Dual Affinity Labeling of the Active Site of Human Lysozyme with an *N*-Acetyllactosamine Derivative: First Ligand Assisted Recognition of the Second Ligand[†,‡]

Michiro Muraki* and Kazuaki Harata

Biomolecules Department, National Institute of Bioscience and Human-Technology, 1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan

Naoki Sugita and Ken-ichi Sato

Faculty of Engineering, Kanagawa University, Yokohama, Kanagawa 221-0802, Japan

Received July 23, 1998; Revised Manuscript Received October 16, 1998

ABSTRACT: Among the three kinds of the 2′,3′-epoxypropyl β-glycoside of disaccharides (GlcNAc-β1,4-GlcNAc, Gal-β1,4-GlcNAc, and Man-β1,4-GlcNAc), the derivative of *N*-acetyllactosamine (Gal-β1,4-GlcNAc-Epo) caused the dual labeling of human lysozyme (HL) most efficiently. The labeled HL was crystallized and analyzed by X-ray diffraction methodology. The X-ray analysis located the two Gal-β1,4-GlcNAc-Epo moieties inside the catalytic cleft of HL. The attachment sites were the side-chain carboxylate groups of the catalytic residues Glu35 and Asp53 in HL. The first Gal-β1,4-GlcNAc-Epo moiety occupied virtually the same position as observed in the HL labeled with single Gal-β1,4-GlcNAc-Epo molecule. The second Gal-β1,4-GlcNAc-Epo moiety was recognized via the carbohydrate−carbohydrate interaction with the first Gal-β1,4-GlcNAc-Epo moiety in addition to the protein-carbohydrate interaction with the "right-side" catalytic cleft of HL through a number of hydrogen bonds including water-mediated ones as well as many van der Waals contacts. The two *N*-acetylglucosamine residues stacked with each other, while the two rings of galactose residues approximately shared the same plane. The dual labeling with two Gal-β1,4-GlcNAc-Epo molecules was supposed to have occurred sequentially, which was accompanied with the alteration to the pK_a of Glu35 derived from the esterification of Asp53 in the first labeling. Both asymmetric carbons in the connection parts between HL and *N*-acetyllactosamine moieties showed the same stereoconfiguration derived from the reaction with (2′*R*) stereoisomer concerning the epoxide group in the labeling reagent. The results demonstrated that the HL labeled with single Gal-β1,4-GlcNAc-Epo was functional as a novel *N*-acetyllactosamine-binding protein, and the second labeling was performed by way of the first-ligand assisted recognition of the second ligand.

A variety of important biological processes depend on the fidelity of the recognition of carbohydrate by proteins (*1*, *2*). Vertebrate or chicken-type (c-type) lysozyme (1,4-β-*N*-acetylmuramidase, EC 3.2.1.17) is a bacteriolytic enzyme existing in widespread biological body fluids. Since the determination of tertiary structure of hen egg-white lysozyme (HEWL) (*3*) and its complex with *N*-acetylglucosamine derivatives (*4*) as the first enzyme molecule, c-type lysozymes have been playing a crucial role as model carbohydrate-binding proteins in the studies on interrelationships between the three-dimensional structure and the enzymatic function (*5*−*7*). Human lysozyme (HL) is one of the most well-characterized c-type lysozymes in terms of both structural and functional aspects. The tertiary structure of HL cocrystallized with hexa-*N*-acetylchitohexaose (*8*) as well as the structure of HL without a ligand (*9*) have been solved by X-ray diffraction methodology. The structural and functional analyses of its site-directed mutants concerning the residues essential for substrate binding have been performed (*10*, *11*).

Approaches using affinity labeling greatly contribute not only to the understanding of catalytic and substrate-binding mechanisms but also to the engineering of glycosidases with altered characteristics (*12*). For such purposes, several types of labeling reagents have been developed by the chemical modification of sugars with the specific affinity to each enzyme. Epoxyalkyl glycosides are the potential irreversible, mechanism-based inhibitors, which have been used for the investigations of the active-site residues of many glycosidases (*12*, *13*). Despite the frequent use of epoxyalkyl glycosides as irreversible inhibitors, only a few kinds of β-glycosidases in complex with an epoxide−saccharide inhibitor have been structurally characterized to date (*14*, *15*). The labeling reagent has been proved to be covalently attached to a carboxylate-group side chain of the targeted enzyme in all instances. The attachment site of the affinity-labeling reagent was first identified as to the HEWL labeled with 2′,3′-

[†]This work was supported by a grant from the Agency of Industrial Science and Technology, MITI, Japan.

[‡] Coordinates have been deposited in the Brookhaven Protein Data Bank (file name 1RE2).

* Author to whom correspondence should be addressed. Phone: 81 298-54-6193. Fax: 81 298-54-6194: E-mail address: muraki@nibh.go.jp.

10.1021/bi981779g CCC: $18.00
Published on Web 12/18/1998

epoxypropyl β-glycoside of *N*-acetylglucosamine by means of amino acid sequence analysis (*16*) and X-ray crystallography (*17*). The employment of X-ray structural analysis has been extended to the discrimination between *R* and *S* stereoisomers at the chiral center of the attachment site by assigning the partial occupancy for each isomer (*18*), and also to the fine characterization of substrate-binding cleft using the labeling reagents with three different alkyl-chain lengths (*19*). The reaction mechanism for the covalent attachment of the epoxyalkyl inhibitor at the active site has been under discussion. The kinetic analysis of the inactivation of *Bacillus subtilis* 1,3;1,4-β-D-glucan-4-glucanohydroase by 3,4-epoxybutyl β-D-cellobioside suggested that the reaction started with the donation of a proton from an acidic residue to the epoxide oxygen atom, and then the attack by another nucleophilic carboxylate-anion bearing residue followed (*20*). In contrast, the putative carboxylic acid−base catalyst reacted with the epoxide compound in the inactivation of *Trichoderma reesei* endo-1,4-xylanase II by 3,4-epoxybutyl β-D-xyloside (*19*).

Recently, we applied the affinity labeling to probe the origin of carbohydrate recognition specificity of human lysozyme using 2′,3′-epoxypropyl β-glycosides of three kinds of disaccharides, GlcNAc-β1,4-GlcNAc,[1] Gal-β1,4-GlcNAc and Man-β1,4-GlcNAc. The structural analyses of the products indicated that all disaccharides in the labeled HLs occupied essentially the same position designated as subsites B and C (*21*, *22*). The lysozyme−saccharide complex has been extensively studied by X-ray crystallography (*6*). However, no example showing multiple ligands interacting with each other in the binding site has been reported. In the present study, we isolated and crystallized the HL labeled with two Gal-β1,4-GlcNAc-Epo. X-ray analysis revealed the presence of both Gal-β1,4-GlcNAc-Epo moieties inside the catalytic cleft, which demonstrated the recognition of the second ligand via the carbohydrate−carbohydrate interaction in addition to the protein−carbohydrate interaction.

