Substrate specificity of Chondroitinase ABC I based on analyses of biochemical reactions and crystal structures in complex with disaccharides

Makoto Takashima², Ippei Watanabe¹³, Akimasa Miyanaga², and Tadashi Eguchi²

² Department of Chemistry, Tokyo Institute of Technology, 2-12-1 O-okayama, Meguro-ku, Tokyo, 152-8551, Japan.
³ Medical Affairs, Seikagaku Corporation, 1-6-1 Marunouchi, Chiyoda-ku, Tokyo 100-0005, Japan.

¹To whom correspondence should be addressed: Tel: (81)3-5220-8595; Fax: (81)3-5220-8975; e-mail: ippei.watanabe@seikagaku.co.jp
Running title: Substrate specificity of Chondroitinase ABC I

Supplementary data included: Table S1, Table SII, Fig. S1 and Fig. S2

Keywords: chondroitinase ABC I / chondroitin sulfate / crystal structure / kinetic analysis / substrate specificity

Abstract

Chondroitinase ABC I (cABC-I) is the enzyme which cleaves the β-1,4 glycosidic linkage of chondroitin sulfate (CS) by β-elimination. To elucidate more accurately the substrate specificity of cABC-I, we evaluated the kinetic parameters of cABC-I and its reactivity with CS isomers displaying less structural heterogeneity as substrates, e.g., approximately 90% of disaccharide units in Chondroitin sulfate A (CSA) or Chondroitin sulfate C (CSC) is D-glucuronic acid and 4-O-sulfated N-acetyl galactosamine (GalNAc) (A-unit) or D-glucuronic acid and 6-O-sulfated GalNAc (C-unit), respectively. cABC-I showed the highest reactivity to CSA and CSC among all CS isomers, and the $k_{cat}/K_m$ of cABC-I was higher for CSA than for CSC. Next, we determined the crystal structures of cABC-I in complex with CS disaccharides, and analyzed the crystallographic data in combination with molecular docking data. Arg500 interacts with 4-O-sulfated and 6-O-sulfated GalNAc residues. The distance between Arg500 and the 4-O-sulfate group was 0.8 Å shorter than that between Arg500 and the 6-O-sulfated group. Moreover, it is likely that the 6-O-sulfated group is electrostatically repulsed by the nearby Asp490. Thus, we demonstrated that cABC-I has the highest affinity for the CSA richest in 4-O-sulfated...
GalNAc residues among all CS isomers. Recently, cABC-I was used to treat lumbar disc herniation. The results provide useful information to understand the mechanism of the pharmacological action of cABC-I.

Introduction

As linear polysaccharides composed of repeating disaccharide units of uronic acid and hexosamine, the glycosaminoglycans (GAGs) are classified into, e.g., chondroitin sulfate (CS), dermatan sulfate (DS), hyaluronan (HA), heparan sulfate (HS), and keratin sulfate (KS). CS is a sulfated polysaccharide composed of repeating of D-glucuronic acid (GlcA) and N-acetyl galactosamine (GalNAc) with a β1-3 linkage. It is widely distributed in human or animal tissues and associated with biological and pathological phenomena (Mikami and Kitagawa 2013; Raman et al. 2005; Linhardt and Toida 2004; Gama et al. 2006). CS is classified into six major isomers based on the positions of sulfate groups as follows: chondroitin 4-O-sulfate (CSA), chondroitin 6-O-sulfate (CSC), chondroitin 2,6-O-sulfate (CSD), chondroitin 4,6-O-sulfate (CSE), and non-sulfated CS (CH). DS also known as CSB, is an epimeric form of CSA and contains L-iduronic acid (IdoA) instead of GlcA.

Chondroitin sulfate lyase (CSase) is the enzyme which cleaves the β1-4 glycosidic linkage of GAG polysaccharides by β-elimination. According to its substrate specificity, CSase is mainly categorized into three types: Chondroitinase ABC (cABC), Chondroitinase AC (cAC), and Chondroitinase B. cABC has been used as a GAG-degrading enzyme for more than 30 years (Yoshida et al. 1989; Murata and Yokoyama. 1985a). Two cABC enzymes with different substrate reactivity, cABC-I and cABC-II, have been identified in Proteus vulgaris (Yamagata et al. 1968, Hamai et al. 1997). cABC-I (Hernicore®) was developed as a therapeutic agent for lumbar disc herniation (LDH), which has been used in Japan since 2018. The nucleus pulposus (NP) at the center of lumbar disc contains proteoglycan consisting of CS or KS (Olczyk K. 1993), which forms a gel-like structure with water-retention properties. A bulging NP displaced towards the posterior annulus fibrosus compresses the spinal and nerve roots, causing LDH with various symptoms, e.g., back pain and leg pain. By degradation of CS, cABC-I directly injected into the NP causes reduction in disc pressure and herniation volume, and relieves the symptoms associated with LDH (Sugimura et al. 1996; Takahashi et al. 1996; Yamada et al. 2001).

A 997-amino acid residue monomeric protein, cABC-I, is an endolytic enzyme with broad reactivity to CS isomers as substrates. Although previous studies have shown that cABC-I is highly reactive with CSC, CSA, and DS, these CS molecules are derived from animals and contain considerable amounts of various disaccharides that are characteristic of other CS isomers (Suzuki et al. 1968; Volpi 2007). For example, CSC derived from shark cartilage has approximately 70% of C-unit [GlcA β1-3 GalNAc (6S)] and 20% of A-unit [GlcA β1-3 GalNAc (4S)] and other disaccharide units. Moreover, the molecular weight (Mw) is different for each CS isomer. To evaluate the substrate specificity of cABC-I more accurately, we speculated that it is necessary to use CS molecules with less structural heterogeneity.
Enzyme-substrate complex crystal structure and mutagenesis studies can provide useful information on the function of the amino acid residues in enzymes and their interaction with substrates. However, the crystal structure of cABC-I in complex with its substrates has never been determined. Previous crystallographic and mutagenesis studies of cABC-I alone have demonstrated that the cleft region of this enzyme has six amino acid residues (Arg500, His501, Tyr508, Arg560, His561, Glu653) involved in the reaction with substrates (Prabhakar et al 2005a; Prabhakar et al 2005b, Kawaguchi et al 2013). However, the substrate recognition mechanism of cABC-I described above is only a hypothetical model and is not based on insight gained from study of the crystal structure of cABC-I in complex with its substrates. It is thus still unclear which amino acid residues are involved and how each of them works to determine substrate specificity.

