



## Thioesterase-mediated side chain transesterification generates potent Gq signaling inhibitor FR900359

Cornelia Hermes, René Richarz, Daniel A. Wirtz, Julian Patt, Wiebke Hanke, Stefan Kehraus, Jan Hendrik Voss, Jim Küppers, Tsubasa Ohbayashi, Vigneshwaran Namasivayam, et al.

### ► To cite this version:

Cornelia Hermes, René Richarz, Daniel A. Wirtz, Julian Patt, Wiebke Hanke, et al.. Thioesterase-mediated side chain transesterification generates potent Gq signaling inhibitor FR900359. Nature Communications, 2021, 12 (1), pp.144. 10.1038/s41467-020-20418-3 . hal-03375694

**HAL Id: hal-03375694**

**<https://hal.science/hal-03375694>**

Submitted on 13 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Thioesterase-mediated Side Chain Transesterification**

## **Generates Potent Gq Signaling Inhibitor FR900359**

Cornelia Hermes<sup>1</sup>, René Richarz<sup>1</sup>, Daniel A. Wirtz<sup>1</sup>, Julian Patt<sup>1</sup>, Wiebke Hanke<sup>1</sup>, Stefan Kehraus<sup>1</sup>, Jan Hendrik Voß<sup>2</sup>, Jim Küppers<sup>2</sup>, Tsubasa Ohbayashi<sup>3,4</sup>, Vigneshwaran Namasivayam<sup>2</sup>, Judith Alenfelder<sup>1</sup>, Asuka Inoue<sup>5</sup>, Peter Mergaert<sup>3</sup>, Michael Gütschow<sup>2</sup>, Christa E. Müller<sup>2</sup>, Evi Kostenis<sup>1</sup>, Gabriele M. König<sup>1</sup> and Max Crüsemann<sup>1\*</sup>

1= Institute of Pharmaceutical Biology, University of Bonn, Nussallee 6, 53115 Bonn, Germany

2= Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, An der Immenburg 4, 53121 Bonn, Germany

3= University of Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Avenue de la Terrasse, 91198 Gif-sur-Yvette, 91198, France

4= Institute for Agro-Environmental Sciences, National Agriculture and Food Research Organization (NARO), Tsukuba, 305-8604, Japan

5= Graduate School of Pharmaceutical Sciences, Tohoku University, Sendai, 980-8578, Japan

## ABSTRACT

The potent and selective Gq protein inhibitor depsipeptide FR900359 (FR), originally discovered as the product of an uncultivable plant endosymbiont, is synthesized by a complex biosynthetic system comprising two nonribosomal peptide synthetase (NRPS) assembly lines. Here we characterize a cultivable bacterial FR producer, enabling detailed investigations into biosynthesis and attachment of the functionally important FR side chain. We reconstitute side chain assembly by the monomodular NRPS FrsA and the non-heme monooxygenase FrsH, and characterize intermolecular side chain transesterification to the final macrocyclic intermediate FR-Core, mediated by the FrsA thioesterase domain. We harness FrsA substrate promiscuity to generate FR analogs with altered side chains and demonstrate indispensability of the FR side chain for efficient Gq inhibition by comparative bioactivity, toxicity and docking studies. Finally, evolution of FR and side chain biosynthesis is discussed based on bioinformatics analyses. Side chain transesterification boosts potency and target affinity of selective Gq inhibitor natural products.

## INTRODUCTION

Natural product secondary metabolites comprise a plethora of structurally complex molecules purposed for communication and mediating interactions within or between species.<sup>[1]</sup> Humans take advantage of natural products, e.g. as medicines or research tools, because of their frequently observed superiority in affinity and selectivity for their targets in comparison to synthetic chemicals. This may be explained by the high likeliness of natural products to cover biologically relevant chemical space.<sup>[2,3]</sup>

In bacteria, the complete multienzymatic machinery responsible for natural product synthesis is usually encoded in biosynthetic gene clusters (BGCs). Polyketides (PK) and nonribosomal peptides (NRP) are important sources of bioactive and structurally complex natural products, synthesized from monomeric building blocks by assembly-line megaenzymes.<sup>[4,5]</sup> Nonribosomal peptide synthetases (NRPSs) contain adenylation (A), thiolation (T) and condensation (C) domains, organized into modules, to select and activate the building blocks and to extend the nascent peptide chain.<sup>[5]</sup> A type I thioesterase (TE), the final

domain in canonical NRPS and polyketide synthase (PKS) systems, catalyzes hydrolytic release from the assembly line and, prevalently, the additional formation of macrocyclic scaffolds such as macrolactones and -lactams.<sup>[6]</sup> However, in recent years, individual type I TEs catalyzing noncanonical reactions, e.g.  $\beta$ -lactone formation,<sup>[7]</sup> double bond formation<sup>[8]</sup> or epimerization<sup>[9]</sup> were characterized. While the vast majority of NRPs and PKs are synthesized by one system with a final type I TE, few characterized natural products, e.g. rimosamide<sup>[10]</sup> and the hybrid NRP/PK metabolites necroxime<sup>[11]</sup> and salinamide<sup>[12]</sup> are assembled by the joint action of two NRPS/PKS systems. Here, both systems possess a terminal TE domain, suggesting an intermolecular transesterification reaction conjugating two separately assembled molecules to the final natural product. This novel paradigm for TE functions was first explored for Sln9<sub>TE</sub> from the salinamide BGC, shown to catalyze conjugation of an acylglycine “basket handle” to a cyclic octapeptide intermediate.<sup>[12]</sup>

FR900359 (FR) (**1**, Fig. 1), first isolated in 1988 from the tropical plant *Ardisia crenata*,<sup>[13]</sup> is a cyclic depsipeptide containing several modified building blocks, among them three 3-hydroxyleucine (Hle) residues. This natural product family furthermore comprises the analogs FR1-4, mostly varying in length of Hle acyl chains<sup>[14]</sup> as well as YM-254890 (YM) and analogs thereof. YM is structurally closely related, containing one less bulky residue in the macrocycle (Thr vs. Hle).<sup>[15]</sup> FR is the most potent inhibitor of heterotrimeric Gq family proteins known to date and interacts with its molecular target with high selectivity.<sup>[16-18]</sup> Heterotrimeric G proteins are the immediate downstream effectors of G protein-coupled receptors (GPCRs), the largest family of membrane proteins in mammalian cells. Therefore, inhibition of Gq, one of the four major G protein families, by FR is a very efficient strategy to simultaneously shut down the intracellular signaling of multiple GPCRs.<sup>[16,17]</sup> FR1-4 and YM inhibit Gq with similar potency and selectivity<sup>[14]</sup>, although YM is slightly less active and has altered binding kinetics characterized by a significantly shorter residence time.<sup>[19]</sup>

Due to its high selectivity and unique mechanism of action, FR has proven to be an outstanding tool to address fundamental pharmacological questions in GqPCR-related research.<sup>[16,17,20]</sup> Recent studies suggest that FR even holds promise as potential therapy for the treatment of Gq-related diseases such as asthma<sup>[21]</sup> or uveal melanoma.<sup>[22,23]</sup>

Structure-activity relationship (SAR) studies on over 30 natural and synthetic FR and YM analogs revealed that almost the entire molecule is vital for interacting with residues in the Gαq binding pocket and thus constitutes the pharmacophore.<sup>[18, 24-27]</sup> Even subtle structural changes were associated with remarkable drops of FR potency, suggesting that the natural product scaffold is optimal for Gαq inhibition. The specialized metabolite FR was proposed to protect *A. crenata* against predators, supported by killing of insects through comparably strong inhibition of insect Gq.<sup>[28]</sup> Recent studies revealed, that the actual producer of **1** is the unculturable bacterial endosymbiont “*Candidatus Burkholderia crenata*”, located in *A. crenata* leaf nodules and harboring the *firs* NRPS BGC (Fig. 1a) on an extrachromosomal plasmid.<sup>[29]</sup> Based on bioinformatics, heterologous expression and metabolomics, we have provided a model for FR biosynthesis.<sup>[28]</sup> Accordingly, FR is assembled on two distinct NRPS systems: a cyclic heptapeptide core (**2**) is assembled by the NRPS FrsD-G, while the side chain *N*-propionyl-3-hydroxyleucine (*N*-Pp-Hle) is synthesized by the monomodular NRPS FrsA (Fig. 1b). This model implies the enzymatic conjugation of the two NRPS products as a final biosynthetic step to yield **1**.

Here we discover and characterize a cultivable and genetically tractable bacterial FR producer, enabling the detailed study of FR biosynthesis *in vivo* and *in vitro*. We dissect side chain biosynthesis and demonstrate, that the FrsA TE domain catalyzes intermolecular transesterification of the FR side chain to the cyclic depsipeptide intermediate FR-Core (**2**). This modification substantially improves inhibition of and affinity to Gq, as shown by comparative analyses. Moreover, we harnessed the enzymatic machinery to produce FR analogs with altered side chains *in vitro* and *in vivo* and perform bioinformatics analyses to identify related BGCs and to investigate *firsA* evolution.

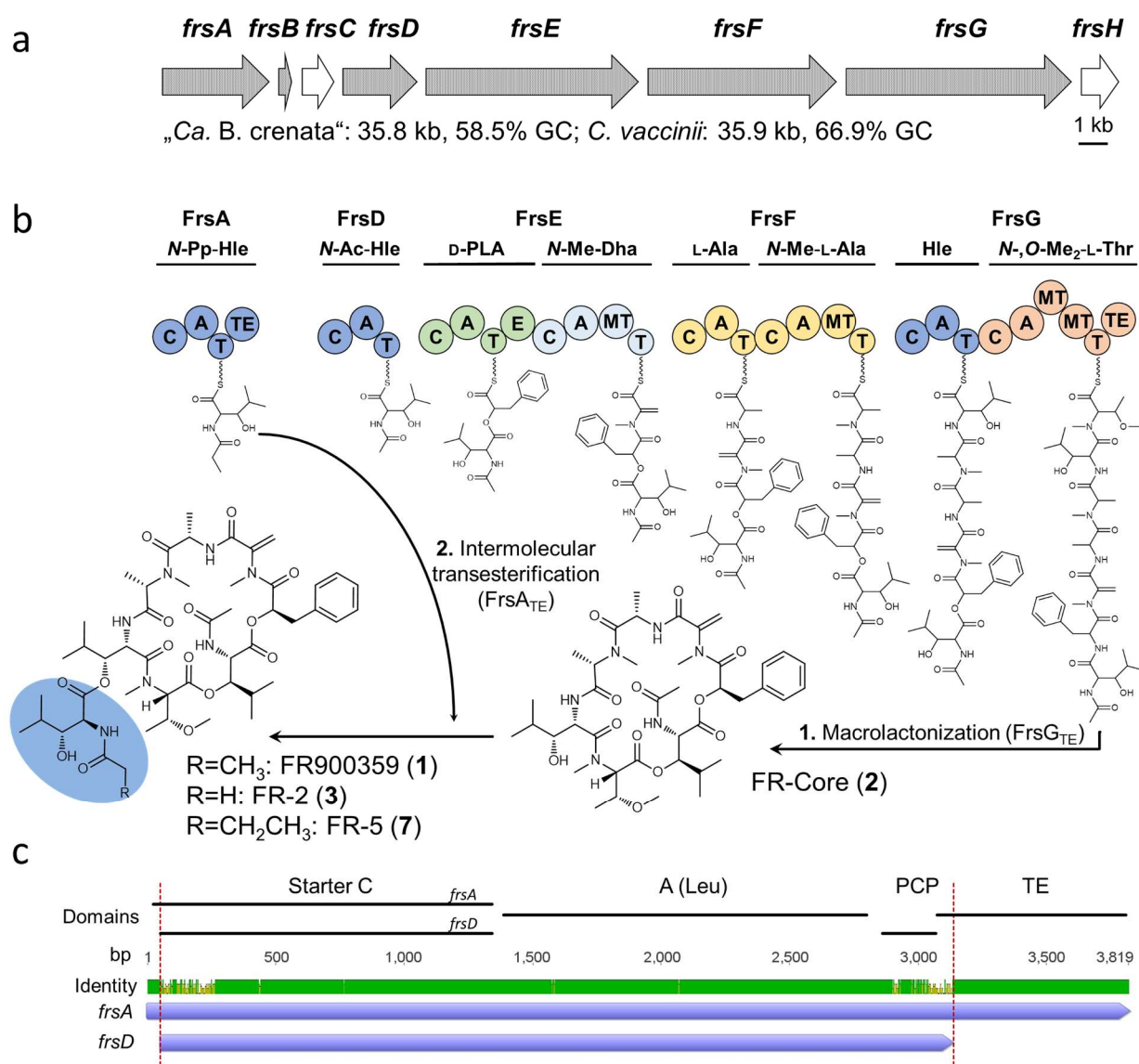
## RESULTS

### Characterization of a cultivable FR producer

To test our biosynthetic model, we aimed at studying FR biosynthesis *in vivo*. Practical use of our heterologous expression system in *Escherichia coli* was hampered by extremely low titers,<sup>[28]</sup> which could not yet be enhanced despite multiple efforts. Thus, to identify cultivable FR producing bacteria, we

