MINI - REVIEW
ADVANCEMENTS OF MICROBIOLOGY
2025, 64, 2, 86-94
https://doi.org/10.2478/am-2025-0008

OVERVIEW OF *AGA GENES* AND THEIR ROLE IN UTILIZATION OF N-ACETYL-D-GALACTOSAMINE AND D-GALACTOSAMINE BY BACTERIA

Leon Petruńko[1*], Dawid Gmiter[2]

[1]Jan Kochanowski University of Kielce, Faculty of Natural Science, Institute of Biology, Uniwersytecka 7, 25-406 Kielce, Poland
[2] University of Warsaw, Faculty of Biology, Institute of Microbiology, I. Miecznikowa 1, 02-096 Warszawa

Submitted in February 2025, accepted in June 2025

Abstract: Bacteria use a variety of mechanisms in order to successfully survive in hosts. To support proliferation sophisticated means of maintenance in host especially metabolism of carbon and nitrogen is needed. One metabolic mechanism is the metabolism of amino sugars which are being used simultaneously as a carbon and nitrogen source. D-galactosamine (GalN) and N-acetyl-D-galactosamine (GalNAc) and their derivatives are widely used by bacteria. For instance, they build parts of cell wall and LPS. The metabolism of both amino sugars is performed and controlled by proteins encoded by *aga* genes, present in a wide variety of *Proteobacteria*. The genetic mechanism underlying this metabolic pathway is, however not yet fully known. Despite this, there is a possibility of using this pathway as a target for therapy, especially in the times of ever-growing danger of bacterial drug resistance. The goal of this article was to present the current knowledge of *aga* genes and their importance in GalN and GalNAc metabolism.

Keywords: *aga genes*, Amino sugars, Antibacterial therapy, Metabolism regulation, Sugar metabolism

1. Introduction

Bacteria are known to be very versatile organisms, able to colonize a variety of environments. One such environment is the human organism, which they colonize and infect opportunistically or non-opportunistically, using their virulence factors, causing a variety of symptoms and diseases. This creates a number of challenges to overcome and treat such infections (Reizer et al. 1996).

However, these infections are possible for the bacteria only due to their mastery of utilizing metabolic pathways, allowing them to be flexible and adapt to the current conditions. Fortunately, we have managed to understand, harness, and apply these bacterial abilities, eg. to design safe and effective strains that generate a wide range of novel biological products. Unfortunately, this features also allows pathogenic bacteria to quickly gain an advantage over our therapy. The most noticeable advantage nowadays being growing bacterial drug resistance, and our relative inability to counter it in clinical conditions (Reizer et al. 1996).

One of the metabolic pathways responsible for creating substrates that are used by bacteria, to develop their virulence factors, is the use of amino sugars which are a variety of complex monosaccharides in which hydroxyl group is replaced by amine group (Reizer et al. 1996). Most of the research on bacterial metabolism of amino sugars comes from studies with *Escherichia coli* (Reizer et al. 1996), *Bacillus subtilis* (Freymond et al. 2006) and *Streptococci* (Afzal et al. 2016). Not enough, however, is known about processes and genetical mechanisms underlying amino sugar metabolism to potentially use them in our advantage.

Metabolism of amino sugars comes in different ways, differing between bacterial species or even strains (Sadler et al. 1979; Brinkkötter et al. 2000; Ray and Larson 2004; Freymond et al. 2006; Leyn et

* Corresponding author: Leon Petruńko, Jan Kochanowski University of Kielce, Faculty of Natural Science, Institute of Biology, Uniwersytecka 7, 25-406 Kielce, Poland, e-mail: leonpetrunko@gmail.com

Cite as:
Overview of *aga genes* and their role in utilization of N-acetyl-D-galactosamine and D-galactosamine by bacteria. Petruńko L. and Gmiter D., ADV MICROBIOL-NY, 2025, 64, 2, 86-94, https://doi.org/10.2478/am-2025-0008

al. 2012). Hence, a vast number of molecular mechanisms controlled by different operon and genes had been developed in the course of evolution. Because of that, at the moment, we are unable to see the whole picture of bacterial amino sugar metabolism (Shen et al. 2025). However, *aga* genes operon responsible for catabolism of such sugars, namely D-galactosamine (GalN) (Fig.1A) and N-acety-D-galactosamine (GalNAc) (Fig.1B) might serve as a good model for future investigations. This operon is widespread, especially in pathogenic bacteria.

In many bacteria, such as *E. coli, Enterobacter* and *Shewanella* spp. an *aga* operon consists mostly of around twelve genes. This number, however, may vary due to other bacteria developing new genes, or adapting the genes responsible for other amino sugar catabolism (mainly Glucosamine) (Ray and Larson 2004; Leyn et al. 2012; Zhang et al. 2015). The characterization of *aga* genes as well as metabolic pathway controlled by proteins encoded by them are the subject of this article. We focused primarily on the mode of action, as well as the role of the *aga* genes in amino sugar metabolism. Finally, we rationalize why these pathways have a great potential as a target for therapy or diagnostics.

2. Role of N-acetyl-D-galactosamine and D-galactosamine in eukaryotic and prokaryotic cells

Both amino sugars are common components and build a variety of cell structures in both eukaryotic and prokaryotic domains. In bacteria, GalNAc is not only a component of the cell wall but is also found in lipopolysaccharide (LPS). For instance, *Pseudomonas aeruginosa* and *Campylobacter jejuni* both have GalNAc residuses in their LPS core, as well as require UDP-GalNAc in order to be able to build full length LPS molecule.(Masoud et al. 1995; Sadovskaya et al. 1998; Bernatchez et al. 2005). Some bacteria even use it to mimic our immune cells (De Jong et al. 2022). In mammals, it links carbohydrate chains in mucins (Carraway and Hull 1991). Both GalNAc and GalN are found in glycosylated proteins of both domains (Sadler et al. 1979; Abu-Qarn et al. 2008). Given these versatile roles we expect GalNAc/GalN metabolism to be closely linked to bacterial virulence (Zhang et al. 2015).

