Conformational preferences of the aglycon moiety in models and analogs of GlcNAc-Asn linkage: Crystal structures and ab initio quantum chemical calculations of N-(beta-D-glycopyran...

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · AUGUST 2008

Impact Factor: 12.11 · DOI: 10.1021/ja800335m · Source: PubMed

J|A|C|S
ARTICLES

Published on Web 06/07/2008

Conformational Preferences of the Aglycon Moiety in Models and Analogs of GlcNAc-Asn Linkage: Crystal Structures and ab Initio Quantum Chemical Calculations of *N*-(β-D-Glycopyranosyl)haloacetamides

Mohamed Mohamed Naseer Ali,[†] Udayanath Aich,[‡] Babu Varghese,[‡] Serge Pérez,[†] Anne Imberty,*[,†] and Duraikkannu Loganathan*[,‡]

CERMAV-CNRS (affiliated to Université Joseph Fourier and ICMG), BP 53, 38041 Grenoble cedex 9, France, Department of Chemistry and Sophisticated Analytical Instrumentation Facility, Indian Institute of Technology Madras, Chennai, 600036, India

Received January 15, 2008; Revised Manuscript Received March 30, 2008

Abstract: The biological addition of oligosaccharide structures to asparagine residues of *N*-glycoproteins influences the properties and bioactivities of these macromolecules. The linkage region constituents, 2-acetamino-2-deoxy-β-D-glucopyranose monosaccharide (GlcNAc) and L-asparagine amino acid (Asn), are conserved in the *N*-glycoproteins of all eukaryotes. In order to gain information about the structure and dynamics of glycosylated proteins, two chloroacetamido sugars, Glcβ$NHCOCH_2Cl$ and Manβ$NHCOCH_2Cl$, have been synthesized, and their crystal structures have been solved. Structural comparison with a series of other models and analogs gives insight about the influence of the *N*-acetyl group at position C2 on the conformation of the glycan-peptide linkage at C1. Interestingly, this *N*-acetyl group also influences the packing and network of hydrogen bonds with involvement in weak hydrogen bonds C−H···X that are of biological importance. DFT ab initio calculations performed on a series of models and analogs also confirm that the GlcNAc derivatives present different preferred conformation about the $N-CO-CH_2-X$ (χ_2) torsion angle of the glycan-peptide linkage, when compared to other monosaccharide derivatives. The energy profiles that have been obtained will be useful for parametrization of molecular mechanics force-field. The conjunction of crystallographic and computational chemistry studies provides arguments for the structural effect of the *N*-acetyl group at C2 in establishing an extended conformation that presents the oligosaccharide away from the protein surface.

Introduction

Glycan components of glycoproteins play vital roles in many biological processes as both recognition determinants, encoding information necessary for cell−cell as well as cell−matrix interactions, and modulators of the intrinsic properties, including folding of proteins.[1−4] Understanding the structure−function correlations of the protein-linked glycans has been a challenging problem in glycobiology. The linkage region constituents, 2-acetamino-2-deoxy-β-D-glucopyranose monosaccharide (GlcNAc) and L-asparagine amino acid (Asn), are conserved in the *N*-glycoproteins of all eukaryotes.[5] Since dynamics of the GlcNAc-Asn linkage can influence the presentation of the glycan chain on the cell surface, elucidation of the structure and conformation of the linkage region of glycoproteins is of fundamental importance. Structural investigations of glycoproteins, however, have been hindered due to the complexity of the glycan chains, their microheterogeneity, flexibility and the nonavailability of these compounds in sufficient quantities. Use of model compounds and analogs would be a valuable approach to understand the effect of structural variations on the linkage region conformation.

As part of a major research program aimed at understanding the structural significance of the constituents of the *N*-glycoprotein linkage region, Loganathan et al.,[6,7] have earlier reported the results of a systematic investigation among several β-*N*-glycopyranosylamido derivatives, pertaining to the effect of structural variation in the linkage region sugar and its aglycon moiety on the conformation (Figure 1) when compared to the crystal structures of GlcNAc-Asn.[8,9] Crystal structures of propionamido derivatives of several monosaccharides were also obtained.[10] A comprehensive

[†] CERMAV-CNRS, Grenoble, France.
[‡] Department of Chemistry, IIT Madras, India.

(1) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683.
(2) Gabius, H. J.; Siebert, H. C.; Andre, S.; Jimenez-Barbero, J.; Rudiger, H. *ChemBioChem* **2004**, *5*, 740.
(3) Varki, A. *Glycobiology* **1993**, *3*, 97.
(4) Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J. *Essentials of Glycobiology*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1999.
(5) Spiro, R. G. *Glycobiology* **2002**, *12*, 43R
(6) Loganathan, D.; Aich, U.; Lakshmanan, T. *Proc. Indian Nat. Sci. Acad.* **2005**, *71A*, 213.
(7) Sriram, D.; Lakshmanan, T.; Loganathan, D.; Srinivasan, S. *Carbohydr. Res.* **1998**, *309*, 227.
(8) Delbaere, L. T. *Biochem. J.* **1974**, *143*, 197.
(9) Ohanessian, J.; Avenel, D.; Neuman, A.; Gillier-Pandraud, H. *Carbohydr. Res.* **1980**, *80*, 1.
(10) Lakshmanan, T.; Sriram, D.; Priya, K.; Loganathan, D. *Biochem. Biophys. Res. Commun.* **2003**, *312*, 405.

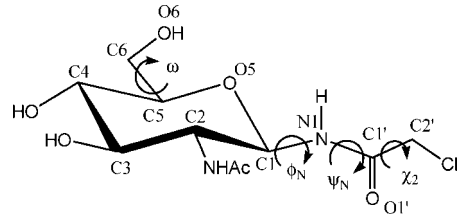

Figure 1. Schematic representation of GlcNAcβNHCOCH$_2$Cl together with the various torsion angles of interest ϕ_N = O5−C1−N1−C1′, ψ_N = C1−N1−C1′−C2′, ω = O5−C5−C6−O6, and χ_2 = N1−C1′−C2′−Cl.

analysis of both the regular hydrogen bonds and the weak interactions involving C−H···O/N interactions present in the crystal structures of 12 *N*-glycoprotein models and analogs revealed a cooperative network of bifurcated hydrogen bonds consisting of N−H···O and C−H···O interactions seen uniquely for the model, GlcNAcβNHAc, and not for any analog including the propionamide derivative, GlcNAcβNHPr (Figure 2).[11] The observed differences in the *N*-glycosidic torsion (ϕ_N) among the models and analogs could partly be attributed to the varied interactions underlying the molecular assembly, such as the hydrogen bonds between neighboring *N*-acetyl groups. In addition, inherent conformational preferences of these compounds need to be examined in order to obtain a better understanding of the intra- and intermolecular interactions and eventually of the structural significance of the constituents of the *N*-glycoprotein linkage region.

