*J. Biochem.* **140**, 221–227 (2006)
doi:10.1093/jb/mvj142

# Construction of Enzyme-Substrate Complexes between Hen Egg-White Lysozyme and *N*-Acetyl-D-Glucosamine Hexamer by Systematic Conformational Search and Molecular Dynamics Simulation

**Hideki Hirakawa[1], Yoshihiro Kawahara[2], Atsuko Ochi[2], Shigeru Muta[2], Shunsuke Kawamura[3], Takao Torikata[3] and Satoru Kuhara[1,2,*]**

*[1]Graduate School of Systems Life Sciences and [2]Graduate School of Genetic Resource Technology, Kyushu University, Hakozaki, Higashi-ku, Fukuoka 812-8581; and [3]Department of Bioscience, School of Agriculture, Kyushu Tokai University, Aso, Kumamoto 869-1404*

Received May 11, 2006; accepted June 7, 2006

**We constructed the complexes between HEWL and $(GlcNAc)_6$ oligomer in order to investigate the amino acid residues related to substrate binding in the productive and nonproductive complexes, and the relationship between the distortion of the GlcNAc residue D and the formation of the productive complexes. We obtained 49 HEWL-$(GlcNAc)_6$ complexes by a systematic conformational search and classified the each one to the three binding modes; left side, center, or right side. Furthermore we performed the molecular dynamics simulation against 20 HEWL-$(GlcNAc)_6$ complexes (8: chair model, 12 : half-chair model) selected from the 49 complexes to investigate the interaction between HEWL and $(GlcNAc)_6$. As results, we confirmed that it is necessary for GlcNAc residue D to be half-chaired form to bind toward the right side to form productive complexes. We found newly that eight amino acid residues interact with the $(GlcNAc)_6$ oligomer, as follows, Arg73, Gly102, Asn103 for GlcNAc residue A; Asn103 for GlcNAc residues B and C; Leu56, Ala107, Val109 for GlcNAc residue D; Ala110 for GlcNAc residue E; and Lys33 for GlcNAc residue F. We also indicated that GlcNAc residue F does not interact with Thr47 and rarely interacts with Phe34 and Asn37.**

**Key words: docking, lysozyme, molecular dynamics (MD) simulation, structure of complex, systematic conformational search.**

Abbreviations: HEWL, hen egg-white lysozyme; GlcNAc, *N*-acetyl-D-glucosamine; MurNAc, *N*-acetylmuramic acid; MD, molecular dynamics.

Lysozyme is a well-studied enzyme that hydrolyzes the β-(1,4)-glycosidic linkage of *N*-acetyl-D-glucosamine (GlcNAc) oligomer. Lysozyme also has transglycosylation activity. The X-ray structure of hen egg-white lysozyme (HEWL) revealed a cleft that could accommodate six sugar residues, named subsites A to F. There are two carboxylate groups, Glu35 and Asp52. The position of Glu35 suggested it would act as a general acid, and that of Asp52 suggested it would be a negative counter-ion to a putative glycosyl cation. A cleavage of the sugar residue occurs between subsites D and E (*1*). Here the GlcNAc residues in subsites A to F are called GlcNAc residues A to F, respectively.

The widely known "Phillips mechanism" has been proposed as a lysozyme catalytic mechanism (*1*). It has been considered that, in the reaction mechanism, GlcNAc residue D would be distorted so as to allow the hexasaccharide to bind with the correct orientation and to allow the effective hydrolysis of a productive complex.

A previous study determined the amino acid residues that participate in substrate binding. According to the X-ray structure of HEWL-$(GlcNAc)_3$ complex, Asp101 and Trp62 participate in substrate binding at subsites A and B, and Asn59, Trp62, Trp63, and Ala107 participate in substrate binding at subsite C. However, the residues participate in substrate binding at subsites D to F were only estimated using the model building (*2*). Scheraga *et al.* constructed the HEWL-$(GlcNAc)_6$ complexes using conformational energy calculations and predicted that there were two distinct binding modes at subsites E and F: a left-sided binding mode and a right-sided one. The former involves such residues as Arg45, Asn46, and Thr47, while the latter involves such residues as Asn37 and Arg114. However, in this treatment, the enzyme structure was held rigidly, and the substrate was allowed to move only within the region of the active site and to change conformation only during energy minimization in order to reduce the computation time (*3–5*). Recently, we have been able to perform the molecular mechanics (MM) calculation, molecular dynamics (MD) simulation, or docking simulations including all atoms. Many approaches have been proposed for docking a ligand onto its receptor using the stochastic or Monte Carlo method (*6*), including AutoDock (*7*), GOLD (*8*), and DOCK (*9*). Although these methods have proven to be efficient for selecting a reasonable number of candidates for the complex, they are difficult to apply in cases of many degrees of freedom in the ligand. On the other hand,

*To whom correspondence should be addressed. Tel: +81-92-642-3043, Fax: +81-92-642-3043, E-mail: kuhara@grt.kyushu-u.ac.jp

 

