

N-glycolyl chondroitin synthesis using metabolically engineered *E. coli*

Adeola E. Awofiranye

Rensselaer Polytechnic Institute

Sultan N. Baytas

Rensselaer Polytechnic Institute

Ke Xia

Rensselaer Polytechnic Institute

Abinaya Badri

Rensselaer Polytechnic Institute

Wenqin He

Rensselaer Polytechnic Institute

Ajit Varki

University of California San Diego

Mattheos Koffas

Rensselaer Polytechnic Institute

Robert J Linhardt (✉ Linhar@rpi.edu)

Rensselaer Polytechnic Institute <https://orcid.org/0000-0003-2219-5833>

Original article

Keywords: Sialic acid, biotransformation, *N*-glycolyl chondroitin, metabolite, *N*-glycolyl glucosamine

Posted Date: August 5th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-35611/v2>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at AMB Express on August 17th, 2020. See the published version at <https://doi.org/10.1186/s13568-020-01084-6>.

Abstract

N-glycolyl chondroitin (Gc-CN) is a metabolite of *N*-glycolylneuraminic acid (Neu5Gc), a sialic acid that is commonly found in mammals, but not humans. Humans can incorporate exogenous Neu5Gc into their tissues from eating red meat. Neu5Gc cannot be biosynthesized by humans due to an evolutionary mutation and has been implicated in causing inflammation causing human diseases, such as cancer. The study Neu5Gc is important in evolutionary biology and the development of potential cancer biomarkers. Unfortunately, there are several limitations to detecting Neu5Gc. The elimination of Neu5Gc involves a degradative pathway leading to the incorporation of *N*-glycolyl groups into glycosaminoglycans (GAGs), such as Gc-CN. Gc-CN has been found in humans and in animals including mice, lamb and chimpanzees. Here, we present the biosynthesis of Gc-CN in bacteria by feeding chemically synthesized *N*-glycolylglucosamine to *Escherichia coli*. A metabolically engineered strain of *E. coli* K4, fed with glucose supplemented with GlcNGc, converted it to *N*-glycolylgalactosamine (GalNGc) that could then be utilized as a substrate in the chondroitin biosynthetic pathway. The final product, Gc-CN was converted to disaccharides using chondroitin lyase ABC and analyzed by liquid chromatography-tandem mass spectrometry with multiple reaction monitoring detection. This analysis showed the incorporation of GalNGc into the backbone of the chondroitin oligosaccharide.

Key Points

1. *N*-glycolyl chondroitin (Gc-CN) is a stable metabolite of Neu5Gc
2. Metabolic engineering of *E. coli* K4
3. *E. coli* has promiscuous enzymes involved in CN biosynthesis
4. Feeding *E. coli* K4 with GlcNGc affords Gc-CN

Introduction

Sialic acids constitute a family of acidic sugars with a 9-carbon backbone found at the terminal end of glycan chains attached to many soluble glycoproteins (Schauer 2000; Wang and Brand-Miller 2003). Their presence at the terminal end of sugars gives them an advantage in performing their biological roles, providing structure, serving as a ligand for intrinsic and extrinsic receptors, as a binding site for pathogens and toxins, and in molecular mimicry for host invasion (Varki 2008). Sialic acids are mainly expressed in vertebrates and certain bacteria (Angata and Varki 2002). In mammals, the two most common sialic acids found are *N*-acetylneuraminic acid (Neu5Ac) and *N*-glycolylneuraminic acid (Neu5Gc). An enzyme called cytidine monophosphate *N*-acetylneuraminic acid hydroxylase (Cmah), encoded by the *CMAH* gene, is responsible for converting CMP-Neu5Ac to CMP-Neu5Gc through the addition of an oxygen atom (Kozutsumi et al. 1990; Bergfeld et al. 2012)) (Fig. 1). A few million years ago, a mutation event that occurred in our last common ancestor with the apes, resulted in humans

losing the *CMAH* gene and being unable to convert Neu5Ac to Neu5Gc (Chou et al. 1998; Chou et al. 2002).

However, small amounts of Neu5Gc are metabolically incorporated into human tissues from exogenous dietary sources of Neu5Gc like red meat (Samraj et al. 2015). Exogenous sources include other mammals that still synthesize Neu5Gc. Since the human body does not biosynthesize Neu5Gc, a reaction leading to inflammation is triggered when ingested Neu5Gc is incorporated into tissues. The immune system recognizes Neu5Gc as a foreign molecule (xeno-autoantigen) and produces anti-Neu5Gc antibodies (xeno-autoantibodies) present postnatally due to a commensal bacteria (Taylor et al. 2010). These antibodies recognize these foreign molecules (Padler-Karavani et al. 2013) and triggers inflammation (xenosialitis), which is hypothesized to be a key contributor to diseases associated with red meat consumption (Higashi et al. 1977; Samraj et al. 2014). The accumulation of dietary Neu5Gc can lead to local chronic inflammation largely in epithelial and endothelial tissues, and contribute to human pathologies. (Tangvoranuntakul et al. 2003; Soulillou et al. 2020). Some carcinomas that have been shown to possess accumulation of Neu5Gc due to its incorporation into epithelial tissues, leading to cancers like lung, gastric, ovarian, prostate and colorectal cancers (Marquina et al. 1996; Carr et al. 2000; Padler-Karavani et al. 2012). Apart from cancers, Neu5Gc when incorporated into endothelial cell can lead to inflammation and cause diseases such as Kawasaki, atherosclerosis and other cardiovascular diseases (Arita et al. 1982; Padler-Karavani et al. 2013; Fernández-Ruiz 2019; Yehuda and Padler-Karavani 2020; Kawanishi et al. 2019).

Moreover, when *Cmah*^{-/-} mice (human-like Neu5Gc-deficient mice) were fed with dietary Neu5Gc and challenged with anti-Neu5Gc antibodies, they developed inflammation, and prolonged exposure to these conditions led to significantly higher cases of cancer (Samraj et al. 2015).

Despite all these finding, proving this hypothesis connecting dietary Neu5Gc to diseases in humans is difficult due to the limitations in directly studying the trace amount of Neu5Gc incorporated. There is no reliable non-invasive method of detecting Neu5Gc incorporation in humans, and the most commonly used methods for detecting sialic acid involves tissue processing (serum does not contain detectable amounts of Neu5Gc) (Tangvoranuntakul et al. 2003). Using antibodies to detect Neu5Gc is also challenging. For example, the Neu5Gc antibody-binding site accommodates 4-5 monosaccharides along with Neu5Gc making it difficult for antibodies to recognize and bind free Neu5Gc not part of a glycan (Padlan and Kabat 1988; Dhar et al. 2019). Another challenge with the detection of Neu5Gc using antibodies is that while some anti-Neu5Gc monoclonal antibodies specifically bind Neu5Gc containing gangliosides (Samraj et al. 2014), gangliosides do not survive paraffin embedding required for immunohistochemistry. This means that some molecules carrying Neu5Gc might not survive to the point of detection.