In the present study, reactivity and kinetic parameters of cABC-I were estimated using CS isomers with less structural heterogeneity as substrates. Further, we succeeded in structural determination of three complexes of cABC-I-CS disaccharide units designated: A-unit, C-unit, and non-sulfated unit. In addition, we performed docking analysis of CSA or CSC tetrasaccharide. These data reveal the structural basis of the substrate specificity of cABC-I, which is useful information to understand the mechanism of the pharmacological action of cABC-I in the NP.

Results

Reactivity of cABC-I with CS isomers and non-sulfated GAGs

We used GAGs shown in Table I to evaluate the substrate reactivity of cABC-I. No evidence has yet indicated a relationship between the Mw of CS and the reactivity of cABC-I. To investigate the influence of the Mw of CS on the reactivity of cABC-I, firstly, we used CSCs of Mw: 8 kDa, 14 kDa, and 64 kDa. As shown in Fig. 1A, the reaction of these CSs with cABC-I proceeded over time. The smaller the Mw of the CSC, the higher was the reactivity. It suggested that in order to investigate substrate reactivity using various CS isomers, the substrates should have the same Mw. Next, we evaluated the reactivity of cABC-I toward five CS isomers adjusted to a Mw of almost 25 kDa, and non-sulfated GAGs as substrates: CSA, DS, CSC, CSD, CSE, CH, and HA (Fig. 1B). cABC-I had a high reactivity to CSA and CSC compared with other CS isomers, and barely recognized the non-sulfated GAGs CH and HA (Fig. 1B and D). Both CS oligosaccharides, CSA and CSC, with a Mw of 1 kDa showed almost the same reactivity as those with a Mw of 25 kDa described above (Fig. 1C). The reactivity of DS at 60 minutes was about one-third of that of CSA and CSC. Although the reactivity of CSD and CSE was similar to that of DS up to 30 minutes, the reaction then almost stopped. In the quantification of generated disaccharides units of each CS isomer by enzymatic reaction, the D-unit \([\text{GlcA}\beta1-3 \text{GalNAc}(2,6-O\text{-disulfate})]\) and E-unit \([\text{GlcA}\beta1-3 \text{GalNAc}(4,6-O\text{-disulfate})]\) mostly were not detected (Fig. 2).

Moreover, we extended the reaction time to 24 hours, and evaluated the reactivities of cABC-I to CSA-26K, CSC-25K, and homogeneous CSA (hCSA) and CSC (hCSC) (Fig. 1E and F). The reaction of cABC-I with CSA-26K or CSC-25K almost stopped in about 6 hours after the start of the reaction,
and the OD values of CSA-26K and CSC-25K at 24 hours was 2.4 and 1.8, respectively. Generated disaccharide of each CS substrate also reached a plateau in about 6 hours and the amounts of generated disaccharides derived from CSA tended to be more than that of CSC (Fig. 2F and G), which correlated well with the profile of the absorbance of reaction mixture of each CS. Although there was no significant difference between the reactivities of cABC-I to CSA-26K and CSC-25K in the first 60 minutes (Fig. 1B), after extending the reaction time to 24 hours, it became clear that cABC-I had a higher reactivity with CSA-26K compared with CSC-25K. We used hCSA and hCSC as CS molecules with less structural heterogeneity, containing approximately 90% ΔDi-4S or ΔDi-6S, respectively (Table I). As shown in Fig. 1F, the OD values of hCSA and hCSC at 24 hours were 2.6 and 2.3, respectively. The reactivity of cABC-I was dramatically improved for hCSC compared with that for CSC-25K. For generated disaccharide of each hCS, ΔDi-4S was generated more than ΔDi-6S over the entire measurement point (Fig. 2H). The reaction of cABC-I with hCSA began to slow down after 12 hours, but unlike CSA-25K, the OD value continued to rise gradually.

Kinetic analysis of cABC-I with CSA and CSC

The $k_{\text{cat}}/K_m$ for CSA and CSC showed high values in the range of pH 7.0-7.5 (Fig. 3A and B). The $k_{\text{cat}}/K_m$ peak was reached at around pH 7.25, and the values started to decrease when the pH exceeded 7.5, respectively. It suggested that the optimum pH range for cABC-I is from 7 to less than 8, which is close to the body’s own pH environment. As shown in Fig. 3, irrespective of Mw, CSA showed higher $k_{\text{cat}}/K_m$ compared to CSC as follows. The $k_{\text{cat}}/K_m$ for CSA-26K was slightly higher than that for CSC-25K. The $k_{\text{cat}}/K_m$ for CS oligosaccharides (with Mw of 1K) also exhibited pH dependence similar to that for CS with Mw of 25K, and the $k_{\text{cat}}/K_m$ for CSA-1K was slightly higher than that for CSC-1K. The $k_{\text{cat}}/K_m$ for CS oligosaccharides was approximately one-fifth that for CS-25K, indicating that the reaction rate constant varied depending on Mw (Fig. 3B). As shown in Table II, the $k_{\text{cat}}/K_m$ for hCSA and hCSC at pH 7.25 was 964 ($\mu$M$^{-1}$ sec$^{-1}$) and 599 ($\mu$M$^{-1}$ sec$^{-1}$), respectively. The $k_{\text{cat}}/K_m$ for each of the hCSs was almost 2-3 times higher than that for CSA-26K and CSC-25K.