The preliminary work on the structural investigation of two chloroacetamido sugars, GlcβNHCOCH$_2$Cl and GalβNHCOCH$_2$Cl, with Cl in place of the hydrophobic and electron donating methyl group of the propionamide analogs has appeared earlier.[12] Besides serving as analogs of the *N*-glycoprotein linkage region, such chloroacetamido sugars are potential inhibitors of many enzymes involved in carbohydrate metabolism. *N*-(β-D-Glucopyranosyl)chloroacetamide is an efficient inhibitor of glycogen phosphorylase.[13,14] The acetylated derivative of *N*-(2-deoxy-2-acetamido-β-D-glucopyranosyl)chloroacetamide was found to inhibit growth of mouse mammary adenocarcinoma cells in vitro.[15] A comparative analysis of the crystal structure of the free form of GlcβNHCOCH$_2$Cl and that of its complex with glycogen phosphorylase demonstrated a large conformational change for the chloroacetamido moiety resulting in enhanced binding to the enzyme.[12]

The *N*-glycoprotein linkage has been the subject of several conformational analyses.[16,17] In most cases, the studies have been performed with the combined use of experimental data and empirical molecular mechanics method. In recent years, huge development of more efficient computers and computational codes has led to the feasibility of performing conformational analysis of carbohydrates using Quantum Mechanical (QM) theory.[18] Simulation of large *N*-glycoproteins of biological

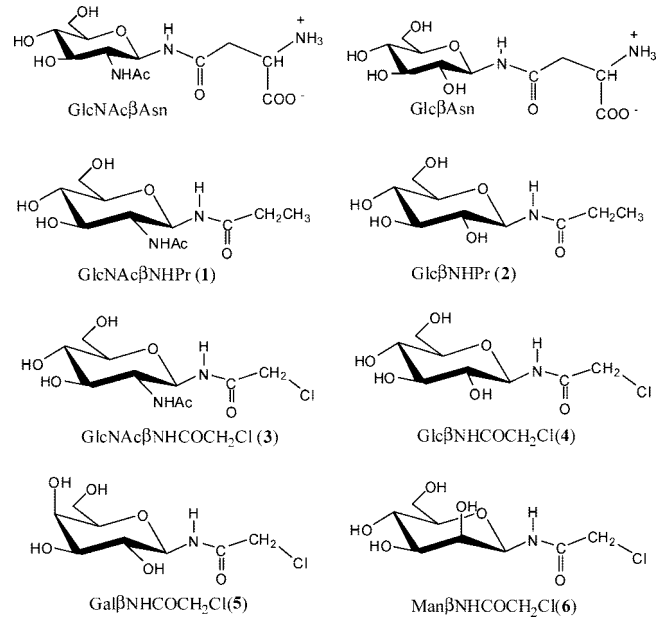

Figure 2. Schematic representation of models and analogs of the GlcNAc-Asn linkage.

interest with Molecular Mechanics method require to define appropriate parameters since accuracy of MM methods depends mostly on the parameters derived. Most programs are well-suited for modeling the protein part and some of them have tools for modeling of oligosaccharides part,[19,20] but the linkage between sugar and peptide is rarely correctly taken into account. However, as ab initio QM methods are based on first principles, they can be used to study any system including those for which no experimental information is available.

The only calculations performed using quantum chemistry consisted in investigating the ϕ_N preference of *N*-acetyl derivative of tetrahydropyran.[21] There is therefore a lack of theoretical knowledge on the influence of both the nature of the glycosidic ring and the nature of the aglycon on the conformation of the *N*-glycoprotein linkage region. The precise description of the energy dependence on conformation is required for future design of potent inhibitors.

Prompted by the above-mentioned considerations, the present work was undertaken to gather structural data and to perform ab initio calculations. We report herein the crystal structures of two *N*-chloroacetamido sugars, GlcNAcβNHCOCH$_2$Cl and ManβNHCOCH$_2$Cl (Figure 2). The details of molecular conformation and packing of four chloroacetamido analogs derived from GlcNAc, Glc, Gal, and Man are compared. The major focus of this study is on the conformation of the amido aglycon group. Due to stereoelectronic effects, ϕ_N and Ψ_N are somewhat limited in their conformational freedom and we focus on the conformational behavior of the χ_2 angle using ab initio calculations on six propionamido and chloroacetamido analogs (Figure 2). Comparison between solid state and in silico conformations

(11) Loganathan, D.; Aich, U. *Glycobiology* **2006**, *16*, 343.
(12) Aich, U.; Lakshmanan, T.; Varghese, B.; Loganathan, D. *J. Carbohydr. Chem.* **2003**, *22*, 891.
(13) Oikonomakos, N. G. *Curr. Protein Pept. Sci.* **2002**, *3*, 561.
(14) Watson, K. A.; Mitchell, E. P.; Johnson, L. N.; Cruciani, G.; Son, J. C.; Bichard, C. J.; Fleet, G. W.; Oikonomakos, N. G.; Kontou, M.; Zographos, S. E. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1995**, *51*, 458.
(15) Paul, B.; Bernacki, R. J.; Korytnyk, W. *Carbohydr. Res.* **1980**, *80*, 99.
(16) Imberty, A.; Pérez, S. *Protein Eng.* **1995**, *8*, 699.
(17) Wormald, M. R.; Petrescu, A. J.; Pao, Y. L.; Glithero, A.; Elliott, T.; Dwek, R. A. *Chem. Rev.* **2002**, *102*, 371.
(18) Da Silva, C. O. *Theor. Chem. Acc.* **2006**, *116*, 137.
(19) Imberty, A.; Pérez, S. *Chem. Rev.* **2000**, *100*, 4567.
(20) Pérez, S.; Imberty, A.; Engelsen, S. B.; Gruza, J.; Mazeau, K.; Jiménez-Barbero, J.; Poveda, A.; Espinosa, J. F.; van Eyck, B. P.; Jonhson, G.; French, A. D.; Kouwijzer, M. L. C. E.; Grootenhuis, D. J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G. K. R. *Carbohydr. Res.* **1998**, *314*, 141.
(21) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832.
(22) Momany, F. A.; Willett, J. L. *J. Comput. Chem.* **2000**, *21*, 1204.

