





Complete biosynthesis of a sulfated chondroitin in *Escherichia coli*

Abinaya Badri¹, Asher Williams¹, Adeola Awofiranye², Payel Datta ², Ke Xia¹, Wenqin He¹, Keith Fraser², Jonathan S. Dordick ^{1,2}, Robert J. Linhardt ^{1,2,3} & Mattheos A. G. Koffas ^{1,2}✉

Sulfated glycosaminoglycans (GAGs) are a class of important biologics that are currently manufactured by extraction from animal tissues. Although such methods are unsustainable and prone to contamination, animal-free production methods have not emerged as competitive alternatives due to complexities in scale-up, requirement for multiple stages and cost of co-factors and purification. Here, we demonstrate the development of single microbial cell factories capable of complete, one-step biosynthesis of chondroitin sulfate (CS), a type of GAG. We engineer *E. coli* to produce all three required components for CS production—chondroitin, sulfate donor and sulfotransferase. In this way, we achieve intracellular CS production of ~27 µg/g dry-cell-weight with about 96% of the disaccharides sulfated. We further explore four different factors that can affect the sulfation levels of this microbial product. Overall, this is a demonstration of simple, one-step microbial production of a sulfated GAG and marks an important step in the animal-free production of these molecules.

¹Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA. ²Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY, USA. ³Department of Chemistry and Chemical Biology, Rensselaer Polytechnic Institute, Troy, NY, USA. ✉email: koffam@rpi.edu

Western medicine relies heavily on sulfated glycosaminoglycans (GAGs)¹. The two most commonly used GAGs are heparin and chondroitin sulfate (CS). Since the 1940s, heparin has predominated as the primary anticoagulant used in medicine². CS, the most abundant GAG in the human body, is an important component of joint tissues and is frequently used in the treatment of osteoarthritis^{3,4}. Since they are exclusively present in animals, all GAGs are currently commercially manufactured by extraction from animal tissues, primarily from bovine trachea and porcine intestinal mucosa. Prime producers of pig and cattle, such as China, dominate the manufacturing and marketing of GAGs. These sulfated polysaccharides always occur as mixtures in tissues with individual components varying slightly in stereochemistry, length, and sulfation pattern⁵. Such small analytical differences are both a blessing and a curse; they result in remarkably distinct biological function and in vivo behavior; however, they also make their adulteration very hard to detect. Contamination incidents like the heparin adulteration crisis of 2008^{6–8} and FDA's warning about questionable crude GAG sources in 2017⁹ have evoked a major conversation about the deficiencies of current production methods, regulatory practices, and analytical detection methods of adulterants/contaminants in GAGs.

GAGs also have complicated structures necessitating sophisticated analytical instrumentation for verifying their purity¹⁰. GAG activity and specificity are dependent upon their functional group pattern. Specific interactions of GAGs with important biomolecules bring about their physiological roles like anticancer¹¹ and anti-diabetic properties¹². The potentials of such properties have created additional demands for the sustainable availability of pure, chemically defined GAGs. Difficulties in downstream purification, complex and expensive quality control steps, risk of cross-viral contaminations, non-sustainability, and heterogeneity in GAGs from animal tissues and cultural trends against animal-sourced products are key forces driving innovation in GAG manufacturing towards sustainable, microbial-based processes.

Chondroitin sulfate (CS), the first GAG to have been extracted and identified¹³, is extensively prescribed in human and veterinary joint health⁴. CS is composed of [\rightarrow 4]- β -D-GlcA-(1 \rightarrow 3)- β -D-GalNAc-(1 \rightarrow) repeating disaccharide units with various combinations of sulfation and epimerization generating different CS types. CS is extracted primarily from cartilaginous tissues,

with bovine trachea and shark cartilage serving as major raw materials for CS manufacturing^{14,15}. Complex CS structures in proteoglycans have myriad functional group patterns that allow specific interactions with biomolecules¹⁶. Such interactions regulate many important cellular processes, including differentiation and development,^{17–19} and determine the role of CS in health and disease. For example, specific patterns of fucosylated CS from sea cucumbers were shown to possess anti-obesity, anti-diabetic and immunomodulatory activities²⁰. Systematic studies identifying such favorable, physiological activities of specific CS types are severely limited by the availability of pure CS.

A major solution to the limitations associated with the current state-of-art CS production is production using animal-free methods^{21–23}. Chemical, chemoenzymatic, microbial, and mammalian cell-culture based CS production systems are being developed worldwide²¹. Chemical synthesis methods are tedious, involving multiple steps, and are difficult to scale up. Similarly, synthesis from mammalian cell cultures is not ideal due to handling complexities, high media cost, low cell densities, and interference from other GAG pathways.

Complete biosynthesis of CS in *E. coli* requires three factors; unsulfated chondroitin precursor, the sulfate donor 3'-phosphoadenosine-5'-phosphosulfate (PAPS), and chondroitin sulfotransferase. Chondroitin and PAPS are the substrates for CS and each is produced from distinct metabolic pathways. Overall, a total of 13 metabolic reactions are required for chondroitin and PAPS synthesis. Of the 11 metabolic steps to synthesize chondroitin, nine are already present in commonly used *E. coli* strains (Fig. 1a). Enzymes for the two remaining steps (UDP-N-acetylglucosamine-/UDP-glucosamine-4-epimerase and chondroitin synthase) have been identified in *E. coli* K4 (Serovar O5:K4:H4)²⁴. Incidentally, other bacteria like *Pseudomonas aeruginosa* serotype O6²⁵, *Yersinia enterocolitica* serotype O8²⁶, *Pasteurella multocida* Type F²⁷, and *E. coli* O86:B7²⁸ possessing either or both of these chondroitin-specific enzymes have also been characterized.

Complete microbial synthesis of CS holds great promise as it represents a one-step, sustainable process for the production of structurally homogeneous CS. However, this has not been practically accomplished until now. Since the enzymatic route for CS production is well known, multi-step chemoenzymatic methods have served as a placeholder for biotechnological production. *Escherichia coli* is a model organism heavily used in the

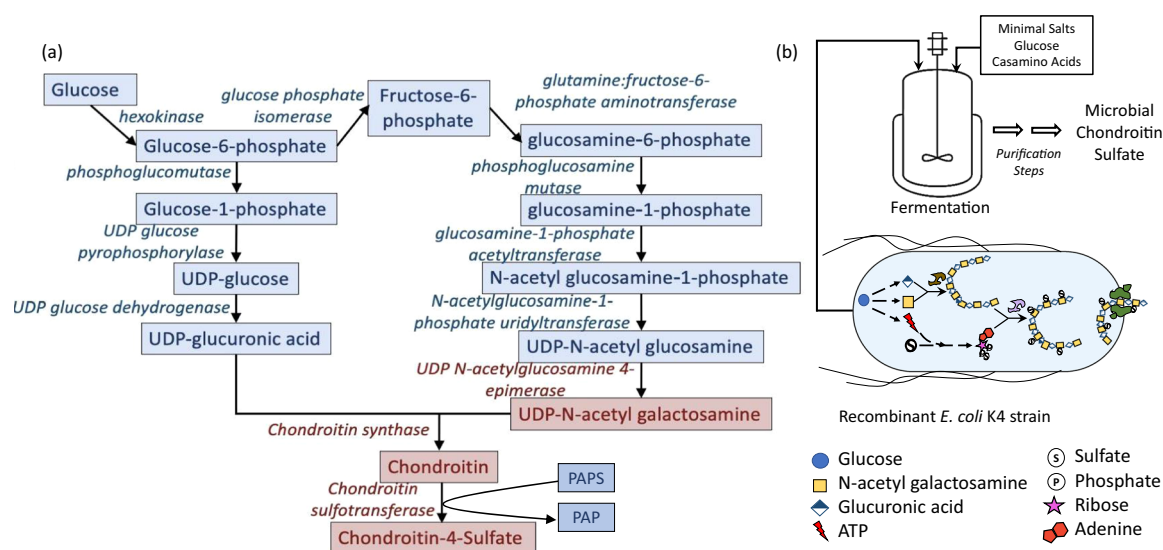


Fig. 1 In vivo production of different CS types. **a** Biosynthetic pathway for chondroitin production from glucose. The parts marked in red are absent in common lab strains of *E. coli*. **b** Schematic of microbial CS production from minimal nutrient media using recombinant *E. coli*.

production of recombinant products and non-native metabolites due to its genetic tractability and ease of scale-up. In this study, we describe the development of a metabolically engineered *E. coli* strain for the complete biosynthesis of animal-free CS (Fig. 1b). We accomplish this by identifying the factors that control CS sulfation levels and engineering efficient chondroitin sulfotransferases. This study represents a one-step, in vivo, microbial synthesis of animal-free sulfated glycosaminoglycans.

