Chinese Chemical Letters 35 (2024) 109028

Contents lists available at ScienceDirect

Chinese Chemical Letters

journal homepage: www.elsevier.com/locate/cclet

A multi-enzymatic cascade reaction for the synthesis of bioactive *C*-oligosaccharides

Kebo Xie[1,*], Qian Zhang[1], Fei Ye, Jungui Dai[*]

State Key Laboratory of Bioactive Substance and Function of Natural Medicines, CAMS Key Laboratory of Enzyme and Biocatalysis of Natural Drugs and NHC Key Laboratory of Biosynthesis of Natural Products, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China

ARTICLE INFO

Article history:
Received 17 July 2023
Revised 30 August 2023
Accepted 31 August 2023
Available online 3 September 2023

Keywords:
C-Oligosaccharides
C-Glycosyltransferase
Cyclomaltodextrin glucanotransferase
Glucoamylase
Enzymatic synthesis
SGLT2 inhibitory activity

ABSTRACT

C-Oligosaccharides are rare in nature and possess diverse bioactivities. However, their chemical synthesis faces many challenges. In this work, enzymatic introduction of *C*-linked sugar chains to target aglycones was successfully achieved by multi-enzymatic cascade reactions. A *C*-glycosyltransferase from *Aloe barbadensis* was employed to introduce the first *C*-linked glucose and then a cyclomaltodextrin glucanotransferase from *Bacillus licheniformis* was used to extend the sugar chain. A total of twenty *C*-oligosaccharides with 2–6 sugars were synthesized from scale-up reactions and exhibited good water solubility and sodium-dependent glucose transporter 2 (SGLT2) inhibitory activity. Furthermore, a glucoamylase was used to control the length of the sugar chain and the *C*-maltosides were efficiently synthesized. These findings not only expanded the structural diversity of *C*-oligosaccharides, but also provided a strategy for the modification of *C*-glycoside drugs to improve the druggability.

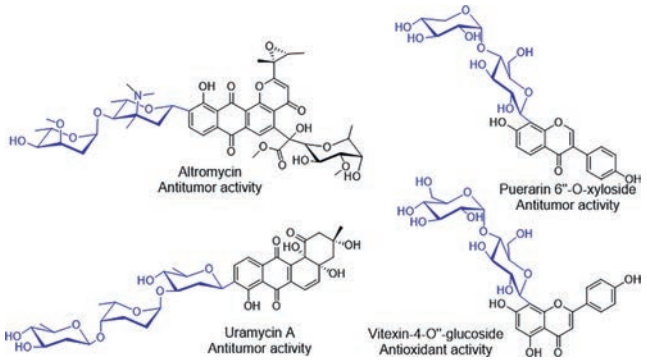
The aryl *C*-glycoside structure is a distinct motif among the plenty of biologically active natural products. *C*-Glycosides have attracted considerable interest due to their potential bioactivities, stable glycosidic bonds and synthetic challenges [1,2]. Most of the natural *C*-glycosides are *C*-monosaccharides which often face the challenge of insufficient druggability due to the poor solubility. *C*-Oligosaccharides are *O*-linked sugar derivatives of *C*-monosaccharides at the primary *C*-sugar moiety, and usually exhibit good water solubility and pharmacological activity (Scheme 1) [2–6]. Nonetheless, *C*-oligosaccharides are rare in nature and chemical synthesis of oligosaccharides faces challenges such as regio- and stereo-selectivity and protection and deprotection of functional groups, which limit their application in drug discovery [7–9]. Therefore, the shortage of *C*-oligosaccharides with different skeletons greatly limited the research of this type of compounds. It is necessary to develop an approach that can efficiently and economically generate diverse types of *C*-oligosaccharides.

