communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-025-08987-9

Identification of the D-glucuronyl C5-epimerase that introduces iduronic acid into N-linked glycans decorating archaeal glycoproteins

Zlata Vershinin[1], Marianna Zaretsky[1], Anna Notaro[2], Dandan Yu[2], Antonio Molinaro[2], Shahar Sofer[1], Iris Grossman-Haham[1,3], Cristina De Castro[2] & Jerry Eichler[1]

In *Halobacterium salinarum*, halophilic archaea that grow in saturated salt solutions, glycoproteins are modified by an N-linked tetrasaccharide that includes iduronic acid (IdoA). This represents the only known example of IdoA, a sugar better known as a component of glycosaminoglycans found in various eukaryal tissues, in N-glycosylation. Although much of the pathway used to assemble this tetrasaccharide is defined, the epimerase that presumably converts glucuronic acid (GlcA) into IdoA had yet to be described. In silico predictions assign this role to VNG1058H. In *Hbt. salinarum* deleted of *VNG1058H*, an N-linked tetrasaccharide appeared in which IdoA is replaced by GlcA and greatly reduced sulfation of the last two tetrasaccharide sugars is noted. The absence of VNG1058H, moreover, affects cell physiology. Furthermore, *Hbt. salinarum* differentially transcribe *VNG1058H* as a function of growth temperature, suggestive of a response to environmental change. VNG1058H is the first D-glucuronyl C5-epimerase to be identified in Archaea.

N-glycosylation, the covalent linkage of sugar assemblies, or glycans, to selected asparagine residues of target proteins, is a post-translational modification performed across all three domains of life, namely, Eukarya, Bacteria, and Archaea. Despite having been studied in only a limited number of species, the archaeal version of N-glycosylation has nonetheless been shown to involve aspects of the process either unique to this domain of life or which have yet to be observed elsewhere, such as the lipid carrier upon which the N-linked glycan is assembled[1,2] and the stereochemistry of glycan linkage to these lipid carriers[3,4]. Archaeal N-linked glycans, moreover, present unparalleled diversity in terms of both the variety of sugars incorporated, as well as the architectures of glycans assembled from these sugars[5–8]. The unique nature of archaeal N-glycosylation is exemplified by *Halobacterium salinarum*, halophilic archaea best known as the source of the light-driven proton pump bacteriorhodopsin[9] and which provided the first example of N-glycosylation outside the Eukarya[10]. In *Hbt. salinarum*, glycoproteins have been reported as being modified by an N-linked glycan comprising ten to fifteen repeats of a sulfated pentasaccharide assembled on a dolichol pyrophosphate carrier, with N-acetylgalactosamine serving as the lipid-linking sugar[11,12] and/or by the N-linked tetrasaccharide β-D-GlcA(2S)-(1 → 4)-α-L-IdoA(3S)-(1 → 4)-β-D-GlcA-(1 → 4)-β-D-Glc, where GlcA is glucuronic acid, IdoA is iduronic acid, Glc is glucose and S represents sulfation[13], assembled on a dolichol phosphate (DolP) carrier[14,15]. Whereas nothing is known about the assembly of the former glycan, much of the pathway used to assemble and attach the latter has been delineated. The glycosyltransferases Agl28, Agl25, Agl26 and Agl27 sequentially add four sugars to a common DolP carrier[16,17], with the sulfotransferases Agl30 and Agl31 then respectively modifying the third and fourth tetrasaccharide sugars[18]. The flippase (or flippase-related protein) Agl29 delivers the lipid-linked sulfated tetrasaccharide across the membrane[17], at which point the archaeal oligosaccharyltransferase AglB transfers the lipid-bound glycan to selected asparagine residues in target proteins[19]. Despite this progress, additional components of this N-glycosylation pathway remain to be defined, such as the enzyme(s) responsible for the appearance of IdoA at tetrasaccharide position three[13,20].

L-IdoA, the C5 epimer of D-GlcA, is best known as a component of the repeating disaccharides that comprise glycosaminoglycans (GAGs) such as heparin, heparan sulfate and dermatan sulfate, polysaccharides found in a variety of animal tissues[21,22]. L-IdoA is also a component of rare non-canonical polymers that contribute to the cell walls of lower eukaryotes[23–25], and has been detected in a limited number of bacteria, where it can serve as a

[1]Department of Life Sciences, Ben-Gurion University of the Negev, Beersheva, Israel. [2]Department of Chemical Sciences, University of Napoli Federico II, Naples, Italy. [3]The Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Beersheva, Israel. e-mail: jeichler@bgu.ac.il

component of a capsular polysaccharide[26], the O-antigen[27] or an extracellular polysaccharide[28]. Accordingly, D-glucuronyl C5-epimerases that catalyze the epimerization of D-GlcA to generate L-IdoA have been characterized in Eukarya and Bacteria[29–33]. To the best of our knowledge, the L-IdoA found in the N-linked tetrasaccharide decorating *Hbt. salinarum* glycoproteins represent the only reported use of this hexuronic acid in an N-linked glycan. Still, the putative *Hbt. salinarum* D-glucuronyl C5-epimerase that converts GlcA into IdoA has yet to be identified, assuming *Hbt. salinarum* indeed synthesizes IdoA as do Eukarya and Bacteria.

In the present study, computational, genetic, biochemical, physiological, and structural biology approaches were adopted to determine whether VNG1058H is responsible for the appearance of IdoA found at the third position of the tetrasaccharide N-linked to *Hbt. salinarum* glycoproteins, thus corresponding to the first D-glucuronyl C5-epimerase to be described in Archaea.

Results

The N-linked tetrasaccharide added to glycoproteins in Δ*VNG1058H* cells is far less sulfated than in the parent strain

In looking to identify currently undefined components of the *Hbt. salinarum* N-glycosylation pathway responsible for the assembly of an N-linked tetrasaccharide, genes of unknown function found adjacent to those previously shown to encode pathway components were considered. This gene cluster, spanning from *VNG1048G* (encoding AglM[34]) to *VNG1068G* (encoding AglB[19]), includes *VNG1058H*. As a first step in determining if and how *VNG1058H* contributes to *Hbt. salinarum* N-glycosylation, a strain deleted of this gene was created, with gene deletion being confirmed by quantitative PCR (qPCR). Reintroducing *VNG1058H* into the deletion strain was able to reverse the growth defect associated with *VNG1058H* deletion (see below and Supplementary Fig. 1A). Moreover, PCR amplification (Supplementary Fig. 1B) and sequencing confirmed that neighboring genes (i.e., *VNG1056C*, *VNG1057C* and *VNG1059C*) were unaffected by *VNG1058H* deletion. Subsequently, the impact of *VNG1058H* deletion on N-linked tetrasaccharide assembly was considered by performing liquid chromatography-electrospray ionization mass spectrometry (LC-ESI MS) on tryptic peptides derived from the surface (S)-layer glycoprotein. As previously reported[15,18], the mass spectrometry profile from the parent strain included a peak at *m/z* 1298.83, corresponding to the $[M + 3H]^{3+}$ ion of the ^{475}SDAVNSSGGVKDNIDTSDFNQGVSSTSSIR^{504} peptide in which Asn-479 is modified by a tetrasaccharide comprising a hexose, a hexuronic acid and two sulfated hexuronic acids (calculated *m/z* 1298.82) (Fig. 1, left column, top panel). The same profile also included the same peptide modified by the same tetrasaccharide containing only one sulfate group (peak at *m/z* 1272.18; calculated *m/z* 1272.15) or no sulfate groups (peak at *m/z* 1245.52; calculated *m/z* 1245.48) (Fig. 1, left column, middle and bottom panels, respectively). When the same S-layer glycoprotein-derived peptide obtained from cells deleted of *VNG1058H* was similarly addressed, the same peaks at *m/z* 1298.83, 1272.18 and 1245.52 were observed (Fig. 1, right column, top, middle and bottom panels, respectively). It thus appears that the absence of VNG1058H does not prevent assembly of the N-linked tetrasaccharide decorating *Hbt. salinarum* glycoproteins.

Although the N-linked tetrasaccharide decorating the S-layer glycoprotein in both parent and Δ*VNG1058H* strain cells seemingly bears two, one or no sulfate groups, the relative degree of sugar sulfation differed greatly between the two strains. To compare the extent of tetrasaccharide sulfation in each strain, the heights of the relevant peaks were normalized to the height of that peak corresponding to the non-sulfated tetrasaccharide-modified peptide in the same strain, considered as 1.0 (Fig. 2). In the case of the parent strain, some 1.6-fold more of the tetrasaccharide decorating the S-layer glycoprotein-derived Asn-479-containing peptide bore one sulfate, relative to the same peptide presenting the non-sulfated tetrasaccharide, while some ten-fold more peptide bore two sulfates, values in agreement with those previously reported[18]. When the same peptide from the Δ*VNG1058H* strain was similarly analyzed, it was noted that the amount of peptide-bound tetrasaccharide bearing either one or two sulfate groups was in either case less than 1% of the level of the non-sulfated tetrasaccharide-modified peptide. These results show that in the absence of VNG1058H, considerably less sulfation of the N-linked tetrasaccharide decorating the S-layer glycoprotein occurred.

To determine whether the observed impact of *VNG1058H* deletion on tetrasaccharide sulfation was unique to the N-linked glycan decorating the S-layer glycoprotein, N-glycosylation of *Hbt. salinarum* archaellins, namely, the building blocks of the archaeal swimming device, the archaellum[35], previously shown to be N-glycosylated by the same tetrasaccharide as the S-layer glycoprotein[15,36,37], was considered as above. Specifically, the levels of mono- and di-sulfated N-linked tetrasaccharide decorating the archaellin ArlB1-derived TASGTDTVDYA^{84}NLTVR peptide, the ArlA1/ArlA2/ArlB2-derived QAAGADNI$^{97/69/73}$NLSK peptide and the ArlB2-derived VVNYA^{60}NLTVR peptide (with the position of the modified Asn residue in each archaellin-derived peptide numbered according to the amino acid sequence of the intact protein(s)) were assessed, relative to the level of the same non-sulfated tetrasaccharide N-linked to each peptide. As with the S-layer glycoprotein, *VNG1058H* deletion did not prevent tetrasaccharide addition to archaellins, yet drastically decreased the extent of tetrasaccharide sulfation (Supplementary Fig. 2).

Protein structure modeling predicts VNG1058H to be a sugar epimerase

Previous efforts revealed sulfation of the third and last sugars of the tetrasaccharide N-linked to *Hbt. salinarum* glycoproteins, namely, IdoA and GlcA, respectively[13,18]. Towards understanding the connection between *VNG1058H* deletion and the loss of N-linked tetrasaccharide sulfation, in silico tools were consulted to predict the role of VNG1058H. At present, the NCBI listing for VNG1058H does not include a predicted function, while UniProt lists VNG1058H (UniProt entry Q9HQQ1) as being a member of a protein family defined by *Haloferax volcanii* AglQ, a predicted sugar epimerase thought to convert GlcA into galacturonic acid (GalA)[38]. To gain further insight into VNG1058H function, AlphaFold was consulted to predict a three-dimensional model of VNG1058H based on the 372 residue-long sequence of the protein. The model, visualized using UCSF Chimera (Fig. 3A), was compared with solved protein structures in the Protein Data Bank to identify the highest scoring homolog of the AlphaFold-generated model, using the DALI server[39]. This strategy determined that VNG1058H best resembles *Spirochaeta thermophila* cellobiose 2-epimerase (PDB 5ZIG) at the structural level, with a *Z*-score of 24.1. Cellobiose 2-epimerases (EC 5.1.3.11) catalyze the epimerization of the D-glucose residue at the reducing end of the β(1 → 4)-linked disaccharide cellobiose to D-mannose and have also been reported to perform other related reactions[40,41]. UCSF Chimera, used to visualize the quality of the structural alignment of VNG1058H and *S. thermophila* cellobiose 2-epimerase, reported an overall RMSD value of 2.3 Å over 259 residue pairs (Fig. 3B). These in silico findings encouraged additional efforts aimed at determining whether VNG1058H acts as a sugar epimerase.

