

## Article

# Novel Insights into the Existence of the Putative UDP-Glucuronate 5-Epimerase Specificity

Ophelia Gevaert, Stevie Van Overtveldt, Matthieu Da Costa, Koen Beerens and Tom Desmet \*

Centre for Synthetic Biology, Department of Biotechnology, Ghent University, Coupure Links 653, 9000 Gent, Belgium; ophelia.gevaert@ugent.be (O.G.); stevie.vanovertveldt@ugent.be (S.V.O.); matthieu.dacosta@ugent.be (M.D.C.); koen.beerens@ugent.be (K.B.)

\* Correspondence: tom.desmet@ugent.be; Tel.: +32-9264-9920

Received: 15 January 2020; Accepted: 9 February 2020; Published: 12 February 2020

**Abstract:** C5-epimerases are promising tools for the production of rare L-hexoses from their more common D-counterparts. On that account, UDP-glucuronate 5-epimerase (UGA5E) attracts attention as this enzyme could prove to be useful for the synthesis of UDP-L-iduronate. Interestingly, L-iduronate is known as a precursor for the production of heparin, an effective anticoagulant. To date, the UGA5E specificity has only been detected in rabbit skin extract, and the respective enzyme has not been characterized in detail or even identified at the molecular level. Accordingly, the current work aimed to shed more light on the properties of UGA5E. Therefore, the pool of putative UGA5Es present in the UniProt database was scrutinized and their sequences were clustered in a phylogenetic tree. However, the examination of two of these enzymes revealed that they actually epimerize UDP-glucuronate at the 4- rather than 5-position. Furthermore, *in silico* analysis indicated that this should be the case for all sequences that are currently annotated as UGA5E and, hence, that such activity has not yet been discovered in nature. The detected L-iduronate synthesis in rabbit skin extract can probably be assigned to the enzyme chondroitin-glucuronate C5-epimerase, which catalyzes the conversion of D-glucuronate to L-iduronate on a polysaccharide level.

**Keywords:** UDP-glucuronate 5-epimerase; UDP-glucuronate 4-epimerase; CEP1 family; L-iduronate; biocatalysis; phylogenetic tree; recombinant expression

## 1. Introduction

Carbohydrate epimerases (CEP, EC 5.1.3) catalyze the inversion of the configuration of a chiral carbon atom in molecules with more than one stereocenter [1]. Accordingly, they are able to create shortcuts in the production of rare sugars starting from their abundant counterparts [2]. Over 30 different carbohydrate epimerases have been described to date, covering various substrate and product specificities. Despite their seemingly simple net reaction, the mechanisms employed by these enzymes are diverse and sometimes quite complex. Recently, a comprehensive classification was proposed for epimerases, which consists of 14 families (CEP1–CEP14) in which structure and mechanism are conserved [3]. Thereof, the CEP1 family is the largest, containing 12 different specificities on nucleoside diphosphate (NDP) sugars that share a mechanism based on a transient keto intermediate (Table 1). This family is a rich pool of epimerases displaying diverse substrate and product specificities. Interestingly, this means that comparable active site architectures are able to bind different substrate molecules and yield varying reaction products.

Based on the CEP classification, two C5-epimerases are present in the CEP1 family. These enzymes are of interest as they might contribute to the production of L-hexoses and derivatives, with applications as for instance antibiotics or nucleosides [4–6]. On the one hand, the double

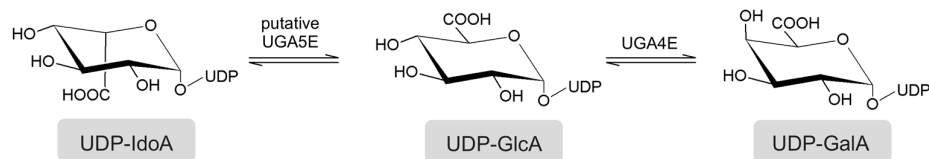
epimerization of GDP-mannose yielding GDP-L-galactose and GDP-L-gulose is catalyzed by GDP-mannose 3,5-epimerase (GM35E, EC 5.1.3.18), an enzyme that has already been thoroughly characterized [7,8]. On the other hand, the UDP-glucuronate 5-epimerase specificity (UGA5E, EC 5.1.3.12) would ensure the interconversion of UDP-D-glucuronic acid (UDP-GlcA) and UDP-L-iduronic acid (UDP-IdoA) (Figure 1) [9]. This enzyme might be industrially relevant as UDP-IdoA is known as a precursor of heparin, a glycosaminoglycan (GAG) consisting of sulfated disaccharide repeating units of uronic acids (GlcA or IdoA) and D-glucosamine residues. This compound displays anticoagulant properties and is already practiced clinically [10,11]. Currently, the drug is isolated from animal tissue, yielding a heterogenous mixture [12]. However, research is focusing on the establishment of a (bio)chemical production process. One of the hurdles there is the synthetic access to L-iduronic acid and L-idose as building blocks [13]. In that respect, UGA5E might prove to be a very useful biocatalyst. In addition, the structural characterization of UGA5E might contribute to a better understanding of the CEP1 specificity determinants. Indeed, the comparison of the reaction mechanisms of GM35E and UGA5E could give more insight into the specific parameters that potentially promote or prevent a double epimerization, respectively.

