SCIENTIFIC REPORTS

OPEN

QSAR-based molecular signatures of prenylated (iso)flavonoids underlying antimicrobial potency against and membrane-disruption in Gram positive and Gram negative bacteria

Carla Araya-Cloutier[1], Jean-Paul Vincken[1], Milou G. M. van de Schans[1,2], Jos Hageman[3], Gijs Schaftenaar[4], Heidy M. W. den Besten[5] & Harry Gruppen[1]

Prenylated flavonoids and isoflavonoids are phytochemicals with remarkable antibacterial activity. In this study, 30 prenylated (iso)flavonoids were tested against *Listeria monocytogenes* and *Escherichia coli* (the latter in combination with an efflux pump inhibitor). Minimum inhibitory concentrations of the most active compounds ranged between 6.3–15.0 µg/mL. Quantitative structure-activity relationships (QSAR) analysis was performed and linear regression models were proposed with R^2 between 0.77–0.80, average R^2_m between 0.70–0.75, Q^2_{LOO} between 0.66–0.69, and relatively low amount of descriptors. Shape descriptors (related to flexibility and globularity), together with hydrophilic/hydrophobic volume and surface area descriptors, were identified as important molecular characteristics related to activity. A 3D pharmacophore model explaining the effect of the prenyl position on the activity of compounds was developed for each bacterium. These models predicted active compounds with an accuracy of 71–88%. With regard to the mode of action, good antibacterial prenylated (iso)flavonoids with low relative hydrophobic surface area caused remarkable membrane permeabilization, whereas those with higher relative hydrophobic surface area did not. Based on the QSAR and membrane permeabilization studies, the mode of action of antibacterial prenylated (iso)flavonoids was putatively rationalized.

Interest in natural compounds for antimicrobial discovery has increased during the last decade due to the rise in highly resistant bacteria[1]. Natural products are a potentially rich source of compounds for antimicrobial discovery and development, as they usually have a high degree of stereochemistry with an extensive variety of (ring) scaffolds and generally low toxicity[2,3].

In the Fabaceae plant family, prenylated phenolic compounds have been identified as promising antimicrobials in extracts from elicited seedlings[4,5]. These phenolic compounds mainly belong to the flavonoid (2-phenyl benzopyrans), isoflavonoid (3-phenyl benzopyrans) and stilbenoid (1,2-diphenyl ethylene) classes. Besides the skeleton, they can differ in configuration and position of prenylation. Prenyl groups can be attached to a phenolic compound as a chain (3,3-dimethylallyl substituent) or as a five- or six-membered ring after undergoing enzymatic cyclisation with an *ortho*-phenolic hydroxyl group[6]. Prenylation has been shown to increase the antibacterial activity of phenolic compounds[7], partly because hydrophobicity increases. Thereby, the partitioning to biological targets, such as membranes, increases[8].

[1]Laboratory of Food Chemistry, Wageningen University & Research, Wageningen, The Netherlands. [2]RIKILT – Wageningen University & Research, Wageningen, The Netherlands. [3]Biometris, Applied Statistics, Wageningen University & Research, Wageningen, The Netherlands. [4]Nijmegen Centre for Molecular Sciences, Radboud University Medical Centre, Nijmegen, The Netherlands. [5]Laboratory of Food Microbiology, Wageningen University & Research, Wageningen, The Netherlands. Correspondence and requests for materials should be addressed to J.-P.V. (email: jean-paul.vincken@wur.nl)
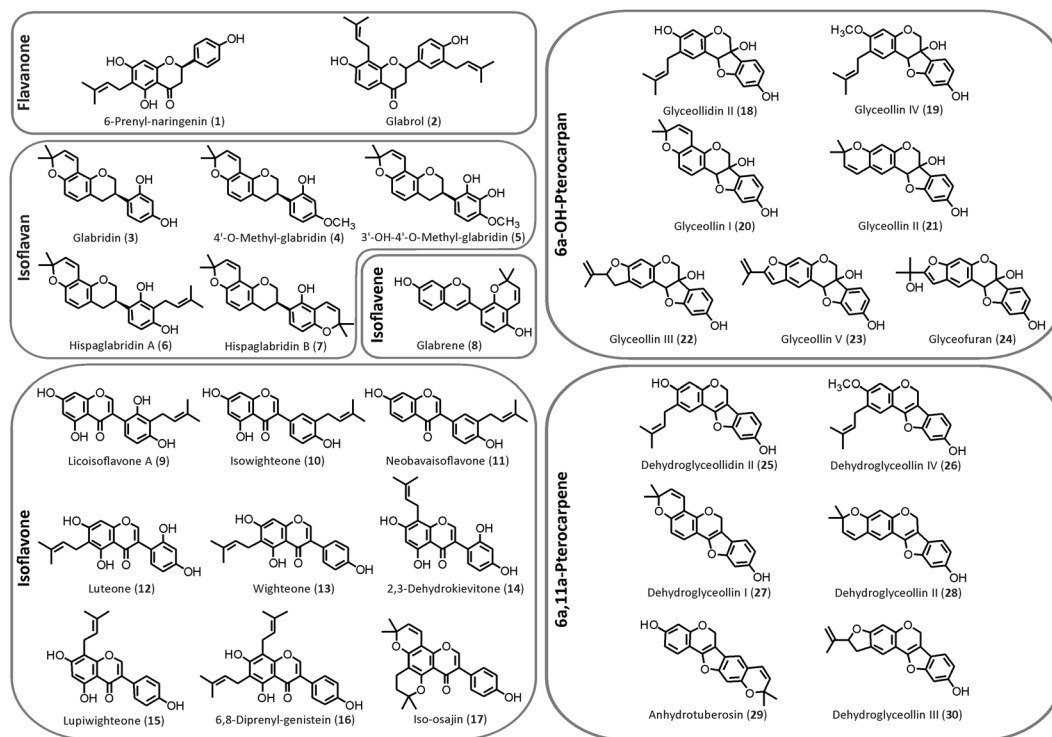
Figure 1. Overview of the structures of the prenylated (iso)flavonoids, and subclass to which they belong, tested in this study.

Structure-antibacterial activity relationships (SAR) of prenylated (iso)flavonoids have been obtained by comparing the activity data and structures of a few analogues (usually ≤ 10 compounds)[9,10]. Studies have focused on describing the substituents and positions within a particular phenolic skeleton that provide high or low antibacterial activity[11,12]. The antibacterial activity of prenylated and non-prenylated (iso)flavonoids from licorice against methicillin-resistant *Staphylococcus aureus* (MRSA) and Gram negative bacteria has been studied[13]. None of the compounds were active against the Gram negative bacteria (minimum inhibitory concentration (MIC) >128 μg/mL), but some were very active against MRSA, especially diprenylated (iso)flavonoids (MIC 8.0 μg/mL). It was also observed that the presence of a D-ring in a coumestan was detrimental for the activity, the effect of which was attributed to the rigidity of the skeleton. In contrast, a diprenylated 6a,11a-pterocarpene and a monoprenylated pterocarpan, both having rigid skeletons, were the most antibacterial compounds against MRSA (MIC 6.3 μg/mL for both)[14]. In another study[15], a diprenylated pterocarpan and diprenylated isoflavan were found to be very active (MIC 3.1–6.3 μg/mL) against vancomycin-resistant enterococci, whereas a diprenylated isoflavanone was not. Further studies showed that the position of the prenyl group(s) differently affected the antibacterial activity of different subclasses of (iso)flavonoids[16,17]. For flavanones, prenylation at *C*8 provided a higher antibacterial activity (i.e. MIC 12.5–25 μg/mL) than prenylation at *C*6 (MIC >100 μg/mL)[18], whereas for isoflavones the opposite result was found[10]. Regarding other substituents, the presence of hydroxyl groups appeared to be essential for antibacterial activity of prenylated (iso)flavonoids[7]. In contrast, the influence of methoxyl groups was more ambiguous, with reduced activity in some isoflavonoids[14,19], and enhanced activity in others[12,17,20].

Increase in hydrophobicity by the presence of prenyl groups has been the traditional explanation given for the increase in antibacterial activity observed upon prenylation of (iso)flavonoids[8]. However, not all prenylated (iso)flavonoids have shown antibacterial activity[18,21], highlighting the importance of other, yet to be defined, molecular characteristics for antibacterial activity. Furthermore, it has not been investigated in any large detail how differences in the overall molecular configuration relate to the mode of action. In this context, *in silico* approaches, like quantitative SAR (QSAR)[22] and pharmacophore modelling[23], can help to elucidate the physicochemical properties important for the antibacterial activity and mode of action. Despite numerous descriptive SAR studies of prenylated (iso)flavonoids, frequently with relatively few compounds, no QSAR analysis of prenylated (iso)flavonoids as antibacterials has been performed so far.

