

Metagenome derived carbohydrate active enzymes for directed modification of polyphenols

Development of a TLC-based functional screening method for identification of flavonoid modifying enzymes and the isolation and characterization of flavonoid modifying glycosyltransferases and glycoside hydrolases

Zur Erlangung des Doktorgrades der Naturwissenschaften (*Dr. rer. nat.*)
im Fachgebiet Mikrobiologie und Biotechnologie des Fachbereichs Biologie
der Fakultät für Mathematik, Informatik und Naturwissenschaften,
der Universität Hamburg

vorgelegt von

Ulrich Rabausch, geb. Köhler

aus Hamburg

Hamburg 2013

Genehmigt vom Fachbereich Biologie
der Fakultät für Mathematik, Informatik und Naturwissenschaften
an der Universität Hamburg
auf Antrag von Professor Dr. W. STREIT
Weiterer Gutachter der Dissertation:
Professor Dr. B. BISPING
Tag der Disputation: 7. Juni 2013

Hamburg, den 18. Juni 2013

A handwritten signature in black ink, consisting of a stylized, cursive 'L' followed by a horizontal line extending to the right.

Professor Dr. C. Lohr
Vorsitzender des
Fach-Promotionsausschusses Biologie

Vorwort zum Stand der Wissenschaft

Polyphenole und Flavonoide

Flavonoide gehören zu den Polyphenolen und sind sekundäre Pflanzeninhaltsstoffe, die über den Shikimisäureweg biosynthetisiert werden (1). Von vielen Polyphenolen sind für den Menschen gesundheitsfördernde Wirkungen bekannt (2). Neben der antioxidativen und Radikalbindenden Funktion können die Verbindungen auf den Organismus antiallergen, antibakteriell, antifungal, antiviral, entzündungshemmend, schmerzstillend, gefäßstabilisierend, kreislauf-fördernd, hormonell und sogar krebsvorbeugend wirken (3). Aufgrund der vielfältigen Wirk-samkeiten werden Flavonoide daher zunehmend in der Kosmetik-, der Nahrungs- und Nah-rungsergänzungsmittelindustrie und der Pharmazie eingesetzt und unterliegen deshalb einer stark wachsenden Nachfrage (4-6).

Flavonoide leiten sich chemisch von der Grund-struktur des Phenylchromans ab, wobei die C2-C3 Bindung des Propanoidrestes konjugiert und C4 eine Ketogruppe tragen kann. Viele C-Atome des Ringsystems können mit Hydroxylgruppen besetzt oder diese substituiert sein. Als sekundäre Modifi-kationen kommen in der Natur Acylierungen, Gly-kosylierungen, Methylierungen, Phosphorylierun-gen und Sulfatierungen des Ringsystems, aber auch Oligomerisierungen vor. All diese führen ins-gesamt zu einer großen Vielfalt an natürlichen Derivaten dieser Stoffklasse. Harborne und Baxter berichteten 1999 bereits fast 6.500 beschriebene Verbindungen (7). An Polyphenolen waren 1998 bereits über 8.000 Verbindungen bekannt (8).

Zuckermodifikationen kommen dabei häufig in der Natur vor. Sie können die Löslichkeit und Funktion der Stoffe stark beeinflussen (9, 10). Auch wird die Bioverfügbarkeit, die Fähigkeit des Organismus diese Verbindungen über Epithelzellen aufzunehmen, stark durch die Glykosylierung bestimmt (11).

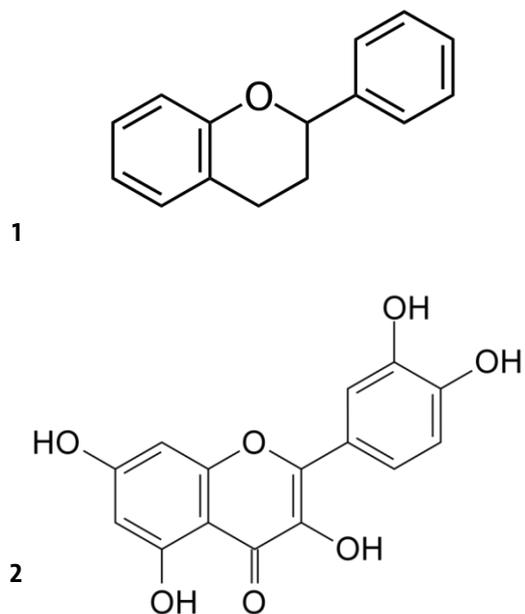


Abbildung 1: Flavonoidgrundstruktur des 2-Phenylbenzopyrans (1) und des Flavonols Quercetin (2).

Bekannte Flavonoid-modifizierende Enzyme

Als Flavonoid-modifizierende Enzyme waren Glykosyltransferasen (EC 2.4.1) von besonderem Interesse. An pflanzlichen Enzymen, die für die sekundäre Modifikation von Flavonoiden verantwortlich sind, wurden speziell Glykosyltransferasen gut untersucht. Auch war die Kristallstruktur einiger pflanzlichen Glykosyltransferasen, teilweise in Bindung mit ihren Donor- und Akzeptorsubstraten, bereits aufgeklärt (12-14). Die Flavonoid-Glykosyltransferasen gehören zur Familie 1 der Glykosyltransferasen (GT1) (15). In GT1 werden Enzyme klassifiziert, die Zuckerreste von aktivierten Donorsubstraten, meist UDP-Glukose, auf lipophile Akzeptorverbindungen übertragen (16). Sie katalysieren einheitlich eine Transferreaktion von Zuckerresten bei der die Bindung des anomeren C-Atoms invertiert wird und besitzen stets eine als GT-B Faltung bezeichnete Tertiärstruktur (17). In der Primärstruktur lässt sich das so genannte PSPG-Motiv, ein konserviertes Sequenzmotiv, aufzeigen, das im C-terminalen Bereich der Enzyme lokalisiert ist und maßgeblich die Donorbinde-Domäne konstituiert (18, 19).

Bei Bakterien war die Fähigkeit Flavonoide zu glykosylieren in einem *Bacillus cereus* Stamm bereits Anfang der achtziger Jahre des letzten Jahrhunderts gezeigt worden (20). Bis zu Beginn dieser Dissertation waren allerdings nur wenige weitere Studien zur Glykosylierung von Flavonoiden durch bakterielle Enzyme publiziert worden. Bakterielle Glykosyltransferasen, denen eine Rolle in der Detoxifizierung von Xenobiotika zugeschrieben wird, waren als Flavonoid-modifizierende Enzyme bekannt aus *B. cereus*, *Xanthomonas campestris* und *Streptomyces* spp. (21-23). Von diesen fallen BcGT-1 aus *B. cereus* ATCC10987, OleD aus *Str. antibioticus* und Mgt aus *Str. lividans* in die Subfamilie der sog. Makrosid-Glykosyltransferasen (TIGR01426) eine Unterfamilie der UDP-Glykosyltransferasen (PFAM00201) innerhalb GT1.

Zudem waren einige bakterielle Enzyme aus der Glykosidhydrolase Familie GH13 bekannt (24), die Zuckerreste von Di- und Oligosacchariden auf polyphenolische Verbindungen übertragen können. Zu ihnen zählen die Cyclomaltoextrin-Glukanotransferasen, α -Amylasen aus *Bacillus* spp. und *Thermotoga maritima* (25-28), sowie Sucrasen aus den *Lactobacillaceae*, *Leuconostoc mesenteroides* und *Streptococcus*-Stämmen (29-33).

Biotechnisch bedeutende Flavonoid-modifizierende Enzyme sind die α -L-Rhamnosidasen (34). Die bakteriellen α -L-Rhamnosidasen sind Glykosidhydrolasen (EC 3.2.1.40), die L-Rhamnose von Flavonoidrutinosiden (-6-O- α -L-Rhamnopyranosyl- β -D-Glukopyranoside) und Flavonoidneohesperidosiden (-2-O- α -L-Rhamnopyranosyl- β -D-Glukopyranoside) abspalten können. Biotech-

nisch stehen dahinter verschiedene Interessen. Zum einen werden Flavonoidglukoside besser aufgenommen als Disaccharide und sind z. T. weniger bitter im Geschmack, was bei der Fruchtsaftproduktion von Interesse ist (11, 34). Andererseits ist die frei werdende L-Rhamnose eine begehrte chirale Verbindung als Ausgangssubstanz für die chemische Synthese (34).

Die bakteriellen α -L-Rhamnosidasen gruppieren sich in den Glykosidhydrolase-Familien GH78 und GH106, wobei RhaM aus *Sphingomonas paucimobilis* FP2001 das bis dato einzig charakterisierte bakterielle Enzym der GH106 darstellt (35). In GH78, die ausschließlich α -L-Rhamnosidasen des A und B Typs umfasst, waren Enzyme aus *Bacillus* sp. GL1, *Bacteroides* spp., *Clostridium stercorarium*, *Fusobacterium* sp. K-60 und *Thermomicrobium* sp. PRI-1686 untersucht worden, die Flavonoide derhamnosylierten (36-41).

Metagenomik

Seit Beginn der Metagenomik-Ära Ende der neunziger Jahre sind mit dieser Technik viele neuartige Enzyme entdeckt worden (42-45). Allerdings handelt es sich bei den in funktionellen Analysen entdeckten Enzymen vornehmlich um Oxidoreduktasen (EC 1) und Hydrolasen (EC 3). Ein limitierender Faktor für die Entdeckung anderer Enzymklassen ist der Mangel an funktionellen Untersuchungsmethoden (46, 47). Diese Tatsache wird u.a. in der Einleitung des Manuskripts zur Veröffentlichung "Functional Screening of Metagenome and Genome Libraries for Detection of Novel Flavonoid Modifying Enzymes" beleuchtet (48). Speziell für die Enzymklasse der Transferasen (EC 2) wurde bis dato keine metagenomische Anwendung beschrieben (49), auch wenn Testsysteme wie z.B. für den Nachweis der Aktivität von Glykosyltransferasen im Mikrotiter-Maßstab bekannt waren (50).

Literaturverweise

1. **Andersen, Ø. M., and Markham, K. R.** 2006. Flavonoids: Chemistry, biochemistry, and applications. CRC, Taylor & Francis, Boca Raton, FL.
2. **Ross, J. A., and Kasum, C. M.** 2002. DIETARY FLAVONOIDS: Bioavailability, Metabolic Effects, and Safety. *Annu Rev Nutr* **22**:19-34.
3. **Ververidis, F., Trantas, E., Douglas, C., Vollmer, G., Kretzschmar, G., and Panopoulos, N.** 2007. Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part II: Reconstruction of multienzyme pathways in plants and microbes. *Biotechnol J* **2**:1235-1249.
4. **Schütz, K., Muks, E., Carle, R., and Schieber, A.** 2006. Quantitative determination of phenolic compounds in artichoke-based dietary supplements and pharmaceuticals by high-performance liquid chromatography. *J Agric Food Chem* **54**:8812-8817.
5. **Wang, M., Liang, C., Pei, Wu, Q., Li, Simon, J., E., and Ho, C., (ed.)** 2006. Instrumental analysis of popular botanical products in the U.S. market, p. 25–38. *In Herbs: challenges in chemistry and biology.* American Chemical Society, Washington, DC.
6. **Leonard, E., Yan, Y., Fowler, Z. L., Li, Z., Lim, C. G., Lim, K. H., and Koffas, M. A.** 2008. Strain improvement of recombinant *Escherichia coli* for efficient production of plant flavonoids. *Mol Pharm* **5**:257-265.
7. **Harborne, J. B., and Baxter, H.** 1999. The handbook of natural flavonoids. Volume 1 and Volume 2. Wiley, Chichester, UK.
8. **Bravo, L.** 1998. Polyphenols: chemistry, dietary sources, metabolism, and nutritional significance. *Nutr Rev* **56**:317-333.
9. **Graefe, E. U., Wittig, J., Mueller, S., Riethling, A. K., Uehleke, B., Drewelow, B., Pforte, H., Jacobasch, G., Derendorf, H., and Veit, M.** 2001. Pharmacokinetics and bioavailability of quercetin glycosides in humans. *J Clin Pharmacol* **41**:492-499.
10. **Kren, V., and Martinkova, L.** 2001. Glycosides in medicine: The role of glycosidic residue in biological activity. *Curr Med Chem* **8**:1303-1328.
11. **Hollman, P. C. H., Bijlsman, M. N. C. P., van Gameren, Y., Cnossen, E. P. J., de Vries, J. H. M., and Katan, M. B.** 1999. The sugar moiety is a major determinant of the absorption of dietary flavonoid glycosides in man. *Free Radic Res* **31**:569-573.
12. **Li, L., Modolo, L. V., Escamilla-Trevino, L. L., Achnine, L., Dixon, R. A., and Wang, X.** 2007. Crystal structure of *Medicago truncatula* UGT85H2--insights into the structural basis of a multifunctional (iso)flavonoid glycosyltransferase. *J Mol Biol* **370**:951-963.
13. **Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B. G., Tarling, C. A., Ford, C. M., Bowles, D. J., and Davies, G. J.** 2006. Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *Embo J* **25**:1396-1405.
14. **Shao, H., He, X., Achnine, L., Blount, J. W., Dixon, R. A., and Wang, X.** 2005. Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. *Plant Cell* **17**:3141-3154.
15. **Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B.** 2003. An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol* **328**:307-317.
16. **Bowles, D., Lim, E. K., Poppenberger, B., and Vaistij, F. E.** 2006. Glycosyltransferases of lipophilic small molecules. *Annu Rev Plant Biol* **57**:567-597.
17. **Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G.** 2008. Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* **77**:521-555.
18. **Paquette, S., Moller, B. L., and Bak, S.** 2003. On the origin of family 1 plant glycosyltransferases. *Phytochemistry* **62**:399-413.

19. **Hughes, J., and Hughes, M. A.** 1994. Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. *DNA Seq* **5**:41-49.
20. **Rao, K. V., and Weisner, N. T.** 1981. Microbial Transformation of Quercetin by *Bacillus cereus*. *Appl Environ Microbiol* **42**:450-452.
21. **Ko, J. H., Gyu Kim, B., and Joong-Hoon, A.** 2006. Glycosylation of flavonoids with a glucosyltransferase from *Bacillus cereus*. *FEMS Microbiol Lett* **258**:263-268.
22. **Kim, H. J., Kim, B. G., Kim, J. A., Park, Y., Lee, Y. J., Lim, Y., and Ahn, J. H.** 2007. Glycosylation of flavonoids with *E. coli* expressing glucosyltransferase from *Xanthomonas campestris*. *J Microbiol Biotechnol* **17**:539-542.
23. **Yang, M., Proctor, M. R., Bolam, D. N., Errey, J. C., Field, R. A., Gilbert, H. J., and Davis, B. G.** 2005. Probing the breadth of macrolide glucosyltransferases: in vitro remodeling of a polyketide antibiotic creates active bacterial uptake and enhances potency. *J Am Chem Soc* **127**:9336-9337.
24. **Henrissat, B., and Davies, G.** 1997. Structural and sequence-based classification of glycoside hydrolases. *Current Opinion in Structural Biology* **7**:637-644.
25. **Shimizu, R., Shimabayashi, H., and Moriwaki, M.** 2006. Enzymatic production of highly soluble myricitrin glycosides using beta-galactosidase. *Biosci Biotechnol Biochem* **70**:940-948.
26. **Suzuki, Y., and Suzuki, K.** 1991. Enzymatic formation of 4G-alpha-D-glucopyranosyl-rutin. *Agric Biol Chem* **55**:181-187.
27. **Kometani, T., Nishimura, T., Nakae, T., Takii, H., and Okada, S.** 1996. Synthesis of neohesperidin glycosides and naringin glycosides by cyclodextrin glucanotransferase from an alkalophilic *Bacillus* species. *Biosci Biotechnol Biochem* **60**:645-649.
28. **Li, D., Park, J. H., Park, J. T., Park, C. S., and Park, K. H.** 2004. Biotechnological production of highly soluble daidzein glycosides using *Thermotoga maritima* maltosyltransferase. *J Agric Food Chem* **52**:2561-2567.
29. **Kitao, S., Ariga, T., Matsudo, T., and Sekine, H.** 1993. The Syntheses of Catechin-Glucosides by Transglycosylation with *Leuconostoc mesenteroides* Sucrose Phosphorylase. *Biosci Biotechnol Biochem* **57**:2010-2015.
30. **Moon, Y.-H., Jin-Ha Lee, Deok-Young Jhon, Woo-Jin Jun, Seong-Soo Kang, Jeonggu Sim, Heungsic Choi, Jae-Hak Moon and Doman Kim.** 2007. Synthesis and characterization of novel quercetin- α -d-glucopyranosides using glucansucrase from *Leuconostoc mesenteroides*. *Enzyme Microbial Technol* **40**:1124-1129.
31. **Meulenbeld, G. H., Zuilhof, H., van Veldhuizen, A., van den Heuvel, R. H., and Hartmans, S.** 1999. Enhanced (+)-catechin transglucosylating activity of *Streptococcus mutans* GS-5 glucosyltransferase-D due to fructose removal. *Appl Environ Microbiol* **65**:4141-4147.
32. **Nakahara, K., Kontani, M., Ono, H., Kodama, T., Tanaka, T., Ooshima, T., and Hamada, S.** 1995. Glucosyltransferase from *Streptococcus sobrinus* Catalyzes Glucosylation of Catechin. *Appl Environ Microbiol* **61**:2768-2770.
33. **Sato, T., Nakagawa, H., Kurosu, J., Yoshida, K., Tsugane, T., Shimura, S., Kirimura, K., Kino, K., and Usami, S.** 2000. Alpha-anomer-selective glucosylation of (+)-catechin by the crude enzyme, showing glucosyl transfer activity, of *Xanthomonas campestris* WU-9701. *J Biosci Bioeng* **90**:625-630.
34. **Manzanares, P., Vallés, S., Ramòn, D., and Orejas, M.** 2007. α -L-rhamnosidases: Old and New Insights, p. 117-140. *In Industrial Enzymes*. Springer Netherlands.

35. **Miake, F., Satho, T., Takesue, H., Yanagida, F., Kashige, N., and Watanabe, K.** 2000. Purification and characterization of intracellular α -L-rhamnosidase from *Pseudomonas paucimobilis* FP2001. *Arch Microbiol* **173**:65-70.
36. **Hashimoto, W., Miyake, O., Nankai, H., and Murata, K.** 2003. Molecular identification of an alpha-L-rhamnosidase from *Bacillus* sp strain GL1 as an enzyme involved in complete metabolism of gellan. *Arch Biochem Biophys* **415**:235-244.
37. **Zverlov, V. V., Hertel, C., Bronnenmeier, K., Hroch, A., Kellermann, J., and Schwarz, W. H.** 2000. The thermostable alpha-L-rhamnosidase RamA of *Clostridium stercoarium*: biochemical characterization and primary structure of a bacterial alpha-L-rhamnoside hydrolase, a new type of inverting glycoside hydrolase. *Mol Microbiol* **35**:173-179.
38. **Jang, I. S., and Kim, D. H.** 1996. Purification and characterization of alpha-L-rhamnosidase from *Bacteroides* JY-6, a human intestinal bacterium. *Biol Pharm Bull* **19**:1546-1549.
39. **Bokkenheuser, V. D., Shackleton, C. H., and Winter, J.** 1987. Hydrolysis of dietary flavonoid glycosides by strains of intestinal *Bacteroides* from humans. *Biochem J* **248**:953-956.
40. **Park, S.-Y., Kim, J.-H., and Kim, D.-H.** 2005. Purification and Characterization of Quercitrin-Hydrolyzing alpha-L-Rhamnosidase from *Fusobacterium* K-60, a Human Intestinal Bacterium. *J Microbiol Biotechnol* **15**:519-524.
41. **Birgisson, H., Hreggvidsson, G. O., Fridjónsson, O. H., Mort, A., Kristjánsson, J. K., and Mattiasson, B.** 2004. Two new thermostable alpha-L-rhamnosidases from a novel thermophilic bacterium. *Enzyme Microb Technol* **34**:561-571.
42. **Ferrer, M., Beloqui, A., Timmis, K. N., and Golyshein, P. N.** 2009. Metagenomics for mining new genetic resources of microbial communities. *J Mol Microbiol Biotechnol* **16**:109-123.
43. **Tuffin, M., Anderson, D., Heath, C., and Cowan, D. A.** 2009. Metagenomic gene discovery: How far have we moved into novel sequence space? *Biotechnol J* **4**:1671-1683.
44. **Steele, H. L., Jaeger, K. E., Daniel, R., and Streit, W. R.** 2009. Advances in recovery of novel biocatalysts from metagenomes. *J Mol Microbiol Biotechnol* **16**:25-37.
45. **Simon, C., and Daniel, R.** 2009. Achievements and new knowledge unraveled by metagenomic approaches. *Appl Microbiol Biotechnol* **85**:265-276.
46. **Uchiyama, T., and Miyazaki, K.** 2009. Functional metagenomics for enzyme discovery: challenges to efficient screening. *Curr Opin Biotechnol* **20**:616-622.
47. **Taupp, M., Mewis, K., and Hallam, S. J.** 2011. The art and design of functional metagenomic screens. *Curr Opin Biotechnol* **22**:465-472.
48. **Rabausch, U., Juergensen, J., Ilmberger, N., Böhnke, S., Fischer, S., Schubach, B., Schulte, M., and Streit, W. R.** 2013. Functional Screening of Metagenome and Genome Libraries for Detection of Novel Flavonoid Modifying Enzymes. *Appl Environ Microbiol* **79**:4551-4563.
49. **Perner, M., Ilmberger, N., Köhler, H. U., Chow, J., and Streit, W. R.** 2011. Emerging Fields in Functional Metagenomics and Its Industrial Relevance: Overcoming Limitations and Redirecting the Search for Novel Biocatalysts, p. 481-498. *In Handbook of Molecular Microbial Ecology II*. John Wiley & Sons, Inc.
50. **Aharoni, A., Thieme, K., Chiu, C. P., Buchini, S., Lairson, L. L., Chen, H., Strynadka, N. C., Wakarchuk, W. W., and Withers, S. G.** 2006. High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat Methods* **3**:609-614.