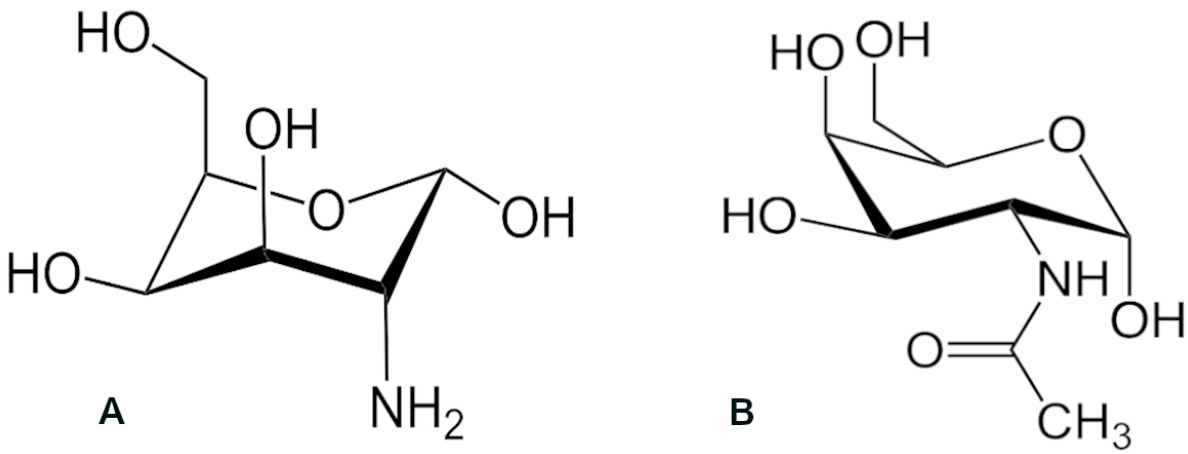

Fig. 1. The chemical structure of glycans. A-D-galactosamine (GalN), B- N-acety-D-galactosamine (GalNAc).

Such abundance and role of these amino sugars makes it no surprise that various mammalian pathogens use them to their advantage. For example, the amino sugars are known to be able to support the growth as carbon and nitrogen sources (Reizer et al. 1996). Brinkköter et al. presented in their study, that Wild-Type strains of *E. coli*, *Klebsiella pneumoniae* and *Salmonella enterica* Typhimurium could effectively use GalNAc and GalN as their main source of carbon and nitrogen. Compared to the laboratory strains, *E. coli* K-12, which only after some time developed mutations which allowed to use this metabolic pathway. These mutants had a phenotype suppressing the *gat* and *nag* genes responsible for galactitol and N-acetyl-D-glucosamine metabolism respectively. They could not however use GalN as their main source of carbon and nitrogen (Brinkkötter et al. 2000). This may suggest that an ability to utilize amino sugars is important for Wild-Type strains to maintain their virulence.

Another example of the usefulness of GalNAc for bacterial virulence may be the case of *Camplylobacter jejuni*, which can produce GalNAc-terminal Lipooligosacharides (LOS). These LOS can bind to the Macrophage Galactose-type lectin found on immature dendritic cells, which successfully mimics our immune system (De Jong et al. 2022).

3. The structure of the *aga* operon

The *aga* operon consists mainly of twelve genes encoding proteins serving various functions regarding GalNAc and GalN metabolism. However, as was already mentioned, there exist several orthologic groups of *aga* genes.