In this study we report on the application of QSAR analysis to elucidate the main molecular characteristics of prenylated (iso)flavonoids as antibacterials. For this, we used the Gram positive *Listeria monocytogenes* and the Gram negative *Escherichia coli* as model organisms. We previously showed that intrinsic resistance in *E. coli* can be overcome by inhibiting the efflux pump systems[5]. Consequently, we here report the antibacterial activity of prenylated (iso)flavonoids against *E. coli* in combination with an efflux pump inhibitor. We tested 30 mono- and diprenylated (iso)flavonoids (Fig. 1), belonging to 6 different (iso)flavonoid subclasses, prenylated at different positions of the main skeleton and with chain, pyran or furan prenyl configurations. The majority of these molecules has never been tested before against *L. monocytogenes*, or against *E. coli* in the presence of an efflux pump

No.	Name compound	*L. monocytogenes*		*E. coli* + EPI	
		MIC	MBC	MIC	MBC
Flavanone					
1	6-Prenyl-naringenin	>50	>50	25	25
2	Glabrol	6.3	6.3	>50	>50
Isoflavan					
3	Glabridin	15	20	10	15
4	4′-*O*-methyl-glabridin	10	12.5	20	20
5	3′-OH-4′-*O*-methyl-glabridin	12.5	25	>50	>50
6	Hispaglabridin A	15	25	>50	>50
7	Hispaglabridin B	10	15	>50	>50
Isoflavene					
8	Glabrene	35	50	50	50
Isoflavone					
9	Licoisoflavone A	35	50	20	25
10	Isowighteone	20	20	25	25
11	Neobavaisoflavone	50	50	25	25
12	Luteone	20	50	10	15
13	Wighteone	10	15	15	15
14	2,3-Dehydrokievitone	>50	>50	>50	>50
15	Lupiwighteone	>50	>50	>50	>50
16	6,8-Diprenyl-genistein	6.3	6.3	>50	>50
17	Iso-osajin	>50	>50	>50	>50
6a-OH-Pterocarpan					
18	Glyceollidin II	>50	>50	>50	>50
19	Glyceollin IV	50	>50	50	>50
20	Glyceollin I	>50	>50	>50	>50
21	Glyceollin II	>50	>50	>50	>50
22	Glyceollin III	>50	>50	>50	>50
23	Glyceollin V	>50	>50	>50	>50
24	Glyceofuran	>50	>50	>50	>50
6a,11a-Pterocarpene					
25	Dehydroglyceollidin II	12.5	15	50	>50
26	Dehydroglyceollin IV	25	35	>50	>50
27	Dehydroglyceollin I	20	20	>50	>50
28	Dehydroglyceollin II	20	25	>50	>50
29	Anhydrotuberosin	>50	>50	n.t.[c]	n.t.
30	Dehydroglyceollin III	>50	>50	>50	>50

Table 1. Antibacterial activity of prenylated (iso)flavonoids. Minimum inhibitory concentration (MIC, μg/mL) and minimum bactericidal concentration (MBC, μg/mL) of prenylated (iso)flavonoids against *L. monocytogenes* EGD-e and against *E. coli* K12, the latter one in the presence of efflux pump inhibitor PaβN (48 μM). No antibacterial activity was observed against *E. coli* without the efflux pump inhibitor (MIC >50 μg/mL). n.t. Not tested.

inhibitor. We determined the minimum inhibitory and bactericidal concentrations of all compounds, and analyzed the membrane permeabilization capacity of the antibacterial compounds.

The main goals of this study were (i) to develop a multiple linear regression model with high predictive capabilities for the activity of other prenylated (iso)flavonoids by using a QSAR approach, and (ii) to understand the mode of action of prenylated (iso)flavonoids against Gram positive and Gram negative bacteria by linking the QSAR findings to the proposed mode of action of prenylated (iso)flavonoids, i.e. disruption of membrane integrity[5,24–26].

Results

Table 1 shows the MICs and MBCs determined for each compound against each bacterium. Diprenylated glabrol (**2**) and 6,8-diprenylgenistein (**16**) were very good antibacterials against *L. monocytogenes* with both MIC and MBC values of 6.3 μg/mL. *E.coli* was not susceptible to any of the tested (iso)flavonoids up to 50 μg/mL. When the efflux pumps were inhibited by PAβN, *E. coli* became susceptible to a similar extent as *L. monocytogenes* (control samples with *E. coli* demonstrated no growth inhibition by PaβN alone, data not shown). In this way, monoprenylated glabridin and luteone showed very good antibacterial activity against *E. coli* with MICs of 10 μg/mL and MBCs of 15.0 μg/mL.
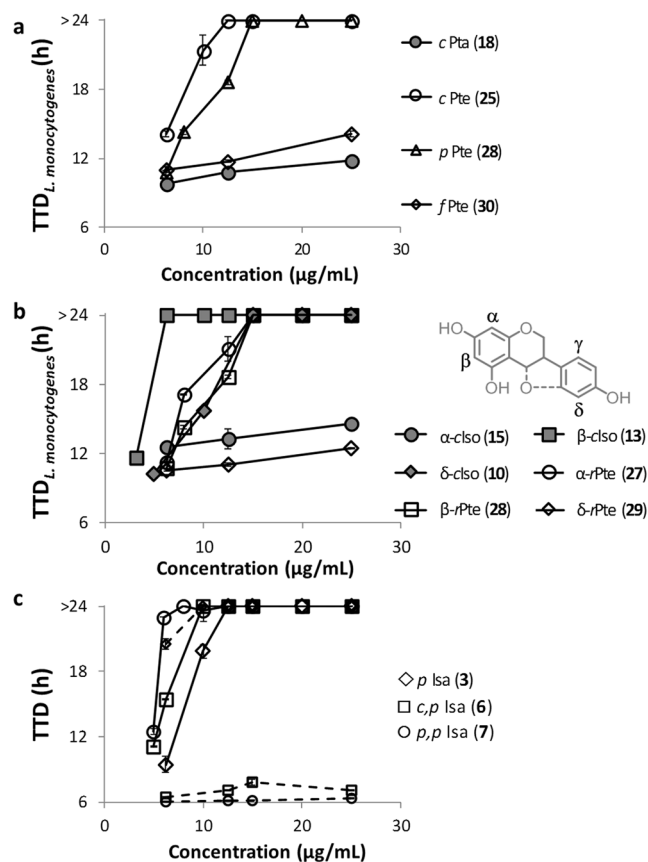
Figure 2. Time to detection (TTD) of bacterial growth upon exposure to prenylated (iso)flavonoids. (**a**) TTD of *L. monocytogenes* growth in the presence of different prenylated 6a-hydroxy-pterocarpans (glyceollidin II, **18**) and 6a,11a-pterocarpenes (dehydroglyceollidin II, **25**; dehydroglyceollin II, **28**; dehydroglyceollin III, **30**). (**b**) TTD of *L. monocytogenes* growth in the presence of chain prenylated isoflavones (closed symbols): isowighteone (**10**), wighteone (**13**), lupiwighteone (**15**); and ring-closed prenylated pterocarpenes (open symbols): dehydroglyceollin I (**27**), dehydroglyceollin II (**28**), anhydrotuberosin (**29**). Prenyl positions are annotated based on the inset. (**c**) TTD of *L. monocytogenes* (solid line) and *E. coli* (with EPI, dashed line) growth in the presence of different prenylated isoflavans: glabridin (**3**), hispaglabridin A (**6**) and hispaglabridin B (**7**). Error bars represent the standard deviation of two biologically independent reproductions. TTD of control cells (i.e. cells not exposed to prenylated compounds) was 11.1 ± 2.9 h for *L. monocytogenes* and 6.9 ± 0.13 for *E. coli*.

Effect of (iso)flavonoid subclass.

In general, isoflavans and isoflavones were among the most active compounds against both *L. monocytogenes and E. coli* (Table 1). Interestingly, most 6a-hydroxy-pterocarpans did not show antibacterial activity up to the highest concentration tested (MIC >50 μg/mL) against the two target bacteria, whereas 6a,11a-pterocarpenes were active against *L. monocytogenes*. Figure 2a illustrates the inhibition of growth of *L. monocytogenes*, when exposed to a chain prenylated 6a-hydroxy-pterocarpan (*c*Pta, **18**) and its equivalent chain prenylated 6a,11a-pterocarpene (*c*Pte, **25**) (the higher the TTD, the stronger the inhibition with a maximum of 24 h). As observed, the dehydration and formation of a double bond at position 6a in dehydroglyceollidin II (**25**) enhanced the antibacterial activity in comparison with glyceollidin II (**18**).

Effect of the configuration of prenyl group.

Figure 2a also compares the inhibitory activity of three 6a,11a-pterocarpenes sharing the same structure but different configuration of prenyl group. The chain prenylated 6a,11a-pterocarpene (*c*Pte, **25**) inhibited the growth of the bacteria at a minimum concentration of 12.5 μg/mL (Fig. 2a). The pyran prenylated 6a,11a-pterocarpene (*p*Pte, **28**) was the next most active compound with a MIC of 20.0 μg/mL. Last, the furan prenylated 6a,11a-pterocarpene (*f*Pte, **30**) did not inhibit the growth at any concentration tested. Also, none of the furan prenylated compounds (**22–24**, **30**) showed antibacterial activity against the target bacteria (Table 1).

Effect of the position of the prenyl group.