Human metabolism of Neu5Gc-containing glycans can result in the incorporation of the *N*-glycolyl group into the glycosaminoglycans (GAGs) chondroitin sulfate and unsulfated chondroitin (CN). These GAGs are relatively stable and easy to isolate from serum and analyze making these better candidates for

testing the role of dietary Neu5Gc in human diseases. The catabolism of Neu5Gc containing glycans involves six metabolic steps leading to the production of uridine diphosphate *N*-glycolylglucosamine (GlcNGc-UDP) and ultimately uridine diphosphate *N*-glycolylgalactosamine (GalNGc-UDP) (Bergfeld et al. 2017). GalNGc-UDP can then be incorporated into Gc-CN and Gc-CS (Bergfeld et al. 2017) (Fig. 2).

These Gc-CN and Gc-CS derivatives are better suited for analysis because they can be detected from even small volumes of serum and plasma (Lu et al. 2010). Over 95% of GAGs, including CS, can form complexes with plasma proteins (Calatroni et al. 1992) and, therefore, can be isolated from the plasma. Moreover, bikunin is a major CS/CN containing proteoglycan found in both plasma and urine (Chi et al. 2008). A direct method of measuring circulating Gc-CS in these biological samples could serve as an indicator of Neu5Gc incorporation. These samples allow easy detection using reliable analytical methods including high performance liquid chromatography (HPLC) and mass spectrometry (MS), avoiding the need for antibodies (Lu et al. 2010). Gc-CS and its metabolites are very stable as demonstrated by detectable amounts Gc-CS in a 4 million-year-old fossil (Bergfeld et al. 2017).

A microbial source of CN was selected to avoid the risk of sample impurities commonly associated with an animal-derived GAG. Chondroitin sulfate (CS) has been traditionally extracted from animals. Recent efforts to produce CS in microorganism have been successful. CS has been biosynthesized from *E. coli* K4, *Pasteurella multocida*, and *Bacillus megaterium* (Sugiura et al. 2002; He et al. 2015; Jin et al. 2016; Cimini et al. 2018; Restaino et al. 2019). Prior studies have shown that the capsular polysaccharide (CPS) of *E. coli* K4 serotype O5:K4:H4 has a backbone with a repeating disaccharide unit of $\rightarrow 4) -\beta$ -D glucuronic acid (GlcA) (1 \rightarrow 3)- β -D-N-acetylgalactosamine (GalNAc), just like chondroitin. However, this CPS also contains an unwanted linked β -fructofuranose branching from the C3 position of GlcA (Rodriguez et al. 1988). The deletion of the *kfoE* gene, responsible for fructosylation, affords an engineered strain that has been optimized for increased chondroitin production (He 2017). In the current study we hypothesized that feeding GlcNGc to *E. coli* could drive the incorporation of *N*-glycolylgalactosamine into chondroitin, through a pathway similar to that involved in *N*-glycolyl incorporation into animal chondroitin sulfate.

Materials And Methods

Plasmid construction

E. coli K4 serotype O5:K4:H4 (U141, 11307) was engineered for the synthesis of chondroitin sulfate. The fructosyltransferase encoded by *kfoE* was deleted using λ red recombinase (Datsenko and Wanner 2000), resulting in strain K4_Δ*kfoE*. The FRT-flanked kanamycin resistance cassette was PCR amplified from pKD4 by deletion primers with 40 nucleotides homologous regions with a target gene on the genome. The PCR product was purified by a PCR cleanup kit (Cycle Pure Kit, Omega) and transformed into the λ red recombinase expressing *E. coli* K4 strain by electroporation. This system enabled the deletion of the *kfoE* gene and its replacement with an antibiotic resistance gene. Finally, positive knockout strains were screened by colony PCR. Two primers were used in this study. The k4_dkfoE_F primer was 5'

TGCAATATGACCTTAGAAGAGATTTCTAATATGTTAGAACAGGAGAAAAACACGTCTTGAGCGATTGTG3'. The k4_dkfoe_R primer was 5' ATATCCAGCCTTGAAAAACGCGAACTCATCCCCGCCATTGGAATTATAAACGGCTGACATGGGAATTAG3'.

Media

Shake flask fermentations utilized rich defined medium developed from modified protocols (Cirino et al. 2006; Neidhardt et al 1974) (5.0 g/L K₂HPO₄, 3.5 g/L KH₂PO₄, 3.5 g/L (NH₄)₂HPO₄, 100 mL of 10X MOPS buffer, (83.7 g/L MOPS, 7.2 g/L Tricine, 28 mg/L FeSO₄·7H₂O, 29.2 g/L NaCl, 5.1 g/L NH₄Cl, 1.1 g/L MgCl₂, 0.5 g/L K₂SO₄, 0.2 mL micronutrient stock), 1 mL of 1 M MgSO₄, 1 mL of 0.5 g/L thiamine HCl, 0.1 mL of 1 M CaCl₂, 20 g/L glucose, with 12.5 mM GlcNGc. Micronutrient Stock consisted of 0.2 g/L (NH₄)₆Mo₇O₂₄, 1.2 g/L H₃BO₃, 0.1 g/L CuSO₄, 0.8 g/L MnCl₂, and 0.1 g/L ZnSO₄. *E. coli* K4 serotype O5:K4 (L):H4 was from American Type Culture Collection (ATCC 23502). All reagents for medium preparation were from Sigma Chemical Co. (St. Louis, MO).

Shake flask experiments

Shake flask experiments were used to evaluate the *N*-glycolyl glucosamine feeding experiments. *E. coli* K4 Δ kfoE, cells from 15% glycerol stock were streaked on an agar plate containing 50 mg/ml of kanamycin and grown overnight. Two colonies from the plate were picked for duplicate sample analysis, and pre-cultures were grown overnight at 37 °C. The samples were then diluted to 100 ml at an optical density (OD) 0.05 and transferred to a 250 ml Erlenmeyer flask and incubated at 37 °C with shaking at 220 rpm. The cultures were left to grow under the same conditions for an additional 48 h. The cell growth and chondroitin production was similar to that previously reported in our laboratory (He et al. 2015).

N-glycolyl chondroitin purification

CPS was purified from the cell pellet by re-suspending in water and autoclaving in the liquid cycle for 15 min. The autoclaved solution was centrifuged, and the supernatant was collected. Autoclaved supernatant from the cell pellet and cell culture supernatant were precipitated with 80 vol% cold ethanol and stored in an explosion-proof refrigerator overnight at 20 °C. Both intracellular and extracellular chondroitin were recovered. The pellet was collected and re-suspended in buffer (100 mM Tris, pH 7.5, 50 mM MgCl₂, 10 mM CaCl₂) and DNase (1 mg/L, Sigma) was added. The sample was incubated at 37 °C for 1 h after which protease K (2.5 mg/mL, Sigma) was then added, and the sample was incubated at 56 °C for 2 h. After a second precipitation from 80% cold ethanol, the dry pellet was collected, re-dissolved in water (~1 mL), and filtered using a 10 KDa spin column (Amicon Ultra, Millipore) to remove small peptides and salt. The chondroitin obtained was sufficiently free of other impurities or other polysaccharides to undertake disaccharide analysis.