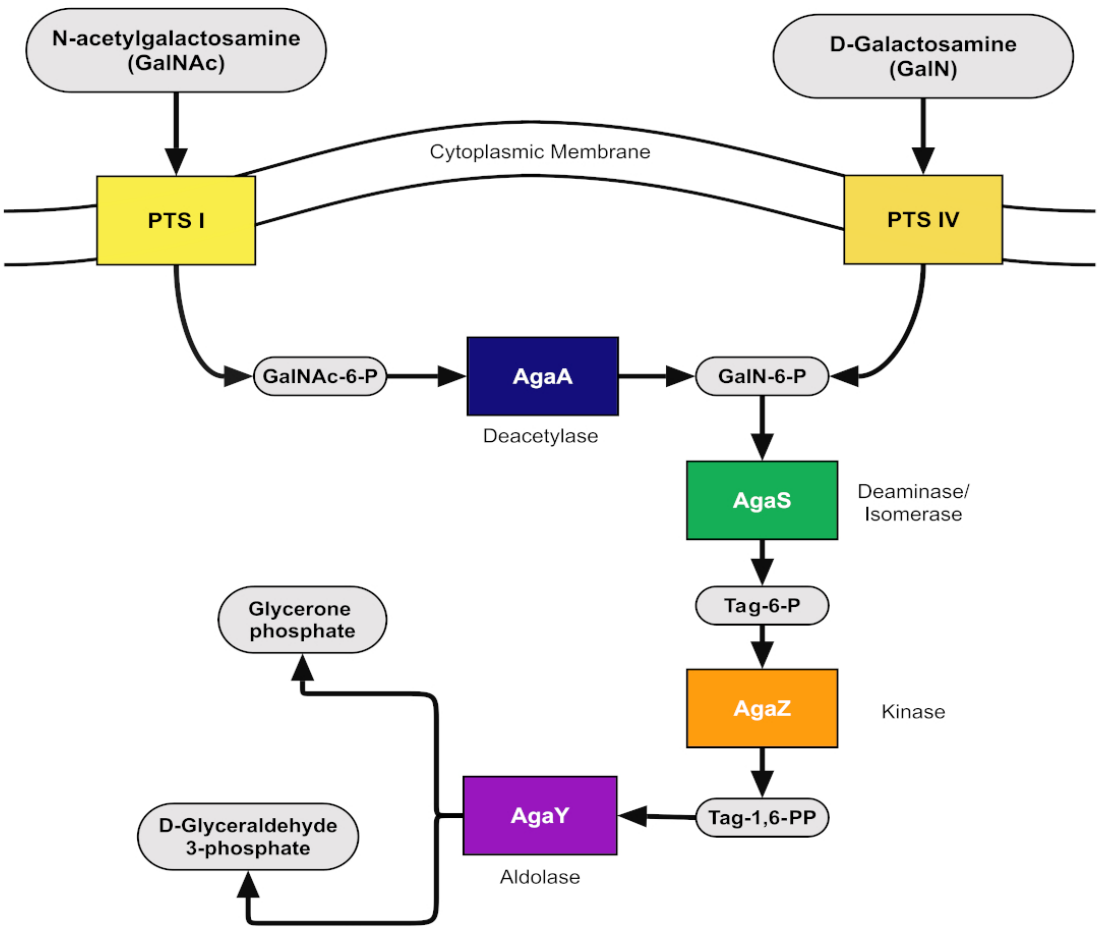

Fig. 2. The catabolic pathway that uses GalNAc and GalN.

The catabolic pathway that uses GalNAc and GalN consists of five steps, as presented in Figure 2. Transport (1) of the substrates is realized by two PTS systems encoded by *agaBCD* and *agaVWEF* genes. Succeeded by the following phosphorylation by AgaK kinase to GalNAc-6-P and GalN-6-P respectively. Subsequent deacetylization (2) of GalNAc-6-P to GalN-6-P catalyzed by AgaA. Deamination and isomerization (3) of GalN-6-P (galactosamine-6-phosphate) by di-functional AgaS enzyme to tagatose-6-phosphate (Tag-6-P) which then undergoes another phosphorylation (4) catalyzed by AgaZ to tagatose-1,6-diphosphate (Tag-1,6-PP). The final step of this GalNAc pathway is Tag-1,6-PP cleavage, catalyzed by AgaY aldolase which leads to the production of glyceraldehyde-3-phosphate and glycerone phosphate (PEP) (Ray and Larson 2004; Leyn et al. 2012).

Interestingly, comparative genomics study by Leyn et al. suggested a vast diversity regarding the utilization of GalNAc/GalN pathway by *Proteobacteria*, such as *Shewanella*. This concerned especially the two first steps, with the latter three being more conserved (Leyn et al. 2012).

The *aga* genes in *Proteobacteria* are regulated by AgaR transcriptomal regulator from DeoR family of transcriptional factors, which recognizes specific sequences located in *agaRZS* promoter regions. Not only does it serve as an autoregulator but also negatively controls the expression of *agaZ* and *agaS* genes, binding to the specific palindromes preceding these genes (Ray and Larson 2004).

4. AgaR – transcriptional regulator of the operon

AgaR is a transcriptional regulator belonging to the DeoR family. Members of this family can be found in a variety of species, ranging from Gram-positive (*Lactococcus lactis*, *Streptococcus mutans*) to Gram-negative bacteria (*E. coli*, *K. pneumoniae, P. aeruginosa*), being present in around 58% of species of *Prokaryota* (Pérez-Rueda and Collado-Vides 2000; Ray and Larson 2004). *E. coli* family of DeoR regulators contains at least fourteen members, which usually function as repressors in sugar metabolism (Elgrably-Weiss et al. 2006). Examples being GutR (glucitol metabolism), LacR (lactose metabolism) FucR and FruR responsible for fucose and fructose respectively, as well as other (Pérez-Rueda and Collado-Vides 2000). Proteins in this family exhibit several common features with high degree of size conservation ranging from 240 to 260 amino acids and highly conserved regions, such as the second helix, known as recognition helix of the helix-turnhelix DNA-binding motif in the N-terminus. The first helix of the DNA-binding domain, however, presents little conservation(Pérez-Rueda and Collado-Vides 2000).

Recent studies by Zhang et al. and Leyn et al. show that there is at least five different orthologs of AgaR regulator characterized by different DNA motifs, discovered throughout 21 species of *Proteobacteria*, some of which even carrying two copies of *agaR* genes. Leyn et al. also observed strong tendencies for *agaR* genes to cluster on the chromosome with GalNAc utilization genes, which suggests conservation of AgaR function. All of the AgaR binding motifs, investigated by comparative genomics approach, shared a common CTTTC pattern with a common consensus of a direct repeat of this sequence. Groups (I), (II) and (III) shared a common CTTTC-5nt-CTTTC consensus, with the copy number and orientation differing between species. In contrast to that, group (IV) had an inverted repeat with the consensus CTTTC-15nt-GAAAG, while group (V) had a common structure with two inverted repeats GAAAG-(16-18)nt-GAAAG (Leyn et al. 2012).

Zhang et al., however, discovered only two, distantly relative, groups of AgaR (R1 and R2), sharing only 30% sequence similarity, as well as proposed different binding site motifs for these two orthologs. In most genomes analyzed using comparative genomics approach, they discovered that *agaR1* had a tendency to cluster on a chromosome close with genes for glycoside hydrolase and PTS, while *agaR2* has a similar tendency for co-localization with the genes encoding deacetylase, isomerase, and aldolase (Zhang et al. 2015). The candidate motifs for binding both regulators also shared a similar palindrome structure. AgaR1 binding sites had a length of 20 bp, whereas AgaR2 binding sites had 18bp. In addition, both AgaR binding motifs had an AT-rich central part with AgaR2-binding motif being more conserved and GC-rich in general (Zhang et al. 2015).

