

1

2 **Tables**

<b>MULTIALIGNMENT POSIX</b>	<b>CONSERVATION</b>	<b>QUALITY</b>	<b>CONSENSUS %</b>	<b>CONSENSUS RESIDUE</b>	<b>RHA-P RESIDUE</b>	<b>BT0986 RESIDUE</b>
1106	8	361.92047	88.46	D 88%	D465	D421
1153	10	375.69186	96.15	D 96%	D503	D458
1161	10	375.40494	95.38	E 95%	E506	Q461
1189	2	275.2199	67.69	D 67%	D527	D482
1238	10	334.8154	90.77	D 90%	D552	D507
1296	1	318.931	81.54	D 81%	D603	D556
1301	0	288.4868	79.23	E 79%	A608	E561
1361	3	366.1997	96.15	E 96%	E644	E593
1390	0	304.88068	81.54	D 81%	D665	D614
1477	2	312.0073	80.77	D 80%	D733	D683
1569	4	337.38684	90.00	D 90%	D763	D713

3

**Table 1**

4

5

6

7

8

**Table 1: Highly conserved residues.** Table showing the highly conserved carboxylic residues in the GH106 family sequences used for the alignment. The corresponding selected residues are indicated both in the RHA-P sequence and in the BT0986 sequence (used for the homology model).

**Table 2**

<b>Protein</b>	<b><math>K_M</math> (<math>\mu\text{M}</math>)</b>	<b><math>k_{\text{cat}}</math> (<math>\text{sec}^{-1}</math>)</b>	<b><math>k_{\text{cat}} / K_M</math> (<math>\text{sec}^{-1} \mu\text{M}^{-1}</math>)</b>
<b>Rha-Phis</b>	<b>157 (<math>\pm</math> 22)</b>	<b>692 (<math>\pm</math> 33)</b>	<b>4.4 (<math>\pm</math> 0.9)</b>
<b>Rha-Phis D552A</b>	<b>159 (<math>\pm</math> 23)</b>	<b>458 (<math>\pm</math> 21)</b>	<b>2.9 (<math>\pm</math> 0.5)</b>
<b>Rha-Phis D763A</b>	<b>138 (<math>\pm</math> 15)</b>	<b>126 (<math>\pm</math> 6)</b>	<b>0.9 (<math>\pm</math> 0.2)</b>

**Table 2: *p*NPR kinetic constants of RHA-Phis and mutants.** RHA-Phis, RHA-Phis/D552A and RHA-Phis/D763A kinetic constants using *p*NPR as substrate.

**Table 3**

<b>Flavonoid</b>	<b>Sugar</b>	<b>Glycosidic linkage position</b>	<b>K<sub>M</sub> (mM)</b>	<b>k<sub>cat</sub> (sec<sup>-1</sup>)</b>	<b>k<sub>cat</sub> / K<sub>M</sub> (sec<sup>-1</sup> mM<sup>-1</sup>)</b>
<b>Naringin</b>	<b>neohesperidose (α,1-2)</b>	<b>A ring Carbon 7</b>	<b>0.59 (±0.09)</b>	<b>214 (±10)</b>	<b>360 (±70)</b>
<b>Rutin</b>	<b>rutinose (α,1-6)</b>	<b>C ring Carbon 3</b>	<b>0.031 (±0.01)</b>	<b>18.2 (±1.2)</b>	<b>586 (±172)</b>
<b>Hesperidin</b>	<b>rutinose (α,1-6)</b>	<b>A ring Carbon 7</b>	<b>0.067 (±0.01)</b>	<b>12.7 (±0.4)</b>	<b>190 (±32)</b>
<b>Quercitrin</b>	<b>rhamnose</b>	<b>C ring Carbon 3</b>	<b>0.15 (±0.03)</b>	<b>12.2 (±0.4)</b>	<b>80 (±16)</b>

**Table 3: Kinetic constants of RHA-Phis on natural flavonoids.** RHA-Phis kinetic constants on naringin, rutin, hesperidin and quercitrin are shown.

## Supplementary Data legend

**Figure S1: Synthetic oligosaccharides used as substrates in RHA-P catalyzed reaction.** Structure of the three synthetic oligosaccharides used to investigate RHA-P specificity on different rhamnose-containing oligosaccharides.