Results

Chondroitin sulfation is catalyzed by chondroitin sulfotransferases. These enzymes transferases of animal origin that undergo posttranslational glycosylation and are present in Golgi apparatus with no identified microbial counterparts. This requires that they be expressed as heterologous proteins. While different chondroitin sulfotransferases give rise to different forms of CS, in the current study, we focus on one homolog of this class, chondroitin-4-*O*-sulfotransferase (*S_w*) that results in the production of the 4-*O*-sulfated CS-A.

E. coli K4 was selected to assemble the required elements for CS biosynthesis. This strain has been extensively studied for chondroitin production^{24,29–31}. However, K4 produces a fructosylated chondroitin as part of its capsular polysaccharide. The fructosyltransferase, responsible for fructosylation of chondroitin’s GlcA residues at the 3-position, is unfavorable for CS production. For this reason, we first deleted the fructosyltransferase-encoding gene, thereby resulting in the generation of non-fructosylated chondroitin in *E. coli* K4^{32,33}. We also integrated the T7 RNA polymerase into the genomic *lacZ* site in *E. coli* K4 $\Delta kfoE$, generating K4 $\Delta kfoE$ (DE3), to facilitate the expression of recombinant genes under the stronger T7 promoter system (Supplementary Figure 1).

PAPS accumulation facilitates intracellular CS synthesis. In our first attempt to enable CS synthesis, we simply added the missing

sulfotransferase component to K4 $\Delta kfoE$ (DE3) by expressing *S_w* via the pETM6-*S_w* plasmid. However, no CS synthesis was observed (Fig. 2a). Sulfation of chondroitin to CS requires PAPS, which is the universal sulfate donor responsible for most biological sulfation processes. PAPS biosynthesis is a subset of the ubiquitous cysteine/methionine biosynthetic pathways, and therefore, is present in almost all cell types, including *E. coli*. Due to the presence of high levels of chondroitin and previously demonstrated successful *S_w* expression, we hypothesized that the lack of CS production was due to limitations in intracellular PAPS⁵⁰. PAPS biosynthesis involves a two-step activation of inorganic sulfate onto ATP catalyzed by ATP sulfurylase (*cysDN*) and APS kinase (*cysC*) (Fig. 2b). However, the native pathway also consists of PAPS reductase (*cysH*) that competes with sulfotransferases and reduces PAPS to inorganic sulfite (Fig. 2b).

To address the elimination of such a competitive enzyme expression, we adopted a *cysH* repression or deletion strategy in the K4 $\Delta kfoE$ (DE3)pETM6-*S_w* strain and evaluated PAPS accumulation and CS sulfation to explore whether PAPS was indeed the limiting component in this strain (Fig. 2a). PAPS accumulation achieved in this strain through the manipulation of *cysH* enabled intracellular sulfation of chondroitin, demonstrating that it is indeed possible to completely biosynthesize CS in *E. coli*. This result shows that PAPS pathway intervention is a necessary condition for GAG biosynthesis in *E. coli*.

As a result of *kfoE* and *cysH* deletions and *S_w* overexpression, CS sulfation yield reached ~19%, which is lower than that in animal-extracted CS-A. Thus, there is room for improvement. Overexpression of other PAPS biosynthetic genes (*cysDNCQ*) in the K4 $\Delta kfoE\Delta cysH$ (DE3) significantly reduced the sulfation of chondroitin (Fig. 2c). Lack of success in this strategy also suggests that using a strain that responds to *cysH* deletion with higher PAPS accumulation might lead to higher CS sulfation. Since our previously developed *E. coli* MG1655 $\Delta cysH$ (DE3)³⁴ accumulated about 54-fold higher (~0.8 μ moles/gDCW) PAPS than K4 $\Delta kfoE\Delta cysH$ (DE3), we

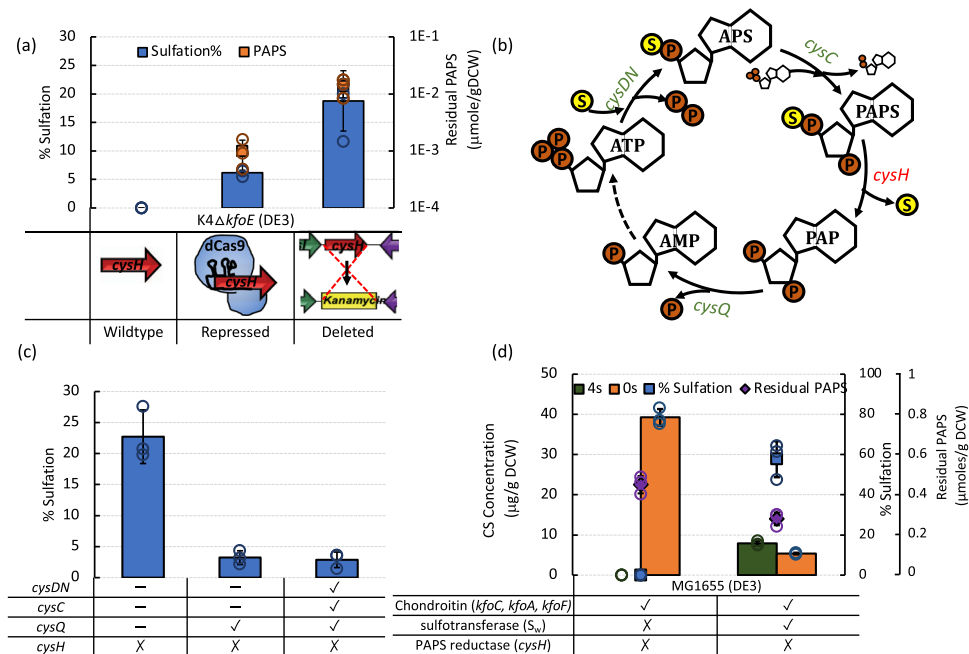


Fig. 2 Effect of PAPS accumulation on CS sulfation. **a** CS sulfation and residual PAPS in K4 $\Delta kfoE$ (DE3)pETM6-*S_w* strains with wildtype, repressed and deleted *cysH*. **b** Biosynthetic route for PAPS synthesis. **c** Effect of overexpression of PAPS biosynthetic enzymes on CS sulfation in K4 $\Delta kfoE$ (DE3)pETM6-*S_w*. **d** CS titer, residual PAPS and sulfation in MG1655 $\Delta cysH$ (DE3) expressing chondroitin synthesis and sulfotransferase enzymes. (Xs = CS disaccharide sulfated in the X position of *N*-acetylglucosamine). Error bars represent standard deviation from three biological replicates. Source data underlying Figs. 2a, c, and d are provided as a Source Data file.

also explored CS synthesis in this strain. Chondroitin production in MG1655 Δ *cysH*(DE3) was enabled by expression of the K4 genes, *kfoC*, *kfoA*, and *kfoF* through the plasmid pETM6-PCAF, previously constructed in our lab³¹. Furthermore, the assembly of all three components in MG1655 Δ *cysH*(DE3) by co-expression of *S_w* (pETM6-PCAF_{*S_w*}) led to a high intracellular CS sulfation level of 58% (Fig. 2d). The metabolic cost of *S_w* expression was also observed in the form of a decrease in total CS.

Improving sulfotransferase activity enhances CS sulfation.