Typically, DeoR binding site is formed by the residues of several amino acids belonging to peripheral subdomains labeled P1 and P2 as well as a double Rossmann fold subdomain at their interface. Although the total sequence identity between different DeoR family regulators may differ even up to 80% difference, the three subdomains forming the binding site have a conservative structure, as well as position and ligand binding orientation (Škerlová et al. 2014). The effector molecule forms a strong covalent linkage with binding sites amino groups, deeply burring itself in the effector-binding site (Škerlová et al. 2014). Škerlová et al. also discovered that C-DeoR was the first discovered bacterial transcriptional regulator, that had its effector module covalently bound (Škerlová et al. 2014), suggesting there can be more DeoR family members expressing this characteristic.

During the study by Ray et al., two promoters for AgaR binding were discovered within *agaR-agaZ* region in *E. coli* K-12, as well as the third one in *agaS-agaA* intergenic region (Ray and Larson 2004). They also discovered that AgaR negatively regulates transcription from each of the three promoters. The absence of this repressor led to growthrate on galactosamine similar to that on glucose without the repressor. In addition to that, *agaZ* promoter was discovered to be a subject not only to repression by AgaR but also catabolite repression. Its expression was discovered to be 10 times higher in cells grown on casamino acids and GalN than on glucose without the functional AgaR. *agaZ* promoter region contained a sequence similar to that recognized by the cAMP-cAMP receptor protein (CRP) complex-(Zhang and Ebright 1990). In mutants with or without the functional *agaR* gene addition of cAMP resulted in 10-fold upregulation (Ray and Larson 2004).

5. AgaS – isomerase right in the heart of the pathway

The most conserved enzyme in GalNAc/GalN pathway discovered by Leyn et al. was AgaS, a hypothetical sugar phosphate isomerase from the SIS family, that was present in all analyzed genomes (Leyn et al. 2012). Enzymes belonging to this family are essential for bacterial proliferation, being often responsible for sugar catabolism, thus regulating general bacterial metabolism, allowing for biogenesis of virulence factors such as LPS, and modifying those factors (for instance altering the O-antigen sugar composition of LPS) as a mean of drug resistance (Gourlay et al. 2010).

The enzymes contain Sugar ISomerase (SIS) domain, responsible for sugar isomerization or sugar binding activities, yet differing with regards to their overall structure and sequences, sometimes even presenting less than 20% sequence similarity (Sommaruga et al. 2009; Gourlay et al. 2010).

Usually, SIS domain in enzymes has a catalytic function as isomerase and binds to phosphorylated sugars (Bateman 1999), in case of AgaS being GalN-6-P. The domain is also present in a family of bacterial transcriptional regulators, such as the ribose-phosphate-isomerase regulator RpiR, which regulates the *rpiB* gene (Sørensen and Hove-Jensen 1996). However, there are no reports of SIS domain regulators demonstrating catabolic activity yet (Bateman 1999).

The deaminase/isomerase activity of AgaS was evaluated by Leyn et al. based on the enzyme overexpressed in *Shewanella*. Its activity with GalN-6-P was approximately 27-fold higher than with GlcN-6-P (glucosamine-6-phosphate) (9,48 μmol/mg x min to 0,35 μmol/mg x min respectively). Their comparative genomic analysis also suggested that *agaS* is a universal catabolic gene in *aga* operon for all *Proteobacteria*. Unfortunately, at the moment of writing there is not much known about *in vitro* enzymatic activities of AgaS derived from other bacteria. Considering AgaS high conservatism, its high specificity towards galactosamine derivatives compared to other amino sugars, and the lack of *in vitro* research of this enzyme, it would be fair to state that, at the moment, enzymatic activity of AgaS isolated from *Shewanella* is representative of AgaS isomerases as a whole. In addition to that *ΔagaS* knockout mutants derived from *Shewanella* completely lost their ability to use GlcNAc as a sole carbon and nitrogen source in minimal medium. This deletion also potentially prevented the transcription of the following *agaY* gene. Therefore Leyn et al. confirmed AgaS to be essential in GlcNac/GalN catabolic pathway (Leyn et al. 2012), possibly inhibiting bacterial virulence.

6. AgaA – deacetylating and merging the processes

AgaA performs the role of GalNAc-6-P deacetylase in this catabolic pathway. This enzyme functions as a catalyst for a bridge reaction of transforming N-acetyl-D-galactosamine-6-phosphate to D-galactosamine-6-phosphate. *aga*A gene is usually closely clustered with other *aga* genes. During their study Leyn et al. found two variants of AgaA enzyme, AgaA^{II} being distinct, and characteristic only for *Shewanella* spp. A number 11 other species of *Proteobacteria* in which *aga*A gene has been found, had been functionally and genetically identical to that of *E. coli*. In *Shewanella* spp. AgaA^{II} shared 50% of similarity to NagA, responsible for N-acetyl-D-glucosamine-6-phosphate (GlcNAc-6-P) deacetylation (Leyn et al. 2012).

Both variants of AgaA, as well as NagA belong to COG1820 protein family of amidohydrolases. Additional phylogenetic analysis by Leyn et al. has confirmed that AgaA^{II} is a close paralog of NagA, which suggested that its development might have been a result of recent gene duplication event. Interestingly, in six *Proteobacteria* species analyzed by them, GalNAc-6-P deacetylases of both types were missing, suggesting that these organisms had the ability to utilize GalN, but not GalNAc (Leyn et al. 2012).