To date, UGA5E activity has only been detected in rabbit skin extracts, where the enzyme was believed to play a role in the synthesis of dermatan sulfate (DS, formerly chondroitin sulfate B) [9]. The structure of this GAG is comparable to heparin, but DS is composed of *N*-acetyl-D-galactosamine (GalNAc) building blocks instead of D-glucosamine residues [14]. Moreover, the IdoA subunits even function as key components in the binding site specificity for GAG-binding proteins [15]. The compound is expressed in various mammalian tissues and is involved in cardiovascular disease, tumorigenesis, infection, wound repair, and fibrosis. Jacobson et al. proposed a three-stage process for the production of this polysaccharide, starting from the conversion of UDP-glucose (UDP-Glc) to UDP-GlcA by the action of an UDP-glucose dehydrogenase. Next, UGA5E would catalyze the synthesis of UDP-IdoA from UDP-GlcA. Eventually, L-iduronic acid would be transferred to DS by a glycosyltransferase [9]. This pathway is comparable to that involving UDP-glucuronate 4-epimerase (UGA4E, EC 5.1.3.6), an enzyme that catalyzes the interconversion of UDP-GlcA and UDP-D-galacturonic acid (UDP-GalA) (Figure 1). The latter serves as a donor for GalA incorporation in plant cell walls, bacterial lipopolysaccharides, or antibiotics [16–20]. UGA4E is also present in the CEP1 family [3].

Despite the detection of UGA5E activity in the tissue extracts, the respective gene was never identified or recombinantly expressed. Nevertheless, sequence databases contain various putative UGA5Es on which their assignment to the CEP1 family is based. On that account, this research focused on the analysis of these putative sequences and on the UGA5E specificity in general.

**Table 1.** Overview of carbohydrate epimerase family 1. The specificities marked with an asterisk have a solved crystal structure.

EC Number	Specificity
5.1.3.2	UDP-glucose 4-epimerase*
5.1.3.10	CDP-paratose 2-epimerase*
5.1.3.6	UDP-glucuronate 4-epimerase*
5.1.3.7	UDP- <i>N</i> -acetylglucosamine 4-epimerase*
5.1.3.16	UDP-glucosamine 4-epimerase
5.1.3.25	dTDP-L-rhamnose 4-epimerase
5.1.3.26	<i>N</i> -acetyl- $\alpha$ -D-glucosaminyl-diphospho-ditrans,octakis-undecaprenol 4-epimerase
5.1.3.5	UDP-L-arabinose 4-epimerase
5.1.3.12	UDP-glucuronate 5-epimerase
5.1.3.20	ADP-L-glyceromanno-heptose 6-epimerase*
/	UDP-2-acetamido-2,6-dideoxy-beta-L-arabino-4-hexulose 3-epimerase
5.1.3.18	GDP-mannose 3,5-epimerase*



**Figure 1.** Epimerization of UDP-glucuronic acid (UDP-GlcA) toward UDP-L-iduronic acid (UDP-IdoA) and UDP-galacturonic acid (UDP-GalA) by UDP-glucuronate 5-epimerase (UGA5E) and UDP-glucuronate 4-epimerase (UGA4E), respectively.

## 2. Results

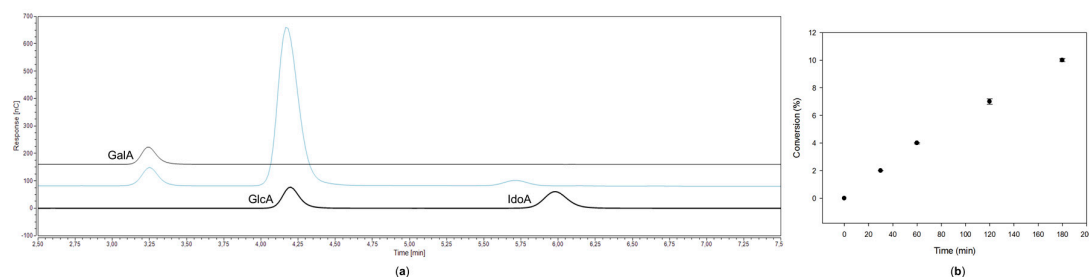
### 2.1. Selection of a UGA5E Sequence

The UniProt database contains 1154 putative ‘UDP-glucuronate 5-epimerase’ sequences, of which 20 are derived from eukaryotes, 10 from archaea, and the others from bacteria. One is even listed as a reviewed sequence, but this enzyme from *Rhizobium meliloti* (RmUGAE, UniProt: O54067) is most likely incorrectly annotated as the literature accompanying this entry actually discusses an UGA4E [21]. Consequently, this sequence was excluded from further analysis.

As the UGA5E specificity might be valuable in an industrial context, a (thermo)stable representative from a bacterial source would be most relevant. Indeed, a bacterial variant should enable facilitated recombinant expression in prokaryotes, and (thermo)stability entails a higher tolerance for the harsh process conditions practiced in industry [22,23]. Therefore, the UniProt database was screened for UGA5E enzymes that meet these conditions. This analysis highlighted three promising sequences, namely, those from *Thermacetogenium phaeum* (UniProt: K4LEK1), *Desulfurobacterium thermolithotrophum* strain DSM 11699 (UniProt: F0S202), and *Thermodesulfobacterium geofontis* (UniProt: F8C4X8). The optimal growth temperatures of these organisms are 58, 70, and 83 °C, respectively [24–27]. For the current study, the highest temperature was selected, i.e., the putative UGA5E from *T. geofontis* (TgUGAE).