searched for *frs* BGC fragments in the available databases. Using BLAST analyses, we found short DNA sequences with high similarity to parts of the *frs* BGC in *Chromobacterium vaccinii*. The identified Gram-negative, rod-shaped bacterium was previously isolated from soil and roots of cranberry plants.<sup>[30]</sup> Subsequent cultivation of *C. vaccinii* in LB medium and extraction with *n*-butanol led to the detection and isolation of **1** (2.5 mg/L, Supplementary Fig. 1) and other derivatives previously identified from *A. crenata*. In contrast to *A. crenata*, the titer of the structural analog FR-2 (**3**, Fig. 1b),<sup>[14]</sup> containing the altered side chain *N*-acetyl-3-hydroxyleucine (*N*-Ac-Hle), was almost as high as that of **1** (Supplementary Fig. 2), suggesting lower availability of propionyl-CoA in *C. vaccinii* or increased promiscuity of the responsible biosynthetic enzymes. Using MS/MS-based GNPS molecular networking,<sup>[31]</sup> we analyzed the molecular family of **1** produced by *C. vaccinii* in LB and M9 minimal medium and compared it to the one produced in *A. crenata*.<sup>[14]</sup> We observed a different profile of FR congeners in *C. vaccinii* depending on the culture medium (Supplementary Fig. 3), opening up possibilities to discover new FR derivatives with altered biological properties.

Next, we aimed at fully sequencing the *frs* BGC in *C. vaccinii*. A draft genome containing eleven (short) contigs with *frs* BGC identity was published in 2015.<sup>[32]</sup> As this sequence was insufficient to determine the total number of genes within the cluster and their structural organization, we performed a new Illumina genome sequencing of *C. vaccinii*. This still did not yield a full *frs* sequence due to the highly identical regions within the BGC, but reduced the number of contigs to six. Assuming a similar organization of the *frs* BGCs from *C. vaccinii* and “*Ca. B. crenata*”,<sup>[28]</sup> we attempted to close the five remaining gaps by employing a PCR-based approach (Supplementary Fig. 4). Ultimately, this enabled assembly of the complete *C. vaccinii frs* BGC, which is identical to the one from “*Ca. B. crenata*” in terms of gene number, organization and domain architecture (Fig. 1a, Supplementary Tables 2, 3). Only slight differences in GC content (58.5% vs. 66.9%) and the length of intergenic regions (“*Ca. B. crenata*”: 120 bp, *C. vaccinii*: 159 bp) were observed. In both *frs* BGCs, the sequences of *frsA<sub>CAT</sub>* and *frsD*, both involved in synthesis of acylated hydroxyleucines, are almost identical (94.3%) (Fig. 1c, Supplementary Fig. 5, 6).



**Figure 1: FR biosynthesis.** **a.** Organization, size and GC content of the *frs* BGCs from “Ca. B. crenata” and *C. vaccinii* (grey = NRPS, white = modifying enzyme). **b.** Biosynthetic pathway of **1**. First, the NRPS FrsDEFG together with FrsH forms a seven-membered linear peptide chain which is then hydrolyzed and cyclized by FrsG<sub>TE</sub> to **2**. Then, FrsA<sub>TE</sub> catalyzes intermolecular transesterification of *N*-Pp-Hle synthesized by FrsA<sub>CAT</sub> and FrsH (see Fig. 3a) onto **2** to yield the final product **1** (or **3**). Pp = Propionyl, Ac = Acetyl, Me = Methyl, Hle= L-3-Hydroxyleucine, PLA = Phenyllactic acid, Dha = Dehydroalanine. **c.** Nucleotide alignment of *frsA* and *frsD*. All predicted domains and identity are indicated. Prediction of domains was conducted with InterPro (EMBL).

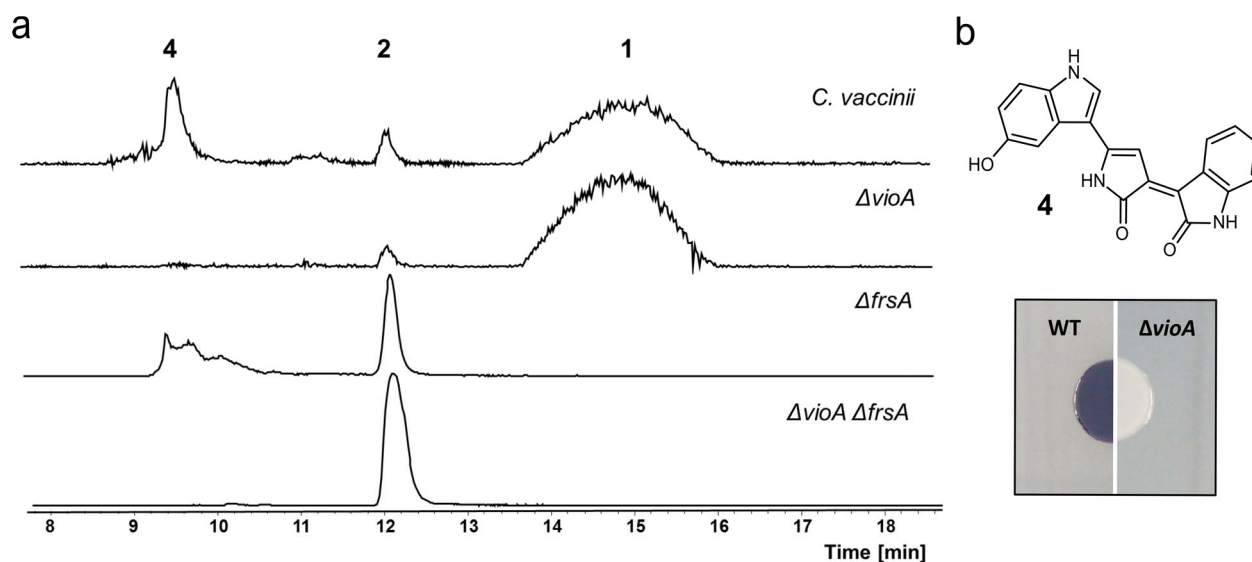
### ***frsA* knockout yields the key intermediate FR-Core**

To confirm experimentally the proposed biosynthetic pathway, we aimed at constructing an *frsA* deletion mutant of *C. vaccinii*. For this, we employed the FRT cassette from pPS858, which has been developed for the construction of *Pseudomonas* deletion mutants.<sup>[33]</sup> In a first step, a 3,708 bp internal region of *frsA* was replaced with the FRT cassette yielding the strain *C. vaccinii*  $\Delta$ *frsA*::FRT. To avoid any polar effects on downstream genes, the 3'-end of *frsA* (51 bp) was left intact. Then, the FRT cassette was removed from the genome by Flp-mediated site-specific recombination, leaving only an 86 bp scar at the site of integration (Supplementary Fig. 7). This resulted in the strain *C. vaccinii*  $\Delta$ *frsA*. Integration and loss of the FRT cassette during this process were verified by site-specific amplification of the *frsA* locus (Supplementary Fig. 8). Analysis of the butanolic extract of *C. vaccinii*  $\Delta$ *frsA* by HPLC-MS revealed the absence of **1**, while the production of **2**, before only found in traces, appeared to be enhanced (Fig. 2a). This supports our biosynthetic hypothesis for FR and confirms the role of **2** as a biosynthetic intermediate.

*C. vaccinii*  $\Delta$ *frsA* offered, however, limited potential to produce sufficient amounts of **2** for further analyses. We thus sought to optimize this strain by interrupting the biosynthesis of the purple pigment violacein (**4**, Fig. 2b), which might enhance production of **2** and would allow for an easier purification from the crude extract. To stop production of **4**, *vioA*, catalyzing the first step of violacein biosynthesis,<sup>[34]</sup> was deleted. Construction of a  $\Delta$ *frsA*/ $\Delta$ *vioA* double mutant was performed by replacing a 932 bp region of *vioA* with the FRT cassette in the *C. vaccinii*  $\Delta$ *frsA* mutant as described above. To investigate any effects of the *vioA* deletion on FR production, a  $\Delta$ *vioA* single mutant was constructed in the same fashion. Both resulting strains showed a white phenotype instead of the purple one reported for the wild type (Fig. 2b).<sup>[30]</sup> Remarkably, the  $\Delta$ *frsA*/ $\Delta$ *vioA* double mutant produced substantially higher amounts of **2** compared to the wild type or the  $\Delta$ *frsA* strain (Fig. 2a). This also applies to the production of **1** from a  $\Delta$ *vioA* strain (Fig. 2a). Subsequently, **2** was isolated and purified in preparative scale from the  $\Delta$ *frsA*/ $\Delta$ *vioA* double mutant strain. MS/MS analyses had indicated **2** to be a derivative of **1** lacking the side chain *N*-Pp-Hle.<sup>[14,28]</sup> The absence of all NMR signals of *N*-Pp-Hle in both the <sup>1</sup>H NMR and <sup>13</sup>C NMR spectra as well as an upfield shifted <sup>1</sup>H resonance of H-33 ( $\delta$  3.42) provided further evidence for this hypothesis. Finally, analyses of the 1D and 2D NMR data unambiguously proved the structure of **2**



(Fig. 1b, Supplementary Fig. 9-15, Supplementary Table 4). The configuration of **2** was deduced in analogy to the known absolute configuration of **1** [35] and in accordance with the biosynthetic pathway (Fig. 1b). We named the compound FR-Core.



**Figure 2: Characterization of *C. vaccinii* knock-out mutants.** **a.** HPLC-MS with extracted ion chromatograms (EIC) of  $m/z$ : 1002.539 (**1**),  $m/z$ : 817.430 (**2**) and  $m/z$ : 344.101 (**4**, violacein) for butanolic extracts of *C. vaccinii* and the  $\Delta vioA$ ,  $\Delta frsA$  and  $\Delta vioA \Delta frsA$  deletion mutant strains. **b.** Colony phenotype of wild type (WT) *C. vaccinii* changes, after deletion of *vioA*, from purple (production of **4**) to white.