IdoA is replaced by GlcA in the N-linked glycans of Δ*VNG1058H* cells

D-glucuronyl C5-epimerases mediate the conversion of D-GlcA into L-IdoA[30,42,43]. To determine whether VNG1058H is a D-glucuronyl C5-epimerase involved in the appearance of the IdoA found at the third position of the N-linked tetrasaccharide decorating *Hbt. salinarum* glycoproteins, previously determined to be β-D-GlcA(2S)-(1 → 4)-α-L-IdoA(3S)-(1 → 4)-β-D-GlcA-(1 → 4)-β-D-Glc-Asn[13], NMR spectroscopy was performed on a tetrasaccharide-containing glycopeptide isolated from archaellins obtained from cells deleted of *VNG1058H*.

The proton spectrum of the tetrasaccharide presented only few signals in the anomeric region (5.3 – 4.4 ppm) in no apparent stoichiometric ratio, while those in a lower field could not be correlated to any carbon, likely due to their low intensity (Supplementary Fig. 3). The proton spectrum, however, included several signals at high-field (3.1–0.7 ppm), likely related to the

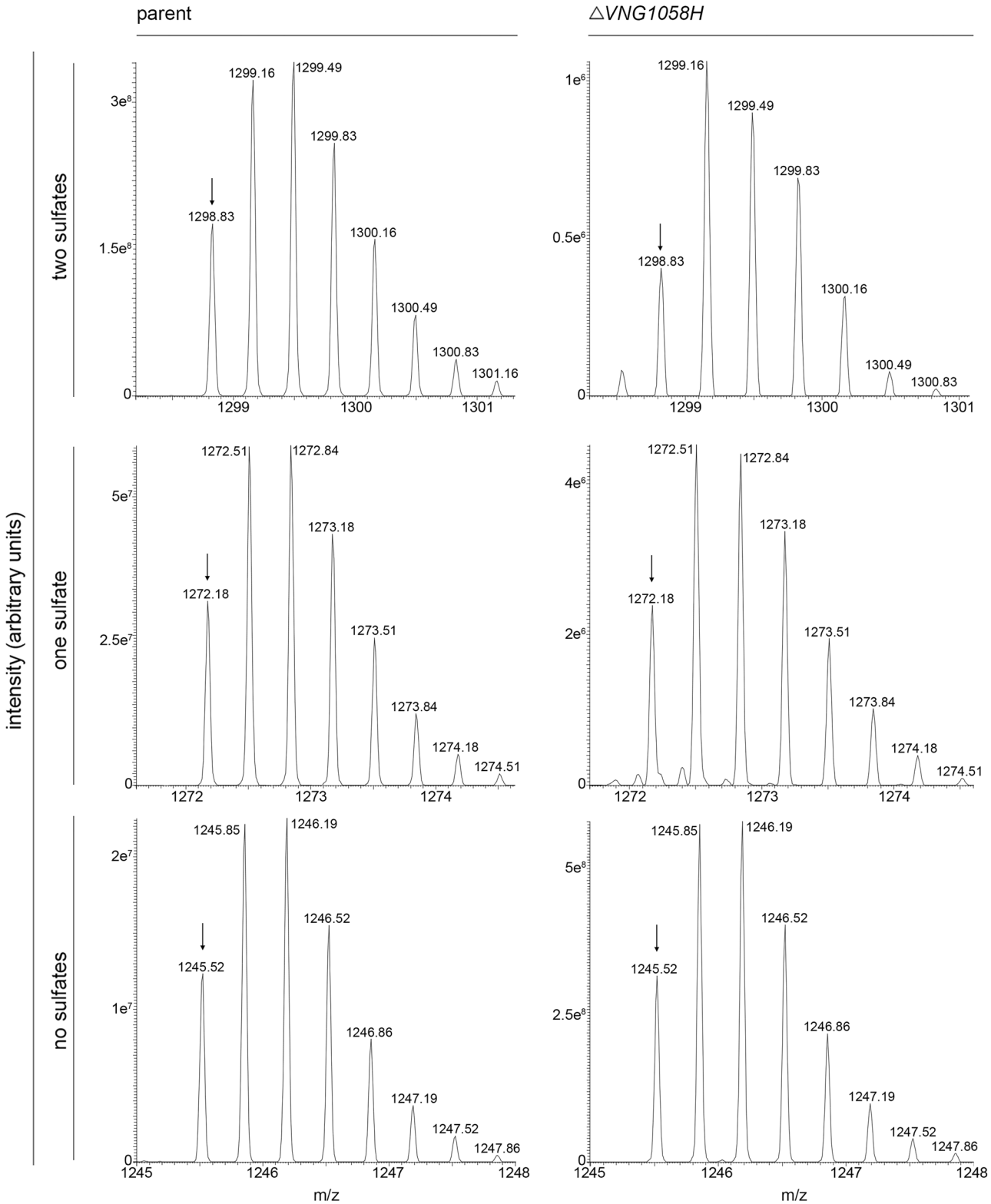


Fig. 1 | Deletion of *VNG1058H* does not prevent assembly of the N-linked tetrasaccharide decorating the *Hbt. salinarum* S-layer glycoprotein. LC-ESI MS profiles from the parent strain (left column) or the Δ*VNG1058H* strain (right column) showing peaks at *m/z* 1298.83, corresponding to the $[M + 3H]^{3+}$ ion of an S-layer glycoprotein-derived Asn-479-containing peptide modified by a tetrasaccharide comprising a hexose, a hexuronic acid and two sulfated hexuronic acids (two sulfates; top row), at *m/z* 1272.18, reflecting the same peptide modified by the same tetrasaccharide presenting only one sulfate (one sulfate; middle row), and at *m/z* 1245.52, reflecting the same peptide modified by the same tetrasaccharide lacking sulfates (no sulfates; bottom row), from parent strain cells. A representative of three biological repeats is shown. The arrows in each panel depict the peak of interest.

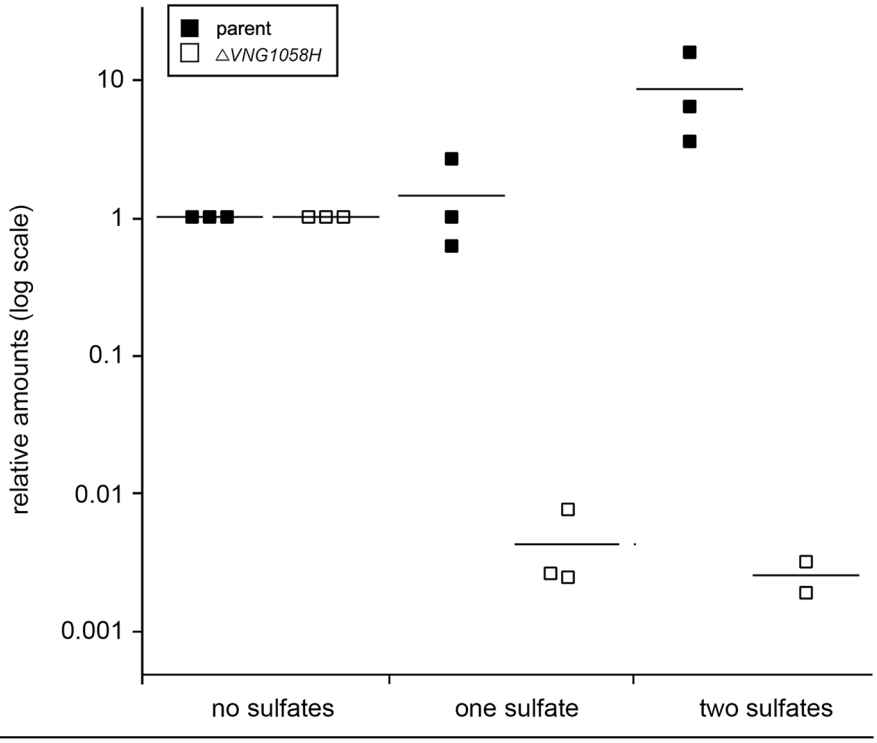


Fig. 2 | Comparison of the degrees of N-linked tetrasaccharide sulfation in parent and Δ*VNG1058H* strain cells. The relative amounts of mono-sulfated and di-sulfated tetrasaccharides N-linked to the peptide shown in Fig. 1 in each strain, relative to amount of the non-sulfated version of the same peptide in that strain, taken as 1.0 ($N = 3$), are shown. Parent strain: One sulfate, 1.58 ± 1.03 (standard deviation), $N = 3$; two sulfates, 10.25 ± 5.69, $N = 3$. Δ*VNG1058H* strain: One sulfate, 0.005 ± 0.003, $N = 3$; two sulfates, 0.002 ± 0.001, $N = 2$. Note that the *Y*-axis is in log scale.

presence of free peptides or of the peptide to which the N-linked glycan is linked. Since these signals were in a region distinct from that of the carbohydrate moiety of the glycopeptide, interpretation of the 2D NMR spectra was performed without further sample purification. Accordingly, the analysis considered the most intense anomeric protons (Fig. 4 and Table 1), and the labels used matched those previously used for the glycopeptide obtained from the parent strain[13].

The HSQC spectrum (Fig. 4B) presented a set of anomeric protons at ~5.0 ppm which correlated to a carbon at ~80 ppm, the value expected for the anomeric carbon atom engaged in a N-glycosidic linkage. This group of signals was collectively labeled **A**, and correlations of the most intense signal (H-1 at 5.02 ppm; Table 1) were analyzed in the 2D NMR spectra recorded. The COSY spectrum determined the positions of H-2 (labeled $\mathbf{A_2}$) and of H-3 ($\mathbf{A_3}$) at 3.43 and 3.69 ppm, respectively (Fig. 5A, C, E), while $\mathbf{A_5}$ was deduced by the NOE effect with $\mathbf{A_1}$ (Fig. 5E), characteristic of 1,3 di-*axial* protons. In the COSY spectrum, $\mathbf{A_5}$ showed weak correlation with a proton at 3.81 ppm, that in turn had an intense cross-peak with another at 3.91 ppm (Fig. 5C). The same correlations were confirmed in the TOCSY spectrum, where the intensities were much higher (Supplementary Fig. 4C). The protons at 3.81 and 3.91 ppm were associated with a hydroxymethyl carbon at 61.0 ppm and were, therefore, labeled as $\mathbf{A_{6'}}$ and $\mathbf{A_6}$, respectively. In correspondence to $\mathbf{A_5}$, the HSQC spectrum presented three densities at 76.6, 77.6, and 79.5 ppm, with the last two assigned to C-5 and C-4 of **A**, based on the values reported for the parent strain sample[13]. This assignment was further supported by the similarities of the H-4 and H-5 values between products, also being coincident in the glycopeptide from the parent strain, where it occurred at a similar chemical shift, i.e., 3.65 ppm.

The HSQC spectrum presented two other main anomeric signals at 4.54 and 4.49 ppm, with the intensity of that at 4.54 ppm being double the other, suggesting that it results from the overlap of two anomeric protons. Analysis of the H-2/H-1 correlation in the COSY spectrum confirmed this hypothesis (Fig. 5D). This correlation was composed of two cross-peaks (enclosed in the dotted boxes), with the left assigned to $\mathbf{D_2}$ (3.41 ppm) and the right to $\mathbf{B_2}$ (3.38 ppm). The corresponding H-3 chemical shifts were deduced from the COSY spectrum, while H-4 values were identified by the intra-residue NOE effects with the respective H-2 signals, since each H-3/H-4 correlation blurred into the diagonal due to the proximity of these protons. It was found that the $\mathbf{B_4}$ and $\mathbf{D_4}$ chemical shifts were slightly different and occurring at 3.68 and 3.69 ppm, respectively (Fig. 5C). It was not possible to distinguish the H-5 signals, with both at ~3.89 ppm. Analysis of the HSQC and HMBC spectra completed the chemical shifts assignment, identifying both **B** and **D** as GlcA units, β-configured at their anomeric centers ($^3J_{H1,H2} = 7.6$ Hz) and substituted at O-4 due to the low-field displacement of the corresponding carbons found at 82.1 and 82.2 ppm in **B** and **D**, respectively[44]. Finally, the HMBC spectrum had a cross peak between the H-5 signals and a carbon at 176.1 ppm (Supplementary Fig. 5), confirming the identity of each of these two units as GlcA.