Publikationsliste

Aus den Forschungsergebnissen der Arbeit sind folgende Publikationen hervorgegangen, die als Nachweis einer kumulativen Dissertation zusammen mit einem Forschungsbericht eingereicht wurden.

1. **Rabausch, U.**, Ilmberger, N., and Streit, W.R. 2014. The metagenome-derived enzyme RhaB opens a new subclass of bacterial B type α -L-rhamnosidases. *J Biotechnol* 191: 38-45.
2. **Rabausch, U.**, Juergensen, J., Ilmberger, N., Böhnke, S., Fischer, S., Schubach, B., Schulte, M., and Streit, W. R. 2013. Functional Screening of Metagenome and Genome Libraries for Detection of Novel Flavonoid Modifying Enzymes. *Appl Environ Microbiol* 79: 4551-4563.
3. Wiegand, S., **Rabausch, U.**, Chow, J., Daniel, R., Streit, W.R., and Liesegang, H. 2013. Complete Genome Sequence of *Geobacillus* sp. GHH01, a Thermophilic Lipase-secreting Bacterium. *Genome Announc* March/April 2013 1 (2) e00092-13
4. Ilmberger, N., Meske, D., Juergensen, J., Schulte, M., Barthen, P., **Rabausch, U.**, Angelov, A., Mientus, M., Liebl, W., Schmitz, R. A., and Streit, W. R. 2012. Metagenomic cellulases highly tolerant towards the presence of ionic liquids-linking thermostability and halotolerance. *Appl Microbiol Biotechnol* 95: 135-146
5. Perner, M., Ilmberger, N., **Köhler, H. U.**, Chow, J., and Streit, W. R. 2011. Emerging Fields in Functional Metagenomics and Its Industrial Relevance: Overcoming Limitations and Redirecting the Search for Novel Biocatalysts, p 481-498, *In Handbook of Molecular Microbial Ecology II*. John Wiley & Sons, Inc.

Nachfolgend sind beiden erstgenannten Publikationen angefügt. Sie enthalten die wesentlichen Ergebnisse und Erkenntnisse der Arbeit.

Ulrich Rabausch



The metagenome-derived enzyme RhaB opens a new subclass of bacterial B type α -L-rhamnosidases

U. Rabausch, N. Ilmberger, W.R. Streit*

Abteilung für Mikrobiologie und Biotechnologie, Biozentrum Klein Flottbek, Universität Hamburg, Ohnhorststraße 18, D-22609 Hamburg, Germany

ARTICLE INFO

Article history:

Received 26 February 2014
Received in revised form 23 April 2014
Accepted 28 April 2014
Available online 6 May 2014

Keywords:

α -L-rhamnosidase
Glycoside hydrolase family 78
Metagenome analysis
Intestinal bacteria
Microbiome

ABSTRACT

A combined sequence- and function-based analysis of a metagenomic library DNA derived from elephant feces led to the identification of a novel bacterial α -L-rhamnosidase belonging to glycoside hydrolase family 78 (GH78). The gene was designated *rhaB* (4095 bp) and encoded for a putative protein of 1364 amino acids. The C-terminal part of the enzyme revealed an amino acid (AA) sequence identity of 58% to a predicted bacterial α -L-rhamnosidase from *Bacteroides nordii*. Interestingly, the N-terminal region of the deduced enzyme RhaB contained a GDSL-like lipase motif and an acetyl-xylan esterase (DAP2) motif. While heterologous expression of the complete *rhaB* failed, subcloning of the gene identified the most active open reading frame (ORF) to be of 3081 bp, which we designated *rhaB1*. The enzyme RhaB1 was overexpressed in *Escherichia coli* BL21 (DE3) and was purified to an amount of 75 mg/L of culture medium. In accordance to the intestinal origin, RhaB1 showed a preference for mesophilic conditions with an optimum activity at a temperature T_{Opt} of 40 °C and a pH_{Opt} of 6.5, respectively. The recombinant protein had a K_m value of 0.79 mM and a specific activity ν_{max} of 18.4 U for pNP- α -L-rhamnose, a calculated K_m of 6.36 mM and ν_{max} of 2.9×10^{-3} U for naringin, and a K_m of 6.75 mM and specific activity ν_{max} of 8.63×10^{-2} U for rutin, respectively. Phylogenetic analysis and amino acid domain architecture comparison revealed that RhaB1 belongs to a new subclass of bacterial B type α -L-rhamnosidases of GH 78. To our knowledge RhaB1 is the first biochemically-characterized enzyme of this subclass.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The α -L-rhamnosidases (EC 3.2.1.40) are inverting glycoside hydrolases that catalyze the cleavage of terminal α -L-rhamnose residues (Yadav et al., 2010). Most known bacterial α -L-rhamnosidases possess the glycoside hydrolase (GH) motif of GH family 78 (GH78) (Henrissat and Davies, 1997). Only few examples are known that are members of GH family 106 (Miyata et al., 2005). The eukaryotic α -L-rhamnosidases RgxB from *Aspergillus niger* CBS51388 and RhaR from *A. niger* DLFCC-90 are grouped within the GH-families GH28 and GH13, respectively (Liu et al., 2012; Martens-Uzunová et al., 2006).

The majority of characterized bacterial α -L-rhamnosidases to date originate from Gram-positive bacteria, i.e. *Bacillus*, *Lactobacillus*, *Staphylococcus*, *Clostridium*, *Streptomyces* and *Thermomicrobium* species (Avila et al., 2009; Beekwilder et al., 2009; Birgisson et al., 2004; Hashimoto et al., 1999; Ichinose et al., 2013; Michlmayr et al.,

2011; Puri and Kaur, 2010; Zverlov et al., 2000). RhaB, one of the two published enzymes of *Bacillus* sp. GL1 was successfully crystallized and its structure determined by X-ray crystallography (Cui et al., 2007).

Some of the so far characterized α -L-rhamnosidases originate from intestinal bacteria including *Bacteroides* spp. isolated of the human gut (Bokkenheuser et al., 1987; Jang and Kim, 1996; Park et al., 2008). Until now the intestinal tract of several animals has been shown to be a rich source for polysaccharide degrading bacteria and their enzymes (Feng et al., 2009; Gloux et al., 2011; Ilmberger et al., 2012; Warnecke et al., 2007). These bacteria and their so called carbohydrate active enzymes (CAZymes, (Cantarel et al., 2009)) play an important role for their hosts in digestion but also affecting even more than alimentation (Hehemann et al., 2012; Maslowski and Mackay, 2011; Nicholson et al., 2012). A predominant bacterial phylum of gut organisms are the *Bacteroidetes* (Thomas et al., 2011). A putative α -L-rhamnosidase from *Bacteroides thetaiotaomicron* VPI-5482 was already crystallized (PDB: 3CIH).

From metagenomic approaches to date most genes of putative GH78 enzymes were annotated though rarely tested for

* Corresponding author. Tel.: +49 40 428 16 463.

E-mail address: wolfgang.streit@uni-hamburg.de (W.R. Streit).

α -L-rhamnosidase activity. For instance, two GH78 genes were identified in one clone from the metagenome of a human fecal sample (Tasse et al., 2010). In a more recent study on the termite abdominal microbiome also sequences possessing GH78 motifs were published but not characterized (Bastien et al., 2013). Further on the functional level, two α -L-rhamnosidases were identified by activity assays in a comprehensive functional study on glycoside hydrolases (GHs) from cow rumen (Ferrer et al., 2012).

The α -L-rhamnosidases are of interest for chemical industry, especially pharmaceutical and food industries (Manzanares et al., 2007). The commercial enzyme preparations naringinase and hesperidinase, both of fungal origin, comprise α -L-rhamnosidase next to β -D-glucosidase activity (Monti et al., 2004; Young et al., 1989). Naringinases and hesperidinases are applied for debittering and clarifying of fruit juices (Ribeiro, 2011; Yadav et al., 2010). The aroma optimization of wine is a further potential field of application of α -L-rhamnosidases (Manzanares et al., 2003; Michlmayr et al., 2011). The derhamnosylating function of α -L-rhamnosidases can enhance the pharmacological properties of drugs like antibiotics as chloropolysporin, steroids as ruscogenin or gypenosides (Di Lazzaro et al., 2001; Feng et al., 2005; Takatsu et al., 1987; Yu et al., 2004). As a byproduct of derhamnosylations L-rhamnose is produced from natural substances. L-rhamnose itself is a demanded precursor as a chiral compound for chemical synthesis but also for glycodiversification of natural compounds (Manzanares et al., 2007; Puri, 2012; Thibodeaux et al., 2007). A rich source of L-rhamnose containing biomolecules are flavonoids. The use of citrus fruit peel waste from juice making as a source of naringin and hence prunin and L-rhamnose was already demonstrated by Kaur and colleagues (Kaur et al., 2010). Flavonoids are well known for their beneficial effects on human health (Ververidis et al., 2007) whereof the mono-glycosylated forms like isoquercitrin and prunin show a higher bioavailability than the respective disaccharides (Hollman et al., 1999). Because of the bioactivity flavonoids meet an increasing demand in the cosmetic, the pharma- and nutraceutical industries (Leonard et al., 2008; Schütz et al., 2006; Wang et al., 2006).

For the discovery of novel α -L-rhamnosidases the metagenome of the intestinal microbiome of an Asian elephant was chosen. In a combined sequence and function based screening approach a metagenome library of fresh feces was analyzed where the putative α -L-rhamnosidase gene *rhaB* was discovered. The gene was subsequently cloned for functional analyses and the active enzyme RhaB1 was characterized biochemically.

2. Material and methods

2.1. Bacterial strains, plasmids and chemical reagents

Bacterial strains and plasmids used in the present work are listed in Table S1 and primers are listed in Table S2. If not otherwise stated *Escherichia coli* was grown at 37 °C in LB media (1% tryptone, 0.5% yeast extract, 1% NaCl) supplemented with appropriate antibiotics. All used reagents and chemicals used were of laboratory grade. Protease Inhibitor cocktail for purification of histidine-tagged proteins as a DMSO solution and *para*-nitrophenyl- α -L-rhamnopyranoside (pNP-rham) and were purchased from Sigma Aldrich. The flavonoids quercetin-3-O- β -D-glucopyranoside (isoquercitrin), naringenin-7-O- α -L-rhamnopyranosyl-(1,2)- β -D-glucopyranoside (naringin), naringenin-7-O- β -D-glucopyranoside (prunin), quercetin, and quercetin-3-O- α -L-rhamnopyranosyl-(1,6)- β -D-glucopyranoside (rutin) were kindly provided by Merck KGaA from the suppliers Roth, ABCR, Sigma Aldrich and AppliChem. Stock solutions of each flavonoid were prepared in DMSO.

2.2. Isolation of DNA and fosmid library construction

The dung sample for construction of the elephant feces library was derived from a healthy six year old female Asian elephant (*Elephas maximus*) named Kandy living in the Hagenbeck Zoo (Hamburg, Germany). Construction of the metagenomic library and storage was performed as described before (Rabausch et al., 2013). The elephant feces library encompassed a total of 20,000 clones. Average insert size of the fosmids was 35 kb.

2.3. Molecular cloning strategies

Amplification of the present open reading frames (ORFs) was performed with purified fosmid pUR16A2 DNA as a template for PCR. The reactions were performed in 35 cycles using *Pfu*-polymerase enzyme. To amplify the different ORFs *rhaB*, *rhaB1*, *rhaB2*, and *rhaB3* the primers *rhaB-Sall*-for, *rhaB1-Sall*-for, *rhaB2-Sall*-for, and *rhaB3-BamHI*-for, respectively, were used each in combination with *rha-HindIII*-rev (see Table S2). The purified PCR fragments were ligated into *Sma*I-prepared pUC19 vector and *E. coli* DH5 α was transformed by heat shock. Insert carrying clones were identified by blue white screening on LB agar plates containing 10 μ M 5-bromo-4-chloro-indolyl- β -D-galactopyranoside (X-gal) and 400 μ M isopropyl- β -D-thiogalactopyranoside (IPTG) after overnight growth. Correct clones were identified by plasmid purification, followed by RFLP analysis. The DNA sequences were verified by di-deoxy DNA sequencing. The resulting plasmids were designated pUC-B, pUC-B1, pUC-B2, and pUC-B3, respectively (Table S1).

Directed ligation of the respective ORF into expression vector pET21a (Merck, Darmstadt, Germany) was achieved applying the primer born endonuclease restriction sites, *Sall* or *BamHI* at the 5' end and the *HindIII*-site at the 3' end. Heat shock transformed *E. coli* DH5 α pET21-clones were analyzed for the correct insert by direct colony PCR using the respective-forward primers *rhaB-Sall*-for, *rhaB1-Sall*-for, *rhaB2-Sall*-for, or *rhaB3-BamHI*-for and T7 terminator primer. Additionally, the inserts of pET21*rhaB*, pET21*rhaB1*, pET21*rhaB2*, and pET21*rhaB3* were sequenced using T7 promoter and T7 terminator primers to certify the constructs.

2.4. Overproduction and purification of enzymes

For overproduction of the hexa-histidin (His₆-) tagged proteins RhaB, RhaB1, RhaB2, and RhaB3 *E. coli* BL21 (DE3) was transformed with the respective pET21*rha* construct. All four strains harboring, respectively, pET21*rhaB*, pET21*rhaB1*, pET21*rhaB2*, or pET21*rhaB3* were treated equally for enzyme production. An overnight preculture was harvested by centrifugation. The expression cultures were inoculated with 1% of resuspended cells and grown at 22 °C until 0.7 OD₆₀₀. The cultures were transferred to 17 °C and overexpression was induced by 400 μ M of IPTG. After 16 h, the cultures were harvested by centrifugation at 7500 \times g and 4 °C. The cells were resuspended in 50 mM phosphate buffer saline (PBS) with 0.3 M NaCl at pH 7.4 and disrupted by ultrasonication with a S2 sonotrode in a UP200S (Hielscher, Teltow, Germany) at a cycle of 0.5 and an amplitude of 75%. To avoid degradation protease inhibitor cocktail (Sigma Aldrich, Basel, Switzerland) was added.

Crude cell extracts were centrifuged at 15,000 \times g and 4 °C. The supernatants were loaded on 1 mL HisTrap FF Crude columns in an ÄKTAprime plus apparatus (GE Healthcare, Freiburg, Germany). The enzymes were purified according to the manufacturer protocol for gradient elution of His-tagged proteins. The fractions were analyzed on a denaturing 10% SDS-PAGE. Fractions of purified protein were pooled and dialyzed twice against 1000 vol. of 50 mM MES,

MOPS, or Tris–HCl buffers, respectively, with 0.3 M NaCl at 4 °C. The concentration of protein was determined by Bradford method using Roti-Quant (Carl Roth, Karlsruhe, Germany).

2.5. Biocatalytic assays

All biocatalytic reactions were performed in triplicate with Ni-affinity purified His₆-tagged enzyme. The physicochemical parameters were determined with pNP-rham as a substrate. If not otherwise stated biocatalytic reactions of 1.0 mL with 1 µg of purified enzyme were performed in 50 mM MES buffer at pH 6.5 with 1 mM of pNP-rham. The assays were pretemperated and incubated at 40 °C for 10 min. The reactions were stopped by addition of 0.5 mL of 1 M Na₂CO₃. After 3 min centrifugation the absorptions were measured at 405 nm. The temperature dependent activity of RhaB1 was determined between 10 °C and 60 °C with addition of 1 mM of CaCl₂ to the reactions. The pH dependency was analyzed from pH 3.7 to 5.5 in NaOAc, from pH 5.5 to 6.5 in MES, from pH 6.5 to 7.5 in MOPS, and from pH 7.2 to 9.0 in Tris–HCl each at 50 mM buffer concentration with addition of 1 mM of CaCl₂ to the reactions. And the effect of divalent metal ions and EDTA was measured at 1 mM concentration of CaCl₂, CdCl₂, CoCl₂, CuCl₂, FeSO₄, MgCl₂, MnCl₂, NiCl₂, ZnCl₂, and Na₂H₂EDTA added from 100 mM stock solutions.

The reaction kinetics on pNP-rham were determined for concentrations of 0.1, 0.2, 0.5, 1.0, 2.0 and 5.0 mM substrate over 10 min. The reaction rates were found to be linear from 15 to 60 s. An extinction coefficient for pNP at 405 nm of $\epsilon = 18.5 \text{ mM}^{-1}/\text{cm}$ was used (Birgisson et al., 2004; Fridjonsson et al., 1999).

The determination of the kinetic parameters of RhaB1 on flavonoid substrates were performed in 1.0 mL reactions of 50 mM MES buffer (pH6.5) at 40 °C. Rutin and naringin were supplemented from stock solutions with a final concentration of 2% (v/v) DMSO. After preheating to 40 °C 100 µg of purified RhaB1–His₆ was added to start the reaction. Samples of 100 µL were taken iteratively after 15, 30, 60, 120, and 240 min reaction. To stop the reaction and for quantification by TLC, the samples were dissolved 1/5 or 1/10 in 3:1 ethyl acetate/acetic acid. The acidified ethyl acetate samples were centrifuged at 10,000 g and the supernatant was used for quantitative TLC analysis as stated below.

2.6. TLC analyses and quantification of flavonoids

The diluted samples from flavonoid activity assays were transferred into HPLC flat bottom vials and used in TLC analysis for quantification of the reaction products. Samples of 10 or 20 µL were applied on 20 × 10 cm² HPTLC silica 60 F₂₅₄ plates (Merck KGaA, Darmstadt, Germany). For quantification of the reaction products from the biocatalytic reactions of rutin and naringin samples were spotted versus different amounts of the product reference substances, isoquercitrin and prunin, respectively. To avoid carryover all samples were spotted with double syringe rinsing in between by the ATS 4 (CAMAG, Muttenz, Switzerland). The sampled TLC plates were developed in ethyl acetate/acetic acid/formic acid/water 100:11:11:27 (Wagner et al., 1983). After separation the TLC plates were dried in an oven at 80 °C for 5 min. Thereafter, the flavonoids were derivatized by dipping the plates in a methanolic solution of 1% (w/v) diphenyl boric acid β -aminoethyl ester (DPBA) (Neu, 1957) for 1 s using a Chromatogram Immersion Device (CAMAG, Muttenz, Switzerland) followed by drying the TLC plates in hot air with a fan. After 2 min the TLC plates were further dipped in an ethanolic PEG4000 solution and dried in hot air fan for another minute. After 30 min at room temperature the fluorescence of the bands was analyzed densitometrically in a TLC Scanner 3 (CAMAG, Muttenz, Switzerland) with the D2 lamp at 350 nm for rutin and 330 nm for naringin reactions, respectively. The results

were analyzed with winCATS v. 1.4.4 software. Regression curves were calculated from the fluorescence peak areas and heights of the applied reference substances to determine the amount of produced flavonoids in the biocatalytic reactions.

2.7. Sequence analyses and GenBank entries

Automated DNA sequencing of small insert plasmids was performed using ABI377 and dye terminator chemistry following the manufacturer's instructions. The sequences were assembled by using Gap 4 software. Fosmid sequences were established by 454 sequencing technology and were assembled by Newbler GS-Assembler version 2.3.

ORF and promotor prediction in fosmid sequences was performed *in silico* with Softberry FGENESB and BPROM (<http://linux1.softberry.com/berry.phtml?topic=index&group=programs&subgroup=gfindb>). Alignments of DNA and protein sequences were done with BLAST and BLASTP (Altschul et al., 1990). Putative ribosome binding sites within ORF *rhaB* of fosmid pUR16A2 were compared to existing consensus sequences of *Bacteroides* sp. (Mastrolopolo et al., 2009; Tribble et al., 1999). The different ORFs *rhaB1*, *rhaB2* and *rhaB3* were processed by Clone Manager 9 Professional edition. All sequences mentioned in this work were deposited at GenBank. The DNA sequence of fosmid pUR16A2 has accession number JX188020, the α -L-rhamnosidase active enzyme RhaB1 has the entry number AGH13557.

Phylogenetic analysis of proteins was conducted using MEGA version 5 software (Tamura et al., 2011). The amino acid sequences and GenBank numbers were retrieved from NCBI. The sequences were aligned by ClustalW (Higgins et al., 1996) in a BLOSUM protein weight matrix. A neighbor joining (NJ) tree was calculated using 100 bootstraps in the Poisson model with a complete gap deletion mode.

3. Results

In the present study we report the identification and characterization of the novel α -L-rhamnosidase RhaB1 derived from a metagenomic clone by combination of sequence- and function-based analysis. The metagenomic DNA was isolated from the dung of a young healthy Asian elephant (*E. maximus*). The *rhaB* gene was identified on fosmid pUR16A2 by 454-sequencing bioactive fosmids from the metagenomic library of elephant feces, that had been identified in a recently published screening assay (Rabausch et al., 2013). Fosmid pUR16A2 carried a 38,761 bp insert (Fig. 1). The 4095 bp *rhaB* gene encoded for a single protein of 1364 amino acids.