### ***In vitro* studies on FrsA to dissect side chain biosynthesis**

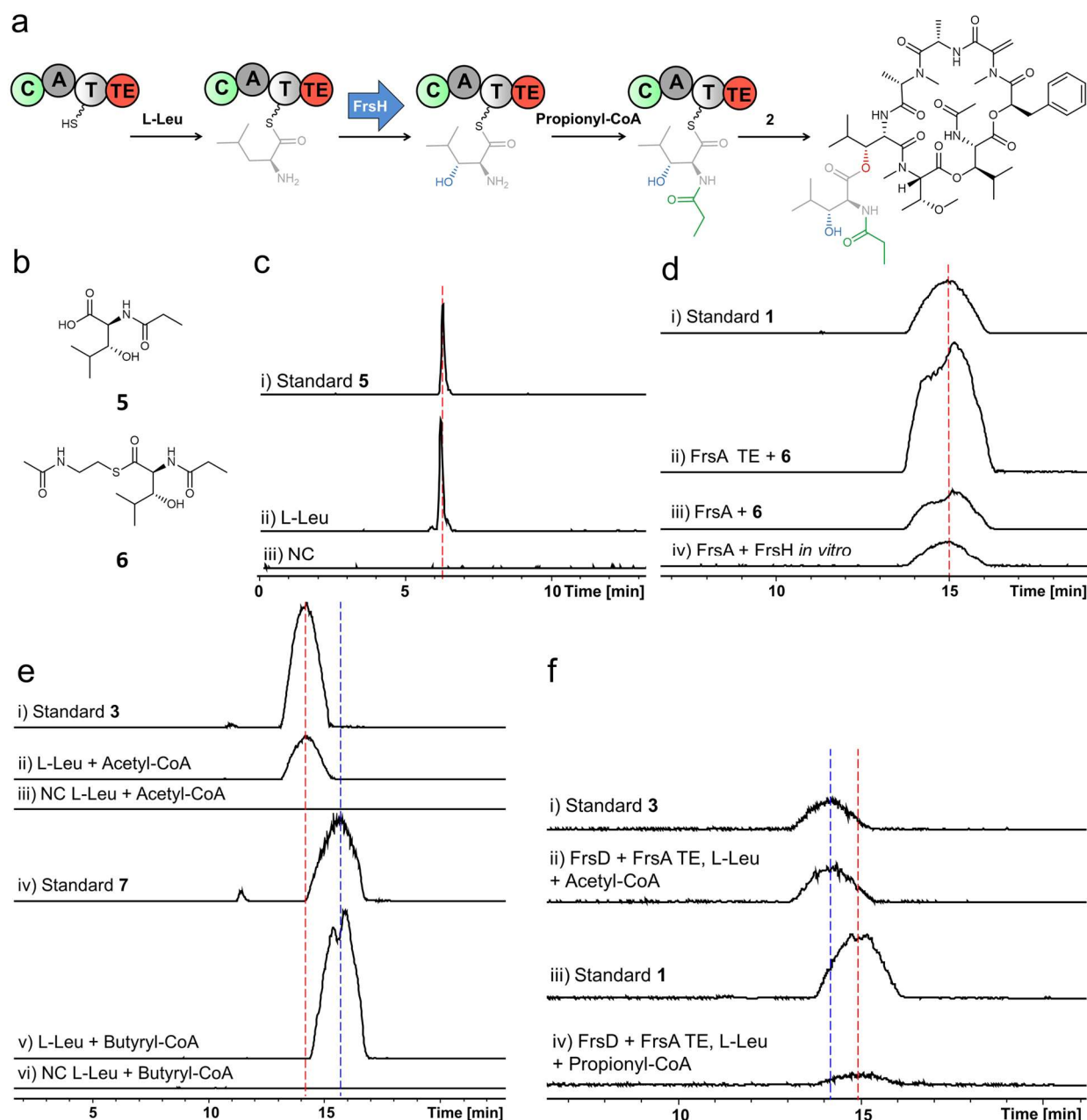
To investigate assembly of *N*-Pp-Hle (**5**) *in vitro* (Fig. 3a), the CAT tridomain of FrsA as well as the almost identical FrsD were expressed in *E. coli* BAP1 with a histidine tag and purified via affinity chromatography. Coexpression with the MbtH-like protein FrsB, resulting in coelution and copurification, was critical for FrsA and FrsD solubility and adenylating activity (Supplementary Fig. 16). The  $\gamma$ - $^{18}\text{O}_4$ -ATP exchange A domain assay [36] with the purified proteins showed favored activation of L-Leu. Despite the exact same sequence of both A domains, FrsA<sub>A</sub> exhibited noticeable broader substrate specificity than FrsD<sub>A</sub>, by also showing moderate activation of the isomers D-Leu and L-Ile (Supplementary Fig. 17). Next, we evaluated the activity of the FrsA C<sub>starter</sub> domain. Using FrsA<sub>CAT</sub>/FrsB

and the substrates L-Leu and propionyl-CoA, acylation of the leucinyl amino group was anticipated. However, only minor traces of propionylleucine were detected after alkaline hydrolysis, suggesting that  $\beta$ -hydroxylation of leucinyl-FrsA<sub>CAT</sub> had to take place before acylation. By bioinformatics, we previously identified the non-heme diiron monooxygenase FrsH as a candidate to catalyze the hydroxylation reaction.<sup>[28]</sup> To date, the only characterized member of this enzyme family is CmlA from the biosynthesis of the antibiotic chloramphenicol.<sup>[37,38]</sup> Although having only 34% sequence identity to CmlA, most residues necessary for overall structure, formation of the phosphopantetheinyl substrate binding pocket and coordination of the diiron cluster are conserved in FrsH.<sup>[28]</sup> We cloned and expressed FrsH in *E. coli* BL21(DE3) to prove its function as a  $\beta$ -hydroxylase (Supplementary Fig. 16). Complexion of iron by the purified enzyme was shown via UV/Vis spectroscopy, revealing a broad band between 300–400 nm, which diminished after transformation to the ferric state with Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> in the presence of methylviologen as a chemical electron transmitter (Supplementary Fig. 18). Upon addition of FrsH to the enzymatic assay mixture, we were able to detect **5** after hydrolysis (Fig. 3b, c) and confirmed its identity by comparison with a synthesized standard (see Supplementary Scheme 1), proving the successful assembly of the FR side chain *in vitro*. We next evaluated the substrate promiscuity of the enzymatic domains by adding different amino acid precursors to the assay. Indeed, reflecting the A domain assay results, LC-MS analyses show the synthesis of altered side chains after addition of the isomers D-Leu and L-Ile, demonstrating that also FrsH and FrsA<sub>C</sub> accept alternative substrates (Supplementary Fig. 19).

### **FrsA<sub>TE</sub> catalyzes side chain transesterification to FR-Core**

We next turned our attention to the unusual TE domain of FrsA that was hypothesized to catalyze an intermolecular transesterification, i.e. the attachment of the side chain assembled by FrsA onto the free hydroxyl group of FR-Core (**2**). To prove this experimentally, we synthesized **6**, an *N*-acetylcysteamine (SNAC) thioester of **5** (Fig. 3b, Supplementary Scheme 1, Supplementary Fig. 20-23), to mimic the phosphopantetheinyl moiety of the T domain. We overexpressed and purified FrsA<sub>TE</sub> from *E. coli* BL21 (DE3) and incubated the domain with the substrates **2** and **6**. Indeed, the formation of **1**, the final product of the biosynthetic pathway, was observed in this experiment, thereby unambiguously confirming the proposed function of FrsA<sub>TE</sub> (Fig. 3d). The same result was observed when incubating the whole

overexpressed and purified module FrsA with **2** and **6**. We next incubated purified coexpressed FrsA/FrsB and FrsH with the precursors L-Leu and propionyl-CoA and added **2** to this reaction mixture. This experimental setup gratifyingly also yielded **1** (Fig. 3d), confirming *in vitro* activity of the whole enzymatic machinery. Next, we sought to utilize the substrate promiscuity of the domains to alter side chain chemistry *in vitro*. First, propionyl-CoA was exchanged for acetyl-CoA leading to the production of **3** (Fig. 1b), an abundant compound in *C. vaccinii*, underlining the high substrate flexibility of FrsA<sub>C</sub> (Fig. 3e). After setting up the reaction with the elongated precursor butyryl-CoA, we observed the formation of a compound with a molecular weight increased by 14 Da (Fig. 3e), suggesting an unprecedented FR derivative with an *N*-butyryl-3-hydroxyleucine side chain, which was supported by MS/MS analysis (Supplementary Fig. 24). We then attempted to obtain this novel molecule in preparative quantities: Satisfyingly, feeding of butyric acid to *C. vaccinii* resulted in the formation of the same compound, which we termed FR-5 (**7**) (Fig. 1b). **7** was then isolated and characterized with 2D NMR spectroscopy (Supplementary Fig. 25-30, Supplementary Table 6) to unambiguously confirm its structure. We furthermore tested if FrsA<sub>CAT</sub> can be replaced by the highly similar FrsD (see Fig. 1c) *in vitro* and if their domains can functionally interact: We incubated FrsD with FrsA<sub>TE</sub> in our assay and also observed the formation of **1** (Fig. 3f), confirming the close relationship of FrsA and FrsD. Exchange of propionyl-CoA for acetyl-CoA in this setup led again to the formation of **3**, but interestingly in higher yields than **1**, highlighting the different substrate promiscuities of the FrsA and FrsD C<sub>starter</sub> domains, despite their very high sequence identities. FrsA<sub>C</sub> tolerates larger carbon chains, which is also reflected in the FR structure (propionyl vs. acetyl) and the precursor-directed generation of **7**.



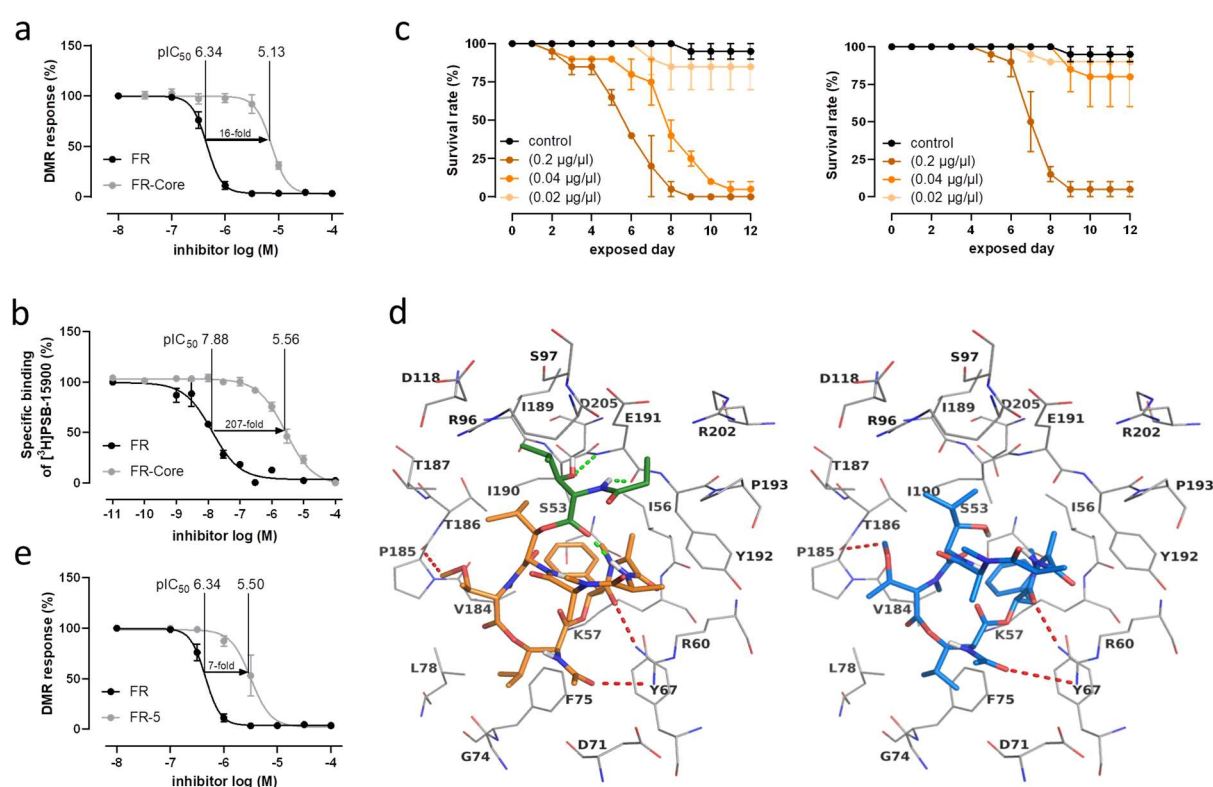
**Figure 3: *In vitro* assays with FrsA/B, FrsH and FrsA<sub>TE</sub>.** **a.** Reaction scheme of *N*-Pp-Hle formation and intermolecular transesterification onto 2 to yield 1. **b.** Structures of synthesized *N*-Pp-Hle (5) and its SNAC (6). **c.** *In vitro* production of 5 (*m/z*: 202.108). i) Synthetic 5 (1 µg/ml). ii) Enzymatic assay with purified FrsA<sub>CAT</sub>/FrsB, FrsH, incubated with L-Leu and propionyl-CoA, hydrolyzed with KOH. iii) Negative control (NC) with heat inactivated protein. **d.** *In vitro* production of 1 (*m/z*: 1002.540). i) 1 (10 µg/ml). ii) Purified FrsA<sub>TE</sub> incubated with 6 and 2. iii) Purified FrsA/B incubated with 6 and 2. iv) Purified FrsA/B, FrsH incubated with propionyl-CoA, L-Leu and 2. **e.** *In vitro* production of FR congeners 3 (*m/z*: 988.530) and 7 (*m/z*: 1016.550). i) 3 (10 µg/ml). ii) Purified FrsA/B, FrsH incubated with L-Leu, acetyl-CoA and 2. iii) Negative control. iv) 7 (10 µg/ml). v) Purified FrsA/B, FrsH incubated with L-Leu, butyryl-CoA and 2. vi) Negative control. **f.** *In vitro* FrsA/D combinations: i) 3 (1 µg/ml). ii) Production of 3 by FrsA<sub>TE</sub> and FrsD, incubated with acetyl-CoA, L-Leu and 2. iii) 1 (1 µg/ml). iv) Production of 1 by FrsA<sub>TE</sub> and FrsD, incubated with propionyl-CoA, L-Leu and 2.