Chondroitin sulfotransferases are Golgi-based, integral transmembrane glycoproteins. These enzymes must be truncated, correctly folded and expressed in a soluble form in order to be effective catalysts for *E. coli*-based CS synthesis³⁵. Based on the hypothesis that increased solubility and stability of *S_w* would improve CS sulfation levels, we relied on a computational approach to identify mutations enhancing these properties. A web-based server called Protein Repair One Stop Shop (PROSS) utilizes sequence and structure information of proteins to identify residue mutations that would result in enhanced expression and solubility³⁶. While the sequence of human *S_w* was known based on our previous work on its expression in *E. coli* and *P. pastoris*, a structure is not yet available in the protein databank. Hence, we constructed a homology model of *S_w* structure to use in PROSS that is based on the previously elucidated structure of the sulfotransferase domain of an olefin synthase from *Synechococcus* PCC 7002 (RMSD of 0.74). The sequence alignment of *S_w* with the template is shown in Supplementary Figure 2a. From our analysis using the PROSS server, we were able to identify three variants of *S_w*, each containing multiple mutations across the sequence (Sequences in Supplementary Table 4). The mutants are named *S_{M1}* (H127E S238Y), *S_{M2}* (K117R H127E S238Y A245G), and *S_{M4}* (I7A R30Q Q106E K117R S118N H127E I146T S226A S237D S238Y A245G E272Q). All of the predicted mutations were at solvent accessible residues, distant from the active site of the enzyme (Fig. 3a). Mutations I7A, R30Q, Q106E, S118N, H127E, I146T, S226A, S237D, S238Y, E272Q resulted in a modified charge distribution at these solvent accessible positions (Sequence Alignments in Supplementary Figure 2a).

The three mutant enzymes (designated as *S_{M1}*, *S_{M2}*, *S_{M4}*) were expressed in *E. coli* BL21(DE3) using a pET32LIC vector and

were then purified. While all three mutants demonstrated enhanced in vitro sulfotransferase activity compared to *S_w* (Fig. 3b and Supplementary Fig. 7), *S_{M2}* demonstrated the highest conversion. Transformation of the mutant plasmids into K4 Δ *kfoE* Δ *cysH*(DE3) (PAPS-accumulator and chondroitin-producer) helped verify the effect of these mutants on in vivo constructs. Expression of *S_{M1}* and *S_{M4}* did not result in significantly different intracellular sulfation levels compared to *S_w*. In contrast, *S_{M2}* produced CS with a three-fold increased sulfation (Fig. 3c). The disparity between the in vitro and the corresponding in vivo activities (*S_w*(in vitro) vs. *S_w*(in vivo); *S_{M1}*(in vitro) vs. *S_{M1}*(in vivo); etc.) could be attributed to differences in the levels of expressed chondroitin-4-*O*-sulfotransferase and/or the chondroitin and PAPS substrates in the respective experiments. Another noteworthy aspect is the difference in the sulfation level of *S_w* while expressed using pET32LIC (Fig. 3c) and pETM6 (Fig. 2a). The intracellular level of sulfation obtained by pET32LIC-*S_w* was 4- to 6-fold lower than that obtained by pETM6-*S_w* depending on expression temperature (Supplementary Figure 3). While both these plasmids have similar backbones, the same origin of replication, and antibiotic resistance, the pET32LIC version expresses the enzyme with N-terminal Trx and His tags. In contrast, pETM6 expresses the enzyme sequence without the tags and is also a plasmid backbone generally used in metabolic engineering studies.

Sulfation level is dependent on sulfotransferase induction.

Growth and induction conditions have been shown to have an effect on biosynthesis of many microbial products^{37–39}. For purposes of GAG production, in vivo sulfation level is a metric of equal, or greater, importance than titer. To explore the effect of induction on in vivo sulfation, K4 Δ *kfoE* Δ *cysH*(DE3) pETM6-*S_w* was grown and induced for *S_w* expression with IPTG at different growth stages (Fig. 4a). Induction at 0.6 OD₆₀₀ and 1.0 OD₆₀₀ was deemed optimal since they both resulted in significantly higher amounts of CS and higher levels of sulfation (Fig. 4b). Hence, both these induction points were further tested for optimal inducer concentration (0.5 mM (low) and 1.0 mM (high) [IPTG]) and optimal expression temperature (37 °C, 20 °C, and 16 °C) (Fig. 4c). Three conditions resulting in relatively high sulfation of chondroitin were identified. These were cultures induced at 1.0 OD₆₀₀

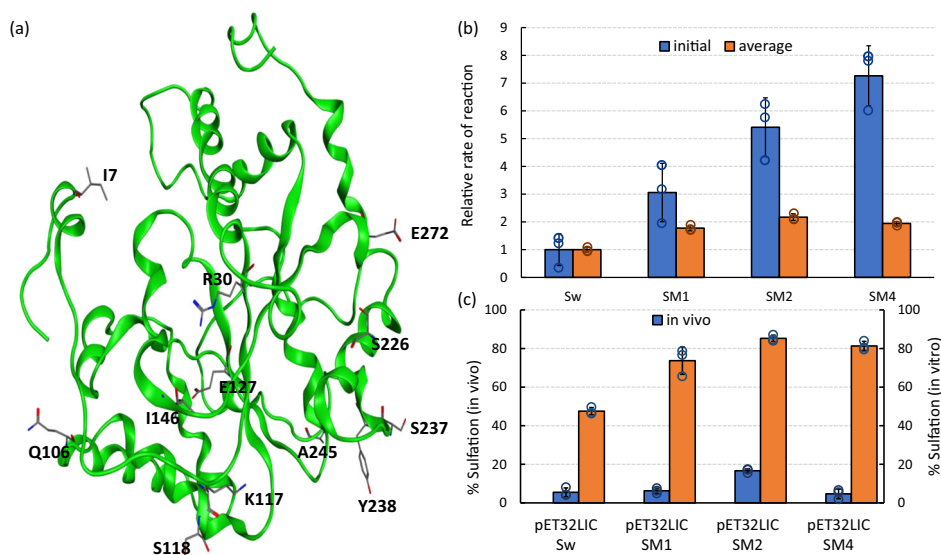


Fig. 3 Effect of sulfotransferase variants on CS sulfation. **a** Locations of predicted mutation displayed on *S_w* structure. **b** Relative initial and average rate of reaction (in vitro) of mutant sulfotransferases compared to *S_w*. **c** In vitro and in vivo (in K4 Δ *kfoE* Δ *cysH*(DE3)) CS sulfation levels of *S_w* and mutants. Error bars represent standard deviation from three biological replicates. Source data underlying Fig. 3b and c are provided as a Source Data file.

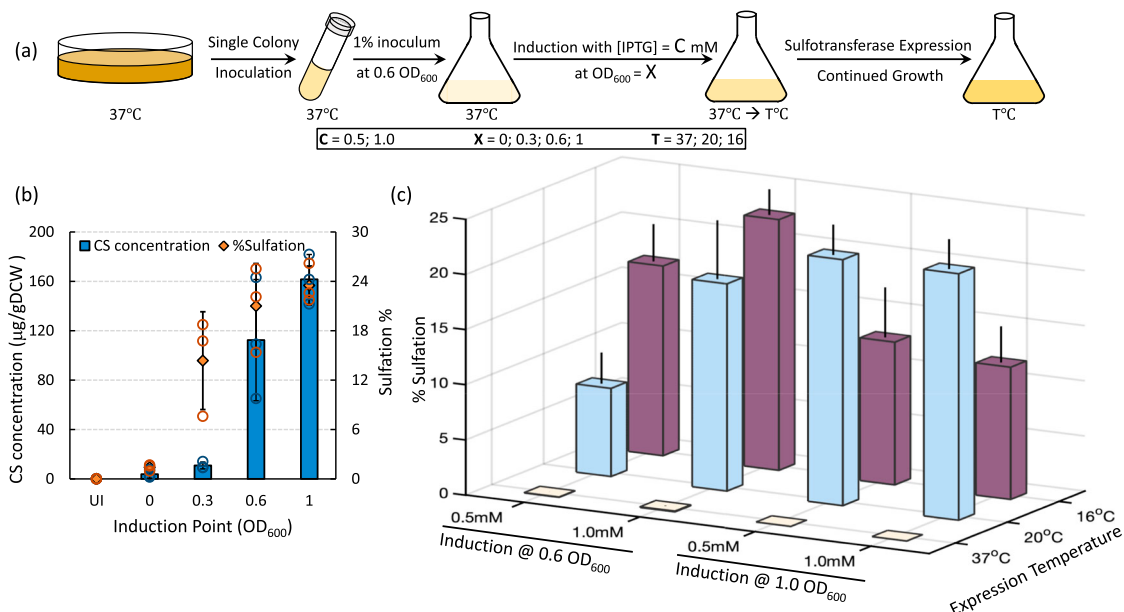


Fig. 4 Effect of fermentation conditions on CS sulfation. **a** Schematic showing growth and cultivation of CS-producing *E. coli* strains. **b** Effect of Induction OD₆₀₀ on CS titer and sulfation in K4Δ*kfoE*Δ*cysH*(DE3)pETM6-S_w. **c** CS sulfation obtained from different inducer concentrations and expression temperatures for 0.6 and 1.0 OD₆₀₀ induction points. Error bars represent standard deviation from three biological replicates. Source data underlying Figs. 4b and c are provided as a Source Data file.

with either low or high IPTG concentration and expressed at 20 °C for 12 h and cultures induced at 0.6 OD₆₀₀ with high IPTG concentration and expressed at 16 °C for 24 h. In contrast to induction at 1.0 OD₆₀₀, induction at 0.6 OD₆₀₀ also made the culture more sensitive to inducer concentration. Overall, employing optimal induction conditions improved the CS sulfation in *E. coli* K4Δ*kfoE*Δ*cysH*(DE3) pETM6-S_w from ~19% to ~23%.