A detailed analysis revealed two conserved regions on the hypothetical protein RhaB. The C-terminal region possessed homology to glycoside hydrolase family 78 domains of bacterial α -L-rhamnosidases (Bac_rhamnosid, PF05592). This homologous region ranged from Val945 to Leu1346 of RhaB (Fig. 1B). Further, a conserved N-terminal domain of bacterial α -L-rhamnosidases (Bac_rhamnosid_N, PF08531) was apparent between Ile710 and Pro850 of RhaB. However, the N-terminal part of hypothetical RhaB showed similarities to putative lipolytic proteins. The region from Leu64 to Asn230 presented the motif of GDSL-like lipase/acylhydrolase family (PFAM13472) proteins. Downstream, the region from Ala268 to Leu671 had similarity to putative acetyl-xylan esterases encompassing a dipeptidyl aminopeptidase/acylaminoacyl-peptidase (DAP2, COG1506) motif between Ile515 and Cys552 (Fig. 1B). Neither the GDSL motif nor the DAP2 motif previously have been reported in conjunction with enzymes of family GH78 which was underlined by a conserved domain architecture (CDART) search at BLAST (Geer et al., 2002).

Fosmid Ele16A2 (38,761 bp, GenBank accession no. JX188020)

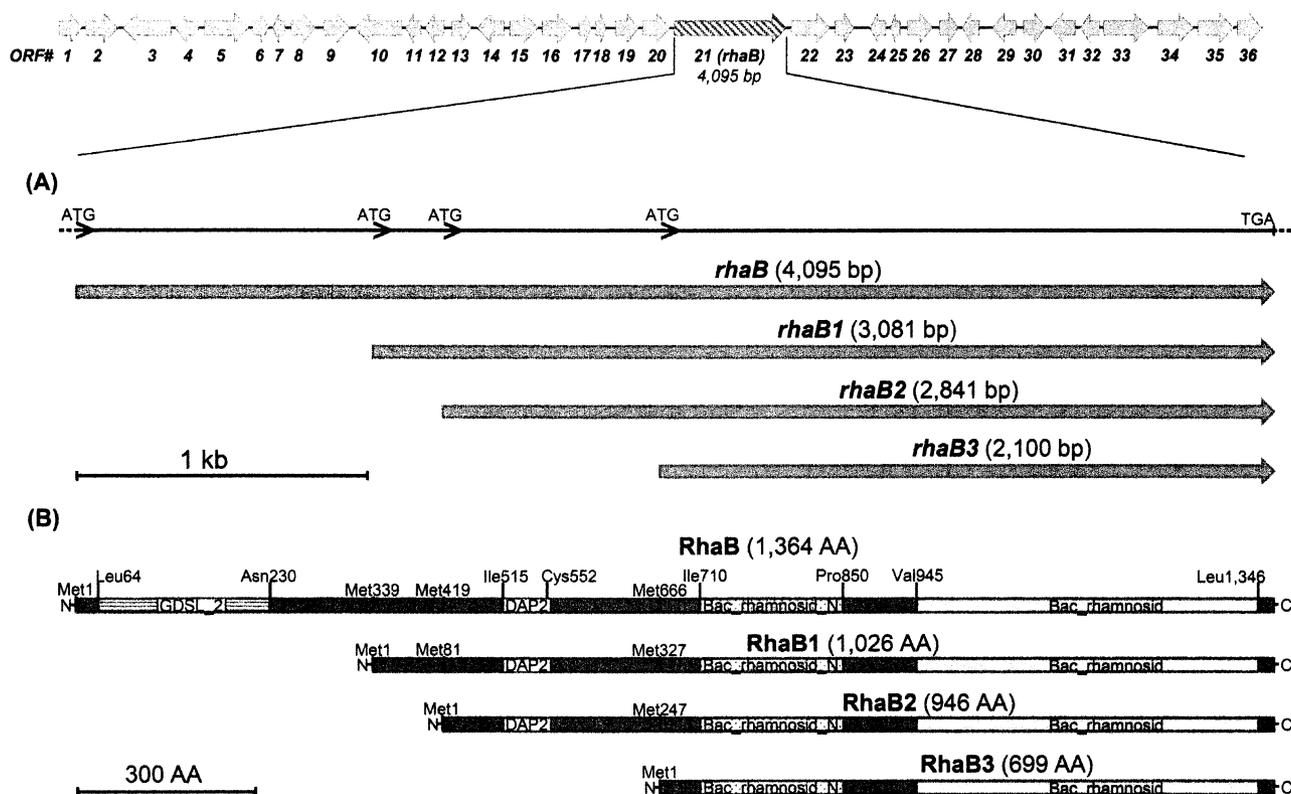


Fig. 1. Graphical illustration of fosmid pUR16A2 comprising the α -L-rhamnosidase encoding region. Annotated ORFs are numbered consecutively. (A) The different ORFs *rhaB*, *rhaB1*, *rhaB2*, and *rhaB3* are depicted as grey arrows, alternative translation start codons and the stop codon are signed above. (B) The deduced proteins RhaB, RhaB1, RhaB2, and RhaB3 are illustrated below as dark grey bars. The different methionine residues are marked as the start of the respective protein variants. Identified homologous motifs in the primary structure are signed as boxes indicating the first and last AA residue. The conserved N-terminal domain (Bac_rhamnosid_N, PFAM08531) and the conserved GH78-family motif of bacterial α -L-rhamnosidases (Bac_rhamnosid, PFAM05592) are present in all four protein variants. The dipeptidyl aminopeptidase/acylaminoacyl-peptidase (DAP2) motif (COG1506) is found in enzyme variants RhaB, RhaB1, and RhaB2, respectively, while the GDSL2 region similar to GDSL-like lipases/acylhydrolases (PFAM13472) is included only in RhaB.

Thus, the AA-analysis gave hint that RhaB might be the member of a novel subclass of type B α -L-rhamnosidases.

Phylogenetic analysis of the C-terminal α -L-rhamnosidase regions of RhaB by neighbor joining method in complete gap deletion mode supported the notion that RhaB belongs to a novel subclass of bacterial B type α -L-rhamnosidases (Fig. 2). RhaB clustered together with three predicted enzymes, all of *Bacteroidetes* origin and forming a new branch in the B type enzymes. Within these, the α -L-rhamnosidase core sequence (Met666 to Tyr1364) of RhaB shared an AA sequence identity of 58% to a predicted bacterial α -L-rhamnosidase from *Bacteroides nordii* (GenBank ZP10377667). Further relatives were the putative enzymes BACCAC02818 from *Bacteroides caccae* ATCC43185 (GenBank ZP01961192) and HMPREF9455_01164 from *Dysgonomonas gadei* ATCC BAA-286 (GenBank ZP08472998) with 57% and 56% sequence identity to the C-terminal region, respectively.

Attempts to express the complete *rhaB* gene in *E. coli* failed. At all, no α -L-rhamnosidase activity was detected in pNP-rham assays of crude cell extracts of clones pUC-B and pET21*rhaB*, respectively. In order to identify an active gene product, the variant ORFs *rhaB1*, *rhaB2*, and *rhaB3* were cloned (Fig. 1A). All probable ORFs showed putative ribosome binding sites (RBS) upstream (Mastropaolo et al., 2009). In database searches the closest *Bacteroidales* homologs were annotated to start approximately 50 AA upstream of the conserved Bac_rhamnosid_N region (Fig. 1B). Accordingly, an alternative start codon for an ORF of 2100 bp length was used. This ORF, designated *rhaB3*, encoded for a 699 AA protein. Simultaneously,

the two clones harboring the variant ORFs *rhaB1* of 3081 bp and *rhaB2* of 2841 bp were generated encoding for 1026 AA and 946 AA proteins, respectively (Fig. 1A). Further, *rhaB1* encoded for a protein in size accordance to the characterized α -L-rhamnosidase from *Bacteroides* sp. JY-6 (Jang and Kim, 1996).

With all constructs pET21*rhaB1*, pET21*rhaB2*, and pET21*rhaB3* heterologous protein expression was observed by SDS-PAGE analysis (Fig. S1). Enzyme activity assays of cell extracts from *E. coli* BL21(DE3) harboring pET21*rhaB3*, and of Ni-NTA purified RhaB3 His₆, respectively, revealed almost no α -L-rhamnosidase activity on pNP-rham of less than 1% compared to the activity of RhaB1. Finally, RhaB1 turned out to be the most active variant being approximately double as active as RhaB2 on pNP-rham.

Subsequently RhaB1 was purified and characterized in more detail. RhaB1 had a calculated molecular weight of 114.8 kDa which was in accordance to denaturing SDS-PAGE analysis of the recombinant protein RhaB1-His₆ (Fig. S1C). By Ni-affinity chromatography RhaB1-His₆ was natively purified to 75 mg/L of *E. coli* BL21(DE3) expression culture on an ÄKTAprime system (GE Healthcare, Freiburg, Germany).

Kinetic parameters of recombinant RhaB1 were determined for pNP-rham, and for the flavonoid substrates rutin and naringin, respectively (Table 1). For pNP-rha a Michaelis Menten kinetic was calculated from concentrations ranging from 0.1 mM up to 5 mM (Fig. 3A). RhaB1-His₆ had a determined K_m of 0.7933 mM, a specific activity v_{max} of 18.40 U, which corresponded to a k_{cat} value of 35.20 s⁻¹ (Table 1). As the solubility of flavonoids in aqueous

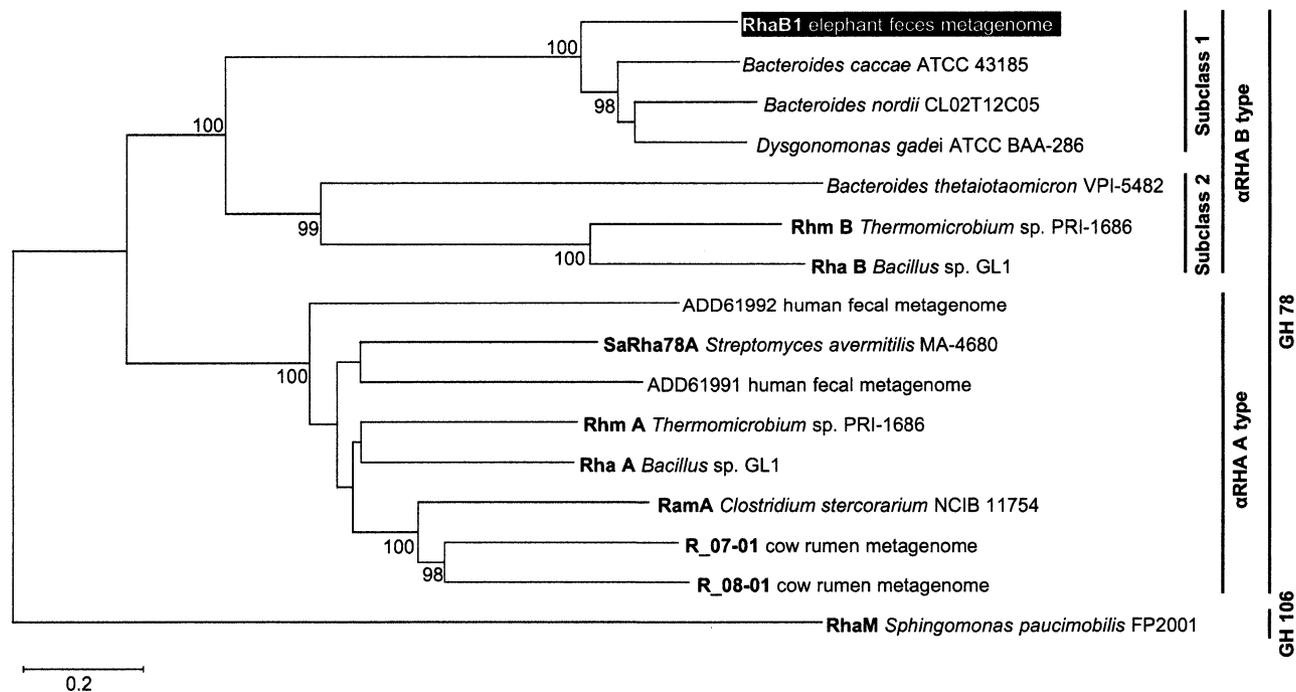


Fig. 2. Phylogenetic dendrogram of the core regions from bacterial α -L-rhamnosidases. The amino acid sequences of the closest sequence homologs of RhaB and of already characterized bacterial A and B type α -L-rhamnosidases were analyzed versus RhaM from *Sphingomonas paucimobilis* FP2001 (GenBank BAD12237) of GH family 106 as an outer group. Bootstrap values higher 95 are indicated next to the branches and the scale indicates amino acid changes per residue. AGH13557, RhaB1 from elephant feces metagenome fosmid pUR16A2; ZP01961192, hypothetical protein BACCAC02818 from *Bacteroides caccae* ATCC43185; ZP10377667, hypothetical protein HMPREF1068 03947 from *Bacteroides nordii* CL02T12C05; ZP08472998, hypothetical protein HMPREF9455 01164 from *Dysgonomonas gadei* ATCC BAA-286; AAO76108, putative α -L-rhamnosidase from *Bacteroides thetaiotaomicron* VPI-5482 (PDB: 3CIH); AAR96047, RhmB from *Thermomicrobium* sp. PRI-1686 (Birgisson et al., 2004); BAB62315, RhaB from *Bacillus* sp. GL1 (Hashimoto et al., 1999); ADD61992, putative GH78 enzyme from human fecal metagenome clone 9 (GenBank GU942952) (Tasse et al., 2010); BAC68538, SaRha78A from *Streptomyces avermitilis* MA-4680 (Ichinose et al., 2013); ADD61991, putative GH78 enzyme from human fecal metagenome clone 9 (GenBank GU942952) (Tasse et al., 2010); AAR96046, RhmA from *Thermomicrobium* sp. PRI-1686 (Birgisson et al., 2004); BAB62314, RhaA from *Bacillus* sp. GL1 (Hashimoto et al., 2003); CAB53341, RamA from *Clostridium stercorarium* NCIB 11754 (Zverlov et al., 2000); AFN57704, α -L-rhamnosidase R.07-01 from cow rumen metagenome clone r07 (Ferrer et al., 2012); AFN57707 α -L-rhamnosidase R.08-01 from cow rumen metagenome clone r08 (Ferrer et al., 2012).

systems is limited the flavonoid activity assays of RhaB1 were performed with concentrations from 0.5 mM up to 2.5 mM for rutin and up to 10 mM for naringin, respectively. Kinetic constants for naringin and rutin were determined from the nonlinear activity regression curves. Taking the solubility limitation of naringin and rutin into account, RhaB1-His₆ had a calculated K_m of 6.36 mM and a v_{max} of 2.9×10^{-3} U for naringin and a K_m of 6.75 mM and a specific activity v_{max} of 8.63×10^{-2} U for rutin, respectively (Table 1). All further experiments concerning the physico-chemical characteristics of RhaB1 were performed with pNP-rham substrate.

The activity of RhaB1-His₆ was determined at temperatures ranging from 10 °C to 60 °C. The recombinant enzyme had highest activity at 40 °C (Fig. 3B). However, at 50 °C just 25% relative activity thereof remained whereas RhaB1 still showed 89% relative activity at 45 °C. This limited capability at elevated temperatures was supported by stability measurements over 24 h. After 1 h preincubation at 50 °C RhaB1-His₆ retained only 2% residual activity compared to 82% at 40 °C after 1 h (data not shown). As a further parameter the pH dependent activity of recombinant RhaB1 was determined. Activity assays were performed from pH 3.7 in sodium acetate up to pH 9.0 in Tris-HCl buffers (Fig. 3C). The optimum pH for enzyme activity was found to be at pH 6.5 in MES and MOPS buffers. Here,

RhaB1-His₆ showed 6% higher activity in MES compared to MOPS. Overall, NaOAc buffers turned out to be the best tested buffer system concerning RhaB1-His₆ activity with a 12% higher activity in NaOAc compared to MES at pH 5.5.

The effect of metal ions and EDTA on RhaB1-His₆ activity was tested each in final concentrations of 1 mM (Fig. 4). Metal ions were used as dichloride salts with the exception of ferric trisulfate due to instability of Fe(II) in assay conditions. EDTA barely influenced on enzyme activity. Cadmium-, copper-, and zinc-ions diminished RhaB1-His₆ activity massively to a residual 12%, 3%, and 8% activity, respectively. Cobalt and nickel reduced enzyme activity significantly to residual 68%, respectively, 42%. Nearly no influence was observed with calcium ions whereas ferric, magnesium and manganese had positive effects with an approx. 20% increased activity in the assay conditions (Fig. 4).

4. Discussion

In the present study the novel α -L-rhamnosidase gene *rhaB* was identified in metagenomic DNA isolated from elephant feces. The ORF *rhaB1* of 3081 bp was successfully overexpressed in *E. coli* BL21(DE3) cells and the recombinant enzyme RhaB1 was purified in high amounts of 75 mg/L of culture.

Table 1
Enzyme kinetics of recombinant RhaB1 using three different substrates.

Substrate	K_m [mM]	v_{max} [μ .mol/min \times mg ⁻¹]	k_{cat} [s ⁻¹]
pNP- α -L-rhamnopyranoside	0.79 \pm 0.0481	18.40 \pm 0.3827	36.38 \pm 0.7568
Naringin	6.36	2.9×10^{-3}	5.7×10^{-3}
Rutin	6.75	8.63×10^{-2}	0.171

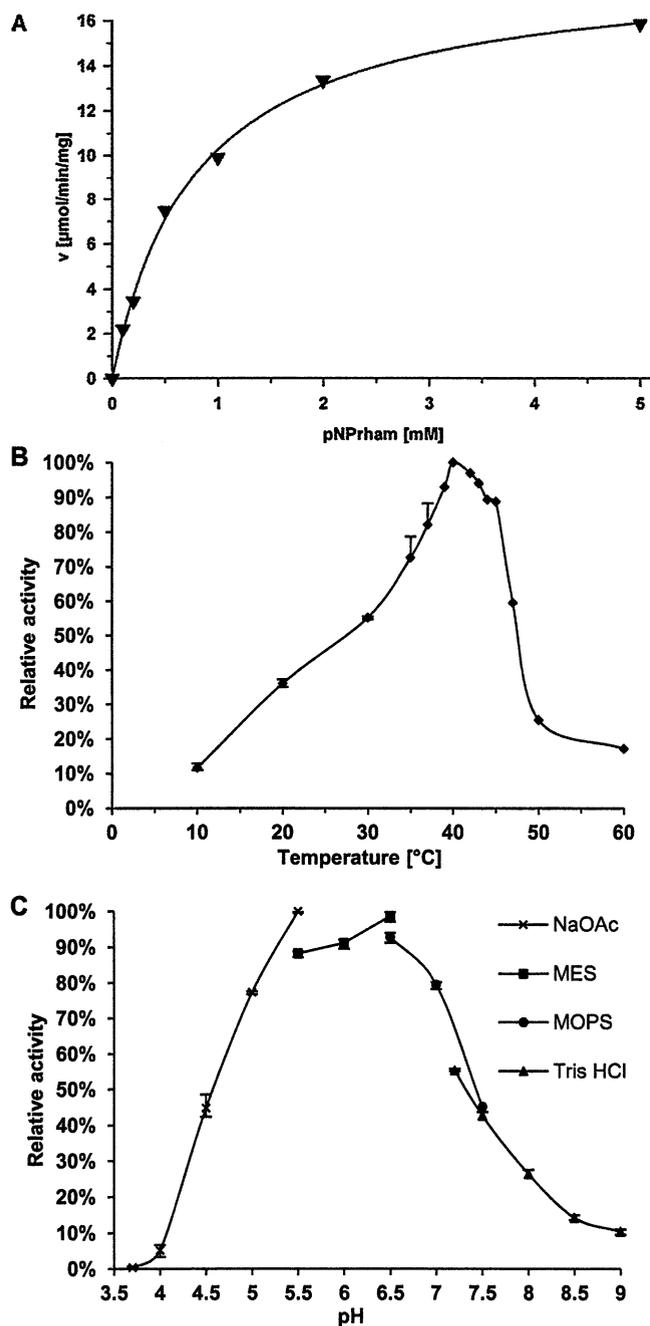


Fig. 3. Physico-chemical properties of recombinant RhaB1 enzyme on pNP-rham. (A) Shows the enzyme kinetics. The reaction rate (v [$\mu\text{mol}/\text{min}$]) is plotted as specific activity in U/mg of enzyme as a function of substrate concentration between 0.1 and 5.0 mM (\blacktriangledown). Linear conversion rates were determined from 15 to 60 s and used for reaction rate calculations. In (B) the relative activity of RhaB1 at different temperatures between 10 $^{\circ}\text{C}$ and 60 $^{\circ}\text{C}$ is depicted, activity at 40 $^{\circ}\text{C}$ was set to 100%. (C) Highlights the dependency of enzyme activity in different buffers from pH 3.7 to pH 9.0. All performed tests were done in triplicate.