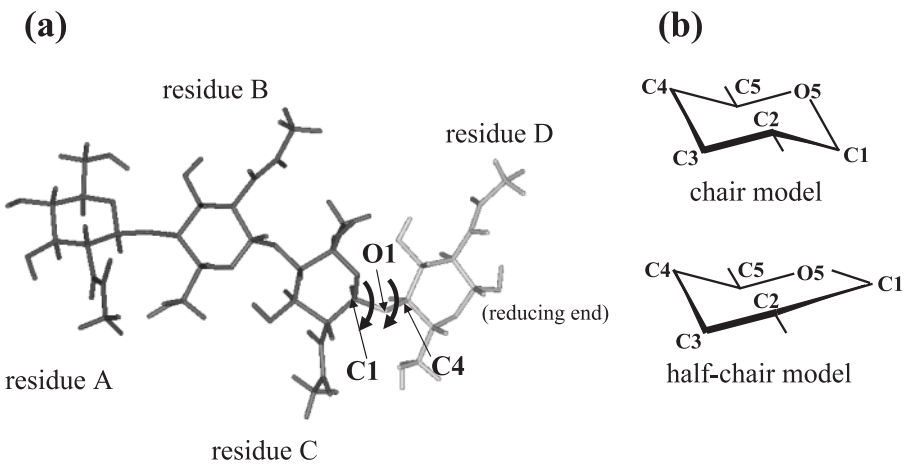


Fig. 1. **Systematic conformational search method in the GlcNAc oligomer.** (a) Two dihederal bonds rotated around the β-(1,4)-glycosidic bond (C1, O1, and C4 atoms). (b) Conformation of GlcNAc residue D in chair and half-chair models.

the systematic conformational search method is applicable when many degrees of freedom are in the ligand (*10*).

The X-ray structure of the complex between a trisaccharide MurNAc-GlcNAc-MurNAc and HEWL bound to subsites B to D indicated the distortion of the sugar ring of GlcNAc residue D (*11*). In order to understand the mechanism of the enzymatic reaction, structural evidence for the HEWL-$(GlcNAc)_6$ complex is necessary. However, the X-ray structure of the complex cannot be determined experimentally, because the lifetime of the intermediates is short, as they are catalyzed rapidly.

Recently, another hydrolysis mechanism of lysozyme was proposed by Vocadlo *et al.* Those authors used the E35Q mutant as HEWL and showed the ESI-MS mass spectra of the HEWL-$(GlcNAc)_2$ complex, as well as the X-ray structure of the intermediate complex bound between GlcNAc residue D and Asp52 by a covalent bond. This reaction model was called the "Koshland mechanism" (*12*).

The current models for the catalytic mechanism of lysozyme are thus established, but several issues remain. First, the binding modes at subsites E and F have not been identified clearly. Second, whether or not the ring of the GlcNAc residue D really becomes distorted on the right side remains controversial. Third, the reaction mechanism of the highly efficient transglycosylation has not been clarified. Since it is known that the binding of the acceptor molecule (GlcNAc oligomer) to subsites E and F would be necessary for transglycosylation, it is important to investigate the amino acid residues related to these subsites (*13*).

In this study, we purpose to construct the complexes between HEWL and $(GlcNAc)_6$ oligomer by a systematic conformational search and to perform the MD simulation in order to investigate the amino acid residues related to substrate binding in the productive and nonproductive complexes, and the relationship between the distortion of the GlcNAc residue D and the formation of the productive complexes.

## MATERIALS AND METHODS

*Obtaining the Initial Structure*—Coordinates for the X-ray crystallographic structure of the complex between HEWL and $(GlcNAc)_4$ were obtained from the Protein Data Bank (PDB code: 1lzc) (*14*). The structure of the complex was determined at high resolution (1.8 Å). Coordinates for the X-ray structure of GlcNAc were obtained from CSD (Cambridge Structure Database (*15*), Accession number: BCHITT10) to add to the reducing end of the GlcNAc oligomer in the systematic conformational search.

*Construction of Complexes Using the Systematic Conformational Search Method*—Hydrogen atoms were added to the heavy atoms in the 1lzc structure using the LEaP module of the AMBER7 package (*16*). Originally, the conformation of GlcNAc residue D in 1lzc was the chair form (colored gray in Fig. 1a). First, the fourth GlcNAc residue in 1lzc was removed, and GlcNAc in the CSD database was added to the reducing end of $(GlcNAc)_3$ by removing an oxygen atom and a hydrogen atom at the β-(1,4) glycosidic bond. To investigate whether or not GlcNAc residue D would be distorted, a systematic conformational search was performed against two cases: the chair and half-chair forms for GlcNAc residue D. The former complex is called the "chair model" (Fig. 1a), and the latter complex is called the "half-chair model" (Fig. 1b).

The systematic conformational search was performed using InsightII (Accelrys Co.). Two dihedral bonds around the β-(1,4)-glycosidic bond (C1, O1, and C4 atoms in Fig. 1a) were stepwise changed at 10 degree intervals for 360 degrees using InsightII, thus resulting in 1,296 (= 36 × 36) structures. To distort the GlcNAc residue from chair to half-chair form, the angle composed of O5, C1, and C2 atoms was adjusted to 0 degrees using the FLAP command in InsightII (Fig. 1b), and its partial charges were assigned using the MOPAC program (*17*). In addition, in the half-chair model, the force constraints of the dihedral angles of the pyranose ring of GlcNAc residue D (C1-C2-C3-C4, C2-C3-C4-C5, C3-C4-C5-O5, C4-C5-O5-C1, C5-O5-C1-C2 and O5-C1-C2-C3) set to 500 kcal/mol. In the systematic conformational search, the coordinates of the lysozyme were fixed. To obtain stable structures, the structures with total energy < 1,000 kcal/mol were selected among all structures formed by the systematic conformational search. Moreover, we calculated RMSDs (root mean square distances) among the Cartesian coordinates for the selected GlcNAc oligomers, and calculated the distance matrix among the coordinates, and clustered them by the neighbor-joining (NJ) method using the Phylip program (*18*). In the distance matrix calculation,

the distance between the position of the oxygen atom and that of the C1 atom in the β-(1,4)-glycosidic bond was weighted by 100, because the position of the glycosidic bond affects the position of the following GlcNAc residues. Structures with RMSDs below the threshold 0.15 Å were considered to belong to the same cluster. A significant energy barrier prevents the *gauche*-to-*trans* conformation of the acetyl group in GlcNAc residue. Therefore, according to the two-dimensional energy map for the two dihedral angles around the glycosidic bond in (GlcNAc)$_2$, the GlcNAc oligomer with disallowed dihedral angles at the glycosidic bond, which leads gauche conformation of the acetyl group, was excluded from the clusters.