Sample digestion into GAG disaccharides

Digestion buffer (50 mM NH₄OAc containing 2 mM CaCl₂ adjusted to pH 7.0) was added to the sample. Recombinant chondroitin lyase ABC (10 mU each, pH optimum 7.4) was added to each sample and mixed. The samples were incubated at 37 °C for 3 days. Under these reaction conditions, chondroitin lyase ABC could depolymerize their GAG substrates (in amounts of over 100 µg) into GAG disaccharides. The samples were washed twice with 100 µL distilled water in 3K MWCO filter unit. The filtrates passing through the filter unit contained disaccharide products, and these were dried using a vacuum centrifuge and stored at -20 °C for AMAC-labeling.

AMAC labeling

The dried samples were AMAC-labeled by adding 10 µL of 0.1 M AMAC in DMSO/acetic acid (17/3, V/V) and incubating at room temperature for 10 min, followed by the addition of 10 µL of 1 M aqueous NaBH₃CN and incubating for 1 h at 45 °C. After the AMAC-labeling reaction, the samples were centrifuged, and each supernatant was recovered. Samples were stored in a light-resistant container at room temperature until analyzed by LC-MS/MS.

LC-MS/MS Analysis

LC was performed on an Agilent 1200 LC system at 45 °C using an Agilent Poroshell 120 ECC18 column (2.7 µm, 3.0 × 50 mm) with a flow rate of 300 µL/min. Mobile phase A (MPA) was 50 mM NH₄OAc aqueous solution, and the mobile phase B (MPB) was pure methanol. The concentration of MPB increased from 5% to 45% for 10 min, then rose to 100% MPB in the next 0.2 min, and a 4 min flow of 100% MPB was applied to elute all compounds. A triple quadrupole mass spectrometer equipped with an ESI source (Thermo Fisher Scientific, San Jose, CA) was used as a detector. The online MS analysis was at the multiple reaction monitoring (MRM) mode. The conditions and collision energies for all of the disaccharides MRM transitions are listed in our previous publication (Sun et al. 2015). The data analysis was performed on Thermo Xcalibur software.

N-glycolyl glucosamine (GlcNGc)-synthesis and chemical characterization

All reagents were purchased from commercial vendors and, unless otherwise noted, used without further purification. "Brine" refers to a saturated aqueous solution of sodium chloride. Thin-layer chromatography (TLC) was performed using Merck Kieselgel 60F254 pre-coated aluminum backed plates. Plates were visualized using 5% H₂SO₄ in methanol. NMR experiments were performed on a Bruker Advance III 600 MHz spectrometer (Bruker Bio Spin, Billerica, MA) with Topspin 3.2 software (Bruker). N-glycolyl glucosamine (GlcNGc) was synthesized following a previously published strategy (Fig. 3). Briefly, the hydrochloride salt of 2-amino-2-deoxy-1,3,4,6-tetra-*O*-acetyl-β-D-glucopyranose **1** was treated with acetoxyacetyl chloride in the presence of triethylamine to afford acetoxyacetyl residue on the amine group. By applying Zemplen conditions, fully protected intermediate **2** was next deprotected, and GlcNGc was achieved in good yield. Spectral properties of the desired products compounds **2**, and GlcNGc matched those described previously (Sinaý 1971; Bergfeld et al. 2012; Macauley et al. 2012).

1,3,4,6-Tetra-O-acetyl-2-acetyloxyacetamido-2-deoxy-glucopyranose 1

The hydrochloride salt of 2-amino-2-deoxy-1,3,4,6-tetra-*O*-acetyl- β -D-glucopyranose **1** (2.0 g, 5.2 mmol) was dissolved in CH₂Cl₂ (20 mL), then triethylamine (2.2 mL, 16 mmol) was added to the solution. The reaction mixture was cooled to 0 °C, and acetoxyacetyl chloride (0.7 mL, 6.5 mmol) was added. The resultant mixture was stirred for 3 h at room temperature. The reaction was monitored by TLC. When the reaction was completed, the mixture was diluted with EtOAc, and the organic phase was washed successively with water, 1 M NaOH, 0.1 M HCl and, brine. The organic phase was dried over MgSO₄, filtered, and concentrated to yield a white crystalline solid. The material was recrystallized using a mixture of ethyl acetate and hexanes to yield the desired compound **2** as a white solid (1.9 g, 81%). ¹H NMR (600 MHz, CDCl₃), δ (ppm) 6.18 (1H, d, J = 9.2 Hz), 5.68 (1H, d, J = 8.8 Hz), 5.17–5.10 (2H, m), 4.52 (1H, d, J = 15.4 Hz), 4.39 (1H, d, J = 15.4 Hz), 4.30 (1H, ddd, J = 9.1 Hz), 4.23 (1H, dd, J = 4.3 Hz, J = 12.3 Hz), 4.10 (1H, dd, J = 2.2 Hz), 3.78 (1H, ddd, J = 9.2 Hz), 2.14 (3H, s), 2.09 (3H, s), 2.07 (3H, s), 2.02 (3H, s), 2.01 (3H, s). ¹³C NMR (150 MHz, CDCl₃), δ (ppm) 171.3, 170.6, 169.7, 169.6, 169.2, 167.5, 92.4, 73.0, 72.1, 67.6, 62.6, 61.6, 52.7, 20.8, 20.7, 20.5. ESI (m/z): [M-H]⁻ calcd. for C₁₈H₂₄NO₁₂ 446.1298, found m/z 446.1296.

2-Deoxy-2-hydroxyacetamido-D-glucopyranose GlcNGc

To a solution of 1,3,4,6-tetra-*O*-acetyl-2-acetyloxyacetamido-2-deoxy- β -D-glucopyranose **2** (1.5 g, 3.35 mmol) in MeOH (50 ml) at 0°C, 0.5 M NaOMe (5 ml) was added dropwise. The solution was stirred 2 h at room temperature. Amberlite IR 120 H⁺ resin was added to neutralize the reaction and was filtered. The filtrate was concentrated under reduced pressure to afford the desired product as a white solid (693 mg, 80%). ¹H NMR (600 MHz, H₂O) δ (ppm) 5.12 (0.7H, d, J = 3.0 Hz), 4.04 (s, 2H), 3.86 (1H, dd, J = 3.0 Hz, J = 10.2 Hz), 3.79 (1H, m), 3.75 (1H, m), 3.72 (1H, m), 3.42 (1H, t, J = 9.6 Hz), 3.39 (1H, m). ESI (m/z): [M-H]⁻ calcd. for C₁₈H₂₄NO₁₂ 236.0770, found m/z 236.0772.