Impact of cellular export mechanism on CS sulfation. *E. coli* K4 naturally exports fructosylated chondroitin through a transporter protein complex. Previous studies have shown that the majority of unsulfated chondroitin produced is exported to the medium⁴⁰. In the current study, we observed that CS, unlike unsulfated chondroitin, was not present in the extracellular medium. Instead, all reported CS production was intracellular. While our experiments point to strategies to improve intracellular sulfation, the bulk of the polysaccharide was both secreted and completely unsulfated (Supplementary Table 1). This led us to examine the possibility that the export of the unsulfated chondroitin is in competition with its sulfation and CS production. We employed CRISPRi-based repression of the transporter genes to verify the effect of decreasing transport activity on intracellular sulfation. The currently accepted mechanism of GAG transport in *E. coli* K4 involves an ABC membrane transporter complex formed by four proteins, KpsT, KpsM, KpsD, and KpsE⁴¹ (Fig. 5a). KpsT is an ATPase that complexes with an inner membrane permease, KpsM. KpsD and KpsE each form dimeric periplasm and membrane-spanning complexes that facilitate the export of the polysaccharide⁴¹.

Five spacers targeting *kpsT* and *kpsM*, designated as dT1, dT2, dT3, dM1 and dM2, were chosen to repress GAG transport (Supplementary Fig. 4a). They were cloned into the CRISPRi plasmid pdCas9, transformed into *E. coli* K4Δ*kfoE*Δ*cysH*(DE3) pETM6-S_w. The strain obtained by transformation with pdCas9-dM₁ was chosen for further investigating the effects of export on sulfation since it demonstrated a 60% decrease in export (Fig. 5b) and Supplementary Fig. 4b). The sulfation of chondroitin in this strain was improved by about 3-fold to ~55%, while no significant

differences were observed in the total GAG titer (Fig. 5c). This suggests that preventing GAG export results in higher levels of sulfation.

High sulfation can be achieved by fermentation optimization.

Upstream fermentation modifications of the constructed CS producer - MG1655Δ*cysH*(DE3) pETM6-PCAFS_w resulted in improvements in CS production. Including a second subculture and fermenting in a bioreactor with controlled pH significantly improved sulfation levels to 96.12 ± 2.55% 4-sulfated CS-A. We also saw minor improvements in titer to about 27 µg/g DCW. Molecular weight (M.W.) analysis (Supplementary Fig. 5) showed that the CS from this MG1655 CS producer is highly polydisperse, with a lower average M.W. Low M.W. CS (1000-3000 Dalton) has potential applications as dietary supplements.

Discussion

Glycosaminoglycans are an important class of pharmaceuticals that are plagued by contamination crises, dependence on extraction from animal tissues and variable quality control^{6–9}. Existing alternatives for animal-sourced GAG production are impeded by the requirement of complex cofactors, poorly active enzymes, poor expression and solubility in bacterial hosts, and difficulties in scale-up²¹. New biological roles of GAGs are still being uncovered and represent a potential paradigm change in their pharmacological applications^{11,12}. Moreover, GAGs represent important co-receptors for pathogens in infectious diseases⁴². For example, the spike protein of SARS-CoV-2 virus tightly binds to the host cell GAG, heparan sulfate, leading to virus interaction with its protein-based ACE2 receptor and resulting in COVID-19⁴³.

Here, we demonstrate the in vivo production of a sulfated GAG in metabolically engineered *E. coli* strains and investigate strategies for improving product levels. Engineered PAPS accumulation was shown to be necessary for in vivo sulfation of chondroitin but improving PAPS regeneration decreased the level of CS production (Figs. 2a and 2c). Optimizing pathway overexpression might overcome the metabolic burden of PAPS regeneration. There

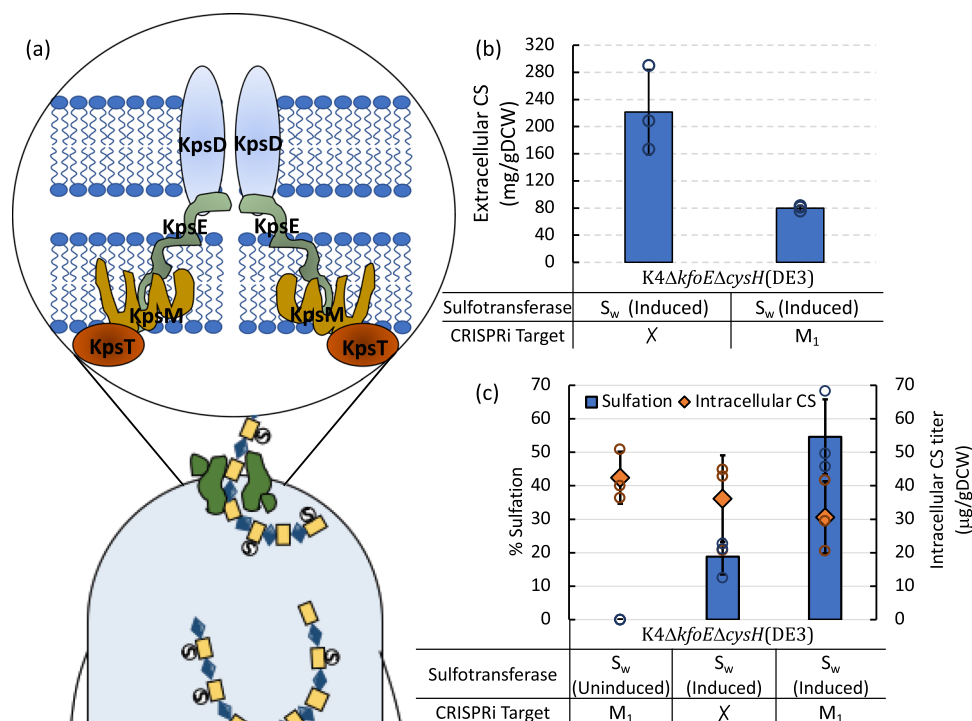


Fig. 5 Effect of export on CS sulfation. **a** Representation of currently accepted CPS transport complex in *E. coli* producing group 2 capsules. **b** Repression of GAG transport using CRISPRi in *K4ΔkfoEΔcysH(DE3)*pETM6- S_w showing 2.5-fold decreased GAG export. **c** CS sulfation in transport repressed *K4ΔkfoEΔcysH(DE3)* expressing S_w with uninduced and unrepressed controls. Error bars represent standard deviation from three biological replicates. Source data underlying Figs. 5b and c are provided as a Source Data file.

might also be unknown transcriptional and metabolic regulation in the PAPS pathway, reducing any further PAPS accumulation. We also showed that selecting a base strain (MG1655(DE3)) that accumulated higher amounts of PAPS on the deletion of *cysH* currently represents the best approach for achieving higher sulfation (Fig. 2d). Engineering PAPS accumulation in *E. coli* is still relatively unexplored and could serve beneficially to a wide range of sulfated biologics with only a couple of studies in recent years^{34,44}.

A second aspect of the current study examines the role of engineering carbohydrate sulfotransferases on GAG production. These Golgi enzymes of animal origin might be further improved through synthetic biology and protein engineering for enhanced expression and activity in prokaryotic cells. While some studies describe such expression attempts with carbohydrate sulfotransferases^{35,51,56}, they have only further employed them in vitro for metabolite production. In the case of chondroitin sulfotransferases, rational protein design is currently limited by the lack of structural information. This study relied on a homology model based on the known structure of a different enzyme, a sulfotransferase component of olefin synthase⁵³. This enzyme has a low sequence similarity to chondroitin sulfotransferase. Despite this limitation, the PROSS server predicted improved mutants (Fig. 3). It is noteworthy that the PROSS-predicted mutations, aimed at increasing stability and solubility, provided a three-fold enhancement of in vivo CS sulfation. This highlights the potential of more highly targeted enzyme engineering to further improve in vivo sulfation.