MATERIALS AND METHODS

Reaction of Lysozyme with the Affinity-Labeling Reagents. All materials and the reaction condition used for the labeling in analytical-scale experiments were as described in the previous papers (*21*, *22*). An aliquot of the reaction mixture was analyzed by a cation-exchange chromatography (Mono S 5/5, Pharmacia) with a linear gradient of 0 to 0.5 M NaCl in 25 mM sodium phosphate buffer (pH 7.0). For the large-scale preparation of the HL doubly labeled with Gal-β1,4-GlcNAc-Epo, 0.9 mM HL (total 25 mg) was incubated with 75 mM Gal-β1,4-GlcNAc-Epo at 30 °C, for 72 h. After the reaction, the reaction mixture was immediately diluted with 50 vol of ice-cold 25 mM sodium phosphate buffer (pH 7.0) and separated with the same gradient condition as performed in the analytical-scale experiments using a preparative cation-exchange HPLC column (Mono S 10/10, Pharmacia). The peak fraction eluted at 0.35 M NaCl (peak 2 in Figure 1) was collected. The pooled fraction was dialyzed against 0.13 mM sodium acetate buffer (pH 4.5) at 4 °C and then lyophilized. For the purpose of measuring molecular weight, a portion of the peak fraction (peaks 1 and 2 in Figure 1) was desalted with a reversed-phase column chromatography conducted on proRPC 5/2 (Pharmacia) and subjected to a MALDI-TOF/MS analysis (Voyager, PerSeptive Biosystems). α-Cyano-4-hydroxycinnamic acid was used as the matrix.

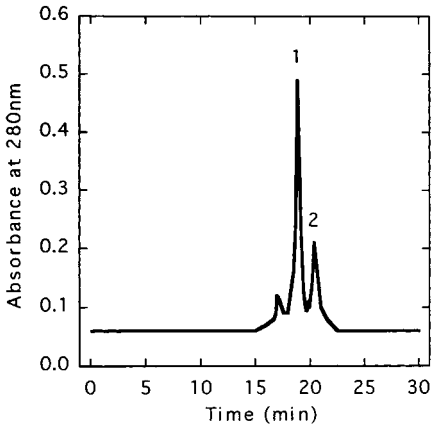

FIGURE 1: Cation-exchange HPLC elution profile of the reaction mixture of HL with Gal-β1,4-GlcNAc-Epo. The linear gradient consisted of 0−50% NaCl in 25 mM sodium phosphate buffer (pH 7.0) for 30 min at a flow rate of 4.0 mL/min. The reaction condition was described in the text.

Crystallization and Structure Determination. The lyophilized HL labeled with two Gal-β1,4-GlcNAc-Epo was dissolved in 3 M ammonium nitrate (pH 4.5, in 25 mM sodium acetate buffer) at a protein concentration of 20 mg/mL. The sample solution was divided into 50 μL portions and equilibrated with 5 M ammonium nitrate solution. First seeding with a native HL crystal gave small crystals of doubly labeled HL, which were used for the repeated seeding to give prisms with the dimensions of 0.1 × 0.1 × 0.3 mm at the most. The space group of the crystal was $P2_12_12_1$ with $Z = 4$ ($V_m = 1.86$) and cell constants of $a = 56.45$ Å, $b = 62.30$ Å, $c = 32.98$ Å. The crystal of the HL labeled with two Gal-β1,4-GlcNAc-Epo was isomorphous with the crystal of native HL (*23*). However, in contrast to the HL labeled with single Gal-β1,4-GlcNAc-Epo (*21*), the crystal cell had a slightly smaller *a*-axis length and a slightly larger *b*-axis length compared with native HL. The diffraction data was collected on an automated oscillation-camera system, R-AXIS IIc (Rigaku), equipped with an imaging-plate detector and a Rotaflex FR rotating-anode generator (40 kV, 40 mA). The data were processed with the R-AXIS IIc data-processing software package (*24*). A total of 23 021 reflections were collected and reduced to a data set of 5742 independent reflections (96.8% of the independent reflections to 2.25 Å resolution) with a merging *R*-factor of 0.111 on intensities.

A set of coordinates of the HL labeled with single Gal-β1,4-GlcNAc-Epo (Protein Data Bank, entry code 1REY) with solvent molecules omitted was used as an initial model for the structural determination. Crystallographic refinement was carried out using X-PLOR (*25*) alternating with manual intervention using TURBO-FRODO (*26*). The first cycle of standard slow-cooling protocol (*27*) with a starting temperature of 3000 K reduced the initial *R*-value of 0.498 to 0.233

[1] Abbreviations: Gal-β1,4-GlcNAc-Epo, 2′, 3′-epoxypropyl *O*-β-D-galactopyranosyl-(1→4)-β-*N*-acetyl-D-glucosamine; GlcNAc-β1,4-GlcNAc-Epo, 2′, 3′-epoxypropyl *O*-β-*N*-acetyl-D-glucosaminyl-(1→4)-β-*N*-acetyl-D-glucosamine; Man-β1,4-GlcNAc-Epo, 2′, 3′-epoxypropyl *O*-β-D-mannopyranosyl-(1→4)-β-*N*-acetyl-D-glucosamine; MALDI-TOF/MS, matrix-assisted laser desorption ionization − time-of-flight mass spectrometry; HPLC, high-performance liquid chromatography.

against the 8.0 Å to 2.5 Å resolution data. At this stage, the first Gal-β1,4-GlcNAc-Epo moiety attached to Asp53 of HL was unambiguously identified as the well-defined electron density in the ($2F_o - F_c$) difference map, and the ($F_o - F_c$) map showed the presence of almost half of the second Gal-β1,4-GlcNAc-Epo moiety with the best-defined electron-density at the connection region to Glu35 of HL. The resolution was gradually increased to 2.3 Å, and the structural components of the second Gal-β1,4-GlcNAc-Epo were very carefully added to the model after examining ($F_o - F_c$) and ($2F_o - F_c$) electron density maps computed at various stages during the refinement. At the final stage, 10% of the observed data was set aside for cross validation analysis (*28*). Water molecules which could form at least one hydrogen bond with an atom of the model structure were included only if the final B-factor was less than 50 Å^2 in the final model. The refinement converged at the *R*-value of 0.171 (R_{free}, 0.247) for 4223 reflections (78.6% completeness) with $|F_o| > 3\sigma(F)$ in the resolution range of 8.0 to 2.3 Å. The coordinate error was estimated to be less than 0.25 Å according to the method of Luzzati (*29*). All stereochemical parameters of the final model checked by PROCHECK (*30*) were either within or better than the average value ranges obtained from good-quality models. The root-mean-square deviations of bond distances and angles from ideality were 0.011 Å and 1.73 °, respectively. The refined atomic coordinates and the structure factors have been deposited in Protein Data Bank (*31*).