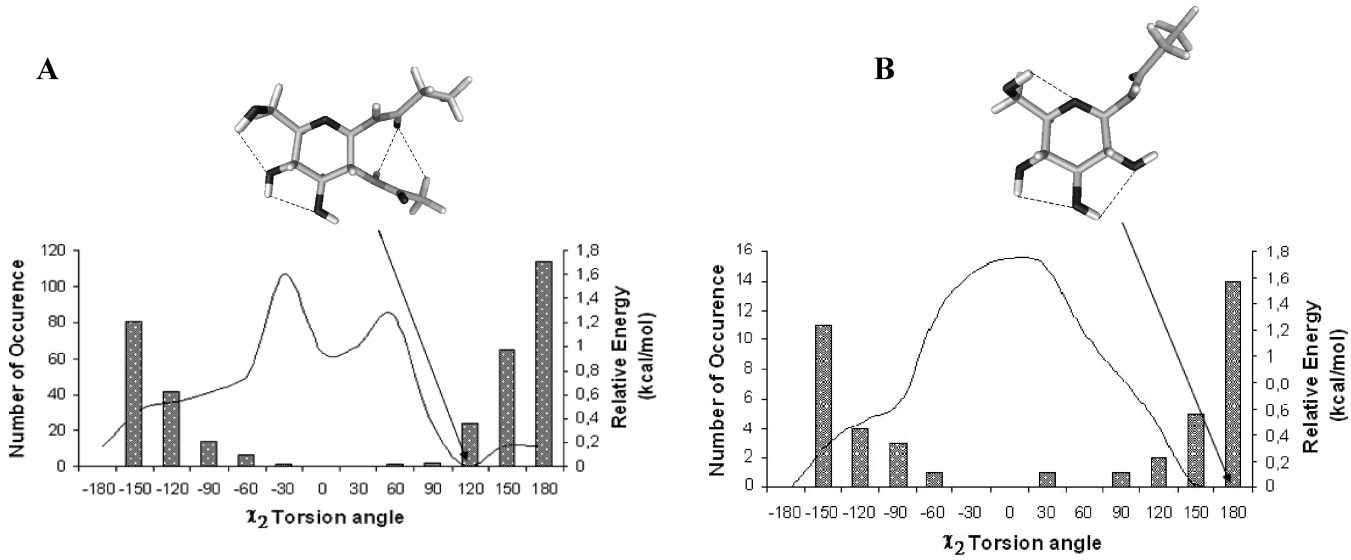

Figure 3. Quantum chemistry HF/6−31+G* energy curve for the χ_2 torsion angle of compound **1** (A) and compound **2** (B). The statistics taken from crystal structures have been extracted from a subset of the Protein Data Bank for compound **1** (A) and from the Cambridge Database for compound **2** (B). A graphical representation of the lowest energy conformation is displayed in each case.

gives some hints on the effects of environment and packing on the *N*-glycoprotein linkage conformational behavior.

Results

Ab Initio Quantum Chemical Calculations. Among the six compounds chosen for ab initio quantum chemical computations, GlcNAcβNHPr (**1**) and GlcβNHPr (**2**) (Figure 2), which possess a propionamido aglycon moiety, are analogs of the respective sugar conjugates of glutamine (Gln) not known to occur in Nature. The chloroacetamido derivative, GlcNAcβNHCOCH$_2$Cl (**3**), is an analog of the conserved *N*-glycoprotein linkage, GlcNAcβAsn, whereas GlcβNHCOCH$_2$Cl (**4**) is an analog of the variant linkage, GlcβAsn, found to occur in certain bacteria.[5] The other two compounds, GalβNHCOCH$_2$Cl (**5**) and ManNHCOCH$_2$Cl (**6**), are chloroacetamido derivatives of Gal and Man, which are commonly found attached to Ser/Thr in *O*-glycoproteins. The inclusion of the latter two analogs in the computational study was an effort toward evaluating their conformational preferences when they are *N*-linked. Calculations were performed using both HF (6−31+G* and 6−31++G**) and DFT (B3LYP/6−31++G**) since this latter method has been widely used by Momany's group for carbohydrates.[22] However, our first DFT calculations on GlcNAcβNAc provided contradictory results with respect to structural data (data not shown), whereas HF reproduces the expected methyl rotation barrier with no significant differences between the basis sets. The presence of *N*-acetyl group at C2 position and hence their influence on C1 *N*-acetyl group severely restricts the method choice and HF/6−31+G* has been selected for this study.

Conformational Behavior of *N*-Glycoprotein Linkage Models. The one-dimensional potential energy surface of propionamide derivative of GlcNAc (**1**) as a function of χ_2 variation, shows the presence of a global minimum at 120° close to a low energy plateau (0.17 kcal/mol) that extends to 180° (Figure 3A and Table 1). A secondary minimum of about 0.95 kcal/mol is also observed for a value of χ_2 of 0° between two energy barriers at −30° and 60°, which may correspond to favorable hydrophobic contacts between the extremities of the two exocyclic groups.

The energy curve is in general agreement with the statistical analysis performed on 350 *N*-glycan linkages selected from the Protein Data Bank (PDB),[23] since the trans conformation presents the higher population (Figure 3A). To refine the position of the energy minimum on the coarse grid potential energy surface (PES), complete optimization has been carried out for the conformers with χ_2 values of 0, 120, 150, and 180°. Upon optimization, the minimum at 120° in PES moves toward 156° with a lowering of energy of −0.18 kcal/mol.

The corresponding structure is characterized by a crown of intramolecular hydrogen bonds created by a clockwise arrangement of the hydroxyl groups. More interestingly, two more contacts are established between the *N*-acetyl group at C2 and the *N*-propionyl group at C1 such that O1′ serves as a bifurcated acceptor to N2−H and C2″−H (Figure 3A). The bond lengths and bond angles of the N2−H···O1′ hydrogen bond and the C2″−H···O1′ interaction are 2.7 Å and 111.2° and 3.1 Å and 127.4°, respectively. The χ_2 conformational preference in propionamide derivative of glucose (**2**) was also explored in order to analyze the effect of acetamido group at C2 position on the χ_2 conformational characteristics. The energy profile (Figure 3B and Table 1) displays a global minimum at 180° that extends to the conformation at 150° (0.05 kcal/mol) with an almost flat potential energy surface. The statistical analysis of crystal structures of glucose derivative from the Cambridge Structural Database (CSD) confirms the strong preference for the trans orientation.

Full optimization from these two grid points has been performed, and the global minimum is located between these two points and stabilized at a χ_2 value of 167°. This conformation that is displayed in Figure 3B also presents a clockwise arrangement of hydroxyl groups forming a cooperative network of intramolecular hydrogen bonds.

The conformational behavior of Glc and GlcNAc propionamide derivatives is rather similar for the localization of the global minimum for χ_2 that is close to 160° in both cases, but

(23) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.

Table 1. Some Selected Torsion Angles (°) of the Computed Models and the Crystal Structures[a]

compound	N1−C1′−C2′−X χ_2	O5−C5−C6−O6 ω	C1−N1−C1′−C2′ ψ_N	C1−N1−C1′−O1′ ψ_N	O5−C1−N1−C1′ ϕ_N
GlcNAcβNHPr	120(0)	−73.0	173.8	−7.0	−141.7
(**1**)	180(0.17)	−74.5	169.1	−12.3	−94.6
	150(0.17)	−74.5	167.1	−13.1	−92.7
	172.4(Cryst)	−66.8	172.5	−6.5	−91.0
	155.9 (Glob.opt)	−73.1	170.9	−8.3	−140.0
GlcβNHPr	180(0)	−57.6	164.7	−17.1	−91.8
(**2**)	150(0.05)	−57.6	163.2	−17.4	−92.1
	114.7(Cryst)	−68.6	166.5	−11.4	−89.5
	166.9 (Glob.opt)	−57.6	163.0	−18.0	−91.6
GlcNAcβNHCOCH$_2$Cl	60(0)	−72.6	171.2	−6.4	−145.0
(**3**)	173.0(Cryst)	−66.7	173.6	−4.1	−91.8
GlcβNHCOCH$_2$Cl	0(0)	−58.1	169.0	−12.7	−94.7
(**4**)	150(3.83)	−58.0	161.7	−17.7	−108.2
	131.3(Cryst)	−69.0	169.9	−8.1	−93.9
	−3.1 (Glob.opt)	−58.0	169.4	−12.7	−94.6
	162 (Sec.opt)	−58.0	161.6	−18.4	−92.2
GalβNHCOCH$_2$Cl	0(0)	63.8	169.2	−12.5	−96.8
(**5**)	153.3 (Cryst)	61.7	176.6	−4.7	−111.5
ManNHCOCH$_2$Cl	0(0)	−76.2	171.4	−10.2	−89.1
(**6**)	−98.8(Cryst)	−57.1	170. 8	−7.2	−115.9

[a] The values in parenthesis represent the relative energies (kcal/mol) for the models.

differs for the occurrence of secondary minima. Indeed, the presence of the *N*-acetyl group at C2 in compound **1** can stabilize the χ_2 in somewhat different orientations. The occurrence of the two hydrogen bonds (N−H···O and C−H···O) involving the *N*-acetyl group of C2 in compound **1** also results in a difference in the preferred orientation of the ϕ_N torsion angle that adopts a value of −140° in the global energy minimum of **1** and −96° in the case of **2**.