Our study also suggests that Golgi sulfotransferases of animal origin are a challenging group of proteins to actively express and conditions such as lower temperature are needed to achieve higher intracellular sulfation of chondroitin. Continuing growth at 37 °C post-induction resulted in no CS sulfation irrespective of induction OD_{600} or inducer concentration indicating poor

expression of active S_w at higher temperatures. Likewise, incubation at lower temperatures from the start of the fermentation resulted in slow growth and in minimal sulfation. Hence, dropping the temperature to 20 °C or 16 °C upon induction was critical for in vivo sulfation. Expression at 16 °C resulted in improved CS sulfation in cultures induced at 0.6 OD_{600} , while expression at 20 °C resulted in improved sulfation for cultures induced at 1.0 OD_{600} (Fig. 4). In contrast to sulfotransferase expression, the production and export of chondroitin is relatively straightforward and is optimal at 37 °C. In addition to engineering sulfotransferases for better in vivo activity, future studies should also aim at engineering precursor production and CS export to work optimally through a range of temperatures.

Another important finding of the current study is the effect of repression of CS export on sulfation levels (Fig. 5). The GAG transporter genes *kpsT* and *kpsM* are located in region III of the capsular polysaccharide gene cluster of *E. coli* K4 under the control of a thermoregulated promoter⁴¹. Hence it is possible that the temperature drop used for sulfotransferase expression represses the export of CS. This temperature effect together with CRISPRi-based *kpsM* repression, could significantly impede CS export, thereby allowing more time for the sulfation of chondroitin. This is consistent with the three-fold increase in sulfation of chondroitin observed in this strain. However, this metabolic design rules out CS secretion, making the approach less attractive for scale-up and downstream processing. Future studies should take into account strain designs that can slow down export while not eliminating it completely.

We also observed that the CS sulfation levels dropped when we tried combining the successful approaches observed above. *K4ΔkfoEΔcysH(DE3)* pETM6- S_{M2} pdCas9-dM1 and MG1655Δ*cysH*(DE3) pETM6-PCAFS S_{M2} showed only about 4% and 5% sulfation of chondroitin. Two *E. coli* strains developed in this study successfully synthesized high levels of 4-*O*-sulfated chondroitin,

CS-A. These were - K4 Δ kfoE Δ cysH(DE3) pETM6-Sw pdCas9-dM1 and MG1655 Δ cysH(DE3) pETM6-PCAFSw. The CS obtained from these microbial strains is comparable to industry-standard, animal-sourced CS-A from bovine trachea. The key take-away is that the level of 4-O-sulfation in the animal (~70%) and microbial CS (~55%) was similar. More importantly, unlike animal-sourced CS-A, microbial CS did not contain the 6-O-sulfated isomer CS-C as an impurity. Details on the disaccharide compositions, along with comparisons to other commercially available CS-A are given in Supplementary Fig. 6.

Overall, this report describes prokaryotic production of a sulfated GAG by fermentation on glucose, minimal salts, sulfate, and casamino acids. This one-step, sustainable process offers a promising alternative to animal extraction in GAG production. This approach also offers the potential to prepare unnatural GAG derivatives and stable isotope-labeled GAGs⁴⁵. In improving the precursors required for CS synthesis, we also demonstrate an increased sulfation level of the intracellular product to ~96%, comparable to commercially available CS-A. Moreover, unlike the animal-derived CS, the microbial product does not contain any impurities such as CS-C. In addition to demonstrating that sulfated GAGs can be produced entirely by metabolically engineered prokaryotes, this study also explores factors that might be improved to achieve even higher sulfation and production levels. The impact of this work can be extended, as our results suggest that this strategy can be extrapolated to develop microbial cell factories for producing other types of GAGs as well.

Methods

Reagents, bacterial strains and plasmids. LB Broth (Lennox), salts, and reagents required for super optimal broth with catabolite repression (SOC) were procured from MilliporeSigma (St. Louis, MO). BD Difco™ M9 minimal media salts and BD Bacto™ casamino acids were procured from BD Biosciences (Franklin Lakes, NJ). Standard lithium salt of 3'-phosphoadenosine-5'-phosphosulfate (PAPS) and reagents required for disaccharide labeling were bought from MilliporeSigma (St. Louis, MO). CS disaccharide standards were purchased from Iduron (Manchester, UK). High performance liquid chromatography (HPLC)-grade solvents and salts used to prepare mobile phases were procured from Fisher Scientific (Springfield, NJ).

Bacterial strains used in this study are *E. coli* DH5 α , *E. coli* BL21Star(DE3), *E. coli* K-12 MG1655(DE3), and *E. coli* K4. ePathBrick vector pETM6 was used to overexpress Chondroitin and PAPS metabolic pathway genes⁴⁶. pETM6 and pET32LIC were used to express chondroitin-4-O-sulfotransferase and its mutants. Transformants were selected using ampicillin resistance that is conferred by the vector backbone, followed by colony polymerase chain reaction (PCR) and Sanger sequencing. CRISPRi repression relied on pdCas9 plasmid, a gift from Luciano Maraffini, carrying a nuclease-null Cas9 from *Streptococcus pyogenes* and a sgRNA scaffold⁴⁷. A list of all the plasmids and primers used in this study are given in Supplementary Table 2 and 3.

Construction of *E. coli* K4 Δ kfoE. *E. coli* K4 (Serovar O5:K4:H4) was engineered for the synthesis chondroitin. The fructosyltransferase encoded by *kfoE* was deleted by λ red recombineering techniques⁴⁸ resulting in K4 Δ kfoE. The FRT-flanked kanamycin resistance cassette was PCR amplified from pKD4 by deletion primers (Supplementary Table 3) with 40 nucleotides homologous regions near *kfoE* on the genome. The PCR product was purified by PCR cleanup kit (Cycle Pure Kit, Omega) and transformed into the λ red recombinase expressing *E. coli* K4. Positive knockout strains were screen by colony PCR and the transformed with pCP20, which expressed the flippase recombination enzyme, to remove the antibiotics resistance marker.

T7 RNA polymerase gene with lacUV5 promoter was integrated into the *LacZ* position in the *E. coli* K4 genome using the method described by Cox and coworkers⁴⁹. Briefly, a small fragment of "landing pad" with a tetracycline resistant marker was amplified from pTKS/CS with flanking 40 bp homologous regions of *LacZ*. Transformation of this purified linear DNA into K4 Δ kfoE expressing λ red recombinase enabled recombination and integration. Positive colonies were verified for successful integration of the landing pad. Next, T7 RNA-polymerase gene was cloned into the pTKIP vector and transformed into K4 Δ kfoE strains with landing-pad integration harboring pKDRED expressing yeast restriction enzyme I-SecI. Induction of I-SecI cuts at the landing pad and also cleaves out the T7-RNA-polymerase insert from pTKIP which is finally integrated into the landing pad region with the aid of λ red recombinases. This resulted in strain *E. coli* K4 Δ kfoE (DE3).

Deletion of PAPS reductase from K4 and MG1655. The *cysH* gene in *E. coli* encoding for PAPS reductase, was deleted using lambda red recombinase by the method described by Datsenko and Wanner⁴⁸. Briefly, a linear kanamycin resistance cassette with 40-bp homology arms to the two ends flanking the chromosomal *cysH* gene was amplified from pKD4 and transformed into host expressing recombinases from pKD46. On recombination, correctly deleted colonies were selected based on: kanamycin resistance; loss of ability to grow on M9 media (without casamino acids); size of chromosomal amplicon around the *cysH* gene region; and Sanger sequencing of the amplicon. Using this method, the *cysH* gene was deleted from *E. coli* strains K4 Δ kfoE(DE3) and MG1655(DE3).