Combinatorial biosynthesis is considered to be an effective approach in synthesizing target compounds with complex structures. According to the structures of *C*-oligosaccharides, enzymes are needed to introduce the first *C*-sugar moiety, extend the length of *C*-sugar moiety and control the length of *C*-oligosaccharides to a designed length. Therefore, to construct the combinatorial biosynthesis approach of *C*-oligosaccharides, in this work, a promiscuous *C*-glycosyltransferase (CGT), a cyclomaltodextrin glucanotransferase (CGTase) and a glucoamylase were combined together. This tandem reaction achieved the efficient synthesis of diverse bioactive new *C*-oligosaccharides.

In our previous work, we reported a new *C*-glycosyltransferase AbCGT which could introduce *C*-sugar moiety to the aglycon **1** to synthesize a sodium-dependent glucose transporter 2 (SGLT2) inhibitor **1a** with the potential in curing diabetes (Fig. 1) [10]. However, the aglycon part brought *C*-monosaccharide **1a** a poor water solubility, which limited its development in further drug discovery. The aglycon part of **1a** has been previously optimized to enhance the bioactivity and in this work, we aimed to modify the sugar part to obtain *C*-oligosaccharides with better druggability. Firstly, we attempt to extend the length of *C*-sugar moiety to improve the water solubility and maintain or enhance the bioactivity. Thus, the *C*-glycosyltransferase (AbCGT) combined with a cyclomaltodextrin glucanotransferase (CGTase) [1,4-α-D-glucan: 4-α-D-(1,4-α-D-glucano)-transferase (cyclizing), EC 2.4.1.19] from *Bacillus licheniformis* were employed to introduce the first *C*-sugar and further extend its length, respectively [11,12]. Recombinant AbCGT and CGTase were expressed in *E. coli* and purified to >90%

* Corresponding authors.
E-mail addresses: keboxie@imm.ac.cn (K. Xie), jgdai@imm.ac.cn (J. Dai).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.cclet.2023.109028
1001-8417/



Scheme 1. Representative natural *C*-oligosaccharides.

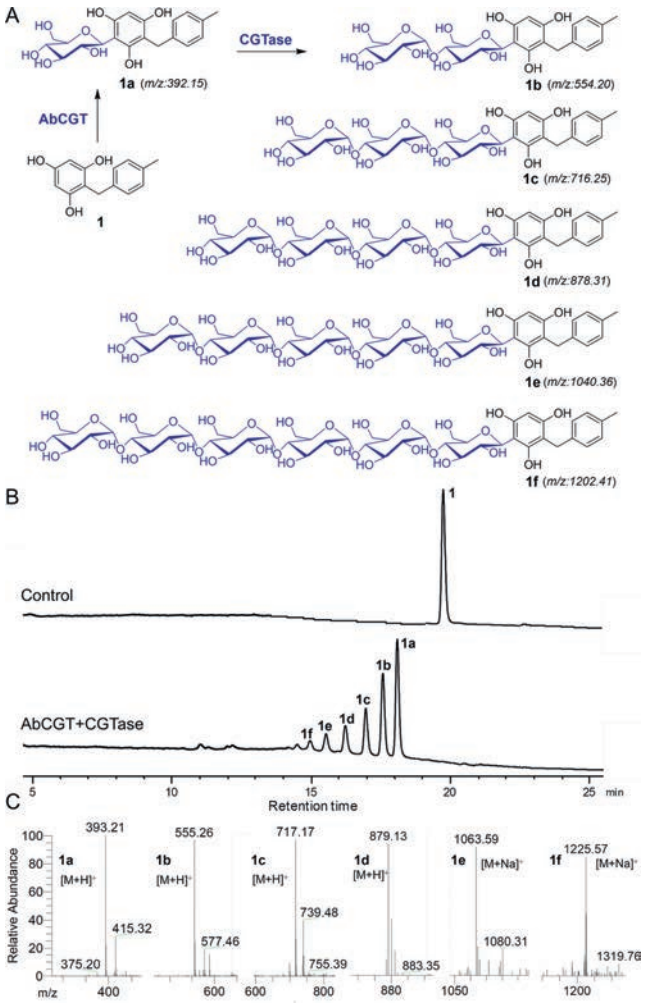


Fig. 1. Combining AbCGT and CGTase to produce *C*-oligosaccharides with different lengths of *C*-sugar chains. (A) Tandem reactions catalyzed by AbCGT and CGTase *in vitro*. UDP-Glc and β-CD were used as sugar donors. (B, C) HPLC-UV/MS analysis of the reactions.

