



Contents lists available at ScienceDirect

# Biochemical and Biophysical Research Communications

journal homepage: [www.elsevier.com/locate/ybbrc](http://www.elsevier.com/locate/ybbrc)



## Crystal structure of chondroitin polymerase from *Escherichia coli* K4

Takuo Osawa<sup>a</sup>, Nobuo Sugiura<sup>b,c</sup>, Hiroaki Shimada<sup>a</sup>, Ryoko Hirooka<sup>d</sup>, Atushi Tsuji<sup>d</sup>, Tadayoshi Shirakawa<sup>e</sup>, Keiichi Fukuyama<sup>e</sup>, Makoto Kimura<sup>a,d</sup>, Koji Kimata<sup>b</sup>, Yoshimitsu Kakuta<sup>a,d,\*</sup>

<sup>a</sup> Laboratory of Structural Biology, Graduate School of Systems Life Sciences, Kyushu University, Fukuoka 812-8581, Japan

<sup>b</sup> Institute for Molecular Science of Medicine, Aichi Medical University, Yazako, Nagakute, Aichi 480-1195, Japan

<sup>c</sup> Central Research Laboratories, Seikagaku Corporation, Tateno, Higashiyamato-shi, Tokyo 207-0021, Japan

<sup>d</sup> Laboratory of Biochemistry, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, 6-10-1, Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

<sup>e</sup> Department of Biological Sciences, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan

### ARTICLE INFO

#### Article history:

Received 16 August 2008

Available online 2 September 2008

#### Keywords:

Bi-functional glycosyltransferases

Crystal structure

Chondroitin polymerase

Biosynthesis of chondroitin

Biosynthesis of glycosaminoglycan

### ABSTRACT

Elongation of glycosaminoglycan chains, such as heparan and chondroitin, is catalyzed by bi-functional glycosyltransferases, for which both 3-dimensional structures and reaction mechanisms remain unknown. The bacterial chondroitin polymerase K4CP catalyzes elongation of the chondroitin chain by alternatively transferring the GlcUA and GalNAc moiety from UDP-GlcUA and UDP-GalNAc to the non-reducing ends of the chondroitin chain. Here, we have determined the crystal structure of K4CP in the presence of UDP and UDP-GalNAc as well as with UDP and UDP-GlcUA. The structures consisted of two GT-A fold domains in which the two active sites were 60 Å apart. UDP-GalNAc and UDP-GlcUA were found at the active sites of the N-terminal and C-terminal domains, respectively. The present K4CP structures have provided the structural basis for further investigating the molecular mechanism of biosynthesis of chondroitin chain.

© 2008 Elsevier Inc. All rights reserved.

Chondroitin, heparin/heparan and hyaluronan are the primary glycosaminoglycans (GAGs) found in humans and are linear polysaccharides consisting of an amino sugar and an uronic acid. Chondroitin chains range from 40 to over 100 repeating units of the disaccharide (GlcUA  $\beta$  (1–3)–GalNAc  $\beta$  (1–4)). Sulfated chondroitins are involved in the regulation of various biological functions such as central nervous system development, wound repair, infection, growth factor signaling, and morphogenesis, in addition to its conventional structural roles [1,2]. In *Caenorhabditis elegans*, chondroitin is an essential factor for the worm to undergo cytokinesis and cell division [3]. Certain pathogenic bacteria yield an extracellular polysaccharide coating composed of GAG or GAG-like polymers [4]. Chondroitin is synthesized as proteoglycans, sulfated and secreted to the cell surface or extracellular matrix. Various glycosyltransferases are involved in the biological synthesis of the chondroitin chain, of which a bi-functional glycosyltransferase catalyzes elongation of the chain. The lack of a 3-dimensional structure of a bi-functional glycosyltransferase, however, has kept us from learning insights of the mechanism.