**Figure S2: RHA-Phis analytical expression experiments.** Analytical experiments evaluating the specific activity of recombinant whole cells expressing RHA-Phis induced in different experimental conditions (named 1-8). These latter involved two different concentrations of IPTG (0.1 and 1 mM), and the addition of both betaine and sorbitol (0, 1, 5 and 10 mM) to the growth medium.

**Figure S3: Structural alignment used to prepare the homology model of rRHA-P.** Secondary structure is shown below the alignment (h, helix; s, extended). Loops labelled 1-3 are the same shown in magenta in Figure 3. Calcium binding residues are highlighted in red; the acid catalyst in cyan; residues contributing to rhamnose sub-site are in green; residues contributing to arabinose sub-site are in magenta; residues at the boundary between rhamnose and arabinose sub-sites are in grey.

**Figure S4: RHA-Phis, RHA-Phis/D552A and RHA-Phis/D763A kinetic behavior using *p*NPR as substrate.** Reaction rate expressed in  $\mu\text{M}/\text{min}$  is plotted as a function of *p*NPR  $\mu\text{M}$  concentration.

**Figure S5: RHA-Phis kinetic behavior using natural flavonoids as substrates.** Reaction rates expressed in  $\mu\text{M}/\text{min}$  are plotted as a function of naringin, rutin, hesperidin and quercitrin mM concentration.

**Table S1**

Primer name	Primer UP sequence 5' – 3'	Primer DOWN sequence 5' – 3'
RHA D504A UP/ RHA D504A DOWN	AGGGCTCTACTCACCGCGTC GATCGAGGTCGGT	ACCGACCTCGATCGACGCGGTGAGT AGAGCCCT
RHA E506A UP/ RHA E506A DOWN	CTCACCGATTGATCGCGGTC GGTGAGGCTAAC	GTTAGCCTCACCGACCGCGATCGAA TCGGTGAG
RHA D552A UP/ RHA D552A DOWN	GACCGGTTCTCTACGCGTAT CGCCGGACCCTC	GAGGGTCCGGCGATACGCGTAGAG GAACCGGTC
RHA E644A UP/ RHA E644A DOWN	AATCTCGTCTGCGGCATCC ATGACCGCATCG	CGATGCGGTCATGGATGCCGCAGC GACGAGATT

<b>RHA D763A UP/ RHA D763A DOWN</b>	<b>GTGCGCTATGCCTACGCGTAT ATTAACCTCAAT</b>	<b>ATTGAAGTTAATATACGCGTAGGCA TAGCGCAC</b>
---	---	---