Phylogenetically the C-terminal rhamnosidase domains of RhaB1 turned out to be somewhat related to putative α -L-rhamnosidases from *Bacteroidales* (Fig. 2). This is in good agreement to the fecal origin of the encoding gene. *Bacteroidetes* are well known affiliates in the gut microbiome of animals (Thomas et al., 2011). RhaB1 and other *Bacteroidales'* α -L-rhamnosidases branch off the cluster of bacterial α -L-rhamnosidase B type enzymes (Fig. 2). To our knowledge RhaB1 is the first biochemically characterized member of this cluster which we now designated subclass 1 of bacterial B type α -L-rhamnosidases. However, the AA-sequence

Metal ion	Relative activity	(%)
None		100
EDTA		104
Ca ²⁺		106
Cd ²⁺		12
Co ²⁺		68
Cu ²⁺		3
Fe ³⁺		126
Mg ²⁺		117
Mn ²⁺		123
Ni ²⁺		42
Zn ²⁺		8

Fig. 4. Effect of EDTA and metal cations on activity of recombinant RhaB1. Enzyme activity without any supplement (None) was set as a reference to 100%.

of a formerly characterized α -L-rhamnosidase with a similar molecular weight of about 120 kDa from *Bacteroides* sp. JY-6 was not deposited in GenBank and thus could not be phylogenetically evaluated (Jang and Kim, 1996). Other so far published bacterial B type α -L-rhamnosidases from *Bacillus* sp. GL1, *Thermomicrobium* sp. PRI-1686, and *B. thetaiotaomicron* VPI-5482 phylogenetically belong to the designated subclass 2 of the B type enzymes (Fig. 2).

The low homology level of RhaB1 to other annotated proteins might be explained by the fact that *Bacteroidetes* possess a great plasticity of their genomes and show frequent gene rearrangements, duplications, inversions or insertions of genes next to high degrees of lateral gene transfer (Xu et al., 2007). This genetic plasticity likely reflects also the organization of a GDSL-like lipase motif as well as a DAP2 domain of acetyl-xylan esterases in the N-terminal part in conjunction with the α -L-rhamnosidase domains in the C-terminal part of the enzyme. This domain architecture is not present in any other known enzyme of family GH78 and favors the speculation of a multi-enzymatic function on complex natural substrates. Unfortunately, the natural substrate of the enzyme remained unclear due to the metagenomic origin.

However, the recombinant enzyme RhaB1 exhibited a specific activity v_{max} of 18.40 U/mg on the synthetic substrate pNP-rham (Table 1). This is about one order of magnitude lower than the enzyme of *Bacteroides* sp. JY-6 (162.57 U/mg) (Jang and Kim, 1996). At the same time the K_m value of RhaB1-His₆ (0.79 mM) was determined to be three times higher than that of the wild type *Bacteroides* sp. JY-6 α -L-rhamnosidase (Fig. 3A). Further, RhaB1-His₆ showed low affinity for the flavonoid rhamnosides naringin and rutin. RhaB1-His₆ had a calculated K_m value of 6.36 mM for naringin and 6.75 mM for rutin, respectively (Table 1). The determined K_m for rutin would be even higher than the apparent solubility of rutin in aqueous environment. In the biocatalytic assays in 50 mM MES buffer the solubility of rutin was observed to be at least 2.5 mM whereas in pure water its solubility is reported to be only about 0.2 mM (Chebil et al., 2007). The low specific activities of 0.0029 U/mg and 0.0863 U/mg RhaB1-His₆ for naringin and rutin, respectively, underline the notion that flavonoids are not the favored natural substrates of RhaB1. In contrast, K_m values were found to be about four times lower for other α -L-rhamnosidases from *Bacteroides* sp. JY-6 and *Lactobacillus plantarum* NCC245, respectively (Avila et al., 2009; Jang and Kim, 1996). Beekwilder and colleagues even reported Ram1_{LP} from *Lactobacillus plantarum*

DSM20205 to have a K_m of 50 μM for rutin. This is about 100 times lower than RhaB1 and in the same range as determined for *Fusobacterium* sp. K-6 α -L-rhamnosidase and for RhaM of *Sphingomonas paucimobilis* FP2001, respectively (Beekwilder et al., 2009; Miake et al., 2000; Park et al., 2008).

Concerning the physico-chemical parameters of temperature and pH RhaB1-His₆ completely fitted to the mesophilic conditions of its origin (Fig. 3). The sensitivity of RhaB1-His₆ to the heavy metal cations Cd²⁺, Co²⁺, Cu²⁺, Ni²⁺, and Zn²⁺ was apparent (Fig. 4). *Bacteroides* sp. JY-6 α -L-rhamnosidase also was demonstrated to be sensitive to Co²⁺, Cu²⁺, and Zn²⁺ but with exception of Co²⁺ far less than RhaB1 (Jang and Kim, 1996). Moreover, the inhibiting effect of metal cations except for Cu²⁺ seems to vary in α -L-rhamnosidases. RhaB1 from *Bacillus* sp. GL1 is strongly affected by Cu²⁺, Fe²⁺, and Hg²⁺ but not Co²⁺ and Zn²⁺ (Hashimoto et al., 1999), *Fusobacterium* sp. K-6 α -L-rhamnosidase was inactivated by Cu²⁺ and Mn²⁺ while Zn²⁺ reduced activity less. The activity of RhaB1 and RhaB2 from *Lactobacillus plantarum* NCC245 was diminished by Cu²⁺, Fe²⁺, and Mn²⁺ while Co²⁺ as Ca²⁺ stimulated activity (Avila et al., 2009). In contrast, SaRha78A from *Streptomyces avermitilis* MA-4680 was reported to be not effected by any cation or EDTA. When tested, EDTA never had a negative effect on the enzyme activities of the α -L-rhamnosidases. This finding supports the hypothesis that α -L-rhamnosidases of GT family 78 do not require divalent metal ion cofactors for their catalytic activity (Cui et al., 2007).

Within this study the α -L-rhamnosidase RhaB1 originating from a gut metagenome of the Asian elephant was characterized. RhaB1 and related enzymes of family GH78 phylogenetically form a separate cluster of bacterial B type α -L-rhamnosidases, which to date only comprise enzymes from gut *Bacteroidales*. RhaB1 now is the first biochemically characterized enzyme of this subclass. Furthermore, RhaB presented an unexpected and novel enzyme domain architecture, comprising typical bacterial α -L-rhamnosidase domains co-organized with lipolytic GDSL-like lipase and acetyl-xylan esterase motifs. These findings again demonstrated that metagenomes are a rich source to increase our knowledge on enzyme biodiversity.

Acknowledgments

This work was funded by the German ministry of education and research "Bundesministerium für Bildung und Forschung" (BMBF) and was embedded in a project of the Biokatalyse2021 cluster.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jbiotec.2014.04.024>.

References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Avila, M., Jaquet, M., Moine, D., Requena, T., Pelaez, C., Arigoni, F., Jankovic, I., 2009. Physiological and biochemical characterization of the two α -L-rhamnosidases of *Lactobacillus plantarum* NCC245. *Microbiology* 155, 2739–2749.
- Bastien, G., Arnal, G., Bozonnet, S., Laguerre, S., Ferreira, F., Faure, R., Henrissat, B., Lefevre, F., Robe, P., Bouchez, O., Noirot, C., Dumon, C., O'Donohue, M., 2013. Mining for hemicellulases in the fungus-growing termite *Pseudacanthotermes militaris* using functional metagenomics. *Biotechnol. Biofuels* 6, 78.
- Beekwilder, J., Marozzi, D., Vecchi, S., de Vos, R., Janssen, P., Francke, C., van Hylckama Vlieg, J., Hall, R.D., 2009. Characterization of Rhamnosidases from *Lactobacillus plantarum* and *Lactobacillus acidophilus*. *Appl. Environ. Microbiol.* 75, 3447–3454.
- Birgisson, H., Hreggvidsson, G.O., Fridjónsson, O.H., Mort, A., Kristjánsson, J.K., Mattiasson, B., 2004. Two new thermostable α -L-rhamnosidases from a novel thermophilic bacterium. *Enzyme Microb. Technol.* 34, 561–571.
- Bokkenheuser, V.D., Shackleton, C.H., Winter, J., 1987. Hydrolysis of dietary flavonoid glycosides by strains of intestinal *Bacteroides* from humans. *Biochem. J.* 248, 953–956.
- Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B., 2009. The carbohydrate-active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* 37, D233–D238.
- Chebil, L., Humeau, C., Anthoni, J., Dehez, F., Engasser, J.M., Ghoul, M., 2007. Solubility of flavonoids in organic solvents. *J. Chem. Eng. Data* 52, 1552–1556.
- Cui, Z., Maruyama, Y., Mikami, B., Hashimoto, W., Murata, K., 2007. Crystal structure of glycoside hydrolase family 78 α -L-Rhamnosidase from *Bacillus* sp. GL1. *J. Mol. Biol.* 374, 384–398.
- Di Lazzaro, A., Morana, A., Schiraldi, C., Martino, A., Ponzzone, C., De Rosa, M., 2001. An enzymatic process for the production of the pharmacologically active glycoside desglucodesrhamnoscin from *Ruscus aculeatus* L. *J. Mol. Catal. B: Enzym.* 11, 307–314.
- Feng, B., Ma, B.-p., Kang, L.-p., Xiong, C.-q., Wang, S.-q., 2005. The microbiological transformation of steroidal saponins by *Curvularia lunata*. *Tetrahedron* 61, 11758–11763.
- Feng, Y., Duan, C.J., Liu, L., Tang, J.L., Feng, J.X., 2009. Properties of a metagenome-derived beta-glucosidase from the contents of rabbit cecum. *Biosci., Biotechnol., Biochem.* 73, 1470–1473.
- Ferrer, M., Ghazi, A., Beloqui, A., Vieites, J.M., Lopez-Cortes, N., Marin-Navarro, J., Nechitaylo, T.Y., Guazzaroni, M.E., Polaina, J., Waliczek, A., Chernikova, T.N., Reva, O.N., Golyshina, O.V., Golyshin, P.N., 2012. Functional metagenomics unveils a multifunctional glycosyl hydrolase from the family 43 catalysing the breakdown of plant polymers in the calf rumen. *PLoS One* 7, e38134.
- Fridjónsson, O., Watzlawick, H., Gehweiler, A., Rohrhirsch, T., Mattes, R., 1999. Cloning of the gene encoding a novel thermostable α -galactosidase from *Thermobrockianus* ITI360. *Appl. Environ. Microbiol.* 65, 3955–3963.
- Geer, L.Y., Domrachev, M., Lipman, D.J., Bryant, S.H., 2002. CDART: protein homology by domain architecture. *Genome Res.* 12, 1619–1623.
- Gloux, K., Berteau, O., El Oumami, H., Beguet, F., Leclerc, M., Dore, J., 2011. A metagenomic beta-glucuronidase uncovers a core adaptive function of the human intestinal microbiome. *Proc. Nat. Acad. Sci. U.S.A.* 108 (Suppl 1), 4539–4546.
- Hashimoto, W., Miyake, O., Nankai, H., Murata, K., 2003. Molecular identification of an alpha-L-rhamnosidase from *Bacillus* sp. strain GL1 as an enzyme involved in complete metabolism of gellan. *Arch. Biochem. Biophys.* 415, 235–244.
- Hashimoto, W., Nankai, H., Sato, N., Kawai, S., Murata, K., 1999. Characterization of α -L-rhamnosidase of *Bacillus* sp. GL1 responsible for the complete depolymerization of gellan. *Arch. Biochem. Biophys.* 368, 56–60.
- Hehemann, J.-H., Kelly, A.G., Pudlo, N.A., Martens, E.C., Boraston, A.B., 2012. Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proc. Nat. Acad. Sci. U.S.A.* 109, 19786–19791.
- Henrissat, B., Davies, G., 1997. Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* 7, 637–644.
- Higgins, D.G., Thompson, J.D., Gibson, T.J., 1996. Using CLUSTAL for multiple sequence alignments. *Methods Enzymol.* 266, 383–402.
- Hollman, P.C.H., Buijsman, M.N.C.P., van Gameren, Y., Cnossen, E.P.J., de Vries, J.H.M., Katan, M.B., 1999. The sugar moiety is a major determinant of the absorption of dietary flavonoid glycosides in man. *Free Radical Res.* 31, 569–573.
- Ichinose, H., Fujimoto, Z., Kaneko, S., 2013. Characterization of an α -L-Rhamnosidase from *Streptomyces avermitilis*. *Biosci. Biotechnol. Biochem.* 77, 213–216.
- Ilmberger, N., Meske, D., Juergensen, J., Schulte, M., Barthen, P., Rabausch, U., Angelow, A., Mientus, M., Liebl, W., Schmitz, R.A., Streit, W.R., 2012. Metagenomic cellulases highly tolerant towards the presence of ionic liquids—linking thermostability and halotolerance. *Appl. Microbiol. Biotechnol.* 95, 135–146.
- Jang, I.S., Kim, D.H., 1996. Purification and characterization of α -L-rhamnosidase from *Bacteroides* JY-6, a human intestinal bacterium. *Biol. Pharm. Bull.* 19, 1546–1549.
- Kaur, A., Singh, S., Singh, R.S., Schwarz, W.H., Puri, M., 2010. Hydrolysis of citrus peel naringin by recombinant α -L-rhamnosidase from *Clostridium stercorarium*. *J. Chem. Technol. Biotechnol.* 85, 1419–1422.
- Leonard, E., Yan, Y., Fowler, Z.L., Li, Z., Lim, C.G., Lim, K.H., Koffas, M.A., 2008. Strain improvement of recombinant *Escherichia coli* for efficient production of plant flavonoids. *Mol. Pharm.* 5, 257–265.
- Liu, T., Yu, H., Zhang, C., Lu, M., Piao, Y., Ohba, M., Tang, M., Yuan, X., Wei, S., Wang, K., Ma, A., Feng, X., Qin, S., Mukai, C., Tsuji, A., Jin, F., 2012. *Aspergillus niger* DLFC-90 rhamnoside hydrolase, a new type of flavonoid glycoside hydrolase. *Appl. Environ. Microbiol.* 78, 4752–4754.
- Manzanares, P., Orejas, M., Gil, J.V., De Graaff, L.H., Visser, J., Ramon, D., 2003. Construction of a genetically modified wine yeast strain expressing the *Aspergillus aculeatus* rhaA gene, encoding an α -L-rhamnosidase of enological interest. *Appl. Environ. Microbiol.* 69, 7558–7562.
- Manzanares, P., Vallés, S., Ramon, D., Orejas, M., 2007. α -L-rhamnosidases: old and new insights. In: Polaina, J., MacCabe, A. (Eds.), *Industrial Enzymes*. Springer, Netherlands, pp. 117–140.
- Martens-Uzunová, E.S., Zandleven, J.S., Benen, J.A.E., Awad, H., Kools, H.J., Beldman, G., Voragen, A.G.J., Van den berg, J.A., Schaap, P.J., 2006. A new group of exo-acting family 28 glycoside hydrolases of *Aspergillus niger* that are involved in pectin degradation. *Biochem. J.* 400, 43–52.
- Maslowski, K.M., Mackay, C.R., 2011. Diet, gut microbiota and immune responses. *Nat. Immunol.* 12, 5–9.
- Mastropaolo, M.D., Thorson, M.L., Stevens, A.M., 2009. Comparison of *Bacteroides thetaiotaomicron* and *Escherichia coli* 16S rRNA gene expression signals. *Microbiology* 155, 2683–2693.

- Miake, F., Satho, T., Takesue, H., Yanagida, F., Kashige, N., Watanabe, K., 2000. Purification and characterization of intracellular α -L-rhamnosidase from *Pseudomonas paucimobilis* FP2001. *Arch. Microbiol.* 173, 65–70.
- Michlmayr, H., Brandes, W., Eder, R., Schumann, C., del Hierro, A.M., Kulbe, K.D., 2011. Characterization of two distinct glycosyl hydrolase family 78 α -L-rhamnosidases from *Pediococcus acidilactici*. *Appl. Environ. Microbiol.* 77, 6524–6530.
- Miyata, T., Kashige, N., Satho, T., Yamaguchi, T., Aso, Y., Miake, F., 2005. Cloning, sequence analysis, and expression of the gene encoding *Sphingomonas paucimobilis* FP2001 α -L-rhamnosidase. *Curr. Microbiol.* 51, 105–109.
- Monti, D., Piscevcova, A., Kren, V., Lama, M., Riva, S., 2004. Generation of an α -L-rhamnosidase library and its application for the selective derhamnosylation of natural products. *Biotechnol. Bioeng.* 87, 763–771.
- Neu, R., 1957. Chelate von diarylborsäuren mit aliphatischen oxyalkylaminen als reagenzien für den nachweis von oxyphenyl-benzo- γ -pyronen. *Naturwissenschaften* 44, 181–182.
- Nicholson, J.K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., Pettersson, S., 2012. Host-gut microbiota metabolic interactions. *Science* 336, 1262–1267.
- Park, S.-Y., Kim, J.-H., Kim, D.-H., 2005. Purification and characterization of quercitrin-hydrolyzing α -L-rhamnosidase from *fusobacterium* K-60, a human intestinal bacterium. *J. Microbiol. Biotechnol.* 15, 519–524.
- Puri, M., 2012. Updates on naringinase: structural and biotechnological aspects. *Appl. Microbiol. Biotechnol.* 93, 49–60.
- Puri, M., Kaur, A., 2010. Molecular identification of *Staphylococcus xylosum* MAK2, a new α -L-rhamnosidase producer. *World J. Microbiol. Biotechnol.* 26, 963–968.
- Rabausch, U., Juergensen, J., Ilmberger, N., Bohnke, S., Fischer, S., Schubach, B., Schulte, M., Streit, W.R., 2013. Functional screening of metagenome and genome libraries for detection of novel flavonoid-modifying enzymes. *Appl. Environ. Microbiol.* 79, 4551–4563.
- Ribeiro, M.H., 2011. Naringinases: occurrence, characteristics, and applications. *Appl. Microbiol. Biotechnol.* 90, 1883–1895.
- Schütz, K., Muks, E., Carle, R., Schieber, A., 2006. Quantitative determination of phenolic compounds in artichoke-based dietary supplements and pharmaceuticals by high-performance liquid chromatography. *J. Agric. Food Chem.* 54, 8812–8817.
- Takatsu, T., Takahashi, S., Takamatsu, Y., Shioiri, T., Iwado, S., Haneishi, T., 1987. Chloropolysporins A, B and C, novel glycopeptide antibiotics from *Faenia interjecta* sp. Nov. IV. Partially deglycosylated derivatives. *J. Antibiot.* 40, 941–945.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739.
- Tasse, L., Bercovici, J., Pizzut-Serin, S., Robe, P., Tap, J., Klopp, C., Cantarel, B.L., Coutinho, P.M., Henrissat, B., Leclerc, M., Doré, J., Monsan, P., Remaud-Simeon, M., Potocki-Veronese, G., 2010. Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome Res.* 20, 1605–1612.
- Thibodeaux, C.J., Melancon, C.E., Liu, H.W., 2007. Unusual sugar biosynthesis and natural product glycodiversification. *Nature* 446, 1008–1016.
- Thomas, F., Hehemann, J.-H., Rebuffet, E., Czek, M., Michel, G., 2011. Environmental and gut *Bacteroidetes*: the food connection. *Front. Microbiol.* 2, Article 93.
- Tribble, G.D., Parker, A.C., Smith, C.J., 1999. Genetic structure and transcriptional analysis of a mobilizable, antibiotic resistance transposon from *Bacteroides*. *Plasmid* 42, 1–12.
- Ververidis, F., Trantas, E., Douglas, C., Vollmer, G., Kretzschmar, G., Panopoulos, N., 2007. Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part I: Chemical diversity, impacts on plant biology and human health. *Biotechnol. J.* 2, 1214–1234.
- Wagner, H., Bladt, S., Zgainski, E.M., 1983. Drogenanalyse, Dünnschichtchromatographische Analyse von Arzneidrogen. Springer-Verlag, Berlin, Heidelberg, New York, NY.
- Wang, M., Liang, C.-P., Wu, Q.-L., Simon James, E., Ho, C.-T., 2006. Instrumental analysis of popular botanical products in the U.S. market. In: *Herbs: Challenges in Chemistry and Biology*. ACS Symposium Series, vol. 925, pp. 25–38, Ch. 4.
- Warnecke, F., Luginbuhl, P., Ivanova, N., Ghassemian, M., Richardson, T.H., Stege, J.T., Cayouette, M., McHardy, A.C., Djordjevic, G., Aboushadi, N., Sorek, R., Tringe, S.G., Podar, M., Martin, H.G., Kunin, V., Dalevi, D., Madejska, J., Kirton, E., Platt, D., Szeto, E., Salamov, A., Barry, K., Mikhailova, N., Kyrpides, N.C., Matson, E.G., Ottesen, E.A., Zhang, X., Hernandez, M., Murillo, C., Acosta, L.G., Rigoutsos, I., Tamayo, G., Green, B.D., Chang, C., Rubin, E.M., Mathur, E.J., Robertson, D.E., Hugenholtz, P., Leadbetter, J.R., 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 450, 560–565.
- Xu, J., Mahowald, M.A., Ley, R.E., Lozupone, C.A., Hamady, M., Martens, E.C., Henrissat, B., Coutinho, P.M., Minx, P., Latreille, P., Cordon, H., Van Brunt, A., Kim, K., Fulton, R.S., Fulton, L.A., Clifton, S.W., Wilson, R.K., Knight, R.D., Gordon, J.I., 2007. Evolution of symbiotic bacteria in the distal human intestine. *PLoS Biol.* 5, e156.
- Yadav, V., Yadav, P.K., Yadav, S., Yadav, K.D.S., 2010. α -L-rhamnosidase: a review. *Process Biochem.* 45, 1226–1235.
- Young, N.M., Johnston, R.A.Z., Richards, J.C., 1989. Purification of the α -L-rhamnosidase of *Penicillium decumbens* and characterization of 2 glycopeptide components. *Carbohydr. Res.* 191, 53–62.
- Yu, H.S., Liu, H., Zhang, C.Z., Tan, D.H., Lu, M.C., Jin, F.X., 2004. Purification and characterization of gypenoside- α -L-rhamnosidase hydrolyzing gypenoside-5 into ginsenoside Rd. *Process Biochem.* 39, 861–867.
- Zverlov, V.V., Hertel, C., Bronnenmeier, K., Hroch, A., Kellermann, J., Schwarz, W.H., 2000. The thermostable α -L-rhamnosidase RamA of *Clostridium stercorearium*: biochemical characterization and primary structure of a bacterial α -L-rhamnosidase hydrolase, a new type of inverting glycoside hydrolase. *Mol. Microbiol.* 35, 173–179.