Microbial polysaccharides are structurally identical or very similar to mammalian GAGs. Using this structural similarity to

their advantages, the K4 and K5 strains of *Escherichia coli* promote their invasiveness and pathogenicity to host cells by masking their presence and avoiding antibody response. K4 chondroitin constitutes extracellular layer of the *E. coli* capsule, in which a number of enzymes from the gene clusters of the group 2 K antigens are utilized to generate chondroitin synthesis [5]. We have previously identified protein K4CP, the K4 capsule gene *kfoc* product, as the bi-functional inverting glycosyltransferase that catalyzes elongation of the chondroitin chain [6]. K4CP transfers *N*-acetylgalactosamine (GalNAc) and glucuronic acid (GlcUA) residues alternatively from UDP-GalNAc and -GlcUA to the non-reducing ends of chondroitin chains, respectively (Supplemental figure 1). In the GalNAc transfer reaction, the acceptor substrates are chondroitin oligomers that exhibit GlcUA at the non-reducing end, whereas in the GlcUA transfer reaction, chondroitin oligomers with GalNAc at the non-reducing end are used as acceptor substrates. After linking fructose at the C3 position of the GlcUA molecules and adding the phospholipid-conjugated 2-keto-3-deoxyoctulosonic acid at the reducing end, the chondroitin translocates across the outer membrane and form the extracellular layer of the capsule.

To obtain the X-ray crystal structure of the K4CP, a truncated K4CP enzyme that fully retains the catalytic activity is constructed, bacterially expressed, purified and crystallized. As the first structure of a bi-functional glycosyltransferase, the X-ray crystal structure of the K4CP enzyme has now been determined. Here, we have

\* Corresponding author. Address: Laboratory of Biochemistry, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, 6-10-1, Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan. Fax: +81 92 642 2854.

E-mail address: [kakuta@agr.kyushu-u.ac.jp](mailto:kakuta@agr.kyushu-u.ac.jp) (Y. Kakuta).

reported the structural features of the bi-functional K4CP enzyme and have discussed their implications in understanding the catalytic mechanism.

## Materials and methods

**Limited trypsin digestion.** Full-length K4CP was expressed as described [6]. Full-length K4CP in 50 mM Tris–HCl (pH 8.0) and 500 mM sodium chloride was incubated with trypsin (Sigma) for 15 min at 293 K in 400,000:1, protein: enzyme ratio. The sample was diluted with SDS buffer, boiled and subjected to SDS–PAGE followed by electroblotting to PVDF and N-terminal sequencing on a PSQ-1 sequencer (Shimadzu).

**Overexpression and purification.** Primers with the sequences 5'-CGGGATCCAAAGCTGTTATTGATATTGATGCAGCAACA-3' and 5'-GGAATTCCTTATAAATCATCTCTATTTTCC-3' were used to remove 57 residues from the N-terminus. The PCR product was inserted into pGEX4T3 vectors, transformed into *E. coli* strain BL21 (DE3) and grown at 30 °C. Overexpression of protein was induced with 0.1 mM IPTG when the OD<sub>600</sub> was ~0.60, and the cells were harvested after overnight incubation at 30 °C. GST-fusion K4CPΔ57 was purified with a Glutathione Sepharose 4B column (Amersham). GST was cleaved at 4 °C overnight with 1:10,000 thrombin (Sigma), leaving two extra residues at the N-terminus (Gly-Ser). The cleaved product was purified by Glutathione Sepharose 4B, followed by Superdex 200 columns (Pharmacia) equilibrated with 50 mM Tris–HCl, pH 8.0 and 500 mM NaCl. Purified protein was concentrated to ~20 mg ml<sup>-1</sup> for crystallization trials. Selenomethionyl K4CPΔ57 was prepared using the described protocols.

**Crystallization and data collection.** A protein mixture containing 20 mg ml<sup>-1</sup> protein, 5 mM MnCl<sub>2</sub> and different reagents (for UDP–GalNAc complex, 10 mM UDP, 20 mM DTT, and 84 mM IPTG; for UDP–GlcUA complex, 0.5 mM UDP, 9.5 mM UDP–GlcUA, 20 mM DTT, and 400 mM NDSB 201; for selenomethionyl K4CPΔ57 complex with UDP, 10 mM UDP, and 800 mM NDSB 201) was used for crystallization at 4 °C. The crystallization condition involves 15–20% PEG 3350 and 200 mM NaCl, and the hanging drop vapor diffusion and batch method were used for both complex and SeMet substituents. Crystals for UDP–GalNAc/UDP complex were soaked with crystallization solution containing 40 mM UDP–GalNAc for 24 h. Data were collected at 100 K with 20% (v/v) glycerol containing solution as a cryoprotectant. A two-wavelength data set of selenomethionyl UDP complex crystal was collected at SPring-8, beamline BL38B1, using an ADSC CCD detector. Data sets of crystals of UDP–GlcUA/UDP and UDP–GalNAc/UDP complex were collected at Photon Factory (PF), beamline NW12A, using a Quantum 210r CCD detector (ADSC) and BL5A, using a Quantum 315 CCD detector (ADSC), respectively. All data were processed using HKL2000 [7]. Statistics for data collection and processing are summarized in Table 1.