## Comparative analyses reveal advantage of side chain attachment

Based on earlier SAR studies,<sup>[18,27]</sup> we reasoned that the side chain of **1** might play an important role in Gq inhibition. Thus, we first compared the Gαq inhibitory activity of **2** in comparison to **1** in CRISPR-Cas9 genome-edited HEK293 cells deficient in Gαq and Gα11 using real-time live-cell phenotypic biosensing based on dynamic mass redistribution (DMR).<sup>[39, 40]</sup> In our experimental setup, endogenous or overexpressed Gq-coupled receptors (herein: carbachol-activated muscarinic M3 receptor) show activity in DMR assays only upon re-expression of Gαq (Supplementary Fig. 31). We first examined **1** and observed the full inhibition of Gq-activity with an IC<sub>50</sub> value of 0.45 μM, which correlates well with previously determined inhibitory activities in related assays.<sup>[41]</sup> Next, we examined **2** and observed significant differences in Gq-inhibition capabilities. Its potency was about 16-fold less than that of **1** (Fig. 4a, Supplementary Fig. 32, Supplementary Table 7). In addition, competition binding assays with **1** and **2** against the radiolabeled FR-derivative [<sup>3</sup>H]PSB-15900,<sup>[19]</sup> performed with human platelet membrane preparations, revealed a 207-fold decrease in binding affinity for **2** compared to **1**. Here, the measured pIC<sub>50</sub>-value decreased from 7.88 ± 0.09 for **1** to 5.56 ± 0.04 for **2** (Fig. 4b). Despite the discrepancies between binding and function, both experiments clearly underline the vital role of side chain **5** for efficient Gq-inhibition by **1** and suggest an even greater role of the side chain as determinant for high-affinity binding of **1** to Gq. To investigate ecological relevance of the side chain, we fed **1** and **2** to nymphs of the stinkbug *Riptortus pedestris*: While high concentrations (200 μM) of both metabolites killed all insects after nine days, at 40 μM, only feeding of **1** had lethal effects, while **2** did not affect the animals (Fig. 4c), demonstrating improved *in vivo* toxicity of **1** compared to **2**.

To get insights into the molecular interaction of **2** to its Gαq binding pocket, a docking-based model of the Gq protein in complex with **2** based on the co-crystal structure of Gαq-βγ bound with YM (PDB ID: 3AH8, resolution: 2.9 Å)<sup>[42]</sup> was generated and compared with the model created for **1**. We found that the macrocyclic core structures likely display identical orientation, and the isopropyl groups of both molecules are anchored inside the binding pocket through hydrophobic interactions. In both structures, *N,O*-dimethylthreonine and D-phenyllactic acid are predicted to form hydrogen bond interactions with the side chain of R60 and the main chain of P185, respectively (Fig. 4d). In addition, the side chain **5**, which is absent in **2**, is proposed to form strong interactions (two hydrogen bonds) with the backbone of E191.

The intramolecular interactions formed between **5** and *N*-methylalanine likely stabilize the binding conformation of **1** inside the pocket. All these additional interactions are believed to be the reason for the much higher binding affinity of **1**. We finally tested the butyryl analog **7** in our DMR assay and compared it to **1** (Fig. 4e, Supplementary Fig. 32, Supplementary Table 7). Lastly, Gq inhibition capacity of **7** was sevenfold compromised as compared to that of FR. This finding corroborates the notion that the acyl chain length is another critical determinant for optimal fit into the Gαq binding pocket and, consequently, efficient inhibition of this G protein.



**Figure 4: Evaluation of FR-Core and FR-5 compared to FR.** **a.** Concentration-dependent inhibition of activated Gαq proteins by **1** and **2** as determined by label-free whole cell DMR biosensing. DMR recordings are representative (mean + s.e.m.) of at least four independent biological replicates conducted in triplicate. **b.** Competition binding experiments of **1** and **2** versus the FR-derived radiotracer [<sup>3</sup>H]PSB-15900 at human platelet membrane preparation (50 μg protein per vial), incubated at 37 °C for 1 h. **c.** Exposure of nymphs of a stink bug (*Riptortus pedestris*) to different concentrations of **1** (left) and **2** (right), survival rate was measured. **d.** Docked poses of **1** (left, represented in sticks and colored in orange, the *N*-Pp-Hle group present only in **1** is colored in green) and **2** (represented in sticks and colored in blue) in the binding pocket of the Gαq protein shown as line representation. Some of the interactions common for **1** and **2** are indicated by red dotted lines, and the interactions specific for **1** are shown as green dotted lines. Oxygen atoms are colored in red, nitrogen atoms in blue and polar

hydrogen atoms in white. **e.** Concentration-dependent inhibition of activated Gαq proteins by **1** and **7** as determined by label-free whole cell DMR biosensing (see a).

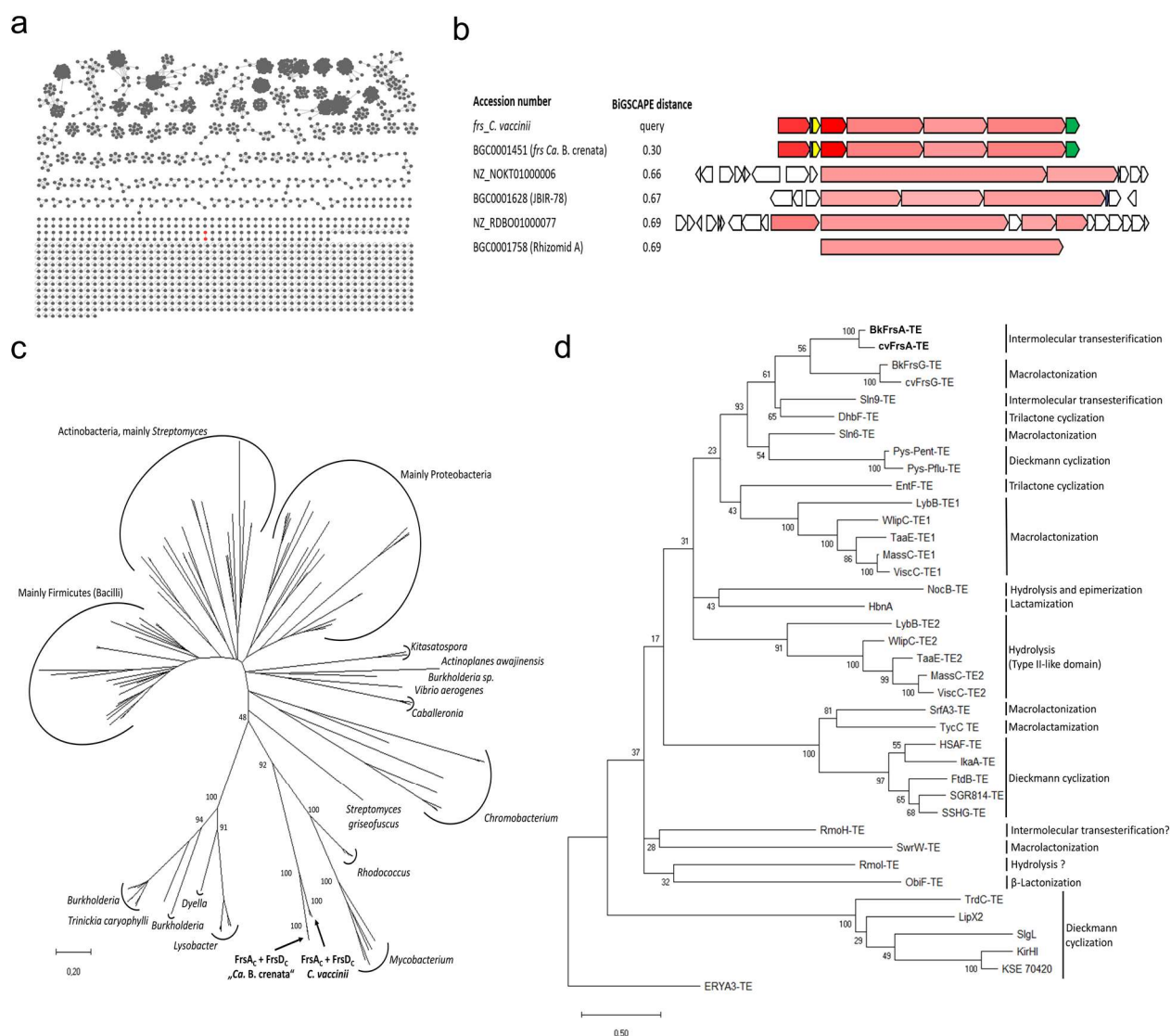
### Global bioinformatics analyses on *frs* and *frsA* evolution

To finally obtain insights into the evolution of *frs* and side chain biosynthesis, we analyzed relationships of the two architecturally identical *frs* BGCs to other NRPS systems. We therefore employed the BiG-FAM database,<sup>[43]</sup> which contains over 1.2 million complete and fragmented BGCs based on a global analysis performed by the recently developed BiG-SLiCE tool.<sup>[44]</sup> We selected all 239,899 non-fragmented NRPS BGCs in BiG-FAM, and then performed a pairwise Euclidean distance measurement based on the BiG-SLiCE algorithm using a distance cutoff  $d=1400$  (Supplementary Fig. 33). From this procedure, we recovered the 2,716 closest matching BGCs to *frs*. These were then subjected to a more sensitive BiG-SCAPE<sup>[45]</sup> analysis, also taking synteny into account. Here, no clustering of the *frs* BGCs with other BGCs was observed at a lenient distance cutoff of 0.5 (Fig. 5a). The most closely related BGCs cluster with *frs* at distances between 0.66 and 0.7, and show significant differences in gene and domain architectures and, at best, weak homologies to *frs* genes (Fig. 5b), highlighting the divergence of the *frs* BGCs from all other natural product BGCs currently in the databases.

We then had a closer look at *frsA* evolution: *frsA<sub>CAT</sub>* and *frsD* share an extremely high sequence similarity, particularly a virtually identical (99.9%) 2640 bp sequence stretch encoding the complete A domain and large parts of the C and T domains (compare Fig. 1c, Supplementary Fig. 5, 6). In addition, the A<sub>7</sub> domain encoding sequence of *frsG* (1634 bp) is also identical to the two abovementioned A domains (Supplementary Fig. 5), indicating more than one duplication event during the evolution of *frs*. We next focused on analyzing the highly similar but not identical C<sub>starter</sub> domains to shed more light onto their evolution and collected the most similar C<sub>starter</sub> domains to *frsA<sub>C</sub>* and *frsD<sub>C</sub>* from BLAST searches. Subsequent removal of redundant sequences from the list yielded a set of 149 domains (Supplementary Table 8). The phylogenetic analysis shows domains mostly clading according to taxonomic origin and irrespective of their substrate specificity, which confirms observations from a recent study (Fig. 5c,

Supplementary Fig. 34).<sup>[46]</sup> All *frs* C<sub>starter</sub> domains form a deeply rooted clade, supporting their close relationship and highlighting differences to the other domains in the database, even from the same taxa. Finally, the evolutionary origin of the transesterifying FrsA<sub>TE</sub> was investigated. Generally, TE domains share low sequence homologies and are known to exhibit broad substrate promiscuities,<sup>[6]</sup> which is regarded as a prerequisite for the evolution of enzymes to new functions and reactions.<sup>[1,47,48]</sup> Indeed, no domain with high sequence homology was found in the database; its sequence identity to FrsG<sub>TE</sub>, catalyzing macrolactonization of **2**, is 42% (Supplementary Fig. 35). Although TEs show a less-defined phylogenetic relationship with the catalytic reaction type,<sup>[6,49]</sup> a recent analysis yielded well separated clades in accordance with the mode of peptide release.<sup>[50]</sup> When we added all Frs TE domains and others to this set of sequences, a new clade emerged in the phylogenetic tree (Fig. 5d, Supplementary Table 9) containing the transesterifying Sln9 and FrsA TEs as well as DhbF<sub>TE</sub>, which catalyzes trimerization in bacillibactin.<sup>[51]</sup> FrsA<sub>TE</sub> was grouped closest to FrsG<sub>TE</sub>, indicating an evolutionary relationship of the two TEs from the same BGC. In support, structural model comparisons of FrsA<sub>TE</sub> and FrsG<sub>TE</sub> show high overall similarities and only subtle differences (Supplementary Fig. 36).





