NDP-Rhamnose Biosynthesis and Rhamnosyltransferases - Building Diverse Glycoconjugates in Nature

Wagstaff, B. A.1,2, Zorzoli, A.2, Dorfmueller, H. C.2

From the 1Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester, United Kingdom and the 2Division of Molecular Microbiology, School of Life Sciences, University of Dundee, Dundee, DD1 5EH, United Kingdom.

E-mail: h.c.z.dorfmueller@dundee.ac.uk

Keywords: rhamnose, biosynthesis, glycosyltransferase, glycoproteins, polysaccharides, pathogens

Abstract

Rhamnose is an important 6-deoxy sugar present in many natural products, glycoproteins, and structural polysaccharides. Whilst predominantly found as the L-enantiomer, instances of D-rhamnose are also found in nature, particularly in the Pseudomonads bacteria. Interestingly, rhamnose is notably absent from humans and other animals, which poses unique opportunities for drug discovery targeted towards rhamnose utilising enzymes from pathogenic bacteria. Whilst the biosynthesis of nucleotide-activated rhamnose (NDP-rhamnose) is well studied, the study of rhamnosyltransferases that synthesise rhamnose-containing glycoconjugates is the current focus amongst the scientific community.

In this review, we describe where rhamnose has been found in nature, as well as what is known about TDP-β-L-rhamnose, UDP-β-L-rhamnose, and GDP-α-D-rhamnose biosynthesis. We then focus on examples of rhamnosyltransferases that have been characterized using both in vivo and in vitro approaches from plants and bacteria, highlighting enzymes where 3D structures have been obtained. The ongoing study of rhamnose and rhamnosyltransferases, in particular in pathogenic organisms, is important to inform future drug discovery projects and vaccine development.

Background

Rhamnose (Rha) is a 6-deoxy hexose sugar found widely throughout nature in plant and microbial natural products (1), including glycoproteins (2) and structural polysaccharides (3). The importance of these glycoconjugates in biological processes leads to rhamnose often being an essential sugar for cell viability in many organisms. Although the L-enantiomer of rhamnose is by far the most abundant throughout nature, D-Rha has been found extensively throughout the Pseudomonas genus (4) as a constituent of their surface exposed lipopolysaccharide (5). Animal production of biological compounds based on L-rhamnose have not been found. Thus, the microbial biosynthesis pathway for L-Rha production has

This research was funded in whole by the Wellcome Trust [Grant number 109357/Z/15/Z and 105606/Z/14/Z]. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.
previously been targeted as a potential therapeutic drug target to fight bacterial infections (3, 6).

The biosynthesis pathways of the nucleotide-activated forms of \( \text{L-} \) and \( \text{D-Rha} \) differ depending on the enantiomer and the domain of life. \( \text{L-Rha} \) is biosynthesized as two variants, thymidine 5’-diphospho-\( \beta \)-L-rhamnose (TDP-\( \beta \)-L-rhamnose) or uridine 5’-diphospho-\( \beta \)-L-rhamnose (UDP-\( \beta \)-L-rhamnose). \( \text{D-Rha} \) is much less common and is biosynthesized as GDP-\( \alpha \)-D-rhamnose, using a different set of enzymes to those utilised in the NDP-\( \beta \)-L-rhamnose biosynthesis pathways. Bacteria (7), some microalgae (8), and archaea (9), have been shown to primarily produce TDP-\( \beta \)-L-rhamnose, whilst the biosynthesis of UDP-\( \beta \)-L-rhamnose is mainly observed in plants or other eukaryotes such as fungi and algae, on large, multifunctional enzymes (10, 11). GDP-\( \alpha \)-D-rhamnose has been observed in the \textit{Pseudomonas} genus (12), \textit{Aneurinibacillus thermoaerophilus} (13), and also the giant algal-infecting virus, \textit{Paramecium bursaria Chlorella} virus 1 (14).

Following nucleotidyl-activation, NDP-Rha becomes a substrate for diverse rhamnosyltransferases that transfer rhamnose onto their glycoconjugate products. Many of the functionally characterized rhamnosyltransferases are responsible for the transfer of \( \text{L-Rha} \) to bacterial LPS (15-18), but other rhamnosyltransferases that transfer rhamnose to plant natural products have also been described (19-21). In this work, we review the presence of rhamnose in glycoconjugates across the domains of life and viruses, with a particular focus on Gram-negative and Gram-positive pathogenic bacteria, which are classified by their rhamnose-rich surface carbohydrate virulence factors, where rhamnose is abundant and found with a variety of linkages. We also review the rhamnosyltransferases that have been characterised structurally and highlight some limitations that currently hinder the progress towards the elucidation of other rhamnosyltransferases from a functional and structural perspective.

**Presence of Rhamnose in Nature**

**Eukaryotes**

Although notably absent from animals and mammals, rhamnose plays a crucial role for plants and a number of lower eukaryotes. Arguably, the most prominent rhamnose glycoconjugate found in land plants is the pectic polysaccharide rhamnogalacturonan-I (RG-I). In this glycoconjugate, \( \alpha \)-L-Rha forms the backbone repeating unit of 4)-\( \alpha \)-D-GalpA-(1,2)-\( \alpha \)-L-Rhap-(1, and is the branching point for the sidechains (Figure 1). Rhamnose is therefore an essential sugar to the polysaccharide structure. Pectic polysaccharides like RG-I, along with other polysaccharides, constitute the major component of the plant cell wall, and are crucial for cell wall flexibility and stability, but also essential for growth and plant development (22). As well as RG-I, rhamnose frequently decorates flavonoids such as kaempferol and quercetin, often found in the flowering tissues of plants (23). The glycosylation of these flavonoids is believed to prevent self-toxicity; importantly, when flavonol glycosylation is genetically blocked, some plants down-regulate flavonol aglycone biosynthesis (24).
The presence of rhamnose in fungal glycoconjugates is less common, but the filamentous human pathogen *Pseudallescheria boydii*, as well as the plant pathogen, *Rhynchosporium secalis* (25), produce a distinct rhamnose-containing cell wall polysaccharide termed ‘peptido-rhamno-mannan’, which has been proposed as a potential diagnostic marker due to its rarity (26). Other eukaryotes such as algae are known to biosynthesise nucleotide-activated rhamnose (8) and produce prominent rhamnose-containing polysaccharides. Worth mentioning is the sulphated structural polysaccharide, ulvan, from the green ‘sea lettuce’ of the *Ulva* genus. It consists of either 4)-β-D-GlcA-(1-4)-α-L-Rha-3S-(1 or alternatively 4)-α-D-IdoA-(1-4)-α-L-Rha-3S-(1 repeating units and has been found to display useful pharmaceutical or nutraceutical properties (27). It has even been recommended as a food source to improve the egg laying ability of hens (28). Rhamnose has also been identified to be a major component of other algal polysaccharides, but their chemical structures and carbohydrate compositions have not been yet elucidated (29, 30). Rhamnose and rhamnose biosynthesis genes and subsequent enzymes are absent in higher eukaryotes, including mammals. This uniqueness suggests that these family of enzymes that are mechanistically distinct to any proteins found in the hosts are potential drug targets in human and veterinary pathogens.