**Structure determination and refinement.** Crystals of selenomethionine-labeled protein were obtained using the same protocol as for native crystals. Anomalous differences in the selenium peak data set were used to find all 28 selenium sites (seven sites per monomer with four monomers in the asymmetric unit) and to calculate the initial phase using peak and remote data with the BNP program [8]. The phases were improved using solvent-flipping and ncs averaging (based on selenium sites) with Resolve [9]. The original model of K4CPΔ57 was built into electron density calculated from these phases using the Coot program [10]. The model was subsequently refined using iterative cycles of model building and Refmac [11].

The crystal structure of the UDP–GlcUA/UDP complex was determined by molecular replacement using the partially refined

**Table 1**

Data collection, phasing and refinement statistics for MAD (SeMet) structures

	UDP–GlcUA/ UDP	UDP–GalNAc/ UDP	SeMet (UDP)	
<b>Data collection</b>				
Space group	P2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub>
Cell dimensions a, b, c (Å)	84.09, 219.83, 85.86	82.37, 109.29, 85.58	88.77, 235.11, 85.77	84.51, 236.09, 86.25
α, β, γ (°)	90.00, 103.07, 90.00	90.00, 103.47, 90.00	90.00, 103.73, 90.00	90.00, 103.91, 90.00
			Peak	Remote
Wavelength	1.0000	1.0000	0.9870	0.9000
Resolution (Å)	50.0–2.4	50.0–3.0	50.00–2.80	50.00–2.80
R <sub>sym</sub>	12.7 (48.0)	10.2 (32.9)	12.6 (51.2)	10.0 (42.2)
I/σI	11.9 (1.5)	11.1 (1.8)	16.7 (1.0)	15.1 (1.8)
Completeness (%)	91.3 (71.3)	86.3 (80.9)	86.8 (43.7)	72.4 (14.1)
Redundancy	3.4 (2.6)	3.5 (2.7)	18.2 (6.4)	6.7 (2.4)
<b>Refinement</b>				
Resolution (Å)	20.0–2.4	20.0–3.0		
No. reflections	97,298	27,587		
R <sub>work</sub> /R <sub>free</sub>	0.2228/ 0.2872	0.2139/ 0.28901		
<b>No. atoms</b>				
Protein	19364	9710		
Mn <sup>2+</sup>	8	4		
UDP	50	50		
UDP–GlcUA	222			
UDP–GalNAc		78		
Water	528	109		
<b>B-factors</b>				
Protein	51.6	49.6		
Mn <sup>2+</sup>	54.2	50.2		
UDP	42.9	44.8		
UDP–GlcUA	56.2			
UDP–GalNAc		53.9		
Water	46.3	35.4		
<b>R.m.s. deviations</b>				
Bond lengths (Å)	0.013	0.015		
Bond angles (°)	1.5	1.6		

<sup>a</sup>Values in parentheses are for highest-resolution shell.

dimmer model as a search model with the Molrep program [12]. The model was built using the Coot program and Arp/warp [13]. The structure was refined with the Refmac and CNS programs [14]. Iterative cycles of refinement and manual rebuilding in Coot were carried out until the R<sub>cryst</sub> factor was 22.8% and R<sub>free</sub> factor was 28.7%.

The crystal structure of K4CPΔ57/UDP–GalNAc/UDP was determined by molecular replacement using the monomer of K4CPΔ57/UDP–GlcUA/UDP as a search model with the Molrep program. The structure was refined with the CNS program. Iterative cycles of refinement and manual rebuilding in Coot were carried out until the R<sub>cryst</sub> factor was 21.9% and R<sub>free</sub> factor was 29.1%. Stereochemical checks were carried out with PROCHECK [15]. In Ramachandran plot analysis of the K4CPΔ57 complex with UDP/UDP–GlcUA, the most favored (%) was 81.4, allowed (%) was 17.8, generously allowed (%) was 0.8, and disallowed (%) was 0. In Ramachandran plot analysis of the K4CPΔ57 complex with UDP/UDP–GalNAc, the most favored (%) was 75.9, allowed (%) was 22.9, generously allowed (%) was 1.2, and disallowed (%) was 0. Refinement statistics are summarized in Table 1.