With respect to the last anomeric signal, labeled **C** (4.49 ppm), the $^3J_{H1,H2}$ value (7.6 Hz) indicated the β configuration of the anomeric center, while the COSY spectrum enabled identification of $\mathbf{C_2}$ and $\mathbf{C_3}$ at 3.34 and 3.51 ppm, respectively. No other correlation was detected from H-3, although it was possible to identify $\mathbf{C_5}$ (3.74 ppm) in the NOESY spectrum due to the NOE effect with H-1 (Fig. 5D). There, $\mathbf{C_5}$ had one correlation in the COSY spectrum with a signal adjacent to $\mathbf{C_3}$, identified as $\mathbf{C_4}$. $\mathbf{C_5}$ had no other correlations, and for this reason, the unit was identified as GlcA, even though no correlation with a carboxylic function was found in the HMBC spectrum. The carbon chemical shift values (Table 1) denoted that **C** was not further substituted and was, therefore, deemed to be located at the non-reducing terminal of the glycan.

Regarding the sequence of the residues, the NOESY spectrum presented a correlation between H-4 of **B** and H-1 of **C** (label $\mathbf{B_4C_1}$; Fig. 5D) and a second at 3.69/4.54 ppm, tentatively assigned to $\mathbf{D_4B_1}$ (Fig. 5D). This assignment was further supported by analyzing the overlap of the HMBC and HSQC spectra. Here, it should be noted that the anomeric carbon chemical shifts of **B** and **D** were slightly different with that of **B** at lower field, giving a better alignment with the density at 3.69 ppm in the HMBC spectrum (Fig. 4F), thus supporting the linkage of **B** to O-4 of **D**. Lastly, **D** was considered linked to O-4 of **A** based on the correlation $\mathbf{D_1A_4}$ detected in the HMBC spectrum (Fig. 4C) that better aligned with the right part of the density. Based on these considerations, the structure of the glycan from the *VNG1058H* deletion strain was identified as β-GlcA-(1 → 4)-β-GlcA-(1 → 4)-β-GlcA-(1 → 4)-β-Glc-peptide, as represented in Fig. 4A.

The tetrasaccharide generated by the deletion strain showed three main differences from that assembled by the parent strain. Firstly, starting from the N-linked Glc (**A**), the third sugar (**B**) of the tetrasaccharide found in the

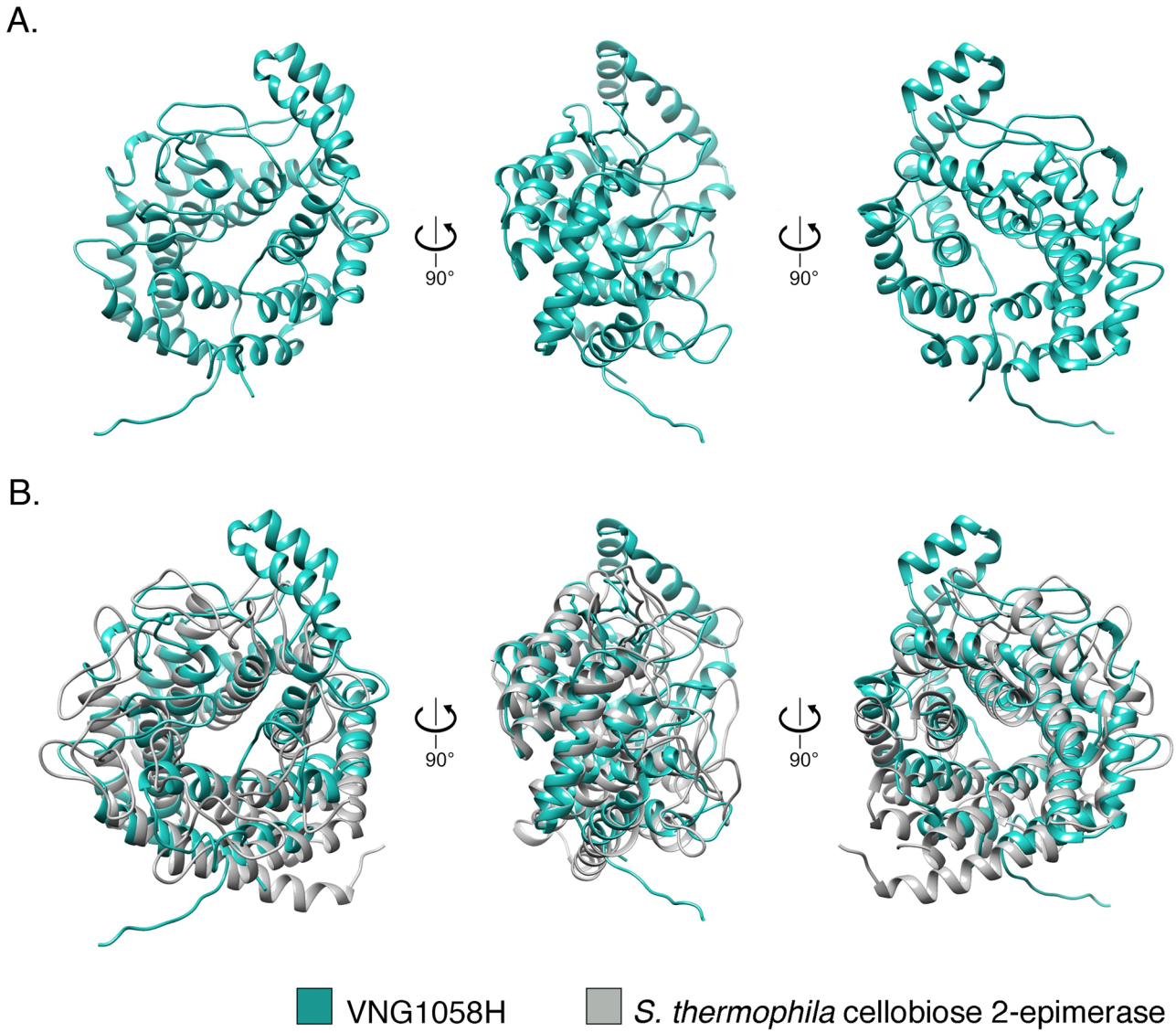


Fig. 3 | Predicted structure of VNG1058H. A The predicted structure of VNG1058H as generated using AlphaFold and visualized with UCSF Chimera. **B** Super-imposition of the AlphaFold-generated VNG1058H structure (sea green) with that of *S. thermophila* cellobiose 2-epimerase (PDB 5ZIG)(gray). In both **A** and **B**, structures at 0, 90 and 180° rotations around the vertical axis are shown.

deletion strain is GlcA, rather than IdoA found at this position in the tetrasaccharide assembled by the parent strain. Secondly, the third sugar of the tetrasaccharide assembled in the parent strain, i.e., IdoA, is found in the α configuation, whereas the GlcA at this position in the *VNG1058H* deletion strain presents the β configuration. Finally, this N-linked glycan generated by the deletion strain is not sulfated, since none of the proton and carbon chemical shift values (Table 1) were shifted at low-field, as occurs upon esterification with a sulfate group[44], such as in the case of the N-linked tetrasaccharide from the parent strain, which is sulfated at O-2 of **C** (the non-reducing GlcA) and at O-3 of the IdoA (equivalent position of the **B** unit in the tetrasaccharide from the deletion strain) (Fig. 4A)[13].

Deletion of *VNG1058H* carries functional implications

Towards assessing the functional impact of *VNG1058H* deletion, resulting in the replacement of IdoA by GlcA in the N-linked tetrasaccharide decorating *Hbt. salinarum* glycoproteins, structural modeling of the glycan was performed. Atomic Gasteiger charges were calculated for the N-linked tetrasaccharides decorating glycoproteins in the parent and *VNG1058H* deletion strains and mapped onto their van der Waals surfaces. The native tetrasaccharide appears bulkier than the tetrasaccharide present in the deletion strain due to the presence of a sulfate group on both the third and fourth sugars in the former (Fig. 6A). Moreover, although both tetrasaccharides present distinct negative charges at the COOH groups on the second, third, and fourth sugars (i.e., sugars **D, B** and **C**, respectively; Fig. 4A), the electrostatic surface of the native tetrasaccharide displays an overall pronounced negative charge over the entire non-reducing end. In the tetrasaccharide from the deletion strain, the negative charge is more localized, with some regions at the non-reducing end being neutral (Fig. 6A).

Given the demonstrated connection between *Hbt. salinarum* N-glycosylation and cell motility[16–19,37], the impact of *VNG1058H* deletion on cell motility was addressed. Cell motility was assessed in a plate assay in which a drop of cell culture is applied to the center of a Petri dish containing semi-solid growth medium, and the distance out to which the cells swim (i.e., the diameter of the halo that appears) after a given period is measured. When motility was assessed four days after plating, the diameter of the halo formed by cells of the parent strain was 5.77 cm on average (standard error of the mean (SEM) = 0.1; $N = 13$), whereas the diameter of the halo of cells of the mutant strain was 4.23 cm on average (SEM = 0.11; $N = 14$) (Fig. 6B). Although these differences are significant (Student's *t* test, two-tailed, unpaired; $p < 0.0001$), and reminiscent of what was seen in cells deleted of either Agl30