**Prokaryotes**

Whilst less common in eukaryotes, rhamnose appears to be ubiquitously incorporated into surface carbohydrates amongst both Gram-negative and Gram-positive bacteria. The Gram-negative O-antigen is often the exclusive component that incorporates rhamnose, with structural diversity arising from the linkage types, repeating units, and the substitution of the core backbone. As well as being found in the *E. coli* O-antigens, L-Rha makes up the major component of the surface carbohydrates in the pathogenic *Shigella* O-antigens. *Shigella* are composed of four subgroups (*S. flexneri, S. sonnei, S. dysenteriae, and S. boydii*), which are further categorised into serotypes; these serotypes are identified based on the carbohydrate structure of the lipopolysaccharide repeat units of the O-antigen. Rhamnose is often the major component of these repeating units and forms the backbone of them, with carbohydrate diversity arising in the glycosidic linkage type and substitutions (31) (Figure 1).

Like *Shigella*, the *Streptococcus* genus is grouped based on the carbohydrate composition of the antigens found on their cell wall and its ability to trigger unique immune responses. This grouping, first described by Rebecca Lancefield in 1933 (32), has prompted the study of the chemical structure and composition of numerous *Streptococcal* Lancefield antigens, which have since been found to contain rhamnose as their major substituent. These rhamnose-rich surface polysaccharides are currently being studied as a potential vaccine candidate against Group A Streptococcus (GAS), both by academics and the pharmaceutical industry (33-35). As well as species from GAS such as *Streptococcus pyogenes*, the *streptococcus* genus contains important pathogens to humans and animals, with GAS being the causative agent for a variety of diseases such as streptococcal pharyngitis (strep throat), scarlet fever and autoimmune disease such as rheumatic heart disease (36). Group B Streptococcus (GBS) is an important human pathogen; although mostly harmless in adults it can be fatal if passed from a pregnant mother to a newborn baby (37). Importantly, GBS is also a bovine
pathogen, similarly to other Lancefield bacteria from groups C, D, F, and G, which are both human and animal pathogens (38). The Group A Carbohydrate (GAC) was the first cell wall carbohydrate reported as a virulence determinant for Group A Streptococcus (33), and similar studies on other Group carbohydrates are eagerly awaited. Some species such as the oral pathogen Streptococcus mutans and other Streptococci from the viridans group do not belong to a conventional Lancefield group due to the lack of a group specific antigen. Nonetheless, these strains also produce cell wall polysaccharides similar to the Lancefield Streptococci that are rich in rhamnose and have been found to be vital for pathogenesis (39) (Figure 1). In addition to the observations made for S. mutans, previous work has emphasized the crucial role of the rhamnose-containing Lancefield antigens from GAS and GBS in pathogenesis (33, 40). Work is ongoing to unravel the complex structures behind some of these carbohydrates, but GAS has previously been shown to contain a rhamnose backbone of alternating α1-2, α1-3 linkages, periodically decorated with N-acetylglucosamine residues (41, 42); more recently it has been found to contain periodic decorations of glycerol-phosphate on the side chains of the carbohydrate GlcNAc (43) (Figure 1). The same negative charge decoration has also been observed in SCC and was proposed to be present in many more Streptococci (41). The Group B Carbohydrate (GBC) is unique from other streptococcal Lancefield antigens as it forms a multiantenna branched structure, with an adapter region to link it to the peptidoglycan layer (44, 45). Importantly, the GBC, like other Lancefield antigens, are rich in rhamnose in both the adapter region and repeat unit of the carbohydrate (Figure 1). Moreover, it was recently shown that the majority of streptococcal rhamnose polysaccharides from almost all Lancefield groups are synthesised via a conserved GlcNAc-rhamnose initiator disaccharide (17). Most annotated rhamnosyltransferase genes and the TDP-rhamnose biosynthesis genes have been reported as being essential in Group A, B and C Streptococcus (33, 46).

Another genus of bacteria found to contain rich rhamnose cell wall polysaccharides is the gram-positive Lactobacillus genus. Famous for their ability to convert sugars into lactic acid, the Lactobacilli spp. form a major component of a healthy gut microbiota (47). Lactobacillus casei BL23, a species used extensively for fermentation purposes, has recently been found to contain a rhamnose rich cell wall polysaccharide structurally related in parts to Streptococcal Lancefield antigens, however synthesised in heptasaccharide repeat units, which are branched on the rhamnose sugars. The GalNAc-rhamnose backbone is decorated with glucose or n-acetylglucosamine side chains forming the core heptasaccharide repeating unit (48). L. lactis has also been found to produce a cell wall polysaccharide containing rhamnose (49), and interestingly the biosynthesis gene cluster for cell wall polysaccharide biogenesis in L. lactis has been shown to share similarities with the Lancefield antigen gene clusters from Streptococcus (3). Each of these gene clusters starts with one or more of the biosynthetic genes for TDP-α-L-rhamnose biosynthesis, highlighting the importance of rhamnose in these diverse cell wall polysaccharides (3). The most recently characterised rhamnose containing polysaccharide is the enterococcal polysaccharide antigen from vancomycin resistant Enterococcus faecalis (50, 51). The reported gene clusters contain in fact three pathways, i) TDP-Rha biosynthesis, ii) the rhamnan backbone biosynthesis cluster and iii) repeat unit biosynthesis.

Several bacteria contain the rhamnose-rich S-layer, an array of abundant proteins or glycoproteins found on the surface (S-layer) of walled bacteria and archaeal envelopes. As
with other post-translational modifications, the proteins found in the S-layer are glycosylated (typically N-glycosylation) to add a layer of further complexity to the proteinaceous structure. Rhamnose has been found as a component of several bacterial S-layers, for instance as polyrhamnan in the S-layer of *Geobacillus stearotherophilus* (2, 52). Another example is that rhamnose is incorporated into the complex N-glycan structure from *Paenibacillus alvei* (53) (Figure 1). Although less widespread, rhamnose has also been found in the S-layer of some archaea as a response to different growth conditions (54).

Whilst L-Rha is much more common across all domains of life than the D-enantiomer, the Gram-negative plant, animal, and human pathogen, *Pseudomonas aeruginosa*, produces a common polysaccharide antigen on its LPS made up solely of D-Rha residues (55). These antigenic regions have been found to contain an average of ~70 D-Rha residues per molecule, making them a major component of the *P. aeruginosa* cell surface. Previous work has shown that *P. aeruginosa* mutants that lack the ability to produce GDP-D-rhamnose are significantly less efficient at adhering to human bronchial epithelial cells, suggesting that these large D-Rha rich antigenic common polysaccharide antigens may play an important role in *P. aeruginosa* adherence to tissues (56).