**Figure 5: Bioinformatic analyses of the *frs* BGC and selected domains. a.** BiG-SCAPE analysis network of *frs* and 2716 NRPS BGCs, obtained from a global BiG-SLiCE query against all non-fragmented NRPS from the BiG-FAM database (distance of  $d < 1400$ ) at distance cutoff 0.5. The *frs* BGCs from *C. vaccinii* and “*Ca. B. crenata*” (red) do not cluster with any other BGC. **b.** Alignment of *frs* and their 4 closest BGCs from the BiG-SCAPE analysis in Fig. 5a. Color intensity denotes the degree of homology to the respective protein of the *frs* BGC. Red: *frsD* (NRPS), blue: *frsB* (MbtH-like protein), yellow: *frsC* (dehydrogenase), green: *frsH* (hydroxylase). None of the listed BGCs contains homologues to *frsC* or *frsH*. Alignment adapted from Corason (Ref. 45). **c.** Phylogenetic tree of  $C_{\text{starter}}$  domains. Taxonomy of clades is indicated. The scale bar represents 20 substitutions per 100 amino acids. **d.** Phylogenetic tree of NRPS TE domains. The observed mode of peptide release is indicated. The scale bar represents 50 substitutions per 100 amino acids.

## DISCUSSION

The cyclic depsipeptide natural product FR (**1**) combines strong inhibition potency and high selectivity for its molecular target Gαq, leading to efficient post-receptor interruption of signal transduction, thereby causing strong effects in the targeted cells and organisms.<sup>[16]</sup> These properties have made FR a valuable pharmacological tool with promise for therapeutic applications.<sup>[16,21-23]</sup> Due to a high degree of conservation in its Gαq binding site across the animal kingdom,<sup>[52]</sup> FR inhibits not only human Gq, but also insect Gq proteins, which may be regarded as ecologically more relevant targets.<sup>[28]</sup> The FR scaffold has been isolated from diverse sources, such as plant leaf nodules,<sup>[14]</sup> soil,<sup>[15]</sup> and a marine ascidian,<sup>[53]</sup> suggesting relevance of Gq inhibitors in different ecological contexts. The cultivable bacterium *C. vaccinii* provides a new source for isolation of **1** and offers opportunities for studying its ecological role in soil. The establishment of a *C. vaccinii* knock-out system, and the subsequent creation of a violacein knock-out mutant, that produces enhanced titers of **1** and derivatives, facilitates our ongoing efforts to fully understand and to engineer FR biosynthesis for the generation of analogs with altered functionalities or G protein specificity.

The obtained experimental data on FrsA and its domains support our biosynthetic model and define **2** as the product of the heptamodular NRPS FrsD-G and substrate for side chain transesterification. A derivative without side chain has also been isolated from the YM-254890 producer,<sup>[54]</sup> suggesting analogous biosyntheses of the closely related FR and YM.

We have dissected the biosynthetic steps leading to the formation of **1** *in vitro*. FrsA<sub>CAT</sub>, in concert with the MbtH-like protein FrsB and the monooxygenase FrsH, synthesizes thioesterified *N*-Pp-Hle. Our data imply that leucynyl-FrsA<sub>T</sub> β-hydroxylation is followed by *N*-acylation. Previous studies on the FrsH homologue CmlA suggest that non-heme diiron monooxygenases are interacting with cognate NRPS domains for amino acid β-hydroxylation, when tethered to their corresponding T domains.<sup>[37,38]</sup> We are currently investigating the details of this specific protein-protein interaction. SAR studies, exemplified by the far less active analog YM-9,<sup>[18]</sup> demonstrate that the side chain hydroxyl group is essential for Gq inhibition, emphasizing the importance of the β-hydroxylase FrsH for inhibitor biosynthesis.

The congener FR-2 (**3**) contains *N*-Ac-Hle as side chain and is consistently slightly less potent than **1**, comprising *N*-Pp-Hle.<sup>[14,27]</sup> Although being almost identical to FrsD<sub>C</sub>, FrsA<sub>C</sub> prefers larger carbon chains (propionyl vs. acetyl). Further acyl chain extension to a butyryl moiety, exemplified by the “unnatural” analogue **7**, is well tolerated in feeding experiments, but not observed in nature. This modification decreases Gαq inhibitory activity significantly, confirming C<sub>2</sub> and C<sub>3</sub> as appropriate acyl chain lengths for effective Gq inhibition.

We have then characterized the conjugation of thioesterified **5** with **2** to yield the final product **1**, catalyzed by FrsA<sub>TE</sub>. The transesterifying thioesterase also interacts with the tridomain FrsD, causing a shift in production yields to **3**, thereby opening possibilities for biosynthetic engineering. In general, such noncanonical intermolecular transesterification reactions provide an opportunity for nature to create depsipeptide ester bonds at polyketide or peptide side chain hydroxyl functions. A similar reaction has been characterized in the salinamide BGC in a marine *Streptomyces* bacterium, where Sln9<sub>TE</sub>-catalyzed transesterification extends the molecular scaffold and gives rise to the potent antibiotic salinamide A.<sup>[12]</sup> The specificity of FrsA<sub>TE</sub> regarding the acceptor substrate of transesterification has yet to be fully explored. *In vitro* experiments show that the free macrocycle **2** is well accepted and converted to **1**, suggesting **2** as the native substrate, also supported by the accumulation of **2** after *frsA* knockout. However, if the enzyme would also process or even prefer a non-cyclized FR-Core substrate, remains to be determined. The minimal substrate Hle as side chain acceptor yielded no transesterification product (data not shown), implying that a larger acceptor substrate is needed for turnover. To this end, a systematic evaluation of diverse cyclic or linear peptide substrates is expected to shed more light on FrsA<sub>TE</sub> acceptor specificity. Our ongoing structural analysis of FrsA<sub>TE</sub> should however reveal its active site and mechanistic details of this reaction, which may guide further exploitation efforts of this enzyme for chemoenzymatic purposes.

The side chain attachment boosts bioactivity of the depsipeptide scaffold, as shown by comparative activity and affinity tests of **1** and **2** on Gαq proteins and insect toxicity assessments. This improvement may be explained by the accompanied docking studies revealing that the extension of chemical space generates additional interactions with the target, as well as intramolecular interactions, thereby highlighting the enormous importance of the side chain for effective Gq inhibition by **1**.

The evolution of natural product BGCs is characterized by both evolution of enzyme promiscuities and genetic events such as horizontal gene transfer, duplications, recombination and gene expansions.<sup>[55,56]</sup> We found that in both *frs* BGCs, the regions encoding the leucine activating A domains of *frsA*, *frsD* and *frsG* are identical in addition to large parts of *frsA<sub>C</sub>* and *frsD<sub>C</sub>*. The close evolutionary relationship of all C<sub>starter</sub> domains is further supported by a detailed phylogenetic analysis. It is thus conceivable that evolution of FR side chain biochemistry involves at least one duplication event, leading to biosynthesis of another acylated Hle moiety. Thioesterase phylogenetic analysis additionally suggests a close evolutionary relationship of the transesterifying FrsA<sub>TE</sub> and the macrocyclizing FrsG<sub>TE</sub>, all together pointing at the evolution of all parts of *frsA* from ancestors within the BGC.

It has been suggested, that some specialized microbial natural products may have evolved from ancestral metabolites, that had once been the end product of a biosynthetic pathway.<sup>[1,47]</sup> These ancestor molecules may have served as templates for structural variations based on evolutionary processes.<sup>[47]</sup> Bioinformatics studies, such as the reconstruction of the evolutionary history of the large natural product families of glycopeptides<sup>[57]</sup> and type II polyketides,<sup>[58]</sup> support this theory. Based on our analyses, it is tempting to speculate that the intermediate macrocycle **2** is the product of the ancestral BGC *frsB-H*, that was extended by duplication, gene expansion and further domain evolution resulting in biosynthesis and incorporation of the acylated Hle side chain, thereby encoding a metabolite with improved Gq inhibition potency. This hypothetical model would be consistent with the recently published dynamic chemical matrix evolution (DCME) hypothesis for natural products evolution,<sup>[1]</sup> when considering improved Gq inhibition properties of the resulting metabolite as fitness advantage for the producer and trait for positive selection of the altered BGC. Yet, this appealing hypothesis remains to be addressed experimentally. Other scenarios for *frsA* evolution, such as the reverse duplication of a previously existing *frsA* to yield *frsD*, its *de novo* generation or its horizontal acquisition from another organism have thus at this stage also to be taken into account and cannot be excluded with certainty.

To conclude, our global BGC analysis revealed no closely related NRPS system in the sequence databases, coining the two *frs* BGCs a rare, small and so far uniform GCF that was likely shaped by intra-BGC evolution to produce a highly specialized metabolite. Sequencing and comparative analyses of more

Gq inhibitor depsipeptide BGCs, such as those of YM and sameuramide, that await discovery, is expected to reveal more details on the evolution of this fascinating family of natural products.

## METHODS

### Cultivation and extraction of *C. vaccinii*

For the cultivation of *Chromobacterium vaccinii* MWU205, two different media (LB and M9) were tested for their influence on the production of FR and congeners. *C. vaccinii* was grown both in LB medium (tryptone 10 g/L, yeast extract 5 g/L, and NaCl 5 g/L), and M9 minimal medium (33.7 mM Na<sub>2</sub>HPO<sub>4</sub>, 22 mM KH<sub>2</sub>PO<sub>4</sub>, 8.55 mM NaCl, 9.35 mM NH<sub>4</sub>Cl, 0.4% glucose, 1 mM MgSO<sub>4</sub>, 0.3 mM CaCl<sub>2</sub>, 4.093 mM biotin, 2.965 mM thiamine HCl, 134 µM EDTA, 31 µM FeCl<sub>3</sub>, 6.2 µM ZnCl<sub>2</sub>, 0.76 µM CoCl<sub>2</sub>, 0.42 µM CoCl, 1.62 µM H<sub>3</sub>BO<sub>3</sub>, and 0.081 µM MgCl<sub>2</sub>), both media supplemented with 50 µg/mL carbenicillin to avoid contamination. Cultivation was performed at 25 °C and 200 rpm for 2 d with samples being taken for extraction every 12 h. For LB/M9 medium, 50 mL of culture were 2x extracted with 50 mL *n*-butanol by shaking, phase separation for 24 h and subsequent centrifugation (4,000 rpm, 5 min).

### HPLC-MS/MS analysis

Mass spectra were recorded on a micrOTOF-QII mass spectrometer (Bruker) with ESI-source coupled with a HPLC Dionex Ultimate 3000 (Thermo Scientific) using an EC10/2 Nucleoshell C<sub>18</sub> 2.7 µm column (Macherey-Nagel). The column temperature was 25 °C. MS data were acquired over a range from 100-3,000 *m/z* in positive mode. Auto MS/MS fragmentation was achieved with rising collision energy (35-50 keV over a gradient from 500-2000 *m/z*) with a frequency of 4 Hz for all ions over a threshold of 100. HPLC begins with 90 % H<sub>2</sub>O containing 0.1% AcOH. The gradient starts after 1 min to 100% acetonitrile (0.1% AcOH) in 20 min. 5 µl of a 1 mg/mL sample solution (MeOH) was injected to a flow of 0.3 mL/min.