homogeneity (Fig. S1 in Supporting information). Reaction mixture containing substrate **1**, UDP-α-D-glucose (UDPG), β-cyclodextrin (β-CD), AbCGT and CGTase was incubated at 30 °C for 12 h. Intriguingly, according to the results of high performance liquid chromatography-ultraviolet/mass spectroscopy (HPLC-UV/MS) analysis of the reaction mixture, combinatorial enzyme-catalyzing system successfully synthesized *C*-oligosaccharides with 2–6 glucosyl substituents (**1b**–**1f**) (Fig. 1). According to the results, the aglycon **1** was firstly *C*-glycosylated with a 100% conversion rate by AbCGT and then the *C*-glycosylated product **1a** was subsequently *O*-glycosylated with 66% conversion rate for one to five times generating the corresponding *C*-oligosaccharides with two to six sugars. As the sugars were added one by one, the yields of *C*-oligosaccharides with different numbers of sugars decreased with the increasing numbers of sugars. Although CGTase has been used for synthesis of bioactive natural products in some cases [12–17], the application of CGTase in the modification of unnatural products and clinical drugs has not been reported.

To further apply the CGTase in the modification of structurally diverse *C*-monosaccharides and synthesis of target *C*-oligosaccharides, different types of *C*-monosaccharides including bioactive unnatural products (**1a**–**5a**), natural products (**6**–**9**) and clinical drugs (**10**–**15**) were selected as substrates (Fig. 2). Substrates **1a**–**5a** are SGLT2 inhibitors, substrates **6**–**9** are bioactive natural products and substrates **10**–**15** are famous diabetic drugs. For some substrates, the poor water solubility greatly limited the development of these compounds in further drug discovery [18–22]. For example, puerarin (**9**), an isoflavone glycoside derived from *Pueraria lobata* (Willd.) Ohwi, has been identified as a pharmacologically active component with diverse benefits [23]. However, puerarin (**9**) cannot be given by injection due to its low solubility in water [18]. In order to modify these bioactive *C*-glycosides and improve their water solubility and bioavailability, sugars were attempted to be introduced. However, the modification of *C*-linked sugar is rarely reported and difficult to be achieved. What is more, as for the clinical drugs dapagliflozin (**10**), ipragliflozin (**11**), empagliflozin (**12**), canagliflozin (**13**), ertugliflozin (**14**) and sotagliflozin (**15**), there are no proper glycosylated sites except the *C*-linked sugar. Therefore, the CGTase, which possesses a large substrate binding pocket might be a proper catalyst in the extension of diverse *C*-linked sugars. Surprisingly, recombinant CGTase was able to recognize all the prepared *C*-glycosides and generated *C*-oligosaccharides with α-1,4-glycosidic bonds and the sugar numbers of two to six. With this approach, twenty *C*-oligosaccharides (**1b**, **1c**, **7a**, **7b**, **9a**, **10a**, **10b**, **10c**, **10d**, **11a**, **11b**, **12a**, **12b**, **13a**, **13b**, **14a**, **14b**, **14c**, **15a** and **15b**) were enzymatically synthesized from the scale-up reactions and the structures were confirmed by MS and nuclear magnetic resonance (NMR) (Fig. 2 and Figs. S2–S47 in Supporting information). Among these twenty *C*-oligosaccharides, nineteen (**1b**, **1c**, **7a**, **7b**, **10a**, **10b**, **10c**, **10d**, **11a**, **11b**, **12a**, **12b**, **13a**, **13b**, **14a**, **14b**, **14c**, **15a** and **15b**) are new compounds.