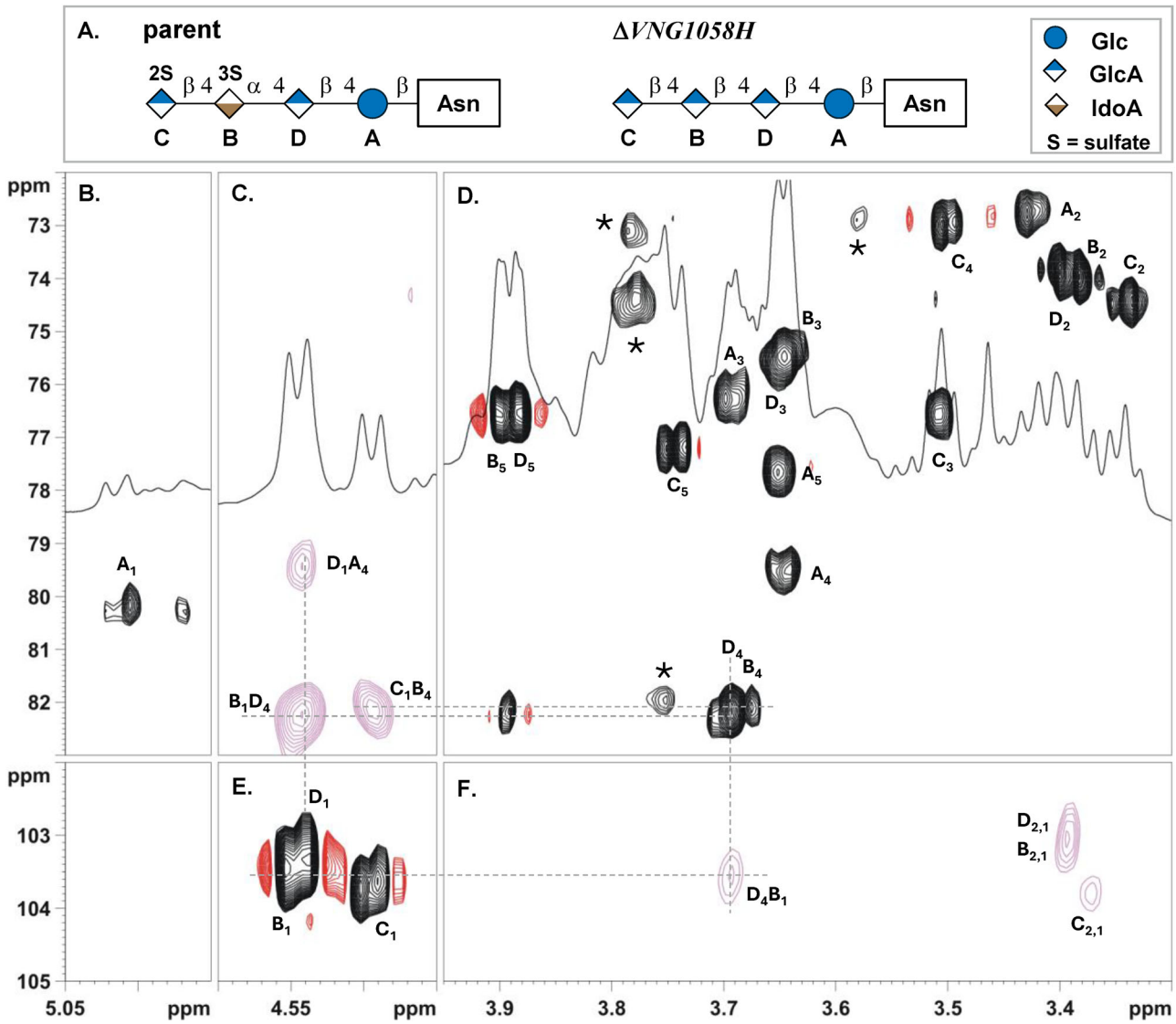


Fig. 4 | HSQC and HMBC spectra. HSQC (black and red) and HMBC (pink) spectra recorded for the glycopeptide isolated from the Δ*VNG1058H* strain. **A** Structures of the N-linked tetrasaccharide assembled by the parent (right) and Δ*VNG1058H* (left) strains. **B**, **E** HSQC expansion of the anomeric regions. **C** HMBC area detailing the long-range correlations of the anomeric protons; **D** HSQC expansion detailing the carbinolic region. **F** HMBC correlations between the carbinolic and the anomeric protons. The alignment between some key correlations is traced with a gray broken line. Letters used to mark densities follow the scheme used in Table 1 and are reported on the N-linked tetrasaccharide structure. "*" indicates a density not related to the glycan.

Table 1 | Proton and carbon (in italics) chemical shifts measured for the glycopeptide isolated from *Hbt. salinarum* Δ*VNG1058H* cells measured at 600 and 150 MHz, respectively, in D_2O at 298 K

Residue	1	2	3	4	5	6
A	5.02	3.43	3.69	3.68	3.68	3.91; 3.81
4-β-Glc-Asn	80.2	72.7	76.2	79.5	77.6	61.0
B	4.54	3.38	3.64	3.68	~3.89	--
4-β-GlcA	103.5	74.0	75.5	82.1	76.6	176.1
C	4.49	3.34	3.51	3.50	3.74	--
4-β-GlcA	103.7	74.4	76.5	73.0	77.2	*ND*
D	4.54	3.41	3.65	3.69	~3.89	--
4-β-GlcA	103.3	73.9	75.5	82.2	76.6	176.1

or Agl31, namely, the sulfotransferases that respectively sulfate the third and final sugars of the N-linked tetrasaccharide[18], the decrease in motility was far less striking than seen with *Hbt. salinarum* mutant strains that cannot perform N-glycosylation or that can only assemble truncated N-linked glycans[16,17,19]. To confirm that the loss of motility observed was not due to a loss of viability upon plating, cells were picked from both parent and mutant strain plates and transferred to liquid growth medium; normal growth of each strain was seen.

To further test whether the differences in glycan composition and conformation carried functional implications, the impact of *VNG1058H* deletion on the protease susceptibility of the S-layer glycoprotein that comprises the S-layer surrounding the *Hbt. salinarum* cell[10] was considered. For this, cells of the parent and deletion strains were challenged with proteinase K and the extent of S-layer glycoprotein proteolysis over time was determined by SDS-PAGE, given the ease of detecting the S-layer glycoprotein in such gels. Whereas over 70% of S-layer glycoprotein in parent strain cells remained intact after 2.5 h of proteolytic digestion (72.8 ± 4.7%

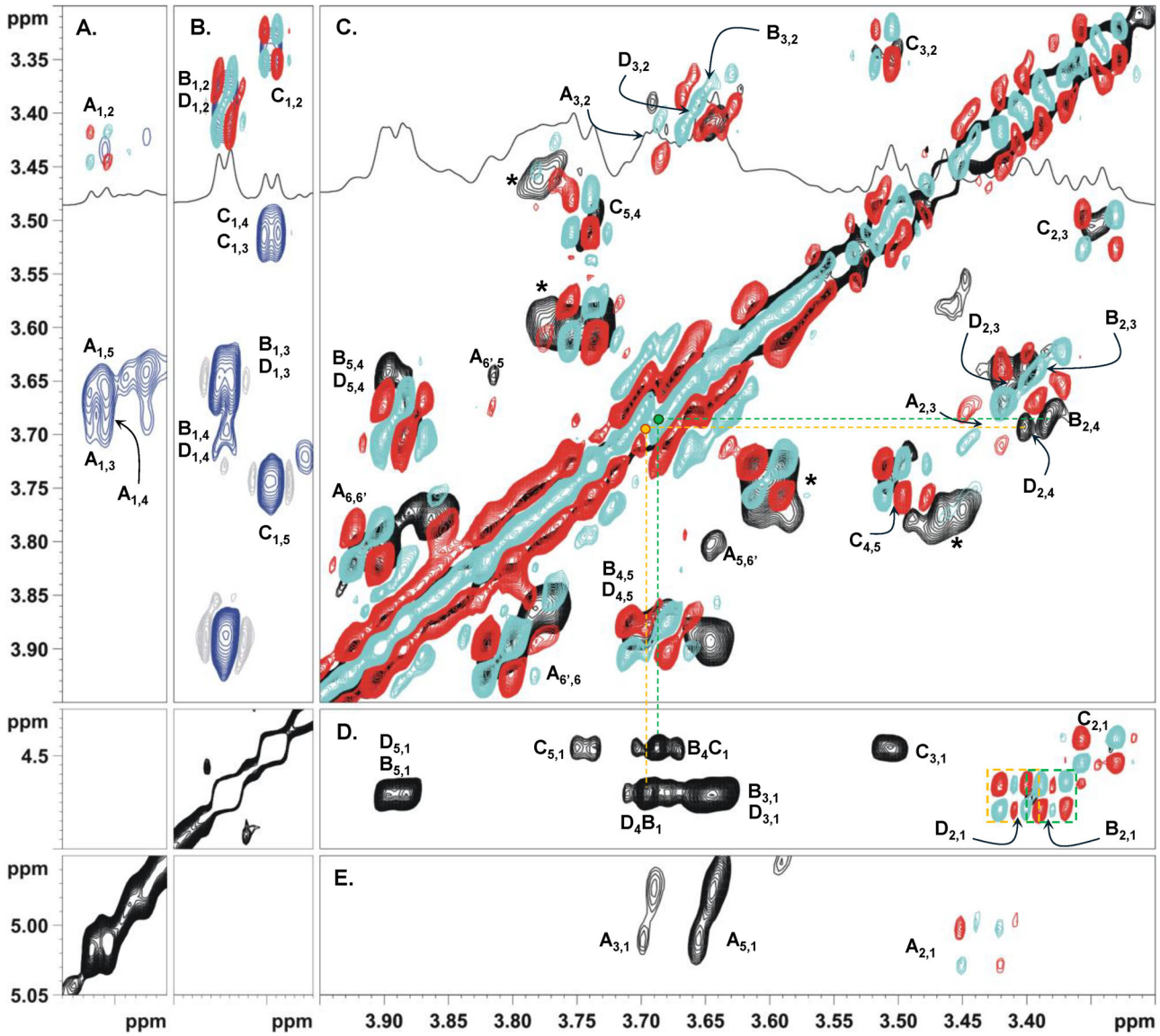
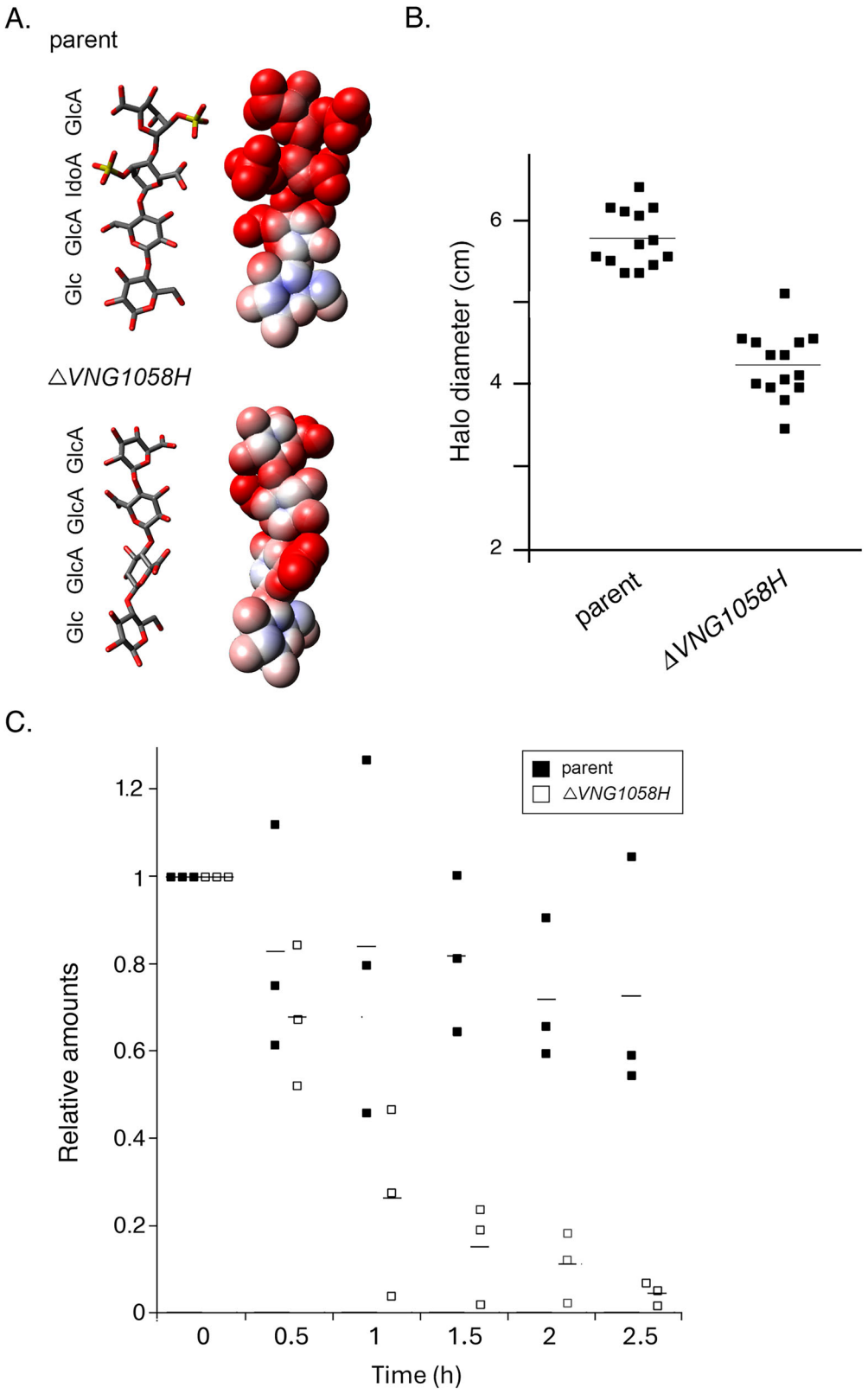


Fig. 5 | NOESY, TOCSY and COSY spectra. NOESY (black), TOCSY (blue) and COSY (cyan and red) spectra recorded from the Δ*VNG1058H* strain. **A**, **B** TOCSY expansion of the anomeric regions. **C** Overlap of the COSY and NOESY spectra detailing the correlation in the carbinolic region, with the broken colored lines traced to facilitate the identification of $\mathbf{B_4}$ and $\mathbf{D_4}$ from the NOEs with the corresponding H-2. **D** NOESY and COSY overlay detailing correlations between carbinolic and **B**, **C**, and **D** anomeric protons. Here, the correlations $\mathbf{B_{2,1}}$ and $\mathbf{D_{2,1}}$ are presented in a dotted colored box. **E** NOESY and COSY overlay detailing correlations between carbinolic and **A** anomeric protons. Letters used to mark densities follow the scheme used in Table 1 and are reported on the N-linked tetrasaccharide structure. "*" indicates a density not related to the glycan.