**Accession numbers.** Protein Data Bank: Coordinates have been deposited with Accession Nos. 2Z87 (K4CPΔ57 complex with UDP/UDP–GalNAc complex) and 2Z86 (K4CPΔ57 complex with UDP/UDP–GlcUA).

## Results and discussion

### Crystallization

Limited trypsin digestion of full-length K4CP produced a truncated K4CP lacking the N-terminal 57 amino acid residues, the K4CP $\Delta$ 57: 58–686 that was found to fully retain the chondroitin polymerase activity. Gel filtration chromatography showed that this N-terminal truncated enzyme is present in a homodimer at the concentration of 20 mg/ml, while the K4CP exists as a monomer under a lower concentration of 1 mg/ml. Given the fact that K4CP $\Delta$ 57: 58–686 is water soluble and retains the activity, a cDNA that encodes K4CP $\Delta$ 57: 58–686 was inserted into pGEX4T3 and was transformed into BL21 (DE3) cells. Using bacterial expression, a large amount of pure K4CP $\Delta$ 57: 58–686 protein (hereafter just called K4CP) was prepared for crystallization. Three forms of the K4CP crystals were produced; the complex of K4CP with inactive donor UDP and those with UDP and UDP-GlcUA as well as with UDP and UDP-GalNAc (Table 1). In the K4CP-UDP and K4CP-UDP-UDP-GlcUA complexes, the K4CP was crystallized as two dimers in the asymmetric unit, while only a single K4CP dimer was present in the asymmetric unit of the K4CP-UDP-UDP-GalNAc complex.

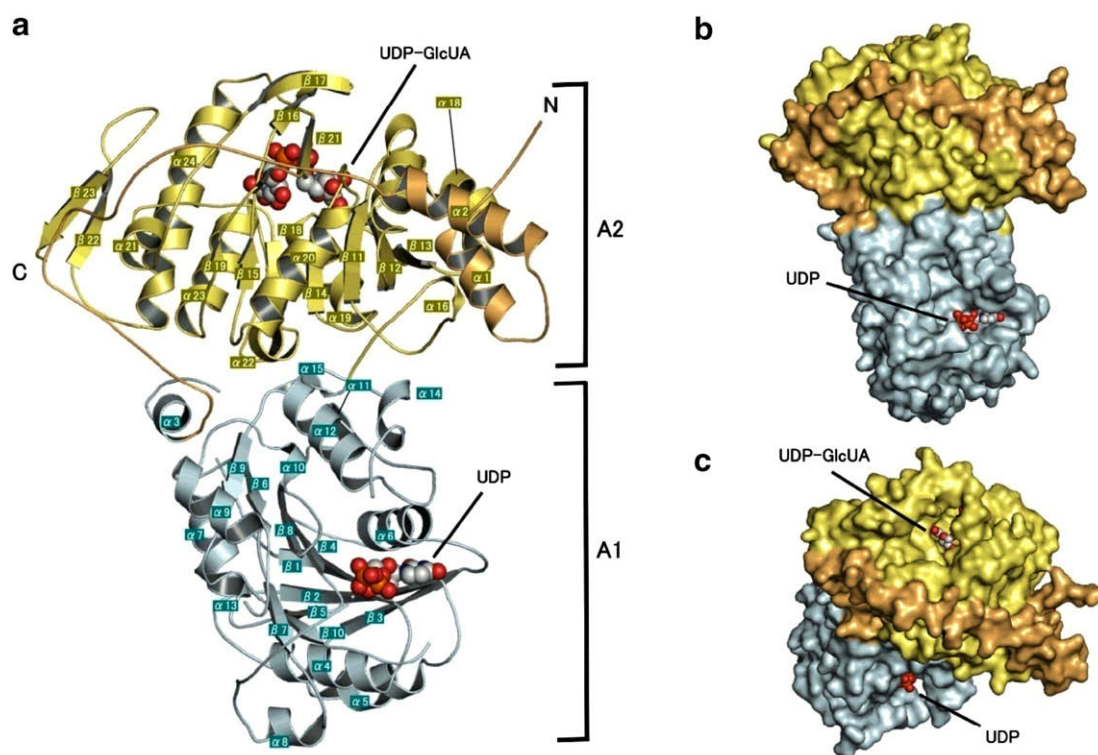
### Overall structure

The present crystal structures revealed two domains of the K4CP molecule with an overall size of approximately 45 Å × 70 Å × 85 Å (Fig. 1a). Seventy-one residues within a coil of the N-terminal region of K4CP (residues 57 to 129) form a linear structure consisting of a random coil and three  $\alpha$  helices ( $\alpha$ 1– $\alpha$ 3). Residues 130–417 constitutes the first domain (A1) and con-