Concerning the (iso)flavonoid skeletons, there are four main positions of *C*-prenylation: *C*6, *C*8, *C*3' and *C*5' (or the equivalents *C*2, *C*4, *C*8 and *C*10 in 6a-hydroxy-pterocarpans and 6a,11a-pterocarpenes). In order to compare different subclasses of (iso)flavonoids, the Greek numbering system of Simons *et al*[6]. was adopted, which is illustrated by the isoflavonoid skeleton in Fig. 2b. The position of the prenyl group clearly affected the antibacterial activity of these compounds. Lupiwighteone, chain prenylated at the α position (α-*c*Iso, **15**), showed no inhibition towards *L. monocytogenes*, whereas β-prenylated

Bacteria	k	*p*-value	R^2	R^2_{adj}	SE	Q^2_{LOO}	$Q^2_{LOO\_adj}$	VIF_{max}
	3	5.2e-06	0.6431	0.6020	0.2807	0.5537	0.5022	1.957
	4	3.2e-05	0.6328	0.574	0.2904	0.6685	0.6155	4.844
L.monocytogenes	**5**	**4.6e-07**	**0.7735**	**0.7263**	**0.2328**	**0.6566**	**0.5850**	**1.592**
	6	4.9e-07	0.8007	0.7487	0.2230	0.6821	0.5991	3.413
	7	3.8e-06	0.7886	0.7213	0.2349	0.7044	0.6103	4.732
	8	1.5e-05	0.7866	0.7053	0.2415	0.6788	0.5564	4.567
	3	2.2e-05	0.6140	0.5677	0.2233	0.4955	0.4373	1.745
	4	7.8e-06	0.6881	0.6361	0.2049	0.5736	0.5054	2.251
E. coli	**5**	**1.8e-07**	**0.8047**	**0.7623**	**0.1656**	**0.6921**	**0.6279**	**4.143**
	6	2.0e-06	0.7879	0.7301	0.1765	0.6620	0.5738	2.317
	7	1.2e-07	0.8597	0.813	0.1469	0.7569	0.6795	4.358
	8	7.0e-07	0.8564	0.7989	0.1523	0.7356	0.6349	3.999

Table 2. Statistical performance of the best OLS models obtained using GA-selection of variables for predicting antibacterial activity of prenylated (iso)flavonoids against *L. monocytogenes* and *E. coli*. k: number of descriptors; R^2: coefficient of determination; R^2_{adj}: adjusted R^2; SE: standard error of residuals; Q^2_{LOO}: leave-one-out cross-validated coefficient of determination; $Q2_{LOO\_adj}$: adjusted Q^2_{LOO}; VIF_{max}: maximum variance inflation factor. The finally chosen OLS models are shown in bold.

wighteone (β-*c*Iso, **13**) had high antibacterial activity, followed by δ-prenylated isowighteone (δ-*c*Iso, **10**). For 6a,11a-pterocarpenes, activity was similar for α- and β-prenylated compounds, i.e. dehydroglyceollin I (α-*r*Pte, **27**) and II (β-*r*Pte, **28**), whereas the δ-prenylated 6a,11a-pterocarpene, i.e. anhydrotuberosin (δ-*r*Pte, **29**), showed the lowest growth inhibition.

Effect of the number of prenyl groups. In Fig. 2c the growth inhibition of *L. monocytogenes* (solid line) and *E. coli* in the presence of the EPI (dashed line) by mono- and diprenylated isoflavans (Isa) is illustrated. Except for iso-osajin (**17**), diprenylated (iso)flavonoids were among the most active ones against the Gram positive bacterium (MIC ≤ 15 μg/mL). Diprenylated hispaglabridin A (*c,p*Isa, **6**) and B (*p,p*Isa, **7**) inhibited bacterial growth faster than monoprenylated glabridin (*p*Isa, **3**). Opposite results were found for the Gram negative bacterium. Diprenylated (iso)flavonoids were not effective antibacterials against *E.coli* (Table 1), whereas monoprenylated compounds, including glabridin (**3**), were very active against this bacterium (MIC ≤ 15 μg/mL).

Effect of other substituents. The number of hydroxyl groups influenced the antibacterial activity of prenylated isoflavones. The presence of 3 OH groups in the isoflavone skeleton as in isowighteone (**10**) was more effective against the Gram positive bacteria (MIC 20 μg/mL) than the presence of 4 OH groups as in licoisoflavone A (**9**, MIC 35 μg/mL) or than the presence of 2 OH groups as in neobavaisoflavone (**11**, MIC 50 μg/mL). For *E. coli*, the prenylated isoflavone with 4 OH groups (**9**, MIC 20 μg/mL) was slightly better antibacterial than prenylated isoflavones with 3 (**10**) or 2 OH groups (**11**, MIC 25 μg/mL). With regard to *O*-methylation, we found no clear effect on the antibacterial activity of prenylated (iso)flavonoids against *L. monocytogenes*. In the case of *E. coli*, *O*-methylation was detrimental for activity, as in dehydroglyceollidin II (**25**, MIC 50 μg/mL) in comparison with *O*-methylated dehydroglyceollin IV (**26**, MIC >50 μg/mL), or as in glabridin (**3**, MIC 15 μg/mL) in comparison with 4'-*O*-methyl-glabridin (**4**, MIC 20 μg/mL).

QSAR analysis. QSAR of prenylated (iso)flavonoids was carried out to interpret the structural features responsible for their antibacterial activity against Gram positive *L. monocytogenes* and Gram negative *E. coli* (in combination with an efflux pump inhibitor). A genetic algorithm (GA) was used to explore which molecular descriptors were best able to predict the antibacterial activity using ordinary least square (OLS) regression. GA is an optimization technique that explores the descriptor space simultaneously by a population of candidate solutions (models), in which solutions compete and recombine[27]. Since the starting point of a GA run is always different, every GA run was repeated 5 times for a fixed subset of descriptors (varied from 3 to 8 descriptors). The best OLS models obtained per number of descriptor are summarized in Table 2 (criteria for selecting the best models are specified in Materials and Methods section). The descriptors used in the best GA-generated OLS models are specified in Table 3.

L. monocytogenes models. The top most frequently used descriptors (i.e. those with a frequency factor larger than 1 in Table 3) in the models of *L.monocytogenes* were related to the fraction of heavy atoms (*rsynth*, *p*-value 0.008), molecular branching (*KierA3*, *p*-value 2×10^{-5}), the relative hydrophilic surface (*vsurf_CW6*, *p*-value 0.002), distribution of hydrophilic/hydrophobic regions (*vsurf_A*, *p*-value 0.003), and the location of hydrophobic interacting atoms or groups (*vsurf_DD12*, *p*-value 0.005).

Additionally, considering the sign (positive/negative) of the coefficient estimates of the descriptors in Table 3, the molecular properties that are important for antibacterial activity of prenylated (iso)flavonoids might be predicted. Specifically, shape related descriptors, i.e. molecular branching (*KierA3*, top descriptor), number of rotational bonds (*b_rotN*) and globularity (*glob*) were positively correlated with antibacterial activity, whereas density (*dens*), atom information content (*a_ICM*) and diameter (*diameter*) were negatively correlated with activity. This gave an indication that small flexible globular molecules were better for antibacterial activity against

L. monocytogenes				E. coli			
Descriptor	F	p-value	Sign	Descriptor	F	p-value	Sign
a_ICM	1	$1.8E^{-06}$	–	a_acc	2	$4.8E^{-06}$–0.0149	–
ast_violation	1	0.0019	+	a_hyd	2	$8.0E^{-06}$–$8.5E^{-05}$	–
b_rotN	1	0.0185	+	b_count	1	$8.4E^{-07}$	–
dens	1	0.0326	–	b_single	1	0.0006	–
diameter	1	0.0088	–	logP$_{o/w}$	1	0.0055	–
glob	1	0.1308	+	npr1	2	0.0003–0.0018	–
h_ema	1	$8.1E^{-05}$	–	PEOE_VSA+0	1	0.0139	–
KierA3	2	1.2–$2.2E^{-05}$	+	petitjeanSC	1	0.0028	–
rsynth	5	$2.1E^{-05}$–0.0085	+	rgyr	2	$1.1E^{-05}$–0.0001	+
vdw_area	1	0.0019	+	RPC+	1	0.0012	+
vsurf_A	2	0.0015–0.0027	–	SlogP_VSA0	1	0.2686	–
vsurf_CP	1	0.1804	+	std_dim3	6	$8.6E^{-08}$–0.0002	+
vsurf_CW2	1	0.6293	–	vsurf_A	1	0.0011	–
vsurf_CW6	3	0.0045–0.0191	+	vsurf_CW1	1	0.0808	+
vsurf_DD12	3	0.0001–0.0049	–	vsurf_EDmin1	1	0.0076	+
vsurf_DW12	1	0.1013	–	vsurf_IW1	1	0.0002	+
vsurf_ID8	1	0.0111	+	vsurf_IW2	3	$5.1E^{-06}$–$7.3E^{-05}$	+
vsurf_IW4	1	0.0078	–	vsurf_IW4	1	0.0004	–
vsurf_IW5	1	0.0484	+	vsurf_IW5	1	0.0023	–
vsurf_IW6	1	0.1961	+	vsurf_IW6	1	0.0005	–
vsurf_IW7	1	0.1350	–	vsurf_W7	1	0.0190	–
vsurf_W4	1	0.0114	+	zagreb	1	0.0002	–
vsurf_W6	1	0.0004	+				

Table 3. Descriptors present in the best OLS regression models obtained for predicting the antibacterial activity against *L. monocytogenes* and *E. coli*. F: selection frequency in best GA-generated models; *p*-value: significance of descriptor coefficient in the models; Sign: sign of descriptor coefficient. Descriptor meaning can be found in Supplementary Table S3.

L. monocytogenes than large flat molecules. The hydrophilic surface (*vsurf_CW6*, top descriptor) and volume (*vsurf_W*) were positively correlated with activity, but having separate hydrophilic and hydrophobic regions within the molecule seems to be detrimental for activity, as the amphiphilic moment (*vsurf_A*, top descriptor) was negatively correlated with activity. The (positive/negative) contribution of other (top) descriptors (e.g. *rsynth*, *vsurf_DD*, *vsurf_IW*) in relation to antibacterial activity remains difficult to interpret and needs further study.

E. coli models. The top most frequently used descriptors (i.e. those with a frequency factor larger than 1 in Table 3) in the models of *E. coli* were related to molecular shape (*std_dim3*, *p*-value 0.0002; *npr1*, *p*-value 0.002), the distribution of hydrophilic regions (*vsurf_IW2*, *p*-value 7×10^{-5}), the number of hydrogen bond acceptors (*a_acc*, *p*-value 0.01), the number of hydrophobic atoms (*a_hyd*, *p*-value 8×10^{-5}), and molecular flexibility (*rgyr*, *p*-value 0.0001). Specifically, globularity, illustrated by the descriptor *std_dim3* (top descriptor), was positively correlated with antibacterial activity against *E. coli*. Together with the negative correlation with activity of the normalized principal moment of inertia 1 (*npr1*, top descriptor), this indicated overall that a globular shape, rather than a flat shape, is better for activity. Flexibility was positively correlated with antibacterial activity, illustrated by the positive correlation of the radius of gyration (*rgyr*, top descriptor). Hydrophobicity, illustrated by the number of hydrophobic atoms (*a_hyd*, top descriptor) and the octanol water partition coefficient (*logP$_{o/w}$*), was negatively correlated with activity against *E. coli*. Other (top) descriptors showed positive and negative coefficient estimates in the different models, making interpretation of their contribution to antibacterial activity difficult.