The representative structure with the lowest energy was selected from each cluster; in that structure unfavorable interactions were relieved by using the steepest descent method followed by conjugate gradient energy minimization until the RMSDs of the elements in the gradient vector were less than 0.01 kcal/(mol Å) or until 3,000 steps had been performed. The GlcNAc residues comprising the binding site of HEWL were allowed to move only during the energy minimizations.

The GlcNAc residue in the chair form was added to the reducing end of the oligomer selected from each cluster to construct the HEWL-(GlcNAc)$_5$ complex by the same procedure performed in the construction of the HEWL-(GlcNAc)$_4$ complex. By repeating these procedures, the structure of complex between HEWL and (GlcNAc)$_6$ were formed.

*MD Simulation of the HEWL-(GlcNAc)$_6$ Complexes*—MD simulations were performed against the HEWL-(GlcNAc)$_6$ complexes selected from each cluster. The amber99 forcefield in the AMBER7 package was used in all simulations (*16*). Glycam was applied to assign the partial charges to the GlcNAc residues (*19*). This system was solvated in a periodic cubic box measuring $62.28 \times 57.79 \times 66.82$ (Å$^3$) and filled with the TIP3 water molecules. Long-range nonbonded interactions were truncated by using a 12.0 Å cutoff. The initial structure for the MD simulation was formed in four steps: 1. optimization of the positions of the water molecules using energy minimization and MD simulation, 2. optimization of the positions of the hydrogen atoms in HEWL using energy minimization, 3. optimization of the conformation of the side-chain using energy minimization, 4. optimization of the positions of all atoms in the system using energy minimization. In these steps, energy minimization was performed by the steepest descents and conjugate gradient minimizations until the gradient vector was less than 0.01 kcal/(mol Å) or until 3,000 steps.

MD simulation was performed as follows. Firstly, solvent in the systems was equilibrated for 20 ps while the temperature was increased from 0 K to 300 K. Following equilibration, MD simulation was performed by running a 500 ps with a 2 fs time step at 300 K in the NVT ensemble. During the MD simulation, the last 200 conformations were sampled at 1 ps intervals, and they were used for structural analysis. The interaction energies of the hydrogen bond, van der Waals bond, and Coulombic bond in the complex between HEWL and (GlcNAc)$_6$ were calculated using the ANAL module of the AMBER7 package (*16*).

## RESULTS AND DISCUSSION

*Construction of HEWL-(GlcNAc)$_6$ Complexes*—The numbers of the HEWL-(GlcNAc)$_4$, HEWL-(GlcNAc)$_5$, and HEWL-(GlcNAc)$_6$ complexes constructed by systematic conformational search are shown in Table 1. The resulting structures were classified into the chair and half-chair models.

The numbers of the complexes whose internal energies were less than 1,000 kcal/mol differed by the oligomer length. The number of the complexes in half-chair model was less than those in chair model in the HEWL-(GlcNAc)$_4$ complexes. Because the half-chaired GlcNAc residue D had a steric overlap with the side chain of Val109. The numbers of the structures were nearly equal between chair and half-chair models in the HEWL-(GlcNAc)$_5$ complexes. The number of the complexes in chair model was less than that in half-chair model in the HEWL-(GlcNAc)$_6$ complexes. These results indicate that the cleft is narrow for the GlcNAc residue D and wide for the GlcNAc residue F in the half-chair model.

*Classification of the Binding Mode of the HEWL-(GlcNAc)$_6$ Complexes*—We classified the 49 HEWL-(GlcNAc)$_6$ complexes (17: chair model; 32: half-chair model) into three binding modes: left side, center, and right side, based on the position of (GlcNAc)$_6$ in the binding site. To classify the position of (GlcNAc)$_6$ in the cleft, the distance LL was calculated between the center of the mass of the sugar ring of GlcNAc residue F and the hydrogen atom bonded to the backbone nitrogen atom (HN) of Arg45, while the distance LR was calculated between the center of the mass of the sugar ring of GlcNAc residue F and the hydrogen atom bonded to $C^{\eta 1}$ of Arg114. The difference L was calculated by LL – LR, and the complexes were classified based on the value of L into left side ($L < -3$), center ($-3 \leqslant L \leqslant 3$), and right side ($L > 3$). The numbers of the complexes are shown in Table 2.

In the chair model, (GlcNAc)$_6$ bound toward the left side and center but did not bind toward the right side. This result corresponded with the catalytic mechanism of lysozyme proposed by Scheraga *et al.*, in which GlcNAc residue D kept its chair form when (GlcNAc)$_6$ bound toward the left side (*5*). On the other hand, in the half-chair model, (GlcNAc)$_6$ bound toward all sides.