Overall structures of cABC-I complexes with disaccharides

The structures of ligand-free cABC-I and cABC-I in complex with disaccharides (ΔDi-4S, ΔDi-6S, or ΔDi-0S), the final product of the enzymatic reaction, have been determined at 1.88 Å, 1.80 Å, 1.92 Å and 2.50 Å resolutions, respectively (Fig. 4 and Table S1). These disaccharides were bound to cABC-I at the active site cleft, specifically at the amino acid residues including Arg500, His501, Tyr508, and Arg560. The overall structures of cABC-I bound to these disaccharides were almost identical to those of the unbound form (Ca r.m.s.d.: cABC-I-ΔDi-4S complex, 0.131 Å; cABC-I-ΔDi-6S complex, 0.161 Å; cABC-I-ΔDi-0S complex, 0.217 Å), indicating that cABC-I was not conformationally changed by interaction with disaccharides. According to the electron density analysis, the number of disaccharides bound to a single molecule of cABC-I was as follows: ΔDi-4S, 3 molecules; ΔDi-6S, 2 molecules; ΔDi-0S, 2 molecules (Supplementary data, Fig. S1).
**Interaction of cABC-I with ΔDi-4S**

Three molecules of ΔDi-4S were bound to the active site cleft of cABC-I, as shown in Fig. 5A. The disaccharide bound to the center of the active site cleft is labeled as ΔDi-4S-c, and the disaccharides located to the left and right of ΔDi-4S-c are labeled as ΔDi-4S-l and -r, respectively. The 4-0-sulfate group in GalNAc of ΔDi-4S-c forms a salt bridge (3.1 Å) with the side chain of Arg500. The interactions between each functional group of ΔGlcA and the amino acid residues of cABC-I involved in the reaction are as follows: the O5 oxygen interacts with the side chain of His501 and Tyr508; the carboxy group interacts with the side chain of His501 and Arg500; the 2-hydroxyl group interacts with the side chain of Arg560. The interactions between the functional groups of ΔDi-4S (-r or -l), and amino acid residues are summarized in Table III.

**Interaction of cABC-I with ΔDi-6S**

Two molecules of ΔDi-6S (-c or -l) were bound to the active site cleft of cABC-I, as shown in Fig. 5B, and the binding position of each disaccharide was similar to that of ΔDi-4S (-c or -l). The 6-0-sulfate group in GalNAc of ΔDi-6S-c forms a salt bridge (3.9 Å) with the side chain of Arg500. The length of the salt bridge with the 6-0-sulfate group is 0.8 Å longer than that with the 4-0-sulfate group. In addition, the 6-0-sulfate group also interacts with the main chain of His501. However, the 6-0-sulfate group is located close to the side chain of Asp490 (3.1 Å), suggesting the existence of electrostatic repulsion. The interactions between each functional group of ΔGlcA and the amino acid residues are almost identical to that of ΔDi-4S-c. The interactions between the functional groups of ΔDi-6S-l and amino acid residues are summarized in Table III.

**Interaction of cABC-I with ΔDi-0S**

Two molecules of ΔDi-0S (-c, -l) were bound to the active site cleft of cABC-I, as shown in Fig. 5C, and the binding position of each disaccharide was similar to that of ΔDi-4S (-c and -l). The GalNAc moiety in ΔDi-0S-c does not interact with any amino acid residues, except for hydrogen bond between the amide nitrogen and the side chain of Tyr655. The interactions between each functional group of ΔGlcA and the amino acid residues are almost identical to those of ΔDi-4S-c and ΔDi-6S-c. The interactions between the functional groups of ΔDi-0S-l and amino acid residues are summarized in Table III.

**Docking analysis**

The result of docking analysis of CSA tetrasaccharide (subsites −2, −1, +1, and +2) is shown in Fig. 6A. The position of the GalNAc moiety of ΔDi-4S-c in the complex structure was well superimposed on subsite +2, suggesting that the binding positions of GalNAc in both analyses were almost identical. These results suggested that the binding positions of GalNAc in subsite −1 could be described more precisely based on the data obtained in the docking study. However, the position of the GlcA moiety...
(subsite +1) differed by approximately 30° compared with that of ΔGlcA of ΔDi-4S-c (Supplementary data, Fig. S2A). This difference is probably due to the fact that ΔGlcA of ΔDi-4S-c has a double bond. The position of GalNAc and ΔGlcA moieties of ΔDi-4S-c were greatly deviated from subsites −1 and −2, respectively, based on the comparison of the binding positions between the crystal structure and docking model. Docking analysis showed that the 4-O-sulfate group of GalNAc (subsite −1) interacts with the side chain of His561.

The result of docking analysis of CSC tetrasaccharide (subsites −2, −1, +1, and +2) is shown in Fig. 6B. As in the result of docking analysis using CSA tetrasaccharide described above, the binding positions of GalNAc in ΔDi-6S-c and subsite +2 well agreed with each other, but the position of GlcA moiety (subsite +1) was slightly different compared with the position of ΔGlcA of ΔDi-4S-c. The position of the GalNAc and ΔGlcA moieties of ΔDi-6S-c were greatly deviated from subsites −1 and −2, respectively, in combined crystal structure and molecular docking analyses (Supplementary data, Fig. S2B). Docking analysis showed that the 6-O-sulfate group of GalNAc (subsite −1) interacts with the side chain of Arg560 and His561; however, the distance to Asp658 was short enough (2.7 Å) to possibly cause electrostatic interaction.

Discussion

Many papers have reported the substrate reactivity of cABC-I, particularly its high reactivity with CSA and CSC (Hamai et al. 1997; Prabhakar et al. 2005b). Various CS isomers derived from animals have a wide range of Mw and the disaccharide composition of individual CS molecules can vary. We hypothesized that differences in CS Mw and disaccharide composition affect the reactivity of cABC-I even when the CS is the same isomer. That is, we thought that in order to evaluate the substrate specificity of cABC-I more accurately, we needed to use CSs that are more homogeneous in structure. When CS is cleaved by cABC-I, an unsaturated bond is generated between the C4 and C5 positions of uronic acid at the nonreducing terminus and shows an ultraviolet absorption maximum of around 230 nm. Following previous test methods (Prabhakar et al. 2005b), we evaluated the reactivity and enzyme kinetics of cABC-I with various CS isomers using a method based on measurement of absorbance at 232 nm.