### 2.2. Evaluation of TgUGAE

First, the inducible expression of TgUGAE in *E. coli* was evaluated and seemed to be the highest at an incubation temperature of 20 °C (Figure S1). As it was found that the enzyme was mainly expressed in the insoluble fraction, activity tests were performed with crude cell lysate. The conversion of UDP-GlcA was followed over time and the formation of one reaction product was observed. The retention time of the hydrolyzed product on HPLC corresponded to that of GalA, the C4-epimer of GlcA (Figure 2). Further analysis confirmed that this detected UGA4E activity was derived from the overexpressed TgUGAE sequence and not from the accompanying *E. coli* background (Figure S2). The possibility that the biocatalyst would display a side activity as C5-epimerase was disproved as no extra product was obtained, even after extensive incubation and spiking the reaction with fresh lysate. Furthermore, activity tests with purified enzyme and 5 mM additional NAD<sup>+</sup> resulted in the same reaction profile. The extra peak in the chromatogram at 5.70 min was also present in control samples and was, thus, not the result of any enzymatic activity (Figure S3). Consequently, it could be concluded that the putative UGA5E sequence from *Thermodesulfobacterium geofontis* is actually an UGA4E, with a specific activity of 1.7 (± 0.2) U/mg (using 10 mM UDP-GlcA at pH 7.5 and 30 °C).



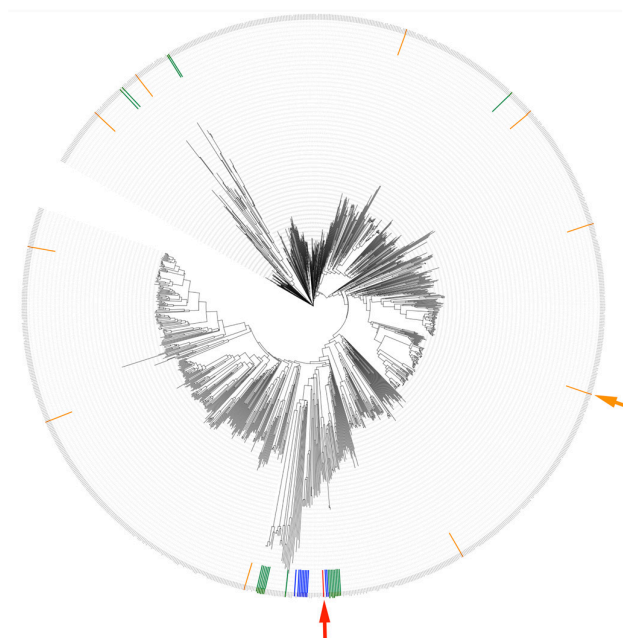
**Figure 2.** Product formation by *TgUGAE* crude cell lysate. (a) Chromatographic analysis of the reaction mixture (blue). The peaks were compared to known standards (black). The nucleoside diphosphate (NDP)-sugars of the reaction mixture were hydrolyzed to their respective monosaccharides prior to analysis. GalA: galacturonic acid; GlcA: glucuronic acid; IdOA: L-iduronic acid. (b) Conversion of UDP-glucuronic acid toward UDP-galacturonic acid over time.

### 2.3. *In Silico* Analysis of the UGA5E Specificity

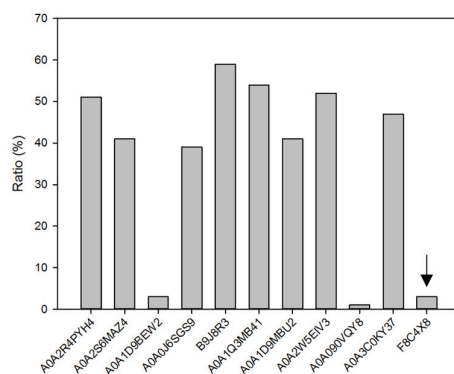
The further quest for a true UGA5E was assisted by an *in silico* approach. Commonly, unknown genes are annotated based on sequence similarities. As no reference sequence is available for UGA5E, the annotation of potential UGA5E sequences is rather arbitrary. The hypothesized UGA5E specificity, and the CEP1 family, in general, is part of the short-chain dehydrogenase/reductase (SDR) superfamily, a large cluster of NAD(P)(H)-dependent oxidoreductases [3,28]. Members of this family display low sequence identities (typically 15–30%), though they share some common motifs such as an GxxGxxG motif for NAD(P)<sup>+</sup> binding within a Rossmann fold and an YxxxK catalytic motif. Based on these features, sequences can be classified in the SDR superfamily. Subsequently, specific conserved residues and secondary structures can contribute to the distinction of various enzymatic specificities, e.g., the conserved catalytic cysteine and lysine in GM35E [8]. Due to the lack of a reference sequence for UGA5E, the determination of these specific conserved residues is thus unachievable at the moment. On that account, the actual function of the putative UGA5E sequences is dubious.

This statement is clearly demonstrated when performing a Basic Local Alignment Search Tool (BLAST) analysis with the *TgUGAE* sequence. Indeed, 19% of the resulting sequences are annotated as ‘NAD-dependent epimerase/dehydratase/oxidoreductase,’ which implies that they are part of the SDR superfamily and this annotation is, thus, most probably based on the sequence motifs. Another 72% of the BLAST outcome consists of ‘NDP-sugar epimerase’ annotated sequences. The remaining results (9%) comprise a variety of specific epimerases, such as UGA5Es and UGA4Es.

In order to find a sequence with a high probability of being an UGA5E, a phylogenetic tree of all putative UGA5E sequences was constructed (Figure 3). Next, ten bacterial sequences from around the tree were randomly selected for further investigation. A BLAST search was performed with each of the sequences and the ratio of ‘UDP-glucuronate 5-epimerase’ as the outcome compared to all results was analyzed (Figure 4). It was hypothesized that enzymes with a high ratio have a bigger chance to display the UGA5E specificity as the annotation is based on a common feature among the putative UGA5E sequences. Accordingly, the UGA5E from *Agrobacterium radiobacter* strain K84 (ArUGAE, UniProt: B9J8R3) was selected as being the sequence with the highest ratio (59%). Interestingly, according to this analysis, *TgUGAE* displayed a very low ratio of only 3%. The fact that this enzyme was then actually a C4-epimerase substantiates the above reasoning.