Thus, the presence or absence of *aga*A gene is crucial for the bacteria to utilize GalNAc. Leyn et al. as well as Zhang et al. have suggested that the presence of *aga*A, thus would also determine PTS transporters specificity (*aga*A lacking bacteria would have GalN-specific PTS), as was the case for *Shewanella* spp. as well as some other species (Leyn et al. 2012). As the result of their comparative genomics study Zhang et al. discovered one PTS previously described as GalN-specific in *Lactobacillaceae*. However, its co-occurrence and co-localization of its genes with *aga*A in most of the studied organisms was contrary to the previous suggestion (Zhang et al. 2015).

Additionally, Leyn et al. discovered the expression of *aga*A^{II} as well as *agaK* and *agaS* in *Shewanella* grown on GalNAc to be elevated over 100-fold compared to cells growing on GlcNAc (Leyn et al. 2012).

Enzymatic activity of AgaA^{II} was also evaluated in the aforementioned study. The enzyme exhibited significantly higher deacetylase activity with GalNAc-6-P than with GlcNAc-6-P (7,98 μmol/mg x min to 0,76 μmol/mg x min respectively) (Leyn et al. 2012).

Further, *E. coli* K-12 strain had a large deletion of several *aga* genes as well as truncation of *aga*A, resulting in GalNAc-negative phenotype (Leyn et al. 2012).

7. AgaZ – the weak link

AgaZ, which functions as Tag-6-P kinase, was present in most genomes studied by Leyn et al. in their comparative genomics study. Contrary to that, Zhang et al. discovered this enzyme and a gene encoding it throughout *Lactobacillaceae*. No paralogs of this enzyme were also found in genomes lacking them. However, it was suggested that this function is compensated by other enzymes, PfkA (phosphofructokinase I) and LacC, both having tagatose-6-phosphate activity (Leyn et al. 2012; Zhang et al. 2015). Brinkkötter et al. even suggesting that AgaZ functions as a non-catalytic subunit of AgaY [9,8].

Despite that, Leyn et al. observed in their study, that patterns of distribution of *agaZ* and *agaY* genes

are different. This observation, as well as an upregulation of *agaZ* in GalNAc environment (see AgaR) (Ray and Larson 2004), however, does not support the previous hypothesis, and suggests that AgaZ has its essential role in GalNAc catabolism independent of AgaY, which has not been discovered. The Catabolic activity of AgaZ has not yet been studied and remains unknown.

8. AgaY – final stitches

AgaY is a tagatose-1,6-diphosphate aldolase, a final enzyme in GalNAc catabolic pathway, catalyzing Tag-1,6-PP breakdown to glycerone phosphate and D-glyceraldehyde-3-phosphate. It was found by Leyn et al. and Zhang et al. to be present in almost every member *Enterobacterales* and *Vibrionales* orders but missing in *Shewanella* and several other *Proteobacteria* (Leyn et al. 2012; Zhang et al. 2015). However, in *Shewanella* its activity is compensated by non-committed enzymes such as class II fructose-bisphosphate aldolase (Fba), present in all its species. Interestingly Fba in *Shewanella* spp. is more similar to AgaY than Fba in *E. coli* (50% to 35% sequence similarity) (Leyn et al. 2012). Moreover, *Haemophilus parasuis*, pathogenic bacteria which cause Glasser disease in pigs, had a unique variant of GalNAc utilization genes including: a different *agaY*^{II} type, encoding an alternative version of tagatose-1,6-biphosphate aldolase belonging to LacC family (Rosey et al. 1991), as well as a unique type V PTS (Leyn et al. 2012).

It is worth mentioning that during their study in *Shewanella* spp. Leyn et al. discovered that the deletion of *agaS* gene prevented *agaY* from expression, thus resulting in a mutant's inability to use GalNAc. Complementation of this effect via restoring both *agaS* and *agaY* resulted in restoration of this ability, which was not the case when only one of them was complemented (Leyn et al. 2012).

9. PTS systems (AgaBCD/VWEF) – not only the transport

agaBCD and *agaVWEF* from the AgaR regulon encode two types of phosphotransferase (PTS) systems. Complex enzyme systems widely used by bacteria for the detection, transport and phosphorylation of various sugar substrates, including monosaccharides, disaccharides, amino sugars, polyols, and other sugar derivatives (Sørensen and Hove-Jensen 1996; Leyn et al. 2012). PTS systems catalyze the uptake of carbohydrates as well as their conversion to their respective phosphoesters during transport. The source of energy for these systems is phosphoenolpyruvate or PEP. The PTS systems have two general components: Histidine Phosphocarrier protein (HPr) with enzyme I and membrane bound sugar specific permeases (Enzymes II). Each enzyme II consists of one or two hydrophobic and two hydrophilic domains. They can exist as distinct proteins, as well as a single multidomain protein. Catalysis of the uptake of sugars and their conversion to phosphoestre is performed by four successive phosphoryl transfers. Initial phosphorylation of enzyme I using PEP as a substrate, transfer of the phosphoryl group from enzyme I to HPr, the self-phosphoryl transfer in HPr catalyzed by enzyme II, after which the phosphoryl group is transferred to histidine or cysteine residues of enzyme II. The sugar is transported through the membrane-bound enzyme II and undergoes phosphorylation (Meadow et al. 1990; Hassan et al. 2014).