Functional Screening of Metagenome and Genome Libraries for Detection of Novel Flavonoid-Modifying Enzymes

U. Rabausch,^a J. Juergensen,^a N. Ilmberger,^a S. Böhnke,^a S. Fischer,^b B. Schubach,^b M. Schulte,^b W. R. Streit^a

Abteilung für Mikrobiologie und Biotechnologie, Biozentrum Klein Flottbek, Universität Hamburg, Hamburg, Germany^a; Merck KGaA, Darmstadt, Germany^b

The functional detection of novel enzymes other than hydrolases from metagenomes is limited since only a very few reliable screening procedures are available that allow the rapid screening of large clone libraries. For the discovery of flavonoid-modifying enzymes in genome and metagenome clone libraries, we have developed a new screening system based on high-performance thin-layer chromatography (HPTLC). This metagenome extract thin-layer chromatography analysis (META) allows the rapid detection of glycosyltransferase (GT) and also other flavonoid-modifying activities. The developed screening method is highly sensitive, and an amount of 4 ng of modified flavonoid molecules can be detected. This novel technology was validated against a control library of 1,920 fosmid clones generated from a single *Bacillus cereus* isolate and then used to analyze more than 38,000 clones derived from two different metagenomic preparations. Thereby we identified two novel UDP glycosyltransferase (UGT) genes. The metagenome-derived *gtfC* gene encoded a 52-kDa protein, and the deduced amino acid sequence was weakly similar to sequences of putative UGTs from *Fibrisoma* and *Dyadobacter*. GtfC mediated the transfer of different hexose moieties and exhibited high activities on flavones, flavonols, flavanones, and stilbenes and also accepted isoflavones and chalcones. From the control library we identified a novel macroside glycosyltransferase (MGT) with a calculated molecular mass of 46 kDa. The deduced amino acid sequence was highly similar to sequences of MGTs from *Bacillus thuringiensis*. Recombinant MgtB transferred the sugar residue from UDP-glucose effectively to flavones, flavonols, isoflavones, and flavanones. Moreover, MgtB exhibited high activity on larger flavonoid molecules such as tiliroside.

For more than a decade, metagenome research has demonstrated that it is a powerful tool for the discovery of novel biocatalysts and other valuable biomolecules by using either function- or sequence-based screening technologies (1–3). Sequence-based approaches allow the identification of candidate genes. In particular, the development of next-generation sequencing (NGS) technology and improved bioinformatic tools have significantly advanced this methodology (4). However, a major drawback of sequence-based screening technologies is that they do not allow direct conclusions about the functionality and biochemical parameters of the encoded enzymes. Furthermore, sequence-based searches are limited to the identification of homologs of already known motifs (5). Yet another problem associated with the sequence-based approach is that it often reveals only partial genes, which make subsequent expression and detailed biochemical analysis of the gene products difficult if not impossible. In contrast, the function-driven approach is usually much slower and more labor-intensive and costly but results in the detection of complete and active enzyme clones. It is of course well known that function-driven metagenomics is hampered due to the problems of expressing genes and incorrect processing of enzymes (6–8). However, the function-based approach allows the identification of truly novel enzymes, and it allows a first judgment on the actual enzyme activities and physicochemical parameters even during the screening process.

It is not surprising that the majority of metagenome-derived enzymes that have been characterized biochemically mainly originated from function-based screenings (9, 10). Interestingly, the majority of biocatalysts that have been identified through functional approaches are hydrolytic enzymes, mainly esterases and glycoside hydrolases (11, 12). This is perhaps linked to the simple plate-based screening procedures required for rapid detection of these hydrolytic enzymes (13). In this context it is noteworthy that

function-based screening technology is limited by the availability of sensitive and reliable assays for enzymes other than hydrolases that are of importance for biotechnology. Since the overall hit rates are usually low, metagenome screens often require high-throughput screening (HTS) technologies to be efficient, and the screenings need to be done under nearly production conditions (6). Further, enzyme screens often require complex substrates and sophisticated chromogenic assays as well as high-performance liquid chromatography (HPLC) or similar analytical methods. Clearly, the setup and development of novel function-driven metagenome screening assays are very tedious and time-consuming. This may be one reason why only a few function-based metagenome screening techniques have been developed during the last decade that focus on enzymes other than hydrolases and on those with relevance to biotechnological processes (14–20). Thus, there is an urgent need to develop function-based screening methods for genes and enzymes that belong to enzyme classes other than hydrolases and that are of relevance to biotechnology.

Flavonoids, as natural substances in fruits and vegetables, are part of our daily nutrition. They are well known for their antioxidative and radical scavenging nature and even more for having various beneficial effects on human health (21). Because of these broad effects, there is an increasing demand for specific flavonoids

Received 5 April 2013 Accepted 13 May 2013

Published ahead of print 17 May 2013

Address correspondence to W. R. Streit, wolfgang.streit@uni-hamburg.de.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AEM.01077-13>.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AEM.01077-13

in the cosmetic industry and the pharma- and nutraceutical industries (22–24). A major problem in meeting this demand arises from their limited availability. Flavonoids are exclusively produced in plants at low levels. The extraction is linked to the use of large quantities of solvents, and the chemical modification is not easily accomplished due to their rather complex structures (25).

The regio-specific modification of flavonoids remains difficult as the directed chemical modification mostly fails. Thus, flavonoid-modifying enzymes have gained interest as they can mediate the regio- and stereochemical modification of flavonoids (26). In particular, the specific glycosylation of flavonoids is the focus of research to influence water solubility and bioavailability of the polyphenolics (27, 28). Enzymes that catalyze this reaction are glycosyltransferases (GTs). Generally, GTs mediate the transfer of sugar residues from a donor substrate to acceptor molecules. Based on their sequence similarities, GTs are currently classified into 94 families (29). GT family 1 (GT1) comprises enzymes that catalyze the glycosylation of small lipophilic molecules (30). These enzymes (EC 2.4.1.x) that use a nucleotide-activated donor belong to the UDP-glycosyltransferase (UGT) superfamily and are also referred to as Leloir enzymes (31, 32). Glycosyltransferases acting on flavonoids also belong to GT1 (33). Enzymes of GT1 possess a GT-B fold structure and present an inverting reaction mechanism concerning the linkage of the transferred sugar moiety (34). Until now very few flavonoid-acting GT1s of prokaryotic origin have been identified and characterized in detail. The currently known flavonoid-accepting UGTs derived from Gram-positive bacteria all belong to the macrolide glycosyltransferase (MGT) subfamily and originate from bacilli and streptomycetes (35–37). Furthermore, a single flavonoid-acting UGT derived from the Gram-negative *Xanthomonas campestris* is known (38).

In the current publication we report on semiautomated thin-layer chromatography (TLC) screening of clone pools from metagenome libraries. The novel method allows the rapid identification of flavonoid-modifying enzyme clones. Using this technology, we have screened more than 40,000 fosmid clones and thereby identified two positive clones that showed significant flavonoid GT activities. The two novel enzymes, designated MgtB and GtfC, belong to GT family 1 and are highly active on flavonoids and similar molecules. While MgtB is highly similar to a hypothetical *Bacillus thuringiensis* MGT, GtfC is weakly similar to a hypothetical protein from *Fibrisoma limi*.

MATERIALS AND METHODS

Bacterial strains, plasmids, and chemical reagents. Bacterial strains and plasmids used in the present work are listed in Table S1 in the supplemental material, and primers are listed in Table S2. If not otherwise stated, *Escherichia coli* was grown at 37°C in LB medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl) supplemented with appropriate antibiotics. *Bacillus* isolates were grown at 30°C in the same medium. All chemical reagents used were of analytical-laboratory grade. Polyphenolic substances were purchased from the following companies located in Germany: Merck KGaA, Darmstadt; Carl Roth GmbH, Karlsruhe; Sigma-Aldrich, Heidelberg; and AppliChem GmbH, Darmstadt. Additional flavonoids were ordered from Extrasynthese (Lyon, France). Stock solutions of the polyphenols were prepared in dimethyl sulfoxide (DMSO) in concentrations of 100 mM.

Isolation of DNA and fosmid library construction. *Bacillus* sp. strain HH1500 was originally isolated from a soil sample of the botanical garden of the University of Hamburg. DNA from *Bacillus* sp. HH1500 was isolated using a peqGOLD Bacterial DNA Kit (PEQLAB Biotechnologie

GmbH, Erlangen, Germany) by following the manufacturer's protocol. The sample for the construction of the elephant feces library was derived from the Hagenbeck Zoo (Hamburg, Germany). Fresh feces of a healthy 6-year-old female Asian elephant (*Elephas maximus*) named Kandy were taken and stored at –20°C in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8) containing 30% (vol/vol) glycerol until DNA extraction. For DNA extraction a QIAamp DNA Stool Mini Kit (Qiagen, Hilden, Germany) was used. The kit was applied according to the manufacturer's protocol. As recommended, we increased the incubation temperature in ASL buffer (Qiagen) to 95°C. Isolation of DNA from Elbe river sediment was performed with sediment samples from the tidal flat zone of the river Elbe near Glückstadt (Germany) at low tide (53°44'40"N, 9°26'14"E). Environmental DNA was extracted using an SDS-based DNA extraction method published by Zhou and coworkers (39).

Construction of the genomic and metagenomic libraries in *E. coli* EPI300 cells harboring fosmid pCC1FOS was achieved with a CopyControl Fosmid Library Production Kit (Epicentre Biotechnologies, Madison, WI) according to the manufacturer's protocol using minor modifications as previously published (40). Clones were transferred into 96-well microtiter plates containing 150 µl of liquid LB medium with 12.5 µg/ml of chloramphenicol and allowed to grow overnight. Libraries were stored at –70°C after the addition of 100 µl of 86% glycerol to each microtiter well. The genomic fosmid library of *Bacillus* sp. HH1500 comprised 1,920 clones; a total of 35,000 clones were obtained for the river Elbe sediment library, and the elephant feces library encompassed a total of 20,000 clones. All libraries contained fosmids with average insert sizes of 35 kb.

Molecular cloning strategies. Fragments of pCC1FOS fosmids were subcloned into pBluescript II SK+ vector using HindIII according to the restriction of the fosmid clones pFOS4B2 and pFOS144C11. The resulting plasmids were designated pSK4B2 and pSK144C11, respectively. Further subcloning of pSK144C11-derived fragments was achieved in pTZ19R-Cm with restriction enzymes EcoRI and PstI. The obtained clones were designated pTZ144E and pTZ144P, respectively. *E. coli* DH5α was transformed with the plasmids by heat shock, and the plasmids carrying subclones were identified by blue-white screening on LB agar plates containing 10 µM 5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-Gal) and 400 µM isopropyl-β-D-thiogalactopyranoside (IPTG) after overnight growth. Different clones were analyzed by plasmid purification, followed by enzymatic digestion and agarose gel electrophoresis and/or DNA sequencing.

PCR amplification of open reading frames (ORFs) was performed with fosmid DNA as a template. The reactions were performed in 30 cycles. To amplify *mgtB*, the primers mgt1-XhoI-for and mgt1-XhoI-rev were used, inserting XhoI endonuclease restriction sites 5' and 3' of the ORF (see Table S2 in the supplemental material). For cloning of *gtfC*, the primer pair gtf-Nde-for and gtf-Bam-rev was used, inserting an NdeI site, including the start codon, 5' of the ORF and a BamHI site 3' of the ORF (see Table S2). PCR fragments were ligated into pDrive using a Qiagen PCR Cloning Kit (Qiagen, Hilden, Germany) and cloned into *E. coli* DH5α. Resulting clones, designated pDmgtB and pDgtfC, respectively, were analyzed for activity in biotransformation and by DNA sequencing for the correct insert. Ligation of *mgtB* and *gtfC* into expression vector pET19b (Merck KGaA, Darmstadt, Germany) was achieved using the inserted endonuclease restriction sites of each ORF. Plasmids containing the correct insert were designated pET19mgtB and pET19gtfC, respectively. *E. coli* DH5α clones harboring the desired plasmids were detected by direct colony PCR using a T7 terminator primer and mgt1-XhoI-for to confirm insertion of *mgtB*, and the T7 terminator primer and gtf-Nde-for were used to verify *gtfC*. Additionally, the inserts of pET19mgtB and pET19gtfC were sequenced using T7 promoter and T7 terminator primers (see Table S2) to verify the constructs.

Overproduction and purification of enzymes. For overproduction of decahistidine (His₁₀)-tagged proteins, *E. coli* BL21(DE3) was transformed with pET19b constructs. An overnight preculture was harvested by centrifugation, and 1% was used to inoculate an expression culture. Cells

carrying pET19*mgtB* were grown at 22°C until an optical density at 600 nm (OD_{600}) of 0.7. The culture was transferred to 17°C and induced by 100 μ M IPTG. After 16 h, the culture was harvested by centrifugation at $7,500 \times g$ at 4°C. Cells were resuspended in 50 mM phosphate-buffered saline (PBS) with 0.3 M NaCl at pH 7.4 and disrupted by ultrasonication with an S2 Sonotrode in a UP200S instrument (Hielscher, Teltow, Germany) at a cycle of 0.5 and an amplitude of 75%.

The overproduction of decahistidine-tagged GtfC was induced at 37°C at an OD_{600} of 0.6 with 100 μ M IPTG. Cells were then incubated for 4 h, harvested, and lysed as stated above for MgtB.

Crude cell extracts were centrifuged at $15,000 \times g$ and 4°C to sediment the cell debris. The clarified extracts were loaded on 1-ml HisTrap FF Crude columns using an ÄKTAprime Plus system (GE Healthcare). The enzymes were purified according to the manufacturer's protocol for gradient elution of His-tagged proteins. Eluted protein solutions were dialyzed twice against 1,000 volumes of 50 mM PBS, pH 7.4, with 0.3 M NaCl at 4°C. The purification was analyzed by 12% SDS-PAGE. The concentration of protein was determined by the Bradford method using Roti-Quant (Carl Roth GmbH, Karlsruhe, Germany).

Biotransformations and biocatalyses. For the detection of flavonoid modifications in bacteria, we used a biotransformation approach. Cultures were grown in LB medium with appropriate antibiotics overnight. Expression cultures were prepared as stated above for overproduction of enzymes. The cells were sedimented by centrifugation at $4,500 \times g$ and resuspended in 50 mM sodium phosphate buffer, pH 7, supplemented with 1% (wt/vol) α -D-glucose. Biotransformations with a final concentration of 100 μ M flavonoid inoculated from stock solutions of 100 mM in DMSO (i.e., 0.1%) were incubated in Erlenmeyer flasks at 30°C and 175 rpm up to 24 h. Samples of 4 ml were withdrawn and acidified with 100 μ l of 1 M H_3PO_4 aqueous (aq) for extraction in 2 ml of ethyl acetate (EtOAc). They were shaken for 1 min and phase separated by centrifugation at $2,000 \times g$ and 4°C. The supernatant was applied in TLC analysis. For quantification, samples of 100 μ l were taken and dissolved 1/10 in ethyl acetate-acetic acid (3:1). These acidified ethyl acetate samples were centrifuged at $10,000 \times g$. The supernatant was used for quantitative TLC analysis as stated below.

Fosmid clones were grown in 96-deep-well plates overnight. Clones were joined in 96, 48, 8, or 6 clones per pool. The pools were harvested by centrifugation at $4,500 \times g$ and resuspended in 50 ml of LB medium containing 12.5 μ g/ml chloramphenicol (CopyControl Autoinduction Solution; Epicentre, Madison, WI) (5 mM arabinose final concentration) and 100 μ M flavonoid for biotransformation. As an alternative to deep-well plates, clones were precultured on agar plates. After overnight incubation, the colonies were washed off with 50 mM sodium phosphate buffer, pH 7, harvested by centrifugation, and resuspended as outlined above. The biotransformations were incubated in 300-ml Erlenmeyer flasks at 30°C with shaking at 175 rpm. Single clones were tested analogously but precultured in 5 ml of LB medium and resuspended in 20 ml of biotransformation medium in 100-ml flasks. Samples of 4 ml were taken from the reaction products after 16, 24, and 48 h, acidified with 40 μ l HCl aq, and prepared for TLC analysis as stated above. Positive pools were verified in a second biotransformation and then systematically downsized to detect the corresponding hit in a smaller pool until the responsible single clone was identified.

Biocatalytic reaction mixtures of 1 ml contained 5 μ g of purified His-tagged enzyme, and reactions were performed in 50 mM sodium phosphate buffer, pH 7, at 37°C. UDP- α -D-glucose or UDP- α -D-galactose was added to a final concentration of 500 μ M as a donor substrate from 50 mM stock solutions in 50 mM sodium phosphate buffer, pH 7. Acceptor substrates were used in concentrations of 100 μ M and were added from stock solutions of 100 mM in DMSO, leading to a final content of 0.1% in the reaction mixture. The reaction was stopped by dissolving 100 μ l of reaction mixture 1/10 in ethyl acetate-acetic acid (3:1). These samples were used directly for quantitative TLC analysis.

TLC analyses. The extracts transferred into HPLC flat-bottom vials were used for TLC analysis. Samples of 20 μ l were applied on 20- by 10-cm² high-performance thin-layer chromatograph (HPTLC) silica gel 60 F₂₅₄ plates (Merck KGaA, Darmstadt, Germany) versus 200 pmol of reference flavonoids. To avoid carryover of substances, i.e., to prevent false positives, samples were spotted with double syringe rinses in between by an Automatic TLC Sampler 4 instrument (ATS 4; Camag, Muttenz, Switzerland). The sampled TLC plates were developed in ethyl acetate-acetic acid-formic acid-water (100:11:11:27) (Universal Pflanzenlaufmittel, or universal plant solvent) (41). After band separation, the TLC plates were dried in an oven at 80°C for 5 min. The absorbance of the separated bands was determined densitometrically depending on the absorbance maximum of the applied educts at 285 to 370 nm using a deuterium lamp in a TLC Scanner 3 (Camag, Muttenz, Switzerland). Subsequently, the substances on developed TLC plates were stained by either dipping or spraying the plates in a 1% (wt/vol) methanolic solution of Naturstoff reagent A, containing diphenyl boric acid β -aminoethyl ester (42), available from Carl Roth GmbH, Karlsruhe, Germany. After immediate drying with a hot air fan, the TLC plates were dipped in or sprayed with a 5% (wt/vol) solution of polyethylene glycol 4000 in ethanol (70%, vol/vol). For dipping, a chromatogram immersion device (Camag, Muttenz, Switzerland) was used. After complete drying the bands were visualized at 365 nm with a UV hand lamp and photographed. Alternatively, fluorescence of the bands was determined densitometrically by the TLC Scanner 3 depending on the absorbance maxima of the applied substances at 320 to 370 nm.

Quantification of flavonoids by TLC. To quantify flavonoids in biotransformation and biocatalytic reactions, samples were diluted 1/10 in ethyl acetate-acetic acid (3:1) and subsequently centrifuged. Samples of 20 μ l were sprayed by an ATS 4 (Camag, Muttenz, Switzerland) on HPTLC silica gel 60 F₂₅₄ plates (Merck KGaA, Darmstadt, Germany) versus different amounts of respective standard educt and product substances. TLC plates were developed, dried, derivatized, and analyzed as stated above. Regression curves were calculated from the peak area of the applied reference substances to determine the amounts of produced and residual flavonoids.

HPLC-ESI-MS analysis. HPLC was carried out on a Purospher Star RP-18e 125-4 column (particle size of 3 μ m; Merck, Darmstadt, Germany) with a Rheos 2000 pump (Flux Instruments, Suisse) and set pressure limits, with a minimum of 0 Pa and a maximum of 400×10^5 Pa. Injection volumes of 10 μ l were separated with solvent A (water supplemented with 0.1% trifluoroacetic acid [TFA]) and solvent B (acetonitrile with 0.1% TFA) under the following HPLC gradient conditions: from 0 min, 0.6 ml/min of 90% A and 10% B; from 14 min, 0.6 ml/min of 75% A and 25% B; from 18 min, 0.6 ml/min of 5% A and B = 95%; from 22 min, 0.6 ml/min of 5% A and 95% B; from 22.1 min, 0.6 ml/min of 90% A and 10% B; and from 28.1 min, 0.6 ml/min of 90% A and 10% B. Elution was monitored with a Finnigan Surveyor photodiode array (PDA) detector, and fractions were collected by an HTC PAL autosampler (CTC Analytics). Mass spectrometry (MS) was performed on a Thermo LCQ Deca XP Plus with an electrospray ionization (ESI) interface in positive ionization.

Sequence analysis. Automated DNA sequencing of small inserted plasmids was performed using an ABI377 instrument and dye terminator chemistry according to the manufacturer's instructions. Large fosmid sequences were established by 454 sequencing technology. The sequences were assembled by using Gap, version 4, software. ORF finding was performed with Clone Manager Professional, version 9, software.