*MD Simulation of the HEWL-(GlcNAc)$_6$ Complexes*—According to the clustering of the RMSD distance matrix in (GlcNAc)$_6$ for the 17 chair models and 32 half-chair models (data not shown), we selected 8 and 12 structures from

**Table 1. Number of complexes formed in each GlcNAc oligomer.**

| Model | HEWL-(GlcNAc)$_4$ | HEWL-(GlcNAc)$_5$ | HEWL-(GlcNAc)$_6$ |
|---|---|---|---|
| Chair | 6 | 6 | 17 |
| Half-chair | 2 | 7 | 32 |

**Table 2. Classification of the binding mode of the HEWL-(GlcNAc)$_6$ complexes.**

| Model | Left side | Center | Right side |
|---|---|---|---|
| Chair | 11 | 6 | 0 |
| Half-chair | 9 | 14 | 9 |

**Table 3. Structural data of the 200 sampled complexes in each simulation.**

| Model | Name | Binding mode | Distance (Å)[a,b] | | | Interaction energy (kcal/mol) | | | Number of hydrogen bonds | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | E35 | D52 ($O^{\delta1}$) | D52 ($O^{\delta2}$) | LL | LP | Total | GlcNAc residue | | | | | |
| | | | | | | | | | A | B | C | D | E | F |
| Chair | C1 | center | 5.0 | 4.9 | 4.0 | 179.8 | –178.5 | 1.3 | 165 | 90 | 463 | 174 | 249 | 10 |
| | C2 | left | 5.4 | 5.4 | 4.6 | 179.2 | –170.1 | 9.1 | 147 | 116 | 356 | 201 | 114 | 79 |
| | C3 | left | 5.1 | 4.7 | 3.9 | 182.1 | –163.1 | 19.0 | 42 | 61 | 485 | 286 | 228 | 6 |
| | C4 | left | 4.9 | 5.0 | 3.9 | 183.3 | –182.7 | 0.6 | 216 | 88 | 481 | 223 | 338 | 3 |
| | C5 | left | 4.9 | 4.8 | 3.9 | 179.6 | –180.3 | –0.7 | 223 | 119 | 441 | 155 | 196 | 15 |
| | C6 | center | 4.9 | 5.0 | 4.1 | 183.7 | –171.3 | 12.4 | 192 | 105 | 422 | 200 | 244 | 4 |
| Half-chair | HC1 | center | 4.4 | 5.3 | 4.7 | 210.9 | –170.3 | 40.6 | 249 | 133 | 363 | 89 | 128 | 18 |
| | HC2 | right | 4.0 | 5.4 | 4.6 | 206.7 | –202.6 | 4.1 | 137 | 81 | 430 | 171 | 32 | 71 |
| | HC3 | right | 4.4 | 5.1 | 4.4 | 200.3 | –173.7 | 26.6 | 229 | 115 | 502 | 85 | 68 | 17 |
| | HC4 | center | 4.6 | 4.8 | 3.8 | 202.8 | –159.9 | 42.9 | 152 | 69 | 487 | 239 | 163 | 9 |
| | HC5 | right | 3.9 | 5.5 | 5.2 | 210.9 | –171.8 | 39.1 | 210 | 114 | 381 | 90 | 161 | 9 |
| | HC6 | left | 4.5 | 5.4 | 4.3 | 204.7 | –185.8 | 18.9 | 141 | 64 | 174 | 74 | 178 | 86 |
| | HC7 | left | 4.8 | 5.4 | 4.1 | 198.9 | –159.8 | 39.1 | 89 | 172 | 283 | 59 | 178 | 42 |
| | HC8 | right | 3.7 | 5.2 | 4.7 | 212.0 | –182.1 | 29.9 | 153 | 105 | 392 | 134 | 120 | 5 |

[a]The distance between the proton (HE2 atom) of Glu35 and the β-(1,4)-glycosidic oxygen (O4) atom linking GlcNAc residues D and E.
[b]The distance between the $O^{\delta1}$ and $O^{\delta2}$ atoms of Asp52 and the β-(1,4)-glycosidic oxygen (O4) atom linking GlcNAc residues D and E.

these models, respectively, to perform the MD simulation. This simulation was carried out during 500 ps against the selected 20 HEWL-$(GlcNAc)_6$ complexes (8: chair model, 12: half-chair model). If the fluctuations in RMSD values and total energy were not stable during the last 200 ps, the simulation was not adopted for further structural analysis. Finally, we obtained 6 (named C1 to C6) and 8 (named HC1 to HC8) initial structures for the MD simulations in the chair and half-chair models, respectively.

In order to investigate the interaction between HEWL and $(GlcNAc)_6$, we counted the number of hydrogen bonds between $(GlcNAc)_6$ and the amino acid residues in each sampled complex, because the intermolecular hydrogen bond has been thought to play a key role in catalysis.

Table 3 shows the structural data of the MD simulation, such as the model name, the name of the complex, the binding mode, the distance between the β-(1,4)-glycosidic oxygen (O4) atom linking GlcNAc residues D and E and the proton (HE2 atom) of Glu35, $O^{\delta1}$ and $O^{\delta2}$ atoms of Asp52, the intermolecular (LL) and intramolecular (LP) energies, and the number of hydrogen bonds between each GlcNAc residue and HEWL in the 200 sampled structures.