The reactivity of cABC-I with CS isomer adjusted to a Mw of almost 25 kDa was as follows: CSA > CSC > DS > CSE ≈ CSD (Fig. 1B and E). As shown in Table I, CSC contains 14% disulfated disaccharides. We speculated that cABC-I is unlikely to react with disulfated disaccharides because the reaction of cABC-I with CSE and CSD reached a plateau in about 30 minutes, and no Di-(4,6)S and Di-(2,6)S disulfated disaccharides contained in CSE or CSD were detected (Fig. 1B, Fig. 2D and E). To remove the influence of disulfated disaccharides, we evaluated the reactivity with hCSA and hCSC as CS molecules with less structural heterogeneity, containing approximately 90% of ΔDi-4S or ΔDi-6S, respectively. Considering the reactivity of cABC-I with CS isomers as shown in Fig. 1 and Fig. 2, CSA containing A-unit were the most reactive with cABC-I.

However, reactivity alone is not sufficient to investigate enzyme-substrate affinity. Generally, the
kinetic parameter $k_{cat}/K_m$ is used as an indicator of enzyme-substrate specificity (Sárkány et al. 2001). In order to evaluate detailed substrate specificity of cABC-I, we created pH- $k_{cat}/K_m$ profiles for CSA and CSC by Mw (Fig. 3). In general, reactivity between a protein and a polysaccharide depends on the Mw of the polysaccharide: the higher the Mw is, the higher the affinity with the protein is (Watanabe et al. 2015). Therefore, we presumed that also in this study, the higher the Mw of CS was, the higher the affinity of cABC-I for the CS substrate was (Fig. 3). Meanwhile, the reactivity increased with decreasing Mw of CS (Fig. 1A), and this suggests higher $k_{cat}$ values. In fact, the $k_{cat}$ of the CS-1kDa was found to be higher than that of the CS-25kDa. The amount (molar number) of enzymatic degradation product per unit time reportedly increased with decreasing Mw of substrate in polysaccharide enzymolysis (Zhang et al. 2001). The cABC-I reaction speed (Vmax) was also reported to increase with the decreasing Mw of CS, suggesting higher $k_{cat}$ values. Judging from these facts, CS isomers of lower Mw were considered to have higher reaction speeds and lower enzyme affinity than those of higher Mw. The kinetic parameters of cABC-I should be linked to properties of CS, e.g., Mw and disaccharide composition. Based on these findings, we evaluated the kinetic parameters at pH 7.25 for hCSs with an adjusted Mw below 30 kDa. As shown in Table II, the $k_{cat}/K_m$ for hCSA was almost 1.6 times higher than that for hCSC. It indicated that the specificity of cABC-I is the highest for A-unit, which is abundant in CSA. Moreover, the $k_{cat}/K_m$ for each hCS was almost 2-3 times higher than that for CSs with greater structural heterogeneity. Structural heterogeneity in CS molecules such as those with disulfated disaccharides may be inferred from the reactivity of cABC-I with A-unit or C-unit.

Structural analysis of complex with substrates is a useful method to understand the details of enzyme-substrate specificity. Although the crystal structure of ligand-free cABC-I has been determined in previous study (Huang et al. 2003), but the crystal structure of cABC-I in complex with substrate or with CS disaccharide, has not been reported. In general, it is difficult to obtain complex with bound substrates such as polysaccharides once enzymatic degradation of the substrate has taken place. In fact, we attempted to soak cABC-I crystals with CS oligosaccharides, but the determined cABC-I structure contained disaccharides, the final product of the enzymatic reaction, in the active site cleft (data not shown). We have successfully determined the crystal structure of cABC-I in complex with the CS disaccharides, ΔDi-4S, ΔDi-6S and ΔDi-0S by soaking cABC-I crystals with these CS disaccharides. As shown in Fig. 5, two or three molecules of each disaccharide were bound to cABC-I at the cleft region, where the amino acid residues (Arg500, His501, Tyr508, His561 and Arg560) exist. A previous analysis of the crystal structure of hyaluronidase in complex with ΔDi-HA indicated that two disaccharide molecules are bound to the cleft region at the “substrate recognition area” and “product release area”, respectively (Ponnuraj and Jedrzejas 2000). Similarly, in cABC-I, ΔDi-X-$c$ may bind to the former and ΔDi-X-$l$ to the latter, playing different roles.

The CS disaccharide used for crystal structural analysis has ΔGlcA with a double bond, which is not present in the undigested polysaccharide CS. As reported in the crystal structure of hyaluronidase in complex with ΔDi-HA (Ponnuraj and Jedrzejas 2000), the presence or absence of the double bond in ΔGlcA may cause a slight shift in the position of the interaction between the functional groups and
the amino acid residues in CS. Meanwhile, the complex with the final product does not mimic the binding of the enzyme to the substrate during the reaction. In order to identify the binding site more precisely, we performed a docking analysis of CSA and CSC tetrasaccharides, both of which do not have the double bonds, and compared the results with those from the complex structure analysis using the unsaturated disaccharide. Previous studies have demonstrated that mutation of the following amino acid residues: Arg500, His501, Tyr508, and Arg560, significantly reduces or abolishes the activity of cABC-I (Table SII). Considering (i), (ii), and (iii) below, the result of the present docking analysis was consistent with that of the previous mutagenesis study, suggesting the usefulness of the analytical method in this study; (i) the distance between the carboxy group of GlcA (subsite +1) and Arg500 is approximately 3 Å, (ii) the distance between the C5 of GlcA (subsite +1) and His501 is approximately 3 Å, (iii) the distance between the oxygen of the glycosidic bond and Tyr508 and Arg560 is 3 to 4 Å. The docking study made it possible to analyze the interactions between sulfate groups of GalNAc (subsite −1) and amino acid residues.

In this study, we demonstrated for the first time that Arg500 interacts with the 4-O- and 6-O-sulfate groups of GalNAc bound to the center of the active site cleft (Fig. 5). The result of docking analysis also suggested that the side chain of His561 interacts with the 4-O-sulfate group, and the side chains of Arg560 and His561 interact with the 6-O-sulfate group (subsite −1, Fig. 6). The amino acid residues of cABC-I presumably recognize GalNAc in a different manner depending on whether the sulfate groups of GalNAc are in position 4 or 6. Furthermore, the affinity of cABC-I for 4-O-sulfate groups was suggested to be higher than that for the 6-O-sulfate groups. The basis for this inference is that: (i) the distance between Arg500 and each sulfate group in GalNAc was 0.8 Å shorter for the 4-O-sulfate group than for the 6-O-sulfate group (Fig. 5), and (ii) the 6-O-sulfate group of GalNAc (subsite +2 or −1) was close to acidic amino acid residues such as Asp490 and Asp658, which may result in electrostatic repulsion (Fig. 6). This electrostatic repulsion may contribute to keeping Arg500 and the 6-O-sulfate groups away from each other. Our results thus suggested that cABC-I has the highest affinity for the A-unit-rich CSA molecule among the CS isomers.