### Molecular networking

All MS/MS data were converted to .mzXML file format using the software Bruker Compass DataAnalysis 4.2 (Bruker Daltonik GmbH). The data were transferred onto the GNPS server

([gnps.ucsd.edu](http://gnps.ucsd.edu)) and molecular networking was performed via the GNPS data analysis workflow using the spectral clustering algorithm.<sup>[31]</sup> The data were clustered with MS-Cluster with a parent mass tolerance of 0.02 Da and an MS/MS fragment ion tolerance of 0.02 Da to create consensus spectra. Further, consensus spectra that contained less than two spectra were discarded. A network was created where edges were filtered to have a cosine score above 0.7 and more than six matched peaks. Further edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. The spectra in the network were searched against GNPS' spectral libraries. The library spectra were filtered in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.7 and at least six matched peaks. Results were visualized in Cytoscape 3.7.1 ([cytoscape.org](http://cytoscape.org)) using the solid layout. Nodes represent parent masses, while the color of the node displays the origin of production, the size of the node is proportional to the number of spectra, and edge thickness corresponds to cosine score. Identified ions for FR and derivatives, when already known and published, are indicated in the respective nodes.

### **Isolation and structure elucidation of FR-Core (2) and FR-5 (7)**

NMR spectra were recorded either on a Bruker Ascend 600 NMR spectrometer with a Prodigy cryoprobe operating at 600 MHz (<sup>1</sup>H) and 150 MHz (<sup>13</sup>C) using acetonitrile-*d*<sub>3</sub> as solvent (Deutero GmbH, 99.8% D) or on a Bruker Avance 300 DPX NMR spectrometer at 300 MHz (<sup>1</sup>H) and 75 MHz (<sup>13</sup>C) using CDCl<sub>3</sub> as solvent (Deutero GmbH, 99.8% D). NMR spectra were processed using Bruker TopspinVersion 1.3 and MestReNova 8.0.1 software packages. Spectra were referenced to residual solvent signals with resonances at  $\delta_{\text{H/C}}$  1.93/117.7 for acetonitrile-*d*<sub>3</sub> or  $\delta_{\text{H/C}}$  7.26/77.0 for CDCl<sub>3</sub>. HPLC was carried out using a Waters HPLC system, controlled by Waters Millennium software, consisting of a 600E pump, a 996 PDA detector, and a 717 plus autosampler or on a Waters Breeze HPLC system equipped with a 1525 $\mu$  dual pump, a 2998 photodiode array detector, and a Rheodyne 7725i injection system. For isolation of **2**, the crude material (butanolic extract from 4 L culture of *C. vaccinii*  $\Delta$ *frsA*/ $\Delta$ *vioA*) was fractionated on a Grace Reveleris X2 flash chromatography system with integrated evaporative light scattering (ELSD) and UV-Vis detection via a Reveleris C<sub>18</sub> flash column (220 g, 40  $\mu$ m). A stepwise gradient solvent system of increasing polarity and a flow rate of 65 mL/min was used starting with 50/50 H<sub>2</sub>O/MeOH for 13 min, then changing to 30/70 H<sub>2</sub>O/MeOH within 1 min and hold again for 13 min. The gradient was then

changed within 1 min to 25/75 H<sub>2</sub>O/MeOH and held for 25 min, then within 1 min to 20/80 H<sub>2</sub>O/MeOH, held for 13 min, then within 1 min to 15/85 H<sub>2</sub>O/MeOH and held for 25 min. Finally, the gradient was changed within 1 min to 100% MeOH and held for additional 10 min. According to the measured ELSD and UV signals, a FR-Core containing fraction was collected at 70 min. Final purification was done by HPLC with a semi-preparative Macherey-Nagel Nucleodur C<sub>18</sub> column (250 x 8 mm, 5 µm) using an isocratic elution with 19/81 H<sub>2</sub>O/MeOH (flow 2.0 mL/min). Pure **2** was isolated as a white powder (*t<sub>R</sub>*: 12 min, 5 mg). For the isolation of **7**, M9 medium was supplemented with 20 mM butyric acid and *C. vaccinii* was cultivated 48 h as described above. The crude material (butanolic extract from 4.5 L culture of *C. vaccinii*) was fractionated as described above. A **7** containing fraction was collected at around 80 min. Final purification was done as described above. Pure **7** was isolated as a white powder (*t<sub>R</sub>*: 25 min, 10 mg).

FR-Core (**2**): white, solid; <sup>1</sup>H NMR (acetonitrile-*d*<sub>3</sub>, 600 MHz) and <sup>13</sup>C NMR (acetonitrile-*d*<sub>3</sub>, 150 MHz) data Supplementary Table 4; (+)-HRESIMS *m/z* 817.4295 [M+H]<sup>+</sup> (calcd for C<sub>40</sub>H<sub>61</sub>N<sub>6</sub>O<sub>12</sub>, 817.4342).

FR-5 (**7**): white, solid; <sup>1</sup>H NMR (CDCl<sub>3</sub> 300 MHz) and <sup>13</sup>C NMR (CDCl<sub>3</sub>, 75 MHz) data Supplementary Table 6; (+)-HRESIMS *m/z* 1016.5507 [M+H]<sup>+</sup> (calcd for C<sub>50</sub>H<sub>78</sub>N<sub>7</sub>O<sub>15</sub>, 1016.5556).

### **Cloning of *frs* genes and domains**

Oligonucleotide primers were synthesized by Eurofins Genomics. Primers listed in Supplementary Table 5 were used for the amplification of the different regions of *frsA* (encoding either FrsA complete, FrsA<sub>CAT</sub> or FrsA<sub>TE</sub>) and the genes *frsB*, *frsD* and *frsH* from genomic DNA of *C. vaccinii*. The resulting fragments were cloned into restriction sites of pCDFDuet-1 (*frsB*) or pET28a (all others) via standard ligation cloning. The resulting constructs were isolated from *E. coli* Alpha-Select Silver (Bioline), analyzed via restriction digest and verified by Sanger DNA sequencing. The final plasmids were then transformed into *E. coli* BL21 Star (DE3) (Thermo Scientific Fisher) or *E. coli* BAP1 via heat shock for overexpression of recombinant hexahistidine-tagged proteins.

### **Protein expression and purification**

Expression of constructs including T domains was performed in *E. coli* BAP1<sup>[59]</sup>, to ensure *in vivo* phosphopantetheinylation of the T domains. All other constructs were expressed in *E. coli* BL21 (DE3).

All constructs containing an A domain were coexpressed with pCDF-Duet-*frsB* to ensure solubility and *in vitro* activity of the A domain. For the expression of FrsH, TB media was supplemented with 25  $\mu$ M Fe(III)-citrate. A glycerol stock culture of *E. coli* BL21 (DE3)/BAP1 harboring the selected plasmids was used to inoculate an overnight culture (37 °C, 200 rpm) of LB, supplemented with appropriate antibiotics. This was then used to inoculate 250 mL terrific broth (TB), supplemented with appropriate antibiotics and incubated at 37 °C, 200 rpm for approx. 4 h ( $OD_{600} = 0.9$ ), then the culture was chilled on ice. Protein expression was induced upon the addition of 0.4 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG). Then the cultures were incubated at 16 °C, 200 rpm for a further 16 h. Finally, the cells were harvested via centrifugation (4,000 g, 4 °C, 15 min). The pellet was resuspended in 2.5 mL 10 mM imidazole lysis buffer per g pellet (50 mM  $NaH_2PO_4$ , 300 mM NaCl, pH 8). The cells were lysed on ice with a sonicator in 10-second intervals. The lysate was centrifuged (10,000 g, 4 °C, 10 min). The supernatant was incubated on ice with 1 mL Ni-NTA-agarose (Qiagen). The suspension was kept in light movement. After 1 h, the suspension was filtered with a propylene column (Qiagen 1 mL). The Ni-NTA agarose was washed with 4 mL of a 20 mM imidazole buffer and 4 mL of a 35 mM imidazole buffer and eluted with 2.5 mL of a 250 mM imidazole buffer. All protein fractions were analyzed by SDS-PAGE with a Tris/Tricine buffer system.<sup>[60]</sup> The elution fraction was rebuffed with PD10 columns (GE). The eluate was concentrated with Vivaspin 500 columns (Sartorius) MWCO 30 kDa or 10 kDa.

#### **$\gamma$ -<sup>18</sup>O<sub>4</sub>-ATP-exchange assay for A domains<sup>[36]</sup>**

Two stock solutions were prepared for the assay. Solution 1: 3 mM of the respective amino acid, 15 mM pyrophosphate, 20 mM Tris pH 7.5; solution 2: 3 mM  $\gamma$ -<sup>18</sup>O<sub>4</sub>-ATP (Cambridge Isotope Laboratories), 15 mM MgCl in 20 mM Tris pH 7.5. 2  $\mu$ L of solution 1, 2  $\mu$ L of solution 2 and 2  $\mu$ L of the 5  $\mu$ M concentrated protein solution (buffer 20 mM Tris, 5 % Glycerol (v/v) pH 7.5) was mixed and incubated at 22 °C for 1.5 h. The reaction was stopped with the addition of 6  $\mu$ L 9-aminoacridine in acetone (10 mg/mL). The samples were analyzed with MALDI-TOF-MS (Bruker AutoFlex III) in negative mode. Absolute substrate conversion was calculated by dividing the peak area at 506 Da through the combined 506, 508, 510, 512, and 514 Da peak areas. Divided by 83.33 for the molar ratio of the equilibrium between labeled and unlabeled PPi, the % exchange was calculated.



### **C<sub>Starter</sub> domain and hydroxylation assay**

Phosphopantetheinylated FrsA was expressed in *E. coli* BAP1 to directly generate the *holo*-form of the T domain,<sup>[59]</sup> isolated with coexpressed FrsB as described above and buffered in 50 mM Tris pH=7.5, 25 mM NaCl, 10 mM MgCl<sub>2</sub> using PD-10 columns (GE) following the manufacturers protocol. Assays were performed in a one-pot-reaction: Purified FrsH was reduced with 5-fold concentration of Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> and equimolar concentration of methylviologen by mixing and incubation on ice for 5 min. To activate the A domain, 1 mM ATP and 1 mM of L-leucine were added. As substrate for the C domain, 500 μM propionyl-S-CoA (CoALA Biosciences, Austin, USA) or acetyl-S-CoA (Sigma) were added. The protein solution of FrsA/B (500 μL, 10 μM) was mixed with 500 μL of the activated FrsH, resulting concentration of the proteins were 5 μM FrsA/B and 25 μM FrsH. The assay was incubated at 20 °C for 3 h. The reaction was stopped by adding 10% (w/v) TCA and incubated on ice for 30 min. Precipitated proteins were pelleted and washed twice with assay buffer. Alkaline hydrolysis was performed with 0.1 M KOH at 70 °C for 20 min. The solution was lyophilized and the pellet solved in a minimal volume of MeOH for LC-MS analysis. For the negative control, the proteins were heat-inactivated at 80 °C for 20 min prior to the addition of the substrates.

### **TE domain assay**

Substrate **2** was dissolved in MeOH to a final concentration of 50 mM, and substrate **6** was dissolved in MeOH to a final concentration of 100 mM. The enzymatic assays were conducted in 50 mM Tris pH=7.5 and 25 mM NaCl, 10 mM MgCl<sub>2</sub>, 2 mM **6** and 0.5 mM **2** in a total volume of 1 mL. Reactions were initiated by the addition of FrsA<sub>TE</sub> (5 μM). The assays were incubated at 22 °C for 5 h and then extracted three times with CH<sub>2</sub>Cl<sub>2</sub>. The organic phase was evaporated and the pellet resuspended in 50 μL MeOH. Alternatively, for the *in vitro* assay, FrsA with coexpressed FrsB and FrsH were used with L-leucine and propionyl-S-CoA to form the side chain under the same conditions as in the C<sub>starter</sub> domain assay. Additionally, 0.5 mM **2** was added and the reaction was incubated and processed like the TE domain assay described above. In this assay mixture the substrates were eventually varied to D-leucine and L-isoleucine for the amino acid and to acetyl-S-CoA (Sigma) and butyryl-S-CoA (CoALA Biosciences, Austin, USA) for the C domain substrate. For the negative control, FrsH was heat-inactivated at 80 °C for 20 min prior to the addition of the substrates.