RESULTS

Affinity Labeling of Lysozymes. The dual labeling of HL was detected during the analytical-scale experiments using a cation-exchange HPLC by chance. Initially, the affinity labeling of HL and HEWL was examined under two different concentrations of Gal-β1,4-GlcNAc-Epo at 37 °C, for 48 h. Under the lower concentration (2.2 mM), the chromatogram of the reaction mixture was essentially composed of a single peak with regard to both HL and HEWL. Whereas, under the higher concentration (6.6 mM), a new but still small second peak (about five percent of peak area compared to the main peak) appeared after the elution position of the original main peak with HL. The shift of the elution position to the higher salt concentration corresponded to the loss of a negative charge (*32*), which was caused by the additional attachment of the labeling reagent to a dissociated carboxylate group in HL. No such additional peak was observed with HEWL under the same condition. Then, more rigorous reaction condition (75 mM) was applied to HL for the purpose of further pursuing the alteration of elution pattern of the reaction mixture. The second peak was amplified distinctly by the increase of the reagent concentration. The area ratio of second peak (peak 2) to main peak (peak 1) reached to approximately 1:3 after 72 h (Figure 1). The peaks 1 and 2 were fractionated and subjected to the analysis by MALDI-TOF/MS. The partial mass spectrum of peak 2 is shown in Figure 2. The observed molecular weights of peaks 1 and 2 were 15 114.1 and 15 561.6, which corresponded to the calculated molecular weight for the HL singly labeled with Gal-β1,4-GlcNAc-Epo (15 132.1) and the HL doubly labeled with Gal-β1,4-GlcNAc-Epo (15 571.6) within the estimated mass accuracy error (ca. ±0.1%), respectively. To investigate whether the appearance of the second peak is specific to Gal-β1,4-GlcNAc-Epo or not, either GlcNAc-β1,4-GlcNAc-Epo or Man-β1,4-GlcNAc-Epo was examined to react with HL under the same condition (75 mM, 37 °C, 72 h). As for Man-β1,4-GlcNAc-Epo, the corresponding peak but much smaller one (less than a few percent of peak area compared with main peak) appeared, and virtually no corresponding peak was observed in the chromatogram of the reaction mixture with GlcNAc-β1,4-GlcNAc-Epo. This indicated the strongest potential of Gal-β1,4-GlcNAc-Epo among the three kinds of labeling reagents and the ineffeciency of GlcNAc-β1,4-GlcNAc-Epo in the dual labeling of HL.

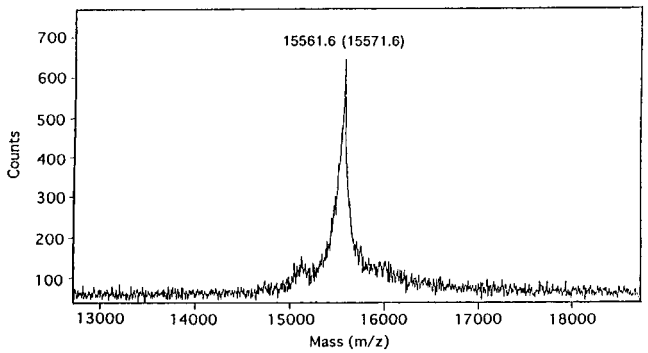

FIGURE 2: MALDI-TOF/MS analysis of the HL labeled with two Gal-β1,4-GlcNAc-Epo. The peak 2 fraction in Figure 1 was analyzed. Calculated value in parentheses.

Structure of the Ligand Part. The three-dimensional structure of the HL doubly labeled with Gal-β1,4-GlcNAc-Epo was determined by X-ray crystallographic analysis. The final model contained the complete HL protein of 130 residues, two Gal-β1,4-GlcNAc-Epo moieties, and 84 water molecules. Figure 3 shows the omit ($F_o - F_c$) electron density maps concerning the bound ligands and the connection regions between protein and ligand (EPO1−Asp53 and EPO2−Glu35). The continuous electron densities confirmed the covalent binding of the labeling reagent to both Asp53 and Glu35 of HL and ensured the conformations of the ligands in the final model. The average *B*-values of atoms of the first ligand and the second ligand were 20.5 Å^2 and 49.0 Å^2, respectively. The side chains of Asp53 and Glu35 slightly altered their conformation through the attachment of Gal-β1,4-GlcNAc-Epo (Figure 4). The root-mean-square displacements of the side-chain atoms from native HL were 1.1 Å and 0.6 Å with Asp53 and Glu35, respectively. The two *N*-acetylglucosamine moieties in the doubly labeled HL located in proximity to each other (Figure 5) with maintaining a lot of close contacts (Table 1). Figure 6 illustrates the conformations of the first ligand (GAL1-NAG1-EPO1) and the second ligand (GAL2-NAG2-EPO2). The two *N*-acetylglucosamine residues and the two galactose residues in the ligand part interacted differently. The two apolar ring planes of the *N*-acetylglucosamine residues stacked with each other, which was mainly conducted by van der Waals interactions between them, whereas the two rings of the galactose residues approximately shared the same plane by interacting via several hydrogen bonds in addition to van der Waals contacts (Figure 6). In Table 2, several stereochemical parameters concerning the glycosidic bond between the *N*-acetylglucosamine residue and the galactose residue in the ligand part were summarized. All the pyranose rings in the ligands took a normal chair conformation. The conformation of the first ligand was very similar to that of the ligand in