**Viruses**

Although the glycobiology of viruses is less understood than eukaryotes or bacteria, with many viruses hijacking host glycosylation machinery for capsid glycosylation, it has recently been shown that viruses do produce rhamnose glycoconjugates. The giant Chlorella virus, *Paramecium bursaria* chlorovirus PBCV-1, synthesises highly complex N-glycans decorated on the major capsid protein that contain both L- and D-Rha (57, 58) and are the first reported viral species that produce both enantiomers (Figure 1). Mimivirus has also been found to contain high amounts of L-Rha in its fibers (59). Interestingly, unlike other viruses, giant viruses such as mimivirus and PBCV-1 appear to encode most of the glycosylation machinery in their large genomes for autonomous glycosylation. PBCV-1 has been found to encode a functional GDP-L-fucose biosynthesis pathway, which also produces the rare GDP-D-rhamnose as a major product (14), and mimivirus has been shown to contain a biosynthesis pathway for TDP-L-rhamnose (60).

**NDP-Rhamnose biosynthesis**

The biosynthesis of NDP-β-L-rhamnose has been studied extensively in bacteria and plants, and more recently microalgae and giant viruses have even had their NDP-β-L-rhamnose biosynthesis pathways discovered (8, 60). Production of nucleotide-activated L-Rha arises from either the Rml pathway (commonly found in bacteria) or the NRS/ER or RHM pathways (commonly found in plants and other eukaryotes) (Fig. 2). Whilst there are examples of some enzyme promiscuity *in vitro* in each pathway (8, 10, 61), the Rml pathway primarily produces TDP-β-L-rhamnose and the NRS/ER or RHM pathway primarily produces UDP-β-L-rhamnose (Figure 2).
TDP-β-L-rhamnose biosynthesis

TDP-β-L-rhamnose is the most commonly found rhamnose nucleotide sugar and is ubiquitous amongst gram-negative and gram-positive bacteria, where it serves as a common component of the O-antigen of lipopolysaccharides, as previously discussed. Biosynthesis of TDP-β-L-rhamnose commonly follows a four-step route, starting with the production of TDP-α-D-glucose by the thymidylyltransferase RmlA. The three other chemical modifications of the substrate TDP-α-D-glucose include 1) dehydration (RmlB – 4,6 dehydratase), 2) epimerization (RmlC – 3,5-epimerase), and subsequent 3) reduction (RmlD – 4-reductase) to produce TDP-β-L-rhamnose. Interestingly, the four Rml genes (rmlA-D) that encode the complete pathway are often found clustered in bacteria. Conversely, the Streptococci from more than 40 different species, including Group A, C and G Streptococci, do not cluster rmlA-D (62). *Streptococcus pyogenes* contains RmlD (GacA) in a separate cluster to the rest of the Rml pathway, at the start of the cluster responsible for Lancefield antigen biosynthesis (62, 63). Van der Beek *et al.*, showed that this novel monomeric homolog of RmlID is essential for Group A Streptococcus, and thus represents a good therapeutic target for drug discovery (63). Homologs of RmlB, RmlC and RmlD have been studied for many years, with all having had their crystal structures solved (64-66). Van der Beek also targeted recombinantly expressed and purified RmlBC and GacA (RmlD) from *Streptococcus pyogenes* in a small compound inhibitor screen, highlighting one compound that inhibited both the growth of *S. pyogenes* but also the rhamnose-dependant *Mycobacterium tuberculosis* (62). Alphey *et al.*, have previously targeted purified RmlA from *Pseudomonas aeruginosa* for inhibitor identification, using a combination of high-throughput screening, and subsequent optimization using protein crystallography, in silico screening, and chemical synthesis. This vigorous approach yielded nanomolar inhibitors of *P. aeruginosa* RmlA (67).

UDP-β-L-rhamnose biosynthesis

Research into UDP-β-L-rhamnose biosynthesis in eukaryotes is more recent, but the multidomain *Arabidopsis thaliana* enzyme, RHM, has drawn particular biotechnological attention for its ability to convert as a single protein UDP-α-D-glucose to UDP-β-L-rhamnose, via the three common enzymatic steps: dehydration, epimerization, and reduction of the Rml pathway (11, 68) (Figure 2). Pei *et al.*, fused an RHM homolog from *Vitis vinifera* to a bifunctional NRS/ER enzyme from *A. thaliana*, resulting in a self-sufficient NADPH-independent enzyme able to convert UDP-α-D-glucose to UDP-β-L-rhamnose (68). As well as the RHM pathway, *A. thaliana* encodes for the NRS/ER pathway, where the three conserved reaction steps are conducted by two enzymes only (10). This pathway is also found in other eukaryotes such as fungi (61), where it is essential for pathogenicity in some species (69). The NRS/ER pathway is also present in giant viruses (60), where both L- and D- enantiomers of rhamnose are found in complex N-glycans (57). Whilst most species of algae appear to encode for either the RHM or NRS/ER pathway, depending on their eukaryotic origin, Wagstaff *et al.* recently reported a novel RmlC/RmlD fusion protein in the Haptophyceae and Alveolata phyla. The authors suggested that this chimera was likely the result of horizontal and endosymbiotic gene transfer events during plastid evolution (8).

GDP-α-D-rhamnose biosynthesis
Unlike L-Rha which is activated with either uridine or thymidine, D-Rha is biosynthesized as guanidine 5’-diphospho-α-D-rhamnose (GDP-α-D-rhamnose) by the actions of a GDP-α-D-mannose 4,6-dehydratase (GMD) (14, 70) and a GDP-6-deoxy-α-lyxo-4-hexulose reductase (13) (Figure 2). The presence of D-Rha and the responsible biosynthetic pathway are much less common across all domains of life. Like the L-Rha pathways, the sugar ring is first dehydrated by the enzyme GDP-mannose 4,6-dehydratase, before undergoing subsequent stereoselective reduction at position 4 to produce GDP-α-D-rhamnose. This pathway has evolved to show some level of promiscuity in the end product, with reduction to the alternative stereoisomer resulting in GDP-α-D-talose, or an additional 3,5-epimerisation step before reduction resulting in GDP-β-L-fucose. GDP-α-D-mannose 4,6-dehydratases (GMD) have been found in many organisms due to their ability to produce the precursors for the biosynthesis of GDP-α-D-rhamnose, GDP-α-D-talose, and GDP-β-L-fucose. However, the characterized GDP-4-keto-rhamnose reductases (Rmd) that produce specifically GDP-α-D-rhamnose are less conserved but have been reported in Anuerinibacillus thermoaerophilus and Pseudomonas aeruginosa (12, 13), which produce D-Rha as a major substituent of their respective LPS, as discussed previously. Interestingly, a GMD enzyme from Paramecium bursaria Chlorella virus 1 has been shown to have both 4,6-dehydratase and 4-reductase activity, leading to a promiscuous production of GDP-α-D-rhamnose and GDP-β-L-fucose from the same enzymatic pathway (14).