Microbial growth conditions. Plate cultures of *E. coli* were grown by streaking glycerol stocks (frozen) onto LB agar plates with appropriate antibiotics. Starter cultures (5 mL) were grown in LB broth by shaking with antibiotics at 37 °C in 14 mL culture tubes until growth reached OD₆₀₀ of 0.6–0.8 (about 6 h). Flask cultures of chondroitinase and sulfotransferase producing strains were grown in 1 L of M9 medium supplemented with 80 μ g/mL ampicillin in PYREX Fernbach Culture Flasks (Corning Life Sciences). Flask cultures of CS-producing strains were grown in 125 mL Erlenmeyer flasks by inoculating 1% starter culture in 25 mL of M9 media supplemented with 1% glucose, 1% casamino acids and containing the appropriate antibiotics. Cellular growth was estimated using optical density of culture at 600 nm in a Biotek plate reader. Cells were grown at 37 °C until reaching an OD₆₀₀ of 0.6 and induced with 1 mM isopropyl-1-thio- β -D-galactopyranoside (IPTG), after which growth was continued at either 16 °C for 24 h or 20 °C for 12 h or 37 °C for 10 h. All liquid cultures were incubated in a rotary air shaker (New-Brunswick Scientific Innova 44 R) at 37 °C, 225 rpm. All CS-producing flask experiments were performed in triplicate.

Bioreactor conditions. Fed-batch fermentations were performed in 1-L Applikon bioreactors. The fermentations consisted of a batch phase, followed by a fed-batch phase. The seed media are LB-Lennox broth and composed of tryptone (10 g/L), yeast extract (5 g/L), sodium chloride (5 g/L), and required antibiotics. The pH was adjusted to 7.0 \pm 0.05. Overnight cultures (starter culture) were incubated at 37 °C (220 rpm). The starter culture was inoculated in seed 1 and the cells were grown at 37 °C (220 rpm) to a cell density (OD₆₀₀) of 0.6. The seed 2 cells were inoculated (10–15% by vol) in batch media. The batch media was composed of tryptone (8–12 g/L, typically 10 g/L), yeast extract (10–20 g/L typically, 17 g/L), glucose (10–20 g/L, typically 20 g/L), casein hydrolysate (8–15 g/L, typically 10 g/L), MgSO₄·7H₂O (0–2 g/L, typically, 1.2 g/L), and required antibiotics. The pH was maintained at 6.9 \pm 0.1, with 5 M NaOH. Feed was initiated when (OD₆₀₀) reached 8.0 \pm 2.5 at an initial rate of 2.5 mL/h/L and the feeding rate was adjusted, based on fluctuations in pH and dissolved oxygen (DO) levels. In some experiments, the feed was initiated 5 h EFT. The cells were fed with casamino acid solution (30–70 %, typically, 50 %) supplemented with glucose (300–500 g/L, typically 350 g/L), magnesium sulfate heptahydrate, (10–20 g/L, typically 12 g/L), yeast extract (0–100 g/L, typically, 50 g/L), and tryptone (0–100 g/L, typically 10 g/L). The sterilized (auto-claved) feed solution was maintained at 37 °C (during the feeding stage). The cell growth was monitored using OD₆₀₀ reading using a spectrophotometer and proBE 3000 Fiber Optic Total Biomass Sensor (BE3000) (BugLab, Concord, CA 94521, USA). The cells were grown at 37 °C for 10–15 EFT. When the OD₆₀₀ reached 22 (6.9–8.6 g/L CDW), the temperature was reduced to 16 °C. The cells were induced with 2 mM IPTG at 16 °C for 16–22 h. The dissolved oxygen was maintained using agitation (400–800 rpm), and air (0.5–1 LPM). Post-induction, the cells were harvested by centrifugation at 4000–4500 \times g for 10–20 min at 4 °C. The cell paste was weighed and stored in Ziploc freezer bags at –80 °C.

Molecular weight analysis. The molecular weight (MW) analysis of intact chondroitin sulfate (CS) chains was performed using a 15% carbohydrate gel, using a previously described method⁵² with some modifications. Briefly, the CS was purified from the cells and freeze dried. The freeze-died sample was dissolved in HPLC-grade water and approximately 10–30 μ g of chondroitin sulfate were loaded in each well. The gel was subjected to electrophoresis at 200 V for 20 min, washed at room temperature in DI water for 30 min, stained in Alcian blue solution for 15–30 min, and destained with several washes of DI water until the gel background was clear. Gel images were acquired using CanoScan LiDE 700 F at 600 dpi and processed for densitometry data using the Un-Scan-it gel 7.1 software (Silk Science Inc., Orem, UT, USA). The data were used to calculate the weight-averaged molecular weights (M_w), number averaged molecular weights (M_n), molecular weight ranges (M_R) and polydispersities (P).

Repression of PAPS reductase and GAG export using CRISPRi. CRISPRi was used to repress the expression of three genes—*cysH* encoding PAPS reductase, *kpsM* encoding the permease component of the capsular export complex and *kpsT* encoding the ATPase component for the capsular transport protein. pdCas9-mCherry⁴⁷ was cloned to incorporate spacer sequences into BsaI sites (golden gate cloning⁴⁰). Spacer sequences were selected based on the region just before the start codon of the genes containing the 5'-NGG PAM sequence for (d)Cas9. Successful clones were selected based on chloramphenicol resistance, colony color and sanger

sequencing. Sequences (Supplementary Table 3) and locations (Supplementary Fig. 4) of the spacers tested are included in the supporting information.

Computational protein redesign of sulfotransferase. The PROSS protein engineering server was used to identify mutations to improve the sulfotransferase³⁶. PROSS predicts mutations that improve protein stability through modification of protein features such as core packing, surface polarity, and backbone rigidity. We used a human chondroitin-4-*O*-sulfotransferase (S_w) sequence with a 60 amino acid truncation in the N-terminus, previously reported by our group⁵¹, to build a homology model structure in the Molecular Operating Environment (MOE) software suite (Chemical Computing Group ULC, Montreal, QC, Canada) using the structure of the sulfotransferase domain from *Synechococcus* PCC 7002 Olefin Synthase (PDB code: 4GOX)⁵³ as a template. Sequence alignment was generated between S_w and 4GOX to assess the similarity between the two sequences. Homology modeling tool in MOE generated 10 models with the following parameters enabled: C-terminal and N-terminal outgap modeling, automatic disulfide bond detection, and side-chain sampling set at 300 K using an Amber 10:EHT force field. Structural alignments and the Ramachandran statistics calculated for the models were used to assess how well the predicted structure conformed to the 4GOX structure and generally well-folded proteins.

Sulfotransferase mutant expression and purification. The three PROSS-predicted mutants of S_w , designated as S_{M1} , S_{M2} , and S_{M4} , were examined for improved activity in *E. coli*. Mutants S_{M1} and S_{M2} were derived from S_w in pET32LIC through multiple rounds of site-directed mutagenesis, while the S_{M4} gene was synthesized by IDT. The genes were cloned into the BamHI and XhoI sites of a pET32LIC vector with N-terminal thioredoxin (Trx) tag (to increase protein solubility) and His-6x tag (for purification). The fusion proteins were estimated to be ~53 kDa with a pI value of 6.85 (ExpPASy). The constructed plasmids were sequence verified and transformed into *E. coli* BL21Star (DE3). Overnight culture (20 mL) was centrifuged at $6800 \times g$ for 10 min at 25 °C and the pellet re-suspended in 1 L of M9. Sulfotransferase expression was induced at an OD_{600} of ~0.8 with 0.2 mM IPTG and the culture was incubated for 16–20 h at 22 °C.

Cells were harvested by centrifugation at $5000 \times g$ for 10 min at 4 °C and the pellet was sonicated upon re-suspension in 20 mL of 50 mM Tris-HCl buffer (pH 8.0, 500 mM NaCl, 30 mM imidazole). Cell debris was removed by centrifugation at $16,000 \times g$ for 1 h at 4 °C. Cell lysate was filtered and applied to a column containing Ni-NTA resin (Qiagen) and washed with buffer A (50 mM Tris-HCl 500 mM NaCl, 30 mM imidazole pH 7.5) and eluted with buffer B (50 mM Tris-HCl 500 mM NaCl, 300 mM imidazole pH 7.5). The imidazole was removed by buffer exchange and replaced with storage buffer (50 mM Tris-HCl 500 mM NaCl, 10% glycerol pH 7.5) and kept at –80 °C until needed. S_w , S_{M1} , S_{M2} , and S_{M4} were expressed and purified under identical conditions. The expression level and the purity of the target proteins were verified by SDS-PAGE using a NuPage 10% Bis-Tris Midi gel (Invitrogen).