Conformational Behavior of *N*-Glycoprotein Linkage Halogenated Analogs. Conformational studies about χ_2 using ab initio HF theory have been extended to the chloroacetamido derivatives of GlcNAc, Glc, Gal, and Man monosaccharides. Similar conformational behavior is observed for the χ_2 torsion angle of GlcβNHCOCH$_2$Cl (**4**), GalβNHCOCH$_2$Cl (**5**), and ManβNHCOCH$_2$Cl (**6**) with a deep global minimum at 0° and an energy barrier with a plateau shape (150° to −150°) varying between 2.9 kcal/mol for Man derivative to 3.9 kcal/mol for Glc derivative (Figure 4 and Table 1). The conformational behavior is therefore very different from nonhalogenated model systems described above, with energy preference for the cis conformation and higher energy barrier. When comparing to crystal structures from the CSD, the highest populated conformation does correspond to the cis orientation of the χ_2 linkage. However, the trans conformation is significantly populated, indicating that the energy plateau observed in the ab initio calculation does correspond to a significant secondary energy minimum.

Full optimization was conducted on global and secondary energy minima of GlcβNHCOCH$_2$Cl (**4**). The result indicates a global minimum at −3° with negligible energy gain and a secondary minimum at 162° with a 0.04 kcal/mol energy gain. The resulting energy difference between the global and secondary minima for compound **4** is found to be 2.8 kcal/mol. DFT calculations conducted at the B3LYP/6−31+G* level yielded the same profile for the three compounds, with energy barrier about 1 kcal/mol higher than the HF results (data not shown).

For GlcNAcβNHCOCH$_2$Cl (**3**), the DFT energy curve has a very different shape with global minimum at 60° and secondary minimum at 0° with a low energy barrier between both conformers (Figure 4 and Table 1). Again, the *N*-acetyl group at C2 position has a strong influence on the conformational preference of the χ_2 torsion angle.

X-Ray Crystallographic Investigation. The single crystals of chloroacetamido sugars **3** and **6** were obtained from aqueous methanol. Compound **3**, GlcNAcβNHCOCH$_2$Cl, crystallized as a monohydrate and the Man analog, **6**, turned out to be anhydrous. The ORTEP representations of their solved structures along with atom numberings are shown in Figure 5.

The pyranose ring in both the compounds exists in the usual $^4C_1$ conformation. A list of selected bond lengths is given in Table 2. The ring C1−O5 bond lengths of both **3** and **6** are shorter than that of C5−O5 as reported for GlcNAcβAsn.[8,9]

The shortening is due to the delocalization of nitrogen lone pair of electrons into the anti bonding orbital of the C1−O5 bond. The C−N bond distances of 1.43 and 1.44 Å of **3** and **6**, respectively, are in good agreement with the value of 1.44 Å

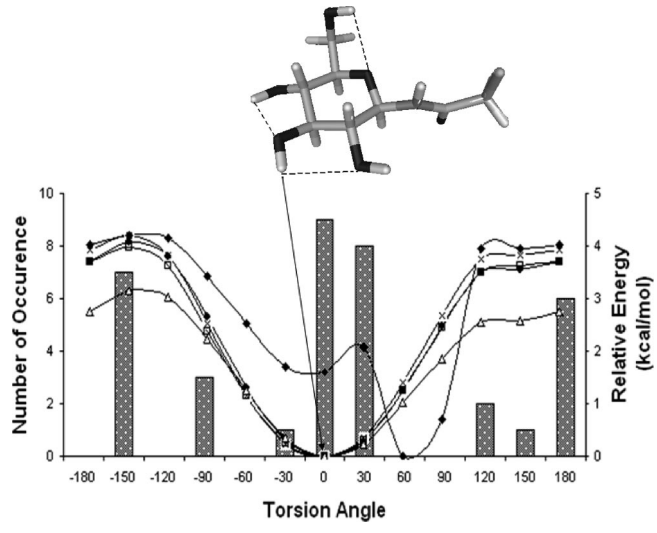

Figure 4. Quantum chemistry HF/6−31+G* energy curve for the χ_2 torsion angle of chloroacetamido derivatives (−♦− GlcNAcβNHCOCH$_2$Cl, −×− GlcβNHCOCH$_2$Cl, −□− GalβNHCOCH$_2$Cl, −△− ManβNHCOCH$_2$Cl, −●− XylβNHCOCH$_2$Cl, gray bars represent statistical distribution of χ_2 torsion for HNCOCCl fragment from CSD). A graphical representation of the lowest energy conformation of compound **4** is also displayed.

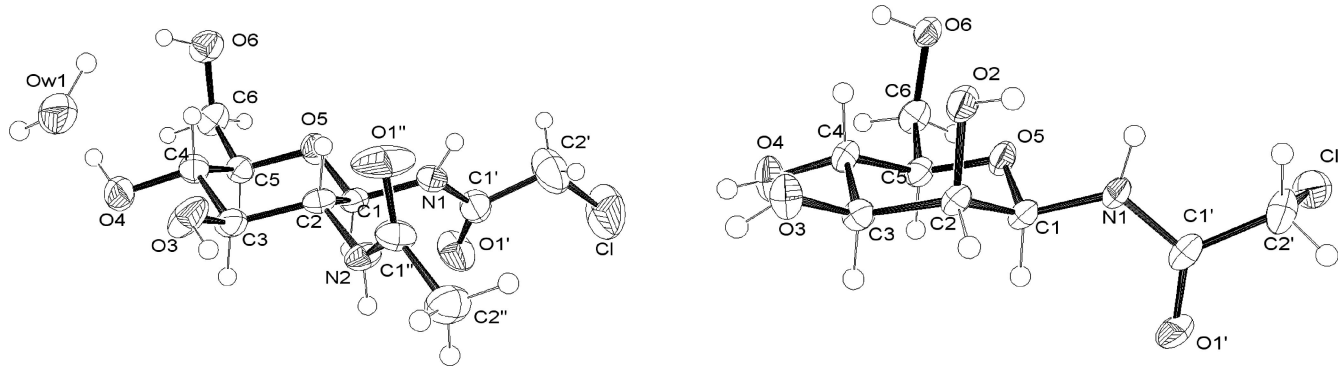

Figure 5. ORTEP representation (with 50% probability level) of GlcNAcβNHCOCH$_2$Cl (**3**), left, and ManβNHCOCH$_2$Cl (**6**), right.