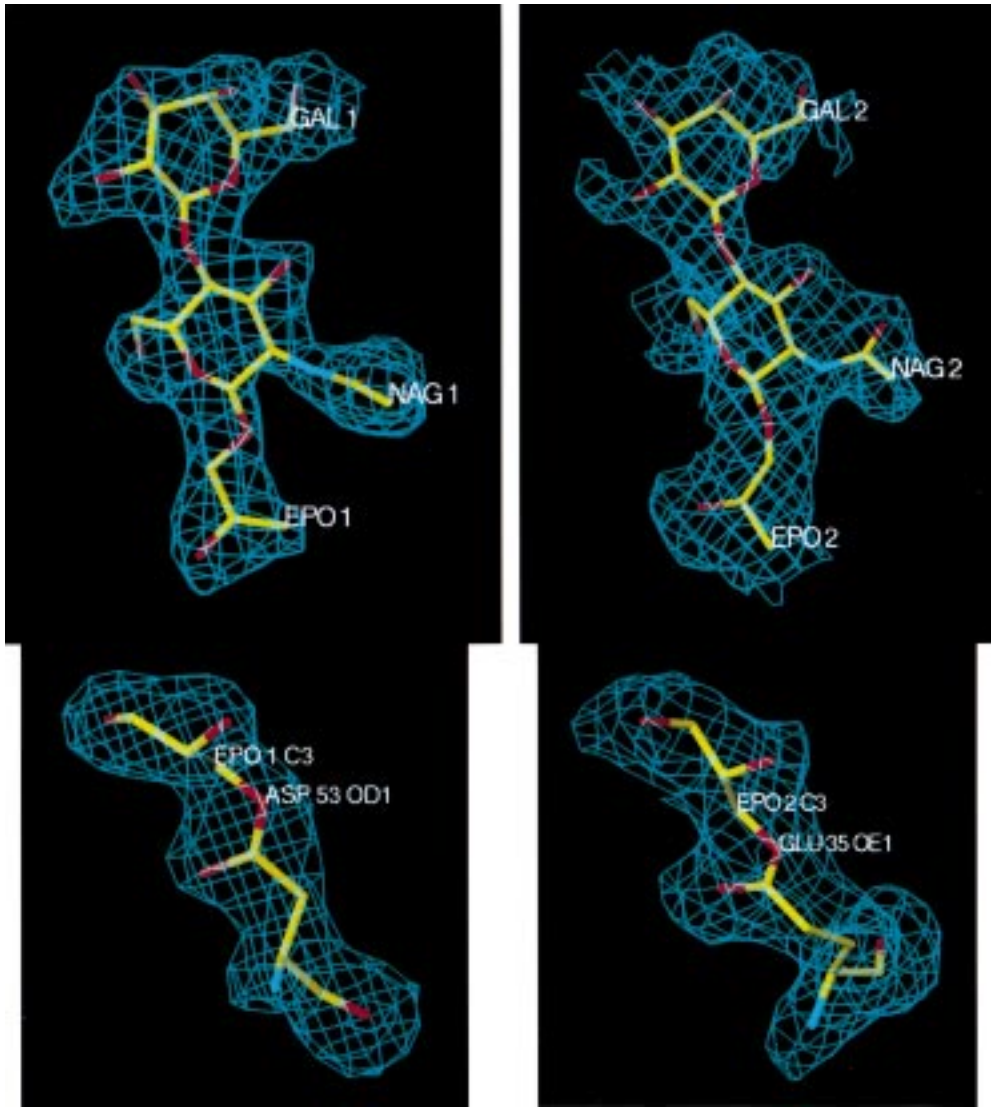

FIGURE 3: ($F_o - F_c$) omit electron density maps for the ligands together with the refined model. Top left, the first ligand; top right, the second ligand; bottom left, EPO1−Asp53 region; bottom right, EPO2-Glu35 region. The maps for the first ligand and the second ligand are contoured at 2.0σ and 1.0σ, respectively.

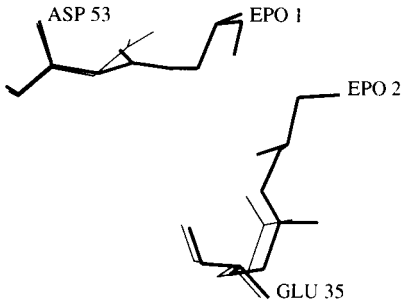

FIGURE 4: Superposition of the connection region between HL and the ligand. Thin lines, native HL; thick lines, the doubly labeled HL.

the singly labeled HL. The linkage torsion angles (ϕ and ψ) of both *N*-acetyllactosamine moieties in the doubly labeled HL coincided with the optimized low-energy conformation region calculated for Gal-β1,4-GlcNAc linkage in histo-blood group oligosaccharides such as Le^y and type-2 H antigen (*35*). Interestingly, the C6−O6 bond in both *N*-acetylglucosamine residues in the doubly labeled HL took a nearly cis conformation toward the C5−O5 bond, which was a +gauche conformation in the singly labeled HL (Table 2). The cis conformation itself should be energetically unfavorable compared to +gauche conformation; however, the following hydrogen-bonding interactions were supposed to compensate for the energetic loss (Table 3). The O6 atom in NAG1, which rather exposed to the solvent region, made a direct hydrogen bond to OH atom of Tyr63 and a couple of water-mediated hydrogen bonds to OD2 atom of Asp49 and ND2 atom of Asn46 in HL. The O6 atom in NAG2 interacted with the residues 107−110 via the direct hydrogen bonds, and if the C6−O6 bonds took a +gauche conformation, the O6 atom in NAG2 would collide with the main-chain atoms of Arg107 and Ala108 residues in HL.

Interaction of the Ligands with Protein. Figure 7 shows the structure of the ligand-binding site. Both Gal-β1,4-GlcNAc-Epo moieties were located inside the substrate-binding cleft of HL. As shown in Figure 5, the first ligand occupied essentially the same position where the Gal-β1,4-GlcNAc-Epo moiety in the singly labeled HL was found, which basically corresponded to the subsites B and C of HL. On the other hand, *N*-acetylglucosamine residue in the second ligand (NAG2) squeezed into the vacant space surrounded by the *N*-acetylglucosamine residue in the first ligand

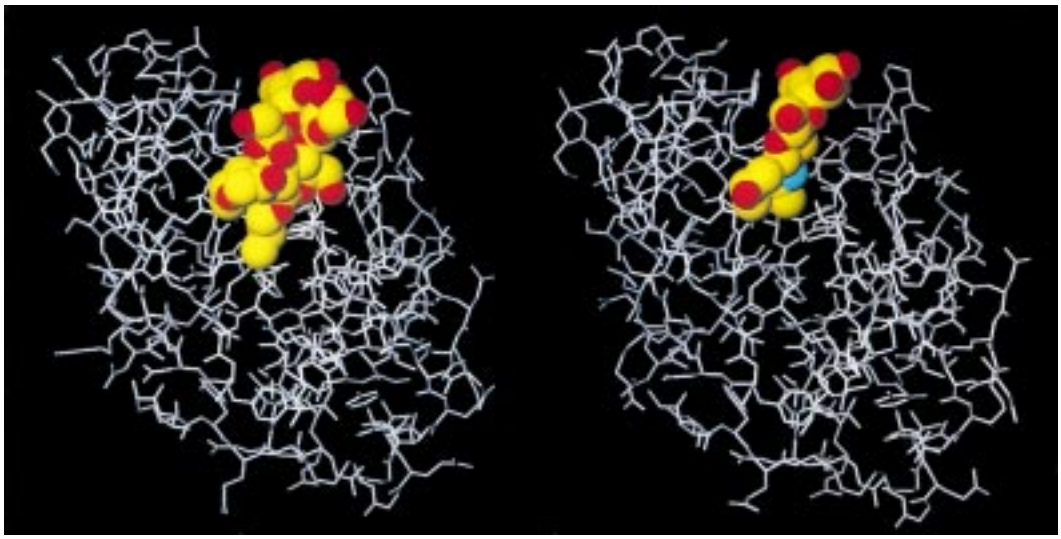

FIGURE 5: Overall structure of HLs labeled with Gal-β1,4-GlcNAc-Epo. Left, the doubly labeled HL; right, the singly labeled HL (*21*). The protein part and the ligand part are drawn in a wire-model and a space-filling model, respectively. Water molecules omitted.

Table 1: Possible Interactions between the Two *N*-Acetyllactosamine Moieties in the Ligand Part

carbohydrate atom 1	carbohydrate atom 2	distance (Å)
(a) Hydrogen-Bonding Interactions[a]		
GAL1 O2	GAL2 O2	2.7
GAL1 O2	GAL2 O3	2.4
GAL1 O3	GAL2 O3	3.4
NAG1 O4	GAL2 O2	2.4
EPO1 O2	NAG2 N2	3.1
(b) van der Waals Interactions[b]		
GAL1 C1	GAL2 O2	3.2
GAL1 O1	GAL2 O1	3.9
GAL1 O1	GAL2 C2	3.4
GAL1 C2	GAL2 C2	3.5
GAL1 C2	GAL2 O2	2.8
GAL1 C2	GAL2 C3	3.9
GAL1 C2	GAL2 O3	3.2
GAL1 O2	GAL2 C2	2.7
GAL1 O2	GAL2 C3	3.1
GAL1 O2	GAL2 C4	4.0
GAL1 C3	GAL2 O3	3.9
NAG1 C1	NAG2 C1	3.8
NAG1 C1	NAG2 C3	3.9
NAG1 C1	NAG2 C5	3.9
NAG1 O1	NAG2 C1	3.9
NAG1 C4	GAL2 O2	3.6
NAG1 C5	GAL2 O2	3.8
NAG1 C5	NAG2 C3	3.8
NAG1 C5	NAG2 C4	4.0
NAG1 C5	NAG2 O4	3.3
NAG1 O5	NAG2 C3	3.8
NAG1 C6	GAL2 O1	3.8

[a] The interactions within the distance of 3.5 Å are listed. [b] The interactions within the distance of 4.0 Å are listed.