**Figure 3.** Phylogenetic tree of the putative UGA5E sequences present in the UniProt database. Eukaryotic sequences are colored in green, whereas the sequences derived from archaea are highlighted in blue. A BLAST search was performed with the sequences represented in orange. *TgUGAE* is presented in red. The arrows indicate the sequences that were experimentally tested in this research.

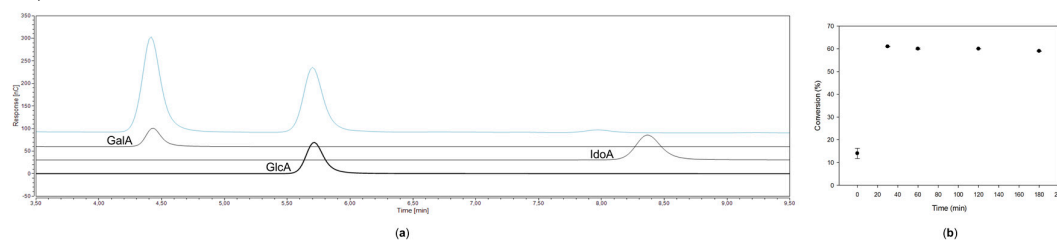


**Figure 4.** Ratio of ‘UDP-glucuronate 5-epimerase’ sequences in the BLAST results of the randomly selected putative UGA5E sequences. The sequences are represented by their UniProt entry. *TgUGAE* is indicated by an arrow.

#### 2.4. Evaluation of *ArUGAE*

The recombinant expression of *ArUGAE* was, again, first evaluated, and an incubation temperature of 20 °C was found to be most suitable (Figure S4). As was the case for *TgUGAE*, the enzyme was mainly present in the insoluble fraction and, thus, the activity tests were performed with crude cell lysate. As can be seen in Figure 5, this enzyme also catalyzes the interconversion of UDP-GlcA and UDP-GalA and thus displays an UGA4E specificity instead of the hypothesized UGA5E specificity. Again, extra controls were included to eliminate the possibility that the UGA4E activity was derived from the *E. coli* background and to assign the extra peak in the chromatogram (Figures

S5 and S6). Moreover, the addition of NAD<sup>+</sup> had no influence on the epimerization specificity. Accordingly, this sequence from *Agrobacterium radiobacter* should be designated as UGA4E. The enzyme displayed a specific activity of 12.6 ( $\pm$  1.2) U/mg (using 10 mM UDP-GlcA at pH 7.5 and 30 °C).



**Figure 5.** Specificity testing of *ArUGAE* crude cell lysate. (a) Product analysis on HPLC of the *ArUGAE* reaction mixture (blue). Standards are represented in black. The reaction mixture was hydrolyzed prior to analysis to obtain free monosaccharides. GalA: galacturonic acid; GlcA: glucuronic acid; IdOA: L-iduronic acid. (b) Conversion of UDP-glucuronic acid to UDP-galacturonic acid over time.

### 2.5. Analysis of the UGA5E Specificity

Both examined putative UGA5E sequences actually appeared to be UGA4Es. These findings raise questions about the specificity of the other sequences in the phylogenetic tree. *ArUGAE* was selected based on its high UGA5E ratio among the BLAST results. However, it is very likely that the presumed UGA5E from *Rhizobium meliloti*, which actually is an UGA4E, was used as a reference for annotation. Consequently, this would mean that UGA4E-specific features are assigned to the hypothesized UGA5E specificity. Accordingly, chances were high that *ArUGAE* displayed the UGA4E specificity. In addition, detailed analysis of the BLAST results of *ArUGAE* showed that *RmUGAE* was one of the outcomes. When a reverse reasoning is applied, this would mean that sequences with a low UGA5E ratio are more likely to actually display UGA5E specificity. Nevertheless, even though the ratio of *TgUGAE* was indeed low, the enzyme still displayed UGA4E activity. Taking into account the fact that both epimerases are located in different clades of the phylogenetic tree, it is hypothesized that the UGA4E specificity can be assigned to all of these sequences.

This hypothesis is supported by a more detailed sequence analysis. Sun et al. recently solved the first crystal structure of an UGA4E (i.e., from *Streptomyces viridosporus*, PDB 6KV9) and found that crucial interactions are formed with the substrate at positions G92 and R192 [20]. Moreover, R192 is believed to favor the sugar ring rotation that is needed for C4-epimerization, and its correct orientation would be ensured by the adjacent D194. The conservation of these residues, and their environment, in all of the UGA4Es characterized to date is displayed in Figure 6. Crucially, the sequence logo of the putative UGA5E sequences is found to be practically identical. In order to make a comparison to a related C5-epimerase, GM35E from the same CEP1 family was analyzed as well. In addition, the sequence logo of UDP-glucose 4-epimerase (Gal4E), the most studied specificity within the CEP1 family, was also included. As a totally different sequence logo was obtained for the latter two epimerases, it can be concluded that the residues are not simply conserved among all CEP1 members or even among enzymes with the same epimerization site. Accordingly, G92, R192, and D194 can probably serve as fingerprints for the UGA4E specificity. In addition, similar differences are observed in the YxxxK catalytic motif (Figure 6). These results indicate that the putative UGA5E sequences are probably just incorrectly annotated UGA4Es.