In the chair model, four complexes had $(GlcNAc)_6$ that bound toward the left side, and two complexes had $(GlcNAc)_6$ that bound toward the center. In the half-chair model, two, two, and four complexes had $(GlcNAc)_6$ that bound toward the left side, center, and right side, respectively. The average numbers of hydrogen bonds from GlcNAc residues A to F in the chair models in which $(GlcNAc)_6$ bound toward the left side (structures C2-5) were 157, 96, 441, 216, 219, and 26, respectively. On the other hand, the average numbers of hydrogen bonds from GlcNAc residues A to F in the half-chair models in which $(GlcNAc)_6$ bound toward the right side (structures HC2, HC3, HC5, and HC8) were 182, 104, 426, 120, 95, and 26, respectively. By the experimental time-course analysis for HEWL, the binding free energies from subsites A to F were –2.0, –3.0, –5.0, 4.5, –2.5, and –1.5 (kcal/mol), respectively (*20*, *21*). We then compared the average number of hydrogen bonds from the GlcNAc residues A to F and the binding free energies from subsites A to F, and found that GlcNAc residue C bound most tightly whereas residue F bound weakly; these results corresponded with the high binding free energy at subsite C and the low binding free energy at subsite F. Thus, the hydrogen bonding energy of each GlcNAc residue agreed with the binding free energy in each subsite.

To investigate whether or not a complex is productive, we calculated the distance between the β-(1,4)-glycosidic oxygen (O4) atom linking GlcNAc residues D and E and the proton bonded to the $O^{\varepsilon2}$ atom of Glu35 as a catalytic site against all of the sampled structures in each MD simulation. As shown in Table 3, the distances of the chair models bound toward the left side ranged from 4.9 to 5.4 Å. On the other hand, the distances of the half-chair models bound toward the right side ranged from 3.7 to 4.4 Å. The complexes in the half-chair model are likely to be catalyzed, because the average distance between the β-(1,4)-glycosidic oxygen (O4) atom linking GlcNAc residues D and E and the proton at Glu35 of the half-chair model was 1.0 Å shorter than that of the chair models, and besides, the conformations of the GlcNAc residue D were distorted. These results indicated that it is necessary for GlcNAc residue D to be half-chaired form to bind toward the right side in order to form productive complexes, which corresponded with the previous studies (*1*, *2*). In the present study, we classified the binding modes into right side, center, and left side. We consider the center binding mode as one of the nonproductive intermediate state, because the average distances between the β-(1,4)-glycosidic oxygen (O4) atom linking GlcNAc residues D and E and the proton at Glu35 of center binding mode were longer than that of the right-sided binding mode in half-chair model.

In Fig. 2, we show the average structures C2 and HC2, whose GlcNAc residue F bound most tightly to HEWL in the chair and half-chair models, respectively. As Table 3 shows, in structure C2, the numbers of hydrogen bonds in

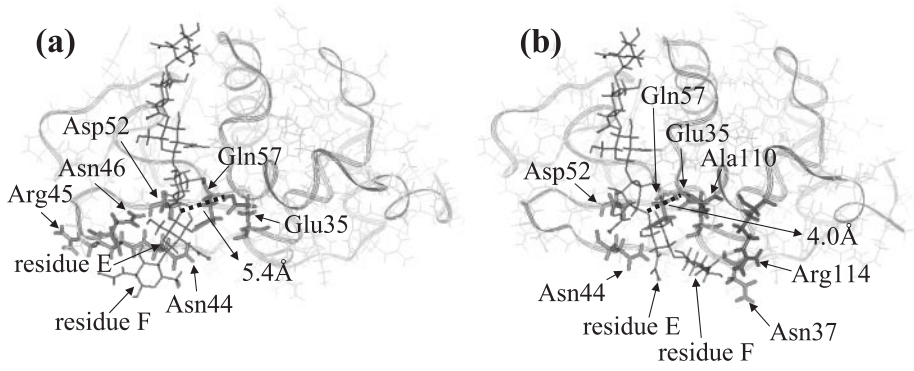


Fig. 2. **Comparison of the average structures in which (GlcNAc)$_6$ bound toward the left and right sides.** (a) Structure C2 (chair model). (b) Structure HC2 (half-chair model). The dotted lines mean the distances between the β-(1,4)-glycosidic oxygen (O4) atom linking GlcNAc residues D and E and the proton (HE2 atom) of Glu35 as a catalytic site. The highlighted residues were shown to form hydrogen bonds in the MD simulation.

GlcNAc residues A to F were 147, 116, 356, 201, 114, and 79, respectively. In structure HC2, the numbers of hydrogen bonds in GlcNAc residues A to F were 137, 81, 430, 171, 32, and 71, respectively.

In the average structure C2 (Fig. 2a), Glu35, Asn44, Arg45, Asn46, Asp52, and Gln57 formed hydrogen bonds with GlcNAc residues E and F. In the average structure HC2 (Fig. 2b), Glu35, Asn37, Asn44, Asp52, Gln57, Ala110, and Arg114 formed hydrogen bonds with GlcNAc residues E and F. According to the average distances between the β-(1,4)-glycosidic oxygen (O4) atom linking GlcNAc residues D and E and the proton bonded to the $O^{\epsilon 2}$ atom of Glu35 as the catalytic site (structures C2: 5.4 Å and HC2: 4.0 Å) and the distortion of the ring of the GlcNAc residue D, structure HC2 was considered productive while C2 was considered nonproductive.