**Rhamnosyltransferases**

Rhamnosyltransferases are the catalytic enzymes that are responsible for the transfer of rhamnose to the target glycoconjugate. Using nucleotide-activated rhamnose (NDP-Rha) as a sugar donor, rhamnosyltransferases can transfer one or multiple rhamnose residues to a diverse range of acceptor molecules. These range from small natural products such as plant flavonoids, to much larger polysaccharides within bacterial LPS (Table 1). Recombinant production and purification of glycosyltransferases are often challenging, due to their nature of being commonly membrane associated, multidomain proteins. Furthermore, the acceptor substrates are often chemically unusual, and the enzymes require elaborate substrates that cannot be easily mimicked or synthesized for *in vitro* studies. Therefore, a lot of work that has been reported on functionally characterized rhamnosyltransferases involves the genetic manipulation of the host organism, and subsequent metabolite profiling to identify changes in levels of rhamnosylation of target products.

**Eukaryotic rhamnosyltransferases characterized using genetic manipulation**

The pectic polysaccharide RG-I is arguably the most notorious rhamnose containing cell wall polysaccharide in land plants. Recently, Takenaka *et al.*, were able to functionally characterize the rhamnosyltransferases from Arabidopsis thaliana that are responsible for the transfer of rhamnose from UDP-ß-L-rhamnose to RG-I oligosaccharides (RRTs) (94). In this work, the authors found that RRT1, which is one of four rhamnosyltransferases responsible for generation of RG-I, is a single-spanning transmembrane protein that is highly expressed during the formation of seed coat mucilage – a specialized cell wall component
abundant with RG-I. In agreement with its proposed function, inactive mutants of RRT1 resulted in a reduction in levels of RG-I in seed coat mucilage.

Other rhamnosyltransferases from plants that have been analysed using genetic manipulation approaches include the β-solaneine/β-chaconine rhamnosyltransferase, SGT3, from Solanum tuberosum (20). Steroidal glycoalkaloids are often harmful speciality metabolites found in Solanaceous plants and include α-solaneine and α-chaconine, two triglycosylated alkaloids that accumulate in potato tubers. Using antisense downregulation of Sgt3, McCue et al., were able to show that SGT3 is responsible for the transfer of rhamnose to the first galactose or glucose residue of β-solaneine or β-chaconine, respectively (20). Downregulation of Sgt3 resulted in accumulation of both β-solaneine and β-chaconine in tubers, with a concomitant reduction in levels of α-solaneine and α-chaconine.

Prokaryotic rhamnosyltransferases characterized using genetic manipulation

Pseudomonas produce glycoconjugates containing both L- and D-Rha, including virulence factors such as extracellular rhamnolipids that contain one (mono-rhamnolipid) or two (di-rhamnolipid) L-Rha residues. Rhamnosyltransferase 1 (RhlAB) is known to transfer the first rhamnose residue from TDP-β-L-rhamnose to β-hydroxydecanoyl-β-hydroxydecanoate. Rahim et al. characterized the function of the rhlC gene product as the second rhamnosyltransferase, responsible for the generation of di-rhamnolipid (85). In this work the authors were able to generate a chromosomal mutant of rhlC which was subsequently unable to produce di-rhamnolipid as judged by TLC analysis of mono- and di-rhamnolipid in the culture supernatants. The authors also showed that this mutant did not affect the levels of mono-rhamnolipid, but the rhlA null mutant was unable to produce both mono- and di-rhamnolipid. Interestingly, levels of the di-rhamnolipid appeared to increase in the stationary phase of growth, suggesting a possible role for it in allelopathy.

The majority of Streptococci produce rhamnose rich lipo-oligosaccharides on their cell-surface as either non-capsular Lancefield antigens or capsule polysaccharides (32). Although recent work has begun to biochemically characterize enzymes responsible for the biosynthesis of these rhamnose-rich carbohydrates (17, 43), the identity of three important rhamnosyltransferases was uncovered over a decade ago for Streptococcus mutans (87). In this work, Shibata et al., were able to analyse LPS profiles of E. coli strains harbouring various combinations of genes from the S. mutans rhamnose-glucose polysaccharide (RGP) pathway. Using Tricine-SDS-PAGE analysis, they showed that an E. coli transformant harbouring rgpA was able to produce a lipooligosaccharide with an additional rhamnose residue and were also able to use similar experiments to show that rgpB and rgpF can further rhamnosylate these lipooligosaccharide samples. Based on these early findings, Shibata et al., were first able to propose that RgpA, RgpB, and RgpF, utilize TDP-rhamnose for rhamnose polysaccharide biosynthesis.

Eukaryotic rhamnosyltransferases characterized \textit{in vitro}

Possibly the first example of successful expression, purification, and biochemical characterization of a rhamnosyltransferase was reported in 1991. Bar-Peled et al. homogenously purified a flavanone-7-O-glucoside-2’’-O-rhamnosyltransferase responsible
for the rhamnosylation of the bitter flavanone glucosides, naringin and neo-hesperidin (97). In this work, Bar-Peled et al. underwent a thorough column chromatography strategy to purify the rhamnosyltransferase >2700 fold from extracts of young pumelo leaves. The purified enzyme was able to selectively transfer rhamnose from UDP-ß-L-rhamnose to flavonoid-7-O-glicosides, but interestingly not to 5-O-glicosides or flavonoid aglycones. The authors then performed a detailed biochemical analysis, showing that the enzyme only rhamnosylates with the donor substrate UDP-ß-L-rhamnose, and detailing the relative activities for a number of other acceptor substrates. Further work from Bar-Peled and co-workers, published in 2004 with improved genomic data, allowed the identification of the Cm1,2RhaT gene from pumelo that was responsible for this rhamnosyltransferase activity observed first in 1991 (19). In this work, Frydman et al. showed that the 1,2-rhamnosyltransferase enzyme responsible for the bitter taste of species such as grapefruit and pumelo, is only found in neohesperidoside-producing citrus species. Using RT-PCR they revealed that the expression of this gene is significantly upregulated in young tissue, which is in agreement with the developmental pattern of flavanone-neohesperidosides accumulating in younger tissues.