tains the 13  $\alpha$  helices ( $\alpha$ 3– $\alpha$ 15) and 10  $\beta$  strands ( $\beta$ 1– $\beta$ 10) (Supplemental figure 2). The second domain (A2) comprised with residues 418–682 and composed of the 11  $\alpha$  helices ( $\alpha$ 1,  $\alpha$ 2,  $\alpha$ 16– $\alpha$ 24) and 13  $\beta$  strands ( $\beta$ 11– $\beta$ 23) (Supplemental figure 2). Both A1 and A2 domains adopt the glycosyltransferase GT-A fold, which nicely superimpose with each other, a r.m.s. deviation of 3.2 Å for 207 C $\alpha$ 's and 2.1 Å for 192 C $\alpha$ 's. Both A1 and A2 domains exhibit nearly identical structure to the human GlcAT1, the mono-functional glycosyltransferase in proteoglycan biosynthesis that has GlcUA transferase activity to the oligosaccharide Gal-Gal-Xyl [17]. In addition, the domains are found to be structurally similar to a large number of the GT-A fold glycosyltransferases with bacterial SpsA being the most homologous. Thus, K4CP appears to be a fusion protein of two mono-functional glycosyltransferases with a GT-A fold. Each of the A1 and A2 domains consists of two subdomains: the signature DXD motif resides in the N-terminal subdomains and the active site pocket is organized in the open cleft between the two subdomains. To our surprise, the active site pockets of the A1 and A2 domains are independently organized and are found at the opposite sides of the K4CP molecule, around 60 Å in distant from each other (between DXD motif in A1 and DXD motif in A2) (Fig. 1b and c). The other intriguing facet is the structure of the N-terminal 71 residues of the K4CP molecule wrapping around the surface of the A2 domain.