Nucleotide sequence accession numbers. All sequences mentioned in this work were deposited in GenBank. The DNA sequence of the *Bacillus* sp. HH1500 16S rRNA gene was deposited in GenBank under accession number [KC145729](#). The fosmid-derived genes from *Bacillus* sp. HH1500 identified on subclone pSK4B2 are *bspA* ([JX157885](#)), *mgtB* ([JX157886](#)), and *bspC* ([JX157887](#)), and their sequences have been deposited under accession numbers [AGH18135](#) and [AGH18137](#), respectively. The Elbe sediment metagenome-derived fosmid subclone pSK144C11 comprised

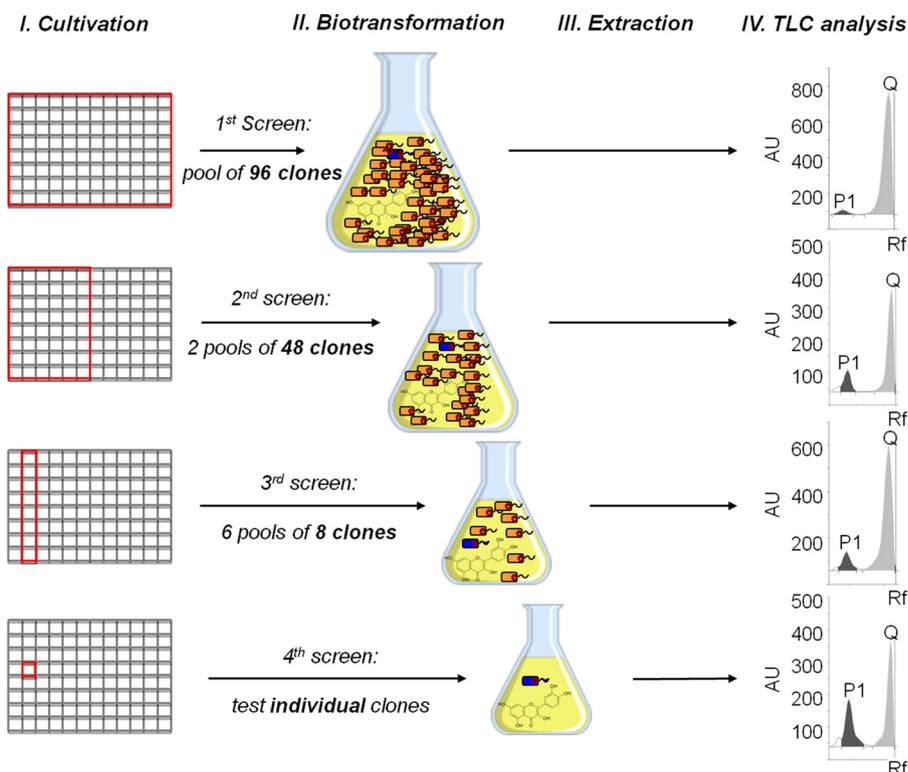


FIG 1 Outline of the metagenome screening for flavonoid-modifying clones. A schematic workflow of the function-based screening procedure for the systematic identification of flavonoid-modifying clones is shown. Initially, pools of 96 clones preincubated in deep-well plates were tested in biotransformation reactions with quercetin, followed by culture extract TLC analysis as described in Materials and Methods. Once a positive pool forming a product peak was observed, the respective plate was divided in two half-plate pools of 48 clones each, which were tested similarly. The positive fraction was divided into six rows of eight clones per column (or, vice versa, 8 by 6 per row) that were tested for activity until candidate single clones were tested and the positive clones were identified. Putative positive clones were then reconfirmed and subcloned for further analysis. UV chromatograms from TLC analysis are displayed in relative absorbance units (AU) versus the R_f -value measured at 365 nm in a densitometric TLC Scanner 3 (Camag, Muttens, Switzerland). Light gray peaks indicate absorbance of the remaining quercetin (Q) substrate, dark gray peaks (P1) indicate the formed product isoquercitrin. Depicted chromatograms show UV chromatograms of culture extracts in EtOAc after 24-h biotransformations of pools with 96, 48, and 8 clones including fosmid clone pFOS19G2 and of the single clone pFOS19G2 (last).

genes *esmA* (JX157626), *gtfC* (JX157627), *esmB* (JX157628), and *esmC* (JX157629). The sequences of the deduced proteins have been deposited under GenBank accession numbers AGH18138 to AGH18141.

RESULTS

Screening method: setup of a TLC-based screening method for the detection of flavonoid-modifying enzyme clones. Since it is known that *Bacillus cereus* and *Bacillus subtilis* encode glycosyltransferases mediating the glucosylation of flavonoids (36), we initially tested several single bacterial isolates from our strain collections with respect to their flavonoid-modifying activities. Biotransformations using whole cells of wild-type isolates confirmed the presence of flavonoid-modifying enzymes in one of the strains. This strain was originally isolated from a soil sample of the botanical garden in Hamburg, Germany, and was designated *Bacillus* sp. HH1500. Sequence analysis of a 16S rRNA gene (GenBank entry KC145729) showed 100% identity to members of the *B. cereus* group (data not shown). In order to use this strain as a positive control, we constructed a fosmid library of its genomic DNA in pCC1FOS. The obtained library contained 1,920 clones with an average insert size of 35 kb. Thus, the library encompassed approximately 67 Mb of cloned genomic DNA (gDNA), or about 10 times the average size of a genome from *B. cereus* group members (43). Further, the sensitivity of the (HP)TLC-based assay was ver-

ified using a serial dilution of isoquercitrin, the 3-O- β -D-glucoside of quercetin, by spraying 10 μ l of a solution of 0.78 μ M up to 100 μ M isoquercitrin on TLC plates and measuring the absorbance at 365 nm (see Table S3 in the supplemental material). In addition, 10- μ l samples of other glycosylated flavonoids were assayed at 10 μ M concentrations and could be detected as clear peaks on the absorbance chromatograms (see Table S3; also data not shown).

Based on the observed sensitivities, we designed a systematic screening scheme, as outlined in Figure 1. Initially 96 fosmid clones were grown in deep-well microtiter plates at 37°C overnight. Cultures were then pooled, and, following this step, the cells were sedimented by centrifugation and resuspended in fresh LB medium containing the appropriate antibiotics and 100 μ M quercetin as an acceptor substrate. After incubation for 16, 24, and 48 h at 30°C, 4-ml samples of the pooled cultures were withdrawn and extracted with half the volume of ethyl acetate. Of these extracts 20 μ l was applied on TLC silica plates and separated using Universal Pflanzenlaufmittel as a solvent. The absorbance of the developed sample lanes was determined densitometrically at 365 nm. Additionally, bands of substrates and modified flavonoids were visualized by staining with Naturstoff reagent A as outlined in the Materials and Methods section. In our hands the sensitivity

TABLE 1 ORFs identified on subclones pSK4B2, derived from the active *Bacillus* sp. HH1500 fosmid clone, and pSK144C11, derived from the river Elbe sediment active fosmid clone

Subclone and ORF	Position (aa)	Homolog (accession no.)	Coverage (%)	% identity	% similarity
pSK4B2					
<i>bspA</i>	221	Putative protein kinase from <i>Bacillus thuringiensis</i> (EEM66464)	100	99	99
<i>mgtB</i>	402	Macrolide glycosyltransferase from <i>Bacillus thuringiensis</i> (EEM96628)	100	98	99
<i>mgtC</i>	261	Hypothetical membrane protein from <i>Bacillus thuringiensis</i> (EAO54527)	100	99	100
pSK144C11					
<i>esmA</i>	80	Putative UDP-N-acetylmuramate-L-alanine ligase from <i>Niabella soli</i> (EHP51575)	99	69	80
<i>gtfC</i>	459	Putative UDP-glucosyltransferase from <i>Fibrisoma limi</i> (CCH52088)	92	51	71
<i>esmB</i>	170	Hypothetical protein from <i>Niastella koreensis</i> (YP005009630)	95	63	77
<i>esmC</i>	150	Putative membrane protein from <i>Solitalea canadensis</i> (YP006258217)	98	68	81

of the assay was high enough to detect a single flavonoid-modifying enzyme clone in a mixture of 96 clones (Fig. 1). After the detection of a positive signal, we divided the 96 fosmid clones into pools of 48 to locate the same peak in one of the resulting two microtiter half-plates. Following this procedure, we divided the 48 clones into six groups of eight clones (Fig. 1) and finally analyzed the eight individual clones. This strategy was applied successfully to identify six overlapping positive clones in the *Bacillus* sp. HH1500 fosmid library testing all 20 microtiter plates with a total of 1,920 clones.

Of these six fosmid clones, one clone, pFOS4B2, of approximately 46 kb was subcloned using the HindIII restriction site of the pBluescript II SK+ vector. The obtained subclones were analyzed using the above-mentioned TLC screening technology. Thereby, a positive subclone designated pSK4B2 was identified and completely sequenced (GenBank entries JX157885 to JX157887). Subclone pSK4B2 carried an insert of 3,225 bp (see Fig. S1A in the supplemental material) and harbored a gene, designated *mgtB*, encoding a protein of 402 amino acids (aa). The identified ORF was subcloned, creating plasmid pD*mgtB*, and again assayed for activity. TLC analysis clearly confirmed the glycosylation activity of the MgtB enzyme in this construct as well. The deduced amino acid sequence of MgtB (GenBank accession number AGH18136) was highly similar to a predicted *B. thuringiensis* macroside glycosyltransferase (Table 1). The *mgtB*-surrounding DNA sequences in plasmid pSK4B2 represented two truncated genes that consistently were almost identical to genes from *B. thuringiensis* (Table 1). This phylogenetic relation was in accordance to the preliminary sequence analysis of the 16S rRNA gene of *Bacillus* sp. HH1500 (see above).

These tests suggested that the screening procedure was suitable for the functional screening of large-insert metagenome libraries. For the function-based screening of metagenomes, we termed this methodology META, for metagenome extract TLC analysis. Although it is not a fully automated high-throughput screening (HTS) technology, META allows screening of about 1,200 clones per TLC plate within a time of 48 h for preculture, biotransformation, and analysis. This number of clones appeared to be feasible if a single person did the screening. Generally, the sampling of about one TLC plate per hour by the ATS 4 is the time-limiting step of the method. But this still allows the pooled screening of several plates a day and, hence, throughput of several thousand clones a day by META.

Identification of a novel glycosyltransferase from a metagenome library. To further apply the screening for enzyme dis-

covery in metagenome libraries, we tested two fosmid libraries constructed in our laboratory. One library was constructed from DNA isolated from river Elbe sediment; the other was from DNA isolated from fresh elephant feces. Altogether both libraries encompassed approximately 50,000 clones with an average insert size of 35 kb. Both libraries were screened using quercetin as a substrate. Using the described strategy, we discovered one positive microtiter plate pool in the river Elbe sediment library. Further screening of this pool resulted in the identification of a single positive fosmid clone, designated pFOS144C11 (Fig. 2). Biotransformations of quercetin (Fig. 2A, Q) with 48 clone pools presented one product peak (P2) by TLC separation with an R_f value comparable to that of quercitrin, the quercetin-3-O- β -L-rhamnoside (Fig. 2A). A second peak (P3) with an R_f value higher than values for the available reference quercetin glycones was observed in conversions with the six-clone pool and the single fosmid clone (Fig. 2B and C, respectively). Clone pFOS144C11 carried a fosmid of approximately 40 kb. Subsequent restriction fragment subcloning into pBluescript II SK+ with HindIII yielded the identification of the positive *E. coli* DH5 α subclone pSK144C11. However, biotransformations with pSK144C11 showed two product peaks, a major one (Fig. 2D, P2) with an R_f value comparable to that of quercitrin and a minor one (P1) similar to isoquercitrin (Fig. 2D). The subclone pSK144C11 still had an insert of approximately 8.5 kb in size. Further sequencing and subcloning of pSK144C11 finally identified the gene putatively responsible for the modifications, which we designated *gtfC* (see Fig. S1B in the supplemental material). The deduced 459-aa sequence of the corresponding enzyme revealed motif similarities to UDP-glucuronosyltransferase/UDP-glucosyltransferases. GtfC (GenBank entry AGH18139) showed a similarity of 71% to the putative glycosyltransferase of the Gram-negative bacterium *Fibrisoma limi* covering 92% of the protein (Table 1). Further cloning of the *gtfC* ORF into the pDrive vector and biotransformation with *E. coli* DH5 α carrying the respective construct pD*gtfC* confirmed the flavonoid-modifying activity of GtfC (Fig. 2E).

In summary, these results demonstrated that the developed screening procedure, META, is sufficiently sensitive to allow the identification of large-insert clones from individual bacterial genomes (i.e., *Bacillus* sp. HH1500) and complex metagenome libraries (i.e., the river Elbe sediment library) showing flavonoid-modifying activities.

Sequence-based classification of MgtB and GtfC. To analyze the affiliation of MgtB and GtfC, we constructed a phylogenetic tree using the MEGA, version 5, software (44). The amino acid

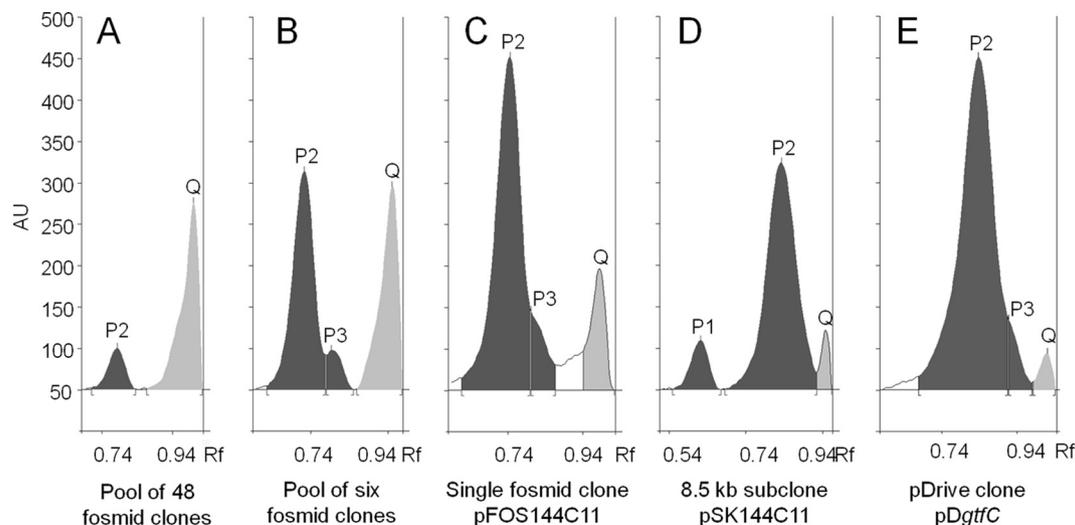


FIG 2 Iterative TLC analyses of culture extracts for the identification of a flavonoid-modifying enzyme. Cells were grown in LB medium with appropriate antibiotics. Biotransformation was performed as described in Materials and Methods. Culture extracts in EtOAc from 24-h biotransformation reactions with 100 μ M quercetin as a substrate were applied on Merck silica gel 60 F₂₅₄ TLC plates. UV chromatograms are displayed in relative absorbance units (AU) versus the R_f value measured at 365 nm on a densitometric TLC Scanner 3 (Camag, Muttenz, Switzerland) for activity determination. Peaks of the remaining quercetin substrate are depicted in light gray near the solvent front (Q); product peaks are shown in dark gray (P1, P2, and P3). TLC analyses of culture extracts from the following biotransformations led to the final isolation of the GT-encoding ORF *gtfC*: pool MT144R of 48 fosmid clones within the positive clone pFOS144C11 (A), pool MT144C of six fosmid clones within the positive clone (B), the positive single fosmid clone pFOS144C11 (C), the positive subclone pSK144C11 derived from pFOS144C11 (D), and the active ORF *gtfC* derived from pFOS144C11 in clone pD*gtfC* (E).

sequences of MgtB and GtfC and their closest sequence-based relatives determined by pBLAST were aligned by ClustalW. Additionally, the sequences of the actually published prokaryotic flavonoid-active GTs were aligned along with those of two eukaryotic enzymes, the flavonoid glucosyltransferase UGT85H2 from *Medicago truncatula* and the flavonoid rhamnosyltransferase UGT78D1 from *Arabidopsis thaliana* as an outer group (45, 46). From these sequences, a neighbor-joining tree with 100 bootstraps was computed (Fig. 3). As expected, MgtB from *Bacillus* sp. HH1500 clustered with other MGTs from the *B. cereus* group. At the time of writing, the MGT of *B. thuringiensis* IBL 200 and the MGT of *B. cereus* G9842 turned out to be the closest relatives, with an amino acid identity to MgtB of 98% each. Both MGTs were annotated as predicted enzymes, and no substrate data were available. From the MGT cluster, five other enzymes were previously reported to mediate the glucosylation of flavonoids. Three of them, BcGT-1, the nearest relative reported to be flavonoid active, BcGT-4, and BcGT-3, all originated from *B. cereus* ATCC 10987 (47–49). Another flavonoid-active MGT, designated BsGT-3, originates from *B. subtilis* strain 168 (36). The remaining flavonoid-active MGT is the well-studied OleD from *Streptomyces antibioticus* (50, 51). GtfC was located in a distinct cluster of UGTs and appeared to be somewhat related to hypothetical enzymes from *Cytophagaceae* bacteria such as *Dyadobacter fermentans* and *Fibrisoma limi* (Fig. 3). Within this cluster only the UGT XcGT-2 is known to accept flavonoid substrates (38). Interestingly, rhamnosyltransferases like BSIG 4748 from *Bacteroides* sp. strain 116 and RtfA from *Mycobacterium avium* phylogenetically also show affiliation to this cluster but form a separate branch (Fig. 3).

To further characterize the identified metagenome-derived GTs, the amino acid residues of the C-terminal donor binding regions were compared to the motifs of the closest relatives and the known flavonoid-active GTs (Fig. 4). Here, the Rossmann fold

$\alpha/\beta/\alpha$ subdomain, the conserved donor-binding region of UGTs, is located (52). Plant UDP-glycosyltransferases like UGT85H2 and UGT78D1 exhibit a highly conserved motif in this region, which is termed the plant secondary product glycosyltransferase (PSPG) motif (45, 46, 53, 54). By alignment we were able to identify key amino acids known to be of importance for NDP-sugar binding. While MgtB revealed a clear UDP-hexose binding motif consisting of highly conserved Gln289 and Glu310 residues for ribose binding and a conserved DQ, GtfC lacked this motif (45, 55, 56). Instead, GtfC presented typical residues Phe336 and Leu357 for deoxyribose nucleotide utilization (57). Moreover, we were able to identify the pyrophosphate binding sites in the MgtB amino acid sequence (Fig. 4). However, GtfC does not possess these conserved phosphate binding residues, suggesting that GtfC and related enzymes have another donor binding mode. In this context GtfC seemed to belong to a novel enzyme class, as underlined by the low level of sequence homology.