Table 2. Selected Bond Lengths (Å) and Bond Angles (°) of GlcNAcβNHCOCH$_2$Cl (**3**) and ManβNHCOCH$_2$Cl (**6**)

parameter	(**3**)	(**6**)
C1−O5	1.421(6)	1.422(4)
C5−O5	1.434(6)	1.430(4)
C6−O6	1.424(8)	1.426(5)
C1−N1	1.430(7)	1.440(5)
C1′−N1	1.337(8)	1.344(4)
C1′−O1′	1.215(7)	1.227(4)
C2′−Cl	1.742(7)	1.775(4)
C4−C5−C6	113.9(4)	113.2(3)
O5−C5−C6	106.9(4)	106.1(3)
O5−C1−N1	107.7(4)	106.1(3)
C2−C1−N1	112.6(4)	111.3(3)
N1−C1′−O1′	124.5(5)	123.8(4)
C1′−C2′−Cl	113.5(5)	108.2(2)

observed for GlcNAcβAsn.[9] As a result of the above-mentioned delocalization of nitrogen lone pair of electrons in these models and analogs, their C−N bond distances are shorter than the alkyl chain C−N distances of 1.50 and 1.55 Å noted in the two independent molecules present in the asymmetric unit of *N*-methylchloroacetamide.[24] Comparison of C2′−Cl bond lengths of **3** and **6** with the values of 1.752(5) and 1.762(3) Å noted for **4** and **5**, respectively,[12] points out the largest difference between those of **3** and **6**.

The exocyclic bond angles, namely C4−C5−C6 and O5−C5−C6, are comparable to each other and agree well with those of other models and analogs reported earlier.[6] The latter angle is found to be consistently smaller than the former, and this is attributed to the effect of the lone pair of electrons causing a bend at the oxygen on the O5−C5−C6 bond angle. Such a difference in these bond angles has been observed in the crystal structures of GlcNAcβAsn[9] and several *N*-linked pyranoses.[25] The *N*-glycosidic valence angle, O5−C1−N1, is smaller than that of C2−C1−N1 for both **3** and **6** and the same trend is consistently seen in all the models and analogs examined. The valence angle involving chlorine atom, C1′−C2′−Cl, of **6** is about 5° shorter than that of **3.**

Interestingly, comparison of the values, 110.9(3) and 111.8(2)°, of the same angle observed for **4** and **5**, respectively, with those of **3** and **6** brings to light the maximum difference exhibited by **3** and **6**. The GlcNAc analog has larger C1′−C2′−Cl bond angle but shorter C2′−Cl bond length than the chloroacetamido derivative of Man.

(24) Koyama, Y.; Shimanouchi, T.; Iitaka, Y. *Acta Crystallogr.* **1971**, *B27*, 940.
(25) Rao, V. S. R.; Qasba, P. K.; Balaji, P. V.; Chandrasekaran, R. *Conformation of Carbohydrates*; Harwood Academic Publishers: The Netherlands, 1998.

Table 3. Selected Torsion Angles of Compounds **1**−**6**

compound code & no.	O5−C1−N1−C1′ (ϕ_N)	C1−N1−C1′−C2′ Ψ_N	N1−C1′−C2′−X χ_2 (X = CH$_3$/Cl)
GlcNAcβAsn·3H$_2$O[8]	−98.9	180.0	−172.2
GlcNAcβNHPr (**1**) [10]	−91.0(5)	172.5(5)	172.4(6)
GlcNAcβNHCOCH$_2$Cl (**3**)	−91.7(6)	173.6(7)	173.0(6)
GlcβAsn·H$_2$O[9]	−83.6	174.0	−166.5
GlcβNHPr (**2**)[10]	−89.3(6)	166.5(6)	114.7(8)
GlcβNHCOCH$_2$Cl (**4**)[12]	−93.9(5)	169.9(4)	131.3(4)
GalβNHCOCH$_2$Cl (**5**)[12]	−111.5(3)	176.6(3)	153.3(2)
ManβNHCOCH$_2$Cl (**6**)	−115.9(3)	170.8(3)	−98.8(3)

Molecular Conformation. Molecular conformation of the chloroacetamido sugars, **3** and **6**, is essentially defined by ω, ϕ_N, Ψ_N, and χ_2 (Figure 1). The conformational preferences of the hydroxyl group about the C5−C6 bond have been investigated by experimental and theoretical methods.[26–28] The values of the torsion angle about C5−C6, ω, indicate that both **3** and **6** adopt a *gg* conformation, which is also found to be case for all the models and analogs derived from GlcNAc and Man. This conformational preference is also observed in the solution phase where the ^{3}J of the H-5 with H-6a and H-6b turn out to be <1 Hz and ∼ 4 Hz respectively, indicating the gauche and gauche conformation (*gg*) of the hydroxyl group (data not shown).

The amide group in compounds **3** and **6** adopts the Z-*anti* conformation as revealed by the values of torsions, ϕ_N and Ψ_N, about C1−N1 and N1−C1′ bonds, respectively (Table 3), as classically observed in *N*-glycosidic linkage models and analogs.[6–9] Furthermore, the C2 acetamido group in the GlcNAc derivatives (**3**) also adopts a *Z-anti* conformation as is the case with GlcNAcβAsn.[8,9] The *N*-glycosidic linkage torsion, ϕ_N (O5−C1−N1−C1′), values of the propionamido (**1**) and chloroacetamido (**3**) derivatives of GlcNAc are nearly identical (Table 3), while the values differ slightly in the case of the corresponding analogs of Glc. The ϕ_N values observed for the chloracetamido derivatives of Gal and Man, **5** and **6**, however, deviate by about 20° and 24°, respectively, from that of the GlcNAc derivative (**3**). Similar deviation was seen earlier among the acetamido analogs of Gal, Man, and GlcNAc.[10] However, the variation in the values of ϕ_N noted for **3** and **4** is rather small and a similar narrow distribution was revealed by the statistical analysis of crystal structures of 26 glycoproteins containing 44 different glycosylation sites.[16]

Variation in the Torsion Angle χ_2. The major focus of the present investigation is the side chain dihedral angle about

(26) Marchessault, R. H.; Pérez, S. *Biopolymers* **1979**, *18*, 2369.
(27) Rockwell, G. D.; Grindley, T. B. *J. Am. Chem. Soc.* **1998**, *120*, 10953.
(28) Tvaroska, I.; Taravel, F. R.; Utille, J. P.; Carver, J. P. *Carbohydr. Res.* **2002**, *337*, 353.

Table 4. Regular Hydrogen Bond Parameters for Compounds **3** and **6**

D−H···A	H···A (Å)	D···A (Å)	D−H···A (°)	symmetry
GlcNAcβNHCOCH$_2$Cl (**3**)				
N1−H1N···O1′	2.16	3.00	175	$x-1, y, z$
N2−H2N···O1″	1.87	2.70	145	$x+1, y, z$
O4−H4O···O1W1	1.96	2.75	163	$x, y+1, z+1$
O6−H6O···O1W1	1.99	2.77	159	$x, y+1, z+1$
O1W1−H2W1···O3	1.87	2.70	166	$-x+1, y-{}^1/_2, -z+1$
OW1−H1W1···O4	1.80	2.71	166	$-x+2, y-{}^1/_2, -z+1$
ManβHCOCH$_2$Cl (**6**)				
N1−H1N···O2	2.06	2.95	175	$-1-x, {}^3/_2+y, {}^1/_2-z$
O2−H2O···O3	1.88	2.72	175	$-1-x, {}^3/_2+y, {}^1/_2-z$
O3−H3O···O6	2.04	2.80	154	${}^3/_2-x, -y, {}^1/_2+z$
O6−H6O···O1′	1.91	2.72	163	${}^1/_2-x, -y, {}^1/_2+z$
O4−H4O···O6	2.05	2.85	166	${}^3/_2-x, -y, {}^1/_2+z$