Based on the statistical performance of the GA-generated OLS models (Table 2) and the number of descriptors (ratio between number of compounds and number of descriptors ≥ 5 for a good model)[28], the models for best predicting the antibacterial activity of prenylated (iso)flavonoids are highlighted in bold in Table 2 and detailed in Table 4. As observed, selected OLS models showed R^2 between 0.77–0.80, average R^2_m between 0.70–0.75, good predictive power ($Q^2 > 0.65$), and no outliers.

Pharmacophore elucidation with antibacterial prenylated (iso)flavonoids. Some prenylated compounds showed different antibacterial activity, but had almost the same structure and chemical properties, except for the position of the prenyl group. To extend the understanding of the effect of the position of prenylation on antibacterial activity we employed ligand-based pharmacophore elucidation to model the structural requirements of prenylated (iso)flavonoids for antibacterial activity against *L. monocytogenes* (Fig. 3a) and *E. coli* (Fig. 3b). The 3D pharmacophore models were similar for both bacteria, consisting of one hydrophobic feature (green, at the location of the prenyl group), two aromatic ring features (orange corresponding to the A- and B-ring of (iso)flavonoids), and one (for *L. monocytogenes*) or two (for *E. coli*) hydrogen bond acceptor projection features

Microorganism	N	R^2	R^2_{adj}	$R^2_{m\_av}$	Q^2_{LOO}	$Q^2_{LOO\_adj}$	Outliers	Descriptor	Coefficient	Standard error	p-value	VIF
								KierA3	0.40	0.07	1.2×10^{-5}	1.26
								rsynth	1.13	0.21	2.1×10^{-5}	1.32
L. monocytogenes	30	0.77	0.73	0.70	0.66	0.59	no	vsurf_DD12	−0.07	0.02	0.0001	1.14
								vsurf_IW4	−0.16	0.05	0.008	1.24
								vsurf_ID8	0.29	0.11	0.01	1.59
								Intercept	2.71	0.36	8.6×10^{-8}	—
								b_count	−0.06	0.01	8.4×10^{-7}	2.39
								std_dim3	1.54	0.23	1.3×10^{-6}	1.76
E. coli	29	0.80	0.76	0.75	0.69	0.62	no	vsurf_IW2	0.74	0.14	3.1×10^{-5}	4.14
								rgyr	0.80	0.17	0.0001	2.19
								vsurf_IW4	−0.28	0.07	0.0004	3.72
								Intercept	1.60	0.63	0.02	—

Table 4. Selected GA-MLR models developed for predicting the antibacterial activity of prenylated (iso) flavonoids. *N*: number of compounds used for building the model; R^2: Correlation coefficient; R^2_{adj}: adjusted R^2; $R^2_{m\_av}$: average modified R^2 (based on scaled activity values)[a]; Q^2_{LOO}: leave-one-out cross-validated correlation coefficient; $Q^2_{LOO\_adj}$: adjusted Q^2; *Outliers*: presence of outliers[b]; *VIF*: variance inflation factor. Descriptor meaning can be found in Supplementary Table S3. [a]Regression plots for calculating the $R^2_{m\_av}$ are in Supplementary Figure S1. [b] The presence of outliers was evaluated with the standardization approach[59] (Supplementary Figure S2).
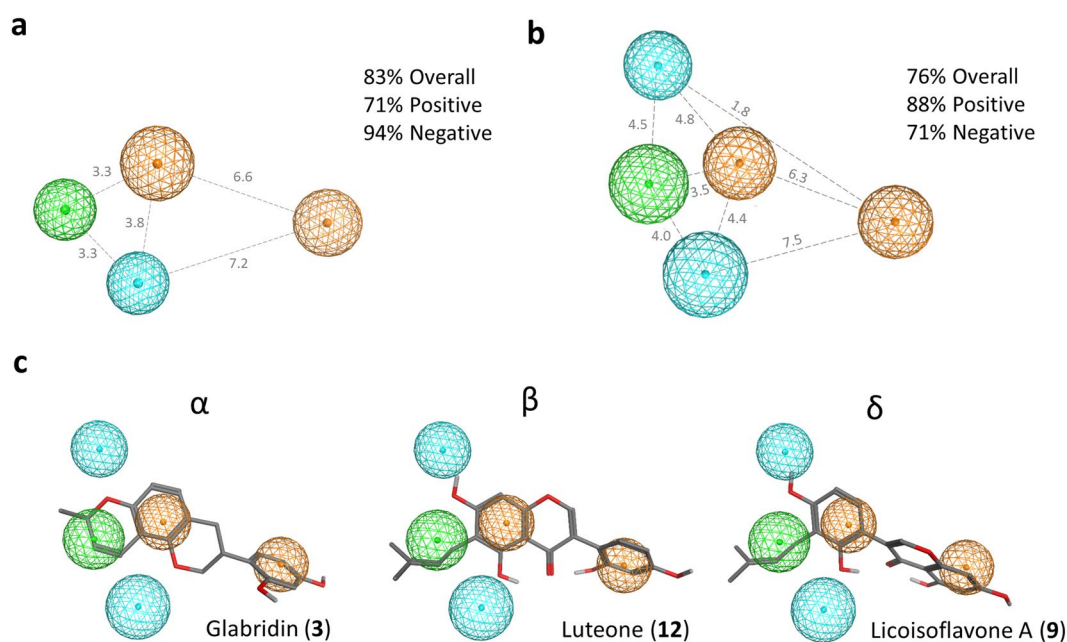

Figure 3. Ligand-based pharmacophore model for antibacterial (MIC ≤ 25 µg/mL) prenylated (iso)flavonoids against *L. monocytogenes* (**a**) and *E. coli* (**b**). The color of the spheres represents the following features: blue spheres are hydrogen bond acceptor projection features (i.e. feature that annotates the location of a possible hydrogen bond partner), whereas green and orange spheres represent hydrophobic areas or aromatic rings, respectively, in the ligand. Distance between the features are in Ångström. Overall accuracy, i.e. proportion of correctly predicted compounds (positives and negatives); positive accuracy, i.e. proportion of correctly predicted positive compounds (those with good antibacterial activity); negative accuracy, i.e. proportion of correctly predicted negative compounds (those with moderate to low antibacterial activity). (**c**) Good antibacterial (iso)flavonoids, prenylated at different positions and with different configurations, fitting *E. coli*'s pharmacophore model.

(blue). A projection feature defines the position of a potential hydrogen bond partner. Using an activity threshold of a MIC value of 25 µg/mL, these two pharmacophore models were able to predict which compounds and, thus prenylation positions, were good for antibacterial activity. These models showed a 76–83% overall accuracy, i.e. proportion of true positives and negatives results (Supplementary Tables S1 and S2). Figure 3c shows examples of good antibacterial compounds (glabridin (**3**), luteone (**12**) and licoisoflavone A (**9**)) prenylated at different positions in relation to their mapping of the *E. coli* pharmacophore model. The aromatic (orange) and hydrogen
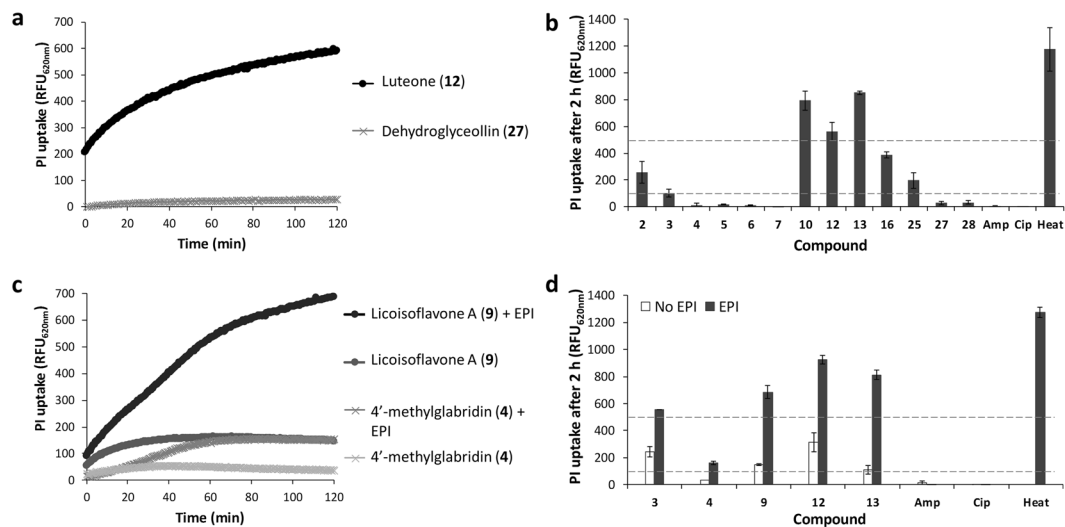
Figure 4. Membrane permeabilization by antibacterial prenylated (iso)flavonoids. PI net uptake by *L. monocytogenes* (**a**) and *E. coli* (**c**) upon exposure to antibacterial prenylated (iso)flavonoids. Signal from the control cells (untreated, with and without EPI has been subtracted). Overall net PI uptake by *L. monocytogenes* (**b**) and *E. coli* (**d**) after 2 h of exposure to antibacterial prenylated (iso)flavonoids and antibiotics (at 20 μg/mL). Dashed lines represent the ranges used to defined non-permeabilizers (PI uptake <100 RFU), poor (PI uptake 100–500 RFU) and good permeabilizers (PI uptake >500 RFU). Data are means of two independent biological reproductions, with standard deviations as error bars. Number of compound refers to those in Table 1. Ampicillin (Amp), ciprofloxacin (Cip), heat-treated cells (Heat).

bond acceptor (blue) features can be mapped by most prenylated (iso)flavonoids. However, inactive compounds are unable to map the hydrophobic feature (green) with their prenyl groups.