(standard deviation)), only some 25% of the intact S-layer glycoprotein was detected after 1 h of proteinase K treatment of the *ΔVNG1058H* strain cells (26.1 ± 21.2%). After 2.5 h of exposure of deletion strain cells to the protease, less than 5% of the S-layer glycoprotein remained intact (4.68 ± 3.5%; Fig. 6C). Thus, the replacement of IdoA with GlcA at the third position of the N-linked tetrasaccharide decorating the *Hbt. salinarum* S-layer glycoprotein led to increased susceptibility to added protease, much as seen upon truncation of the same glycan[16] or of that glycan decorating the *Hfx. volcanii* S-layer glycoprotein[45].

VNG1058H transcript levels are affected by changes in growth temperature but not medium salinity

Given how the absence or presence of VNG1058H has no effect on the modification of *Hbt. salinarum* glycoproteins by an N-linked tetrasaccharide, but rather determines whether the third sugar of this glycan will be GlcA or IdoA, in turn affecting the extent of glycan sulfation, it is conceivable that *Hbt. salinarum* selectively express VNG1058H, thereby modulating N-linked tetrasaccharide composition, possibly in response to specific growth conditions. To test this hypothesis, *VNG1058H* transcript levels as a function of growth conditions were determined by qPCR. Specifically, *VNG1058H* transcript levels were measured in cells grown under standard conditions (i.e., in medium containing 4.2 M NaCl, at 42 °C; see "Methods"), at different growth temperatures (i.e., 37 or 46 °C) or in medium containing reduced salt concentrations (i.e., 2.9 or 3.5 M NaCl). Relative to the normalized level of *VNG1058H* transcription measured in cells grown at 42 °C (1.0 ± 0.096 (SEM); $N = 8$), significantly lower levels of *VNG1058H* transcripts were measured in cells grown at 37 or 46 °C (0.44 ± 0.023, $N = 8$ and 0.49 ± 0.072, $N = 7$, respectively; both differences are significant, namely, $p < 0.0001$ and $p < 0.002$, respectively, according to Student's *t* test, two-tailed, unpaired; Supplementary Fig. 6A, C). In contrast, changing the level of medium salinity did not have any significant impact on *VNG1058H* transcript levels (4.2 M NaCl: 1.0 ± 0.096, $N = 8$; 3.5 M NaCl:



0.88 ± 0.105, $N = 8$; 2.9 M NaCl: 0.95 ± 0.288, $N = 8$; Supplementary Fig. 6B, D).

To assess the impact of *VNG1058H* deletion on growth at the different temperatures, growth rates of parent and *ΔVNG1058H* strain cells grown at 37, 42 or 46 °C were determined (Supplementary Fig. 7). While *VNG1058H* deletion clearly hindered growth at 42 °C, this effect was greatly diminished with cells grown at 37 °C (although differences at several time points remained statistically significant) and seemingly absent with cells grown at 46 °C. As such, it would seem that in cultures growing at non-optimal conditions, there is less need for VNG1058H-induced changes in N-linked tetrasaccharide composition, relative to when cells are grown at 42 °C.

Fig. 6 | *VNG1058H* deletion affects S-layer structure and cell motility. A Atomic models of the tetrasaccharide N-linked to *Hbt. salinarum* glycoproteins in the parent (top pair) and Δ*VNG1058H* (bottom pair) strains. The stick presentations (left in each pair) show carbon atoms in gray, oxygen atoms in red, and sulfur atoms in yellow. Hydrogen atoms were removed for simplicity. In the sphere presentations (right in each pair), spheres, representing the van der Waals radii of each atom, are colored according to atomic Gasteiger charge (red, negative; blue, positive). **B** Aliquots (10 µl) of liquid cultures of the parent and deletion strains grown to logarithmic phase were applied to the center of Petri dishes containing semi-solid medium. After four days at 42 °C, the diameter of the halos that appeared were measured. The halo diameters measured on 13-14 plates of each strain are presented, with the horizontal lines corresponding to the averages (parent strain: 5.77 ± 0.1 (SEM) cm, $N = 13$; Δ*VNG1058H* strain: 4.23 ± 0.11; $N = 14$). **C** Parent and Δ*VNG1058H* strains were exposed to proteinase K and the protein contents of aliquots removed at the indicated times were separated by 8% SDS-PAGE. Following Coomassie staining, the intensities of the S-layer glycoprotein were densitometrically determined and normalized to the intensity of the band immediately prior to protease addition (time = 0), taken as 1. The average intensities of the stained S-layer glycoprotein at each time point from three biological repeats are shown (parent strain: 0.5 h, 0.83 ± 0.26 (standard deviation); 1 h, 0.84 ± 0.41; 1.5 h, 0.82 ± 0.18; 2 h, 0.72 ± 0.16; 2.5 h, 0.73 ± 0.28. Δ*VNG1058H* strain: 0.5 h, 0.68 ± 0.16; 1 h, 0.26 ± 0.21; 1.5 h, 0.15 ± 0.11; 2 h, 0.11 ± 0.08; 2.5 h, 0.05 ± 0.04).

Phylogenetic analysis suggests that haloarchaea other than *Hbt. salinarum* can generate IdoA

Although N-glycosylation appears to be an almost universal trait of Archaea, based on the presence of *aglB*, encoding the archaeal oligosaccharyltransferase[46], essentially nothing is known of the composition of N-linked glycans or the biosynthetic pathways involved in their assembly outside a handful of species[2,7]. With this in mind, VNG1058H was used as a query sequence in a BLAST search (www.uniprot.org/blast; January 2024) designed to identify haloarchaeal homologs of this sugar epimerase, the existence of which could reflect the use of IdoA in other archaeal N-linked glycans.

The list of homologs generated included the only other sugar epimerase for which experimental evidence for function exists, namely, *Hfx. volcanii* AglQ[38,47]. When the list of homologs was assembled into an unrooted neighbor-joining tree, two distinct clades, one containing VNG1058H and the other containing AglQ, were apparent (Fig. 7). As such, the phylogenetic analysis not only distinguished between archaeal sugar epimerases (i.e., VNG1058H and AglQ) deemed as catalyzing the conversion of GlcA into distinct hexuronic acids (IdoA and GalA, respectively), but also revealed the presence of archaeal homologs of VNG1058H, suggesting that IdoA is generated by archaea other than *Hbt. salinarum*. However, of the different species predicted to contain VNG1058H- or AglQ-like proteins, N-linked glycans have only been characterized to date in *Hbt. salinarum* and *Hfx. volcanii*[13,47]. Earlier efforts predicted the presence of D-glucuronyl C5-epimerase in other archaea, however, nothing is known of N-glycosylation in these species[43]. As such, the presence of IdoA in N-linked glycans other than the tetrasaccharide assembled by *Hbt. salinarum* remains an untested prediction.

Discussion

Although many of the components of the pathway responsible for the assembly and attachment of an N-linked tetrasaccharide to *Hbt. salinarum* glycoproteins have been defined[16–19], additional pathway components remain to be described, including the enzyme responsible for the appearance of IdoA, the third sugar of the N-linked tetrasaccharide[13]. In the present study, evidence for VNG1058H being the D-glucuronyl C5-epimerase that catalyzes the conversion of GlcA into IdoA was presented. Specifically, NMR spectroscopy confirmed that in *Hbt. salinarum* cells deleted of *VNG1058H*, GlcA, rather than IdoA, were found as the third sugar of the N-linked tetrasaccharide decorating glycoproteins in this haloarchaeon. As such, VNG1058H is the first D-glucuronyl C5-epimerase to be reported in Archaea, joining the family of enzymes best known for their involvement in the biosynthesis of the GAGs heparin, heparan sulfate and dermatan sulfate in Eukarya[30,42].

Still, open questions regarding the activity of VNG1058H remain. In eukaryotes, D-glucuronyl C5-epimerases modify polymer-bound D-GlcA, rather than the soluble UDP-GlcA precursor[28,42]. The results of in vitro assays using a eukaryal substrate argue that the same holds true for a bacterial D-glucuronyl C5-epimerase[31]. Although it is tempting to assume that VNG1058H also acts on glycan-bound GlcA, rather than soluble UDP-GlcA, this option cannot yet be discounted. Indeed, the predicted structure of VNG1058H differs from the solved structure of the human glucuronyl C5 epimerase GlcE (PDB 6HZZ)[33], possibly reflecting a distinct mechanism of action. Moreover, *Hbt. salinarum* VNG1058H was not recognized in a BLAST search using the human or a bacterial glucuronyl C5 epimerase as query. On the other hand, despite acting on glycan-incorporated GlcA, the D-glucuronyl C5-epimerases involved in biosynthesis of the GAGs heparin, heparan sulfate and dermatan sulfate also do not share sequence homology or domain similarity[42,48]. If one assumes that VNG1058H acts on a glycan-bound GlcA target as does its eukaryal and bacterial counterparts[49], it remains to be determined whether IdoA is generated from GlcA that corresponds to the terminal sugar of the DolP-bound trisaccharide precursor of the tetrasaccharide or from GlcA at position three of the complete DolP-bound tetrasaccharide, although support for the latter scenario exists (see below). It is also not clear why the GlcA at position two of the tetrasaccharide is not epimerized to IdoA. Future in vitro assays using purified VNG1058H and appropriate substrates may help answer these and other mechanistic questions.

The results presented here further revealed that while VNG1058H is not needed for the assembly of a DolP-bound tetrasaccharide, which accordingly lacks IdoA, deletion of the encoding gene also resulted in a loss of glycan sulfation. Recent efforts reported that Agl30 and Agl31, respectively, sulfate the IdoA and GlcA found at positions three and four of the N-linked tetrasaccharide[18]. Although sulfation of IdoA at tetrasaccharide position three and of GlcA at tetrasaccharide position four occurs independently, namely, sulfation of one of these sugars is not needed for sulfation of the second[18], it nonetheless appears that Agl30 cannot sulfate a GlcA found at the third position of the tetrasaccharide and that Agl31 cannot sulfate the terminal GlcA unless IdoA is found at the third position. It thus follows that tetrasaccharide sulfation occurs after the appearance of IdoA at position three, and not when GlcA is found at this position. The claim that Agl30 distinguishes between GlcA and IdoA is further supported by the fact that the N-linked tetrasaccharide IdoA is sulfated at the C3 position, a scenario seemingly unique to *Hbt. salinarum*[13], and not at the C2 position, where GlcA sulfation usually occurs[50–53], including in the case of the terminal GlcA of the N-linked tetrasaccharide. Furthermore, the finding that no glycan sulfation occurs in a strain lacking Agl27, which adds the terminal GlcA to the DolP-bound trisaccharide precursor of the tetrasaccharide[16], argues that the GlcA added to the DolP-bound disaccharide precursor of the tetrasaccharide is only converted to IdoA after the terminal GlcA has been added. At the same time, the relative timing of GlcA epimerization into IdoA and sugar sulfation in Eukarya shows variability. In dermatan sulfate and heparin, a neighboring N-acetylglucosamine must be N-sulfated prior to D-GlcA epimerization, whereas in heparan sulfate, epimerization precedes sulfation of any neighboring N-acetylgalatosamine[42].