PAPS estimation using HPLC/UV. On harvesting, cells were pelleted at 4 °C. Metabolites, including PAPS, were extracted from the pellet with two 30 min washes of 80% methanol solution at –80 °C. Pooled extracts could be stored at –20 °C until further analysis. PAPS concentration in the extract was estimated by HPLC using a 150 × 2 mm Develosil C-30 RPAqueous column (manufactured by Nomura Chemicals, Japan and purchased from Phenomenex, Inc., USA) in an Agilent LC1260 instrument. Potassium phosphate buffer (100 mM, pH 5.8) and 75% acetonitrile (in H₂O) were used as mobile phases A and B respectively. The samples were run on a 40 min protocol (adapted from Furuno and co-workers⁵⁴) at an overall flow rate of 0.2 mL/min. The gradient program was set as follows: 0% B from 0 to 10 min; 0–50% B (linear ramp) from 10 to 12 min; 50% B from 12 to 17 min; 50 to 0% B (linear ramp) from 17 to 20 min and 0% B from 20 to 40 min. Standard PAPS (detected using PDA detector at 260 nm) diluted in mobile phase A elutes at 3.1 min.

GAG extraction and disaccharide analysis using LC/MS. Extracellular GAGs produced in each flask culture were recovered in the solution phase (spent media) after centrifugation. Intracellular GAGs were recovered by re-suspending the cell pellet, autoclaving to prepare cell lysate, and centrifuging to recover the soluble phase. Both solutions containing extracellular and intracellular GAGs were precipitated with 4 volumes ethanol and stored at –20 °C for 12 h in an explosion-proof freezer. The precipitates were collected, dried, and re-dissolved in 0.2 volume sterile water to generate GAG extracts that were stored at –20 °C until further use.

Extracted GAG solutions (100 µL) were passed through a 3 kDa spin column to remove small molecules and to exchange with digestion buffer (50 mM ammonium acetate, 2 mM CaCl₂ (pH 7.4)). GAG solutions were added to 200 µL of digestion buffer and 20 mU purified chondroitinase ABC (25 mM Tris, 500 mM NaCl, 300 mM imidazole buffer (pH 7.4)) and incubated at 37 °C for 12 h for depolymerization. The resulting disaccharides were passed through a 3-kDa spin column, then the filtrate was collected and lyophilized. The freeze-dried disaccharide samples were fluorescently labeled by dissolving in 10 µL of a 0.1 M 2-aminoacridone (AMAC) (17:3 of dimethyl sulfoxide:acetic acid (v/v)). After incubation for 10 min at room temperature, the reaction mixture was

supplemented with 10 µL of 1 M NaBH₃CN, vortex-mixed, and incubated at 45 °C for 1 h. Samples were centrifuged and the supernatant containing the labeled disaccharides was analyzed. The AMAC-labeled disaccharides were separated by HPLC on an Agilent Poroshell 120, EC-C18 column (Agilent Technologies, Inc. Wilmington, DE) using an Agilent 1200 HPLC system with detection by a TSQ Quantum triple quadrupole electron-spray ionization mass spectrometer (Thermo Finnigan, San Jose, CA)⁵⁵. Data were processed to identify disaccharide levels using the Thermo Xcalibur software.

In vitro sulfotransferase assays. Colorimetric activity assay followed a previously published method with some adaptations⁵⁶. Briefly, the total assay volume was 200 µL, consisting of 100 µL 50 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) buffer, 20 µL *p*-nitrophenyl sulfate (PNPS) (20 mM), 20 µL chondroitin (1 mg/mL), 20 µL of 1 mg/mL AST-IV, 20 µL purified C4ST (~1 mg/mL), and 20 µL PAPS (2.5 mM). The assay solution was mixed, with PAPS added immediately before absorbance measurements were started. The temperature controlled SpectraMax plate reader (Molecular Device, Sunnyvale, CA) was pre-incubated at 37 °C, then the formation of PNP was detected at absorbance 400 nm. The reactions were allowed to continue at 37 °C overnight, then processed for disaccharide analysis.

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

Data availability

Data supporting the findings of this work are available within the paper and its Supplementary Information files. A reporting summary for this article is available as a Supplementary Information file. The datasets generated and analyzed during the current study are available from the corresponding author upon request. Source data are provided with this paper.

Received: 27 June 2020; Accepted: 5 February 2021;

Published online: 02 March 2021

References

- Köwitsch, A., Zhou, G. & Groth, T. Medical application of glycosaminoglycans: a review. *Medical application of glycosaminoglycans. J. Tissue Eng. Regen. Med.* **12**, e23–e41 (2018).
- Franchini, M., Liunbruno, G. M., Bonfanti, C. & Lippi, G. Carrion's disease after blood transfusion. *Blood Transfusion* <https://doi.org/10.2450/2015.0096-15> (2015).
- Bishnoi, M., Jain, A., Hurkat, P. & Jain, S. K. Chondroitin sulphate: a focus on osteoarthritis. *Glycoconj. J.* **33**, 693–705 (2016).
- Henrotin, Y., Mathy, M., Sanchez, C. & Lambert, C. Chondroitin sulfate in the treatment of osteoarthritis: from in vitro studies to clinical recommendations. *Therapeut. Adv. Musculoskel.* **2**, 335–348 (2010).
- Rogers, C. J. & Hsieh-Wilson, L. C. in *Carbohydr. Microarrays* (ed. Chevolut, Y.) vol. 808, 321–336 (Humana Press, 2012).
- Liu, H., Zhang, Z. & Linhardt, R. J. Lessons learned from the contamination of heparin. *Nat. Prod. Rep.* **26**, 313 (2009).
- Kishimoto, T. K. et al. Contaminated heparin associated with adverse clinical events and activation of the contact system. *N. Engl. J. Med.* **358**, 2457–2467 (2008).
- Laurencin, C. T. & Nair, L. The FDA and safety—beyond the heparin crisis. *Nat. Biotechnol.* **26**, 621–623 (2008).
- Brennan, Z. *Chinese Heparin Contamination Questions Return With New FDA Warning Letter* (Regul. FocusTM, 2017).
- Kubaski, F. et al. Glycosaminoglycans detection methods: applications of mass spectrometry. *Mol. Genet. Metabol.* **120**, 67–77 (2017).
- Ahn, M. Y. et al. Anti-cancer effect of dung beetle glycosaminoglycans on melanoma. *BMC Cancer* **19**, 9 (2019).
- Chen, Y. et al. Glycosaminoglycan from *Apostichopus japonicus* improves glucose metabolism in the liver of insulin resistant mice. *Mar. Drugs* **18**, 1 (2019).
- Levene, P. A. & La Forge, F. B. On chondroitin sulphuric acid. *J. Biol. Chem.* **15**, 69–79 (1913).
- Volpi, N. Disaccharide mapping of chondroitin sulfate of different origins by high-performance capillary electrophoresis and high-performance liquid chromatography. *Carbohydr. Polym.* **55**, 273–281 (2004).
- Volpi, N. Quality of different chondroitin sulfate preparations in relation to their therapeutic activity. *J. Pharm. Pharmacol.* **61**, 1271–1280 (2009).
- Shipp, E. L. & Hsieh-Wilson, L. C. Profiling the sulfation specificities of glycosaminoglycan interactions with growth factors and chemotactic proteins using microarrays. *Chem. Biol.* **14**, 195–208 (2007).
- Djeral, L., Lortat-Jacob, H. & Kwok, J. Chondroitin sulfates and their binding molecules in the central nervous system. *Glycoconj. J.* **34**, 363–376 (2017).