Table 5. C−H···O/Cl Hydrogen Bonding Pattern of Compounds **3** and **6**

D−H···A	H···A (Å)	D···A (Å)	D−H···A (°)	symmetry
GlcNAcβNHCOCH$_2$Cl (**3**)				
C3−H3···O1″	2.70	3.38	127	$x, -1+y, -1+z$
C1−H1···O1″	2.50	3.21	129	$x, -1+y, -1+z$
ManβNHCOCH$_2$Cl (**6**)				
C2′−H2′A···Cl	2.93	3.50	119	${}^1/_2+x, {}^1/_2-x, -z$
C2−H2···Cl	2.85	3.65	139	$1+x, y, z$
C6−H6A···O1′	2.54	3.45	157	$1-x, {}^1/_2+y, {}^1/_2-z$
C1−H1···O4	2.66	3.54	151	$2-x, -{}^1/_2+y, {}^1/_2-z$
C5−H5···O4	2.64	3.53	152	$2-x, -{}^1/_2+y, {}^1/_2-z$

C1′−C2′ bond, χ_2, which is a key element contributing to the linkage region conformation. A comparison of the χ_2 value determined for the chloroacetamido analog of GlcNAc (**3**) with those of the propionamide analog, **1**, and GlcNAcβAsn (Table 3) shows a negligible effect of functional group variation in the aglycon moiety on this important torsion angle. In contrast, the χ_2 values vary considerably among the Glc derivatives, **2**, **4** and GlcβAsn. The near anti conformation about C1′−C2′ bond exhibited by the variant linkage model of *N*-glycoproteins, GlcβAsn, gets altered to *gauche* conformation seen in the chloroacetamido and propionamido analogs, **2**, **4**, respectively. Clearly the occurrence of the *N*-acetyl group at C2 position is correlated to the minimal variation in χ_2 of the GlcNAc derivatives. This inference is supported by the fairly large spread in χ_2 values among the chloroacetamido analogs of Glc, Gal and Man, all of which carry a hydroxyl group at C2. Also, when analyzing protein crystal structures, the width of variation in χ_2 values is narrower for the *N*-glycosylated Asn residues as compared to the unglycosylated ones.[29] Understandably, another factor to consider in rationalizing the above data pertaining to χ_2 is the influence of differences in molecular packing, the details of which are presented in the following section.

Molecular Packing. The various hydrogen bonding parameters of compounds **3** and **6** are listed in (Table 4 and 5). In both of the compounds, N1 and O1′ are connected through a finite chain of hydrogen bonds. The characteristic feature of the monohydrate of GlcNAc chloroacetamide (**3**) is the occurrence of direct hydrogen bonding between N1 and O1′ accompanied by a similar one involving N2 and O1″ leading to double-pillared molecular packing along the crystallographic *a* axis as shown in Figure 6.

Occurrence of such a pillared pattern was first observed in the model compound GlcNAcβNHAc.[7] The packing of the

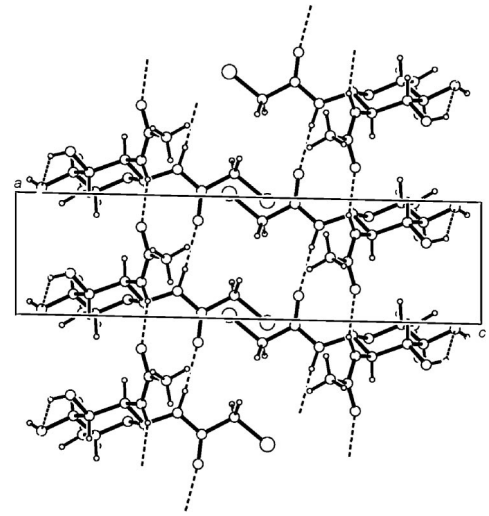

Figure 6. Representation of the packing of GlcNAcβNHCOCH$_2$Cl (**3**) projected along the *b* axis.

chloroacetamido derivative of GlcNAc is additionally stabilized by an infinite chain of hydrogen bonds as seen earlier in the cases of GlcNAcβNHAc and GlcNAcβNHPr (**1**). In this chain, the water molecule is bonded as a donor as well as an acceptor to its neighboring O4 atoms (Figure 7) facilitating the formation of a homodromic cycle of hydrogen-bonded network.

Another finite chain is also seen that starts from O6−H and ends with O3 both of which are hydrogen bonded to the water molecule. Thus, the water molecule in the crystal is fully utilized in terms of its hydrogen bonding potential to stabilize the molecular assembly as it is tetra coordinated donating both the hydrogens to O3 and O4 while serving as acceptor for O4 and O6. The claw-shaped hydrogen bonds created by HO4 and HO6 likely stabilize the *gg* conformation of the hydroxymethyl group. Examination of the C−H···O/N interactions in the crystal structure of **3** reveals the presence of the bifurcated hydrogen bond pattern involving C1−H···O1″···H−N2 (Figure 7A).

This pattern represents the first part of the cooperative network consisting of two antiparallel chains of bifurcated hydrogen bonds observed uniquely for the model, GlcNAcβNHAc, and not for any analog including GlcNAcβNHPr (1).[11] Interestingly, the carbonyl oxygen, O1″, of the C2-acetamido group in the chloroacetamido analog (**3**) is also hydrogen bonded to C3−H (Figure 7A). Such a trifurcated acceptor pattern is novel and not seen in the crystal structures of GlcNAcβAsn, GlcNAcβNHAc, and GlcNAcβNHPr (**1**). Significantly enough, the distinctive C2′−H···O1′ interaction, a key player in the formation of the second part of the above-mentioned antiparallel hydrogen bond network, is clearly missing in **3**, a feature shared by the propionamido analog (**1**).

In contrast to the direct hydrogen bonding between N1 and O1′ seen above, the anhydrous chloroacetamido derivative of Man (**6**) displays a finite chain of regular hydrogen bonds starting at N1, passing through O2, O3, and O6 and ending with O1′ (Figure 7B). This feature is similar to that reported for the corresponding acetamido derivative (ManβNHAc).[10] In addition, O6 of this finite chain in **6** functions as an acceptor for O4−H rendering the former atom tri coordinated. The O4−H···O6 interaction between two neighboring molecules of **6** enables the creation of a seven-membered ring involving C4, C3, O3, and H (O3) atoms. Furthermore, a key acceptor role played by the sugar ring O5 to H−O3 results in a tandem fusion during

(29) Petrescu, A. J.; Milac, A. L.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R. *Glycobiology* **2004**, *14*, 103.