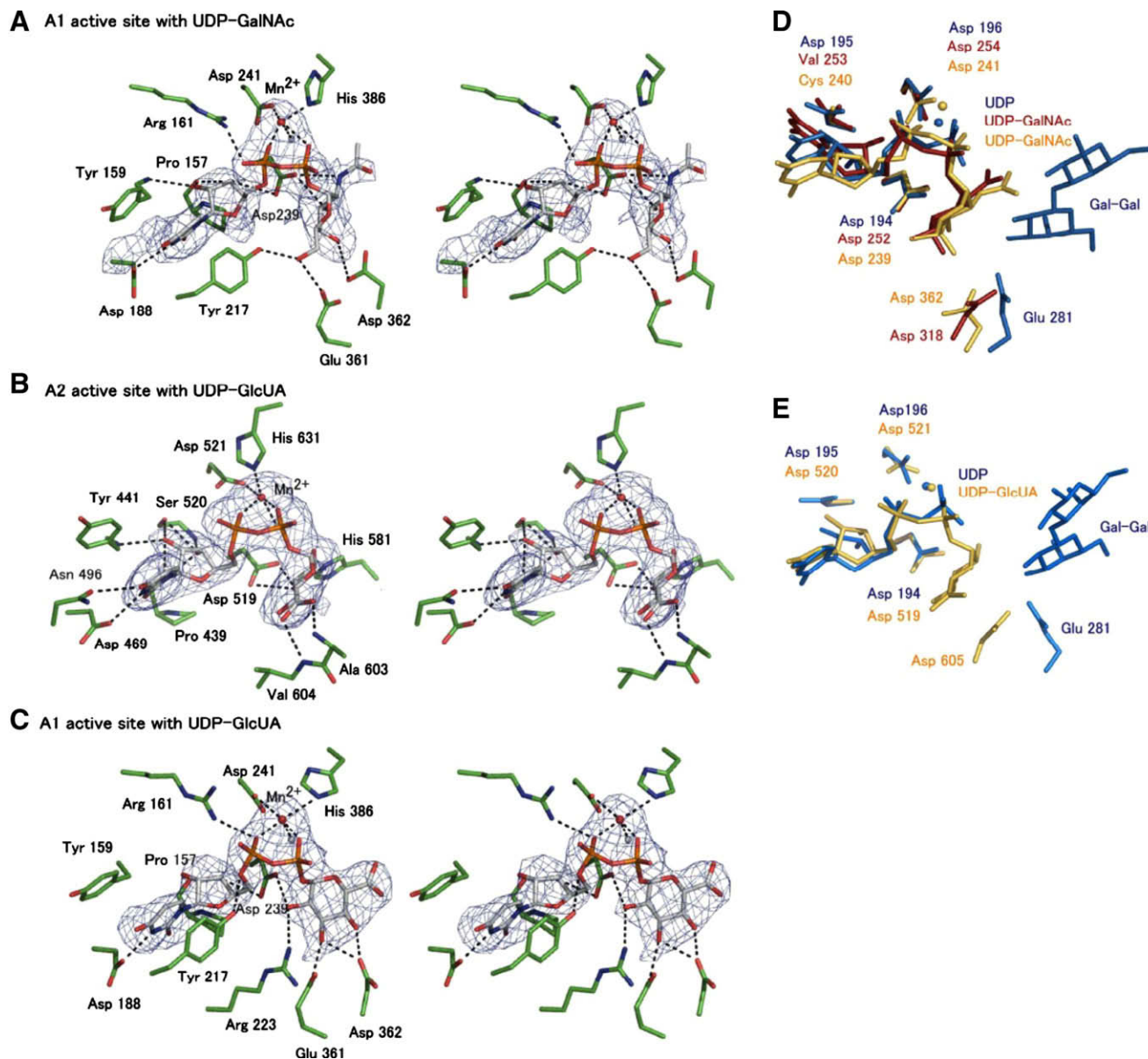
### Binding of the donor substrates

In the crystal structure of the K4CP complex with UDP and UDP-GalNAc, the UDP-GalNAc molecule is accommodated at the activate site of the A1 domain, while the A2 domain accommodates the UDP molecule (Fig. 2a). On the other hand, in the



**Fig. 1.** Overall structure of K4CP $\Delta$ 57 complex with UDP-GlcUA and UDP. (a) Secondary structural elements are indicated by ribbons for  $\alpha$ -helices, arrows for  $\beta$ -strands with a domain color (A1 domain, blue; A2 domain (58–129), orange; A2 domain (418–682), yellow). UDP-GlcUA and UDP are shown as space-fill models with an atomic color scheme (C, white; N, blue; O, red; P, orange). (b) Surface representations with the same color as (a). UDP-GlcUA is shown as a space-fill model. (c) Different orientation of (b). UDP is shown as a space-fill model.





**Fig. 2.** Active site structures with UDP-GalNAc binding and UDP-GlcUA binding. (A) Stereo diagram displaying the active site region of the K4CPΔ57 complex with UDP-GalNAc and  $Mn^{2+}$  ion. Hydrogen bond interactions are drawn as black dashes. The  $F_o - F_c$  omitted electron density map of UDP-GalNAc and  $Mn^{2+}$  ion contoured at  $2.5\sigma$  is displayed in blue mesh. (B) Stereo diagram displaying the active site region of the K4CPΔ57 complex with UDP-GlcUA and  $Mn^{2+}$  ion in A2 domain. Hydrogen bond interactions are drawn as black dashes. The  $F_o - F_c$  omitted electron density map of UDP-GlcUA and  $Mn^{2+}$  ion contoured at  $3\sigma$  is displayed in blue mesh. (C) Stereo diagram displaying the active site region of the K4CPΔ57 complex with UDP-GlcUA and  $Mn^{2+}$  ion in A1 domain. Hydrogen bond interactions are drawn as black dashes. The  $F_o - F_c$  omitted electron density map of UDP-GlcUA and  $Mn^{2+}$  ion contoured at  $3\sigma$  is displayed in blue mesh. (D) Superimposition between the A1 domain of K4CPΔ57 (yellow),  $\beta$ 4Gal-T1 complex with UDP-GalNAc (red), and GlcAT1 complex with UDP and acceptor sugar (blue). (E) Superimposition between A2 domain of K4CPΔ57 (yellow) and GlcAT1 complex with UDP and acceptor sugar (blue).