18. Izumikawa, T., Sato, B. & Kitagawa, H. Chondroitin sulfate is indispensable for pluripotency and differentiation of mouse embryonic stem cells. *Sci. Rep.* **4**, 3701 (2015).
19. Sirkko, S., von Holst, A., Wizenmann, A., Gotz, M. & Faissner, A. Chondroitin sulfate glycosaminoglycans control proliferation, radial glia cell differentiation and neurogenesis in neural stem/progenitor cells. *Dev.* **134**, 2727–2738 (2007).
20. Zhu, Z. et al. Sulfated polysaccharide from sea cucumber and its depolymerized derivative prevent obesity in association with modification of gut microbiota in high-fat diet-fed mice. *Mol. Nutr. Food Res.* **62**, 1800446 (2018).
21. Badri, A., Williams, A., Linhardt, R. J. & Koffas, M. A. The road to animal-free glycosaminoglycan production: current efforts and bottlenecks. *Curr. Opin. Biotechnol.* **53**, 85–92 (2018).
22. Rondanelli, M. et al. Effectiveness of non-animal chondroitin sulfate supplementation in the treatment of moderate knee osteoarthritis in a group of overweight subjects: a randomized, moderate-blind, placebo-controlled pilot study. *Nutrients* **11**, 2027 (2019).
23. Glass, C. A. Recombinant heparin—new opportunities. *Front. Med.* **5**, 341 (2018).
24. Zhu, H.-M. et al. KfoA, the UDP-glucose-4-epimerase of *Escherichia coli* strain O5:K4:H4, shows preference for acetylated substrates. *Appl. Microbiol. Biotechnol.* **102**, 751–761 (2018).
25. Bélanger, M., Burrows, L. L. & Lam, J. S. Functional analysis of genes responsible for the synthesis of the B-band O antigen of *Pseudomonas aeruginosa* serotype O6 lipopolysaccharide. The GenBank accession number for the sequence reported in this paper is AF035937. *Microbiology* **145**, 3505–3521 (1999).
26. Bengoechea, J. A. et al. Functional characterization of Gne (UDP-N-Acetylglucosamine-4-Epimerase), Wzz (chain length determinant), and Wzy (O-Antigen Polymerase) of *Yersinia enterocolitica* Serotype O:8. *J. Bacteriol.* **184**, 4277–4287 (2002).
27. Cunneen, M. M., Liu, B., Wang, L. & Reeves, P. R. Biosynthesis of UDP-GlcNAc, UndPP-GlcNAc and UDP-GlcNAcA Involves Three Easily Distinguished 4-Epimerase Enzymes, Gne, Gnu and GnaB. *PLoS ONE* **8**, e67646 (2013).
28. Guo, H., Li, L. & Wang, P. G. Biochemical characterization of UDP-GlcNAc/Glc4-epimerase from *Escherichia coli* O86:B7 †. *Biochemistry* **45**, 13760–13768 (2006).
29. Xu, S. et al. Chain structure and immunomodulatory activity of a fructosylated chondroitin from an engineered *Escherichia coli* K4. *Int. J. Biol. Macromol.* **133**, 702–711 (2019).
30. Zhang, Q. et al. Enhancing fructosylated chondroitin production in *Escherichia coli* K4 by balancing the UDP-precursors. *Metabol. Eng.* **47**, 314–322 (2018).
31. He, W. et al. Production of chondroitin in metabolically engineered *E. coli*. *Metabol. Eng.* **27**, 92–100 (2015).
32. Liu, J. et al. KfoE encodes a fructosyltransferase involved in capsular polysaccharide biosynthesis in *Escherichia coli* K4. *Biotechnol. Lett.* **36**, 1469–1477 (2014).
33. Trilli, A., Busiello, I., Daly, S. & Bagatin, F. Biotechnological production of chondroitin. US Patent 8609394B2 (2014).
34. Badri, A., Williams, A., Xia, K., Linhardt, R. J. & Koffas, M. A. G. Increased 3'-phosphoadenosine-5'-phosphosulfate levels in engineered *Escherichia coli* cell lysate facilitate the in vitro synthesis of chondroitin sulfate. *biotechnol. J* **14**, 1800436 (2019).
35. Nagai, N. et al. Stem domains of heparan sulfate 6-O-sulfotransferase are required for Golgi localization, oligomer formation and enzyme activity. *J. Cell Sci.* **117**, 3331–3341 (2004).
36. Goldenzweig, A. et al. Automated structure- and sequence-based design of proteins for high bacterial expression and stability. *Mol. Cell* **63**, 337–346 (2016).
37. Jones, J. A. & Koffas, M. A. G. Optimizing metabolic pathways for the improved production of natural products. *Meth. Enzymol.* **575**, 179–193 (2016).
38. Ahmadi, M. K. & Pfeifer, B. A. Improved heterologous production of the nonribosomal peptide-polyketide siderophore yersiniabactin through metabolic engineering and induction optimization. *Biotechnol. Progr.* **32**, 1412–1417 (2016).
39. Jones, J. A. et al. ePathOptimize: a combinatorial approach for transcriptional balancing of metabolic pathways. *Sci. Rep.* **5**, 11301 (2015).
40. He, W. Metabolic Engineering and Applied Enzymology for the Preparation of Nutraceutical/ Pharmaceutical Chondroitin Sulfate. PhD thesis, Rensselaer Polytechnic Institute (2017).
41. Haas, E. K. *Analyses Of The Proteins Kpsm, Kpse And Kpsd In The Group 2 Capsular Polysaccharide Export Complex Of Escherichia Coli* (University of Manchester, 2012).
42. Kamhi, E., Joo, E. J., Dordick, J. S. & Linhardt, R. J. Glycosaminoglycans in infectious disease: Glycosaminoglycans in infectious disease. *Biol. Rev.* **88**, 928–943 (2013).
43. Kim, S. Y. et al. Glycosaminoglycan binding motif at S1/S2 proteolytic cleavage site on spike glycoprotein may facilitate novel coronavirus (SARS-CoV-2) host cell entry. *bioRxiv* <https://doi.org/10.1101/2020.04.14.041459> (2020).
44. Chu, L. L. et al. Metabolic engineering of *Escherichia coli* for enhanced production of naringenin 7-sulfate and its biological activities. *Front. Microbiol.* **9**, e1671 (2018).
45. Cress, B. F. et al. Heavy heparin: a stable isotope-enriched, chemoenzymatically-synthesized, poly-component drug. *Angew. Chem. Int. Ed.* **58**, 5962–5966 (2019).
46. Xu, P., Vansiri, A., Bhan, N. & Koffas, M. A. G. ePathBrick: a synthetic biology platform for engineering metabolic pathways in *E. coli*. *ACS Synth. Biol.* **1**, 256–266 (2012).
47. Bikard, D. et al. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* **41**, 7429–7437 (2013).
48. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97**, 6640–6645 (2000).
49. Kuhlman, T. E. & Cox, E. C. Site-specific chromosomal integration of large synthetic constructs. *Nucleic Acids Res.* **38**, e92–e92 (2010).
50. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* **3**, e3647 (2008).
51. He, W. et al. Expression of chondroitin-4-O-sulfotransferase in *Escherichia coli* and *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* **101**, 6919–6928 (2017).
52. Ly, Mellisa et al. Analysis of *E. coli* K5 capsular polysaccharide heparosan. *Anal. Bioanal. Chem.* **399**, 737–745 (2011).
53. McCarthy, J. G. et al. Structural basis of functional group activation by sulfotransferases in complex metabolic pathways. *ACS Chem. Biol.* **7**, 1994–2003 (2012).
54. Imamura, M., Kumagai, T., Sugihara, N. & Furuno, K. High-performance liquid chromatographic assay of 3'-Phosphoadenosine 5'-phosphosulfate (PAPS) and UDP-glucuronic acid (UDPGA) in cultured hepatic cell extracts. *J. Health Sci.* **49**, 395–400 (2003).
55. Yang, B., Chang, Y., Weyers, A. M., Sterner, E. & Linhardt, R. J. Disaccharide analysis of glycosaminoglycan mixtures by ultra-high-performance liquid chromatography–mass spectrometry. *J. Chromatogr. A* **1225**, 91–98 (2012).
56. Sterner, E. et al. Assays for determining heparan sulfate and heparin O-sulfotransferase activity and specificity. *Anal. Bioanal. Chem.* **406**, 525–536 (2014).

Acknowledgements

This work was supported by the National Science Foundation Grant (CBET- 1604547) and GRL NRF-2016K1A1A2912829.

Author contributions

The study was mainly designed and written by A.B., R.J.L., J.S.D., and M.A.G.K. R.J.L. and M.A.G.K. supervised the work and are equally contributing corresponding authors. K.F. performed the computational analysis of sulfotransferase under the supervision of J.S.D. A.W. constructed and cloned the mutant sulfotransferases and performed in vitro activity assays. W.H. performed constructed the strain K4 $\Delta kfoE$ (DE3). A.B. constructed strains and performed in vivo shake-flask experiments for CS production in *E. coli*. A.A., P.D., and K.X. performed bioreactor experiments, molecular weight analysis, and related CS characterization. All authors participated in the writing of the manuscript and approved the final version.

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/s41467-021-21692-5>.

Correspondence and requests for materials should be addressed to M.A.G.K.

Peer review information *Nature Communications* thanks Hiroshi Kitagawa, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021