The absence of VNG1058H impacted *Hbt. salinarum* cell physiology, as reflected by changes in cell growth, cell motility and S-layer structure in the deletion strain. The impacts of such changes seems, however, to be related to growth temperature. qPCR revealed that *VNG1058H* transcription decreased upon growth at 37 or 46 °C, relative to what was measured at 42 °C. At the same time, *VNG1058H* deletion harmed the growth of cells at 42 °C, yet had much far less, if any, impact on cell growth at 37 or 46 °C. As such, for reasons still unclear, it would appear that at these non-optimal growth temperatures, the importance of VNG1058H is decreased, which could explain why *VNG1058H* transcription is reduced when cells were

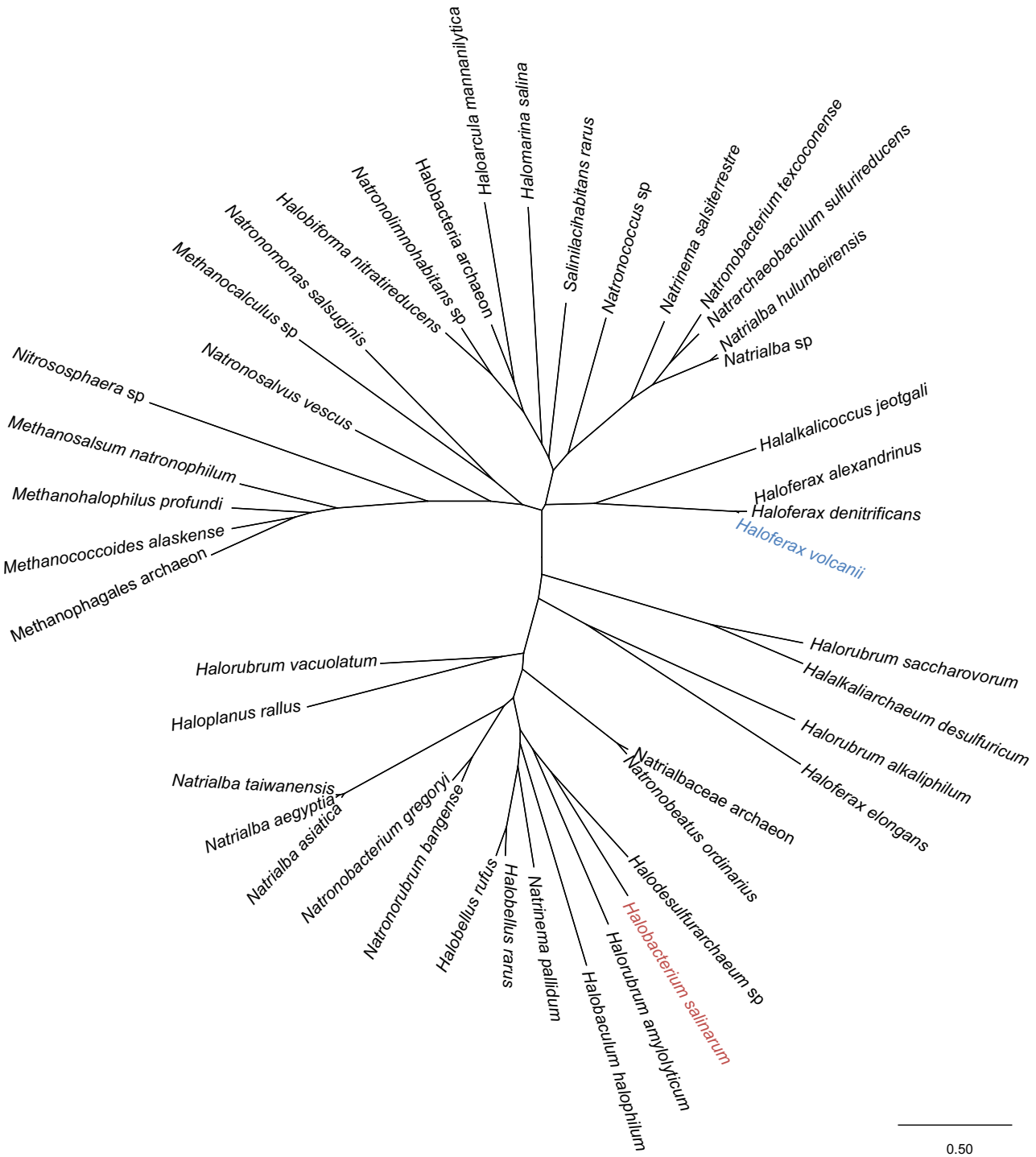


Fig. 7 | Phylogenetic analysis predicts that different epimerases of GlcA appear in numerous archaeal species. A phylogenetic tree of archaeal VNG1058H (GenBank locus tag AAG19462.1) homologs predicts a clade of AglQ-like proteins that includes *Hfx. volcanii* AglQ (blue) and a second clade that includes *Hbt. salinarum* VNG1058H-like proteins (red).

grown at 37 or 46 °C. These findings thus support the concept that *Hbt. salinarum* can control the composition of the N-linked tetrasaccharide in response to the environment. While the presence of L-iduronic acid in GAGs was deemed to be important for protein binding, given how this sugar increases conformational flexibility[54,55], it remains to be determined how *Hbt. salinarum* benefit from replacing GlcA with IdoA at tetrasaccharide position three (and subsequently not sulfating the last two tetrasaccharide sugars).

In summary, although earlier efforts had predicted the existence of archaeal versions of D-glucuronyl C5-epimerase[43], the present investigation provides the first experimental evidence for this prediction. Moreover, VNG1058H represents the first D-glucuronyl C5-epimerase shown to participate in N-glycosylation (Fig. 8). Given the demonstrated role of VNG1058H in *Hbt. salinarum* N-glycosylation, it was renamed Agl32, according to the nomenclature used to annotate components of archaeal N-glycosylation pathways[56].

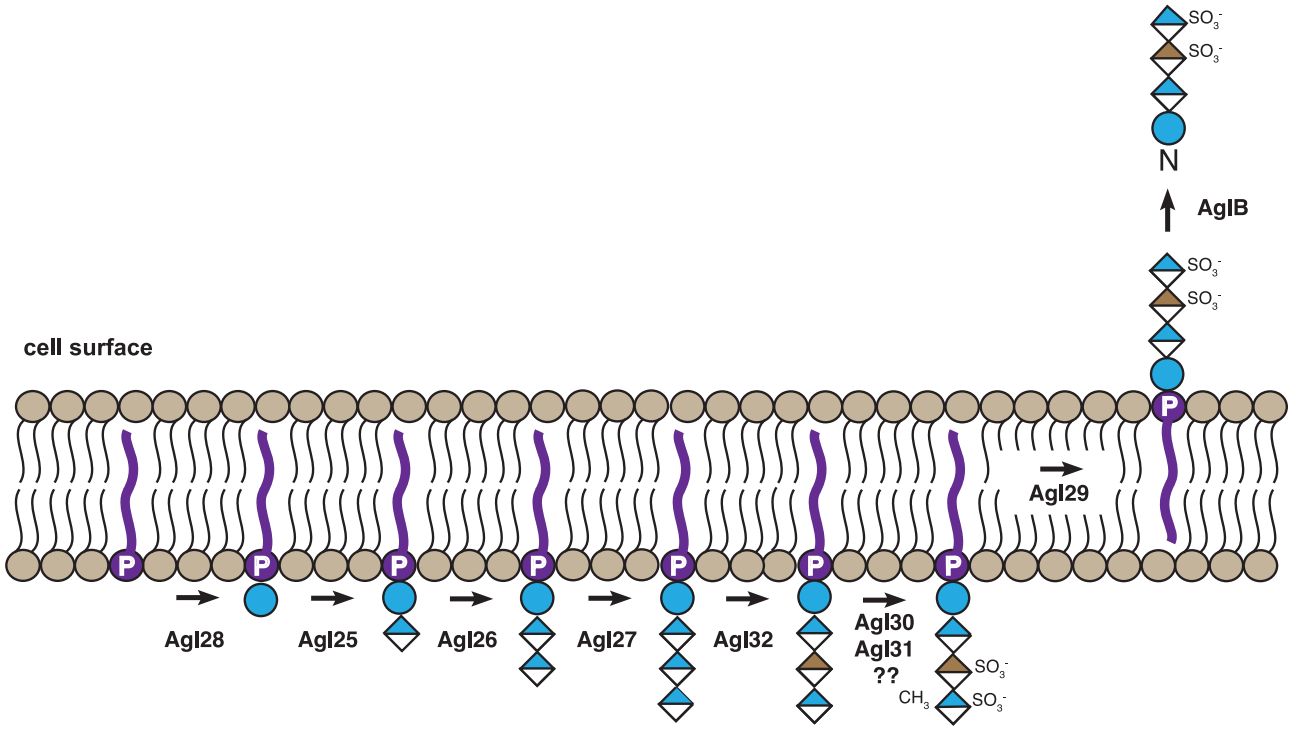


Fig. 8 | Working model of *Hbt. salinarum* N-glycosylation. Assembly of the N-linked tetrasaccharide decorating *Hbt. salinarum* glycoproteins begins with the sequential addition of one Glc and three GlcA units to a common DolP carrier on the cytoplasmic face of the plasma membrane by the glycosyltransferases Agl28, Agl25, Agl26, and Agl27, respectively. The D-glucuronyl C5-epimerase Agl32 then epimerases the GlcA at tetrasaccharide position three into IdoA. The sulfotransferases Agl30 and Agl31 then sulfate IdoA and the terminal GlcA at the O-3 and O-2 positions, respectively. An as yet undefined methyltransferase (indicated by "??") methylates the terminal GlcA. The temporal relation between sulfation and methylation is currently unknown. Likewise, the demethylation enzyme that removes the methyl group and when it acts also remain unclear. Once assembled, the DolP-linked tetrasaccharide is translocated across the membrane in a reaction involving Agl29. Finally, the tetrasaccharide is transferred to selected Asn residues in target proteins by AglB. In the absence of Agl32, a non-sulfated N-linked tetrasaccharide comprising Glc and three GlcA units is added to target proteins. Glc is represented by a blue circle, GlcA by a half blue diamond and IdoA by a half brown diamond. For further details, see refs. 13,16–19.

Methods

Cell growth and *VNG1058H* deletion and complementation

Hbt. salinarum NRC-1 (ATCC strain 700922) parent strain cells were grown in medium containing 250 g NaCl, 20 g $MgSO_4{\cdot}7H_2O$, 3 g sodium citrate, 2 g KCl, 10 g peptone per l, supplemented with 50 µg/ml uracil at 42 °C[57]. Cells deleted of *VNG1058H* were generated and grown as above, with lack of the gene being confirmed by qPCR[16,18]. PCR was also performed to confirm that *VNG1058H* deletion had no detrimental effect on neighboring genes, i.e., *VNG1056C*, *VNG1057C*, and *VNG1059C*. To construct a complementation plasmid, *VNG1058H* was amplified from a gBlock containing the gene encoded to include a 6xHis-Flag tag at the C-terminal end of the product and presenting *Nde*I and *Bam*HI restriction sites at the 5′ and 3′ ends, respectively. The construct was introduced into the plasmid pMTFchis[58] (a kind gift from Amy K. Schmid, Duke University), linearized using the *Nde*I and *Bam*HI restriction enzymes, using a Gibson Assembly Kit (New England BioLabs), according to the manufacturer's instructions. *Hbt. salinarum* parent cells were transformed with plasmid pMTFchis lacking *VNG1058H* (empty vector), while Δ*VNG1058H* cells were transformed with the empty vector or with the construct described above[59]. Transformed cells were grown in medium containing 10 µg/ml mevinolin. All primers used are listed in Table 2.