**Table S1: Mutagenic primer sequences.** Primers used for RHA-P site-directed mutagenesis.

**Table S2**

Sequence number	Accession Number	Identity %	Description	Microorganism
1	CCA90848.1	100	glycoside hydrolase family protein	<i>Novosphingobium sp. PP1Y</i>
2	AIT81997.1	83.07	glycoside hydrolase	<i>Novosphingobium pentaromativorans US6-1</i>
3	ABD26167.1	60.67	glycoside hydrolase family 2, sugar binding protein	<i>Novosphingobium aromaticivorans DSM 12444</i>
4	AGU10471.1	57.23	hypothetical protein, partial	<i>uncultured organism</i>
5	BAD12237.1	55.88	alpha-L-rhamnosidase	<i>Sphingomonas paucimobilis</i>
6	AGU10470.1	53.85	hypothetical protein, partial	<i>uncultured organism</i>
7	CCH88181.1	43.59	Glycoside hydrolase family 2, sugar binding	<i>Modestobacter marinus</i>
8	ADV83038.1	43.3	coagulation factor 5/8 type domain protein	<i>Terriglobus saanensis SPIPR4</i>
9	AHF25049.1	42.92	glycoside hydrolase family 2	<i>uncultured bacterium Contig1514</i>
10	AIT08070.1	42.88	hypothetical protein MC45_07320	<i>Sphingomonas taxi</i>
11	ADB36422.1	42.66	coagulation factor 5/8 type domain protein	<i>Spirosoma linguale DSM 74</i>
12	ACL41491.1	42.61	glycoside hydrolase family 2 sugar binding	<i>Arthrobacter chlorophenolicus A6</i>
13	ACT92766.1	42.46	glycoside hydrolase family 2 sugar binding	<i>Dyadobacter fermentans DSM 18053</i>
14	EAR14584.1	42.25	hypothetical protein RB2501_00871	<i>Robiginitalea bifornata HTCC2501</i>
15	ACO34516.1	41.82	F5/8 type C domain/sugar binding domain protein	<i>Acidobacterium capsulatum ATCC 51196</i>
16	AEV33004.1	41.66	glycosyl hydrolases family 2,F5/8 type C domain-containing protein	<i>Owenweeksia hongkongensis DSM 17368</i>
17	AEI46908.1	41.28	coagulation factor 5/8 type domain protein	<i>Runella slithyformis DSM 19594</i>
18	ACB75560.1	39.94	glycoside hydrolase family 2 sugar binding	<i>Opitutus terrae PB90-1</i>
19	BAJ74682.1	39.67	hypothetical protein MTES_1718	<i>Microbacterium testaceum StLB037</i>
20	AHF25045.1	39.1	glycoside hydrolase family 2	<i>uncultured bacterium Contig1514</i>
21	AIE83958.1	38.32	putative glycosylhydrolase	<i>Fimbriimonas ginsengisoli Gsoil 348</i>
22	AEU35230.1	37.7	glycoside hydrolase family 2 sugar binding protein	<i>Granulicella mallensis MP5ACTX8</i>
23	AHE55710.1	37.65	hypothetical protein NX02_20265	<i>Sphingomonas sanxanigenens DSM 19645 = NX02</i>
24	AEK44383.1	37.38	coagulation factor 5/8 type domain protein	<i>Amycolatopsis mediterranei S699</i>
25	ADB37307.1	35.9	glycoside hydrolase family 2 sugar binding protein	<i>Spirosoma linguale DSM 74</i>
26	AHF15424.1	35.46	glycosyl transferase family 2	<i>Niabella soli DSM 19437</i>