(NAG1) and the residues 107−110 of HL, which made up a part of "right-side" cleft lobe around the subsites C and D. The galactose residue in the second ligand (GAL2) relatively protruded to the outside of the cleft while keeping the contacts with the galactose residue in the first ligand (GAL1) and Gln104 of HL which formed a part of the "right-side" cleft lobe around the subsite B.

In Table 3, the possible ligand−protein hydrogen-bonding interactions in the HL labeled with two Gal-β1,4-GlcNAc-Epo were summarized. All major direct and water-mediated hydrogen-bonding interactions in the singly labeled HL (*21*)

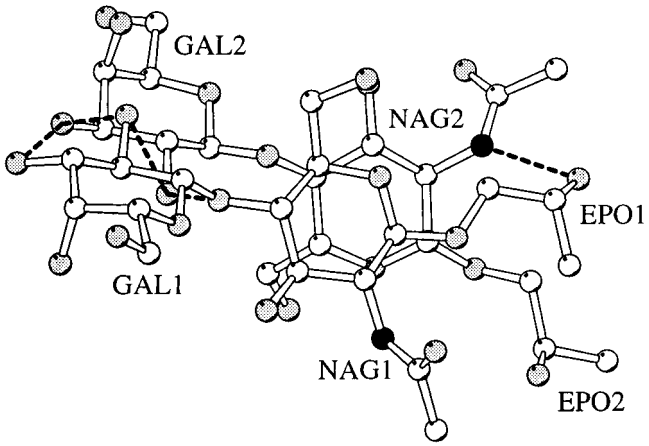

FIGURE 6: Ball and stick representation of the Gal-β1,4-GlcNAc-Epo moieties in the doubly labeled HL. Carbon atoms, oxygen atoms and nitrogen atoms are shown in white, gray and black, respectively. The broken lines indicate the hydrogen-bonding interactions. This figure was drawn with MOLSCRIPT (*33*).

Table 2: Stereochemical Parameters of Disaccharides[a]

linkage	ϕ (deg)	ψ (deg)	ψ_H (deg)	O5−O3′ (Å)	C6−O6 bond[b] GAL	C6−O6 bond[b] NAG
GAL1−NAG1	−67	−116	58	3.0	trans	nearly cis
GAL2−NAG2	−76	−95	68	3.6	trans	nearly cis
GAL131−NAG132[c]	−69	−119	53	3.0	trans	+gauche
GAL−NAG[d]	−73	−102	nd[e]	nd	nd	nd

[a] Values of ϕ and ψ are the torsion angles about C1−O1(O4′) and O1(O4′)−C4 defined by O5−C1−O1(O4′)−C4′ and C1−O1(O4′)−C4′−C5′, respectively. The value of ψ_H is the helical twist parameter defined as the average of the pseudorotation angles ψ_1 = O5−C1−C4′−C3′ and ψ_2 = C2−C1−C4′−C5′ (*34*). [b] The conformation toward the C5−O5 bond. [c] Values are those in the structure determined by Muraki et al. (*21*). [d] The lowest-energy conformation reported for Ley by Imberty et al. (*35*). [e] nd, not described.

were also found in the doubly labeled HL. In addition to the conserved hydrogen bonds, three new direct and six new water-mediated hydrogen-bonding interactions with protein part were detected in the first *N*-acetyllactosamine part (GAL1-NAG1). Six water molecules were identified as the bridging water between ligand atoms and protein atoms (Figure 7). Each oxygen atom of the water molecules showed

Table 3: Possible Ligand−Protein Hydrogen-Bonding Interactions in the HL Labeled with Two Gal-β1,4-GlcNAc-Epo[a]

(a) Direct Ligand-Protein Interactions

ligand atom	protein atom	distance (Å)
GAL1 O4	Asp102 OD2	2.5 *
NAG1 N2	Ala108 O	3.0 *
NAG1 O3	Trp64 NE1	3.2 *
NAG1 O4	Gln104 NE2	3.1
NAG1 O6	Tyr63 OH	3.0
NAG1 O7	Trp64 NE1	2.9
NAG1 O7	Asn60 N	2.8 *
EPO1 O2	Asp53 OD1	2.9
GAL2 O2	Gln104 NE2	2.5
GAL2 O2	Gln104 OE1	2.7
NAG2 O5	Val110 N	2.9
NAG2 O6	Arg107 O	2.2
NAG2 O6	Ala108 O	3.5
NAG2 O6	Trp109 N	3.1
NAG2 O6	Val110 N	3.2
EPO2 O2	Glu35 OE1	2.8
EPO2 O2	Glu35 OE2	3.5
EPO2 O2	Gln58 O	3.5
EPO2 O1	Val110 N	2.8

(b) Water-Mediated Ligand-Protein Interactions

ligand atom	first water atom	second water atom	protein atom	distance (Å)[b]
GAL1 O3	Wat8 O		Gln104 N	3.2/3.0 *
GAL1 O4	Wat8 O		Gln104 N	3.2/3.0 *
GAL1 O4	Wat83 O		Gln104 NE2	2.7/2.6 *
GAL1 O4	Wat83 O	Wat9 O	Val99 O	2.7/2.7/2.8
GAL1 O4	Wat83 O	Wat9 O	Gly105 N	2.7/2.7/3.5
GAL1 O4	Wat83 O	Wat9 O	Gly105 O	2.7/2.7/2.9
GAL1 O5	Wat83 O		Gln104 NE2	3.1/2.6 *
NAG1 O3	Wat83 O		Gln104 NE2	3.0/2.6
NAG1 O6	Wat38 O		Asp49 OD2	2.7/2.6
NAG1 O6	Wat38 O	Wat50 O	Asn46 ND2	2.7/3.0/3.0
NAG2 O6	Wat55 O		Arg107 O	3.5/3.2

[a] The interactions within the distance of 3.5 Å are listed. [b] The distances of the water-mediated hydrogen bonds are shown as the distances of (ligand atom−first water atom)/(first water atom−protein atom) or (ligand atom−first water atom)/(first water atom−second water atom)/(second water atom−protein atom). The asterisks indicate the presence of the same interaction in the HL labeled with a single Gal-β1,4-GlcNAc-Epo.

a *B*-value comparable to the other structural atoms as a fully occupied atom (Wat8 O, 21.7 Å^2; Wat9 O, 10.6 Å^2; Wat38 O, 26.3 Å^2; Wat50 O, 39.8 Å^2; Wat55 O, 35.3 Å^2; Wat83 O, 31.4 Å^2, each denoted in Table 3b). The *N*-acetylglucosamine residues (NAG1 and NAG2) were recognized mainly by direct ligand−protein interactions due to the relatively secluded position from solvent and the full packing around the subsites C and D with the ligands. On the other hand, GAL1 was suggested to form seven water-mediated ligand−protein hydrogen bonds in addition to the direct hydrogen bond between the axial 4-OH group of GAL1 and OD2 atom of Asp102 in HL (Table 3). As described above, the recognition of GAL2 was largely conducted by the interaction with the first ligand (Table 1); however, the side-chain amide group of Gln104 in HL was also responsible for the recognition through a couple of hydrogen bonds with the 2-OH group of GAL2 (Table 3). The formation of hydrogen bonds between the 2-OH groups of EPO1 and EPO2 and the side-chain oxygen atom of either Asp53 or Glu35 might contribute to stiffen the connection region between the protein part and the ligand part.

