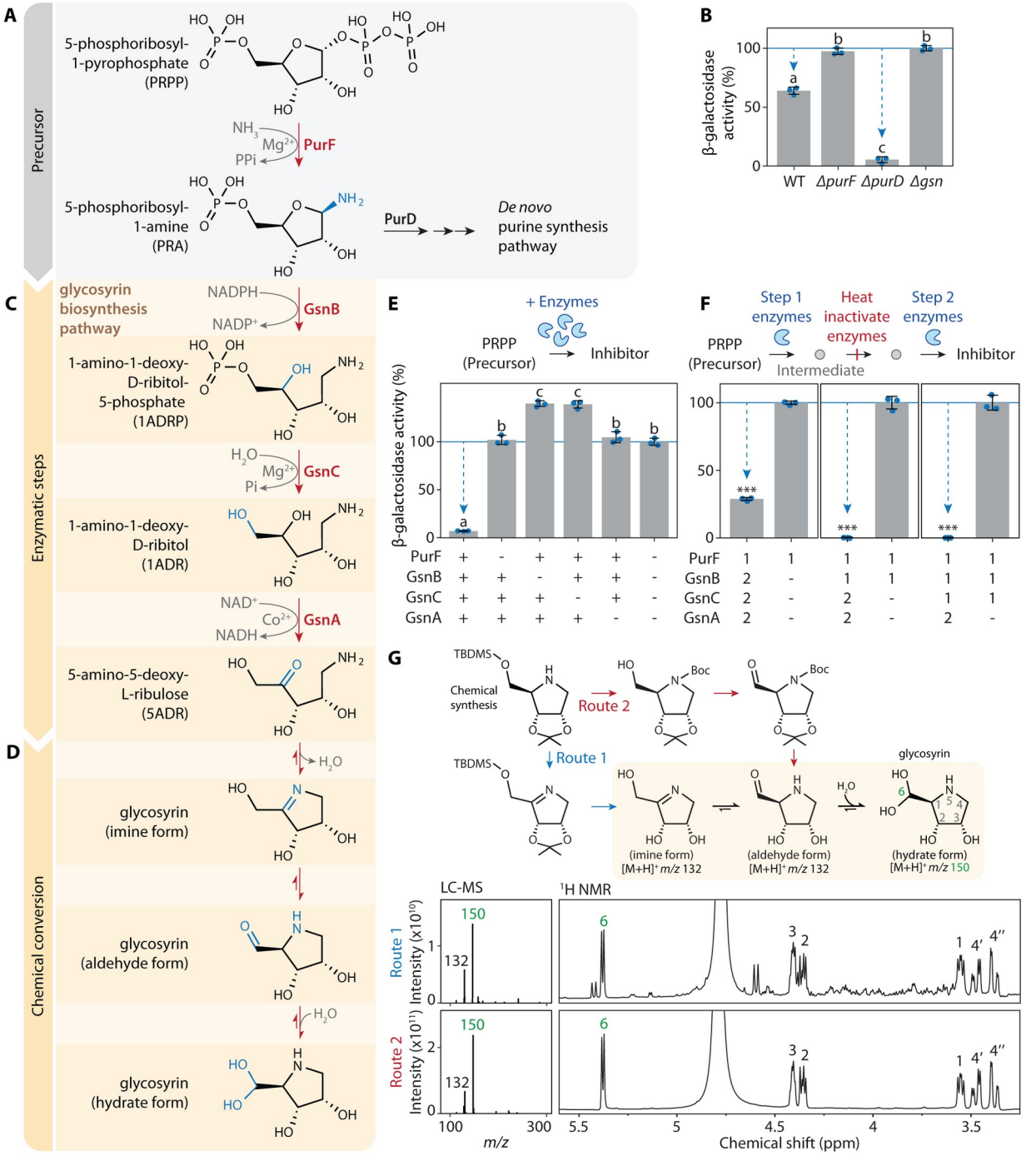


**Fig. 3.** Glycosyrin biosynthesis branches off from purine biosynthesis pathway.

**(A)** Glycosyrin biosynthesis starts from PRA, an intermediate in the purine synthesis pathway. **(B)** *purF* but not *purD* is required for glycosyrin biosynthesis. Bacterial strains (WT or knockout mutants *ΔpurF*, *ΔpurD*, *Δgsn*) were grown in virulence-inducing MG medium containing purines