PTS^{IIGalNAc} (*agaBCD*) being GalN-specific and GalNAc-specific PTS^{IIGalN} (*agaVWEF*) (Reizer et al. 1996; Ray and Larson 2004; Leyn et al. 2012). Both systems belong to the mannose-sorbose family (Leyn et al. 2012). PTS^{II} contains three domains (IIB, IIC and IID) encoded by the three *aga* genes, which are fused to one peptide. *aga*F however, encodes IIA PTS protein, that functions in the transport of both amino sugars (Ray and Larson 2004). During their study Leyn et al. discovered four types of PTS in *Proteobacteria* and distinguished two separate clades of PTS based on the components that are encoded by gene loci containing the adjacent *aga*A deacetylase gene. This may hint that these two clades (labeled PTS^{I} and PTS^{III}) are GalNAc-specific, whereas PTS components not encoded in close vicinity to *aga*A (PTS^{II} and PTS^{IV}), are GalN-specific (Leyn et al. 2012). Zhang et al. discovered a separate uncharacterized PTS^{V} type system which is found in some *Proteobacteria* (Zhang et al., 2015).

Amino sugar specific PTS play a crucial role of defining bacterial ability of their utilizations. Bacteria without them, for example, *E. coli* K-12 strain which lacks GalNAc-specific PTS, cannot use this sugar as a sole carbon, nor nitrogen source like other *Proteobacteria* (Brinkkötter et al. 2000; Mukherjee et al. 2008; Leyn et al. 2012). Some bacteria, however, without this specific PTS systems, for example *Shewanella*, developed unique sets of GalNAc- and GalN-specific permeases (AgaP) and kinases (AgaK) (Leyn et al. 2012).

10. AgaP and AgaK – new additions to the family?

AgaP is an amino sugar transporter protein, which is commonly accompanied by other amino sugar related transporters belonging to Ton-B-dependent receptors. *Shewanella spp.*, as well as *Burkholderia*

cenocepacia and *Caulobacter spp*., which lacks typical GalNAc-specific transporters utilizes sugar uptake through the outer membrane using a TonB-dependent Omp^{Aga} transporter. Transport through the inner membrane is done by AgaP permease. The subsequent phosphorylation of amino sugar is performed by AgaK kinase (Leyn et al. 2012).

AgaP was found to be a close paralog (50% similarity) to NagP, a GlcNAc permease belonging to GGP sugar transporter family (Rodionov et al. 2010; Leyn et al. 2012). AgaK was found by Leyn et al. to be most similar to the *Shewanella spp.* Glk^{II} glucokinase belonging to ROK family (35% similarity) (Rodionov et al. 2010; Leyn et al. 2012).

The *aga* cluster genes in *Burkholderia cenocepacia* and *Caulobacter* encoded different $AgaP^{II}$ and $AgaK^{II}$ variants (belonging to EamA and BcrAD_BadFG families respectively). With the absence of GalNAc-6-P deacetylase, and GalNAc-, GalN-specific PTS in their genomes, it was suggested that these enzymes are involved in GalN uptake and phosphorylation (Leyn et al. 2012).

This hints at the importance of GalNAc utilization for bacteria, which develop alternative routes of its uptake if the classic *aga*-controlled pathway is unavailable.

11. Aga operon as potential target for therapy

Another important aspect of amino sugar metabolism is its practical use by pathogenic bacteria that can be turned to our advantage. As was mentioned earlier, both amino sugars are directly or indirectly connected to bacterial virulence factors, examples being: LPS, murein and protein glycosylation and in some cases molecular mimicry. At the moment, however, there is not much knowledge available, with the exception of a few case and genomic studies, about how GalNAc and GalN utilization influences bacterial virulence.

The example of a case study is one conducted by Mukherjee et al. on several *E. coli* O157:H7 strain isolates derived from 2006 spinach outbreak compared to other *E. coli* strain notably K-12. They discovered that most of the spinach-associated O157:H7 isolates presented similar to K-12 strain $GalNAc^-$ phenotype unlike most of the other available O157:H7 representatives. It was suggested that this change was a result of a neutral mutation in *agaF* gene responsible for encoding Enzyme II of GalNAc-specific PTS. The aforementioned mutation has had a tendency to spread among *E. coli* O157:H7 as an adaptation useful in colonizing new niches (Mukherjee et al. 2008). This hints at the fact that one of the reasons *E. coli* K-12, (in case of the other strains, a widely spread multi-niche pathogen) is unable to secure a successful infection to be a result of its inability to use amino sugars, represented by GalNAc, as a carbon and nitrogen source highly present in animal and human organisms. Knowing this may help us understand strain-specific pathogenicity and perhaps even combat it.

Confirming the previous statement, Zhang et al. in their study conducted several Proteobacterial invasion assays based on animal models, and discovered that in case of *Streptococcus suis*, mostly a cattle-based pathogen with a limited ability to transfer and cause infection in humans, GalNAc/GalN utilization regulated by AgaR2 played a crucial role in its virulence (Zhang et al., 2015).

In the same study authors went as far as to suggest that both amino sugars took part in bacterial recognition and crosstalk, thus through maintenance and regulation of their utilization pathway allowing for successful infections (Zhang et al. 2015).

On the other hand, for *Proteus mirabilis*, a mostly human pathogen responsible for Urinary Tract Infections (mostly catheter-associated urinary tract infections, CAUTIs), but sometimes appearing in veterinary conditions, usage of amino sugars via *aga* genes also seemed to be important. As a result of a genome-wide transposon mutagenesis of Armbruster et al. AgaR regulator was listed as one of the fitness factors for human kidney colonization in this uropathogenic bacteria (Armbruster et al. 2017). However, the exact role and extent of GalN/GalNAc metabolisms influence on uropathogenicity is yet to be investigated.

How will this knowledge allow us to specifically combat pathogen strains which base their pathogenicity on the ability to use amino sugars? This would be possible through targeting key pathway elements in their metabolism, or bacterial constructs using such sugars similarly to as discovered by Gill et al. during their study. According to them GalNAc-linked post-translational modifications were used by pathogenic bacteria to modify host GTPases, through toxins, with α-GalNAc-O-Tyr in order to promote their virulence (Gill et al. 2011). Based on this information Behren et al. prepared vaccine constructs as well as polyclonal antibodies specific to this modification. They suggested that GalNAc protein residues might become a vital target in specific glycoproteomal detection, as well as vaccines and therapy in the future (Behren et al. 2023).