Flavanol glycosides are one of the most prominent natural product classes found in the model plant species, Arabidopsis thaliana. Despite this abundant present in nature, no flavonoid glycosyltransferases had been biochemically characterized from A. thaliana until 2003. Jones et al. conducted an in planta and in vitro approach, to characterize both a UDP-rhamnose:flavonol-3-O-rhamnosyltransferase and UDP-glucose:flavonol-3-O-glucoside-7-O-glycosyltransferase (80). In this work, the authors used LC-MS to analyse the presence or absence of rhamnosylated/glicosylated glycosides in the mutant lines ugt78D1 and ugt73C6, proposing that ugt78D1 was a rhamnosyltransferase responsible for the formation of quercetin-3-O-rhamnoside. The authors also recombinantly expressed UGT78D1 in E. coli as a GST-fusion, but this enzyme lost activity rapidly. Therefore, subsequent assays were performed with crude extracts of E. coli harbouring the rhamnosyltransferase-containing plasmid. Mixtures of recombinant UGT78D1 combined with a UDP-rhamnose preparation, and the acceptor compound quercetin resulted in production of quercetin-3-O-rhamnoside, in agreement with their in planta data that UGT78D1 is flavonol-3-O-rhamnosyltransferase.

Soybean (Glycine max) is another plant containing an abundance of flavonol glycosides. Like Jones et al., in 2014 Rodas et al. took both an in planta and in vitro approach in order to functionally characterize a flavonol 3-O-glucoside-1,6-rhamnosyltransferase (82). In this work the authors used HPLC to investigate the flavonol glycosides produced by a number of recombinant inbred lines, which were developed by crossing Kitakomachi and Koganejiro cultivars of soybean. The authors found that the flavonol glycosides of Koganejiro were rhamnosylated, whereas flavonol glycosides of Kitakomachi did not contain rhamnose. Using this methodology, the authors identified a candidate gene for this specific rhamnosyltransferase activity, designated GmF3G6"Rt (UGT79A6). An in vitro assay was developed for recombinantly produced UGT79A6 and subsequent HPLC and MS analysis of the reaction products confirmed that UGT79A6 was a rhamnosyltransferase, which produced kaempferol 3-O-rutinoside (rhamnosylated) utilising kaempferol 3-O-glucoside and UDP-rhamnose. The authors showed that the activity was specific for UDP-rhamnose, as UGT79A6 could not utilise UDP-glucose, UDP-galactose, or UDP-glucuronic acid as sugar-donors.
To date, only a single rhamnosyltransferase from eukaryotes has been crystallised; the flavonol α-7-O-L-rhamnosyltransferase, UGT89C1, from *A. thaliana*. Zong *et al.*, recently reported the crystal structure of this plant rhamnosyltransferase in complex with the sugar-donor, UDP-rhamnose, and the flavonol acceptor, quecetin (98). Having previously been characterized as a flavonol-7-O-L-rhamnosyltransferase based on a combination of *in planta* and *in vitro* work (21), Zong and co-workers used mutagenesis studies to shed detailed insight into key amino acid residues for sugar donor recognition and specificity for UDP-rhamnose. The mutant H357Q, in particular, was shown to exhibit activity with both UDP-rhamnose and UDP-glucose. Discovery of key active site amino acids, like this histidine sidechain, are important for the bioinformatic-based prediction of whether a glycosyltransferase may be a rhamnosyltransferase or have other sugar-donor specificity. Interestingly, of the six rhamnosyltransferases that have been crystallised to date, all display a GT-B fold (Table 1, Figure 3).

**Prokaryotic rhamnosyltransferases characterized *in vitro***

The first crystal structure of a rhamnosyltransferase was reported by Steiner *et al.*, in 2010, when they successfully solved the structure of the β-1,2-rhamnosyltransferase WsaF from *Geobacillus stearothermophilus* (52) (Figure 3). WsaF is responsible, along with WsaE, for the generation of a trirhamnan repeating unit of the polyrhamnan that decorates proteins in the S-layer in *G. stearothermophilus*. Steiner *et al.* revealed through recombinant expression in 2008, that WsaE is a trifunctional rhamnosyltransferase responsible for both methylation and generation of independent α-1,2 and α-1,3-linked rhamnose residues in the growing polyrhamnan chain (88). Using a clever design of truncated enzyme constructs that included or excluded these individual domains, the authors were able to use TLC analysis with an octyl-linked oligosaccharide acceptor to show that both glycosyltransferase domains were needed for the full activity of the enzyme, and that the C-terminal glycosyltransferase was responsible for the α-1,2-rhamnosyltransferase activity. In the same work, the authors also recombinantly expressed and purified WsaF. Using an octyl-linked oligosaccharide and TLC analysis, they showed that WsaF is a rhamnosyltransferase that adds a single rhamnose to the growing chain. Using NMR spectroscopy, the authors were able to show that the new rhamnose introduced by WsaF was of a β-1,2-linkage, and that the additional rhamnosyltransferase domain of WsaE was transferring rhamnose in an α-1,3-linkage. Alongside WsaE and WsaF, Steiner *et al.* used a similar approach to characterize two further rhamnosyltransferases, WsaC and WsaD. The authors found that these two α-1,3-rhamnosyltransferases were responsible for the addition of a single rhamnose residue each to the adaptor saccharide region (88).

Just two years after the first crystal structure of WsaF was published, the crystal structure of another rhamnosyltransferase was revealed (99). SpnG is a rhamnosyltransferase responsible for the final tailoring of the spinosyns A and D. The spinosyns are complex polyketides produced by the actinomycete *Saccharopolyspora spinose* and make up the active ingredient of pesticides. SpnG crystallised as a homodimer, as it has been observed for other glycosyltransferases belonging to the GT1 family (102, 103). This crystal structure
highlighted the importance of residues H13 and D316 in SpnG activity, in agreement to previous reports in 2009 (76). The crystal structure also highlighted how SpnG was selective for TDP-sugars over UDP-sugars; the amine of the Asn202 sidechain forms a hydrogen bond with the 3-OH from the thymidine moiety whilst also preventing ribose-containing nucleotides from binding due to clashes with 2′-OH.

N-linked glycosylation is a common posttranslational modification of proteins across all domains of life, which contributes to the correct folding, localization, and signalling of cellular proteins. Typically, proteins are glycosylated with either glucose, or more commonly N-acetylglucosamine. Recently, a novel form of arginine glycosylation with rhamnose has been discovered in *Pseudomonas aeruginosa*, *Pseudomonas putida*, *Shewanella oneidensis*, and *Neisseria meningitidis* (90-93). Known as EarP enzymes, these rhamnosyltransferases have been shown to rhamnosylate translation elongation factor P. This translational modification has been shown to be crucial for the rescue of polyproline-stalled ribosomes, and for the pathogenicity of organisms such as *Pseudomonas aeruginosa* (104). A structural basis for this arginine rhamnosylation was uncovered by Krafczyk and coworkers in 2017 when they crystallised EarP from *Pseudomonas putida* (91). This work has been crucial in identifying important active site residues and catalytic residues such as Asp20. This conserved residue was also described by Sengoku et al. in their report of the crystal structure of the EarP homolog from *Neisseria meningitidis* in 2018 (100). The latest of these EarP rhamnosyltransferases to be crystallised was EarP from *Pseudomonas aeruginosa* by He and co-workers (101). Interestingly, EarP from *Pseudomonas aeruginosa* was shown to display important conformational changes upon TDP-rhamnose and translation elongation factor P binding; with sugar-donor binding enhancing acceptor binding. The continued study of these novel rhamnosyltransferases and future development of drug discovery screens should be a key focus for the upcoming years.