Change in the Main-Chain Structure. As shown in Figure 5, the overall structure of the doubly labeled HL was very similar to that of the singly labeled HL. The least-squares superposition indicated that the average displacement value of the main-chain atoms was 0.32 Å. However, some additional conformational changes in the backbone structure from native HL were observed with the doubly labeled HL compared to the singly labeled HL. A fairly large displacement of main-chain atoms in "lower right-side" cleft lobe in the region of residues 118−123 was observed with the doubly labeled HL in addition to the shift in the region of residues 101−103 composing the "upper right-side" cleft lobe of HL. The latter shift was the most significant conformational change with the singly labeled HL (*21*). These conformational changes made the substrate-binding cleft of doubly labeled HL narrower as compared with native HL. The maximum C^{α}-atom deviation for each region was 1.3 Å at Asp120 for the "lower right-side" cleft region and 1.5 Å at Pro103 for the "upper right-side" cleft region, respectively. Another characteristic of the conformational change in HL accompanied with the dual labeling was the displacement of residues 46−49 composing the "upper left-side" cleft lobe with the maximum displacement of 1.0 Å at Gly 48. The shift slightly widened the substrate binding cleft, which was not observed in the single labeling of HL with Gal-β1,4-GlcNAc-Epo. A larger shift of this region has been observed in the labeling of HL with Man-β1,4-GlcNAc-Epo (*22*).

DISCUSSION

Our previous study demonstrated that not only GlcNAc-β1,4-GlcNAc-Epo but also Gal-β1,4-GlcNAc-Epo have the potential to label HL via a covalent attachment to the side-chain carboxylate group of Asp53 (*21*). The inactivation ability of Gal-β1,4-GlcNAc-Epo against the lytic activity of HL was much less than that of GlcNAc-β1,4-GlcNAc-Epo, which reflected the weaker affinity of Gal-β1,4-GlcNAc compared with GlcNAc-β1,4-GlcNAc toward native HL. Nevertheless, the dual attachment of the affinity-labeling reagent to HL was performed much more effectively by Gal-β1,4-GlcNAc-Epo compared with GlcNAc-β1,4-GlcNAc-Epo. This suggested that the second labeling was conducted through the specific recognition of the second Gal-β1,4-GlcNAc-Epo by the singly labeled HL. The recognition mechanism was revealed by X-ray crystallographic analysis of the three-dimensional structure of the HL labeled with two Gal-β1,4-GlcNAc-Epo. The comparison of the X-ray structures indicated that the basic conformation of the ligand and all major hydrogen-bonding interactions between the ligand and the protein in the singly labeled HL were unchanged in the doubly labeled HL (Table 2 and Table 3). The conservative behavior of the first ligand during the incorporation of the second ligand suggested that the first labeling produced the suitable pocket for the recognition of the second ligand. In other words, HL conjugated with a Gal-β1,4-GlcNAc-Epo moiety at Asp53 was functional as the novel carbohydrate-binding protein with the affinity for *N*-acetyllactosamine. The already existed *N*-acetyllactosamine moiety played an essential role in the second recognition step both by forming specific hydrogen bonds to the second *N*-acetyllactosamine moiety and by making a lot of attractive van der Waals contacts (Table 1).

Hydrogen-bonding interaction is a major factor determining the specificity and the affinity of carbohydrate-binding