**Figure 6.** Overview of sequence logos. The conservation of various active site residues, including their environment, and the YxxxK motif were evaluated for UDP-glucuronate 4-epimerase (UGA4E), putative UDP-glucuronate 5-epimerase (UGA5E), GDP-mannose 3,5 epimerase (GM35E), and UDP-glucose 4-epimerase (Gal4E). Residues are numbered according to the UGA4E from *Streptomyces viridosporus*.

### 3. Discussion

Carbohydrate epimerases have recently come into focus as they are a valuable route toward rare sugar production. However, more research on this enzyme class is required in order to fully appreciate their potential. Even though 39 specificities have been discovered to date, in-depth information on these enzymes' physiological roles or catalytic mechanisms is often lacking. Therefore, efforts should be made to characterize representatives in order to fill in these knowledge gaps. Moreover, prominent mechanistic insights might lead to the rational engineering of these biocatalysts, potentially resulting in altered specificities, stabilities, or activities.

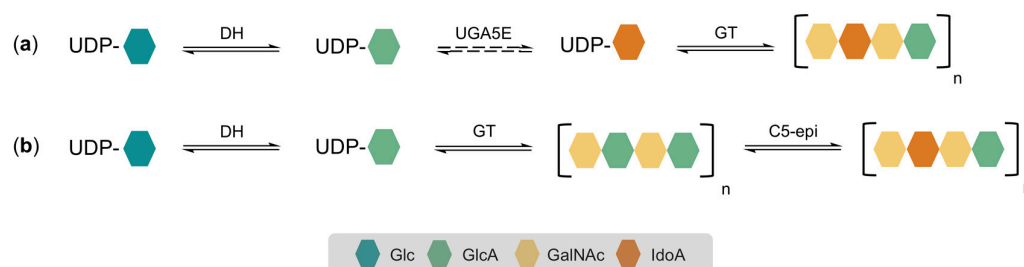
On that account, this study focused on the UGA5E specificity, hypothesized to catalyze the reversible epimerization of UDP-GlcA toward UDP-IdoA [9]. The enzyme's activity was only detected in rabbit tissue before, and there is no report on the recombinant expression or characterization of a putative UGA5E to date. This research clustered all 1154 putative UGA5E sequences from the UniProt database in a phylogenetic tree and the combination of experimental work and *in silico* analysis resulted in the hypothesis that these sequences actually display UGA4E specificity. Subsequently, the existence of the UGA5E specificity in general was questioned, resulting in a physiological issue. Indeed, it was previously hypothesized that UGA5E is part of the enzymatic pathway for the production of IdoA-containing DS starting from UDP-Glc (Figure 7a) [9]. To shed more light on this issue, pathways relevant for other eukaryotic and prokaryotic IdoA-containing polysaccharides, e.g., heparin sulfate, were examined [11,14,29]. These metabolic pathways start with the action of an UDP-Glc dehydrogenase, converting UDP-Glc into UDP-GlcA. However, this sugar acid is then first polymerized before the resulting polysaccharide undergoes modifications such as sulfation or epimerization. In other words, GlcA is, thus, epimerized to IdoA at the polysaccharide level instead of at the NDP-sugar level. For the bacterial GAGs structures and heparin/heparan sulfate, this epimerization step is catalyzed by an enzyme referred to as glucuronyl C5-epimerase (EC 5.1.3.17, CEP8) [30,31]. For dermatan and chondroitin sulfate synthesis, it was found that this reaction is realized by chondroitin-glucuronate C5-epimerase (EC 5.1.3.19, not classified in a CEP family) [32,33].

These results thus contradict the hypothesis of Jacobson et al., although the outcomes can be unified [9]. Indeed, the detected UDP-glucose dehydrogenase activity is anyway present as the first step in the pathway in order to obtain UDP-GlcA. Subsequently, the assumed UGA5E activity can be explained by the fact that they analyzed protein fractions and not purified proteins. The detected UGA5E activity in such a fraction was, thus, most probably the combined action of a glycosyltransferase that catalyzes the polymerization and chondroitin-glucuronate C5-epimerase,



resulting in IdoA subunits (Figure 7b). Subsequently, the polysaccharide can be further sulfated, eventually yielding DS. In general, the probability of the presence of the UGA5E specificity in nature is further decreased by the observation that no additional functions for free (UDP-)L-iduronic acid have been detected to date and no extra evidence on this epimerase specificity has been published since 1962.

In conclusion, this research states that the assumed UGA5E specificity present in sequence databases is not justified, and that the corresponding EC entry should be retracted. The epimerase specificity was initially studied as a potential route toward IdoA production, which can then be used for the synthesis of heparin. On that account, an enzymatic route comprising chondroitin-glucuronate C5-epimerase or glucuronyl C5-epimerase can still be exploited. In addition, this study unlocks a new pool of UGA4E sequences, which can be deployed for UDP-GalA synthesis, and has highlighted fingerprints that could be used for the correct annotation of sequences harboring UGA4E activity.



**Figure 7.** Production of dermatan sulfate. (a) Hypothesized pathway containing the putative UGA5E specificity versus (b) confirmed pathway on a polysaccharide level. Glc: glucose; GlcA: glucuronic acid; GalNAc: N-acetyl-D-galactosamine; IdoA: L-iduronic acid; DH: UDP-glucose dehydrogenase; UGA5E: UDP-glucuronate 5-epimerase; GT: glycosyltransferase; C5-epi: chondroitin-glucuronate C5-epimerase.

## 4. Materials and Methods

### 4.1. Materials

All chemicals were obtained from Sigma-Aldrich (Saint Louis, MO, USA) or Carbosynth (Compton, UK) and were of the highest purity, unless stated otherwise.