27	CDF79916.1	35.33	glycoside hydrolase (GH2)	<i>Formosa agariphila</i> KMM 3901
28	AEV98273.1	34.87	glycoside hydrolase family 2 sugar binding protein	<i>Niastella koreensis</i> GR20-10
29	AAO78018.1	34.25	putative glycosylhydrolase	<i>Bacteroides thetaiotaomicron</i> VPI-5482
30	AHF91789.1	33.94	acyl-CoA thioesterase	<i>Opitutaceae</i> bacterium TAV5
31	AEI49030.1	33.79	glycoside hydrolase family 2 sugar binding protein	<i>Runella slithyformis</i> DSM 19594
32	ACU06077.1	33.67	glycoside hydrolase family 2 sugar binding	<i>Pedobacter heparinus</i> DSM 2366
33	ACU05683.1	33.26	glycoside hydrolase family 2 sugar binding	<i>Pedobacter heparinus</i> DSM 2366
34	ACT93511.1	33.17	glycoside hydrolase family 2 sugar binding	<i>Dyadobacter fermentans</i> DSM 18053
35	AEV99811.1	33.15	glycoside hydrolase family 2 sugar binding protein	<i>Niastella koreensis</i> GR20-10
36	AHF14591.1	32.97	glycoside hydrolase family 2 sugar binding protein	<i>Niabella soli</i> DSM 19437
37	AEE51369.1	32.92	glycoside hydrolase family 2 sugar binding protein	<i>Haliscomenobacter hydrossis</i> DSM 1100
38	AHF14289.1	32.89	glycosyl transferase family 2	<i>Niabella soli</i> DSM 19437
39	ACT93491.1	32.63	glycoside hydrolase family 2 sugar binding	<i>Dyadobacter fermentans</i> DSM 18053
40	ADZ77027.1	32.48	glycoside hydrolase family 2 sugar binding protein	<i>Sphingobacterium</i> sp. 21
41	ACU59305.1	32.44	glycoside hydrolase family 2 sugar binding	<i>Chitinophaga pinensis</i> DSM 2588
42	AHF16888.1	32.38	DNA-binding protein	<i>Niabella soli</i> DSM 19437
43	ACU04950.1	32.09	glycoside hydrolase family 2 sugar binding	<i>Pedobacter heparinus</i> DSM 2366
44	ACU04508.1	32.01	glycoside hydrolase family 2 sugar binding	<i>Pedobacter heparinus</i> DSM 2366
45	ACU06164.1	31.96	glycoside hydrolase family 2 sugar binding	<i>Pedobacter heparinus</i> DSM 2366
46	CBK68178.1	31.96	Glycosyl hydrolases family 2, sugar binding domain.	<i>Bacteroides xylanisolvens</i> XB1A
47	ACU03164.1	31.84	glycoside hydrolase family 2 sugar binding	<i>Pedobacter heparinus</i> DSM 2366
48	CBK65666.1	31.83	Glycosyl hydrolases family 2, sugar binding domain.	<i>Bacteroides xylanisolvens</i> XB1A
49	ADZ78900.1	31.81	glycoside hydrolase family 2 sugar binding protein	<i>Sphingobacterium</i> sp. 21
50	ADV83420.1	31.64	glycoside hydrolase family 2 sugar binding protein	<i>Terriglobus saanensis</i> SP1PR4
51	AEV99885.1	31.63	coagulation factor 5/8 type domain protein	<i>Niastella koreensis</i> GR20-10
52	AGA77964.1	31.35	glycosyl hydrolases family 2,F5/8 type C domain-containing protein	<i>Echinicola vietnamensis</i> DSM 17526
53	AII66314.1	31.04	hypothetical protein GV66_00650	<i>Bacteroides dorei</i>
54	AAO76093.1	30.92	Glycoside hydrolase family 2, sugar binding protein	<i>Bacteroides thetaiotaomicron</i> VPI-5482
55	ADY36872.1	30.53	glycoside hydrolase family 2 sugar binding protein	<i>Bacteroides salanitronis</i> DSM 18170
56	AIM36426.1	30.52	hypothetical protein KO02_06710	<i>Sphingobacterium</i> sp. ML3W
57	ABQ07109.1	30.47	alpha-L-rhamnosidase-like protein	not reported
58	ABR40348.1	30.27	conserved hypothetical protein, putative DNA-binding protein	<i>Bacteroides vulgatus</i> ATCC 8482
59	ADY50922.1	30.26	glycoside hydrolase family 2 sugar binding protein	<i>Pedobacter saltans</i> DSM 12145
60	AEW21961.1	30.24	glycosyl hydrolase family 2, sugar binding domain protein	<i>Tannerella forsythia</i> ATCC 43037
61	AII62838.1	30.17	DNA-binding protein	<i>Bacteroides dorei</i>
62	AII68823.1	30.17	DNA-binding protein	<i>Bacteroides dorei</i>
63	AII67845.1	30.12	glycosyl hydrolase family 2	<i>Bacteroides dorei</i>
64	EDO11490.1	29.81	glycosyl hydrolase family 2, sugar binding domain protein	<i>Bacteroides ovatus</i> ATCC 8483

65	ADY36873.1	29.39	glycoside hydrolase family 2 sugar binding protein	<i>Bacteroides salanitronis</i> DSM 18170
66	AEW21098.1	28.96	glycosyl hydrolase family 2, sugar binding domain protein	<i>Tannerella forsythia</i> ATCC 43037
67	AII65174.1	27.75	glycosyl hydrolase family 2	<i>Bacteroides dorei</i>
68	ADG10252.1	27.2	conserved hypothetical protein	<i>Caulobacter segnis</i> ATCC 21756
69	AII66369.1	27.2	glycoside hydrolase family 2	<i>Bacteroides dorei</i>
70	AII64737.1	27.08	glycoside hydrolase family 2	<i>Bacteroides dorei</i>
71	AEQ20323.1	27.07	glycosyl hydrolase family 2 protein	uncultured bacterium CSLG10
72	AEL26093.1	27.04	glycoside hydrolase family 2 sugar binding protein	<i>Cyclobacterium marinum</i> DSM 745
73	ABR37891.1	26.96	conserved hypothetical protein	<i>Bacteroides vulgatus</i> ATCC 8482

1

2 **Table S2: Identity percent of 130 GH106 protein sequences vs RHA-Phis.**

3