Overexpression and glycosylation patterns of MgtB and GtfC. To further characterize the novel enzymes and verify their functions, we overexpressed and purified MgtB and GtfC as His-tagged proteins in *E. coli* BL21 (DE3). Both genes, *mgtB* and *gtfC*, were ligated into the expression vector pET19b. The recombinant enzymes containing N-terminal His₁₀ tags were purified by Ni affinity chromatography under native conditions and gradient elution. MgtB could be purified to yield more than 5 mg of protein/g of cell pellet (wet weight). The maximum yield of GtfC was 3 mg of protein/g of cell pellet. The molecular weights of the proteins were verified by SDS-PAGE analysis under denaturing conditions according to Laemmli. After Coomassie staining, His₁₀-MgtB was visible as a single band with a molecular mass of approximately 50 kDa on a 12% SDS-PAGE gel (see Fig. S2A in the supplemental material). This was in accordance with the calculated molecular mass of 51.2 kDa including the N-terminal His

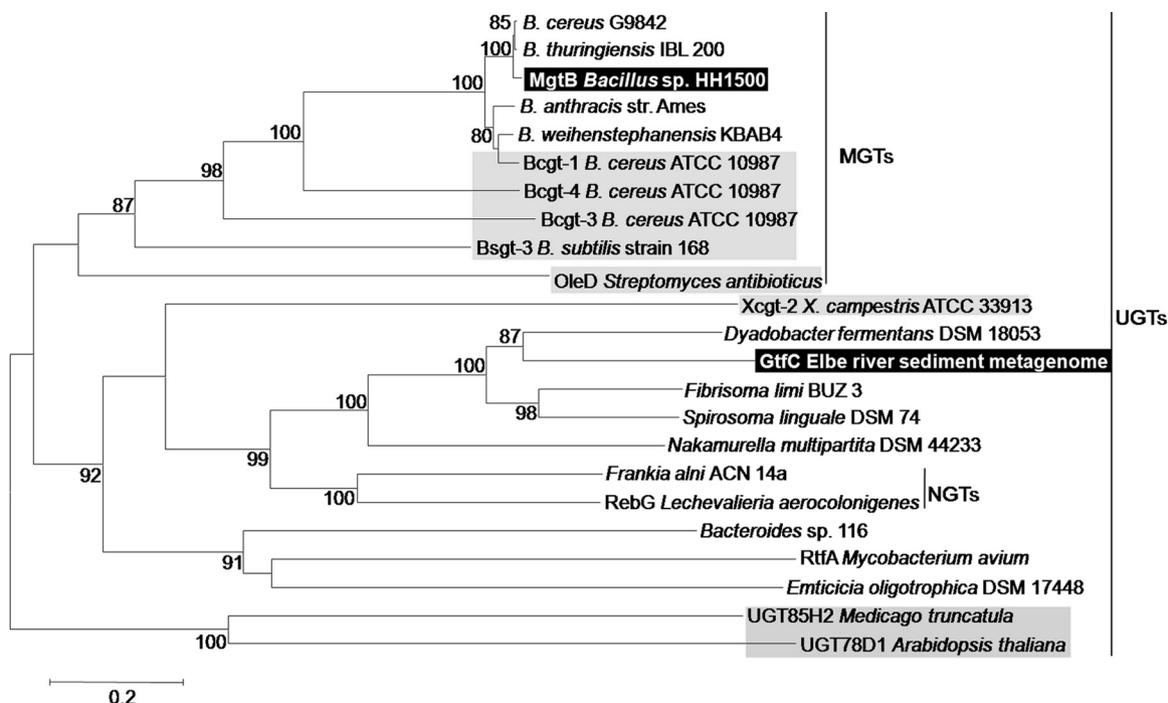


FIG 3 Phylogenetic dendrogram of glycosyltransferases (GTs) related to the two GTs (MgtB and GtfC in black boxes) identified in this study. GTs known to act on flavonoids as acceptors are highlighted in gray. Phylogenetic analysis was conducted using MEGA, version 5 (44), with ClustalW sequence alignment in a BLOSUM protein weight matrix. The neighbor-joining tree was calculated using the bootstrap method with the Poisson model; bootstrap values higher than 75 are indicated next to the branches. The scale represents the number of amino acid changes per residue. All GTs shown belong to the family of Gtf-like GT1 and are further subclassified into MGTs (macroside glycosyltransferases; TIGR01426), NGTs (*N*-glycosyltransferases), and UGT (a UDP-glucuronosyltransferase/-glucosyltransferase; PF00201) as indicated on the right. GenBank numbers were retrieved from NCBI as follows: [YP002445489](#), MGT from *B. cereus* G9842; [ZP04071678](#), MGT from *B. thuringiensis* IBL 200; [AGH18136](#), MgtB from *Bacillus* sp. HH1500; [AAP25969](#) *Bacillus anthracis* strain Ames; [ABY43166](#), MGT from *Bacillus weihenstephanensis* KBAB4; [NP978481](#), BcGT-1 from *B. cereus* ATCC 10987; [NP979441](#), BcGT-4 from *B. cereus* ATCC 10987; [AAS41737](#), BcGT-3 from *B. cereus* ATCC 10987; [NP389104](#), BsGT-3 from *B. subtilis* strain 168; [ABA42119](#), OleD from *S. antibioticus*; [AAM41712](#), XcGT-2 from *X. campestris* ATCC 33913; [YP003086330](#), Dfer 1940 from *Dyadobacter fermentans* DSM 18053; [AGH18139](#), GtfC from Elbe river sediment metagenome; [CCH52088](#), UGT from *Fibrisoma limi* BUZ 3; [YP003388759](#), Slin3970 from *Spirosoma linguale* DSM 74; [ACV78946](#), Namu2594 from *Nakamurella multipartita* DSM 44233; [YP712191](#), FRAAL1959 from *Frankia alni* ACN14a; [AAN01207](#), RebG from *Lechevalieria aerocolonigenes*; [ZP09941874](#), BSIG 4748 from *Bacteroides* sp. 116; [AAC71702](#), RtfA from *Mycobacterium avium*; [AFK05536](#), Emtol 0266 from *Emticia oligotrophica* DSM 17448; [2PQ6A](#), UGT85H2 from *Medicago truncatula*; and [NP564357](#), UGT78D1 from *Arabidopsis thaliana*.

tag. His₁₀-GtfC revealed a molecular mass of about 55 kDa on a 12% SDS-PAGE gel, which was well in accordance with the calculated molecular mass of 54.7 kDa including the N-terminal His tag. While virtually no additional bands were visible on SDS-PAGE gels with purified recombinant MgtB protein, some minor contaminating bands were still visible on the SDS-PAGE gel loaded with purified GtfC (see Fig. S2B in the supplemental material). In summary, both proteins could be purified to allow further biochemical characterization.

The purified His₁₀-MgtB protein was able to use UDP- α -D-glucose as a donor substrate. The recombinant enzyme catalyzed the transfer of α -D-glucose residues to various polyphenols. Biocatalytic reactions were performed with 500 μ M UDP- α -D-glucose as a donor and 100 μ M acceptor substrate. The following flavonoids served as acceptor substrates and were modified with high yields: luteolin, quercetin, kaempferol, tiliroside, naringenin, and genistein (Table 2). Flavonols turned out to be the best acceptor molecules. Generally, the conversion during a 2-h assay ranged from 52% for naringenin to approximately 100% for quercetin and kaempferol. Interestingly, in the presence of quercetin and kaempferol, no residual educts could be monitored by HPTLC analysis. The specific educts and their observed glycones of the

biocatalytic reactions are summarized in Table 2 together with the respective R_f values. MgtB favored the glucosylation at the C-3 hydroxy group, if accessible, as in the aglycone flavonols quercetin and kaempferol. Further, the C-7-OH was attacked and glucosylated by the enzyme, which could be shown for not only the flavone luteolin but also the flavanone naringenin and the isoflavone genistein (Table 2). MgtB glucosylated luteolin also at the C-3' hydroxy group forming the 3',7-di-O-glucoside of luteolin if the C-7-OH was glucosylated previously. Furthermore, MgtB catalyzed the conversion of the kaempferol derivative tiliroside [kaempferol-3-(O-6'-trans-p-coumaryl)-glucoside]. One glucosylated product with an R_f value of 0.54 was detected.

Finally, the chalcone xanthohumol and the stilbene *t*-resveratrol were tested in biotransformation reactions with *E. coli* expressing *mgtB*, but conversions were not quantified (data not shown). Xanthohumol yielded three detectable products, whereas the biotransformation of *t*-resveratrol yielded one observed product by absorbance TLC analysis.

Tests with recombinant and purified GtfC using UDP- α -D-glucose and UDP- α -D-galactose and quercetin as an acceptor molecule suggested that dTDP-activated sugar moieties were transferred by this enzyme. This finding was confirmed by HPLC-ESI-MS analyses

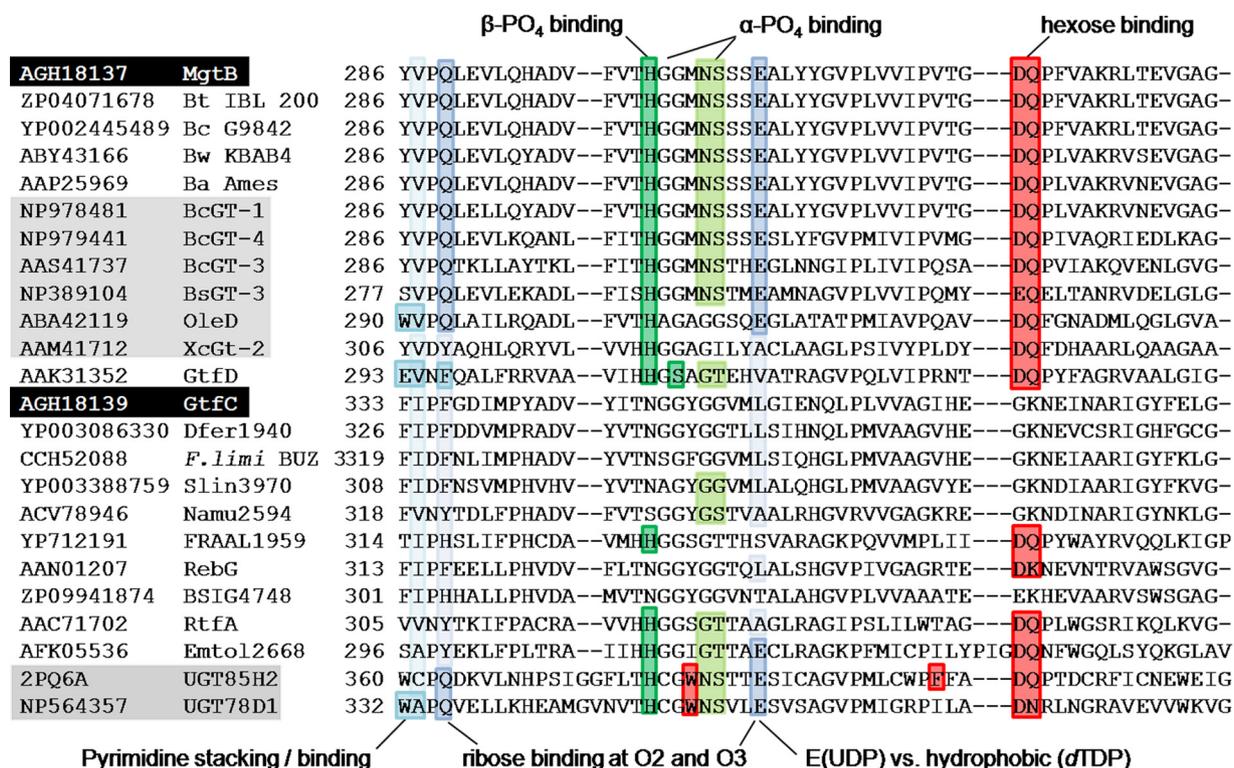


FIG 4 ClustalW alignment of the MgtB and GtfC (black boxes) amino acid sequences, their nearest sequence-based relatives, and other flavonoid-active GTs (gray boxes). The C-terminal region of the Rossmann fold $\alpha/\beta/\alpha$ subdomain, the conserved donor-binding region of GTs, is shown. Plant UDP-glycosyltransferases like UGT85H2 (2PQ6A) from *Medicago truncatula* and UGT78D1 (NP564357) from *Arabidopsis thaliana* exhibit the highly conserved plant secondary product glycosyltransferase (PSPG) motif in this region. Amino acids are boxed according to their roles in donor nucleoside (blue), phosphate group (green), and hexose (red) binding, as reported in the literature (45, 52, 55–57, 73, 74). Dark and light green boxes refer to beta- and alpha-phosphate group binding of the donor molecule, respectively. Analogous binding of the nucleoside part is indicated for base positions (light and full bright blue) and ribose positions (dark blue); for the latter, lighter variants of blue boxes differentiate deoxyribose-specific and ribose-specific positions. GenBank numbers were retrieved from NCBI (see the legend of Fig. 3 for species identifications).

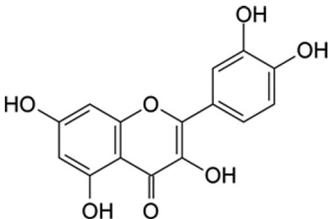
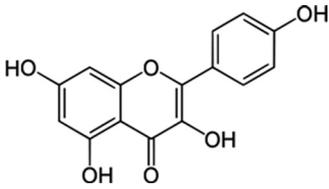
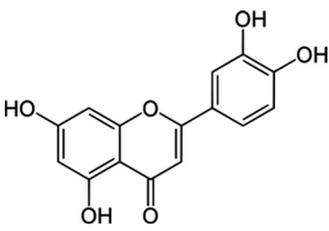
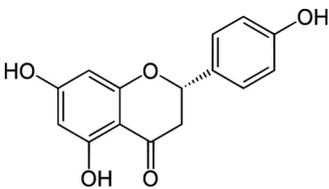
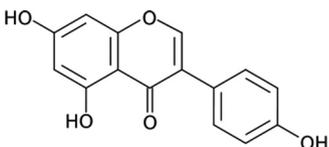
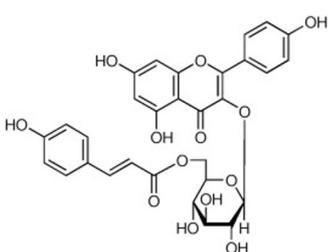
of biotransformation assays (see the following paragraph). Unfortunately, deoxyribose nucleotide-activated hexoses, e.g., dTDP-rhamnoside, were commercially not available to further analyze the obtained reaction products in more detail (58).

Biotransformations with the *E. coli* strain expressing GtfC and using various polyphenols as substrates yielded conversions ranging from 52% for xanthohumol up to almost 100% turnover for most flavonols tested (Table 3). Quercetin was transformed almost completely after 4-h biotransformations and yielded three detectable products (P1 to P3). To further characterize these products, UV absorbance spectra were recorded and compared to the reference glycones of quercetin isoquercitrin and quercitrin (59). P1 revealed an R_f value identical to the value of isoquercitrin. Further, the UV absorbance spectrum of P1 matched the spectrum of isoquercitrin (see Fig. S3A in the supplemental material). P2 revealed an R_f value identical to the one known for quercitrin. P2 also exhibited the same UV absorbance spectrum as quercitrin (see Fig. S3B). P3 revealed an R_f value of 0.82, which clearly differed from the R_f values of known and available quercetin glycones. Compared to isoquercitrin, P3 showed a similar hypsochromic shift of band I to a maximum wavelength (λ_{max}) of 363 nm (see Fig. S3C); however, it revealed a less hypsochromic shift in band II of only 5 nm to 272 nm with a shoulder at 280 nm. It is further notable that the HPLC-ESI-MS analysis of biotransformation products of quercetin consistently identified three distinct

reaction products (see Fig. S4). P1 had a retention time (RT) of 17.93 min in the HPLC analysis and revealed a molecular mass of 464 Da, which is equivalent to isoquercitrin. P2 revealed an RT of 18.06 min and had a molecular mass of 448 Da. This mass corresponds well with the molecular mass of quercitrin. Finally, P3 with a RT of 18.31 min revealed a molecular mass of 446 Da, indicating the formation of a novel, not further characterized, quercetin glycoside.

Glycosylation patterns of GtfC on quercetin suggested a preference to act on the C-3 hydroxy group mediating the transfer of different sugar residues. However, if a C-3-OH group was not available, GtfC efficiently catalyzed the glycosylation of other positions. Flavones lacking the hydroxy function at C-3 were converted depending on the availability of other hydroxy groups. Pratol, which possesses only a single free C-7 hydroxy group, was converted weakly and yielded a single detectable product. Further, the biotransformation of 3',4'-dihydroxyflavone yielded three detectable glycones, and 5-methoxy-eupatorin yielded two products (see Table S4 in the supplemental material; also data not shown); the biotransformation of the mono-4'-hydroxyflavanone yielded one glycosylated product, and the glycosylation of naringenin yielded two products. The major biotransformation product of naringenin revealed the same R_f values and absorbance spectra as prunin, the naringenin-7-O-glucoside (Table 3). The second naringenin glycone could not be further specified due to the lack of

TABLE 2 Flavonoid substrates converted by recombinant MgtB in bioassays

Substrate				
Name	Structure	Conversion (%) ^a	R _f value ^b	Product ^b
Quercetin		~100	0.79 0.64 0.27 0.25	— ^c Isoquercitrin — —
Kaempferol		~100	0.74 0.35	Astragalin —
Luteolin		82	0.65 0.32	Cynaroside -3',7-Di-O-Glc
Naringenin		52	0.76	Prunin
Genistein		72	0.69	Genistin
Tiliroside		83	0.54	—

^a Reactions were carried out at 37°C for 2 h in triplicate using 1 ml of reaction mixture consisting of 500 μM UDP-glucose, a 100 μM concentration of the respective flavonoid, and 5 μg/ml of purified and recombinant MgtB.

^b R_f values and products in bold indicate the main products of the biocatalytic reactions.

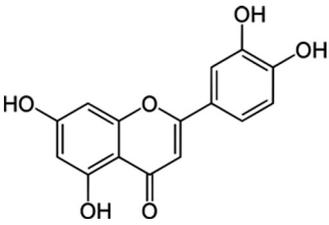
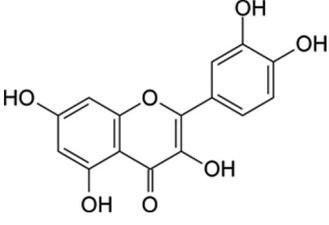
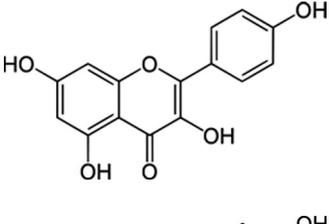
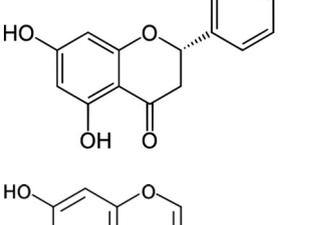
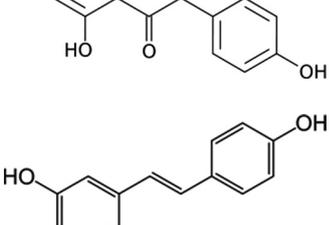
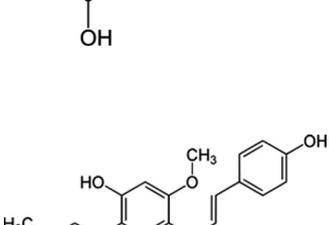
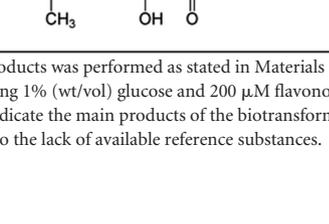
^c Products were not specified due to the lack of available reference substances.

commercially available reference substances. Altogether these results suggested that GtFC acts on the C-3, C-3', C-4', and C-7 hydroxy groups of the flavonoid backbone.

In summary these data demonstrated that MgtB and GtFC pos-

sess interesting biocatalytic properties. While MgtB specifically mediated the transfer of glucose residues, GtFC transferred different hexose moieties. MgtB was capable of catalyzing the glycosylation of already glycosylated flavonoids to form diglycosides

TABLE 3 Flavonoid substrates and products of biotransformation assays with recombinant GtFC

Substrate				
Name	Structure	Conversion (%) ^a	R _f value ^b	Product ^b
Luteolin		86	0.81 0.73 0.68 0.58	— ^c — — —
Quercetin		~100	0.82 0.75 0.64	— Quercitrin Isoquercitrin
Kaempferol		~100	0.85 0.80 0.68	— — Astragalin
Naringenin		76	0.87 0.84 0.77	— — Prunin
Genistein		68	0.83 0.76 0.68	— — Genistin
<i>t</i> -Resveratrol		96	0.83 0.77 0.64 0.58 0.51 0.46	— — — — — —
Xanthohumol		52	0.85 0.48	— —

^a Quantification of the reaction products was performed as stated in Materials and Methods. Triplicate reactions using 50 ml of reaction mixture were performed in 50 mM sodium phosphate buffer, pH 7.0, containing 1% (wt/vol) glucose and 200 μM flavonoid at 30°C.

^b R_f values and products in bold indicate the main products of the biotransformation reactions.

^c Products were not specified due to the lack of available reference substances.

(e.g., formation of luteolin-3',7-di-O-glucoside) and even tiliroside to generate novel glucosides not available from natural resources. In contrast, the glycosylation pattern of GtfC suggested the transfer of single sugar residues to only aglycone flavonoid forms. Interestingly, GtfC seemed to be very variable with respect to its activity at various positions on the flavonoid backbone. This may lead to the formation of truly novel flavonoids not available naturally. Hence, both enzymes might be helpful in the generation of new natural compounds.

DISCUSSION

Within the manuscript, we report on the development of a semi-automated TLC-based detection system for flavonoid-modifying enzyme clones. The screening assay was highly reproducible and highly sensitive. It allowed the detection of micromolar concentrations of glycosylated flavonoids. Isoquercitrin was detectable at a 0.78 μ M concentration in the assay (see Table S3 in the supplemental material). Using this assay we were able to systematically identify one positive clone out of pools of 96 metagenome clones (Fig. 1 and 2). To our knowledge this is the first published TLC-based screening method for functional searches in metagenome libraries. We speculate that slight modifications of this screening system could easily allow the detection of other flavonoid-modifying enzyme clones, for example, through acylation or methylation reactions.

Using this novel screening technology, we identified a macroside glycosyltransferase, MgtB, from a soil isolate (i.e., *Bacillus* sp. HH1500). A fosmid library established with DNA from this strain, which had been isolated from the local botanical garden only recently, was initially used to develop and verify the outlined screening technology; and using the novel screening technology, MgtB was quickly identified from a pool of almost 2,000 clones. Isolation and purification of recombinant MgtB revealed a novel MGT. MgtB shared 89% amino acid identity with BcGT-1 from *B. cereus* ATCC 10987, the closest relative reported to act on flavonoids. BcGT-1 was reported to catalyze the glucosylation of flavones, flavonols, flavanones, and isoflavones (47). On flavonols BcGT-1 acted on C-3, C-7, and C-4' hydroxy groups creating triglucosides of kaempferol (48). In contrast, biocatalysis of kaempferol with MgtB yielded just two detectable glucosylated products. However, reactions with quercetin resulted in three detectable glycones. These data suggested that MgtB acted at the C-3'-OH group. This hypothesis also was supported by the observation that recombinant MgtB converted luteolin to luteolin-3',7-di-O-glucoside as a by-product. These results were in accordance with the glucosylation pattern of BcGT-3, yet another MGT from *B. cereus* ATCC 10987 (49). Interestingly, BcGT-3 shares only 40% amino acid identity with MgtB, but both enzymes act on the same flavonoids, forming diglucosides from flavones and flavonols at the same positions and only monoglucosides from naringenin. The most spectacular conversion observed for MgtB was that of tiliroside. The product is likely to be the 7-O-glucoside, taking the glycosylation pattern of MgtB into account. Tiliroside glycosides, however, have not been reported in scientific literature. This raises the possibility of the generation of new natural compounds. The natural substrates of *Bacillus* MGTs still have not been reported. Other MGTs like OleD usually detoxify macroside antibiotics but often possess broad acceptor tolerance (35, 60).