*Investigation of the Amino Acid Residues That Interact with (GlcNAc)$_6$*—Here we selected structures C2–C5, whose (GlcNAc)$_6$ bound toward the left side in the chair models; and structures HC2, HC3, HC5, and HC8, whose (GlcNAc)$_6$ bound toward the right side in the half-chair models. Table 4a lists the amino acid residues estimated to interact with each GlcNAc residue in the chair and half-chair models. Each number means the hydrogen-bonding between the amino acid residue and each GlcNAc residue in 200 sampled structures. Table 4b lists the amino acid residues estimated to relate to substrate binding in each subsite in the previous studies (*1*, *2*).

As shown in Table 4b, Trp62, Trp108, and Val109 are related to subsites B, C, and E, respectively, by hydrophobic interaction (written in parentheses). However, the energy of the hydrophobic interaction could not be calculated because its energy term was not defined in the forcefield. Therefore, we did not compare the interactions among these amino acid residues in parentheses. As shown in Table 4a, we found newly some amino acid residues to interact with GlcNAc residues using various structures obtained by the MD simulation. GlcNAc residue A interacted with Arg73, Asp101, Gly102, and Asn103 by hydrogen bonding; of these, Asp101 was frequently found in the MD simulation. GlcNAc residue B interacted with Asp101 and Asn103 in subsite B. In the previous studies (*1*, *2*), Asp101 was related to substrate binding to subsites A and B, which corresponded with the results of the present study. GlcNAc residue C interacted with Asn59, Trp62, Trp63, Asn103, and Ala107 by hydrogen bonding. The following eight residues were newly found to be related to substrate binding: Arg73, Gly102, Asn103 for GlcNAc residue A; Asn103 for GlcNAc residues B and C; Leu56, Ala107, Val109 for GlcNAc residue D; Ala110 for GlcNAc residue E; and Lys33 for GlcNAc residue F.

As shown in Table 4a, GlcNAc residue D in the chair model interacted mainly with Asn46 and Gln57. On the other hand, GlcNAc residue D in the half-chair model interacted mainly with Glu35 and Asn46. GlcNAc residue E interacted mainly with Glu35 and Asp52 in the chair model. On the other hand, GlcNAc residue E interacted mainly with Glu35 and Gln57 in the half-chair model. GlcNAc residue F interacted mainly with Asn44, Arg45, and Asn46 in the chair model. On the other hand, GlcNAc residue F interacted mainly with Lys33, Phe34, Asn37, and Arg114 in the half-chair model. Thus the amino acid residues that interacted with GlcNAc residues D to F were rather different between the chair and half-chair models.

The amino acid residues in Table 4b mostly corresponded with the residues that were estimated in this study (Table 4a), except for Thr47. To investigate the distance between Thr47 and GlcNAc residue F, we calculated the distance between the hydrogen atom bonded to $C^{\gamma}$ in Thr47 and the O1 atom in GlcNAc residue F. The average distance of structures C1 to C6 was 11.8 Å in the chair model, and that of structures HC1 to HC8 was 15.9 Å in the half-chair model. From these results, we conclude that GlcNAc residue F does not interact with Thr47.

Asn37 was found to interact only with GlcNAc residue F in structures HC2 and HC8, and Phe34 was found to interact only with GlcNAc residue F in structure HC5. In the case of Asn37, in an earlier study (*2*), a hydrogen bond was estimated to form between the two hydrogen atoms bonded to $N^{\delta 2}$ of Asn37 (HD21 and HD22) and the O6 atom in GlcNAc residue F. In structure HC2, the average distance between the hydrogen atom bonded to $N^{\delta 2}$ (HD21) in Asn37 and the O6 atoms in GlcNAc residue F was 7.2 Å (SD = 0.50), and the average distance between the hydrogen atom bonded to $N^{\delta 2}$ (HD22) of Asn37 and the O6 atoms in GlcNAc residue F was 6.1 Å (SD = 0.56). In structure HC8, the average distance between the hydrogen atom bonded to $N^{\delta 2}$ (HD21) in Asn37 and the O6 atoms in GlcNAc residue F was 7.0 Å (SD = 0.52), and the average distance between the hydrogen atom bonded to $N^{\delta 2}$ (HD22) of Asn37 and the O6 atoms in GlcNAc residue F was 6.0 Å (SD = 0.62). Inoue *et al.* suggested that the side chain of Asn37 may unfavorably interact with the substrate in the transition state (*22*). Therefore, the functional role of Asn37 was regarded as unclear. Kawamura *et al.*

**Table 4. Number of the hydrogen bonding between amino acid residues and $(GlcNAc)_6$ in this study and the previous studies.**
(a) Residues in this study. (b) Residues in the previous studies (*1*, *2*). $E_L$ and $F_L$ mean the subsites E and F in the left side, respectively (*5*).