Results

We exhaustively digested chondroitin samples with chondroitin lyase ABC to afford GAG disaccharides and remove any other minor impurities, such as residual lipopolysaccharides. These digests were then purified and AMAC labeled for LC-MS/MS MRM to verify the presence of Gc-CN. A low capillary temperature and spray voltage was maintained to minimize in-source fragmentation of AMAC-labeled Gc-CN oligosaccharides in MRM (Sun et al. 2015). The 572/396 precursor/product ion pair was used for AMAC-labeled chondroitin disaccharide transition. For AMAC-labeled Gc-CN disaccharide, there is a 16 Da molecular mass increase in the precursor at a mass of 588 used. In the product ion survey, the most sensitive product ion generated had a mass of 412. Therefore, the precursor/product ion pair of 588/412 was used for AMAC-labeled Gc-CN disaccharide transition. From our chondroitin producing construct, serotype O5:K4:H4, only chondroitin was produced as signified by the absence of AMAC-labeled Gc-CN disaccharide (Fig. 5a). From the GlcNGc fed culture, both AMAC-labeled CN and Gc-CN disaccharides were observed at the expected m/z ratio (Fig. 5b). The peak corresponding to the AMAC-labeled Gc-CN

disaccharide had an area of 93312, with a very high signal to noise value of 27389. MRM unequivocally demonstrated that Gc-CN had been synthesized. Chondroitin and Gc-CN may have different ionization efficiencies, and since there is no Gc-CN standard, it is difficult to calculate the absolute concentration of Gc-CN. Thus, we are unable to estimate the percent conversion of chondroitin to Gc-CN, although based on peak areas it is probably produced at a relatively low level (<1%).

Discussion

In this study, we explored the use of metabolically engineered *E.coli* K4 strain, fed with GlcNGc, to produce Gc-CN. The K4 wild type strain was first modified by deleting the *kfoE* gene to prevent the fructosylation of its CPS (He 2017). The resulting organism produces a substantial concentration of chondroitin (~300 mg/L) that is both secreted and intracellular. This bacterial strain produces a cell capsule containing unsulfated chondroitin (Rodriguez et al. 1988). The chondroitin present on the cell capsule and within the cell as well as the shed chondroitin can be isolated from the cell pellet and supernatant, respectively. Moreover, studies have shown that chondroitin is present both intracellularly and extracellularly (He et al. 2015). Small-scale shake flask experiments were carried out on this strain with growth on rich, defined media. The results verified the production of chondroitin using LC-MS/MS-MRM. This method required the conversion of chondroitin to the chondroitin disaccharide, Δ UAGalNAc (where Δ UA is a 4-deoxy- α -L-threo-hex-4-enopyranosyluronic acid), by using chondroitin ABC lyase. The resulting disaccharide was then analyzed by an ultrasensitive and reliable method for disaccharide analysis LC-MS/MS MRM with a limit of detection (LOD) < 1 ng/mL and limit of quantification (LOQ) < 4 ng/mL. (Sun et al. 2015; Yu et al. 2019)

We next synthesized GlcNGc (Fig. 3), a sugar that we hypothesized would be converted to GalNGc and incorporated in chondroitin during its biosynthesis (Fig. 4) to produce Gc-CN. *E. coli* K4 cells in their CPS biosynthetic pathway make both UDP-GlcA and UDP-GalNAc, the two sugars required to biosynthesize chondroitin. We hypothesized that on feeding GlcNGc engineered *E. coli* K4 would synthesize UDP-GalNGc, and together with UDP-GlcA, would be converted to Gc-CN. As expected, UDP-GalNGc was synthesized using this pathway (Fig. 4), and *E.coli* phosphotransferase system (PTS) was able to transport GlcNGc along with Glucose and GlcNAc into the cell. GlcNGc, taken up into the cell through the phosphotransferase system (PTS), was phosphorylated by *N*-acetyl-D-glucosamine kinase (*NagK*) (Uehara and Park 2004) to afford GlcNGc-6-P. Next, phosphoglucosamine mutase (*GlmM*), responsible for the conversion of glucosamine-6-phosphate to glucosamine-1-phosphate, (Jolly et al. 1999) catalyzed the conversion of GlcNGc-6-P to GlcNGc-1-P. Next, GlcNGc was converted by the bifunctional enzyme (*GlmU*) (Mengin-Lecreulx and van Heijenoort 1994), to UDP-GlcNGc. UDP-glucose 4-epimerase (*KfoA*) (Zhu et al. 2018) catalyzed the conversion of UDP-GlcNGc to UDP-GalNGc. UDP-GalNGc and UDP-GlcA were then converted to Gc-CN through the action of a chondroitin polymerase, (*KfoC*) (Zanfardino et al. 2010). This pathway demonstrates that all the enzymes in this biosynthetic pathway (Fig. 4) are able to accept the *N*-glycolyl functional group in place of the natural *N*-acetyl functional group, suggesting enzyme promiscuity.

These preliminary studies showed about 1% conversion to Gc-CN. The exact concentration of synthesized Gc-CN could not be determined due to the absence of standards. Comparatively, a Gc-CN could be chemically synthesized in a 10% yield (Bergfeld et al. 2017). There are two likely reasons for the relatively low levels of Gc-CN afforded through fermentation. The first reason is that the cells were fed with a relatively low concentration (12.5 mM) of GlcNGc, providing only a small amount of *N*-glycolyl for incorporation into CN. Future feeding experiments will examine the use of higher concentrations of GlcNGc or per-acetylation to increase conversion of CN to Gc-CN. The second reason is that there may be competition for resources between the production of UDP-GalNGc from GlcNGc (Fig. 4) with the production of UDP-GalNAc from glucose (He et al. 2015; Cimini et al. 2015). *E. coli* may be diverting most of its resources to the familiar and preferred pathway from glucose to UDP-GalNAc. We hypothesize that it may be possible to knock out the bifunctional *glmU* gene to prevent this competition. This gene encodes a bi-functional enzyme responsible for both the uridyltransferase and acetyltransferase activity (Mengin-Lecreulx and van Heijenoort 1994). The acetyltransferase catalyzes the acetylation of glucosamine-1-phosphate (GlcN-1-P), forming *N*-acetylglucosamine-1-phosphate (GlcNAc-1-P) (Mengin-Lecreulx and van Heijenoort 1993) and, the uridyltransferase catalyzes the formation of UDP-*N*-acetylglucosamine (UDP-GlcNAc) (Mengin-Lecreulx and van Heijenoort 1994). In the production of Gc-CN, the acetyltransferase domain is not required. Therefore, it may be possible to mutate the *glmU* gene to inactivate the acetyltransferase while maintaining uridyltransferase activity. Such a mutation should stop the cells from making GalNAc, so that all of the cell's resources can be diverted to GalNGc synthesis. Modulating this gene could also be useful in generating Gc-CN standards. Such modulation has been reported (Pompeo et al. 1998).

In conclusion, we demonstrate the first production of Gc-CN in an engineered strain of *E. coli*. While the levels of Gc-CN residues are relatively low, they are similar to the Gc-CN content expected on human consumption of red meat. Enhancement of incorporation of GalNGc into chondroitin might be achieved through optimized feeding of GlcNGc, by knocking out enzymes in the biosynthetic pathway or through the protein engineering of bifunctional biosynthetic enzymes. Future studies will examine all three of these approaches.