The SGLT2 inhibitory activity of the derivatives of compounds **1** and **10**–**15** was investigated. Interestingly, all the prepared *C*-oligosaccharides exhibited SGLT2 inhibitory activity. With the sugar number increasing, the inhibitory activity of *C*-oligosaccharides was slightly decreased except the derivative (**14c**) of ertugliflozin (**14**) (Table S1 in Supporting information). The architecture of the SGLT2 protein binding with empagliflozin (**12**) has been obtained and the inhibitory mechanism of SGLT2 inhibitors was also elucidated [24]. Empagliflozin (**12**) occupies both the substrate sugar-binding site and the external vestibule of SGLT2 resulting in blocking the function of SGLT2. Therefore, to explore the inhibitory mechanism of the *C*-oligosaccharides with various sugars such as derivatives (**12a** and **12b**) of empagliflozin (**12**), molecular docking was performed (Fig. S48 in Supporting information). *C*-Oligosaccharides **12a** and **12b** showed similar binding mode with **12**. However, the different lengths of sugar-chains could lead to the differences in the substrates binding with SGLT2, which might affect the inhibitory activity of the *C*-oligosaccharides with various sugars. Reasonably, the water solubility of *C*-oligosaccharides is greatly improved due to the introduction of hydrophilic groups. The water solubility of the representative *C*-oligosaccharides was further investigated and the water solubilities of *C*-oligosaccharides (**9a** and **10a**) of puerarin and dapagliflozin were increased by 32 and 57 times (Table S2 in Supporting information). Therefore,