| (a) | | Chair model | | | | Half-chair model | | | | (b) | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GlcNAc | AA | C2 | C3 | C4 | C5 | HC2 | HC3 | HC5 | HC8 | Subsite | AA |
| A | Arg73 | 1 | 3 | 1 | 3 | 0 | 1 | 4 | 0 | A | Asp101 |
| | Asp101 | 141 | 39 | 206 | 217 | 133 | 216 | 205 | 153 | | |
| | Gly102 | 0 | 0 | 2 | 0 | 4 | 3 | 0 | 0 | | |
| | Asn103 | 5 | 0 | 7 | 3 | 0 | 9 | 1 | 0 | | |
| B | Asp101 | 116 | 61 | 77 | 118 | 68 | 89 | 106 | 105 | B | (Trp62) |
| | Asn103 | 0 | 0 | 11 | 1 | 13 | 26 | 8 | 0 | | Asp101 |
| C | Asn59 | 200 | 200 | 200 | 200 | 200 | 200 | 176 | 201 | C | Asn59 |
| | Trp62 | 25 | 57 | 34 | 50 | 18 | 28 | 33 | 53 | | Trp62 |
| | Trp63 | 33 | 72 | 49 | 39 | 32 | 73 | 26 | 33 | | Trp63 |
| | Asn103 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | | Ala107 |
| | Ala107 | 98 | 156 | 198 | 152 | 180 | 200 | 146 | 105 | | (Trp108) |
| D | Glu35 | 1 | 2 | 3 | 31 | 125 | 13 | 69 | 14 | D | Glu35 |
| | Asn46 | 0 | 200 | 200 | 105 | 37 | 59 | 19 | 64 | | Asp52 |
| | Asp52 | 0 | 0 | 0 | 0 | 3 | 1 | 0 | 12 | | Gln57 |
| | Leu56 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | | |
| | Gln57 | 200 | 83 | 16 | 2 | 0 | 9 | 0 | 43 | | |
| | Ala107 | 0 | 0 | 4 | 7 | 0 | 0 | 0 | 0 | | |
| | Val109 | 0 | 1 | 0 | 10 | 6 | 3 | 2 | 0 | | |
| E | Glu35 | 2 | 16 | 199 | 61 | 6 | 62 | 99 | 35 | E | Glu35 |
| | Asn44 | 2 | 11 | 0 | 20 | 1 | 1 | 20 | 39 | | Asn44 |
| | Arg45 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | | Gln57 |
| | Asn46 | 92 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | | (Val109) |
| | Asp52 | 6 | 200 | 139 | 113 | 2 | 5 | 1 | 3 | | |
| | Gln57 | 12 | 0 | 0 | 0 | 22 | 0 | 41 | 43 | | |
| | Ala110 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | | |
| F | Lys33 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | F | Phe34 |
| | Phe34 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | | Asn37 |
| | Glu35 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | | Arg114 |
| | Asn37 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | $E_L$ & $F_L$ | Arg45 |
| | Asn44 | 0 | 6 | 3 | 15 | 0 | 0 | 0 | 0 | | Asn46 |
| | Arg45 | 62 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | Thr47 |
| | Asn46 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| | Arg114 | 0 | 0 | 0 | 0 | 70 | 16 | 5 | 1 | | |

investigated the functional role of Asn37 by replacing it with Gly37 or Ser37 by site-directed mutagenesis. According to the parameters estimated for the binding free energy in subsites and the rate constant of transglycosylation, Asn37 is involved not only in subsite binding at subsite F but also in transglycosylation activity. Moreover, those authors analyzed the X-ray structure of N37G and N37S mutants and found that Asn37 formed a hydrogen bond with Lys33. In the N37S mutant, the hydrogen bond interacting with Lys33 was lost, and the side chain of the substituted Ser37 was rotated to GlcNAc residue F (*23*). From these results, we conclude that the hydrogen bond between Asn37 and GlcNAc residue F is weak, because the distance between them is a little far to interact, but it is possible that Asn37 binds with (GlcNAc)6. In the case of Phe34, the average distance between the backbone oxygen atom of Phe34 and the H6O atom of GlcNAc residue F was 4.7 Å (SD = 0.51) in structure HC5. We conclude that Phe34 bind with the substrate more frequently than would Asn37.

Vocadlo *et al*. supported the Koshland mechanism, in which a covalent bond forms between GlcNAc residue D and Asp52 in the productive complex (*12*). In the Koshland mechanism, Asp52 is suggested to interact with the half-chaired GlcNAc residue D in HEWL-$(GlcNAc)_6$ complex in the transition state same as in the Phillips mechanism. Interestingly, we found hydrogen bonding between Asp52 and GlcNAc residue E in the chair model (structures C2–C5), and between Asp52 and GlcNAc residue D in the half-chair model (structures HC2, HC3 and HC8) as shown in Table 4a. From these results, when the $(GlcNAc)_6$ bound to left side, the GlcNAc residue E could interact with Asp52 by hydrogen bonding. We could not construct the complex whose chaired GlcNAc residue D bound to Asp52 with covalent bond after the breakdown of the intermediate, which was suggested in the Koshland mechanism. To investigate the formation of the covalent bond between the GlcNAc residue D and Asp52, an analysis using quantum mechanics should be applied. When the $(GlcNAc)_6$ bound to right side with distortion of GlcNAc residue D, the GlcNAc residue D could interact with Asp52 by hydrogen bonding. According to the right-sided complexes in the half-chair model, the intermediate structure could be

stabilized through the electrostatic interaction suggested in the Phillips mechanism, because the hydrogen bonding were formed between the half-chaired GlcNAc residue D and Asp52.

Recently, many approaches have been proposed to simulate docking between ligand and receptor. But we could not obtain a suitable complex using the docking program Autodock (data not shown), because there were many degrees of freedom in (GlcNAc)$_6$. Besides, the distortion of the ring of GlcNAc residue D is difficult to form using a stochastic method such as Autodock. Applying both the systematic conformational search and the MD simulation, we found newly eight amino acid residues that interact with (GlcNAc)$_6$. Especially, the interaction between the amino acid residues in subsites E and F with each residue in (GlcNAc)$_6$ oligomer can now be investigated in detail. In order to clarify the transglycosylation reaction mechanism, three-dimensional structures of complex between lysozyme and acceptor molecule should be formed, but there is no structural information about the positions to bind for acceptor. Our future studies will construct the lysozyme-acecptor complexes to explore the amino acid residues which interact with acceptor using the HEWL-(GlcNAc)$_6$ complexes constructed in this study. These results will be helpful for understanding of the mechanism of transglycosylation.