Declarations

Ethics approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication

All authors read and approved the final manuscript and give their consent for publication.

Availability of data and materials

All data and materials are available upon request.

Competing interests

The authors declare that have no conflict of interest.

Funding

This work was supported by a US National Science Foundation award CBET-1604547 awarded to MK and by National Institutes of Health grant CA231074 and GM32373.

Authors' contributions

AEA, AB, RJL and MK wrote the manuscript. RJL and MK provided the funding for this study. AV and RJL designed the study. AEA and WH performed the metabolic engineering and fermentation. SB performed the synthesis. KX performed the MS analysis.

Acknowledgements

The authors acknowledge the Center for Biotechnology and Interdisciplinary Studies for access to their core facilities.

References

1. Angata T, Varki A (2002) Chemical Diversity in the Sialic Acids and Related α -Keto Acids: An Evolutionary Perspective. *Chem Rev* 102:439–470 . <https://doi.org/10.1021/cr000407m>
2. Arita K, Ikuta K, Nishi Y, Kato S, Yamauchi E, Maki S, Naiki M (1982) Heterophile Hanganutziu-Deicher antibodies in sera of patients with Kawasaki diseases. *Biken J* 25:157–62
3. Bergfeld AK, Lawrence R, Diaz SL, Pearce OMT, Ghaderi D, Gagneux P, Leakey MG, Varki A (2017) N-glycolyl groups of nonhuman chondroitin sulfates survive in ancient fossils. *Proc Natl Acad Sci* 114:E8155–E8164 . <https://doi.org/10.1073/pnas.1706306114>
4. Bergfeld AK, Pearce OMT, Diaz SL, Pham T, Varki A (2012) Metabolism of vertebrate amino sugars with N-glycolyl groups: elucidating the intracellular fate of the non-human sialic acid N-glycolylneuraminic acid. *J Biol Chem* 287:28865–28881 . <https://doi.org/10.1074/jbc.M112.363549>
5. Calatroni A, Vinci R, Ferlazzo AM (1992) Characteristics of the interactions between acid glycosaminoglycans and proteins in normal human plasma as revealed by the behaviour of the protein-polysaccharide complexes in ultrafiltration and Chromatographie procedures. *Clin Chim Acta* 206:167–180 . [https://doi.org/10.1016/0009-8981\(92\)90086-6](https://doi.org/10.1016/0009-8981(92)90086-6)
6. Carr A, Mullet A, Mazorra Z, Vázquez AM, Alfonso M, Mesa C, Rengifo E, Pérez R, Fernández LE (2000) A Mouse IgG 1 Monoclonal Antibody Specific for N-Glycolyl GM3 Ganglioside Recognized Breast and Melanoma Tumors. *Hybridoma* 19:241–247 . <https://doi.org/10.1089/02724570050109639>

7. Chi L, Wolff JJ, Laremore TN, Restaino OF, Xie J, Schiraldi C, Toida T, Amster IJ, Linhardt RJ (2008) Structural Analysis of Bikunin Glycosaminoglycan. *J Am Chem Soc* 130:2617–2625 .
<https://doi.org/10.1021/ja0778500>
8. Chou H-H, Hayakawa T, Diaz S, Krings M, Indriati E, Leakey M, Paabo S, Satta Y, Takahata N, Varki A (2002) Inactivation of CMP-N-acetylneuraminic acid hydroxylase occurred prior to brain expansion during human evolution. *Proc Natl Acad Sci* 99:11736–11741 .
<https://doi.org/10.1073/pnas.182257399>
9. Chou H-H, Takematsu H, Diaz S, Iber J, Nickerson E, Wright KL, Muchmore EA, Nelson DL, Warren ST, Varki A (1998) A mutation in human CMP-sialic acid hydroxylase occurred after the Homo-Pan divergence. *Proc Natl Acad Sci* 95:11751–11756 . <https://doi.org/10.1073/pnas.95.20.11751>
10. Cimini D, Carlino E, Giovane A, Argenzio O, Dello Iacono I, De Rosa M, Schiraldi C (2015) Engineering a branch of the UDP-precursor biosynthesis pathway enhances the production of capsular polysaccharide in *Escherichia coli* O5:K4:H4. *Biotechnol J* 10:1307–1315 .
<https://doi.org/10.1002/biot.201400602>
11. Cimini D, Restaino OF, Schiraldi C (2018) Microbial production and metabolic engineering of chondroitin and chondroitin sulfate. *Emerg Top Life Sci* 2:349–361 .
<https://doi.org/10.1042/ETLS20180006>
12. Cirino PC, Chin JW, Ingram LO (2006) Engineering *Escherichia coli* for xylitol production from glucose-xylose mixtures. *Biotechnol Bioeng* 95:1167–1176 . <https://doi.org/10.1002/bit.21082>
13. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci* 97:6640–6645 .
<https://doi.org/10.1073/pnas.120163297>
14. Dhar C, Sasmal A, Varki A (2019) From “Serum Sickness” to “Xenosialitis”: Past, Present, and Future Significance of the Non-human Sialic Acid Neu5Gc . *Front. Immunol.* 10:807
15. Fernández-Ruiz I (2019) Enzyme loss during evolution linked to atherosclerosis predisposition. *Nat Rev Cardiol* 16:580–580 . <https://doi.org/10.1038/s41569-019-0251-9>
16. He W (2017) Metabolic engineering and applied enzymology for the preparation of nutraceutical/ pharmaceutical chondroitin sulfate. Rensselaer Polytechnic Institute
17. He W, Fu L, Li G, Andrew Jones J, Linhardt RJ, Koffas M (2015) Production of chondroitin in metabolically engineered *E. coli*. *Metab Eng* 27:92–100 .
<https://doi.org/10.1016/j.ymben.2014.11.003>
18. Higashi H, Naiki M, Matuo S, Ōkouchi K (1977) Antigen of “serum sickness” type of heterophile antibodies in human sera: Identification as gangliosides with N-glycolylneuraminic acid. *Biochem Biophys Res Commun* 79:388–395 . [https://doi.org/https://doi.org/10.1016/0006-291X\(77\)90169-3](https://doi.org/https://doi.org/10.1016/0006-291X(77)90169-3)
19. Jin P, Zhang L, Yuan P, Kang Z, Du G, Chen J (2016) Efficient biosynthesis of polysaccharides chondroitin and heparosan by metabolically engineered *Bacillus subtilis*. *Carbohydr Polym.*
<https://doi.org/10.1016/j.carbpol.2015.12.065>