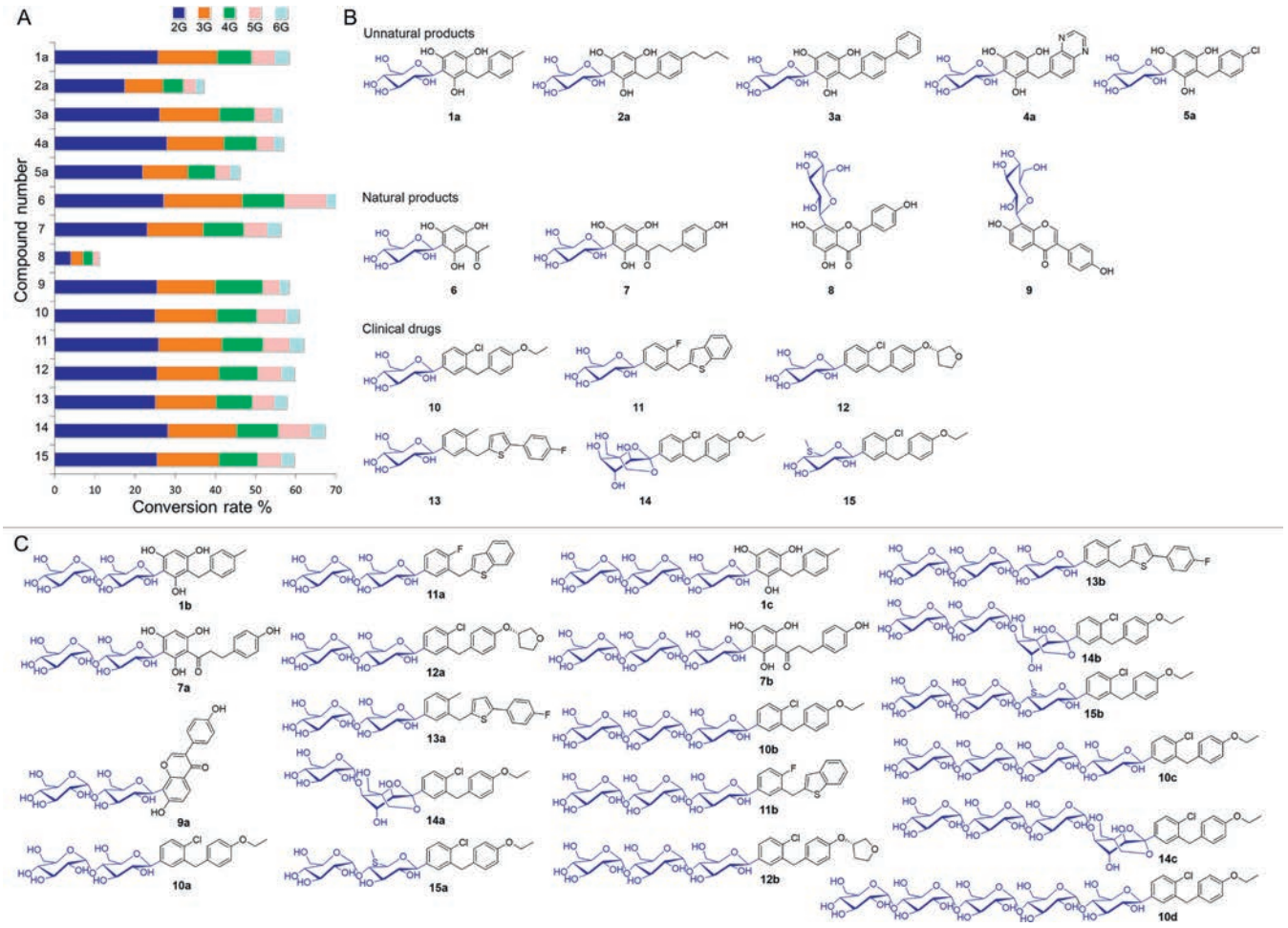


Fig. 2. The application of promiscuous CGTase in synthesizing diverse *C*-oligosaccharides. (A) The conversion rates and the product distributions of the CGTase catalyzing reactions with the substrates shown in B. 2G–6G represent the sugar numbers of the glycosylated products are 2–6, respectively. (B, C) The structures of the substrates and the isolated products of the CGTase catalyzing reactions.

comparing with the aglycons, the *C*-sugar chain improved both SGLT2 inhibitory activity and water solubility.

In this tandem enzyme reaction, UDPG is expensive in the market and cannot be used as easily available raw material and enzymes in this mixture need to be purified, which brings great inconvenience and high cost. Thus, to get rid of the limitation of high cost and complex protein purification process, an engineered *E. coli* strain harboring its native biosynthetic pathway of UDPG and pCDFDuet-*AbCGT*-*CGTa* was developed as a catalyst. As shown in Fig. 3, the whole-cell catalytic system successfully produced *C*-monoglycoside **1a** and *C*-maltoside **1b** with the total conversion rate of only 51%. Then to enhance the catalytic efficiency of whole-cell catalyst for large-scale preparation, the system including the cell density, substrate concentrations, and incubation time was optimized. This engineered *E. coli* exhibited the highest conversion rate at the OD_{600} of 6.0 (Fig. 3A). When the mole ratio of sugar donor (β-CD) and acceptor (**1**) reached five, the conversion rate only slightly increased (Fig. 3B). According to the time course assays, five hours is an efficient incubation time (Fig. 3C). Therefore, at the optimized conditions, the conversion rate of acceptor **1** was greatly improved and reached more than 90% (Fig. 3D). Thus, an efficient and economical method for generating *C*-oligosaccharides was developed and optimized, which will be further applied for the scale-up reactions to produce diverse bioactive *C*-oligosaccharides due to the high promiscuity of AbCGT and CGTase.

The successful establishment of the low-cost combinatorial whole-cell method brings tremendous potential to the application of the enzymatic tandem reactions in industrial production. It not only solves the problem of insufficient structural diversity of *C*-oligosaccharides, but also has important significance for efficient, large-scale, low-cost production of *C*-oligosaccharides, greatly promoting the practical application of combinatorial enzyme catalysts in industrial production. In addition, the whole-cell catalysis method showed great optimization potential. We believe that the yield and productivity of target molecules could be greatly improved, through further systemic optimization and the use of scaled-up fermentation for cultivation.