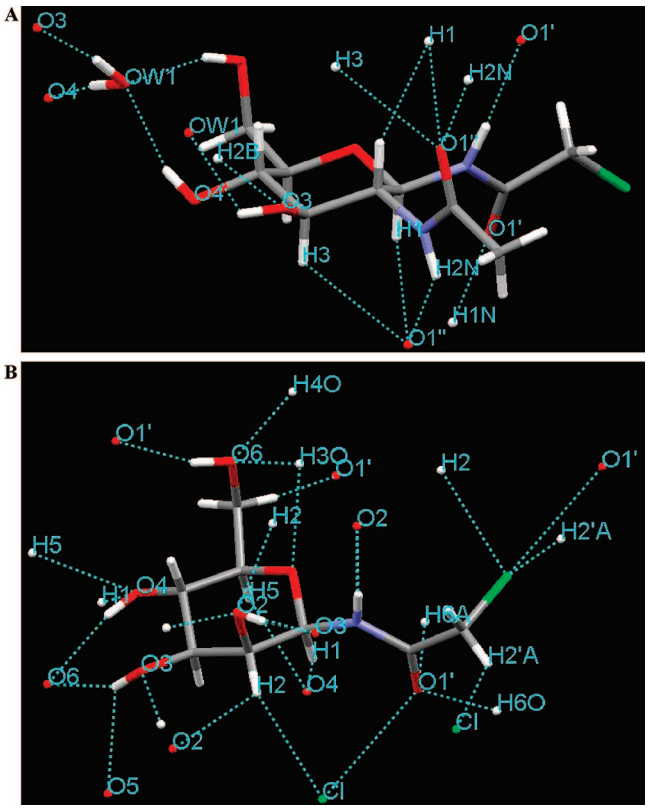

Figure 7. Representation of the hydrogen bonds and interaction network in the crystalline environment of GlcNAcβNHCOCH$_2$Cl (**3**), above, and ManβNHCOCH$_2$Cl (**6**), below.

molecular assembly connecting the above-described 7-membered ring with the hexopyanose via a 5-membered ring involving C5, C6, and O6 atoms and the concomitant stabilization of the *gg* conformation of the hydroxymethyl group. As compared to that of **3**, the molecular assembly of **6** is stabilized by a greater number of C−H···O/Cl interactions (Figure 7B). One of these involves the Cl atom of the aglycon moiety serving as a bifurcated acceptor for the symmetry related C2′−H and C2−H. In turn, C2−H is hydrogen bonded to O2 making the former (C2−H) a bifurcated donor. The carbonyl oxygen of the chloroacetamido group, O1′ lends further stability to the conformation of the CH$_2$Cl group through van der Waals contact with Cl atom, besides being an acceptor for C6−H. Another bifurcated acceptor type hydrogen bonding pattern is also noticed that involves C1−H and C5−H as donors and O4 as acceptor.

Discussion and Conclusions

The amount of data obtained in the present work allow for a comparison between the conformations observed in the crystal structures with the ones predicted by quantum chemistry calculations. All observed and calculated orientations for the χ_2 torsion angles have been reported in Figure 8 and Table 1.

The preferred trans orientation that is predicted for nonhalogenated compounds is somewhat altered by the packing forces for GlcNAcβNHPr and GlcβNHPr. For halogenated compounds, the differences are much stronger since all solid state conformations are strongly different from ab initio predicted orientation. Indeed, the calculations are correct, since they are validated by the statistics obtained from large number of compound in CSD, but the conformation about C1′−C2′ bond in each of the crystalline compounds appears to be accounted to a large extent by the presence or absence of intermolecular interactions with respect to CH$_2$Cl moiety. The chloromethyl group in GlcNAc has no inter or intramolecular contacts, while the Glc, Gal, and Man derivatives have several contacts involving C−H···O/Cl interactions and van der Waals contacts which stabilize the higher energy conformers (Supporting Information).

The double-pillared bifurcated hydrogen bond network in GlcNAc analogs is seemingly a strong and dominant force in molecular packing. As a result, the CH$_2$Cl group of GlcNAc analogs need not lend their atoms for packing, and take up the anti conformation (χ_2 value of 173°) that corresponds to the secondary energy minimum in the calculations. The increasing order in the number of interactions involving the CH$_2$Cl group (Figure 7 and Supporting Information) also correlates well with the observed order, **3**< **4** < **5** < **6**, of their C2′−Cl bond lengths (Table 2).

The present study has also shown the importance of weak hydrogen bonds known as unconventional hydrogen bonds involving C−H···X (O/N/Cl) interactions, in the conformation and packing. These forces are now considered as important as the conventional hydrogen bonds in crystal packing and molecular recognition.[30] The readiness with which the weaker C−H···X interactions can be formed and broken constantly at physiological temperature makes their role in biological processes very special.

Finally, the crystallographic and computational results of the present study have been performed on models and analogs that are different from the "real" situation in *N*-glycoproteins where the χ_2 bond is part of the aglycon-peptide linkage. We have therefore to be careful when extending our conclusion to *N*-glycoproteins. Nevertheless, we now have evidence pointing out that the C2−NHAc group of GlcNAc controls the rotational freedom about the C1′−C2′ bond and helps to ensure minimal variations with respect to χ_2. Statistical analysis of crystallographic data on glycoproteins has shown that nonglycosylated Asn residues display a large range of conformations around the χ_2 linkage as compared to the glycosylated Asn residues that populate the trans conformer.[29] Similarly, the conservation of the extended side chain conformation (*t*) seen among the GlcNAc derivatives, GlcNAcβAsn, GlcNAcβNHPr (**1**) and GlcNAcβNHCOCH$_2$Cl (**3**) is also in sharp contrast to the larger spread of χ_2 values and the near gauche conformation observed in the present study for the sugar derivatives - GlcβNHPr (**2**), GlcβNHCOCH$_2$Cl (**4**), GalβNHCOCH$_2$Cl (**5**), and ManNHCOCH$_2$Cl (**6**)—that lack the NHAc group at C2 position. Though there is a large body of literature pointing out that *N*-glycosylation leads to decreased flexibility of peptide chain, the present study, for the first time, indicates through the combined use of X-ray crystallography and quantum chemical calculations, the critical role of the C2−NHAc group of GlcNAc in controlling χ_2 at the linkage region of *N*-glycoproteins and conserving the extended side chain conformation.

Materials and Methods

Ab Initio Quantum Chemical Calculations. Ab initio and Density Functional Theory have been utilized as complementary tools to explore the χ_2 conformational characteristics of *N*-glycoprotein linkage region models. The starting geometry for the quantum mechanical analyses has been considered from the crystal structure of the investigated systems. All crystal geometries were

(30) Desiraju, G. R.; Steiner, T. D., *The Weak Hydrogen Bond: In Structural Chemistry and Biology* The Oxford University Press: Oxford, 1999.