20. Jolly L, Ferrari P, Blanot D, van Heijenoort J, Fassy F, Mengin-Lecreulx D (1999) Reaction mechanism of phosphoglucosamine mutase from *Escherichia coli*. *Eur J Biochem* 262:202–210 .
<https://doi.org/10.1046/j.1432-1327.1999.00373.x>
21. Kawanishi K, Dhar C, Do R, Varki N, Gordts PLSM, Varki A (2019) Human species-specific loss of CMP- N -acetylneuraminic acid hydroxylase enhances atherosclerosis via intrinsic and extrinsic mechanisms. *Proc Natl Acad Sci* 116:16036– 16045 . <https://doi.org/10.1073/pnas.1902902116>
22. Kozutsumi Y, Kawano T, Yamakawa T, Suzuki A (1990) Participation of Cytochrome b5 in CMP-N-Acetylneuraminic Acid Hydroxylation in Mouse Liver Cytosol1. *J Biochem* 108:704–706 .
<https://doi.org/10.1093/oxfordjournals.jbchem.a123268>
23. Lu H, McDowell LM, Studelska DR, Zhang L (2010) Glycosaminoglycans in Human and Bovine Serum: Detection of Twenty-Four Heparan Sulfate and Chondroitin Sulfate Motifs Including a Novel Sialic Acid-modified Chondroitin Sulfate Linkage Hexasaccharide. *Glycobiol Insights* 2010:13–28
24. Macauley MS, Chan J, Zandberg WF, He Y, Whitworth GE, Stubbs KA, Yuzwa SA, Bennet AJ, Varki A, Davies GJ, Vocadlo DJ (2012) Metabolism of Vertebrate Amino Sugars with N -Glycolyl Groups. *J Biol Chem* 287:28882–28897 . <https://doi.org/10.1074/jbc.M112.363721>
25. Marquina G, Waki H, Fernandez LE, Kon K, Carr A, Valiente O, Perez R, Ando S (1996) Gangliosides expressed in human breast cancer. *Cancer Res* 56:5165–71
26. Mengin-Lecreulx D, van Heijenoort J (1994) Copurification of glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridyltransferase activities of *Escherichia coli*: characterization of the glmU gene product as a bifunctional enzyme catalyzing two subsequent steps in the pat. *J Bacteriol* 176:5788–5795 . <https://doi.org/10.1128/JB.176.18.5788-5795.1994>
27. Mengin-Lecreulx D, van Heijenoort J (1993) Identification of the glmU gene encoding N-acetylglucosamine-1-phosphate uridyltransferase in *Escherichia coli*. *J Bacteriol* 175:6150–6157 .
<https://doi.org/10.1128/JB.175.19.6150-6157.1993>
28. Neidhardt FC, Bloch PL, Smith DF (1974) Culture Medium for Enterobacteria. *J Bacteriol* 119:736 LP – 747
29. Padlan EA, Kabat EA (1988) Model-building study of the combining sites of two antibodies to alpha (1—6)dextran. *Proc Natl Acad Sci* 85:6885–6889 . <https://doi.org/10.1073/pnas.85.18.6885>
30. Padler-Karavani V, Deng L, Garcia A, Karp F, Naidu N, Choudhury B, Varki N, Varki A (2012) Expression of the tumor-associated antigen Neu5Gc-Sialyl-Tn in human carcinomas (74.6). *J Immunol* 188:74.6 LP-74.6
31. Padler-Karavani V, Tremoulet AH, Yu H, Chen X, Burns JC, Varki A (2013) A simple method for assessment of human anti-Neu5Gc antibodies applied to Kawasaki disease. *PLoS One* 8:e58443–e58443 . <https://doi.org/10.1371/journal.pone.0058443>
32. Pompeo F, van Heijenoort J, Mengin-Lecreulx D (1998) Probing the role of cysteine residues in glucosamine-1-phosphate acetyltransferase activity of the bifunctional GImU protein from *Escherichia coli*: site-directed mutagenesis and characterization of the mutant enzymes. *J Bacteriol* 180:4799–4803

33. Restaino OF, D'ambrasio S, Cassese E, Barbuto S, Ferrainuolo S, Alfano A, Ventriglia R, Marrazzo A, Schiraldi C, Cimini D, (2019) Molecular weight determination of heparosan-and chondroitin-like capsular polysaccharides: figuring out differences between wild-type and engineered *E. coli* strains. *Appl. Microbiol. Biotechnol.* 103:6771-6782.
34. Rodriguez M-L, Jann B, Jann K (1988) Structure and serological characteristics of the capsular K4 antigen of *Escherichia coli* O5:K4:H4, a fructose-containing polysaccharide with a chondroitin backbone. *Eur J Biochem* 177:117–124 . <https://doi.org/10.1111/j.1432-1033.1988.tb14351.x>
35. Samraj A, Läubli H, Varki N, Varki A (2014) Involvement of a Non-Human Sialic Acid in Human Cancer . *Front. Oncol.* 4:33
36. Samraj AN, Pearce OMT, Läubli H, Crittenden AN, Bergfeld AK, Banda K, Gregg CJ, Bingman AE, Secrest P, Diaz SL, Varki NM, Varki A (2015) A red meat-derived glycan promotes inflammation and cancer progression. *Proc Natl Acad Sci* 112:542–547 . <https://doi.org/10.1073/pnas.1417508112>
37. Schauer R (2000) Achievements and challenges of sialic acid research. *Glycoconj J* 17:485–499 . <https://doi.org/10.1023/a:1011062223612>
38. Sinaÿ P (1971) Synthé du 3-O-(-1-carboxyéthyl)-2-désoxy-2-glycolamido- -glucose (acide N-glycolylmuramique). *Carbohydr Res* 16:113–122 . [https://doi.org/10.1016/S0008-6215\(00\)86104-2](https://doi.org/10.1016/S0008-6215(00)86104-2)
39. Soulillou J-P, Cozzi E, Bach J-M (2020) Challenging the Role of Diet-Induced Anti-Neu5Gc Antibodies in Human Pathologies. *Front Immunol* 11:834 . <https://doi.org/10.3389/fimmu.2020.00834>
40. Sugiura N, Tawada A, Sugimoto K, Watanabe H (2002) Molecular cloning and characterization of chondroitin polymerase from *Escherichia coli* strain K4. *J Biol Chem.* <https://doi.org/10.1074/jbc.M201719200>
41. Sun X, Li L, Overdier KH, Ammons LA, Douglas IS, Burlew CC, Zhang F, Schmidt EP, Chi L, Linhardt RJ (2015) Analysis of Total Human Urinary Glycosaminoglycan Disaccharides by Liquid Chromatography–Tandem Mass Spectrometry. *Anal Chem* 87:6220–6227 . <https://doi.org/10.1021/acs.analchem.5b00913>
42. Tangvoranuntakul P, Gagneux P, Diaz S, Bardor M, Varki N, Varki A, Muchmore E (2003) Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc Natl Acad Sci* 100:12045–12050 . <https://doi.org/10.1073/pnas.2131556100>
43. Taylor RE, Gregg CJ, Padler-Karavani V, Ghaderi D, Yu H, Huang S, Sorensen RU, Chen X, Inostroza J, Nizet V, Varki A (2010) Novel mechanism for the generation of human xeno-autoantibodies against the nonhuman sialic acid N-glycolylneuraminic acid. *J Exp Med* 207:1637–1646 . <https://doi.org/10.1084/jem.20100575>
44. Uehara T, Park JT (2004) The N-Acetyl-d-Glucosamine Kinase of *Escherichia coli* and Its Role in Murein Recycling. *J Bacteriol* 186:7273–7279 . <https://doi.org/10.1128/JB.186.21.7273-7279.2004>
45. Varki A (2008) Sialic acids in human health and disease. *Trends Mol Med* 14:351–360 . <https://doi.org/10.1016/j.molmed.2008.06.002>
46. Wang B, Brand-Miller J (2003) The role and potential of sialic acid in human nutrition. *Eur J Clin Nutr* 57:1351–1369 . <https://doi.org/10.1038/sj.ejcn.1601704>