Only the WsaF enzyme from *Geobacillus stearothermophilus* has a retaining mechanism of action, whilst all other reported structural studies reported an inverting mechanism of catalysis, alternating the stereochemistry of the nucleotide sugar donor during catalysis to generate the product (Figure 3). Four of the reported structures are functional dimers, with only WsaF and SpnG, from *Saccharopolyspora spinosa*, being functional homodimers. A major structural difference between the retaining WsaF enzyme and the inverting SpnG, from *Saccharopolyspora spinosa* is that the dimerization interface appears to be on the opposite site (Figure 3). Further structural and functional characterisation is required to reveal if this is a common tertiary structure feature for retaining and inverting rhamnosyltransferase.

In *Mycobacterium tuberculosis*, the conserved gene wbbL encodes an inverting rhamnosyltransferase responsible for the transfer of α-1,3-rhamnose from TDP-β-L-rhamnose to the GlcNAc moiety of decaprenyldiphosphoryl-α-D-GlcNAc. This growing lipo-oligosaccharide chain is then further elongated with galactofuranose and arabinofuranose units before being transferred onto peptidoglycan. Grzegorzewicz et al. were able to recombinantly assay WbbL’s activity by overexpression in *E. coli* and purifying the WbbL-containing membrane fractions (105). Using this methodology, the authors created a plate-based assay using the natural substrate, GlcNAc-PP-decaprenyl, and were able to characterize pH optima, metal dependency, and kinetic parameters for WbbL.
An enzyme with a similar rhamnosyltransferase activity, GacB, has been characterized from the human exclusive pathogen *Streptococcus pyogenes* (Group A Streptococcus). Zorzoli, Meyer *et al.* reported that GacB, like WbbL, is responsible for the rhamnosylation of a GlcNAc-PP-undecaprenyl substrate *in vivo* (17). In Group A Streptococcus, this primed lipo-dissacharide is then further elongated with α-1,3- and α-1,2-rhamnose residues, before having GlcNAc side-chains incorporated that are further decorated with glycerol-phosphate modifications to build the Group A Carbohydrate virulence factor (43). Zorzoli *et al.* genetically incorporated streptococcal polyrhamnose biosynthesis gene clusters into *E. coli* using parts of the gene cluster from *Streptococcus mutans* and used antibodies against the polyrhamnose product as a marker to investigate the effects of manipulating individual genes. Using this strategy, GacB was shown to be essential for polyrhamnose biosynthesis (RhaPS). Using radiolabelling experiments, GacB was proposed to transfer a single rhamnose residue to the growing lipid-linked oligosaccharide, which was confirmed *in vitro* by detergent extracting the GacB enzyme from *E. coli* membrane fractions and assaying using MALDI-TOF. This *in vitro* assay was also used to show that GacB is metal independent, and NMR analysis of the reaction product confirmed that GacB is a novel retaining rhamnosyltransferase producing a β-1,4 linkage, unlike WbbL, which is an α-1,3-rhamnosyltransferase (105). Strikingly, the authors revealed that the GacB enzyme produces the first committed step not only in Group A Streptococcus and *S. mutans*, but also is conserved in the RhaPS biosynthesis gene cluster of Group B, Group C, and Group G Lancefield Streptococci. This conservation and the essentiality of the gacB gene and its homologs has significant implications for the future development of novel streptococcal drug candidates.

**Viral rhamnosyltransferases characterized *in vitro***

Very recently, Speciale *et al.* biochemically characterized the rhamnosyltransferase activity of a multifunctional enzyme from the algal-infecting giant virus, PBCV-1 (58). In this work, the authors showed that A064R contains 3 domains; 2 of which have β-1,4- and α-1,2-rhamnosyltransferase activity, with the third responsible for 2-′O-methyltransferase activity which installs the 2′-O-Me group on the terminal rhamnose residue (58). In this work, the authors used several synthetic lipo-oligosaccharides and constructs of A064R varying in length, together with HPLC analysis, to characterize the activity of each domain. Both rhamnosyltransferases domains were shown to use UDP-β-L-rhamnose as the L-rhamnose nucleotide sugar donor. A064R-D1 was shown to encode the β-1,4-rhamnosyltransferase activity required to install the first rhamnose at the 4′-OH position of the xylose residue. A064R-D2 was shown to encode an α-1,2-rhamnosyltransferase responsible for installing the terminal rhamnose residue onto the previously installed L-rhamnose. After both rhamnose residues are installed, A064R-D3 was shown to methylate the terminal rhamnose at the 2′-OH position. This breakthrough work is, to our knowledge, the first example of a biochemically characterized viral rhamnosyltransferase, and sheds light on how giant viruses have evolved to perform N-glycosylation.

**Concluding remarks and perspectives**

Rhamnose, in both the L- and D-form, is found widespread in plant and microbial natural products and surface glycans. It is particularly prevalent in lipopolysaccharides and surface
carbohydrate in Gram-positive and Gram-negative bacteria. Strikingly, rhamnose and any nucleotide-rhamnose biosynthesis enzymes are entirely absent from animals and higher eukaryotes. This distinction has often led to the proposal to screen for small molecule inhibitors to target those enzymes that synthesise or utilise rhamnose-nucleotide sugars, which are ideal drug targets as they are often essential in pathogenic bacteria and absent in humans. A number of projects have focused on bacterial enzymes involved in the biosynthesis of TDP-β-L-rhamnose, including inhibitor screens and optimization to target successfully the first enzyme of the pathway, RmlA (62, 67). However, it is also noteworthy to highlight that there is no report of any drug discovery programme operated by academia or industry to develop these small molecule inhibitors into lead compounds for drug discovery pipelines. It is commonly known that the financial costs to develop novel antimicrobial drugs is hindering those developments. Researchers and private as well as governmental funding bodies need to continue working together to increase the opportunities to successfully develop a generation of new antimicrobial drugs, including those targeting the human and veterinary absent sugars L- and D-Rha.