Some compounds were falsely predicted as positive during the pharmacophore search (Supplementary Table S2). Diprenylated compounds were able to map *E. coli*'s pharmacophore features and consequently predicted as active, although no MIC value was found against *E. coli*. The same result was observed for the α-prenylated 6a-hydroxy-pterocarpan (incorrectly predicted as active). To reach (i.e. permeate) the inner membrane, hydrophobicity is known to be determinant. These false positive results might be explained by an inappropriate hydrophobicity of these molecules: hydrophobicity might be too high for the diprenylated molecules to cross the OM of *E. coli*, or too low to permeate to the inner membrane for the 6a-OH-pterocarpan.

Membrane permeabilization by antibacterial prenylated (iso)flavonoids.

Permeabilization of the cytoplasmic membrane of *L. monocytogenes* and *E. coli*, the latter in the presence and absence of EPI, was assayed by measuring the uptake of the fluorescent probe propidium iodine (PI). Figure 4a,c show the net PI uptake by *L. monocytogenes* and *E. coli* upon exposure to different antibacterial prenylated (iso)flavonoids (i.e. MIC <25 μg/mL) and traditional antimicrobials. As observed, membrane permeabilization was immediate when treating the cells with some antibacterial prenylated compounds (e.g. luteone and licoisoflavones A). Figure 4b,d show the net PI uptake by *L. monocytogenes* and *E. coli* after 2 h of exposure to different prenylated (iso)flavonoids. The monoprenylated isoflavones isowighteone (**10**), luteone (**12**) and wighteone (**13**) were good membrane permeabilizers (i.e. PI uptake >500 RFU) of *L. monocytogenes*. Unexpectedly, other active compounds, including some diprenylated compounds, showed little change in fluorescence (PI uptake <100 RFU). In the case of *E. coli*, glabridin (**3**), licoisoflavone A (**9**), luteone (**12**) and wighteone (**13**) were good membrane permeabilizers of *E. coli* in the presence of the EPI (i.e. PI uptake >500 RFU). Interestingly, these compounds were able to permeabilize without the EPI, although to a lesser extent than when EPI was present (i.e. PI uptake <500 RFU). This indicated that monoprenylated isoflavones were able to permeabilize *E. coli* cells (as some PI uptake occurred), but no lethal permeabilization took place as no MIC or MBC were found under those conditions. It cannot be excluded that recovery of PI permeable cells has occurred, as this has been observed previously for other microorganisms[29].

Because hydrophobicity and shape were important molecular characteristics correlated with antibacterial activity according to the QSAR analysis, these properties were also expected to affect the membrane permeabilization properties of antibacterial prenylated (iso)flavonoids. Interestingly, we found a significant (p 1.6×10^{-7}) negative correlation (Fig. 5a) between permeabilization and relative hydrophobic surface area of the compounds. The good antibacterials that showed poor or no membrane permeabilization had a relatively larger hydrophobic surface area (>85%) than the good antibacterials that showed good membrane permeabilization. Figure 5b shows examples of good antibacterial compounds against *L. monocytogenes* with different permeabilization capacity, colored according to hydrophobicity and polarity.
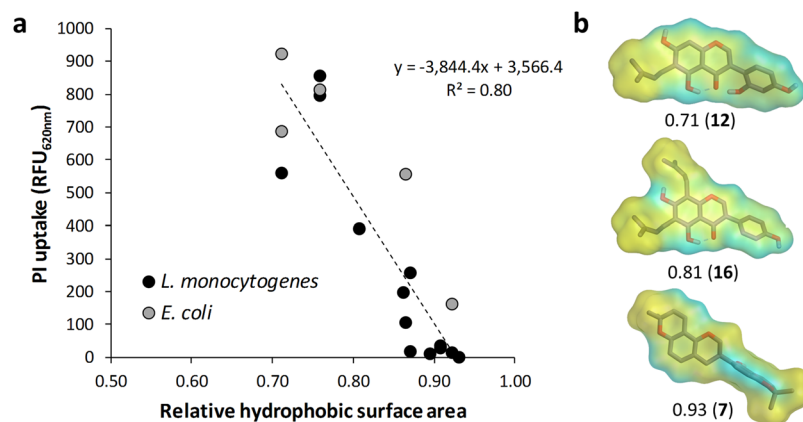
Figure 5. Correlation between the membrane permeabilization capacity (indicated by the PI uptake) of antibacterial prenylated (iso)flavonoids (i.e. MIC <25 μg/mL) and the relative hydrophobic surface area of the molecules (**a**). Molecular surface of active prenylated (iso)flavonoids with different membrane permeabilization properties (**b**). Surface color is yellow for hydrophobic areas and light blue for polar areas. Number below each molecule refers to the relative hydrophobic surface area, as calculated by MOE (*PEOE_VSA_FHYD*).

Discussion

Not all prenylated (iso)flavonoids tested in this study showed potent antibacterial activity, suggesting that increase of hydrophobicity by prenylation alone does not warrant antibacterial activity and that structural details might have a profound effect on antibacterial potency. From the QSAR analysis, shape (flexibility and globularity) and hydrophilicity/hydrophobicity related descriptors were found to be important for the antibacterial activity of prenylated (iso)flavonoids against both *L. monocytogenes* and *E. coli* (in the presence of an EPI). Additionally, a specific distribution of functional groups was illustrated by a 3D pharmacophore model, which can be used for predicting the position of prenylation required for antibacterial activity.

Importance of hydrophobicity on antibacterial activity of prenylated (iso)flavonoids.

The effect of hydrophobicity of compounds on antibacterial activity was bacterium-dependent. Diprenylated (iso)flavonoids were effective against Gram positive *L. monocytogenes* but not against *E. coli*. Analyzing the *logD* of the effective compounds against these bacteria (Figure S3), the minimum hydrophobicity for prenylated (iso)flavonoids to be active against *L. monocytogenes* and *E. coli* corresponded to a *logD* of 3.6, whereas for *E. coli* there was also a maximum corresponding to a *logD* of 4.5. These differences in hydrophobicity requirements between *L. monocytogenes* and *E. coli* might be explained by their different influx pathways. In *E. coli* the entrance to the periplasmic space of small (MW <600–700) and hydrophilic antimicrobials is usually through the nonspecific porins. These are narrow (6–15 Å diameter) hydrophilic channels with exposed negatively charged residues[30–32]. Because of the porin's charged lining, hydrophobicity of antimicrobials has been shown to strongly reduce their influx rates through porins[33]. Alternatively, diffusion across the lipid bilayer of the outer membrane (OM), composed of lipopolysaccharides packed in a highly ordered fashion, is known to be very slow for hydrophobic compounds[34]. There are other substrate-specific (active) transporters in the OM, however, they require substrate recognition[32,35]. Flexibility and "3-dimensionality" (including globularity) were also important characteristics in a previous study affecting the accumulation of small traditional antimicrobials inside *E. coli*[36]. Having a low number of rotatable bonds (<6) and a globularity lower than 0.29 (where 0 represents a plane and 1 a sphere) improved the accumulation inside Gram negative bacteria. All prenylated (iso)flavonoids tested in the present study complied with those requirements (active prenylated compounds against *E. coli* had ≤ 3 rotatable bonds and a globularity of ≤ 0.1). Therefore, shape does not seem to be the limiting factor for entrance of prenylated compounds into *E. coli* cell envelope.

There are no studies regarding the pathway of influx of prenylated (iso)flavonoids into Gram negative bacterial cells. Nevertheless, based on (i) the high substrate specificity of the specialized OM transporters[35,37], (ii) the known porin-mediated entrance of antimicrobials of similar size, shape (flexibility and globularity) and hydrophobicity as our (active) prenylated (iso)flavonoids[34,36,38], (iii) the observed detrimental effects of hydrophobicity, and (iv) the beneficial effects of extra polar groups in prenylated (iso)flavonoids regarding *E. coli* antibacterial activity (Table 1), we postulate that prenylated (iso)flavonoids mainly enter via the unspecific porins. Consequently, the rate of influx through porins and thus the efficiency in inhibiting *E. coli* decreases as the hydrophobicity of the prenylated compound increases (e.g. as in diprenylated compounds). In contrast, because of the absence of the OM in Gram positives, (di)prenylated (iso)flavonoids have no hydrophobicity restrictions to enter the cell envelope[39].

Importance of molecular shape on antibacterial activity of prenylated (iso)flavonoids.

Besides antibacterial uptake and accumulation, effective interactions with the target site are essential for antibacterial action. Because of the relatively high hydrophobicity of prenylated (iso)flavonoids (usually *logD* >3), it is expected that these compounds have high affinity towards the cytoplasmic membrane. Flexibility has been shown to be a mechanical determinant for membrane-interacting antimicrobial peptides. Peptides require some degree

of conformational freedom to partition into membranes as well as to arrange effective intermolecular interactions by rearranging their flexible side chains[40–42]. In line with this, we propose that chain prenylated (iso)flavonoids are able to rearrange more effectively than ring-closed prenyl groups, as prenyl chains are more flexible than prenyl rings. These rearrangements will improve their intermolecular interactions with a potential target. The positive correlation of globularity with antibacterial activity of prenylated (iso)flavonoids is in accordance with our previous results[5], where flatness of prenylated stilbenoids was associated with their relatively low antibacterial activity as opposed to the relatively good antibacterial activity of elbow-shaped prenylated (iso)flavonoids.