crystal structure of the K4CP complex with UDP and UDP-GlcUA, the UDP-GlcUA and UDP molecules are found at the active sites of the A2 and A1 domains, respectively (Fig. 2b and c). At the two different active sites, the UDP-GalNAc and UDP-GlcUA molecules are accommodated by the nearly identical conformation and orientation. These specific donor bindings suggested that the A1 and A2 domains catalyze the UDP-GalNAc and UDP-GlcUA transfer reactions, respectively. The DXD motif is the key binding element for  $Mn^{2+}$  ion for all known  $Mn^{2+}$  requirements of GT-A fold glycosyltransferases. In both the A1 and A2 domains, one of the residues in each DXD motif binds  $Mn^{2+}$  ion directly: Asp241 in the  $^{239}DCD^{241}$  motif of the A1 domain, and Asp521 in the  $^{519}DSD^{521}$  motif of the A2 domain directly coordinate with the  $Mn^{2+}$  ion (Fig. 2a and b).

#### Identification of the catalytic base residues

Both glycosyltransfer reactions catalyzed by the K4CP enzyme are the inverting reaction, inverting  $\alpha$  configuration of the C1 bond of the UDP-sugar molecule to the  $\beta$  configuration in the product after transferred. This inverting reaction undergoes a  $SN_2$  type in-line displacement mechanism that requires a catalytic base residue to deprotonate the acceptor group of acceptor substrate. Such base residues (generally Glu or Asp) have been identified in various glycosyltransferases [5,16,17]. To search the catalytic bases of the K4CP enzyme, we superimposed the A1 domain with the  $\beta$ 4Gal-T1 (bovine  $\beta$ 4-galactosyltransferase) as well as with the human GlcAT1 (Fig. 2d). Both catalytic bases Asp318 (in  $\beta$ 4Gal-T1) and Glu281 are (in GlcAT1) superimposed on Asp362 of the A1 domain.

Similarly Asp605 of the A2 domain is superimposed on Glu281. In addition to these base residues, the DXD motifs of the A1 and A2 domains perfectly overlapped with those of  $\beta$ 4Gal-T1 and GlcAT1 (Fig. 2e).

#### Unusual binding of UDP-GlcUA to the A1 domain

In addition to the abovementioned X-ray crystal structure in which UDP and UDP-GlcUA are bound to the A1 and A2 domains, respectively, the other crystal structure that accommodates UDP-GlcUA and UDP in the A1 and A2 active sites was also obtained. At the A1 active site, the UDP-GlcUA molecule adopts an unusual conformation and orientation (Fig. 2c): the C1 carbon of the GlcUA moiety is positioned away from the possible nucleophilic attack by the acceptor group. Whereas the carboxyl group of the GlcUA moiety at the A2 active site interacts with Asp205, Ala603, and Val604, at the A1 active site it forms only a single interaction with Arg268. These differences in the interaction may have positioned the UDP-GlcUA to adopt the conformation of a non-productive substrate unable to transfer the GlcUA moiety. Presently, it remains enigmatic whether this non-productive binding of UDP-GlcUA at the A1 active site plays any roles in the polymerase reaction. The non-productive binding could simply be an artifact caused during crystallization; it may be the mechanism to stop chain elongation or may help the K4CP enzyme to stop the GalNAc transfer during the period in which the enzyme is catalyzing the GlcUA transfer reaction.

#### Conclusion

We have solved the X-ray crystal structures of the bacterial bi-functional chondroitin polymerase K4CP in the presence of the donor substrates UDP, UDP-GlcUA and UDP-GalNAc. The K4CP molecule consists of two GT-A domains bearing the GalNAc and GlcUA transfer activities to the N- and C-terminal A1 and A2 domains, respectively. The two active sites are independently organized to the distant areas of the K4CP molecule and do not share the same space. The present K4CP structures have provided us with the molecular basis for further investigating the catalytic mechanism by which the K4CP enzyme accurately alternates the transfer of GalNAc and GlcUA. Biosynthesis of glucosaminylglycans such as chondroitin and heparin in mammals requires the additional three types of glycosyltransferases that catalyzes the distinct transfer reactions: inverting and retaining reactions as well as the reaction transferring a mono-saccharine to the protein core. For the mammalian glycosyltransferases that catalyze these three reactions, the X-ray crystal structures have been available [18]. Our K4CP structure now enables us to investigate the polymerization reaction of the bi-functional glycosyltransferases.