### 4.2. Gene Cloning and Transformation

The codon-optimized putative UGA5E genes from *Thermodesulfobacterium geofontis* (UniProt: F8C4X8) and *Agrobacterium radiobacter* strain K84 (UniProt: B9J8R3) were synthesized and subcloned into the pET21 vector at *Nde*I and *Xho*I restriction sites, providing a C-terminal His<sub>6</sub>-tag, by GeneArt Gene Synthesis (Thermo Fisher Scientific, Waltham, MA, USA). The constructs were transformed in *E. coli* BL21(DE3) electrocompetent cells for protein expression.

### 4.3. Enzyme Production

First, 250 mL lysogeny broth (LB) medium (10 g/L trypton, 5 g/L yeast extract and 5 g/L NaCl) supplemented with 100 µg/mL ampicillin was inoculated with 2% (v/v) stationary culture and incubated at 37 °C and 200 rpm until the OD<sub>600</sub> equaled 0.6. Subsequently, enzyme production was induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and the culture was incubated overnight at 20 °C and 200 rpm. Cells were harvested by centrifuging for 30 min at 9000 rpm and 4 °C. The obtained pellet was frozen and stored at −20 °C for at least 1 h.

In order to obtain crude cell lysate for the initial specificity tests, the pellet was resuspended in 8 mL of lysis buffer (10 mM imidazole and 100 µM phenylmethane sulfonyl fluoride (PMSF) in 100



mM MOPS pH 7.5) and cooled on ice for 30 min. Subsequently, the cells were further disrupted by sonication (3 times 2.5 min, Branson sonifier 250, level 3, 50% duty cycle).

On the other hand, purified enzyme was obtained by first lysing the cells by resuspending the pellet in 8 mL of lysis buffer (500 mM NaCl, 10 mM imidazole, 100  $\mu$ M PMSF, and 1 mg/mL lysozyme in 50 mM sodium phosphate buffer pH 7.4) and cooling them on ice for 30 min. Next, the cells were subjected to 3 times 2.5 min of sonication (Branson sonifier 250, level 3, 50% duty cycle). Cell debris was collected by 15 min of centrifugation at 12000 rpm and 4 °C. Next, the His<sub>6</sub>-tagged proteins present in the supernatant were purified to apparent homogeneity by nickel-nitrilotriacetic acid chromatography, with small variations to the supplier's description (Thermo Scientific, Waltham, MA, USA). The resin was washed twice with 8 mL wash buffer (500 mM NaCl, 20 mM imidazole in 50 mM sodium phosphate buffer pH 7.4). Protein was eluted with 3 times 4 mL elution buffer (500 mM NaCl, 250 mM imidazole in 50 mM sodium phosphate buffer pH 7.4). As a final step, buffer was exchanged to 100 mM MOPS pH 7.5 by using Amicon Ultra-15 centrifugal filter units with 30 kDa cut-off (Merck Millipore, Darmstadt, Germany). For TgUGAE, the purification yield was only 0.2 mg/L culture, whereas for ArUGAE, this was about 45 mg/L (Figures S7 and S8).

The protein content was determined by measuring the absorbance at 280 nm with the NanoDrop2000 Spectrophotometer (Thermo Scientific, Waltham, MA, USA). The extinction coefficient and molecular weight of His<sub>6</sub>-tagged enzymes were calculated using the ProtParam tool on the ExPASy server (<https://web.expasy.org/protparam/>). Molecular weight and purity of the protein were verified by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; 12% gel). As the expression of TgUGAE was low, the measured protein concentration was adjusted using the ImageJ software [34].

#### 4.4. Activity Testing

For the examination of the putative UGA5Es, 10 mM UDP-GlcA was incubated together with 100  $\mu$ L crude cell lysate in 100 mM MOPS pH 7.5 at 30 °C (500  $\mu$ L total reaction volume). A substrate control (10 mM UDP-GlcA in 100 mM MOPS pH 7.5), as well as an enzyme control (100  $\mu$ L crude cell lysate in 100 mM MOPS pH 7.5) were included and treated in the exact same way as the reaction mixture. Tests were performed in triplicate. In addition, the potential activity of the background proteins was tested by setting up a reaction with 10 mM UDP-GlcA and 100  $\mu$ L crude cell lysate obtained from induced *E. coli* BL21(DE3) cells containing an empty pET21 vector in 100 mM MOPS pH 7.5 at 30 °C. The influence of NAD<sup>+</sup> on the enzyme's specificity was tested by incubating 0.1 mg/mL enzyme with 10 mM UDP-GlcA and 5 mM NAD<sup>+</sup> in 100 mM MOPS pH 7.5 at 30 °C (400  $\mu$ L total reaction volume). The enzymes' specific activities were determined with 10 mM UDP-GlcA and 0.05 mg/mL purified enzyme in a total reaction volume of 300  $\mu$ L 100 mM MOPS pH 7.5 at 30 °C. For all activity tests, samples were taken at defined time points and inactivated in 90 mM acetic acid. The samples were then hydrolyzed for 1 h at 95 °C in order to obtain the respective monosaccharides. The hydrolyzed reaction mixtures were analyzed using high-performance anion exchange chromatography and pulsed amperometric detection (HPAEC-PAD; Dionex ICS-3000 system, Thermo Fisher Scientific, Waltham, MA, USA). The separation of the various uronic acids was realized by an isocratic method using a mixture of 150 mM NaOAc and 100 mM NaOH for 10 min.