LC-ESI MS

LC-ESI MS analysis of the *Hbt. salinarum* S-layer glycoprotein and archaellins was performed[16]. Peptides were analyzed using a Q Exactive HF mass spectrometer (Thermo) fitted with a capillary HPLC (easy nLC 1200, ThermoFisher (HF)). The peptides were loaded in solvent A (0.1% formic acid in water) on a homemade capillary column (30 cm, 75-micron ID) packed with Reprosil C18-Aqua (Dr. Maisch HPLC, Ammerbuch, Germany). The peptide mixture was resolved with a 6–34% linear gradient of solvent B (80% acetonitrile with 0.1% formic acid) for 60 min followed by a gradient of 15 min of 34 to 95% and 15 min at 95% solvent B at flow rates of 0.15 µl/min. Mass spectrometry was performed in a positive mode (m/z 300–1500, resolution 60,000 for MS1 and 15,000 for MS2) using repetitively full MS scan followed by high collision dissociation (HCD, at 27 normalized collision energy) of the 18 most dominant ions (>1 charges) selected from the first MS scan. The AGC settings were 3×10^6 for the full MS and 1×10^5 for the MS/MS scans. The intensity threshold for triggering MS/MS analysis was 1.3×10^5. A dynamic exclusion list was enabled with an exclusion duration of 20 s.

Structural modeling

To predict the structure of VNG1058H, the AlphaFold computational modeling tool (https://alphafold.ebi.ac.uk/)[60] was employed in the UCSF ChimeraX (version 1.4) platform[61]. The top-ranked AlphaFold-generated model was then compared against entries in the Protein Data Bank (January, 2025) using the DALI protein structure comparison server (http://ekhidna2.biocenter.helsinki.fi/dali/)[39]. The top hit provided by the DALI server was visualized using UCSF Chimera[62], with the two structures being superimposed using the default settings to determine the quality of the match.

PDB files of glycans were generated using GlycoGlyph[63]. Sticks and van der Waals surface presentations were created using Avogadro version 2.0[64]. Electrostatic potential surface was depicted by mapping atomic Gasteiger charges onto the molecular van der Waals surface.

Purification and NMR analysis of an N-linked tetrasaccharide-containing glycopeptide

Hbt. salinarum archaellum filaments were enriched from the spent growth medium of Δ*VNG1058H* strain cells[19] and the N-linked glycans from the

Table 2 | Primers used in this study

Primers	Sequences
Primers for gene deletion	
VNG1058H-Hind-Gibs-F	CGAGCAGACGCATCTGGATCCACGAAGCTTGAGATGTACGTGTTCGCTGAGAATC
VNG1058H-up-R	CGATAACAGGACAGATACTCTTCAGTGGGGCATGGGTGCGAACAGGCCGCTCG
VNG1058H-down-F	CGAGCGGCCTGTTCGCACCCATGCCCCACTGAAGAGTATCTGTCCTGTTATCG
VNG1058H-Nco-Gibs-R	AGGTATCTAGAACCGGTGACGTCACCATGGGATACGACGTCGAGTATCACATCATCG
Primers for qPCR	
VNG1058H FW	CTGTACTGGATCGTGTGAAGAG
VNG1058H Rev	GGACGTGCGTAGTGCTTTAT
VNG0657G FW	CGGATTCGGTCGAGTTTCAT
VNG0657G Rev	CACATCGTGGTGATCCAGTT
Primers for complementation	
FW1058pMTF	CGGAAGCCGAACTCTGCACATATGCCGTCTAGTTCGAACCGCCA
REV1058pMTFhis-flag	GACTCTAGAACTAGTGGATCCTCACTTATCGTCGTCATCCTTGTAATCCTCGAG
Primers for *VNG1058* and neighboring genes	
VNG1056-up	ATGGCGGGAAATCTCCGT
VNG1056-down	CGAGATGATCGTTGATGTCGGT
VNG1057C-up	CTGAATCGGGTAGCGACATC
VNG1057C-down	ACTCCTCCAGTTGCTCGTA
VNG1058-up	CACATTCGCGGACCTATTGCGCATATGCCGTCTAGTTCGAACCGCCATC
VNG1058-down	ACCGTCTCGTGACAGCCGAATTCTCAGTGTAGGTCTACGTTCGGAATGTC
VNG1059C-up	ATGCCCAGCATCGGAATTCTGT
VNG1059-down	TCAGATCTCGGATCGAAACACGG

archaellins comprising these filaments were isolated using a method adapted from that of Notaro et al.[13]. Briefly, archaellins were digested with proteinase K and the resulting N-glycosylated peptides (1 mg) were separated from the other peptide fragments by size exclusion chromatography using a Biogel P2 column (total volume 20 ml, flow rate of 12 ml/h, DDW as eluent). The eluate was monitored with a refractive index detector, pooled accordingly and analyzed by proton NMR to identify those fractions enriched in glycopeptide. Additional purification was achieved by anion exchange chromatography Q-Sepharose fast-flow (Cytiva, Buccinasco, Italy; total volume 0.5 ml), eluted using NaCl solutions of increasing ionic strength (10 mM, 100 mM, 200 mM, 400 mM, 600 mM, 1000 mM; three volumes of each), followed by a final wash with 1 M NaOH, immediately neutralized to pH 7 after collection. Eluates with the same ionic strength were pooled and desalted on a Biogel P2 column (same conditions as above). This additional purification step efficiently removed traces of starch, as well as other peptides that co-eluted with the glycopeptide. Starch was recovered at low ionic strength (10 mM NaCl), while the N-linked tetrasaccharide-containing peptide (400 µg) eluted at higher ionic strength (100 mM NaCl).

Samples were dissolved in deuterated water (550 µl) and spectra were recorded on a Bruker 600 DRX apparatus equipped with a CryoProbe at 298 K. Homonuclear ^{1}H–^{1}H 2D experiments (COSY, TOCSY, NOESY) were recorded using 512 FIDs of 2048 complex points with 48 scans per FID. Mixing times of 100 and 500 ms were used for TOCSY and NOESY spectra, respectively. ^{1}H–^{13}C HSQC and HMBC spectra were acquired with 512 FIDs of 2048 complex points by accumulating 80 and 50 scans, respectively. All spectra were calibrated on internal acetone (^{1}H 2.225 ppm, ^{13}C 31.35 ppm). Transformation and analysis of the spectra were performed using Bruker Topspin 4.0 software.

Motility assays

Cell motility on semi-solid medium was assayed according to Vershinin et al.[16]. Briefly, liquid cultures of parent and deletion strain cells were grown to logarithmic phase (OD_{600} ~ 0.8) and aliquots (10 µl) were applied to the center of plates of semi-solid medium containing 0.3% agar (w/v). After four days at 42 °C, the diameters of the motility halos, corresponding to the distance the cells had swum, were measured in 13–14 plates per strain. Statistical significance was determined using Student's unpaired t-test.

Proteolytic digestion of the S-layer

To assess the impact of *VNG1058H* deletion on the susceptibility of the *Hbt. salinarum* S-layer to proteolytic digestion, cells (1 ml) of the parent and deletion strains were grown to OD_{600} = 1.0 and challenged with proteinase K (1 mg/ml, final concentration; Sigma) at 37 °C. Aliquots (100 µl) were removed immediately prior to proteinase K addition (considered as time = 0) and at 30 min intervals following addition of the protease for up to 2.5 h. Following separation by 8% SDS-PAGE, the protein contents of each aliquot were Coomassie-stained. The intensities of the stained bands corresponding to the S-layer glycoprotein were determined densitometrically using the Fiji platform[65] and normalized to the intensity of that band in the aliquot taken at time = 0, considered as 1.

Assessing *VNG1058H* transcript levels as a function of growth conditions

To quantify *VNG1058H* transcript levels as a function of growth conditions, parent and Δ*VNG1058H* strain cells were grown to logarithmic phase (OD_{600} ~ 0.8) in medium containing 2.9, 3.5 or 4.2 M NaCl at 42 °C or in medium containing 4.2 M NaCl at 37, 42 or 46 °C. Transcript levels were determined by qPCR analysis using a CFX384TM Real Time System (Bio Rad). The reaction mix contained 5 µl of SYBR green mix (Applied Biosystems), 0.3 µM of primers (listed in Table 2), 5 ng cDNA and DDW in a total reaction volume of 10 µl. The following parameters were used: 95 °C for 3 min, 40 cycles of 15 s at 95 °C and 1 min at 60 °C for annealing, extension and read fluorescence, respectively. Melting curve analysis was performed after each run to ensure the specificity of the products. The efficiency of each

primer set was calculated using five to six serial dilutions of the parent strain sample. Using this efficiency value for each primer set, relative expression was calculated using the standard $2^{-\Delta\Delta ct}$ formula, with the *VNG0657G* housekeeping gene as reference[58] and normalized to the transcript level measured from cells grown in medium containing 4.2 M NaCl at 42 °C, taken as 1.0.

Phylogenetic analysis

An unrooted neighbor-joining phylogenetic tree of *Hbt. salinarum* VNG1058H and *Hfx. volcanii* AglQ homologs was constructed using MEGA software version 11.0.13[66]. The VNG1058H protein sequence was downloaded from the UniProt database (accession: Q9HQQ1) and submitted to BlastP in the MEGA software. The search parameter was set to Archaea (taxid: 2157). Selected DNA sequences were added to the alignment explorer tool in MEGA and aligned using the MUSCLE algorithm, set to "align codons". A tree was constructed using the "general time reversible + gamma distribution (GTR + G)" model.

Statistics and reproducibility

Where relevant, averages were calculated using Student's *t* test. Standard deviation or standard error of the mean was calculated for each average, as described in the text or relevant figure legend. The numbers of biological repeats are also provided.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are found in Supplementary Data 1 or are available from the corresponding author upon reasonable request.

Received: 3 June 2025; Accepted: 2 October 2025;
Published online: 17 November 2025