Membrane permeabilization by prenylated (iso)flavonoids.

We measured the level of membrane permeabilization by antibacterial prenylated (iso)flavonoids as an indicator of membrane integrity. Interestingly, we observed a significant ($p < 0.01$) negative correlation between membrane permeabilization and the relative hydrophobic surface area of antibacterial prenylated compounds. Hydrophobicity has been proposed to affect the localization of (prenylated) flavonoids within model membranes (liposomes)[26,43]. The difference in distribution along the phospholipids, and consequently, along the membrane might be responsible for the effectiveness in permeabilization observed in this study. Compounds interacting closer to the surface of the membrane, near the polar head groups (i.e. those with relatively lower hydrophobic surface area), might permeabilize the membrane more efficiently than compounds interacting deeper in the membrane (those with relatively higher hydrophobic surface area).

In general, these findings indicate that permeabilization is not the sole mechanism of action of these antibacterial compounds. Because of their high hydrophobicity, it is still expected that these compounds have high affinity towards the cytoplasmic membrane. Membrane integrity might also be disrupted by other mechanisms, such as alteration of membrane fluidity or elasticity, alteration of hydration of membrane surface, alteration of membrane conductivity, or induction of lipid peroxidation[43,44].

Mode of action of prenylated (iso)flavonoids revisited.

The proposed effects of hydrophobicity and shape of (iso)flavonoids on their partitioning, accumulation in the bacterial inner membrane, and disruption of membrane integrity (which includes permeabilization but also other undefined mechanisms), are summarized in Fig. 6. In the Gram positive bacterium, prenylated compounds are not expected to encounter considerable barrier effects by the peptidoglycan layer to reach the cytoplasmic membrane where they can integrate, provided that the minimum hydrophobicity is met (as is the case for all compounds shown in Fig. 6). In the Gram negative bacterium, highly hydrophobic compounds (e.g. 6,8-diprenyl-genistein, dark green compound) encounter influx restrictions due to porins, but compounds meeting the hydrophobicity restrictions of porins can cross the outer membrane of Gram negatives. In the presence of a broad efflux pump inhibitor, such as PaβN, these prenylated (iso)flavonoids can pass the outer membrane and remain inside the periplasm. Flexible globular compounds (e.g. chain prenylated isoflavones, such as wighteone or 6,8-diprenyl-genistein, represented in Fig. 6 by deformed circles inside the membrane) are able to adapt their conformation to make effective intermolecular interactions with their target (proposed here to be in the cytoplasmic membrane). Less flexible and flat compounds (e.g. pyran prenylated pterocarpenes, such as dehydroglyceollin III) will have more limitations to interact effectively with the target site (as for the red compound). Consequently, they show less antibacterial activity, despite their hydrophobicity. Fast permeabilization by prenylated compounds might occur when they interact closer to the polar head groups of the phospholipids (indicated by orange shadow in the head groups). When the compound permeates deeper into the membrane due to their higher relative hydrophobic surface area, other disruption effects (e.g. alteration of membrane fluidity) might occur.

This is the first report on the QSAR of prenylated (iso)flavonoids as antibacterials against Gram positive and Gram negative bacteria. Important molecular characteristics identified in this study can be used for the design and development of novel antibacterials from these natural compounds.

Materials and Methods

Materials.

Prenylated isoflavonoids (glabrene, 3′-hydroxy-4′-O-methyl-glabridin, 4′-O-methyl-glabridin, hispaglabridin A, hispaglabridin B, glyceofuran, glyceollidin II, glyceollin I, glyceollin II, glyceollin III, glyceollin IV, glyceollin V, dehydroglyceollidin II, dehydroglyceollin I, dehydroglyceollin II, dehydroglyceollin, III, dehydroglyceollin IV) and one prenylated flavone (glabrol) were previously purified and chemically characterized[45,46]. 6-Prenylnaringenin, propidium iodine (PI), Phe-Arg β-naphthylamide dihydrochloride (PaβN), ampicillin and ciprofloxacin were purchased from Sigma Aldrich (St. Louis, MO, USA). Isowighteone and anhydrotuberosin were purchased from ChemFaces (Wuhan, Hubei, China). Wighteone, lupiwighteone, isowighteone, luteone, 2,3-dehydrokievitone, licoisoflavone A, neobavaisoflavone, iso-osajin and 6,8-diprenygenistein were purchased from Plantech UK (Reading, UK). Bacto brain heart infusion (BHI) broth was purchased from BD (Franklin Lakes, NJ, USA), tryptone soya broth (TSB) and bacteriological agar from Oxoid Ltd (Basingstoke, UK), and peptone physiological salt solution (PPS) from Tritium Microbiologie (Eindhoven, The Netherlands). Ethanol absolute (EtOH) was purchased from Biosolve (Valkenswaard, The Netherlands).

Antibacterial activity.

Compounds were tested for their antimicrobial activity against Gram positive *Listeria monocytogenes* EGD-e and Gram negative *Escherichia coli* K12. For the latter, the experiments were conducted in the presence and absence of the efflux pump inhibitor PaβN.

Bacteria were streaked from a −80 °C glycerol stock onto a BHI agar plate and incubated 24 h at 37 °C. Next, one colony was transferred to BHI broth (10 mL) and further incubated for 18 h at 37 °C. These overnight cultures were diluted with TSB (final inoculum concentration $3.4 ± 0.4$ $\log_{10}$ CFU/mL for *L. monocytogenes* and $4.2 ± 0.1$ $\log_{10}$ CFU/mL for *E. coli*). Stock solutions of the different prenylated compounds were prepared in EtOH (96% v/v) and subsequently diluted with TSB. Equal volumes (100 μL) of bacteria and prenylated compound solutions
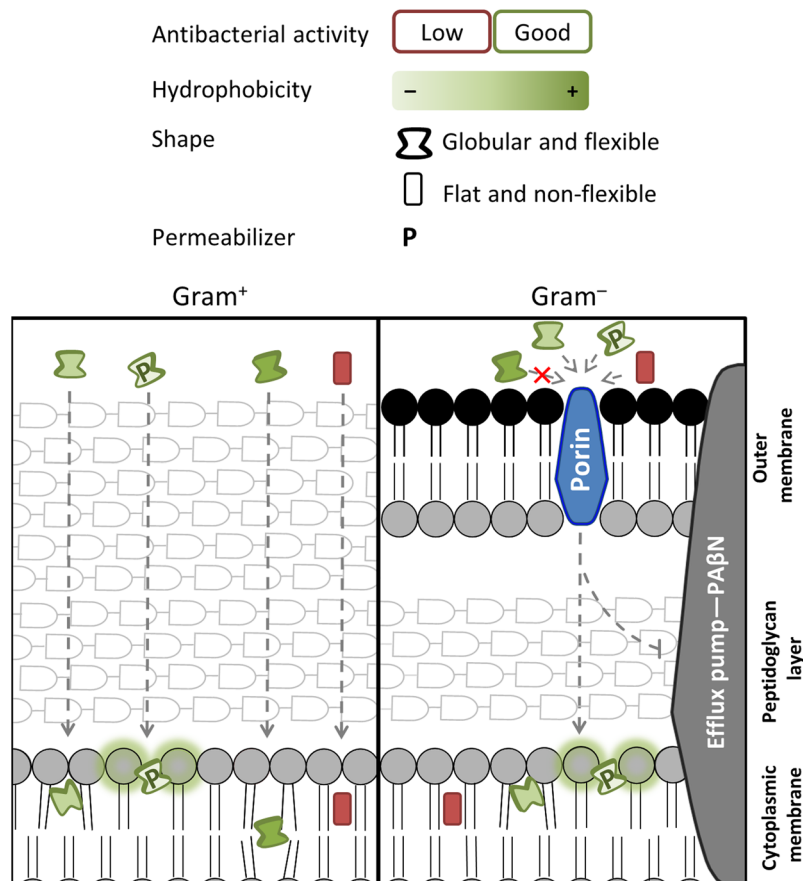
10

Figure 6. Schematic overview of the main molecular characteristics found in this study to influence the antibacterial action of prenylated (iso)flavonoids against Gram positive and Gram negative bacteria. Compounds with green outline are good antibacterials (i.e. MIC <25 µg/mL). The intensity of the green shading in the compounds indicates the extent of hydrophobicity (with a minimum hydrophobicity of $logD$ 3.6). Compound with red outline and red shading indicates a hydrophobic molecule ($logD \geq 3.6$) with moderate or low antibacterial activity. Globular and flexible compounds were good antibacterials (indicated by the deformed shape). The compound with moderate/low antibacterial activity is a non-flexible planar molecule. Green shading around the phospholipids represents permeabilization.

(final concentrations tested of prenylated (iso)flavonoids 5–50 µg/mL, 2.5% v/v EtOH max.) in TSB were mixed into a 96-well plate. For *E. coli*, the efflux pump inhibitor (EPI) PAβN was added in the medium to a concentration of 25 µg/mL (48 µM)[47]. The 96-well plate was incubated in a SpectraMax M2e (Molecular Devices, Sunnyvale, CA, USA), at 37 °C with constant linear shaking. The optical density (OD) at 600 nm was measured every 5 min for 24 h.