#### 4.5. Sequence Analysis

All amino acid sequences annotated as 'UDP-glucuronate 5-epimerase' were retrieved from the UniProt database (<https://www.uniprot.org>) and subsequently aligned with Clustal Omega using default parameters (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) [35]. The concomitant phylogenetic tree was visualized with iTOL (<https://itol.embl.de>). BLAST analyses were performed at the UniProt server using default parameters. For the sequence logos, the sequences of the characterized UGA4Es, GM35Es, and Gal4Es were extracted from the UniProt database and aligned with Clustal Omega using default parameters. Subsequently, the sequence logos were generated by WebLogo [36].

**Supplementary Materials:** The following are available online at [www.mdpi.com/xxx/s1](http://www.mdpi.com/xxx/s1), Figure S1. Optimization of the recombinant expression of TgUGAE; Figure S2. Analysis of potential UGA4E activity in the *E. coli* background; Figure S3. Chromatographic analysis of the reaction controls for TgUGAE; Figure S4. Optimization of ArUGAE overexpression in *E. coli*; Figure S5. Exclusion of potential UGA4E activity in the *E. coli* background; Figure S6. Chromatographic analysis of the reaction controls for ArUGAE; Figure S7. SDS-PAGE analysis of TgUGAE purification; Figure S8. SDS-PAGE analysis of ArUGAE purification.

**Author Contributions:** All authors have read and agreed to the published version of the manuscript. Conceptualization, O.G., T.D., S.V.O., M.D.C. and K.B.; methodology, O.G.; validation, S.V.O., M.D.C., T.D. and K.B.; formal analysis, O.G.; investigation, O.G.; data curation, O.G.; writing—original draft preparation, O.G.; writing—review and editing, T.D., S.V.O., M.D.C. and K.B.; visualization, O.G.; supervision, T.D. and K.B.; project administration, O.G.

**Funding:** This research was funded by the Fund for Scientific Research Flanders (FWO-Vlaanderen) via a doctoral scholarship for O.G. (grant n° 1S16817N), S.V.O. (grant n° FWO.3F0.2016.0056.01) and M.D.C. (grant n° 1S03020N), as well as the EpiSwitch project (grant n° G0F3417N).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Samuel, J.; Tanner, M. Mechanistic aspects of enzymatic carbohydrate epimerization. *Nat. Prod. Rep.* **2002**, *19*, 261–277.
- Beerens, K.; Overtveldt, S.; Van Desmet, T. The “epimerring” highlights the potential of carbohydrate epimerases for rare sugar production. *Biocatal. Biotransform.* **2017**, *35*, 230–237.
- Van Overtveldt, S.; Verhaeghe, T.; Joosten, H.J.; van den Bergh, T.; Beerens, K.; Desmet, T. A structural classification of carbohydrate epimerases: From mechanistic insights to practical applications. *Biotechnol. Adv.* **2015**, *33*, 1814–1828.
- Yu, Z.; Schmaltz, R.M.; Bozeman, T.C.; Paul, R.; Rishel, M.J.; Tsosie, K.S.; Hecht, S.M. Selective tumor cell targeting by the disaccharide moiety of bleomycin. *J. Am. Chem. Soc.* **2013**, *135*, 2883–2886.
- Mathé, C.; Gosselin, G. L-Nucleoside enantiomers as antiviral drugs: A mini-review. *Antivir. Res.* **2006**, *71*, 276–281.
- Beerens, K.; Desmet, T.; Soetaert, W. Enzymes for the biocatalytic production of rare sugars. *J. Ind. Microbiol. Biotechnol.* **2012**, *39*, 823–834.
- Gevaert, O.; Van Overtveldt, S.; Beerens, K.; Desmet, T. Characterization of the first bacterial and thermostable GDP-mannose 3,5-epimerase. *Int. J. Mol. Sci.* **2019**, *20*, 3530.
- Major, L.L.; Wolucka, B.A.; Naismith, J.H. Structure and function of GDP-mannose-3',5'-epimerase: An enzyme which performs three chemical reactions at the same active site. *J. Am. Chem. Soc.* **2005**, *127*, 18309–18320.
- Jacobson, B.; Davidson, E.A. UDP-D-glucuronic acid-5-epimerase and UDP-N-acetylglucosamine-4-epimerase of rabbit skin. *BBA—Enzymol. Subj.* **1962**, *73*, 145–151.
- Damus, P.S.; Hicks, M.; Rosenberg, R.D. Anticoagulant action of heparin. *Nature* **1973**, *246*, 355–357.
- Rabenstein, D. Heparin and heparan sulfate: Structure and function. *Nat. Prod. Rep.* **2002**, *19*, 312–331.
- Liu, H.; Zhang, Z.; Linhardt, R.J. Lessons learned from the contamination of heparin. *Nat. Prod. Rep.* **2009**, *26*, 313–321.
- Dulaney, S.B.; Huang, X. Strategies in synthesis of heparin/heparan sulfate oligosaccharides: 2000-present. *Adv. Carbohydr. Chem. Biochem.* **2012**, *67*, 95–136.
- Sugahara, K.; Mikami, T.; Uyama, T.; Mizuguchi, S.; Nomura, K.; Kitagawa, H. Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. *Curr. Opin. Struct. Biol.* **2003**, *13*, 612–620.
- Trowbridge, J.M.; Gallo, R.L. Dermatan sulfate: New functions from an old glycosaminoglycan. *Glycobiology* **2002**, *12*, 117R–125R.
- Neufeld, E.F. UDP-D-galacturonic acid 4-epimerase from radish roots. *Methods Enzymol.* **1966**, *8*, 276–277.
- Dalessandro, G.; Northcote, D.H. Changes in enzymic activities of nucleoside diphosphate sugar interconversions during differentiation of cambium to xylem in pine and fir. *Biochem. J.* **1977**, *162*, 267–279.
- Regué, M.; Hita, B.; Pique, N.; Izquierdo, L.; Merino, S.; Fresno, S.; Benedi, V.J.; Tomás, J.M. A gene, *uge*, is essential for *Klebsiella pneumoniae* virulence. *Infect. Immun.* **2004**, *72*, 54–61.