The absence of any rhamnose in humans has also been exploited for vaccine development to prevent infections from pathogenic bacteria, including Shigella and the Streptococcus genus. A particularly well investigated glycoconjugate vaccine candidate has been developed over the past two decades for Shigella dysenteriae serotype 1. This vaccine candidate is composed of a carrier protein that carries recombinantly produced Shigella dysenteriae serotype 1 surface polysaccharides, which contain a rhamnose repeat unit (Fig 1). Clinical trials were found safe and immunogenic in a phase I clinical trial (106). Preclinical trials in animals have shown that the Group A Carbohydrate backbone is also a strong immunogenic vaccine candidate and is currently exploited by several research groups and industry (33). This unique opportunity will hopefully benefit the vaccine development of a number of pathogenic bacteria, that produce similar, rhamnose-containing glycoconjugates on their surface.

The biological activity of many rhamnosyltransferases has been revealed and the CAZy database continues to be an invaluable resource for the cataloguing of glycosyltransferases. The current difficulties in characterizing the rhamnosyltransferase function in vitro are often linked to the biological and physiochemical properties of the enzymes, in particular the membrane association due to the lipid-linked nature of their acceptor substrates. These features significantly reduce the chances to crystallize a full length or functional rhamnosyltransferase. Another major limitation is the access to affordable quantities of TDP- or UDP-rhamnose. In addition, radiolabelled derivatives of TDP- or UDP-rhamnose are not available – in comparison to other nucleotide sugars - challenging the researchers to develop their bespoke analytical systems and tools. The huge diversity of the acceptor substrates, in particular with regards to the lipid linked carbohydrates, is a significant obstacle to reveal functional insights into the enzymes under study. These tools need to be continuously developed and made widely accessible to study rhamnosyltransferases. These current constraints result in fewer studies conducted on these classes of enzymes and the majority of studies has relied on genetic manipulation of host organisms or semi-crude purifications of the target enzyme. Only half a dozen of rhamnosyltransferases have been purified to homogeneity and crystallised. As more structures of rhamnosyltransferase families become available, detailed insights into their biochemistry will be revealed and can
be exploited to develop better tools to study related rhamnosyltransferases. This progress and development will ultimately benefit the discovery of the first inhibitors to target pathogenic bacteria via their rhamnosyltransferases.