Positive controls (ampicillin at 10 µg/mL), negative controls (TSB suspension of bacteria with 2.5% (v/v) EtOH) and blanks (extracts and TSB medium with no bacteria) were considered for optical comparison and sterility control. Inhibition of growth was assessed by measuring the time to detection (TTD), i.e. the time to reach a change in OD of 0.05 units[48]. When no change in OD (i.e. ΔOD <0.05) was observed after the 24 h of incubation, cell viability was verified by plate counting. Briefly, 100 µL of the well with no change in OD was decimally diluted in PPS solution and 100 µL of each dilution was spread onto BHI agar plates. Plates were incubated for 24 h at 37 °C and colonies were counted. The minimum inhibitory concentration (MIC) was defined as the lowest concentration of compound that resulted in a bacterial count equal or lower than that of the initial inoculum. The minimum bactericidal concentration (MBC) was defined as the lowest concentration of compound that resulted in >99% bacterial inactivation from the initial bacterial inoculum. Prenylated compounds were tested in two independent biological reproductions, each performed in duplicate.

Cell membrane permeability. To investigate the effects on the cytoplasmic membrane permeability, the fluorescent probe propidium iodine (PI) was used. Bacteria (−80 °C glycerol stock) were streaked onto a BHI agar plate and incubated overnight at 37 °C. Subsequently, one colony was transfer to 50 mL of BHI broth and incubated at 25 °C for 13 h. Cells were harvested by centrifugation (4696 g, 4 °C, 20 min) and washed twice with PPS (pH 7.2). The final cell pellet was suspended in 5 mL PPS to obtain an inoculum of 9.9 ± 0.3 log$_{10}$ CFU/mL. Stock solutions of PI and prenylated compounds were diluted in PPS to concentrations of 60 µM and 40–80 µg/mL, respectively. A volume of 50 µL of each solution (final concentration in test of PI 15 µM and of compounds

10–20 μg/mL, 2% (v/v) EtOH max.) and 100 μL of inoculum were added to a black with clear bottom 96-well plate (Greiner Bio One, Kremsmünster, Austria). Emission of fluorescence was measured every 30 s at 620 nm (bottom read mode), while exciting the sample at 520 nm, using a Spectramax M2e. For the positive control, cells were treated for 10 min at 95 °C in a Thermomixer (Eppendorf, Hamburg, Germany) and added as inoculum, whereas blank control contained 100 μL of PPS instead of inoculum. Tests with the antimicrobials ciprofloxacin and ampicillin (10–20 μg/mL) were included for comparison. Intrinsic fluorescence or quenching effects of the compounds and of PI were considered by the use of blanks (compounds with PI, without cells; PI without cells). Compounds were tested in two independent biological reproductions, each performed in duplicate.

QSAR modeling. Molecular Operating Environment (MOE, version 2016.0802, Chemical Computing Group, Montreal, QC, Canada) was used to calculate 2D and internal 3D (i3D) molecular descriptors of prenylated (iso)flavonoids (full list of descriptors is in Supplementary Table S3), after structure preparation (MOPAC PM3 energy minimization, gradient 0.01 kcal/Å). Highly correlated ($|R| \geq 0.99$) and constant descriptors were eliminated from the list. Antibacterial activity, expressed as pMIC (i.e. –logMIC, in molar)[49,50], was used as the dependent variable. To take into account the complete structural diversity of the tested molecules, inactive molecules with a MIC >50 μg/mL were included by imputing a MIC 100 μg/mL (i.e. the next 2-fold higher concentration to be tested) during the ordinary least squares (OLS) regression analysis. As the amount of compounds was relatively low (30 compounds), separation of the dataset into training and test sets is not recommended due to higher risks of chance correlation and overfitting[51,52]. Based on Tropsha (2010), the minimum number of compounds should be no less than 40 for splitting the dataset[53]; therefore all compounds tested in our study were used for building the QSAR models.

A genetic algorithm (GA)[54] was used to select a fixed subset of predictors best able to predict the antibacterial activities using OLS. The GA parameters were optimized using a full factorial experiment design and were determined to be: population size = 100, cross-over rate = 0.6, mutation rate = 0.5. The maximum number of iterations was set to 200 and elitism set to 10. The number of predictors to be selected during a GA run was varied between 3 and 8. To exclude lucky or unlucky GA runs, every run was repeated 5 times with different starting seeds (Supplementary Figure S4). Models were assessed based the following statistical parameters: (i) significance (p-value); (ii) coefficient of determination (R^2); (iii) adjusted coefficient of determination (R^2_{adj}); (iv) residual standard error (SE); (v) leave-one-out cross-validation (LOOCV) coefficient of determination (Q^2_{LOO}), to assess the model's internal predictivity; (vi) the difference between R^2 and $Q^2 < 0.30$, to avoid overfitting;[55] (vii) adjusted LOOCV coefficient of determination ($Q^2_{LOO\_adj}$), to allow comparison between models with different number of variables;[56,57] (viii) variance inflation factor (VIF), to assess descriptor inter-correlation or multicollinearity. When VIF values were >5 the combination of predictors was penalized during the GA, effectively removing the solution from the population. Last, the finally selected models per bacterium were further internally validated by calculating: (i) the average modified coefficient of determination ($R^2_{m\_av}$), based on the scaled values of the observed and predicted response data[58]; (ii) the applicability domain to identify outliers in the dataset, based on the standardization approach[59].

Ligand-based pharmacophore elucidation. The pharmacophore elucidation query module of MOE was used to build a pharmacophore model. This is done by aligning the multiple ligands in the training set (i.e. energy minimized 3D structures) and determining the essential common chemical features to construct the model[23]. All compounds (i.e. 30 compounds tested for *L. monocytogenes* and 29 tested compounds for *E. coli*) were part of the training set and an activity of MIC 25 μg/mL was set to discriminate good antibacterial activity and moderate/low activity. As the exact interactions and target site of these compounds are unknown, the relevance of each feature (i.e. hydrogen bond acceptor, hydrogen bond donor, hydrophobic atoms and aromatic rings) was considered equal.

Models obtained were ranked and selected based on the accuracy of the pharmacophore in discriminating positive (i.e. good antibacterial activity) and negative (moderate/low activity) compounds. Accuracies of the pharmacophore models were calculated as follows[60]:

Overall accuracy = proportion of total correct predictions (true positives and true negatives);

Positive accuracy (sensitivity) = proportion of correctly predicted positive results;

Negative accuracy (specificity) = proportion of correctly predicted negative results.

Data availability. The descriptor dataset generated for tested prenylated (iso)flavonoids is available on request.

References
1. World Health Organization. (2014). Antimicrobial resistance: global report on surveillance 2014. Geneva, Switzerland.
2. 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).
3. Upadhyay, A., Upadhyaya, I., Kollanoor-Johny, A. & Venkitanarayanan, K. Combating pathogenic microorganisms using plant-derived antimicrobials: A minireview of the mechanistic basis. *BioMed Res. Int.* **2014**, 1–18 (2014).
4. Araya-Cloutier, C., den Besten, H. M. W., Aisyah, S., Gruppen, H. & Vincken, J.-P. The position of prenylation of isoflavonoids and stilbenoids from legumes (Fabaceae) modulates the antimicrobial activity against Gram positive pathogens. *Food Chem.* **226**, 193–201 (2017).
5. Araya-Cloutier, C., Vincken, J.-P., van Ederen, R., den Besten, H. M. W. & Gruppen, H. Rapid membrane permeabilization of *Listeria monocytogenes* and *Escherichia coli* induced by antibacterial prenylated phenolic compounds from legumes. *Food Chem.* **240**, 147–155 (2018).