The roles of these amino acid residues based on the mutagenesis studies are proposed to be as follows: (i) His561 interacts with the sulfate group, (ii) Arg500 plays a role in the charge neutralization of the carboxy group, (iii) His501 abstracts the proton from C5, and (iv) Tyr508 and Arg560 donate a proton to the oxygen of the glycosidic bond (Table SII). Hence, the catalytic mechanisms were estimated on the basis of a comprehensive evaluation of published data, and our results for the demonstrated positional relationships between amino acids and substrate functional groups are shown in Fig. 7. Arg500 and His561 mutations were reported to reduce the substrate reactivity of cABC-I but, unlike other amino acid mutations, they did not completely deactivate cABC-I (Table SII). Hence, the two amino acid residues were suggested to be important to the recognition of the sulfate group of GalNAc by CS, and not to be involved directly in the cleavage of the CS chain, like other amino acid residues involved in the reaction. The reaction between cABC-I and the substrate was suggested to begin with the interaction of the sulfate group of CS with Arg500 and His561, and to proceed to the
above-described reactions (iii) and (iv).

The information on the substrate specificity of cABC-I obtained from this study may be useful in understanding the mode of action of the enzyme, but this study has several limitations. The reactivity for the substrate of cABC-I in in vitro experiments does not completely reproduce real activity in the NP because of differences between the laboratory environment and the actual bio-environment (Fig. 1 and 2). To assess quantitatively the pharmacodynamics of cABC-I in the living body, ex vivo testing by using NP is needed. CS molecules had structures that were resistant to cABC-I attack according to Sugahara et al. (1994). Such structures may have affected the reactivity of cABC-I with CS, and resulted in stopping reaction as shown in Fig. 1E and F. Evaluating the reactivity of cABC-I to CS, including the enzyme’s resistant structures in the NP, would provide information on the real activity of cABC-I in the living body. Further studies are needed to elucidate the detailed mode of action of the enzyme.

This study is the first to report the substrate specificity of cABC-I, based on the reactivity with CS of low structural heterogeneity and the structural analysis of complex with disaccharides. cABC-I is used as a therapeutic agent for LDH. We mainly used CSA and CSC to evaluate the substrate specificity of this enzyme, because A- and C-units are the dominant forms of GalNAc residues in the NP (Olczyk K. 1993). The in vitro test results from the present study would serve as a basis for understanding the pharmacokinetics of cABC-I in the living body.

Materials and methods
Glycosaminoglycans

GAGs used in this study are summarized in Table I. CSA from a whale cartilage was purchased from CarboSynth Ltd. (Berks, UK), DS from a chicken’s cockscomb, CSC from shark cartilage, CSD from shark cartilage, CSE from squid cartilage, HA from a chicken’s cockscomb, and CH from recombinant E. coli were obtained from Seikagaku Corp. (Tokyo, Japan). Homogeneous CSA (hCSA) and CSC (hCSC) prepared from CH according to the method of Sugiura et al. (2011). GAGs, except for CSA-26K and DS, were hydrolyzed with hydrochloric acid to reduce the Mw below 30 kDa. The average Mw and disaccharide composition of GAGs were estimated by gel exclusion column chromatography according to the method of Watanabe et al. (2015). The average Mw of CSA-1K and CSC-1K were estimated by LC-MS. The LC-MS system used for this study consisted of an LC-20AD Prominance liquid chromatograph connected to a LC-MS-IT-TOF (Shimadzu Corporation, Japan). Instrument control, data acquisition, and processing were accomplished using Shimadzu LCMS solution version 3.80. Briefly, the analysis was performed on a YMC Diol-300 column (8.0 mm × 500 mm) with 10 mM ammonium formate/MeOH (90:10, v/v) as a mobile phase at a flow rate of 0.5 mL/min at 25°C, and in negative ion mode polarity. The MS recording range was 200–1200 in MS mode and 75–2000 in MS/MS mode. Other MS parameters are shown below: detector voltage; 1.5 kV, CDL; 300°C, heat block; 200°C, nebulizer gas; 1.5 L/min.
Unsaturated disaccharides for crystal structural analysis

Unsaturated disaccharides, ΔDi-4S, ΔDi-6S and ΔDi-0S were prepared according to the method of Murata et al (1985b; 1986). Each disaccharide composition was more than 99%.

Enzyme purification

Extraction and purification of cABC-I was performed according to the method of Hamai et al (1997). Briefly, Proteus vulgaris NCTC 4636 was cultured at 30°C in the medium containing peptone, meat extract, yeast extract, and NaCl with CSC as the inducer. The cells were harvested by centrifugation at the end of logarithmic phase. Then cABC-I was extracted using polyoxyethylene lauryl ether and then purified by ion-exchange resin.

Reactivity of cABC-I with CS isomers and non-sulfated GAGs

Reactivity of cABC-I was measured by absorbance at 232 nm as described by Yamagata et al (1968). Exactly 1.0 mL of 0.05% substrate in 50 mM Tris-HCl Buffer was mixed with 2.0 mU of cABC-I and incubated for 60 min or for 24 hour at 37°C (n = 3). Time-dependent absorbance at 232 nm of the mixture was measured using a spectrophotometer (UV-1900i, Shimadzu Corporation). Moreover, each disaccharide generated (µg/mL) by the enzymatic reaction within 60 min was quantified according to the method of Watanabe et al (2015).