In the catalysis, the CGTase seems to keep adding glycosyl moieties to its previous products as long as there are enough initial substrates and sugar donors. Although the introduction of sugars could greatly improve the water solubility, the bioactivity of some *C*-oligosaccharides is gradually decreased with the increasing number of sugars (Table S1). Therefore, there is a balance point between bioactivity and water solubility, which means that the length of the sugar chain should be under control. In order to control the extension of *C*-oligosaccharide chains, commercially available glucoamylase was employed.

Glucoamylase, which exhibits high hydrolytic specificity toward α-(1→4)-glucosidic linkages, is an important industrial enzyme and is widely used in starch saccharification, brewing and distilling industry [25,26]. In this present case, glucoamylase was applied to hydrolyze redundant sugars in *C*-oligosaccharides to produce *C*-maltosides, which can keep the balance between maintaining bioactivity and enhancing water solubility.

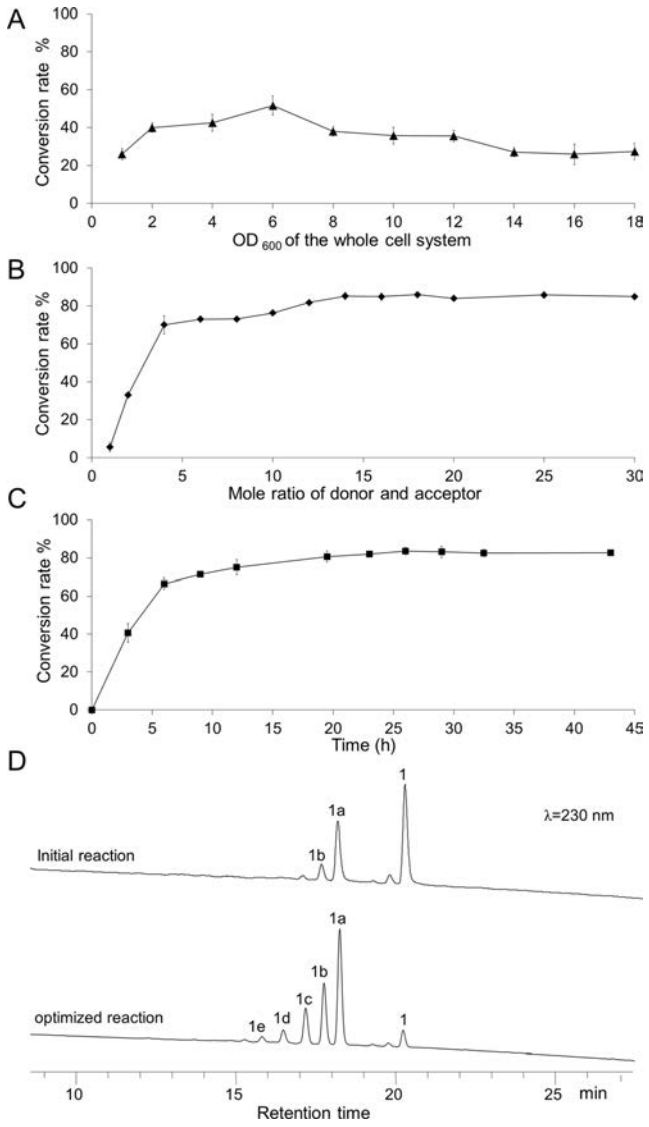


Fig. 3. Constructing and optimizing the whole cell-*AbCGT-CGTa* catalyzed reactions. Optimization of OD_{600} of the cultured cells (A), mole ratio of donor and acceptor (B), and reaction time (C). (D) HPLC analysis of the initial reaction and optimized reaction of the whole cell-*AbCGT-CGTa* catalytic reactions.