47. Yehuda S, Padler-Karavani V (2020) Glycosylated Biotherapeutics: Immunological Effects of N-Glycolylneuraminic Acid. *Front Immunol* 11:21 . <https://doi.org/10.3389/fimmu.2020.00021>
48. Yu Y, Zhang F, Colón W, Linhardt RJ, Xia K (2019) Glycosaminoglycans in human cerebrospinal fluid determined by LC-MS/MS MRM. *Anal Biochem* 567:82–84 . <https://doi.org/10.1016/j.ab.2018.12.013>
49. Zanfardino A, Restaino OF, Notomista E, Cimini D, Schiraldi C, De Rosa M, De Felice M, Varcamonti M (2010) Isolation of an *Escherichia coli* K4 kfoC mutant over-producing capsular chondroitin. *Microb Cell Fact* 9:34 . <https://doi.org/10.1186/1475-2859-9-34>
50. Zhu H-M, Sun B, Li Y-J, Meng D-H, Zheng S, Wang T-T, Wang F-S, Sheng J-Z (2018) KfoA, the UDP-glucose-4-epimerase of *Escherichia coli* strain O5:K4:H4, shows preference for acetylated substrates. *Appl Microbiol Biotechnol* 102:751–761 . <https://doi.org/10.1007/s00253-017-8639-0>

Figures

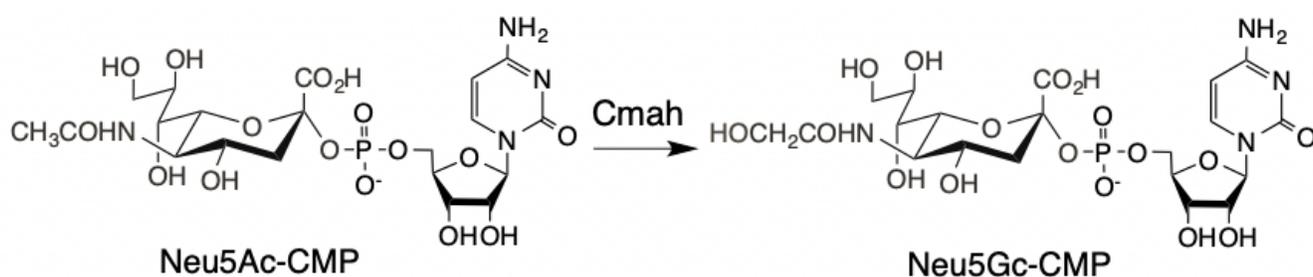


Figure 1

The Cmah enzyme catalyzes the conversion of CMP-Neu5Ac to CMP-Neu5Gc through the addition of an oxygen atom.

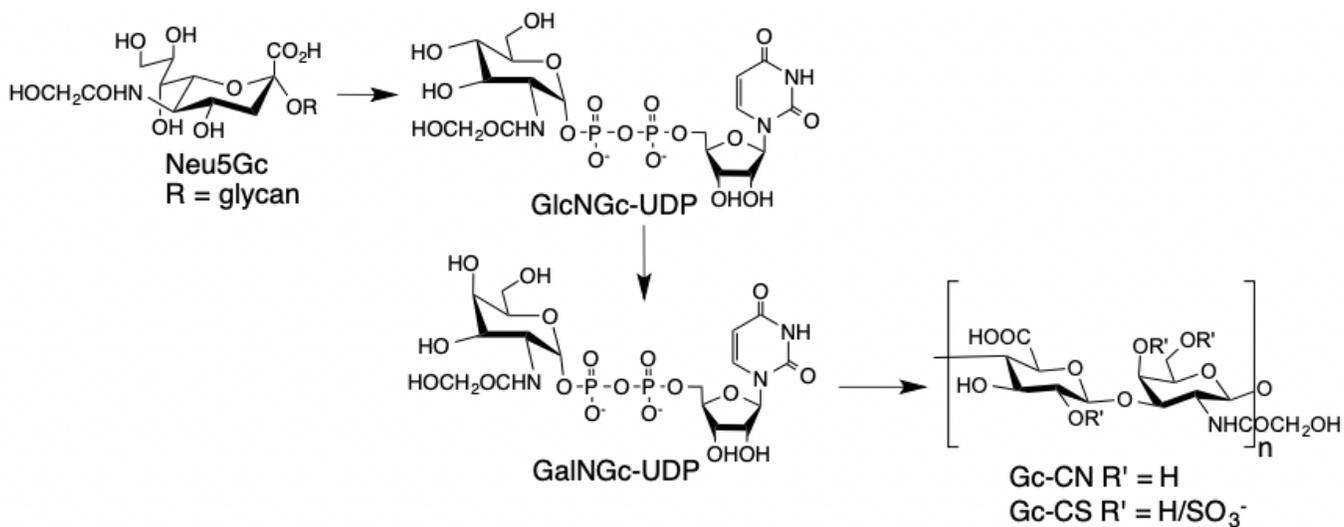


Figure 2

Neu5Gc is converted to UDP-GlcNGc and ultimately UDP-GalNGc. UDP-GalNGc can then be incorporated into Gc-CN and Gc-CS.

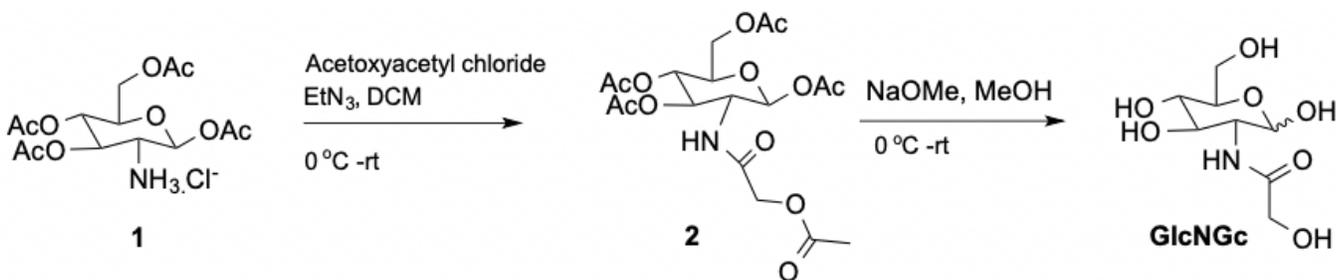


Figure 3

Synthetic route for GlcNGc synthesis.