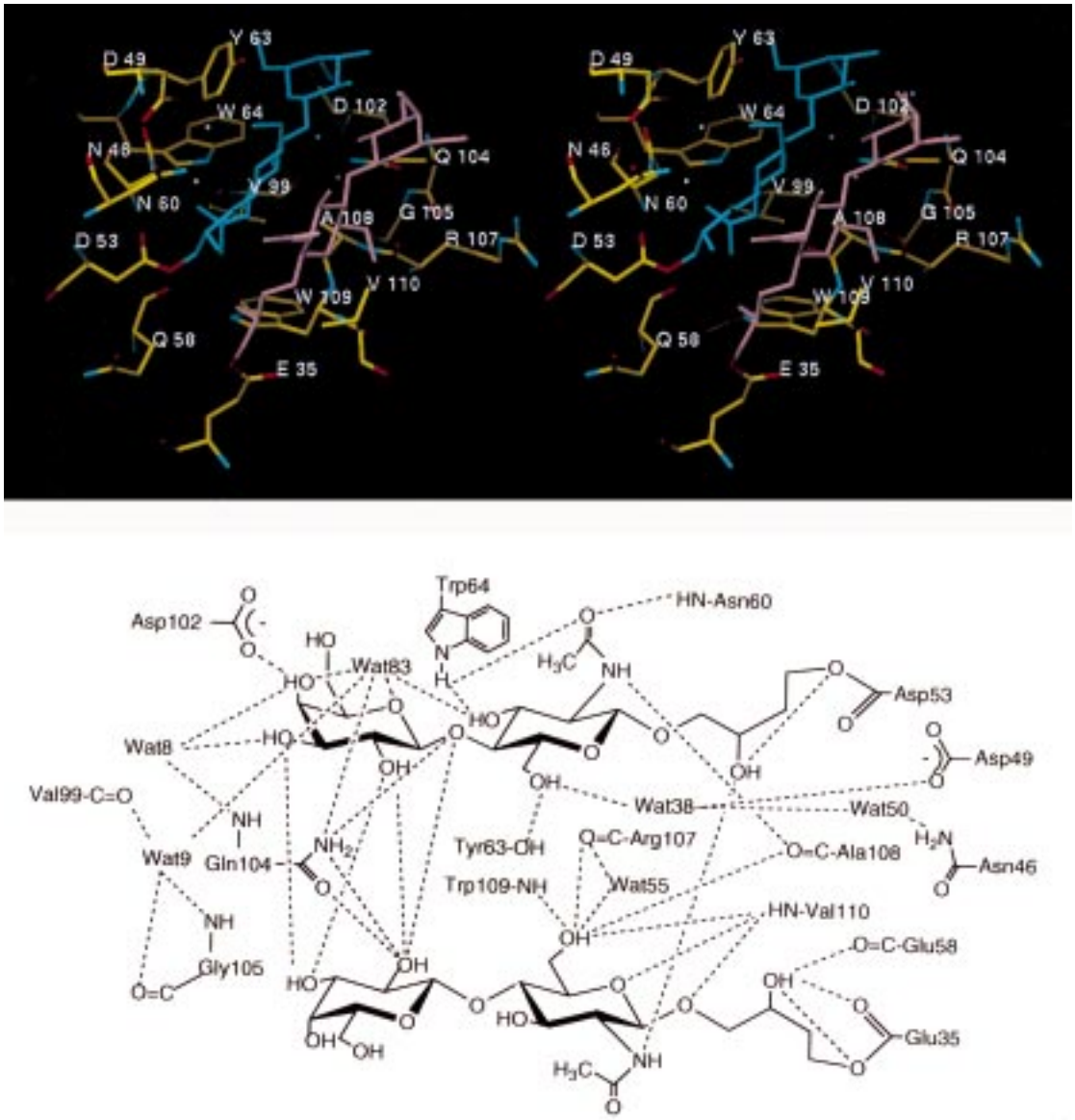

FIGURE 7: Ligand-binding structure in the active site of the doubly labeled HL. The amino acid residues and water molecules involved in the hydrogen-bonding interactions with the ligands are included. (top) Stereoview. The protein part, the first ligand part, the second ligand part and water molecules are drawn with atom colored wires, blue wires, pink wires and white balls, respectively. (bottom) Schematic drawing. The possible hydrogen-bonds are shown in broken lines.

proteins (*1*, *2*). The 2-OH group of galactose residue was a key structural element in the direct hydrogen-bonding interactions concerning the recognition of the second ligand (Tables 1 and 3). Therefore, the replacement of the equatorial 2-OH group in galactose residue with either the bulkier *N*-acetylamino group of *N*-acetylglucosamine residue or the axial OH group of mannose residue should damage the affinity toward the second ligand, which accounted for the predominance of Gal-β1,4-GlcNAc-Epo over GlcNAc-β1,4-GlcNAc-Epo or Man-β1,4-GlcNAc-Epo in dual labeling. The localized conformational change induced by the oligosaccharide binding via direct and water-mediated hydrogen bonds is a common feature in the recognition events involving carbohydrate-binding proteins (*36*−*38*). Several additional direct and water-mediated hydrogen-bonding interactions between the first *N*-acetyllactosamine (GAL1-NAG1) moiety and the protein were observed in the doubly labeled HL compared with the corresponding part in the singly labeled HL (Table 3). This indicated that the binding of the second Gal-β1,4-GlcNAc-Epo moiety refined the conformation of the protein together with that of the first *N*-acetyllactosamine moiety in the singly labeled HL to the more energetically stabilized form in the doubly labeled HL. The conformational changes of the structural part involved in the additional hydrogen-bonding interactions are shown in Figure 8.

If the assistance of recognition by the first ligand alone is taken into account, there might be an alternative way for the dual labeling, in which the first attachment of affinity label occurred to the carboxylate group of Glu35 instead of Asp53. However, this mechanism would not be probable under the pH condition (pH 5.4) used in this study. The two catalytic carboxyl groups in c-type lysozymes (e.g., Glu35 and Asp53 in HL and Glu35 and Asp52 in HEWL) have quite different pK_a values. The disagreement in pK_a has been ascribed to the electrostatic interaction between the two carboxyl groups as well as the hydrophobic environment around Glu35 (*39*, *40*). The pK_a value of Glu35 of HL in

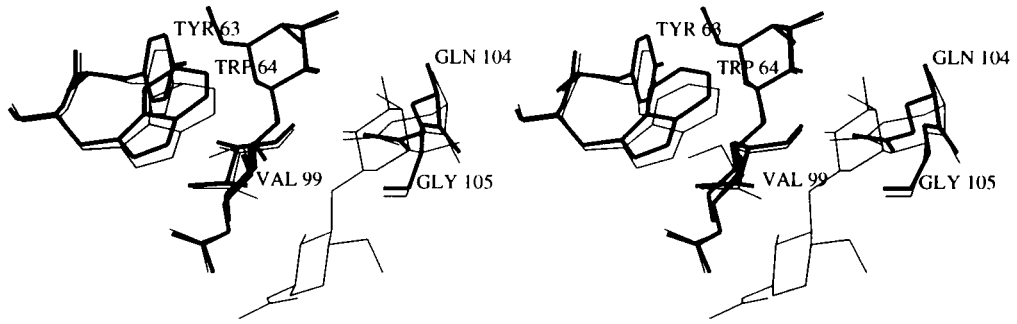

FIGURE 8: Stereoview of the conformational changes on labeling with the second Gal-β1,4-GlcNAc-Epo to the singly labeled HL. The amino acid residues involved in the additional hydrogen bonding interactions are shown. Thick lines, the singly labeled HL; thin lines, the doubly labeled HL.

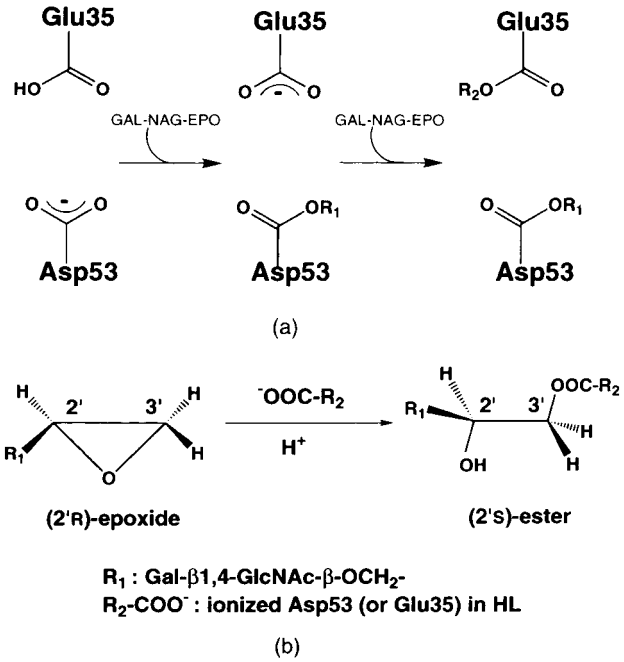

FIGURE 9: Labeling reaction of HL with Gal-β1,4-GlcNAc-Epo. (a) Reaction scheme. R_1 and R_2 indicate Gal-β1,4-GlcNAc-β-OCH_2-$CH(OH)CH_2$ residues. (b) Stereochemistry of the reaction.

the uncomplexed state is determined as 6.8, which was abnormally higher than that of Asp53 (3.4) (*41*). The difference would make most Glu35 remain as neutral carboxylic acid form and most Asp53 as negatively charged carboxylate anion form under pH 5.4. Epoxides react readily with ionized carboxylate anions, but not with undissociated carboxylic acids (*42*). The reaction is suggested to proceed via a bimolecular S_N2 mechanism. Supportingly, both carboxylate groups in this study were bound to the sterically less hindered methylene carbon (C3) in the epoxide group (Figure 3). The esterification of Asp52 in HEWL with GlcNAc-β1,4-GlcNAc-Epo alters the pK_a of Glu35 from 6.1 to 5.2 (*39*). The pK_a of Glu35 in HL is calculated as 4.4 when the ionization of Asp53 was blocked (*41*). These results suggested that once Asp53 in HL was esterified with the first Gal-β1,4-GlcNAc-Epo, considerable part of Glu35 turned to the ionized state, which made Glu35 be ready to react with the second Gal-β1,4-GlcNAc-Epo. Consequently, we propose the sequential reaction scheme shown in Figure 9a.