Kinetic analysis of cABC-I with CSA and CSC

The kinetic parameters of cABC-I reacting with CSA-26K, CSC-25K, CSA-1K, and CSC-1K were calculated based on the Michaelis-Menten equation. To investigate the pH dependence of the kinetic parameters, each substrate was dissolved in 50 mM Tris-HCl Buffer adjusted to pH 7.0, 7.25, 7.5, 7.75, and 8.0, respectively (n = 3). The kinetic parameters of cABC-I reacting with hCSA and hCSC at pH 7.25 were calculated by the same procedure. The amount of CS (µM) in each solution was measured by a carbazole-sulfuric acid method (Dische 1947). Protein concentrations were determined by the Lowry method (Lowry et al. 1951) using bovine serum albumin as the standard. cABC-I was mixed with each substrate solution, and the absorbance at 232 nm (P) was measured several times (t) at 37°C. The value of the initial velocity (v₀) was calculated by using the following equation [1].

\[ v_0 = \frac{d[P]}{dt} = \frac{(P_1 - P_0)}{(t_1 - t_0)} \cdot \cdot \cdot [1] \]

The 1/v₀ was plotted versus the reciprocal of substrate concentration (1/[S]) for five different concentrations and the data were fitted to a straight line. The maximum rate (Vₘₐₓ) and Michaelis-Menten constant (Kₘ) were extracted from the slope and y-intercept of the Lineweaver-Burk plot generated by plotting 1/v₀ and 1/[S] according to equation [2].

\[ \frac{[S]}{v_0} = \frac{1}{V_{\text{max}}} \cdot \frac{[S]}{[S]} + \frac{K_{md}}{V_{\text{max}}} \cdot \cdot \cdot [2] \]

Turnover number (k₉) was calculated by dividing Vₘₐₓ by the concentration of enzyme [E]₀.
\[ k_{cat} = \frac{V_{max}}{[E]_0} \cdots [3] \]

**Crystallization, data collection and structural determination**

Crystals of cABC-I were grown from a 1:1 mixture of a protein solution (7.5 mg/mL in 20 mM Tris-HCl [pH 7.4]) and a reservoir solution (200 mM magnesium acetate, 200 mM ammonium acetate, 17.5% [vol/vol] polyethylene glycol 3350, and 10 mM HEPES-Na [pH 7.5]) by using the sitting-drop vapor-diffusion method at 20°C. Crystals were then soaked in cryoprotectant solution (200 mM magnesium acetate, 200 mM ammonium acetate, 30% [vol/vol] polyethylene glycol 3350, and 10 mM Tris-HCl [pH 7.5]) and flash-frozen in a liquid nitrogen stream. Crystals of cABC-I in complex with ΔDi-4S, ΔDi-6S or ΔDi-0S were prepared by soaking ligand-free cABC-I crystals in cryoprotectant solution containing 100 mM of these disaccharide for 30–90 min and were then flash-frozen in a liquid nitrogen stream. X-ray data were collected by using beamlines AR-NW12A and BL5A (Photon Factory, Tsukuba, Japan). All diffraction data were indexed, integrated, and scaled by using the program XDS (Kabsch W 2010). The initial phases were determined by the molecular replacement method using the program MOLREP (Vagin et al. 2010), with the previously reported structure of cABC-I (PDB ID 1HNO; Huang et al. 2003) used as a search model. Coot (Emsley and Cowtan 2004) was used for visual inspection and manual rebuilding of the model. Refmac (Murshudov et al. 1997) was used for refinement. The figures were prepared by using PyMOL. Data-collection and processing statistics are summarized in supplemental material (Table S1). The atomic coordinates and structure factors have been deposited in the Protein Data Bank: PDB entry 7EIP, ligand-free cABC-I; 7EIQ, cABC-I in complex with ΔDi-4S; 7EIR, cABC-I in complex with ΔDi-6S; 7EIS, cABC-I in complex with ΔDi-0S.

**Docking analysis**

The docking study was carried out by using the AutoDock program (Version 4.2; Morris et al. 2009). The CSA tetrasaccharide and CSC tetrasaccharide molecules were generated by using the PRODRG2 Server (Schüttelkopf et al. 2004). The crystal structure of ligand-free cABC-I was used for the docking study. All of the water molecules were removed for the docking study. By using AutoDockTools, polar hydrogen atoms were added to amino acid residues, and Gasteiger charges were assigned to all atoms of the protein. The sugar rings of the ligand molecule were fixed in the 4C1 conformation, but the other rotatable bonds were set to be flexible for the calculation. All of the protein residues were kept rigid. The cubic energy grid was centered at the substrate binding pocket and had an extension of 40 Å in each direction. A total of 256 docking runs were performed with the Lamarckian genetic algorithm.

**Funding**

No funding to declare.

**Acknowledgements**
The authors are grateful to Kiyoshi Suzuki (Seikagaku Corp.), Yuniko Shibata (Seikagaku Corp.), Hirotaka Hoshi (Seikagaku Corp.) and Hideyuki Futatsumori (Seikagaku Corp.) for skillful technical support in the handling of the enzyme and helpful discussion. We also thank Taichi Chisuga (Tokyo Institute of Technology) for helpful discussion. This work was performed with the approval of the Photon Factory Program Advisory Committee (Proposal 2020G533).

Conflicts of interest
The authors declare no conflicts of interest.

Abbreviations

cABC, Chondroitinase ABC; cABC-I, Chondroitinase ABC I; cAC, Chondroitinase AC; CS, chondroitin sulfate; CSase, Chondroitin sulfate lyase; CSA, Chondroitin sulfate A; CSC, Chondroitin sulfate C; DS, Dermatan sulfate; GAG, Glycosaminoglycans; GalNac, N-acetyl-D-galactosamine; GlcA, D-glucuronic acid; HA, Hylaruronan; homogeneous CSA, hCSA; homogeneous CSC, hCSC; HS, Heparan sulfate; IdoA, L-iduronic acid; KS, keratin sulfate; LDH, Lumbar disc herniation; Mw, Molecular weight; NP, Nucleus pulposus.

References

Dische Z. 1947. A new specific color reaction of hexuronic acids. *J Biol Chem*. 167(1):189–198.

Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60(12):2126–2132.

Gama CI, Tully SE, Sotogaku N, Clark PM, Vaidehi N, Goddard 3rd, Nishi A, Hsieh-Wilson LC. 2006. Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nature Chemical Biology*. 2(9):467–473.