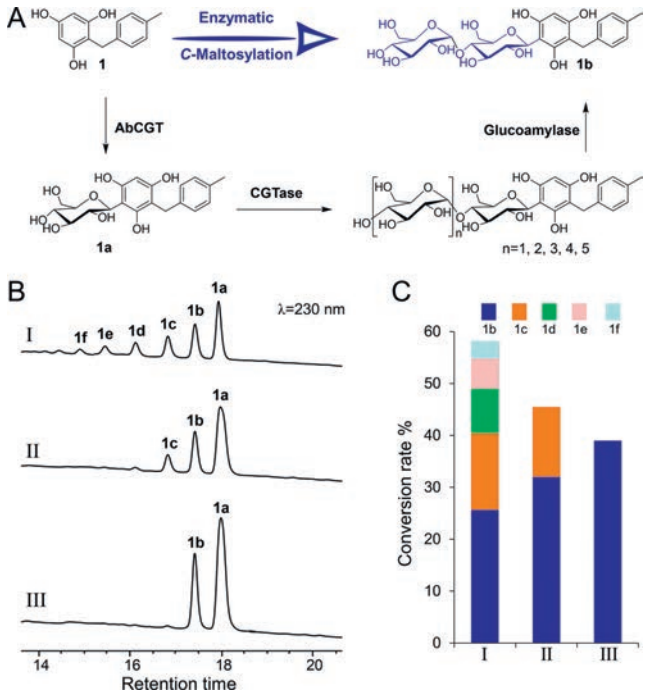


Fig. 4. Enzymatic *C*-maltosylation system. (A) Combining AbCGT, CGTase and Glucoamylase to produce target *C*-maltosides **1b**. (B) HPLC analysis of reactions catalyzed by AbCGT and CGTase (I), AbCGT, CGTase and glucoamylase with the concentrations of 1 U/μL (II) and 1.5 U/μL (III), respectively. (C) The conversion rates of the *C*-oligosaccharides with different lengths of sugar chains corresponding to the reaction I, II and III.

Therefore, after the glycosylation reaction catalyzed by CGTase, different concentrations of the glucoamylase were added to the reaction mixtures and incubated for 10 min. According to the HPLC analysis as shown in Fig. 4, when 1 U/μL of glucoamylase was added to the reaction mixtures, the *C*-oligosaccharides with sugar numbers of four (**1d**), five (**1e**), and six (**1f**) were transformed to those with sugar numbers of one (**1a**), two (**1b**) and three (**1c**). And when the concentration of glucoamylase was increased to 1.5 U/μL, all the glycosylated products were transformed and the *C*-maltoside (**1b**) was generated with the conversion rate of 41%. So, with the glycosylation enzyme CGTase and the deglycosylation enzyme glucoamylase and their proper concentrations and reaction times, an enzymatic *C*-maltosylation system was constructed. In addition, the glucoamylase showed broad substrate spectra and was also able to catalyze all the other *C*-oligosaccharides with different structures. Given the substrate promiscuity of these two enzymes, this enzymatic *C*-maltosylation system could be broadly applied in natural and unnatural products and even drug modification.

In summary, enzymatic introduction of *C*-linked sugar chains to target aglycones was successfully achieved by tandem reactions. A *C*-glycosyltransferase was employed to introduce the first *C*-linked glucose and then a CGTase was used to extend the sugar chain. With this method, *C*-oligosaccharides were introduced to bioactive natural products, SGLT2 inhibitors and clinical anti-diabetes drugs to improve their water solubility and druggability. With this method, nineteen representative new *C*-oligosaccharides with different skeletal structures were successfully generated and some of the derivatives exhibited better druggability. Finally, a glucoamylase was also successfully used to control the length of the sugar chain and target *C*-maltosides were efficiently synthesized. The control of the length of the sugar chain under accurate numbers is vital for the further application of this method. These findings expanded the diversity of *C*-oligosaccharides and provided a strategy for modifying *C*-glycoside drugs to improve their druggability. In further study, engineering of the glucoamylase in molecular level to produce mutants with the ability of accurately hydrolyzing *C*-oligosaccharides will be very necessary. Our work will be helpful to solve the problem of the shortage of bioactive *C*-oligosaccharides, which can be further used in drug discovery.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is financially supported by the National Key Research and Development Program of China (No. 2020YFA0908000), CAMS Innovation fund for Medical Sciences (No. 2021-I2M-1–029).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cclet.2023.109028.