19. Muñoz, R.; Lopez, R.; De Frutos, M.; García, E. First molecular characterization of a uridine diphosphate galacturonate 4-epimerase: An enzyme required for capsular biosynthesis in *Streptococcus pneumoniae* type 1. *Mol. Microbiol.* **1999**, *31*, 703–713.
20. Sun, H.; Ko, T.P.; Liu, W.; Liu, W.; Zheng, Y.; Chen, C.C.; Guo, R.T. Structure of an antibiotic-synthesizing UDP-glucuronate 4-epimerase MoeE5 in complex with substrate. *Biochem. Biophys. Res. Commun.* **2019**, *521*, 31–36.
21. Kereszt, A.; Kiss, E.; Reuhs, B.L.; Carlson, R.W.; Kondorosi, Á.; Putnoky, P. Novel *rkp* gene clusters of *Sinorhizobium meliloti* involved in capsular polysaccharide production and invasion of the symbiotic nodule: The *rkpK* gene encodes a UDP-glucose dehydrogenase. *J. Bacteriol.* **1998**, *180*, 5426–5431.
22. Bloom, J.D.; Labthavikul, S.T.; Otey, C.R.; Arnold, F.H. Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5869–5874.
23. Eijssink, V.G.H.; Gaseidnes, S.; Borchert, T.V.; van den Burg, B. Directed evolution of enzyme stability. *Biomol. Eng.* **2005**, *22*, 21–30.
24. Elkins, J.G.; Hamilton-Brehm, S.D.; Lucas, S.; Han, J.; Lapidus, A.; Cheng, J.F.; Goodwin, L.A.; Pitluck, S.; Peters, L.; Mikhailova, N.; et al. Complete genome sequence of the hyperthermophilic sulfate-reducing bacterium *Thermodesulfobacterium geofontis* OPF15T. *Genome Announc.* **2013**, *1*, 1–2.
25. Hattori, S.; Kamagata, Y.; Hanada, S.; Shoun, H. *Thermacetogenium phaeum* gen. nov., sp. nov., a strictly anaerobic, thermophilic, syntrophic acetate-oxidizing bacterium. *Int. J. Syst. Evol. Microbiol.* **2000**, 1601–1609.
26. Göker, M.; Daligault, H.; Mwirichia, R.; Lapidus, A.; Lucas, S.; Deshpande, S.; Pagani, I.; Tapia, R.; Cheng, J.F.; Goodwin, L.; et al. Complete genome sequence of the thermophilic sulfur-reducer *Desulfurobacterium thermolithotrophum* type strain (BSAT) from a deep-sea hydrothermal vent. *Stand. Genomic Sci.* **2011**, *5*, 407–415.
27. Radke-Mitchell, L.C.; Sandine, W.E. Influence of temperature on associative growth of *Streptococcus thermophilus* and *Lactobacillus bulgaricus*. *J. Dairy Sci.* **1986**, *69*, 2558–2568.
28. Kavanagh, K.L.; Jörmvall, H.; Persson, B.; Oppermann, U. The SDR superfamily: Functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell. Mol. Life Sci.* **2008**, *65*, 3895–3906.
29. Raedts, J.; Kengen, S.W.M.; Van Der Oost, J. Occurrence of L-iduronic acid and putative D-glucuronyl C5-epimerases in prokaryotes. *Glycoconj. J.* **2011**, *28*, 57–66.
30. Raedts, J.; Lundgren, M.; Kengen, S.W.M.; Li, J.P.; Van Der Oost, J. A novel bacterial enzyme with D-glucuronyl C5-epimerase activity. *J. Biol. Chem.* **2013**, *288*, 24332–24339.
31. Debarnot, C.; Monneau, Y.R.; Roig-Zamboni, V.; Delaunay, V.; Le Narvor, C.; Richard, E.; Hénault, J.; Goulet, A.; Fadel, F.; Vivès, R.R.; et al. Substrate binding mode and catalytic mechanism of human heparan sulfate D-glucuronyl C5 epimerase. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 6760–6765.
32. Pacheco, B.; Maccarana, M.; Goodlett, D.R.; Malmström, A.; Malmström, L. Identification of the active site of DS-epimerase 1 and requirement of N-glycosylation for enzyme function. *J. Biol. Chem.* **2009**, *284*, 1741–1747.
33. Maccarana, M.; Olander, B.; Malmström, J.; Tiedemann, K.; Aebersold, R.; Lindahl, U.; Li, J.P.; Malmström, A. Biosynthesis of dermatan sulfate: Chondroitin-glucuronate C5-epimerase is identical to SART2. *J. Biol. Chem.* **2006**, *281*, 11560–11568.
34. Schindelin, J.; Rueden, C.T.; Hiner, M.C.; Eliceiri, K.W. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol. Reprod. Dev.* **2015**, *82*, 518–529.
35. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; Thompson, J.D.; Higgins, D.G.; McWilliam, H.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 1–6.
36. Crooks, G.E.; Hon, G.; Chandonia, J.-M.; Brenner, S.E. WebLogo: A sequence logo generator. *Genome Res.* **2004**, *14*, 1188–1190.