The metagenome-derived GtfC turned out to be a completely novel enzyme. Only seven flavonoid-active UGTs have been re-

ported so far that originate from five different prokaryotes (35, 36, 38, 47, 49). With the exception of XcGT-2 from the Gram-negative *X. campestris* ATCC 33913, all are MGT enzymes from Gram-positive *Bacilli* and *Streptomyces*. MGTs play an important role in xenobiotic defense mechanisms of prokaryotes and thus show broad acceptor specificities (55, 60). This also applies to eukaryotic UGTs, pointing to a biological principle of detoxification (61). To our knowledge GtfC is the first metagenome-derived GT acting on flavonoids. Moreover, it is also the first bacterial enzyme reported to transfer various dTDP-activated hexose sugars to polyphenols (see below), in contrast to the usually stringent donor specificities of Gtf-like enzymes such as GtfD (57). With respect to the notion that many NDP-sugars in prokaryotes are dTDP and not UDP activated, GtfC might be a promising biocatalyst in glyco-diversification approaches (58, 62, 63). GtfC is similar to predicted GTs from *Cytophagaceae* bacteria (64–66). These Gram-negative bacteria have large genomes, suggesting extensive secondary metabolic pathways, and they are well known for the presence of resistance mechanisms to antibiotics such as trimethoprim and vancomycin (67, 68). As commonly known, glycosylation of xenobiotics is a ubiquitous detoxification process in all kingdoms of life. The phylogenetically diverse members of *Cytophagaceae* have only recently become an object of research, and concrete estimation of the phylogenetic breadth of this family and exact taxonomic ranking still remain unclear (65, 69). Thus, the identification of the metagenome-derived GtfC and its partial characterization suggest that this group of microorganisms is perhaps a highly promising resource for novel GTs and also other enzymes.

A ClustalW alignment of the donor-binding region of GtfC suggested that the activated donor substrates are of deoxy-thymidine nucleoside origin. GtfC possesses the typical amino acid residues Phe336 for thymine base stacking and hydrophobic Leu357 for deoxyribose fitting (57). Concerning the donor binding of GTs, GtfC appears not to exhibit the known amino acid residues for pyrophosphate binding (Fig. 4). Instead of the conserved residue His/Arg in the current solved protein structures, GtfC contains an Asn at amino acid position 349 (52, 70). This applies also for the nearest GtfC relatives Dfer1940, UGT of *F. limi* BUZ 3, and Slin3970 as well as the NGTs RebG and BSG4748. Further, GtfC does not show the conserved Ser/Thr residue responsible for α -phosphate binding. Instead, the Gly354 appears to be of importance for the α -phosphate binding, similar to structure of the OleD transferase (55).

The assumption of dTDP-activated cosubstrates used by GtfC was supported by the observation that glucose, rhamnose, and a third sugar residue with a molecular weight of 446 were transferred by GtfC in biotransformations using intact *E. coli* cells. Moreover, biocatalytic approaches with purified GtfC and either UDP- α -D-glucose or -galactose as a donor substrate failed. In bacteria, the activated sugars, dTDP- α -D-glucose, -4-keto-6-deoxy- α -D-glucose or -4-keto- β -L-rhamnose, and - β -L-rhamnose, are part of the dTDP-sugar biosynthesis pathway and are present in *E. coli* (71). Moreover, levels of dTDP-sugars are allosterically regulated by dTDP-rhamnose levels through activity of RmlA (72).

In summary, the screening protocol described in this report is a very helpful tool for the identification of truly novel enzymes for the modifications of flavonoids and related substrates. In this study, we have used the technology to identify two novel flavonoid-modifying enzymes. Both of these enzymes would perhaps not have been detected without the above-developed screen-

ing technology. The partial biochemical characterizations using either biocatalysis or biotransformation suggest that MgtB and GtfC are both very interesting enzymes with high potential for biotechnological applications with respect to flavonoid modifications on an industrial scale. Thus, future work will now refine this technology to also identify other enzymes linked to flavonoid modifications.

ACKNOWLEDGMENTS

This work was funded by the Ministry of Education and Research Bundesministerium für Bildung und Forschung (BMBF) within the network program Biokatalyse2021 as project P18 (FKZ 0315166C).

The Bacillus Genetic Stock Center (Columbus, OH, USA) kindly provided *Bacillus cereus* ATCC 10987 used as a positive control.

REFERENCES

- Perner M, Ilmberger N, Köhler HU, Chow J, Streit WR. 2011. Emerging fields in functional metagenomics and its industrial relevance: overcoming limitations and redirecting the search for novel biocatalysts, p 483–498. In de Bruijn FJ (ed), Handbook of molecular microbial ecology II: metagenomics in different habitats. John Wiley & Sons, Inc., Hoboken, NJ.
- Schmeisser C, Steele H, Streit WR. 2007. Metagenomics, biotechnology with non-culturable microbes. Appl. Microbiol. Biotechnol. 75:955–962.
- Iqbal HA, Feng Z, Brady SF. 2012. Biocatalysts and small molecule products from metagenomic studies. Curr. Opin. Chem. Biol. 16:109–116.
- Simon C, Daniel R. 2011. Metagenomic analyses: past and future trends. Appl. Environ. Microbiol. 77:1153–1161.
- Daniel R. 2005. The metagenomics of soil. Nat. Rev. Microbiol. 3:470–478.
- Streit WR, Daniel R, Jaeger K-E. 2004. Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms. Curr. Opin. Biotechnol. 15:285–290.
- Warren RL, Freeman JD, Levesque RC, Smailus DE, Flibotte S, Holt RA. 2008. Transcription of foreign DNA in *Escherichia coli*. Genome Res. 18:1798–1805.
- Uchiyama T, Miyazaki K. 2009. Functional metagenomics for enzyme discovery: challenges to efficient screening. Curr. Opin. Biotech. 20:616–622.
- Ferrer M, Beloqui A, Timmis KN, Golyshin PN. 2009. Metagenomics for mining new genetic resources of microbial communities. J. Mol. Microbiol. Biotechnol. 16:109–123.
- Tuffin M, Anderson D, Heath C, Cowan DA. 2009. Metagenomic gene discovery: how far have we moved into novel sequence space? Biotechnol. J. 4:1671–1683.
- Simon C, Daniel R. 2009. Achievements and new knowledge unraveled by metagenomic approaches. Appl. Microbiol. Biotechnol. 85:265–276.
- Steele HL, Jaeger KE, Daniel R, Streit WR. 2009. Advances in recovery of novel biocatalysts from metagenomes. J. Mol. Microbiol. Biotechnol. 16: 25–37.
- Taupp M, Mewis K, Hallam SJ. 2011. The art and design of functional metagenomic screens. Curr. Opin. Biotechnol. 22:465–472.
- Beloqui A, Pita M, Polaina J, Martinez-Arias A, Golyshina OV, Zumarraga M, Yakimov MM, Garcia-Arellano H, Alcalde M, Fernandez VM, Elborough K, Andreu JM, Ballesteros A, Plou FJ, Timmis KN, Ferrer M, Golyshin PN. 2006. Novel polyphenol oxidase mined from a metagenome expression library of bovine rumen: biochemical properties, structural analysis, and phylogenetic relationships. J. Biol. Chem. 281:22933–22942.
- Collier AC, Tingle MD, Keelan JA, Paxton JW, Mitchell MD. 2000. A highly sensitive fluorescent microplate method for the determination of UDP-glucuronosyl transferase activity in tissues and placental cell lines. Drug Metab. Dispos. 28:1184–1186.
- DeSantis G, Zhu Z, Greenberg WA, Wong K, Chaplin J, Hanson SR, Farwell B, Nicholson LW, Rand CL, Weiner DP, Robertson DE, Burk MJ. 2002. An enzyme library approach to biocatalysis: development of nitrilases for enantioselective production of carboxylic acid derivatives. J. Am. Chem. Soc. 124:9024–9025.
- Knietsch A, Waschowitz T, Bowien S, Henne A, Daniel R. 2003. Metagenomes of complex microbial consortia derived from different soils as sources for novel genes conferring formation of carbonyls from short-chain polyols on *Escherichia coli*. J. Mol. Microbiol. Biotechnol. 5:46–56.
- Schipper C, Hornung C, Bijtenhoorn P, Quitschau M, Grond S, Streit WR. 2009. Metagenome-derived clones encoding two novel lactonase family proteins involved in biofilm inhibition in *Pseudomonas aeruginosa*. Appl. Environ. Microbiol. 75:224–233.
- Uchiyama T, Abe T, Ikemura T, Watanabe K. 2005. Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. Nat. Biotechnol. 23:88–93.
- Williamson LL, Borlee BR, Schloss PD, Guan C, Allen HK, Handelsman J. 2005. Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor. Appl. Environ. Microbiol. 71:6335–6344.
- Ververidis F, Trantas E, Douglas C, Vollmer G, Kretzschmar G, Panoopoulos N. 2007. Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part II: reconstruction of multienzyme pathways in plants and microbes. Biotechnol. J. 2:1235–1249.
- Schütz K, Muks E, Carle R, Schieber A. 2006. Quantitative determination of phenolic compounds in artichoke-based dietary supplements and pharmaceuticals by high-performance liquid chromatography. J. Agric. Food Chem. 54:8812–8817.
- Leonard E, Yan Y, Fowler ZL, Li Z, Lim CG, Lim KH, Koffas MA. 2008. Strain improvement of recombinant *Escherichia coli* for efficient production of plant flavonoids. Mol. Pharm. 5:257–265.
- Wang M, Liang CP, Wu QL, Simon JE, Ho CT (ed.). 2006. Instrumental analysis of popular botanical products in the U.S. market, p. 25–38. In Herbs: challenges in chemistry and biology. American Chemical Society, Washington, DC.
- Manach C, Scalbert A, Morand C, Remesy C, Jimenez L. 2004. Polyphenols: food sources and bioavailability. Am. J. Clin. Nutr. 79:727–747.
- Das S, Rosazza JP. 2006. Microbial and enzymatic transformations of flavonoids. J. Nat. Prod. 69:499–508.
- Graefe EU, Wittig J, Mueller S, Riethling AK, Uehleke B, Drewelow B, Pforte H, Jacobasch G, Derendorf H, Veit M. 2001. Pharmacokinetics and bioavailability of quercetin glycosides in humans. J. Clin. Pharmacol. 41:492–499.
- Kren V, Martinkova L. 2001. Glycosides in medicine: the role of glycosidic residue in biological activity. Curr. Med. Chem. 8:1303–1328.
- Coutinho PM, Deleury E, Davies GJ, Henrissat B. 2003. An evolving hierarchical family classification for glycosyltransferases. J. Mol. Biol. 328: 307–317.
- Bowles D, Lim EK, Poppenberger B, Vaistij FE. 2006. Glycosyltransferases of lipophilic small molecules. Annu. Rev. Plant Biol. 57:567–597.
- Mackenzie PI, Owens IS, Burchell B, Bock KW, Bairoch A, Belanger A, Fournel-Gigleux S, Green M, Hum Iyanagi DWT, Lancet D, Louisot P, Magdalou J, Chowdhury JR, Ritter JK, Schachter H, Tephly TR, Tipton KF, Nebert DW. 1997. The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. Pharmacogenetics 7:255–269.
- Lairson LL, Henrissat B, Davies GJ, Withers SG. 2008. Glycosyltransferases: structures, functions, and mechanisms. Annu. Rev. Biochem. 77: 521–555.
- Osmani SA, Bak S, Møller BL. 2009. Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. Phytochemistry 70:325–347.
- Breton C, Snajdrova L, Jeanneau C, Koca J, Imbert A. 2006. Structures and mechanisms of glycosyltransferases. Glycobiology 16:29R–37R.
- Yang M, Proctor MR, Bolam DN, Errey JC, Field RA, Gilbert HJ, Davis BG. 2005. Probing the breadth of macrocyclic glycosyltransferases: *in vitro* remodeling of a polyketide antibiotic creates active bacterial uptake and enhances potency. J. Am. Chem. Soc. 127:9336–9337.
- Jeon Y, Kim B, Kim J, Cheong Y, Ahn J-H. 2009. Enzymatic glycosylation of phenolic compounds using BsGT-3. J. Korean Soc. Appl. Biol. Chem. 52:98–101.
- Rao KV, Weisner NT. 1981. Microbial transformation of quercetin by *Bacillus cereus*. Appl. Environ. Microbiol. 42:450–452.
- Kim HJ, Kim BG, Kim JA, Park Y, Lee YJ, Lim Y, Ahn JH. 2007. Glycosylation of flavonoids with *E. coli* expressing glycosyltransferase from *Xanthomonas campestris*. J. Microbiol. Biotechnol. 17:539–542.
- Zhou J, Bruns MA, Tiedje JM. 1996. DNA recovery from soils of diverse composition. Appl. Environ. Microbiol. 62:316–322.
- Ilmberger N, Meske D, Juergensen J, Schulte M, Barthen P, Rabausch U, Angelov A, Mientus M, Liebl W, Schmitz RA, Streit WR. 2012.

- Metagenomic cellulases highly tolerant towards the presence of ionic liquids-linking thermostability and halotolerance. *Appl. Microbiol. Biotechnol.* 95:135–146.
41. Wagner H, Bladt S, Zgainski EM. 1983. Drogenanalyse, Dünnschichtchromatographische Analyse von Arzneidrogen. Springer, Berlin, Germany.
 42. Neu R. 1957. Chelate von Diarylborsäuren mit aliphatischen Oxyalkylaminen als Reagenzien für den Nachweis von Oxyphenyl-benzo-äpyronen. *Naturwissenschaften* 44:181–182.
 43. Tourasse NJ, Helgason E, Oslash OA, Hegna IK, Kolstø AB. 2006. The *Bacillus cereus* group: novel aspects of population structure and genome dynamics. *J. Appl. Microbiol.* 101:579–593.
 44. Tamura K, Peterson P, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28:2731–2739.
 45. Li L, Modolo LV, Escamilla-Trevino LL, Achnine L, Dixon RA, Wang X. 2007. Crystal structure of *Medicago truncatula* UGT85H2-insights into the structural basis of a multifunctional (iso)flavonoid glycosyltransferase. *J. Mol. Biol.* 370:951–963.
 46. Jones P, Messner B, Nakajima J, Schäffner AR, Saito K. 2003. UGT73C6 and UGT78D1, glycosyltransferases involved in flavonol glycoside biosynthesis in *Arabidopsis thaliana*. *J. Biol. Chem.* 278:43910–43918.
 47. Hyung Ko JH, Gyu Kim B, Joong-Hoon A. 2006. Glycosylation of flavonoids with a glycosyltransferase from *Bacillus cereus*. *FEMS Microbiol. Lett.* 258:263–268.
 48. Jung NR, Joe EJ, Kim BG, Ahn BC, Park JC, Chong Y, Ahn JH. 2010. Change of *Bacillus cereus* flavonoid O-triglycosyltransferase into flavonoid O-monoglycosyltransferase by error-prone polymerase chain reaction. *J. Microbiol. Biotechnol.* 20:1393–1396.
 49. Ahn BC, Kim BG, Jeon YM, Lee EJ, Lim Y, Ahn JH. 2009. Formation of flavone di-O-glucosides using a glycosyltransferase from *Bacillus cereus*. *J. Microbiol. Biotechnol.* 19:387–390.
 50. Choi SH, Ryu M, Yoon YJ, Kim DM, Lee EY. 2012. Glycosylation of various flavonoids by recombinant oleandomycin glycosyltransferase from *Streptomyces antibioticus* in batch and repeated batch modes. *Biotechnol. Lett.* 34:499–505.
 51. Williams GJ, Zhang C, Thorson JS. 2007. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* 3:657–662.
 52. Hu Y, Walker S. 2002. Remarkable structural similarities between diverse glycosyltransferases. *Chem. Biol.* 9:1287–1296.
 53. Hughes J, Hughes MA. 1994. Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. *DNA Seq.* 5:41–49.
 54. Paquette S, Moller BL, Bak S. 2003. On the origin of family 1 plant glycosyltransferases. *Phytochemistry* 62:399–413.
 55. Bolam DN, Roberts S, Proctor MR, Turkenburg JP, Dodson EJ, Martinez-Fleites C, Yang M, Davis BG, Davies GJ, Gilbert HJ. 2007. The crystal structure of two macrolide glycosyltransferases provides a blueprint for host cell antibiotic immunity. *Proc. Natl. Acad. Sci. U. S. A.* 104:5336–5341.
 56. Offen W, Martinez-Fleites C, Yang M, Kiat-Lim E, Davis BG, Tarling CA, Ford CM, Bowles DJ, Davies GJ. 2006. Structure of a flavonoid glycosyltransferase reveals the basis for plant natural product modification. *EMBO J.* 25:1396–1405.
 57. Mulichak AM, Lu W, Losey HC, Walsh CT, Garavito RM. 2004. Crystal structure of vancosaminyltransferase GtFD from the vancomycin biosynthetic pathway: interactions with acceptor and nucleotide ligands. *Biochemistry* 43:5170–5180.
 58. Lim E-K, Ashford DA, Bowles DJ. 2006. The synthesis of small-molecule rhamnosides through the rational design of a whole-cell biocatalysis system. *ChemBioChem* 7:1181–1185.
 59. Mabry TJ, Markham KR, Thomas MB. 1970. The systematic identification of flavonoids. Springer-Verlag, New York, NY.
 60. Gantt RW, Goff RD, Williams GJ, Thorson JS. 2008. Probing the aglycon promiscuity of an engineered glycosyltransferase. *Angew. Chem. Int. Ed. Engl.* 47:8889–8892.
 61. Vogt T, Jones P. 2000. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends Plant Sci.* 5:380–386.
 62. Williams GJ, Gantt RW, Thorson JS. 2008. The impact of enzyme engineering upon natural product glycodiversification. *Curr. Opin. Chem. Biol.* 12:556–564.
 63. Yoon JA, Kim BG, Lee WJ, Lim Y, Chong Y, Ahn JH. 2012. Production of a novel quercetin glycoside through metabolic engineering of *Escherichia coli*. *Appl. Environ. Microbiol.* 78:4256–4262.
 64. Filippini M, Kaech A, Ziegler U, Bagheri HC. 2011. *Fibrisoma limi* gen. nov., sp. nov., a filamentous bacterium isolated from tidal flats. *Int. J. Syst. Evol. Microbiol.* 61:1418–1424.
 65. Lail K, Sikorski J, Saunders E, Lapidus A, Glavina Del Rio T, Copeland A, Tice H, Cheng JF, Lucas S, Nolan M, Bruce D, Goodwin L, Pitluck S, Ivanova N, Mavromatis K, Ovchinnikova G, Pati A, Chen A, Palaniappan K, Land M, Hauser L, Chang YJ, Jeffries CD, Chain P, Brettin T, Detter JC, Schutze A, Rohde M, Tindall BJ, Goker M, Bristow J, Eisen JA, Markowitz V, Hugenholtz P, Kyrpides NC, Klenk HP, Chen F. 2010. Complete genome sequence of *Spirosoma linguale* type strain (1). *Stand. Genomic Sci.* 2:176–185.
 66. Lang E, Lapidus A, Chertkov O, Brettin T, Detter JC, Han C, Copeland A, Glavina Del Rio T, Nolan M, Chen F, Lucas S, Tice H, Cheng JF, Land M, Hauser L, Chang YJ, Jeffries CD, Kopitz M, Bruce D, Goodwin L, Pitluck S, Ovchinnikova G, Pati A, Ivanova N, Mavromatis K, Chen A, Palaniappan K, Chain P, Bristow J, Eisen JA, Markowitz V, Hugenholtz P, Goker M, Rohde M, Kyrpides NC, Klenk HP. 2009. Complete genome sequence of *Dyadobacter fermentans* type strain (NS114). *Stand. Genomic Sci.* 1:133–140.
 67. Chelius MK, Triplett EW. 2000. *Dyadobacter fermentans* gen. nov., sp. nov., a novel gram-negative bacterium isolated from surface-sterilized *Zea mays* stems. *Int. J. Syst. Evol. Microbiol.* 50:751–758.
 68. Finster KW, Herbert RA, Lomstein BA. 2009. *Spirosoma spitsbergense* sp. nov. and *Spirosoma luteum* sp. nov., isolated from a high Arctic permafrost soil, and emended description of the genus *Spirosoma*. *Int. J. Syst. Evol. Microbiol.* 59:839–844.
 69. Filippini M, Svercel M, Laczko E, Kaech A, Ziegler U, Bagheri HC. 2008. 2011. *Fibrella aestuarina* gen. nov., sp. nov., a filamentous bacterium of the family *Cytophagaceae* isolated from a tidal flat, and emended description of the genus *Rudanella* Weon et al. 2008. *Int. J. Syst. Evol. Microbiol.* 61:184–189.
 70. Ha S, Gross B, Walker S. 2001. *E. coli* MurG: a paradigm for a superfamily of glycosyltransferases. *Curr. Drug Targets Infect. Disord.* 1:201–213.
 71. Samuel G, Reeves P. 2003. Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydr. Res.* 338:2503–2519.
 72. Giraud MF, Naismith JH. 2000. The rhamnose pathway. *Curr. Opin. Struct. Biol.* 10:687–696.
 73. Modolo LV, Li L, Pan H, Blount JW, Dixon RA, Wang X. 2009. Crystal structures of glycosyltransferase UGT78G1 reveal the molecular basis for glycosylation and deglycosylation of (iso)flavonoids. *J. Mol. Biol.* 392:1292–1302.
 74. Shao H, He X, Achnine L, Blount JW, Dixon RA, Wang X. 2005. Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. *Plant Cell* 17:3141–3154.