It is reported that the 1,3-1,4-β-D-glucan 4-Glucanohydrolases from *B. subtilis* and *barley* (*Hordeum vulgare*) are inactivated predominantly by the optically pure (3*S*)- and (3*R*)-3,4-epoxybutyl β-cellobioside, respectively (*43*). The affinity-labeling reagent, Gal-β1,4-GlcNAc-Epo, used in this study was a mixture of the stereoisomers (2′*R* and 2′*S*) concerning the asymmetric carbon in the epoxide group. The configurations of both asymmetric carbons in the connection regions between protein and saccharide were modeled as the same (2′*S*) configuration from the shape of the electron density of this region (Figure 3). The (2′*R*) epoxide isomer was suggested to be used predominantly in the esterification of both Asp53 and Glu35, since the (2′*R*) epoxide produces the (2′*S*) ester in the labeling reaction (Figure 9b).

In conclusion, the present study has provided a rational mechanism for the dual labeling of the active site of HL with Gal-β1,4-GlcNAc-Epo. The results also suggested the possibility of creating a novel carbohydrate-binding protein which shows the different recognition specificity from that of the original protein using the preconjugation of a saccharide moiety.

ACKNOWLEDGMENT

We acknowledge Mr. Toshiaki Hiraki (Hokkaido System Science Co., Ltd.) for his assistance in MALDI-TOF/MS analyses.

REFERENCES

1. Quiocho, F. A. (1986) *Ann. Rev. Biochem. 55*, 287−315.
2. Vyas, N. K. (1991) *Curr. Opin. Struct. Biol. 1*, 732−740.
3. Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., and Sarma, V. R. (1965) *Nature (London) 206*, 757−761.
4. Johnson, L. N., and Phillips, D. C. (1965) *Nature (London) 206*, 761−763.
5. Jollès, P., and Jollès, J. (1984) *Mol. Cell. Biochem. 63*, 165−189.
6. Strynadka, N. C. J., and James, M. N. G. (1996) In *EXS* (Jollès, P., Ed.) Vol. 75, pp 185−222, Birkhauser Verlag, Basel.
7. Imoto, T. (1996) in *EXS* (Jollès, P., Ed.) Vol. 75, pp 163−181, Birkhauser Verlag, Basel.
8. Song, H., Inaka, K., Maenaka, K., and Matsushima M. (1994) *J. Mol. Biol. 244*, 522−540.
9. Artymiuk, P. J., and Blake, C. C. F. (1981) *J. Mol. Biol. 152*, 737−762.
10. Muraki, M., Harata, K., and Jigami, Y. (1992) *Biochemistry 31*, 9212−9219.
11. Harata, K., Muraki, M., and Jigami, Y. (1993) *J. Mol. Biol. 233*, 524−535.
12. Withers, S. G., and Aebersold, R. (1995) *Protein Sci. 4*, 361−37.
13. Mooser, G. (1992) in *The Enzymes* (Sigman, D. S., Ed.) 3rd ed., Vol. 20, pp 187−233.
14. Keitel, T., Simon, O., Borris, R., and Heinemann, U. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5287−5291.

15. White, A., and Rose, D. R. (1997) *Curr. Opin. Struct. Biol. 7*, 645−651.
16. Eshdat, Y., McKelvy, J. F., and Sharon, N. (1973) *J. Biol. Chem. 248*, 5892−5898.
17. Moult, J., Eshdat, Y., and Sharon, N. (1973) *J. Mol. Biol. 75*, 1−4.
18. Sulzenbacher, G., Schülein, M., and Davies, G. J. (1997) *Biochemistry 36*, 5902−5911.
19. Havukainen, R., Törrönen, A., Laitinen, T., and Rouvinen, J. (1996) *Biochemistry 35*, 9617−9624.
20. Høj, P. B., Rodriguez, E. B., Stick, R. V., and Stone B. A. (1989) *J. Biol. Chem. 25*, 4939−4947.
21. Muraki, M., Harata, K., Sugita, N., and Sato, K.-I. (1996) *Biochemistry 35*, 13562−13567.
22. Muraki, M., Harata, K., Sugita, N., and Sato, K.-I. (1998) *Acta Crystallogr. Sect. D 54*, 834−843.
23. Osserman, E. F., Cole, S. J., Swan, I. D. A., and Blake C. C. F. (1969) *J. Mol. Biol. 46*, 211−212.
24. Sato, M., Yamamoto, M., Imada, K., Katsube, Y., Tanaka, N., and Higashi, T. (1992) *J. Appl. Crystallogr. 25*, 348−357.
25. Brünger, A. T. (1992) *X-PLOR* ver. 3.1 Manual, pp 1−269, Yale University Press, New Heaven and London.
26. Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268−272.
27. Brünger, A. T., Krukowski, A., and Erickson, J. W. (1990) *Acta Crystallogr. Sect. A 46*, 585−593.
28. Brünger, A. T. (1992) *Nature (London) 355*, 472−475.
29. Luzzati, V. (1952) *Acta Crystallogr. 5*, 802−810.
30. Laskowski R. A., McArthur M. W., Moss D. S., and Thornton J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
31. Bernstein F. C., Koetzla, T. F., Williams, G. J. B., Meyer, E. F. Jr., Brice M. D., Rodgers, J. R., Kennard O., Shimanouchi, T., and Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.
32. Muraki, M., Morikawa, M., Jigami, Y., and Tanaka, H. (1988) *Protein Eng. 2*, 49−54.
33. Kraulis, P. (1991) *J. Appl. Crystallogr. 11*, 268−272.
34. Mo, F., and Jensen, L. H. (1978) *Acta Crystallogr., Sect. B, 34*, 1562−1569.
35. Imberty, A., Mikros, E., Koća, J., Mollicone, R., Oriol, R., and Pérez, S. (1995) *Glycoconjugate J. 12*, 331−349.
36. Strynadka, N. C. J., and James, M. N. G. (1991) *J. Mol. Biol. 220*, 401−424.
37. Bourne Y., Rougé, P., and Cambillau C. (1992) *J. Biol. Chem. 267*, 197−203.
38. Quiocho, F. A., Spurlino, J. C., and Rodseth, L. E. (1997) *Structure 5*, 997−1015.
39. Parsons, S. M., and Raftery, M. A. (1972) *Biochemistry 11*, 1623−1629.
40. Inoue, M., Yamada, H., Yasukochi T., Kuroki, R., Miki, T., Horiuchi T., and Imoto T. (1992) *Biochemistry 31*, 5545−5553.
41. Kuramitsu, S., Ikeda, K., Hamaguchi, K., Fujio, H., Amano, T. Miwa, S., and Nishina T. (1974) *J. Biochem. (Tokyo) 76*, 671−683.
42. Ross, W. C. J. (1950) *J. Chem. Soc.* 2257−2273.
43. Høj, P. B., Rodriguez, R. B., Iser, J. R., Stick, R. V., and Stone, B. A. (1991) *J. Biol. Chem. 266*, 11628−11631.

BI981779G