Huang W, Lunin VV, Li Y, Suzuki S, Sugiura N, Miyazono H, Cygler M. 2003. Crystal structure of Proteus vulgaris chondroitin sulfate ABC lyase I at 1.9 Å resolution. *J. Mol. Biol.* 328(3):623-634.

Kabsch W. 2010 XDS. *Acta Crystallogr Sect. D: Biol. Crystallogr.* 66(2):125–132.

Kawaguchi Y, Sugiura N, Kimata K, Kimura M, Kakuta Y. 2013. The crystal structure of novel chondroitin lyase ODV-E66, a baculovirus envelope protein. *FEBS Letters* 587(24): 3943–3948.

Linhardt RJ, Toida T. 2004. Role of Glycosaminoglycans in cellular communication. *Acc. Chem. Res.* 37(7):431–438.
Lowry O H, Rosebrough N J, Farr A L, Randall R J. 1951. Protein measurement with the Folin phenol reagent. *J Biol Chem.* 193(1):265–275.

Mikami T, Kitagawa H. 2013. Biosynthesis and function of chondroitin sulfate. *Biochim Biophys Acta - Gen Subj.* 1830(10):4719–4733.

Morris GM, Huey R, Lindstrom W, Sanner FM, Belew RK, Goodsell DS, Olson AJ. 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.* 30(16), 2785–2791.

Murata K, Yokoyama Y. 1985a. Enzymatic analysis with chondrosulfatases of constituent disaccharides of sulfated chondroitin sulfate and dermatan sulfate isomers by high-performance liquid chromatography. *Anal Biochem.* 149(1):261–268.

Murata K, Yokoyama Y. 1985b. A high-performance liquid chromatography for constituent disaccharides of chondroitin sulfate and dermatan sulfate isomers. *Anal Biochem.* 146(2):327–335.

Murata K, Yokoyama Y. 1986. Analysis of hyaluronic acid and chondroitin by high-performance liquid chromatography of the constituent disaccharide units. *J Chromatogr.* 374:37–44.

Murshudov G N, Vagin A A, Dodson E J. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53(3):240–255.

Olczyk K. 1993. Age-related changes in glycosaminoglycans of human intervertebral discs. *Folia Histochem Cytobiol.* 31(4):215-220.

Ponnuraj K, Jedrzejas MJ. 2000. Mechanism of Hyaluronan Binding and Degradation: Structure of Streptococcus pneumoniae Hyaluronate Lyase in Complex with Hyaluronic Acid Disaccharide at 1.7 Å Resolution. *J. Mol. Biol.* 299(4):885-895.

Prabhakar V, Capila I, Bosques CJ, Pujasek K, Sasisekharan R. 2005a. Chondroitinase ABC I from Proteus vulgaris: cloning, recombinant expression and active site identification. *Biochemical J.* 386 (1): 103-112.

Prabhakar V, Raman R, Capila I, Bosques CJ, Pujasek K, Sasisekharan R. 2005b. Biochemical characterization of the chondroitinase ABC I active site. *Biochemical J.* 390 (2):395-405.

Raman R, Sasisekharan V, Sasisekharan R. 2005. Structural Insights into Biological Roles of Protein-
Glycosaminoglycan Interactions. *Chemistry & Biology.* 12(3):267-277.

Sárkány Z, Szeltner Z, Sárkány LP. 2001. Thiolate-Imidazolium Ion Pair Is Not an Obligatory Catalytic Entity of Cysteine Peptidases: The Active Site of Picornain 3C. *Biochemistry.* 40(35):10601–10606.

Schüttelkopf AW, van Aalten DMF. 2004. PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* 60 (8):1355–1363.

Sugahara K, Shigeno K, Masuda M, Fujii N, Kurosaka A, Takeda K. 1994. Structural studies on the chondroitinase ABC-resistant sulfated tetrasaccharides isolated from various chondroitin sulfate isomers. *Carbohydr Res.* 255:145–163.

Sugimura T, Kato F, Mimatsu K, Takenaka O, Iwata H. 1996. Experimental chemonucleolysis with chondroitinase ABC in monkeys. *Spine.* 21(2):161-165.

Sugiura N, Setoyama Y, Chiba M, Kimata K, Watanabe H. 2011. Baculovirus Envelope Protein ODV-E66 Is a Novel Chondroitinase with Distinct Substrate Specificity. *J Biol Chem.* 286(33):29026–29034.

Suzuki S, Saito H, Yamagata T, Anno K, Seno N, Kawai Y, Furuhashi T. 1968. Formation of Three Types of Disulfated Disaccharides from Chondroitin Sulfates by Chondroitinase Digestion. *J. Biol. Chem.* 243(7): 1543–1550.

Takahashi T, Kurihara H, Nakajima S, Kato T, Matsuzaka S, Sekiguchi T, Onaya M, Miyauchi S, Mixuno S, Horie K, Fujita Y, Hirose T. 1996. Chemonucleolytic effects of chondroitinase ABC on normal rabbit intervertebral discs. Course of action up to 10 days postinjection and minimum effective dose. *Spine.* 21(21):2405-2411.

Vagin A, Teplyakov A. 2010. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* 66(Pt 1):22–25.

Volpi N. 2007. Analytical Aspects of Pharmaceutical Grade Chondroitin Sulfates. *J. Pharma. Sci.* 96(12): 3168-3180.

Watanabe I, Hikita T, Mizuno H, Sekita R, Minami A, Ishii A, Minamisawa Y, Suzuki K, Maeda H, Hidari IPJK, Suzuki T. 2015. Isolation and characterization of monoclonal antibodies specific for chondroitin sulfate E. *Glycobiology.* 25(9):953-962.

Yamada K, Tanabe S, Ueno H, Oinuma A, Takahashi T, Miyauchi S, Shieno S, Hirose T, Miyahara K,
Sato M. 2001. Investigation of the short-term effect of chemonucleolysis with chondroitinase ABC. *J Vet Med Sci.* 63(5):521-525.