References

[1] K. Kitamura, Y. Ando, T. Matsumoto, K. Suzuki, Chem. Rev. 118 (2018) 1495–1598.
[2] T. Bililign, B.R. Griffith, J.S. Thorson, Nat. Prod. Rep. 22 (2005) 742–760.
[3] M. Dou, X. He, Y. Sun, et al., Chin. Chem. Lett. 26 (2015) 255–258.
[4] X. Ying, R. Wang, J. Xu, et al., J. Chromatogr. Sci. 47 (2009) 201–205.
[5] A. Benmerache, A.A. Magid, A. Kabouche, et al., Nat. Prod. Res. 34 (2020) 2887–2893.
[6] A. Trefzer, D. Hoffmeister, E. Künzel, et al., Chem. Biol. 7 (2000) 133–142.
[7] K. Shimoda, D. Sato, H. Hamada, Chem. Lett. 38 (2009) 930–931.
[8] Z. Hou, J. Wang, X. Zhang, et al., Chin. Chem. Lett. 34 (2023) 107804.
[9] Q. Zhao, S. Zhou, Y. Wang, et al., Chin. Chem. Lett. 34 (2023) 107982.
[10] K. Xie, X. Zhang, S. Sui, et al., Nat. Commun. 11 (2020) 5162.
[11] D.E. Hill, R. Aldape, J.D. Rozzell, Nucleic Acids Res. 18 (1990) 199.
[12] W. Huang, C. Wen, Z. Zhou, et al., Adv. Synth. Catal. 361 (2019) 3114–3119.
[13] N. Khummanee, P. Rudeekulthamrong, J. Kaulpiboon, Appl. Biochem. Microbiol. 55 (2019) 360–370.
[14] D. Svensson, S. Ulvenlund, P. Adlercreutz, Biotechnol. Bioeng. 104 (2009) 854–861.
[15] M.Y. Rather, K.Z.G. Ara, E.N. Karlsson, P. Adlercreutz, Process Biochem. 50 (2015) 722–728.
[16] P. Aramsangtienchai, W. Chavasiri, K. Ito, P. Pongsawasdi, J. Mol. Catal. B: Enzym. 73 (2011) 27–34.
[17] W. Huang, Q. He, Z. Zhou, et al., ACS Omega 5 (2020) 12251–12258.
[18] D. Li, S.H. Park, J.H. Shim, et al., Carbohydr. Res. 339 (2004) 2789–2797.
[19] N.C. Patel, H.A. Patel, Folia Med. 64 (2022) 488–500.
[20] I. El-Bagory, N.K. Alruwaili, M.H. Elkomy, et al., J. Drug Deliv. Sci. Technol. 54 (2019) 101279.
[21] Q. Li, Q. Cao, Z. Yuan, et al., Nanomedicine 42 (2022) 102534.
[22] H. Sims, K.H. Smith, P. Bramlage, J. Minguetet, Diabet. Med. 35 (2018) 1037–1048.
[23] Y. Zhou, H. Zhang, C. Peng, Phytother. Res. 28 (2014) 961–975.
[24] Y. Niu, R. Liu, C. Guan, et al., Nature 601 (2022) 280–284.
[25] C.R. Soccol, P.J. Rojan, A.K. Patel, A. Pandey, C. Webb, C.R. Soccol, C. Larroche, et al., Glucoamylase, in: Enzyme Technology, Springer-Verlag, New York, 2006, pp. 221–237.
[26] A. Pandey, Starch 47 (1995) 439–445.