References

1. Taguchi, Y., Fujinami, D. & Kohda, D. Comparative analysis of archaeal lipid-linked oligosaccharides that serve as oligosaccharide donors for Asn glycosylation. *J. Biol. Chem.* **291**, 11042–11054 (2016).
2. Eichler, J. & Guan, Z. Lipid sugars carriers at the extremes: the phosphodolichols Archaea use in N-glycosylation. *Biochim. Biophys. Acta* **1862**, 589–599 (2017).
3. Larkin, A., Chang, M. M., Whitworth, G. E. & Imperiali, B. Biochemical evidence for an alternate pathway in N-linked glycoprotein biosynthesis. *Nat. Chem. Biol.* **9**, 367–373 (2013).
4. Eichler, J. & Imperiali, B. Stereochemical divergence of polyprenol phosphate glycosyltransferases. *Trends Biochem. Sci.* **43**, 10–17 (2018).
5. Schwarz, F. & Aebi, M. Mechanisms and principles of N-linked protein glycosylation. *Curr. Opin. Struct. Biol.* **21**, 576–582 (2011).
6. Eichler, J. Extreme sweetness: protein glycosylation in Archaea. *Nat. Rev. Microbiol.* **11**, 151–156 (2013).
7. Jarrell, K. F. et al. N-linked glycosylation in Archaea: a structural, functional, and genetic analysis. *Microbiol. Mol. Biol. Rev.* **78**, 304–341 (2014).
8. Notaro, A., Zaretsky, M., Molinaro, A., De Castro, C. & Eichler, J. N-glycosylation in Archaea: unusual sugars and unique modifications. *Carbohydr. Res.* **534**, 108963 (2023).
9. Oesterhelt, D. & Stoeckenius, W. Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. *Nat. N. Biol.* **233**, 149–152 (1971).
10. Mescher, M. F. & Strominger, J. L. Purification and characterization of a prokaryotic glycoprotein from the cell envelope of *Halobacterium salinarium*. *J. Biol. Chem.* **251**, 2005–2014 (1976).
11. Paul, G., Lottspeich, F. & Wieland, F. Asparaginyl-N-acetylgalactosamine. Linkage unit of halobacterial glycosaminoglycan. *J. Biol. Chem.* **261**, 1020–1024 (1986).
12. Lechner, J. & Wieland, F. Structure and biosynthesis of prokaryotic glycoproteins. *Annu. Rev. Biochem.* **58**, 173–194 (1989).
13. Notaro, A., Vershinin, Z., Guan, Z., Eichler, J. & De Castro, C. An N-linked tetrasaccharide from *Halobacterium salinarum* presents a novel modification, sulfation of iduronic acid at the O-3 position. *Carbohydr. Res.* **521**, 108651 (2022).
14. Lechner, J., Weiland, F. & Sumper, M. Biosynthesis of sulfated saccharides N-glycosidically linked to the protein via glucose. Purification and identification of sulfated dolichyl monophosphoryl tetrasaccharides from halobacteria. *J. Biol. Chem.* **260**, 860–866 (1985).
15. Vershinin, Z., Zaretsky, M., Guan, Z. & Eichler, J. Revisiting N-glycosylation in *Halobacterium salinarum*: characterizing a dolichol phosphate- and glycoprotein-bound tetrasaccharide. *Glycobiology* **31**, 1645–1654 (2021).
16. Vershinin, Z., Zaretsky, M., Guan, Z. & Eichler, J. Identifying components of a *Halobacterium salinarum* N-glycosylation pathway. *Front. Microbiol.* **12**, 779599 (2021).
17. Vershinin, Z., Zaretsky, M., Guan, Z. & Eichler, J. Agl28 and Agl29 are key components of a *Halobacterium salinarum* N-glycosylation pathway. *FEMS Microbiol. Lett.* **370**, fnad017 (2023).
18. Zaretsky, M., Vershinin, Z., Erez, L., Grossman-Haham, I. & Eichler, J. Two different sulfotransferases modify sugars of the N-linked tetrasaccharide decorating *Halobacterium salinarum* glycoproteins. *mBio* **16**, e0353424 (2025).
19. Zaretsky, M., Darnell, C. L., Schmid, A. K. & Eichler, J. N-glycosylation is required for archaellin transcription and translation, archaella assembly and cell motility in *Halobacterium salinarum*. *Front. Microbiol.* **10**, 1367 (2019).
20. Wieland, F., Lechner, J. & Sumper, M. Iduronic acid: constituent of sulphated dolichyl phosphate oligosaccharides in halobacteria. *FEBS Lett.* **195**, 77–81 (1986).
21. Vallet, S. D., Berthollier, C. & Ricard-Blum, S. The glycosaminoglycan interactome 2.0. *Am. J. Physiol. Cell Physiol.* **322**, C1271–C1278 (2022).
22. Khan, S. A. et al. Glycosaminoglycans in mucopolysaccharidoses and other disorders. *Adv. Clin. Chem.* **122**, 1–52 (2024).
23. Miyasaki, T., Tsuchihashi, H., Yamada, H. & Yodomae, T. Detection and identification of L-iduronic acid in the glycuronan "protuberic acid" from *Koboyasia nipponica*. *Carbohydr. Res.* **77**, 281–284 (1979).
24. Lahaye, M. & Robic, A. Structure and functional properties of ulvan, a polysaccharide from green seaweeds. *Biomacromolecules* **8**, 1765–1774 (2007).
25. Carroll, A. R., Duffy, S. & Avery, V. M. Citronamides A and B, tetrapeptides from the Australian sponge *Citronia astra*. *J. Nat. Prod.* **72**, 764–768 (2009).
26. Lee, L. & Cherniak, R. Identification of iduronic acid as a constituent of the "type-specific" polysaccharide of *Clostridium perfringens* Hobbs 10. *Carbohydr. Res.* **33**, 387–390 (1974).
27. Hanniffy, O. M. et al. Structure of a highly acidic O-specific polysaccharide of lipopolysaccharide of *Pseudoalteromonas haloplanktis* KMM 223 (44-1) containing L-iduronic acid and D-QuiNHb4NHb. *Carbohydr. Res.* **307**, 291–298 (1998).
28. Stack, R. J., Plattner, R. D. & Cote, G. L. Identification of L-iduronic acid as a constituent of the major extracellular polysaccharide produced by *Butyrivibrio fibrisolvens* strain X6C61. *FEMS Microbiol. Lett.* **51**, 1–5 (1988).
29. Hagner-McWhirter, A. et al. Biosynthesis of heparin/heparan sulfate: kinetic studies of the glucuronyl C5-epimerase with N-sulfated

derivatives of the *Escherichia coli* K5 capsular polysaccharide as substrates. *Glycobiology* **10**, 159–171 (2000).
30. Li, J. P. Glucuronyl C5-epimerase an enzyme converting glucuronic acid to iduronic acid in heparan sulfate/heparin biosynthesis. *Prog. Mol. Biol. Transl. Sci.* **93**, 59–78 (2010).
31. Raedts, J., Lundgren, M., Kengen, S. W., Li, J. P. & van der Oost, J. A novel bacterial enzyme with D-glucuronyl C5-epimerase activity. *J. Biol. Chem.* **288**, 24332–24339 (2013).
32. Qin, Y. et al. Structural and functional study of D-glucuronyl C5-epimerase. *J. Biol. Chem.* **290**, 4620–4630 (2015).
33. Debarnot, C. et al. Substrate binding mode and catalytic mechanism of human heparan sulfate d-glucuronyl C5 epimerase. *Proc. Natl. Acad. Sci. USA* **116**, 6760–6765 (2019).
34. Kandiba, L. & Eichler, J. Deciphering a pathway of *Halobacterium salinarum* N-glycosylation. *Microbiologyopen* **4**, 28–40 (2015).
35. Jarrell, K. F. & Albers, S. V. The archaellum: an old motility structure with a new name. *Trends Microbiol.* **20**, 307–312 (2012).
36. Wieland, F., Paul, G. & Sumper, M. Halobacterial flagellins are sulfated glycoproteins. *J. Biol. Chem.* **260**, 15180–15185 (1985).
37. Sofer, S. et al. Perturbed N-glycosylation of *Halobacterium salinarum* archaellum filaments leads to filament bundling and compromised cell motility. *Nat. Commun.* **15**, 5841 (2024).
38. Arbiv, A., Yurist-Doutsch, S., Guan, Z. & Eichler, J. AglQ is a novel component of the *Haloferax volcanii* N-glycosylation pathway. *PLoS ONE* **8**, e81782 (2013).
39. Holm, L. Dali server: structural unification of protein families. *Nucleic Acids Res.* **50**, W210–W215 (2022).
40. Park, C. S. et al. Characterization of a recombinant mannobiose 2-epimerase from *Spirochaeta thermophila* that is suggested to be a cellobiose 2-epimerase. *Biotechnol. Lett.* **35**, 1873–1880 (2013).
41. Saburi, W. Functions, structures, and applications of cellobiose 2-epimerase and glycoside hydrolase family 130 mannoside phosphorylases. *Biosci. Biotechnol. Biochem.* **80**, 1294–1305 (2016).
42. Valla, S., Li, J., Ertesvåg, H., Barbeyron, T. & Lindahl, U. Hexuronyl C5-epimerases in alginate and glycosaminoglycan biosynthesis. *Biochimie* **83**, 819–830 (2001).
43. Raedts, J., Kengen, S. W. & van der Oost, J. Occurrence of L-iduronic acid and putative D-glucuronyl C5-epimerases in prokaryotes. *Glycoconj. J.* **28**, 57–66 (2011).
44. Speciale, I. et al. Liquid-state NMR spectroscopy for complex carbohydrate structural analysis: a hitchhiker's guide. *Carbohydr. Polym.* **277**, 118885 (2022).
45. Tamir, A. & Eichler, J. N-glycosylation is important for proper *Haloferax volcanii* S-layer stability and function. *Appl. Environ. Microbiol.* **83**, e03152–16 (2017).
46. Nikolayev, S., Cohen-Rosenzweig, C. & Eichler, J. Evolutionary considerations of the oligosaccharyltransferase AglB and other aspects of N-glycosylation across Archaea. *Mol. Phylogenet. Evol.* **153**, 106951 (2020).
47. Kandiba, L., Lin, C. -w, Aebi, M., Eichler, J. & Guerardel, Y. Structural characterization of the N-linked pentasaccharide decorating glycoproteins of the halophilic archaeon *Haloferax volcanii*. *Glycobiology* **26**, 745–756 (2016).
48. Hasan, M. et al. The structure of human dermatan sulfate epimerase 1 emphasizes the importance of C5-epimerization of glucuronic acid in higher organisms. *Chem. Sci.* **12**, 1869–1885 (2020).
49. Lindahl, U., Bäckström, G., Malmström, A. & Fransson, L. Å Biosynthesis of L-iduronic acid in heparin: epimerization of D-glucuronic acid on the polymer level. *Biochem. Biophys. Res. Commun.* **46**, 985–991 (1972).
50. Jones, C. J., Beni, S., Limtiaco, J. F., Langeslay, D. J. & Larive, C. K. Heparin characterization: challenges and solutions. *Annu. Rev. Anal. Chem.* **4**, 439–465 (2011).
51. Pomin, V. H. & Mourão, P. A. Specific sulfation and glycosylation-a structural combination for the anticoagulation of marine carbohydrates. *Front. Cell. Infect. Microbiol.* **4**, 33 (2014).
52. Casu, B., Naggi, A. & Torri, G. Re-visiting the structure of heparin. *Carbohydr. Res.* **403**, 60–68 (2015).
53. Pretorius, D., Richter, R. P., Anand, T., Cardenas, J. C. & Richter, J. R. Alterations in heparan sulfate proteoglycan synthesis and sulfation and the impact on vascular endothelial function. *Matrix Biol.***16**, 100121 (2022).
54. Casu, B. & Lindahl, U. Structure and biological interactions of heparin and heparan sulfate. *Adv. Carbohydr. Chem. Biochem.* **57**, 159–206 (2001).
55. Thelin, M. A. et al. Biological functions of iduronic acid in chondroitin/dermatan sulfate. *FEBS J.* **280**, 2431–2446 (2013).
56. Eichler, J., Jarrell, K. F. & Albers, S. A proposal for the naming of N-glycosylation pathway components in Archaea. *Glycobiology* **23**, 620–621 (2013).
57. Darnell, C. L., Tonner, P. D., Gulli, J. G., Schmidler, S. C. & Schmid, A. K. Systematic discovery of archaeal transcription factor functions in regulatory networks through quantitative phenotyping analysis. *mSystems* **2**, e00032–17 (2017).
58. Darnell, C. L. et al. The ribbon-helix-helix domain protein CdrS regulates the tubulin homolog ftsZ2 to control cell division in archaea. *mBio.* **11**, e01007–e01020 (2020).
59. Peck, R. F., DasSarma, S. & Krebs, M. P. Homologous gene knockout in the archaeon *Halobacterium salinarum* with *ura3* as a counterselectable marker. *Mol. Microbiol.* **35**, 667–676 (2000).
60. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
61. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
62. Pettersen, E. F. et al. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
63. Mehta, A. Y. & Cummings, R. D. GlycoGlyph: a glycan visualizing, drawing and naming application. *Bioinformatics* **36**, 11 (2020).
64. Hanwell, M. D. et al. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **4**, 17 (2012).
65. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
66. Tamura, K., Stecher, G. & Kumar, S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).

Acknowledgements

This work supported by the Israel Science Foundation (grant 414/20 to J.E. and grant 1691/23 to I.G.-H.). D.Y. is a Ph.D. student supported by a fellowship from the H2020 Marie Skłodowska-Curie Actions Doctoral Network (project GLYCO-N, grant 101119499).

Author contributions

Z.V. and M.Z. performed the genetics, mass spectrometry, cell physiology and phylogenetic studies; Z.V., S.S. and I.G.-H. performed the modeling studies; and A.N., D.Y. and C.D.C. performed the NMR studies. A.M., I.G.-H., C.D.C. and J.E. obtained funding and J.E. supervised the research. All authors reviewed and approved the submitted text and figures.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-025-08987-9.

Correspondence and requests for materials should be addressed to Jerry Eichler.

Peer review information *Communications Biology* thanks Takahiro Yamasaki, Satoshi Nakagawa and the other anonymous reviewer(s) for their contribution to the peer review of this work. Primary handling editor: Ophelia Bu.