Yamagata T, Saito H, Habuchi O, Suzuki S. 1968. Purification and Properties of Bacterial Chondroitinases and Chondrosulfatases. *J. Biol. Chem.* 243(7):1523–1535

Yoshida K, Miyauchi S, Kikuchi H, Tawada A, Tokuyasu K. 1989. Analysis of unsaturated disaccharides from glycosaminoglycuronan by highperformance liquid chromatography. *Anal Biochem.* 177(2):327–332.

Zhang H, Neau S.H. 2001. In vitro degradation of chitosan by a commercial enzyme preparation: effect of molecular weight and degree of deacetylation. *Biomaterials.* 22(12):1653-1658.

**Legends to figures**

Fig. 1. Reactivity of cABC-I with various GAGs. Reactivity of CSCs with various Mws (A), CS isomers with a Mw of 25 kDa (B), CSA and CSC with a Mw of 1 kDa (C), CH and HA (D), enzymatic reaction of CSA and CSC with a Mw of 25 kDa for 24 hour (E), enzymatic reaction of hCSA and hCSC for 24 hour (F) were determined by time-dependent absorbance at 232 nm of the reaction mixture. **The values represent the mean ± standard deviations (n = 3).**
Fig. 2. Quantification of generated disaccharides units. The disaccharides produced by enzymatic reaction of CSA, DS, CSC, CSD, and CSE with Mw of almost 25 kDa for 60 min (A, B, C, D, and E, respectively). The disaccharides produced by enzymatic reaction of CSA and CSC with Mw of almost 25 kDa (F and G, respectively), and hCSA or hCSC for 24 hour (H). The disaccharides were determined by HPLC with post column fluorescence detection (n = 1).

Fig. 3. pH- $k_{cat}/K_{m}$ profile for the reaction of cABC-I with CSA and CSC. The $k_{cat}/K_{m}$ for CSA and CSC with Mw of almost 25 kDa at various pH values (A) and that for Mw of 1 kDa (B). The values represent the mean ± standard deviations (n = 3).
Fig. 4. Overall structure of cABC-I. The crystal structure of cABC-I (PDB ID: 7EIP), complex with ΔDi-4S (PDB ID: 7EIQ), ΔDi-6S (PDB ID: 7EIR), and ΔDi-0S (PDB ID: 7EIS) are shown as a cartoon model (A, B, C, and D, respectively). Colors: N-terminal domain (25 - 234) – magenta, middle domain (235 - 617) – green, C-terminal domain (618 - 1021) – cyan; the disaccharides are shown as yellow sticks.
Fig. 5. Interaction of cABC-I with ΔDi-4S, ΔDi-6S and ΔDi-0S (A, B, and C, respectively). The disaccharides and the residues involved in interactions are shown as yellow and green sticks, respectively. Key interactions are shown in red dashed lines.

Fig. 6. Structure of cABC-I active site in the docking model with CSA and CSC tetrasaccharides (A and B, respectively). Each tetrasaccharide and residues involved in interactions are shown as slate blue and green sticks, respectively. Key interactions are shown in red dashed lines.

Fig. 7. Schematic illustration of the reaction mechanism of cABC-I. (a) CSA substrate in the active site of cABC-I. (b) CSC substrate in the active site of cABC-I.

Tables

| Table I | Mw and disaccharide compositions of GAGs used in enzyme assay. |
|---------|---------------------------------------------------------------|
| GAG     | Mw (kDa) | 0S  | 6S  | 4S  | (2,6)S | (4,6)S | (2,4,6)S | Sulfur content (%) |
| CSA-1K  | 1.3      | 1.1 | 14.3| 84.6| 0.0    | 0.0    | 0.0      | 6.3            |
| CSA-26K | 26.7     | 6.3 | 16.1| 77.3| 0.1    | 0.3    | 0.0      | 6.1            |
Table II  Kinetic parameters of cABC-I with hCSA and hCSC.

| Substrate | $K_m$ (μM) | $k_{cat}$ (sec$^{-1}$) | $k_{cat} / K_m$ (μM$^{-1}$ sec$^{-1}$) |
|-----------|------------|------------------------|--------------------------------------|
| hCSA      | 1.56 ± 0.01| 1510 ± 130             | 964 ± 74                              |
| hCSC      | 2.11 ± 0.54| 1210 ± 83              | 599 ± 153                             |

The values represent the mean ± standard deviations ($n = 3$).

Table III Distance between the functional groups of disaccharides and interacting amino acid residues.

| Region | Sugar     | protein | Distance (Å) |
|--------|-----------|---------|--------------|
|        |           | Atom    | ΔDi-0S | ΔDi-4S | ΔDi-6S |
| hCSA   | O2        | Asn564  | ND2    | -      | 3.1    | 3.4   |
|        | O2        | Asn564  | ND2    | -      | 3.1    | 3.4   |
|        | N         | Gly507  | O      | -      | 3.3    | 3.2   |
|        | O6        | Asn447  | ND2    | -      | 3.4    | -     |
|        | O6        | Asn276  | ND2    | -      | 3.3    | -     |
|        | O6        | Arg395  | NH1    | -      | 3.5    | -     |
|        | O7        | His561  | NE2    | 3.1    | 2.8    | 2.8   |
| hCSC   | O2        | Asn564  | ND2    | -      | 2.8    | 2.8   |
|        | O5        | Arg500  | NH2    | 3.0    | 3.2    | 3.0   |
|        | O6B  | Arg500 | NH2 | 2.8 | 2.9 | 3.0 |
|--------|------|--------|-----|-----|-----|-----|
|        |      |        | NE  | 2.6 | 2.7 | 2.9 |
|        |      |        | NE2 | 2.9 | 2.7 | 2.7 |
| N      | Tyr655 | OH    | 3.4 | 3.1 | 3.5 |
| GalNAc (sulfate) | 4O3S | Arg500 | NH2 | -   | 3.1 | -   |
|        | 6O2S | Arg500 | NH1 | -   | -   | 3.9 |
|        |      | His501 | N   | -   | -   | 2.9 |
| ΔGlcA  | O6A  | Thr713 | N   | -   | 3.0 | -   |
|        | O6B  |        | OG1 | -   | 2.8 | -   |
| GalNAc (sulfate) | 4O2S | Arg966 | NH2 | -   | 3.0 | -   |
|        | O6   | Met715 | N   | -   | 2.8 | -   |