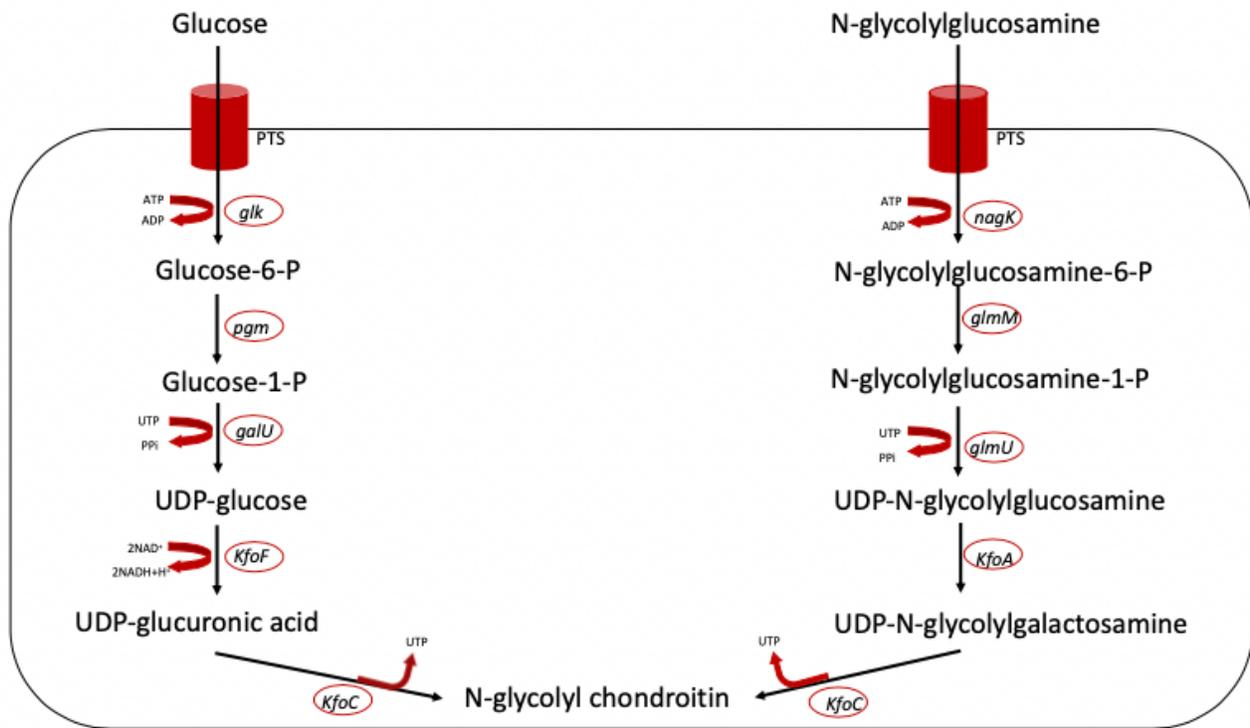


Figure 4

Metabolic pathway for N-glycolyl chondroitin synthesis in *E. coli* K4.

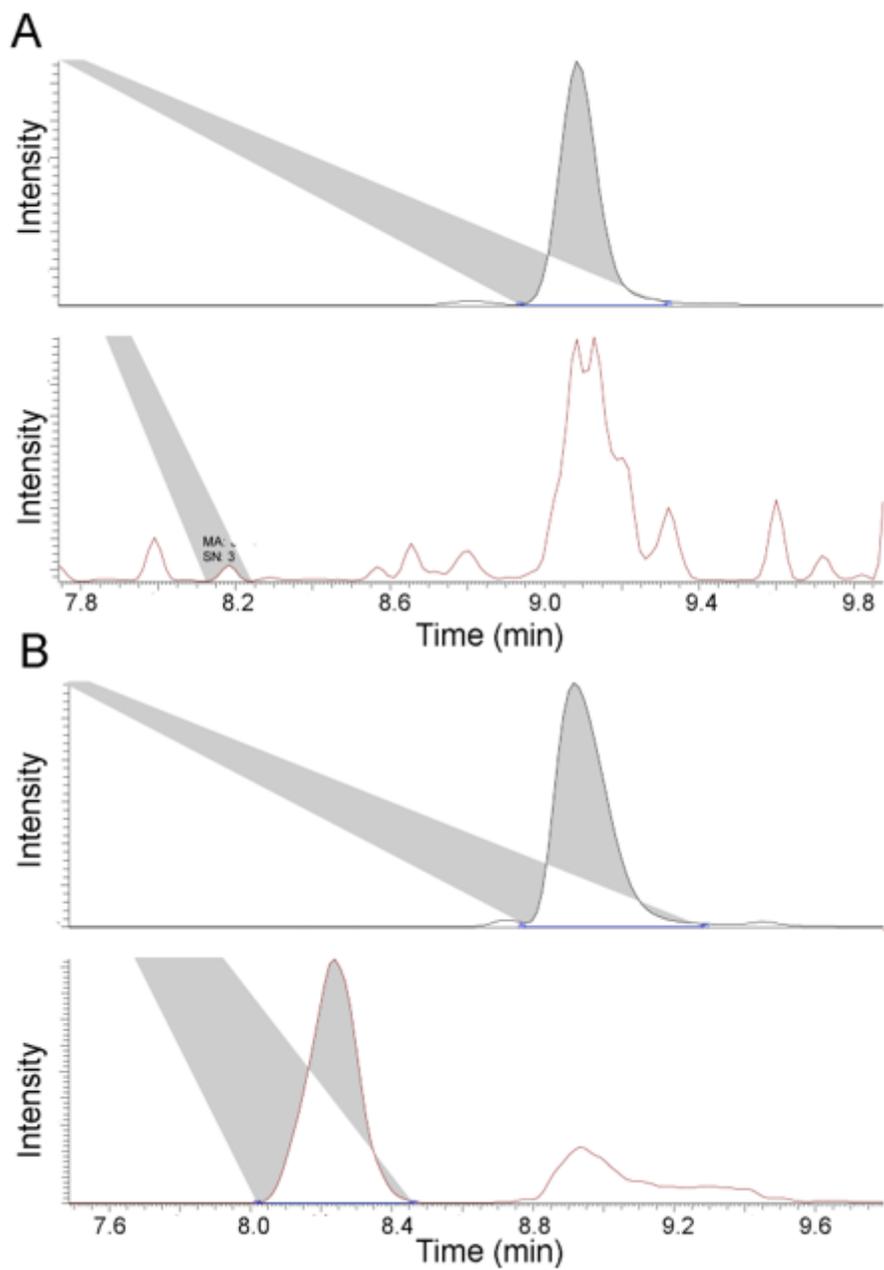


Figure 5

From our chondroitin producing construct, serotype O5:K4:H4, only chondroitin was produced as signified by the absence of AMAC-labeled Gc-CN disaccharide (Fig. 5a). From the GlcNGc fed culture, both AMAC-labeled CN and Gc-CN disaccharides were observed at the expected m/z ratio (Fig. 5b).