overnight. **(C)** Three *gsn*-encoded enzymes convert PRA into 5ADR. **(D)** A chemical conversion pathway of 5ADR into the hydrate form of glycosyrin. **(E)** PurF, GsnB, GsnC and GsnA are required and sufficient for the biosynthesis of glycosyrin from PRPP *in vitro*. PRPP was mixed with and without purified enzymes and their cofactors. **(F)** PurF, GsnB, GsnC and GsnA act consecutively in glycosyrin biosynthesis. PRPP was mixed with purified enzymes and their cofactors in two separate steps: [1] enzymes added in the first step to produce an intermediate followed by heat inactivation of the enzymes, [2] enzymes added in the second step to complete glycosyrin biosynthesis. **(B, E, F)** Inhibitor production was tested in an FDG hydrolysis assay with LacZ. β-galactosidase activity is reported as a percentage relative to the mean of no-inhibitor-control (*Δgsn* for B, all enzymes omitted for E, enzyme 2 omitted for F). Arrows highlight inhibition. Data are mean +/- SD (N = 3). Different letters indicate groups with significant difference ($P < 0.001$) from one-way ANOVA and post-hoc Tukey HSD test. Asterisks indicate significant difference ($P < 0.001$) from Welch's t-test. **(G)** Both imine and aldehyde forms of glycosyrin spontaneously convert into the hydrate form in water. (Top) Chemical synthesis of glycosyrin using two routes, via imine or aldehyde forms. (Bottom) Products were analysed with LC-MS (left) and $H^1$-NMR (right), showing the spectra that correspond to the hydrate form. Positions within the structure of the hydrate form are numbered and labelled on the corresponding signals in NMR spectra.

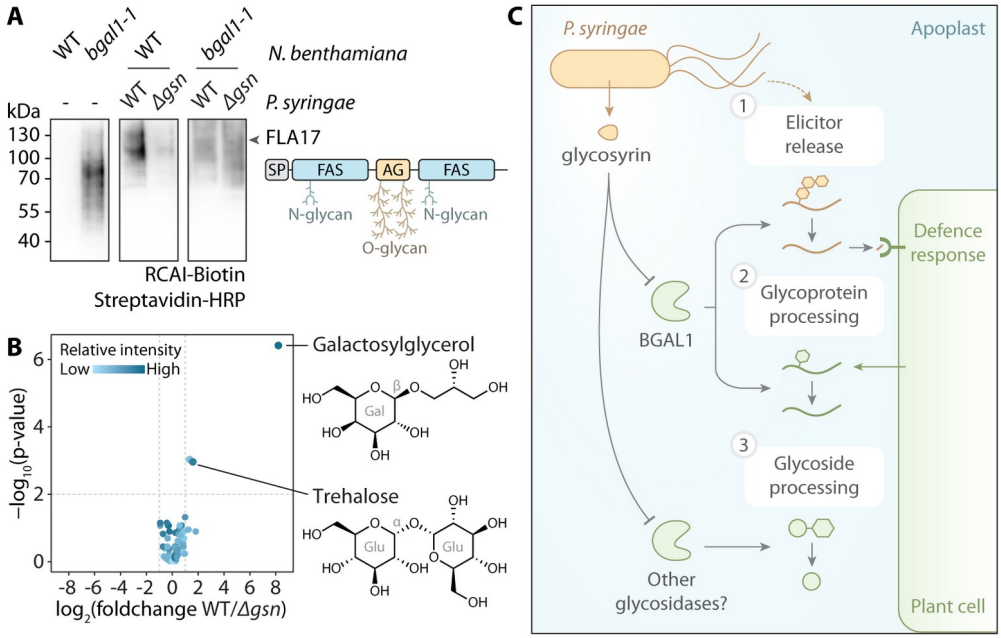


**Fig. 4.** Glycosyrin manipulates multiple aspects of extracellular plant glycobiology.

**(A)** Accumulation of RCAI-positive glycoproteins in the apoplast upon infection is dependent on BGAL1 and glycosyrin production. Proteins were extracted by acetone precipitation of apoplastic fluids from *N. benthamiana* (wild-type (WT) or BGAL1 knockout mutant (*bgal1-1*) with or without infection by *P. syringae* (WT or *Δgsn*), then separated on SDS-PAGE, blotted and probed with the RCAI lectin to detect galactose. FLA17 (fasciclin-like arabinogalactan protein17, NbL16g25050) was identified as the differentially accumulating glycoprotein (**Fig. S20**). FLA17 sequence features: signal peptide (SP), fasciclin domain (FAS, PF02469) with putative N-glycosylation sites, and disordered region with arabinogalactan protein motif (AG) for O-glycosylation. The full gel with loading control is shown in **Fig. S20B**. **(B)** Glycosyrin-dependent acculation of galactosylglycerol and trehalose in the apoplast. Volcano plot of soluble metabolites detected by GC-MS of apoplastic fluids from leaves infected with WT or *Δgsn*

mutant (**Data S3**). Glucoside components are shown with galactose (Gal), glucose (Glu) and bond configuration (α or β). **(C)** *P.* syringae produces glycosyrin to manipulate multiple aspects of glycobiology inside host plants by inhibiting plant glycosidases in the apoplast. [1] One major target of glycosyrin is BGAL1, the β-galactosidase that was previously shown to initiate the release of immunogenic fragments from flagellin (*2*). [2] Inhibition of BGAL1 by glycosyrin also interrupts apoplastic glycoprotein processing resulting in the accumulation of galactose-containing glycoproteins. [3] Glycosyrin also disrupts processing of glycosides by glycosidases other than BGAL1, resulting in the accumulation of specific glycosides in the apoplast.