6. Simons, R., Gruppen, H., Bovee, T. F. H., Verbruggen, M. A. & Vincken, J.-P. Prenylated isoflavonoids from plants as selective estrogen receptor modulators (phytoSERMs). *Food Funct.* **3**, 810–827 (2012).
7. Tsuchiya, H. *et al.* Comparative study on the antibacterial activity of phytochemical flavanones against methicillin-resistant *Staphylococcus aureus*. *J. Ethnopharmacol.* **50**, 27–34 (1996).
8. Botta, B. *et al.* Prenylated isoflavonoids: Botanical distribution, structures, biological activities and biotechnological studies. An update (1995-2006). *Curr. Med. Chem.* **16**, 3414–3468 (2009).
9. Mbaveng, A. T. *et al.* Antimicrobial activities of the methanol extract and compounds from the twigs of *Dorstenia mannii* (Moraceae). *BMC Complement. Altern. Med.* **12**, 1–6 (2012).
10. Sato, M., Tanaka, H., Tani, N., Nagayama, M. & Yamaguchi, R. Different antibacterial actions of isoflavones isolated from *Erythrina poeppigiana* against methicillin-resistant *Staphylococcus aureus*. *Lett. Appl. Microbiol.* **43**, 243–248 (2006).
11. Mukne, A. P., Viswanathan, V. & Phadatare, A. G. Structure pre-requisites for isoflavones as effective antibacterial agents. *Pharmacogn. Rev.* **5**, 13–18 (2011).
12. Eerdunbayaer, MohamedA. A., Orabi, Hiroe, Aoyama, Kuroda, T. & Hatano, T. Structures of new phenolics isolated from licorice, and the effectiveness of licorice phenolics on vancomycin-resistant *Enterococci*. *Molecules* **19**, 13027–13041 (2014).
13. Hatano, T. *et al.* Phenolic constituents of licorice. VIII. Structures of glicophenone and glicoisoflavanone, and effects of licorice phenolics on methicillin-resistant *Staphylococcus aureus*. *Chem. Pharm. Bull.* **48**, 1286–1292 (2000).
14. Tanaka, H. *et al.* Antibacterial activity of isoflavonoids isolated from *Erythrina variegata* against methicillin-resistant *Staphylococcus aureus*. *Lett. Appl. Microbiol.* **35**, 494–498 (2002).
15. Sato, M. *et al.* Antibacterial activity of phytochemicals isolated from *Erythrina zeyheri* against vancomycin-resistant enterococci and their combinations with vancomycin. *Phytother. Res.* **18**, 906–910 (2004).
16. Rao, S. A., Merugu, R. & Rao, M. J. Antibacterial and antifungal studies of prenylated isoflavones and prenylated 3-aryl coumarins isolated from *Derris scandens* BENTH. *J. Phcog.* **3**, 51–54 (2012).
17. Yin, S., Fan, C.-Q., Wang, Y., Dong, L. & Yue, J.-M. Antibacterial prenylflavone derivatives from *Psoralea corylifolia*, and their structure-activity relationship study. *Bioorg. Med. Chem.* **12**, 4387–4392 (2004).
18. Zhou, B. & Wan, C.-X. Phenolic constituents from the aerial parts of *Glycyrrhiza inflata* and their antibacterial activities. *J. Asian Nat. Prod. Res.* **17**, 256–261 (2014).
19. Nanayakkara, N. P. D., Burandt, C. L. & Jacob, M. R. Flavonoids with activity against methicillin-resistant *Staphylococcus aureus* from *Dalea scandens* var. *paucifolia*. *Planta Med.* **68**, 519–522 (2002).
20. Wu, T., Zang, X., He, M., Pan, S. & Xu, X. Structure-activity relationship of flavonoids on their anti-*Escherichia coli* activity and inhibition of DNA gyrase. *J. Agric. Food Chem.* **61**, 8185–8190 (2013).
21. Villinski, J. R. *et al.* Pyrano-isoflavans from Glycyrrhiza *uralensis* with antibacterial activity against *Streptococcus mutans* and *Porphyromonas gingivalis*. *J. Nat. Prod.* **77**, 521–526 (2014).
22. Kar, S. & Roy, K. QSAR of phytochemicals for the design of better drugs. *Expert Opin. Drug Discov.* **7**, 877–902 (2012).
23. Yang, S.-Y. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Dis. Today* **15**, 444–450 (2010).
24. Hendrich, A. B. *et al.* Differential interaction of *Sophora* isoflavonoids with lipid bilayers. *Eur. J. Pharm. Sci.* **16**, 201–208 (2002).
25. Tsuchiya, H. & Iinuma, M. Reduction of membrane fluidity by antibacterial sophoraflavanone G isolated from *Sophora exigua*. *Phytomedicine* **7**, 161–165 (2000).
26. Wesołowska, O., Gąsiorowska, J., Petrus, J., Czarnik-Matusewicz, B. & Michalak, K. Interaction of prenylated chalcones and flavanones from common hop with phosphatidylcholine model membranes. *Biochim. Biophys. Acta* **1838**, 173–184 (2014).
27. Paterlini, S., Minerva, T. Regression model selection using genetic algorithms. In: *Proceedings of the 23rd workshop of the Italian Neural Networks Society* (ed^(eds). World Scientific and Engineering Academy and Society (WSEAS) (2010).
28. Xu, L. & Zhang, W.-J. Comparison of different methods for variable selection. *Analytica Chimica Acta* **446**, 475–481 (2001).
29. Davey, H. M. & Hexley, P. Red but not dead? Membranes of stressed *Saccharomyces cerevisiae* are permeable to propidium iodide. *Environ. Microbiol.* **13**, 163–171 (2011).
30. Li, X.-Z., Plésiat, P. & Nikaido, H. The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. *Clin. Microbiol. Rev.* **28**, 337–418 (2015).
31. Galdiero, S. *et al.* Microbe-host interactions: structure and role of Gram-negative bacterial porins. *Curr. Protein Peptide Sci.* **13**, 843–854 (2012).
32. Braun, V., Bös, C., Braun, M. & Killmann, H. Outer membrane channels and active transporters for the uptake of antibiotics. *J. Infect. Dis.* **183**, S12–S16 (2001).
33. Nikaido, H. & Rosenberg, E. Y. Porin channels in *Escherichia coli*: studies with liposomes reconstituted from purified proteins. *J. Bacteriol.* **153**, 241–252 (1983).
34. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* **67**, 593–656 (2003).
35. Black, P. N. & Zhang, Q. Evidence that His'10 of the protein FadL in the outer membrane of *Escherichia coli* is involved in the binding and uptake of long-chain fatty acids: possible role of this residue in carboxylate binding. *Biochem. J.* **310**, 389–394 (1995).
36. Richter, M. F. *et al.* Predictive compound accumulation rules yield a broad-spectrum antibiotic. *Nature* **545**, 299–304 (2017).
37. van den Berg, B. The FadL family: unusual transporters for unusual substrates. *Curr. Opin. Struct. Biol.* **15**, 401–407 (2005).
38. Hirai, K., Aoyama, H., Irikura, T., Iyobe, S. & Mitsuhashi, S. Differences in susceptibility to quinolones of outer membrane mutants of *Salmonella* Typhimurium and *Escherichia coli*. *Antimicrob. Agents Chemother.* **29**, 535–538 (1986).
39. Lambert, P. A. Cellular impermeability and uptake of biocides and antibiotics in Gram-positive bacteria and mycobacteria. *J. Appl. Microbiol.* **92**, 46S–54S (2002).
40. Liu, L., Fang, Y. & Wu, J. Flexibility is a mechanical determinant of antimicrobial activity for amphipathic cationic α-helical antimicrobial peptides. *Biochim. Biophys. Acta* **1828**, 2479–2486 (2013).
41. Reißer, S., Strandberg, E., Steinbrecher, T., Ulrich & Anne, S. 3D Hydrophobic moment vectors as a tool to characterize the surface polarity of amphiphilic peptides. *Biophys. J.* **106**, 2385–2394 (2014).
42. Kier, L. B. An index of molecular flexibility from kappa shape attributes. *Quant. Struct.-Act. Relat.* **8**, 221–224 (1989).
43. Selvaraj, S., Krishnaswamy, S., Devashya, V., Sethuraman, S. & Krishnan, U. M. Influence of membrane lipid composition on flavonoid–membrane interactions: Implications on their biological activity. *Prog. Lipid Res.* **58**, 1–13 (2015).
44. Hendrich, A. B. Flavonoid-membrane interactions: possible consequences for biological effects of some polyphenolic compounds. *Acta Pharmacol. Sin.* **27**, 27–40 (2006).
45. van de Schans, M. G. M. *et al.* Involvement of a hydrophobic pocket and helix 11 in determining the modes of action of prenylated flavonoids and isoflavonoids in the human estrogen receptor. *ChemBioChem* **16**, 2668–2677 (2015).
46. van de Schans, M. G. M. *et al.* Glyceollins and dehydroglyceollins isolated from soybean act as SERMs and ER subtype-selective phytoestrogens. *J. Steroid Biochem. Mol. Biol.* **156**, 53–63 (2016).
47. Bohnert, J. A. & Kern, W. V. Selected arylpiperazines are capable of reversing multidrug resistance in *Escherichia coli* overexpressing RND efflux pumps. *Antimicrob. Agents Chemother.* **49**, 849–852 (2005).
48. Aryani, D. C., den Besten, H. M. W., Hazeleger, W. C. & Zwietering, M. H. Quantifying strain variability in modeling growth of *Listeria monocytogenes*. *Int. J. Food Microbiol.* **208**, 19–29 (2016).
49. Dolezal, R. *et al.* Towards understanding the mechanism of action of antibacterial N-alkyl-3-hydroxypyridinium salts: Biological activities, molecular modeling and QSAR studies. *Eur J Med Chem* **121**, 699–711 (2016).

50. Khazaei, A., Sarmasti, N. & Seyf, J. Y. Quantitative structure–activity relationship of the curcumin-related compounds using various regression methods. *J. Mol. Struct.* **1108**, 168–178 (2016).
51. Westerhuis, J. A. *et al.* Assessment of PLSDA cross validation. *Metabolomics* **4**, 81–89 (2008).
52. Gramatica, P. Principles of QSAR models validation: internal and external. *QSAR Comb. Sci.* **26**, 694–701 (2007).
53. Tropsha, A. Best practices for QSAR model development, validation, and exploitation. *Mol. Inform.* **29**, 476–488 (2010).
54. Leardi, R. Genetic algorithms in chemistry. *J. Chromatogr. A* **1158**, 226–233 (2007).
55. Leach, A. *Molecular modelling principles and applications.* Pearson Education Limited (2001).
56. Gajo, G. C., de Assis, T. M., Assis, L. C., Ramalho, T. C. & da Cunha, E. F. F. Quantitative structure-activity relationship studies for potential Rho-associated protein kinase inhibitors. *Journal of Chemistry* **2016**, 12 (2016).
57. Assis, T. M. *et al.* QSAR models guided by molecular dynamics applied to human glucokinase activators. *Chem. Biol. Drug Des.* **87**, 455–466 (2016).
58. Roy, K. *et al.* Some case studies on application of "rm2" metrics for judging quality of quantitative structure–activity relationship predictions: Emphasis on scaling of response data. *J. Comput. Chem.* **34**, 1071–1082 (2013).
59. Roy, K., Kar, S. & Ambure, P. On a simple approach for determining applicability domain of QSAR models. *Chemometrics Intellig. Lab. Syst.* **145**, 22–29 (2015).
60. Singh, N. *et al.* QSAR classification model for antibacterial compounds and its use in virtual screening. *J. Chem. Inf. Model.* **52**, 2559–2569 (2012).

Author Contributions

C.A-C., J-P.V., H.M.W.d.B. and H.G. conceptualized the study; C.A-C. and M.G.M.v.d.S. designed and performed the experimental work (purification and antibacterial properties of compounds); C.A-C., J.H. and G.S. performed the data analysis (molecular modelling and statistical analysis); C.A-C., J-P.V., H.M.W.d.B. and H.G. drafted the manuscript. All the authors reviewed the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

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