12. Conclusion

Not much is still known about amino sugars, represented by N-acetyl-d-galactosamine and D-galactosamine, role in bacterial pathogenicity. Recent studies, however, show that human and animal pathogens,

able to successfully secure infections, developed highly sophisticated and specific processes of their utilization, especially in comparison to non-pathogens who often lack the ability (Shen et al. 2025). Amino sugars utilization often appears alongside other, already known to be important, virulence factors is sure to be directly or indirectly involved in developing bacterial pathogenicity. Hence, it is not to our surprise that there are more and more real possibilities for these amino sugars, their utilization pathways as well as the *aga* operon responsible for it, to be used as targets in bacterial classification, pathogen detection and possibly even therapy start to appear. To achieve this, however, more studies regarding GalNAc/GalN utilization pathway in bacteria should be conducted in the future.

ORCID
Leon Petruńko https://orcid.org/0009-0004-8346-5440
Dawid Gmiter https://orcid.org/0000-0001-8663-129X

Conflict of interest

The authors do not report any financial or personal connections with other persons or organizations, which might negatively affect the contents of this publication and/or claim authorship rights to this publication.

Acknowledgements

This work was supported by the Polish National Science Centre (Grant No. 2019/33/N/NZ6/02406–D.G.

References

1. Abu-Qarn M., Eichler J., Sharon N.: Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. *Curr Opin Struct Biol*. **18**(5), 544-50 (2008) DOI: 10.1016/j.sbi.2008.06.010
2. Afzal M., Shafeeq S., Ahmed H., Kuipers O.P.: N-acetylgalatosamine-mediated regulation of the aga operon by AgaR in Streptococcus pneumoniae. *Front Cell Infect Microbiol*. **6**(SEP), 101 (2016) DOI:10.3389/fcimb.2016.00101
3. Armbruster C.E., Forsyth-DeOrnellas V., Johnson A.O., Smith S.N., Zhao L., Wu W., Mobley H.L.T.: Genome-wide transposon mutagenesis of Proteus mirabilis: Essential genes, fitness factors for catheter-associated urinary tract infection, and the impact of polymicrobial infection on fitness requirements. *PLoS Pathog*. **13**(6), e1006434 (2017) DOI:10.1371/journal.ppat.1006434
4. Bateman A.: The SIS domain: A phosphosugar-binding domain. *Trends Biochem Sci*. **24**(3), 94-5 (1999) DOI:10.1016/S0968-0004(99)01357-2
5. Behren S., Schorlemer M., Schmidt G., Aktories K., Westerlind U.: Antibodies Directed Against GalNAc- and GlcNAc-O-Tyrosine Posttranslational Modifications – a New Tool for Glycoproteomic Detection. *Chemistry - A European Journal*. **29**(29), e202300392 (2023) DOI:10.1002/chem.202300392
6. Bernatchez S., Szymanski C.M., Ishiyama N., Li J., Jarrell H.C., Lau P.C., Berghuis A.M., Young N.M., Wakarchuk W.W.: A single bifunctional UDP-GlcNAc/Glc 4-epimerase supports the synthesis of three cell surface glycoconjugates in Campylobacter jejuni. *Journal of Biological Chemistry*. **280**(6), 4792-802 (2005) DOI:10.1074/jbc.M407767200
7. Brinkkötter A., Klöß H., Alpert C.A., Lengeler J.W.: Pathways for the utilization of N-acetyl-galactosamine and galactosamine in Escherichia coli. *Mol Microbiol*. **37**(1), 125-35 (2000) DOI:10.1046/j.1365-2958.2000.01969.x
8. Carraway K.L., Hull S.R.: Cell surface mucin-type glycoproteins and mucin-like domains. *Glycobiology* **1**(2), 131-8 (1991) DOI:10.1093/glycob/1.2.131
9. Elgrably-Weiss M., Schlosser-Silverman E., Rosenshine I., Altuvia S.: DeoT, a DeoR-type transcriptional regulator of multiple target genes. *FEMS Microbiol Lett*. **254**(1), 141-8 (2006) DOI: 10.1111/j.1574-6968.2005.00020.x
10. Freymond P.P., Lazarevic V., Soldo B., Karamata D.: Poly(glucosyl-N-acetylgalactosamine 1-phosphate), a wall teichoic acid of Bacillus subtilis 168: Its biosynthetic pathway and mode of attachment to peptidoglycan. *Microbiology* (NY) **152**(6), 1709-1718 (2006) DOI:10.1099/mic.0.28814-0
11. Gill D.J., Clausen H., Bard F.: Location, location, location: New insights into O-GalNAc protein glycosylation. *Trends Cell Biol*. **21**(3) (2011) doi:10.1016/j.tcb.2010.11.004
12. Gourlay L.J., Sommaruga S., Nardini M., Sperandeo P., Dehò G., Polissi A., Bolognesi M.: Probing the active site of the sugar isomerase domain from E. coli arabinose-5-phosphate isomerase via X-ray crystallography. *Protein Science* **19**(12), 2430-2439 (2010) DOI: 10.1002/pro.525
13. Hassan K.A., Paulsen I.T., Elbourne L.D.H., Ren Q., Cameron A.D., Henderson P.J.F. Microbial Solute Transporters . In: Reference Module in Biomedical Sciences. (2014)
14. De Jong H., Wösten M.M.S.M., Wennekes T.: Sweet impersonators: Molecular mimicry of host glycans by bacteria. *Glycobiology* **32**(1), 11-22 (2022) DOI: 10.1093/glycob/cwab104
15. Leyn S.A., Gao F., Yang C., Rodionov D.A.: N-acetylgalactosamine utilization pathway and regulon in proteobacteria: Genomic reconstruction and experimental characterization in Shewanella. *Journal of Biological Chemistry* **287**(33), 28047-56 (2012) DOI: 10.1074/jbc.M112.382333
16. Masoud H., Sadovskaya I., De Kievit T., Altman E., Richards J.C., Lam J.S.: Structural elucidation of the lipopolysaccharide core region of the O- chain-deficient mutant strain A28 from Pseudomonas aeruginosa serotype 06 (International Antigenic Typing Scheme). *J Bacteriol*. **177**(23), 6718-26 (1995) DOI: 10.1128/jb.177.23.6718-6726.1995
17. Meadow N.D., Fox D.K., Roseman S.: The bacterial phosphoenol-pyruvate: Glycose phosphotransferase system. *Annu Rev Biochem*. **59**(1), 497-542 (1990) DOI: https://doi.org/10.1146/annurev.bi.59.070190.002433
18. Mukherjee A., Mammel M.K., LeClerc J.E., Cebula T.A.: Altered utilization of N-acetyl-D-galactosamine by Escherichia coli O157:H7 from the 2006 spinach outbreak. *J Bacteriol*. **190**(5), 1710-7 (2008) DOI: 10.1128/JB.01737-07
19. Pérez-Rueda E., Collado-Vides J.: The repertoire of DNA-binding transcriptional regulators in Escherichia coli K-12. *Nucleic Acids Res*. **28**(8), 1838–1847 (2000) DOI: 10.1093/nar/28.8.1838