#### Acknowledgments

We thank the staffs of the SPring-8 (beamline 38B1) and the PF (beamline BL5A and NW12A) for help with data collection. This

work was supported by a special research fund from the Seikagaku Corporation, and by a Grant-in-Aid for Scientific Research on Priority Areas 14082206 from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. This work was also supported by a research grant from the National Project on Protein Structural and Functional Analyses from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2008.08.121.

#### References

- [1] J.V.C. Hascall, D.K. Heinegard, T.N. Wight, in: E.D. Hay (Ed.), *Cell Biology of Extracellular Matrix*, Plenum Press, New York, 1991, pp. 149–176.
- [2] T.N. Wight, D.K. Heinegard, V.C. Hascall, in: E.D. Hay (Ed.), *Cell Biology of Extracellular Matrix*, Plenum Press, New York, 1991, pp. 45–78.
- [3] S. Mizuguchi, T. Uyama, H. Kitagawa, K.H. Nomura, K. Dejima, K. Gengyo-Ando, S. Mitani, K. Sugahara, K. Nomura, Chondroitin proteoglycans are involved in cell division of *Caenorhabditis elegans*, *Nature* 423 (2003) 443–448.
- [4] C. Whitfield, I.S. Roberts, Structure, assembly and regulation of expression of capsules in *Escherichia coli*, *Mol. Microbiol.* 31 (1999) 1307–1319.
- [5] B. Ramakrishnan, P.V. Balaji, P.K. Qasba, Crystal structure of beta1,4-galactosyltransferase complex with UDP-gal reveals an oligosaccharide acceptor binding site, *J. Mol. Biol.* 318 (2002) 491–502.
- [6] T. Ninomiya, N. Sugiura, A. Tawada, K. Sugimoto, H. Watanabe, K. Kimata, Molecular cloning and characterization of chondroitin polymerase from *Escherichia coli* strain K4, *J. Biol. Chem.* 277 (2002) 21567–21575.
- [7] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol.* 276 (1997) 307–326.
- [8] H. Xu, C.M. Weeks, Rapid and automated substructure solution by shake-and-bake, *Acta Crystallogr. D Biol. Crystallogr.* 64 (2008) 172–177.
- [9] T.C. Terwilliger, J. Berendzen, Automated MAD and MIR structure solution, *Acta Crystallogr. D Biol. Crystallogr.* 55 (1999) 849–861.
- [10] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr.* 60 (2004) 2126–2132.
- [11] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D Biol. Crystallogr.* 53 (1997) 240–255.
- [12] A. Vagin, A. Teplyakov, An approach to multi-copy search in molecular replacement, *Acta Crystallogr. D Biol. Crystallogr.* 56 (2000) 1622–1624.
- [13] R.J. Morris, A. Perrakis, V.S. Lamzin, ARP/wARP and automatic interpretation of protein electron density maps, *Methods Enzymol.* 374 (2003) 229–244.
- [14] A.T. Brunger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Crystallography & NMR system: a new software suite for macromolecular structure determination, *Acta Crystallogr. D Biol. Crystallogr.* 54 (1998) 905–921.
- [15] A.A. Vaguine, J. Richelle, S.J. Wodak, SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model, *Acta Crystallogr. D Biol. Crystallogr.* 55 (1999) 191–205.
- [16] L.C. Pedersen, K. Tsuchida, H. Kitagawa, K. Sugahara, T.A. Darden, M. Negishi, Heparan/chondroitin sulfate biosynthesis. Structure and mechanism of human glucuronyltransferase I, *J. Biol. Chem.* 275 (2000) 34580–34585.
- [17] L.N. Gastinel, C. Bignon, A.K. Misra, O. Hindsgaul, J.H. Shaper, D.H. Joiasse, Bovine alpha1,3-galactosyltransferase catalytic domain structure and its relationship with ABO histo-blood group and glycosphingolipid glycosyltransferases, *EMBO J.* 20 (2001) 638–649.
- [18] M. Negishi, J. Dong, T.A. Darden, L.G. Pedersen, L.C. Pedersen, Glucosaminylglycan biosynthesis: what we can learn from the X-ray crystal structures of glycosyltransferases GlcAT1 and EXTL2, *Biochem. Biophys. Res. Commun.* 303 (2003) 393–398.