We thank Dr. Hideki Takehara, Dr. Yoichi Aso, and Dr. Masatsune Ishiguro for valuable advices, suggestions and discussions. We acknowledge the Cambridge Crystallographic Data Centre (CCDC) for using the X-ray crystallographic data. Computation time for MOPAC was provided by Computing and Communications Center, Kyushu University.

## REFERENCES

1. Phillips, D.C. (1966) The three-dimensional structure of an enzyme molecule. *Sci. Am.* **215**, 78–90
2. Blake, C.C., Johnson, L.N., Mair, G.A., North, A.C., Phillips, D.C., and Sarma, V.R. (1967) Crystallographic studies of the activity of hen egg-white lysozyme. *Proc. R. Soc. Lond. B. Biol. Sci.* **167**, 378–388
3. Pincus, M.R., Zimmerman, S.S., and Scheraga, H.A. (1976) Prediction of three-dimensional structures of enzyme-substrate and enzyme-inhibitor complexes of lysozyme. *Proc. Natl. Acad. Sci. USA* **73**, 4261–4265
4. Pincus, M.R., Zimmerman, S.S., and Scheraga, H.A. (1977) Structures of enzyme-substrate complexes of lysozyme. *Proc. Natl. Acad. Sci. USA* **74**, 2629–2633
5. Pincus, M.R. and Scheraga, H.A. (1979) Conformational energy calculations of enzyme-substrate and enzyme-inhibitor complexes of lysozyme. 2. Calculation of the structures of complexes with a flexible enzyme. *Macromolecules* **12**, 633–644
6. Chang, G., Guida, W.C., and Still, W.C. (1989) An internal coordinate Monte Carlo method for searching conformational space. *J. Am. Chem. Soc.* **111**, 4379–4386
7. Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A.J. (1999) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**, 1639–1662
8. Verdonk, M.L., Cole, J.C., Hartshorn, M.J., Murray, C.W., and Taylor, R.D. (2003) Improved protein-ligand docking using GOLD. *Proteins* **52**, 609–623
9. Ewing, T.J., Makino, S., Skillman, A.G., and Kuntz, I.D. (2001) DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *J. Comput. Aided Mol. Des.* **15**, 411–428
10. Saunders, M., Houk, K.N., Wu, Y.D., Still, W.C., Lipton, M., Chang, G., and Guida, W.C. (1990) Conformations of cycloheptadecane. A comparison of methods for conformational searching. *J. Am. Chem. Soc.* **112**, 1419–1427
11. Strynadka, N.C. and James, M.N. (1991) Lysozyme revisited: crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. *J. Mol. Biol.* **220**, 401–424
12. Vocadlo, D.J., Davies, G.J., Laine, R., and Withers, S.G. (2001) Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature* **412**, 835–838
13. Smith-Gill, S.J., Rupley, J.A., Pincus, M.R., Carty, R.P., and Scheraga, H.A. (1984) Experimental identification of a theoretically predicted "left-sided" binding mode for (GlcNAc)6 in the active site of lysozyme. *Biochemistry* **23**, 993–997
14. Maenaka, K., Matsushima, M., Song, H., Suenada, F., Watanabe, K., and Kumagai, I. (1995) Dissection of protein-carbohydrate interactions in mutant hen egg-white lysozyme complexes and their hydrolytic activity. *J. Mol. Biol.* **247**, 281–293
15. Allen, F.H. (2002) The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr. B.* **58**, 380–388
16. Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham III, T.E., DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995) AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Commun.* **91**, 1–41
17. Stewart, J.J.P. (1990) MOPAC: a semiempirical molecular orbital program. *J. Comput. Aided Mol. Des.* **4**, 1–105
18. Felsenstein, J. (1989) PHYLIP—Phylogeny inference package (Version 3.2). *Cladistics* **5**, 164–166
19. Woods, R.J., Dwek, R.A., and Edge, C.J. (1995) Molecular mechanical and molecular dynamical simulations of glycoproteins and oligosaccharides. 1. GLYCAM_93 parameter development. *J. Phys. Chem.* 99, 3832–3846
20. Masaki, A., Fukamizo, T., Otakara, A., Torikata, T., Hayashi, K., and Imoto, T. (1981) Estimation of rate constants in lysozyme-catalyzed reaction of chitooligosaccharides. *J. Biochem.* **90**, 1167–1175
21. Fukamizo, T., Minematsu, T., Yanase, Y., Hayashi, K., and Goto, S. (1986) Substrate size dependence of lysozyme-catalyzed reaction. *Arch. Biochem. Biophys.* **250**, 312–321
22. Inoue, M., Yamada, H., Yasukochi, T., Miki, T., Horiuchi, T., and Imoto, T. (1992) Left-sided substrate binding of lysozyme: Evidence for the involvement of asparagines-46 in the initial binding of substrate to chicken lysozyme. *Biochemistry* **31**, 10322–10330
23. Kawamura, S., Eto, M., Imoto, T., Ikemizu, S., Araki, T., and Torikata, T. (2004) Functional and structural effects of mutagenic replacement of Asn37 at subsite F on the lysozyme-catalyzed reaction. *Biosci. Biotechnol. Biochem.* **68**, 593–601