20. Ray W.K., Larson T.J.: Application of AgaR repressor and dominant repressor variants for verification of a gene cluster involved in N-acetylgalactosamine metabolism in Escherichia coli K-12. *Mol Microbiol*. **51**(3), 813-26 (2004) DOI: 10.1046/j.1365-2958.2003.03868.x
21. Reizer J., Ramseier T.M., Reizer A., Charbit A., Saier M.H.: Novel phosphotransferase genes revealed by bacterial genome sequencing: A gene cluster encoding a putative N-acetylgalactosamine metabolic pathway in Escherichia coli. *Microbiology* (N Y) **142**(2), 231-250 (1996) DOI: 10.1099/13500872-142-2-231
22. Rodionov D.A., Yang C., Li X., Rodionova I.A., Wang Y., Obraztsova A.Y., Zagnitko O.P., Overbeek R., Romine M.F., Reed S., *et al.*: Genomic encyclopedia of sugar utilization pathways in the Shewanella genus. *BMC Genomics*. **11**(1), 494 (2010) DOI:10.1186/1471-2164-11-494
23. Rosey E.L., Oskouian B., Stewart G.C.: Lactose metabolism by Staphylococcus aureus: Characterization of lacABCD, the structural genes of the tagatose 6-phosphate pathway. *J Bacteriol*. **173**(19), 5992-8 (1991) DOI: 10.1128/jb.173.19.5992-5998.1991
24. Sadler J.E., Paulson J.C., Hill R.L.: The role of sialic acid in the expression of human MN blood group antigens. *Journal of Biological Chemistry* **254**(6), 2112-2119 (1979) DOI: 10.1016/s0021-9258(17)37773-6
25. Sadovskaya I., Brisson J.R., Lam J.S., Richards J.C., Altman E.: Structural elucidation of the lipopolysaccharide core regions of the wild-type strain PAO1 and O-chain-deficient mutant strains AK1401 and AK1012 from Pseudomonas aeruginosa serotype O5. *Eur J Biochem*. **255**(3), 673-84 (1998) DOI: 10.1046/j.1432-1327.1998.2550673.x
26. Shen W., Jan Freark de B., Folkert K., and Fu J.: New insights in amino sugar metabolism by the gut microbiome. *Gut Microbes*. **17**(1) 2510462 (2025) DOI:10.1080/19490976.2025.2510462
27. Škerlová J., Fábry M., Hubálek M., Otwinowski Z., Řezáčová P.: Structure of the effector-binding domain of deoxyribonucleoside regulator DeoR from Bacillus subtilis. *FEBS Journal*. **281**(18), 4280-92 (2014) DOI: 10.1111/febs.12856
28. Sommaruga S., Gioia De L., Tortora P., Polissi A.: Structure prediction and functional analysis of KdsD, an enzyme involved in lipopolysaccharide biosynthesis. *Biochem Biophys Res Commun*. **388**(2), 222-7 (2009) DOI: 10.1016/j.bbrc.2009.07.154
29. Sørensen K.I., Hove-Jensen B.: Ribose catabolism of Escherichia coli: Characterization of the rpiB gene encoding ribose phosphate isomerase B and of the rpiR gene, which is involved in regulation of rpiB expression. *J Bacteriol*. **178**(4), 1003-11 (1996) DOI: 10.1128/jb.178.4.1003-1011.1996
30. Zhang H., Ravcheev D.A., Hu D., Zhang F., Gong X., Hao L., Cao M., Rodionov D.A., Wang C., Feng Y.: Two novel regulators of N-acetyl-galactosamine utilization pathway and distinct roles in bacterial infections. *Microbiologyopen* **4**(6), 983-1000 (2015) DOI:10.1002/mbo3.307
31. Zhang X., Ebright R.H.: Substitution of 2 base pairs (1 Base pair per DNA half-site) within the Escherichia coli lac promoter DNA site for catabolite gene activator protein places the lac promoter in the FNR regulon. *Journal of Biological Chemistry* **265**(21), 12400-12403 (1990) DOI: 10.1016/s0021-9258(19)38360-7