### Label-free dynamic mass redistribution (DMR) assay

Cell culture materials were purchased from Invitrogen (Carlsbad, CA, USA). Generation of genetically engineered HEK293 cells using CRISPR/Cas9 technology to knockout the subunits of Gαq and Gα11 (Gαq/Gα11-null cells) is described elsewhere.<sup>[16]</sup> Cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) fetal calf serum (FCS, PAN biotech, Germany), penicillin (100 units/mL) and streptomycin (0.1 mg/mL), and kept in a humidified atmosphere with 5% CO<sub>2</sub> at 37 °C. All cell lines were screened by PCR on a monthly basis for mycoplasma contamination and were tested negative. DMR measurements were performed as previously described<sup>[39,40]</sup> using the Epic DMR reader (Corning, NY, USA) together with the Cybi-SELMA semi-automated electronic pipetting system (Analytik Jena AG, Jena, Germany). In brief, 24 h after transfection with the G protein αq subunit, HEK293 cells were counted and seeded at a density of 17,000 cells per well on 384-well fibronectin-coated biosensor plates. On the next day, cells were washed twice with Hanks' buffered salt solution containing 20 mM HEPES (HBSS+ HEPES) and incubated for 1 h at 37 °C in the EPIC reader. **1**, **2**, **3** or **7** were added 1 h before the measurement in HBSS+HEPES. The sensor plate was scanned to record a baseline optical signature (no change in basal DMR), and after agonist addition DMR alterations were monitored for 3,000 s at 37 °C. All data were analyzed using GraphPad Prism 8.0.0 (GraphPad Inc, La Jolla, CA, USA). Quantification of DMR signals was performed by calculation of the maximum responses. Data points from inhibition curves of individual functional experiments were fitted to a four-parameter logistic function:

$$Y = bottom + \frac{(top - bottom)}{1 + 10^{(logEC_{50} - x)slope}}$$

Concentration-effect curves represented in Fig. 4a-c were normalized by setting each experimental maximal effect as 100% response. All data are expressed as mean + or ± standard error of the mean (s.e.m.) of at least three independent experiments performed in technical triplicates.

### Competition binding assay

Competition binding assays of **1** and **2** were conducted versus the radiolabeled FR derivative [<sup>3</sup>H]PSB-15900 utilizing human platelet membrane preparations that natively express Gαq proteins as previously

described.<sup>[19]</sup> The experiments were performed in 50 mM Tris, pH=7.4, in a final assay volume of 200  $\mu$ l. Each assay tube contained 50  $\mu$ l [<sup>3</sup>H]PSB-15900 solution in the buffer (final concentration: 5 nM), 50  $\mu$ l of human platelet membrane preparation in buffer (50  $\mu$ g of protein), 95  $\mu$ l of buffer, and 5  $\mu$ l of **1** or **2** solution, respectively, in varying concentrations dissolved in DMSO. Non-specific-binding was determined in the presence of 5  $\mu$ M FR dissolved in DMSO. Incubation was performed at 37 °C for 1 h with gentle shaking and was started by the addition of the membrane preparation. The incubation was terminated by rapid filtration through GF/C glass fiber filters using a Brandel 24-well-harvester. Filters were dried, punched out, transferred to scintillation vials, and incubated for at least 6 h in scintillation cocktail (2.5 ml ProSafe FC+®, Meridian Biotechnologies Ltd, UK) prior to measurement in a liquid scintillation counter at an efficiency of 53%. Data were evaluated using GraphPad Prism 7 (GraphPad Inc., San Diego, CA, USA), and IC<sub>50</sub> values were determined by the “One site – Fit logIC<sub>50</sub>”-equation.

### **Molecular docking**

The heterotrimeric G protein (G $\alpha$ i/q $\beta$  $\gamma$ ) had been co-crystallized with the Gq selective inhibitor YM (3AH8.pdb, resolution 2.9 Å).<sup>[42]</sup> The structure was downloaded from the Protein Data Bank and prepared using the structure preparation tool implemented in Molecular Operating Environment (MOE 2019.01; Chemical Computing Group, Montreal, Quebec, Canada, 2019). After structure preparation, the hydrogen atoms were assigned according to the Protonate-3D module in MOE 2019.01. The Gq-inhibitors **1** and **2** were docked into the binding site of the Gq protein using Autodock 4.2.<sup>[61]</sup> The atomic partial charges were added using AutoDockTools,<sup>[62,63]</sup> and three-dimensional energy scoring grids of 60  $\times$  60  $\times$  60 points with a spacing of 0.375 Å were computed. The energy grids were centered based on the co-crystallized ligand YM. During the docking simulations, the ligands **1** and **2** were fully flexible while the residues of the receptor were treated as rigid. Fifty independent docking calculations using the varCPSO-ls algorithm from PSO@Autodock implemented in AutoDock4.2 were performed and terminated after 500,000 evaluation steps.<sup>[63]</sup> Parameters of the varCPSO-ls algorithm, the cognitive and social coefficients c1 and c2, were set at 6.05 with 60 individual particles as swarm size. All the other parameters of the algorithm were set at their default values. A plausible binding mode of **1** or **2** was selected on the basis of the lowest binding energy and based on visual inspection of the interactions.

### Feeding and effect determination of **1** and **2** on insects

A small cotton pad soaked with 500 µl of **1** or **2**-containing water was served to first instar nymphs of the stinkbug *R. pedestris* so that the insects orally intake FR. **1** and **2** were each dissolved in 100% MeOH, diluted with water to 0.2, 0.04, and 0.02 µg/µl concentrations, and given to the insects. As a control, distilled water containing MeOH (2%) was used. For each experiment, 20 individuals of the stinkbug were used. Insects were reared in a moisture chamber at 25 °C under a long day regimen (16 h light, 8 h dark), supplied with dried soybean seeds. The **1** or **2**-containing water was exchanged every 3 d, and survival rate was monitored every day at the same time for 9 d.

### BiG-FAM, BiG-SLiCE and BiG-SCAPE analyses

A pairwise Euclidean distance measurement based on the BiG-SLiCE algorithm.<sup>[44]</sup> was done to compare the *frs* BGCs of *C. vaccinii* and “*Ca. B. crenata*” (MiBIG BGC0001451) against all 239,899 non-fragmented NRPS BGCs retrieved from the BiG-FAM database ([bigfam.bioinformatics.nl](http://bigfam.bioinformatics.nl)).<sup>[43]</sup> Using a distance cutoff  $d=1400$ , 2,716 BGCs (Supplementary Data 1) were selected for more sensitive BGC network analysis using BiG-SCAPE<sup>[45]</sup> (parameters: --include\_gbk\_str region BGC --include\_singletons --mix --no\_classify --clans-off --cutoff 0.5, 0.7 --mode global, using Pfam library version 33.1). The network was visualized with Cytoscape 3.7.1.

### Alignments, structural model, and phylogenetic trees

All sequence alignments of genes and proteins were performed with the online platform EMBOSS needle.<sup>[64]</sup> The structural models of the TE domains were calculated by the I-TASSER server<sup>[65,66]</sup> and aligned with the Swiss-PDB-viewer.<sup>[67]</sup> For the phylogenetic trees of C and TE domains, protein sequences were retrieved from public databases (for source organisms and accession numbers see Supplementary Tables 8 and 9) and aligned using the MUSCLE algorithm<sup>[68]</sup> in MEGA 10.<sup>[69]</sup> In case of multi-domain proteins, only the C, resp. TE domain was analyzed. With this program maximum likelihood trees were calculated as follows: The evolutionary history was inferred by using the Maximum Likelihood method and Le\_Gascuel\_2008 model.<sup>[70]</sup> The tree with the highest log likelihood (C: -58208.75, TE: -8887.94) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT

model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter C= 1.6801, TE = 1.9188)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], C: 0.61%, TE: 1.14% sites). The trees are drawn to scale, with branch lengths measured in the number of substitutions per site. C domain analysis involved 149; TE analysis 39 amino acid sequences. For C domain analysis, all positions with less than 95% site coverage were eliminated, i.e., fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position (partial deletion option). For TE analysis, all positions containing gaps and missing data were eliminated (complete deletion option). There were a total of 409 (C) and 132 (TE) positions in the final datasets.

## DATA AVAILABILITY

*frs* sequences have been deposited at GenBank, accession number MT876545. MS/MS data is deposited at MassIVE ([massive.ucsd.edu](https://massive.ucsd.edu)), dataset MSV000085975. The molecular network can be found under: [gnps.ucsd.edu/ProteoSAFe/status.jsp?task=8e524f93dc66405b98c9d77cc0f9c596](https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=8e524f93dc66405b98c9d77cc0f9c596). The list of BGCs for the BiG-SCAPE analysis is in Supplementary data 1. All other data are available from the corresponding author upon request.

## REFERENCES

- [1] Chevette, M.G, et al. Evolutionary dynamics of natural product biosynthesis in bacteria. *Nat. Prod. Rep.* **37**, 566-599 (2020).
- [2] Newman, D.J., Cragg, G.M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770-803 (2020).
- [3] Harvey, A.L., Edrada-Ebel, R., Quinn, R.J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* **14**, 111-129 (2015).

- [4] Hertweck, C. The Biosynthetic Logic of Polyketide Diversity. *Angew. Chem. Int. Ed.* **48**, 4688–4716 (2009).
- [5] Süssmuth, R.D. & Mainz, A. Nonribosomal Peptide Synthesis-Principles and Prospects. *Angew. Chem. Int. Ed.* **56**, 3770–3821 (2017).
- [6] Horsman, M.E. Hari, T.P. & Boddy, C.N. Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: logic gate or a victim of fate? *Nat. Prod. Rep.* **33**, 183–202 (2016).
- [7] Schaffer, J.E. Reck, M.R. Prasad, N.K. & Wenciewicz, T.A.  $\beta$ -Lactone formation during product release from a nonribosomal peptide synthetase. *Nat. Chem. Biol.* **13**, 737–744 (2017).
- [8] He, H.Y., Tang, M.C., Zhang, F., Tang, G.L. Cis-Double bond formation by thioesterase and transfer by ketosynthase in FR901464 biosynthesis. *J. Am. Chem. Soc.* **136**, 4488–4491 (2014).
- [9] Gaudelli, N.M., Townsend, C.A. Epimerization and substrate gating by a TE domain in  $\beta$ -lactam antibiotic biosynthesis. *Nat. Chem. Biol.* **10**, 251–258 (2014).
- [10] McClure, R.A., et al. Elucidating the Rimosaamide-Detoxin Natural Product Families and Their Biosynthesis Using Metabolite/Gene Cluster Correlations. *ACS Chem. Biol.* **11**, 3452–3460 (2016).
- [11] Niehs, S.P., et al. Mining Symbionts of a Spider-Transmitted Fungus Illuminates Uncharted Biosynthetic Pathways to Cytotoxic Benzolactones. *Angew. Chem. Int. Ed.* **59**, 7766–7771 (2020).
- [12] Ray, L., Yamanaka, K., Moore, B.S. A Peptidyl-Transesterifying Type I Thioesterase in Salinamide Biosynthesis. *Angew. Chem. Int. Ed.* **55**, 364–367 (2016).
- [13] Fujioka, M., Koda, S. Morimoto, Y., Biemann, K. Structure of FR900359, a cyclic depsipeptide from *Ardisia crenata* Sims. *J. Org. Chem.* **53**, 2820–2825 (1988).
- [14] Reher, R. et al. Applying Molecular Networking for the Detection of Natural Sources and Analogues of the Selective Gq Protein Inhibitor FR900359. *J. Nat. Prod.* **81**, 1628–1635 (2018).
- [15] Taniguchi M., et al. YM-254890, a novel platelet aggregation inhibitor produced by *Chromobacterium* sp. QS3666. *J. Antibiot. (Tokyo)* **56**, 358–363 (2003).

- [16] Schrage, R. et al. The experimental power of FR900359 to study Gq-regulated biological processes. *Nat. Commun.* **6**, 10156 (2015).
- [17] Zhang, H., Nielsen, A.L., Strømgaard, K. Recent achievements in developing selective Gq inhibitors. *Med. Res. Rev.*, **40**, 135-157. (2019).
- [18] Xiong, X.F. et al. Total synthesis and structure-activity relationship studies of a series of selective G protein inhibitors. *Nat. Chem.* **8**, 1035–1041 (2016).
- [19] Kuschak, M., et al. Cell-permeable high-affinity tracers for Gq proteins provide structural insights, reveal distinct binding kinetics, and identify small molecule inhibitors. *Br. J. Pharmacol.* **177**, 1898-1916 (2020).
- [20] Grundmann, M. et al., Lack of beta-arrestin signaling in the absence of active G proteins. *Nat. Commun.* **9**, 341 (2018).
- [21] Matthey, M. et al. Targeted inhibition of Gq signaling induces airway relaxation in mouse models of asthma. *Sci. Transl. Med.* **9**, eaag2288 (2017).
- [22] Annala, S. et al. Direct targeting of Gαq and Gα11 oncoproteins in cancer cells. *Sci. Signal.* **12**, eaau5948 (2019).
- [23] Kostenis, E., Pfeil, E.M., Annala, S. Heterotrimeric Gq proteins as therapeutic targets? *J. Biol. Chem.* **295**, 5206-5215 (2020).
- [24] Zhang H. et al. Structure-Activity Relationship Studies of the Cyclic Depsipeptide Natural Product YM-254890, Targeting the Gq Protein. *ChemMedChem.* **12**, 830-834 (2017).
- [25] Zhang, H., et al. Structure-activity relationship and conformational studies of the natural product cyclic depsipeptides YM-254890 and FR900359. *Eur. J. Med. Chem.* **156**, 847-860 (2018).
- [26] Xiong, X.F., et al. Structure-Activity Relationship Studies of the Natural Product Gq/11 Protein Inhibitor YM-254890. *ChemMedChem.* **14**, 865-870 (2019).