References


Figure 1 – Examples of structurally characterised rhamnose-containing glycoconjugates found across nature. The pectic polysaccharide in plant cell walls, RG-I; O-antigens of Shigella dysenteriae Type I and Shigella flexneri 5a (RU = repeat units); Polyrhamnan from the S-layer glycan of G. stearothermophilus NRS 2004/3a and an S-layer glycan of Paenibacillus alvei CCM 2051; Lancefield antigens from Group A and B Streptococcus, Streptococcus mutans serotype C; N-glycan from Chlorovirus PBCV-1.
Figure 2 – Biosynthesis pathways of nucleotide activated L- and D-Rha. NDP-β-L-rhamnose is the most common form of activated rhamnose in nature. Successive dehydration, epimerization, and reduction steps are either conducted via three enzymes (bacteria, archaea, phages and some microalgae - RmlB to RmlD), or by a single or two enzymes (plants, fungi, viruses, and other eukaryotic algae - NRS/ER pathway or RHM). GDP-α-D-rhamnose biosynthesis is less prevalent and has only been reported in Aneurinibacillus, chlorovirus PBCV-1, and Pseudomonas, where D-Rha is widespread throughout the genus (71).
Figure 3 – Timeline of rhamnosyltransferase crystal structures deposited to the PDB database with their PDB IDs. Dark domains refer to the N-terminal sequence, whilst bright coloured domains highlight the C-terminal domain. (a) crystal structure of the 5′,1,2-rhamnosyltransferase, WsaF, from Geobacillus stearothermophilus (52); (b) crystal structure of Spinosyn 9-O-α-L-rhamnosyltransferase, SpnG, from Saccharopolyspora spinosa (99); (c) crystal structure of arginine α-L-rhamnosyltransferase, EarP, from Pseudomonas putida (91); (d) crystal structure of arginine α-L-rhamnosyltransferase, EarP, from Neisseria meningitidis (100); (e) crystal structure of arginine α-L-rhamnosyltransferase, EarP, from Pseudomonas aeruginosa (101); (f) crystal structure of the flavonol α-7-O-L-rhamnosyltransferase, UGT89C1, from Arabidopsis thaliana (98).
<table>
<thead>
<tr>
<th>Family</th>
<th>Name</th>
<th>Activity</th>
<th>Host organism</th>
<th>Genbank</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROKARYOTE</td>
<td>GT1</td>
<td>TiaG1</td>
<td>tiacumin 5-C-methyl-o-rhamnosyltransferase</td>
<td>Doctylosphorangium aurantiacum</td>
<td>ADU86026.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>TiaG2</td>
<td>tiacumin 2-O-methyl-o-rhamnosyltransferase</td>
<td>Doctylosphorangium aurantiacum</td>
<td>ADU85989.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>RtfA</td>
<td>GPL/6-deoxytalose α-1,1,2-rhamnosyltransferase A</td>
<td>Mycobacterium avium</td>
<td>AAD44209.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>GtfA</td>
<td>GPL/3,4-di-O-Me-rhamnose α-1,2-(3-O-methyl)-rhamnosyltransferase</td>
<td>Mycobacterium smegmatis</td>
<td>AAN28688.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>Rv2962c</td>
<td>p-hydroxybenzoic acid ethyl ester rhamnosyltransferase</td>
<td>Mycobacterium tuberculosis</td>
<td>CAB05415.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>RhlB</td>
<td>rhamnosyltransferase</td>
<td>Pseudomonas aeruginosa</td>
<td>AAG06866.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>SpnG</td>
<td>spinosyn 9-O-α-l-rhamnosyltransferase</td>
<td>Saccharopolyspora spinosa</td>
<td>AAG23268.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>AraGT</td>
<td>α-1-l-rhamnosyltransferase (aranciarcin synthase)</td>
<td>Streptomyces coelicolor</td>
<td>ABL09968.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>EImGT</td>
<td>β-dimethyl-tetracenomycin C α-1-l-rhamnosyltransferase</td>
<td>Streptomyces olivaceus</td>
<td>CAC16413.2</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>StfG</td>
<td>Steffimycin α-1-rhamnosyltransferase</td>
<td>Streptomyces steffisburgensis</td>
<td>CAJ42338.1</td>
</tr>
<tr>
<td>EUKARYOTE</td>
<td>GT1</td>
<td>UGT78D1</td>
<td>flavonol 3-O-rhamnosyltransferase</td>
<td>Arabidopsis thaliana</td>
<td>OAP13716.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>UGT89C1</td>
<td>flavonol α-7-O-l-rhamnosyltransferase</td>
<td>Arabidopsis thaliana</td>
<td>AAF80123.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>C12R1</td>
<td>flavonoid α-1,2-l-rhamnosyltransferase</td>
<td>Citrus maxima</td>
<td>AAI99997.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>UGT79A6</td>
<td>flavonol 3-O-glucoside α-1,6-l-rhamnosyltransferase</td>
<td>Glycine max</td>
<td>BAN91401.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>UGT772</td>
<td>anthocyanidin-3-glucoside α-1-l-rhamnosyltransferase</td>
<td>Petunia x hybrida</td>
<td>CAA50376.1</td>
</tr>
<tr>
<td></td>
<td>GT1</td>
<td>Sgt3</td>
<td>β-solanie/β-chaconine rhamnosyltransferase</td>
<td>Solanum tuberosum</td>
<td>ABB84472.1</td>
</tr>
<tr>
<td>PROKARYOTE</td>
<td>GT2</td>
<td>WbbL</td>
<td>α-GlNAc-PP-CN5 α-1,3-l-rhamnosyltransferase</td>
<td>Mycobacterium tuberculosis</td>
<td>AFN51271.1</td>
</tr>
<tr>
<td></td>
<td>GT2</td>
<td>WapR</td>
<td>LPS α-1,3-l-rhamnosyltransferase</td>
<td>Pseudomonas aeruginosa</td>
<td>AAG08385.1</td>
</tr>
<tr>
<td></td>
<td>GT2</td>
<td>RhlC</td>
<td>l-Rhaβ-1-hydroxydecanoylβ-hydroxydecanoate rhamnosyltransferase</td>
<td>Pseudomonas aeruginosa</td>
<td>AAG04519.1</td>
</tr>
<tr>
<td></td>
<td>GT2</td>
<td>MigA</td>
<td>LPS α-1,6-l-rhamnosyltransferase</td>
<td>Pseudomonas aeruginosa</td>
<td>AAG04094.1</td>
</tr>
<tr>
<td></td>
<td>GT2</td>
<td>RfbF</td>
<td>l-Rha α-1,2-l-rhamnosyltransferase</td>
<td>Shigella flexneri O2A</td>
<td>CAA50772.1</td>
</tr>
<tr>
<td></td>
<td>GT2</td>
<td>RgpB</td>
<td>α-1,3-rhamnosyltransferase*</td>
<td>Streptococcus mutans</td>
<td>Q840W4</td>
</tr>
<tr>
<td></td>
<td>GT2</td>
<td>WsaC</td>
<td>α-1,3-rhamnosyltransferase</td>
<td>Geobacillus stearotherophilus</td>
<td>Q7BS65</td>
</tr>
<tr>
<td></td>
<td>GT2</td>
<td>WsaD</td>
<td>α-1,3-rhamnosyltransferase</td>
<td>Geobacillus stearotherophilus</td>
<td>Q7BS65</td>
</tr>
<tr>
<td></td>
<td>GT2</td>
<td>WsaE</td>
<td>Multifunctional α-1,3- and α-1,2-rhamnosyltransferase</td>
<td>Geobacillus stearotherophilus</td>
<td>Q7BS65</td>
</tr>
<tr>
<td>PROKARYOTE</td>
<td>GT4</td>
<td>RgpA</td>
<td>α-1-GlNAc-β-1,4-rhamnosyltransferase*</td>
<td>Streptococcus mutans</td>
<td>Q840W5</td>
</tr>
<tr>
<td></td>
<td>GT4</td>
<td>WsaF</td>
<td>β-1,2-rhamnosyltransferase</td>
<td>Geobacillus stearotherophilus</td>
<td>Q7BSG50</td>
</tr>
<tr>
<td></td>
<td>GT4</td>
<td>GacB</td>
<td>α-1-GlNAc-β-1,4-rhamnosyltransferase</td>
<td>Streptococcus pyogenes</td>
<td>N/A</td>
</tr>
<tr>
<td>PROKARYOTE</td>
<td>GT102</td>
<td>WbbB</td>
<td>Multimodular glycosyltransferase catalyzing glycan polymerization and chain termination</td>
<td>Raoultella terrigena</td>
<td>AAQ82931.1</td>
</tr>
<tr>
<td>PROKARYOTE</td>
<td>GT104</td>
<td>EarP</td>
<td>arginine α-1-rhamnosyltransferase</td>
<td>Pseudomonas aeruginosa</td>
<td>AAG06240.1</td>
</tr>
<tr>
<td></td>
<td>GT104</td>
<td>EarP</td>
<td>arginine α-1-rhamnosyltransferase</td>
<td>Pseudomonas putida</td>
<td>AAN67476.1</td>
</tr>
<tr>
<td></td>
<td>GT104</td>
<td>EarP</td>
<td>arginine α-1-rhamnosyltransferase</td>
<td>Shewanella oneidensis</td>
<td>AAN5363.1</td>
</tr>
<tr>
<td></td>
<td>Genotype</td>
<td>Accession</td>
<td>Species/Sequence</td>
<td>Database ID</td>
<td>(Protein ID)</td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
<td>-----------</td>
<td>------------------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td><strong>GT104</strong></td>
<td>EarP</td>
<td>arginine α-L-rhamnosyltransferase</td>
<td>Neisseria meningitidis</td>
<td>BAU19337.1</td>
<td>(93)</td>
</tr>
<tr>
<td><strong>EUKARYOTE</strong></td>
<td>GT106</td>
<td>RRT1</td>
<td>rhamnogalacturonan-I rhamnosyltransferase</td>
<td><em>Arabidopsis thaliana</em></td>
<td>CAC01773.1</td>
</tr>
<tr>
<td></td>
<td>GT106</td>
<td>RRT2</td>
<td>rhamnogalacturonan-I rhamnosyltransferase</td>
<td><em>Arabidopsis thaliana</em></td>
<td>AEE73784.1</td>
</tr>
<tr>
<td></td>
<td>GT106</td>
<td>RRT3</td>
<td>rhamnogalacturonan-I rhamnosyltransferase</td>
<td><em>Arabidopsis thaliana</em></td>
<td>AEC05682.1</td>
</tr>
<tr>
<td></td>
<td>GT106</td>
<td>RRT4</td>
<td>rhamnogalacturonan-I rhamnosyltransferase</td>
<td><em>Arabidopsis thaliana</em></td>
<td>AEE29099.1</td>
</tr>
<tr>
<td><strong>PROKARYOTE</strong></td>
<td>Unknown</td>
<td>RgpF</td>
<td>Multidomain α-1,3- and α-1,2-rhamnosyltransferase*</td>
<td><em>Streptococcus mutans</em></td>
<td>Q840W3</td>
</tr>
<tr>
<td><strong>VIRAL</strong></td>
<td>Unknown</td>
<td>A064R</td>
<td>Multidomain β-1,4- and α-1,2-rhamnosyltransferase</td>
<td>PBCV-1</td>
<td>Q89399</td>
</tr>
</tbody>
</table>