- [27] Reher, R. et al. Deciphering Specificity Determinants for FR900359-Derived Gq  $\alpha$  Inhibitors Based on Computational and Structure-Activity Studies. *ChemMedChem* **13**, 1634–1643 (2018).
- [28] Crüsemann, M. et al. Heterologous Expression, Biosynthetic Studies, and Ecological Function of the Selective Gq-Signaling Inhibitor FR900359. *Angew. Chem. Int. Ed.* **57**, 836–840 (2018).
- [29] Carlier, A. et al. The genome analysis of *Candidatus Burkholderia crenata* reveals that secondary metabolism may be a key function of the *Ardisia crenata* leaf nodule symbiosis. *Environ. Microbiol.* **18**, 2507–2522 (2016).
- [30] Soby, S.D., Gadagkar, S.R., Contreras, C., Caruso, F.L. *Chromobacterium vaccinii* sp. nov., isolated from native and cultivated cranberry (*Vaccinium macrocarpon* Ait.) bogs and irrigation ponds. *Int. J. Syst. Evol. Microbiol.* **63**, 1840–1846 (2013).
- [31] Wang, M. et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **34**, 828–837 (2016).
- [32] Vöing, K. Harrison, A. Soby, S.D. Draft Genome Sequence of *Chromobacterium vaccinii*, a Potential Biocontrol Agent against Mosquito (*Aedes aegypti*) Larvae. *Genome Announc.* **3**, e00477-15 (2015).
- [33] Hoang, T. T., Karkhoff-Schweizer, R. R., Kutchma, A. J., Schweizer, H. P. A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. *Gene* **212**, 77–86 (1998).
- [34] Sánchez, C., Braña, A.F., Méndez, C., Salas, J.A. Reevaluation of the violacein biosynthetic pathway and its relationship to indolocarbazole biosynthesis. *Chembiochem* **7**, 1231–1240 (2006).
- [35] Miyamae, A., Fujioka, M., Kosa, S., Morimoti, Y. Structural Studies of FR900359, a Novel Cyclic Depsipeptide from *Ardisia crenata* Sims (Myrsinaceae) *J. Chem. Soc. Perkin Trans.* **1**, 673-678 (1989).



- [36] Phelan, V. V., Du, Y., McLean, J. A., Bachmann, B. O. Adenylation enzyme characterization using  $\gamma$ - $^{18}\text{O}_4$ -ATP pyrophosphate exchange. *Chem. Biol.* **16**, 473–478 (2009).
- [37] Makris, T.M., Chakrabarti, M., Münck, E., Lipscomb, J.D. A family of diiron monooxygenases catalyzing amino acid beta-hydroxylation in antibiotic biosynthesis. *Proc. Natl. Acad. Sci U S A* **107**, 15391–15396 (2010).
- [38] Kaniusaite, M., Goode, R.J.A., Schittenhelm, R.B., Makris, T.M., Cryle, M.J. The Diiron Monooxygenase CmlA from Chloramphenicol Biosynthesis Allows Reconstitution of  $\beta$ -Hydroxylation during Glycopeptide Antibiotic Biosynthesis. *ACS Chem Biol.* **14**, 2932-2941 (2019).
- [39] Schröder, R. et al. Deconvolution of complex G protein-coupled receptor signaling in live cells using dynamic mass redistribution measurements. *Nat. Biotechnol.* **28**, 943–949 (2010).
- [40] Schröder, R. et al. Applying label-free dynamic mass redistribution technology to frame signaling of G protein-coupled receptors noninvasively in living cells. *Nat. Protoc.* **6**, 1748–1760 (2011).
- [41] Malfacini, D. et al. Rational design of a heterotrimeric G protein  $\alpha$  subunit with artificial inhibitor sensitivity. *J. Biol. Chem.* **294**, 5747–5758 (2019).
- [42] Nishimura, A. et al. Structural basis for the specific inhibition of heterotrimeric Gq protein by a small molecule. *Proc. Natl. Acad. Sci. U S A* **107**, 13666-13671 (2010).
- [43] Kautsar, S.A., Blin, K., Shaw, S., Weber, T., Medema, M.H. BiG-FAM: the biosynthetic gene cluster families database. *Nucleic Acids Res.* gkaa812, doi: 10.1093/nar/gkaa812 (2020).
- [44] Kautsar, S.A., van der Hooft, J.J.J., de Ridder, D., Medema, M.H. BiG-SLiCE: A highly Scalable Tool Maps the Diversity of 1.2 Million Biosynthetic Gene Clusters. *bioRxiv* (2020).
- [45] Navarro-Muñoz, J. C., et al. A computational framework to explore large-scale biosynthetic diversity. *Nat. Chem. Biol.* **16**, 60-68 (2020).

- [46] Niehs, S.P., Dose, B., Scherlach, K., Roth, M., Hertweck, C. Genomics-Driven Discovery of a Symbiont-Specific Cyclopeptide from Bacteria Residing in the Rice Seedling Blight Fungus. *Chembiochem* **19**, 2167-2172 (2018).
- [47] Fewer, D.P., Metsä-Ketelä, M. A pharmaceutical model for the molecular evolution of microbial natural products. *FEBS J.* **287**, 1429-1449 (2020).
- [48] Noda-Garcia, L., Tawfik, D.S. Enzyme evolution in natural products biosynthesis: target- or diversity-oriented? *Curr. Opin. Chem. Biol.* **59**, 147-154 (2020).
- [49] Adamek, M., Alanjary, M., Ziemert, N. Applied evolution: phylogeny-based approaches in natural products research. *Nat. Prod. Rep.* **36**, 1295–1312 (2019).
- [50] Klapper, M., Braga, D., Lackner, G., Herbst, R., Stallforth, P. Bacterial Alkaloid Biosynthesis: Structural Diversity via a Minimalistic Nonribosomal Peptide Synthetase. *Cell Chem. Biol.* **25**, 659–665 (2018).
- [51] May, J.J., Wendrich, T.M., Marahiel, M.A. The *dhb* operon of *Bacillus subtilis* encodes the biosynthetic template for the catecholic siderophore 2,3-dihydroxybenzoate-glycine-threonine trimeric ester bacillibactin. *J. Biol. Chem.* **276**, 7209–7217 (2001).
- [52] Krishnan, A., Mustafa, A., Almén, M.S., Fredriksson, R., Williams, M.J., Schiöth, H.B. Evolutionary hierarchy of vertebrate-like heterotrimeric G protein families. *Mol. Phylogenet. Evol.* **91**, 27-40 (2015).
- [53] Machida, K. et al. Sameuramide A, a new cyclic depsipeptide isolated from an ascidian of the family Didemnidae. *Bioorg. Med. Chem.* **26**, 3852-3857 (2018).
- [54] Taniguchi, M. et al. YM-254890 analogues, novel cyclic depsipeptides with Gαq/11 inhibitory activity from *Chromobacterium* sp. QS3666. *Bioorg. Med. Chem.* **12**, 3125-3133 (2004).

- [55] Medema, M.H., Cimermancic, P., Sali, A., Takano, E., Fischbach, M.A. A systematic computational analysis of biosynthetic gene cluster evolution: lessons for engineering biosynthesis. *PLoS Comput Biol.* **10**, e1004016 (2014).
- [56] Fischbach, M.A., Walsh, C.T., Clardy, J. The evolution of gene collectives: How natural selection drives chemical innovation. *Proc. Natl. Acad. Sci. U S A* **105**, 4601-4608 (2008).
- [57] Waglechner, N., McArthur, A.G., Wright, G.D. Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. *Nat. Microbiol.* **11**, 1862-1871 (2019).
- [58] Hillenmeyer, M.E., Vandova, G.A., Berlew, E.E., Charkoudian, L.K. Evolution of chemical diversity by coordinated gene swaps in type II polyketide gene clusters. *Proc. Natl. Acad. Sci. U S A* **112**, 13952-13957 (2015)
- [59] Pfeifer, B. A., Admiraal, S.J., Gramajo, H., Cane, D.E., Khosla, C. Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of *E. coli*. *Science* **291**, 1790–1792 (2001).
- [60] Schägger, H. Tricine-SDS-PAGE. *Nat. Protoc.* **1**, 16–22 (2006).
- [61] Morris, G.M. et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785-2791 (2009).
- [62] Sanner, M. F. Python: a programming language for software integration and development. *J. Mol. Graph. Model.* **17**, 57-61 (1999).
- [63] Namasivayam, V. & Gunther, R. pso@autodock: a fast flexible molecular docking program based on Swarm intelligence. *Chem. Biol. Drug Des.* **70**, 475-484 (2007).
- [64] Madeira, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 2019, **47**, W636–W641 (2019).
- [65] Yang, J., Zhang, Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res.* **43**, W174–W181 (2015).

- [66] Zhang, C., Freddolino, P.L., Zhang, Y. COFACTOR: improved protein function prediction by combining structure, sequence and protein–protein interaction information. *Nucleic Acids Res.* **45**, W291–W299 (2017).
- [67] Guex, N., Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).
- [68] Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
- [69] Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Bio. Evol.* **35**, 1547–1549 (2018).
- [70] Le, S.Q., Gascuel, O. An Improved General Amino Acid Replacement Matrix. *Mol. Biol. Evol.* **25**, 1307–1320 (2008).

## ACKNOWLEDGEMENTS

We thank the DFG (FOR 2372 grants CR464/7-1 to M.C., KO 902/17-1 and 17-2 to G.M.K., KO 1582/10-1 and 10-2 to E.K., MU 1667/7-1 und 7-2 to C.E.M., GU 345/3-1 to M.G.) and the DBU (PhD scholarship 20018/568 to W.H.) for funding. A.I. was funded by the LEAP JP19gm0010004 from the Japan Agency for Medical Research and Development (AMED). Furthermore, we thank Prof. Dr. Matthias Ullrich, Jacobs University Bremen for the supply of vectors pEX18Tc, pFlp2 and pPS858, Goran Grujicic for help with extractions and Göttingen Genomics Laboratory (G2L) for *C. vaccinii* genome sequencing. Satria Kautsar is acknowledged for advice on bioinformatics and for help with BiG-SLiCE analyses.

## AUTHOR CONTRIBUTIONS

C.H. performed all *in vitro* experiments with FrsA and FrsD, and isolated compounds **2** and **7**. R.R. assembled the *frs* gene cluster and performed knock-outs in *C. vaccinii*. D.A.W. expressed and

characterized FrsH. C.H. and D.A.W. performed bioinformatics analyses. W.H. performed culturing experiments of *C. vaccinii* and mutants. J.P. and J.A. performed Gq inhibition experiments under supervision of E.K., S.K. elucidated the structures of compounds **2** and **7**. J.-H. V. performed competition experiments, V.N. performed docking studies, both under supervision of C.E.M. J.K. synthesized compounds **5** and **6** together with C.H. under supervision of M.G. T.O. and P.M. performed stinkbug feeding experiments. A.I. contributed valuable analytical tools ('Gαq/Gα11-null' cells). G.M.K. and M.C. designed and supervised the research. The manuscript was written by M.C. with contributions by all authors.

## **COMPETING INTERESTS**

The authors declare no competing interests.

## **MATERIALS & CORRESPONDENCE**

Correspondence and material requests should be addressed to Max Crüsemann ([cruesemann@uni-bonn.de](mailto:cruesemann@uni-bonn.de)).