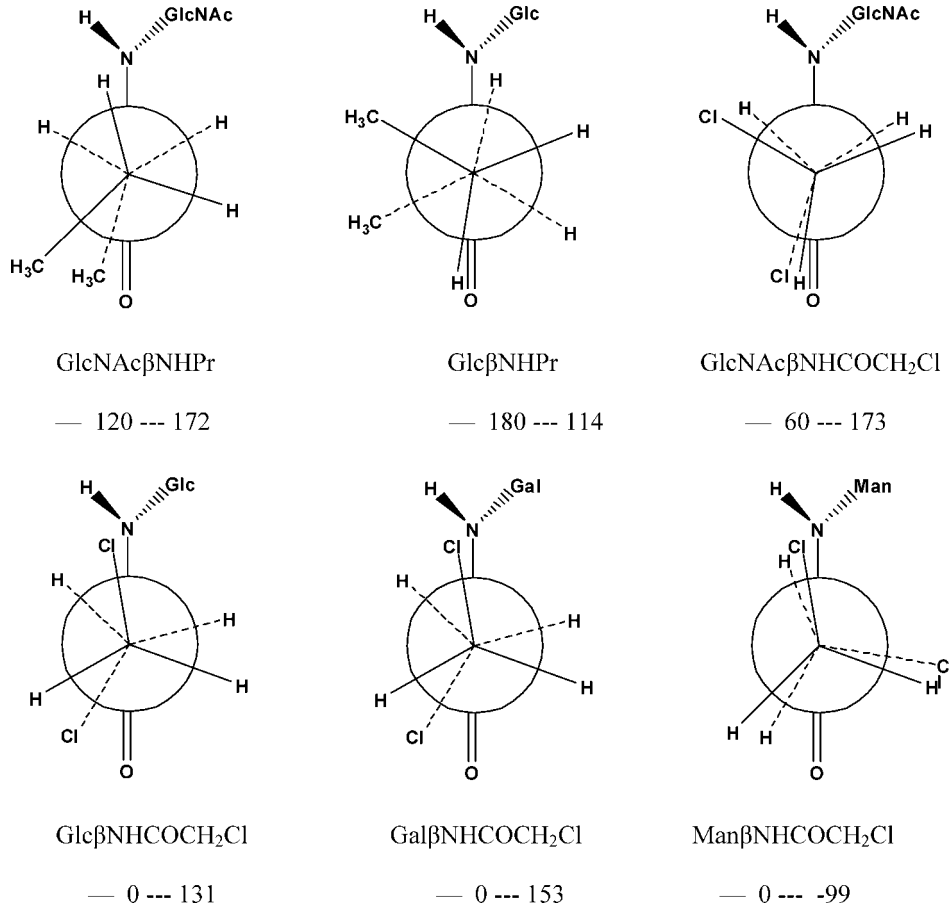

Figure 8. Newman projections for calculated (solid lines) and crystallographic (dashed lines) orientations of the χ_2 torsion angle for compounds **1** to **6**.

Table 6. Data Collection and Refinement Statistics for Compounds **3** and **6**

parameter	GlcNAcβNHCOCH$_2$Cl (**3**)	ManβNHCOCH$_2$Cl (**6**)
empirical formula	$C_{10}H_{19}N_2O_7Cl$	$C_8H_{14}NO_6Cl$
formula weight	314.72	255.65
wavelength	0.71073 Å	0.71073 Å
crystal system	monoclinic	orthorhombic
space group	$P2_1$	$P2_12_12_1$
cell dimensions	a = 4.919(3)	a = 6.3408(19)
	b =7.856(8)	b = 8.5671(8)
	c = 18.784(7)	c = 19.352(3)
	α = 90	α = 90
	β=91.27(4)	β = 90
	γ = 90	γ = 90
volume (Å^3)	725.7(9)	1051.2(4)
absorption coefficient (mm^{-1})	0.101	0.378
F(000)	332	536
crystal size (mm)	0.3 × 0.2 × 0.2	0.3 × 0.2 × 0.2
theta range (°)	2.17 to 24.98	2.10 to 24.96
index ranges	$0 \le h \le 5$, $0 \le k \le 9$, $22 \le l \le 22$	$0 \le h \le 7$, $0 \le k \le 10$, $0 \le l \le 22$
reflections collected /unique	1537/1365 [R(int) = 0.2288]	1105/1105 [R(int) = 0.0000]
data/restraints/parameters	1365/4/203	1105/2/166
final R indices [$I > 2$sigma(I)]	$R1$ = 0.0509, w$R2$ = 0.1413	$R1$ = 0.0398, w$R2$ = 0.1002
R indices (all data)	$R1$ = 0.0852, w$R2$ = 0.1848	$R1$ = 0.0458, w$R2$ = 0.1078

fully optimized using HF and B3LYP with 6−31+G* basis set level with Berny optimization algorithm. Systematic conformational search has been done around χ_2 (N1−C1′−C2′−C3′) in 12 steps of 30° each using the HF/6−31+G*//HF/6−31+G* and B3LYP/6−31+G*//6−31+G* levels of theory. To locate the true minimum on the potential energy surface, full optimization has also been performed for certain selected minima. All of the quantum chemical calculations were carried using Gaussian03.[31]

Search in Crystal Structures Databases. Statistical analysis of torsional distribution were carried out using the data obtained both from PDB[23] and CCDC-ConQuest.[32] Database analyses of *N*-glycoproteins were done with high resolution (<2 Å) X-ray structures selected using the database-mining tool Gly Vicinity.[33] A total of 469 fragments were identified within 4 Å radius from

(31) Frisch, M. J.; et al. *Gaussian 03*, Revision C 02; Gaussian, Inc.: Wallingford, CT, 2004.
(32) Allen, F. H.; Kennard, O. *Chem. Des. Autom. News* **1993**, *8*, 31.
(33) Lutteke, T.; Frank, M.; Von der Lieth, C. W. *Nucleic Acids Res. 33*, 242.

GlcNAc search criteria. Among these, only 350 fragments were chosen for the statistical analysis after neglecting the remaining as distorted fragments. Database search for N−C−C−C and N−C−C−Cl torsions were done using CSD version 5.28 (November 2006) plus two updates with $HNCOCH_2CH_3$ and $HNCOCH_2Cl$ as search fragments.

Preparation of Haloacetamido Sugars. Both GlcNAcβ-$NHCOCH_2Cl$ (**3**) and Man$\beta NHCOCH_2Cl$ (**6**) were prepared in 72 and 60% isolated yield from their corresponding β-D-glycopyranosylamines by selective *N*-chloroacetylation following the same procedure reported earlier by us for making the Glc and Gal analogs.[12]

Crystal Structures Solution and Refinement. Crystallization of compound **3** and **6** was done in aqueous methanol by the slow evaporation method at room temperature. X-ray diffraction data were collected at room temperature in the ω-2θ scan mode on an Enfra-Nonius CAD4 diffractometer and the relevant details of data collection and refinement are given in Table 6. The structure was solved by direct methods using SHELXS-97[34] and the refinement was done by full matrix using SHELXL-97. The packing diagrams were drawn using Mercury 1.4.2.

Complete structural data of the chloroacetamido sugars derived from GlcNAc and Man (CCDC Nos. 232234 and 232235, respectively) have been deposited at the Cambridge Crystallographic Data Centre, which can be obtained free of charge via www.ccdc.cam.ac.uk/data_request/cif (or from the Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB21EZ, United Kingdom; Fax: (+44) 1123−336−033. or E-mail: deposit@ccdc.cam.ac.uk).

Acknowledgment. We are grateful to the Indo-French Centre for Promotion of Advanced Research (IFCPAR), New Delhi, for the financial support and encouragement. Funding provided by the Department of Science and Technology, New Delhi, for the purchase of the 400 MHz NMR under IRPHA Scheme and ESI−MS under the FIST program to the Department of Chemistry, IIT Madras is gratefully acknowledged. The authors (U.A., B.V., and D.L.) thank the Sophisticated Analytical Instrumentation Facility (SAIF), IIT Madras for the X-ray data collection. One of us (U.A.) is thankful to the Council of Scientific and Industrial Research (CSIR), New Delhi, for award of a Junior Research Fellowship. We also thank Cambridge Crystallographic Data Centre (CCDC), United Kingdom, for making the program Mercury 1.4.2 available for use.

Supporting Information Available: Representation of the crystallographic environment directing the orientation of χ_2 for **1**, **2**, **4**, and **5**; crystallographic conformation of **1**–**6**; X-ray crystallographic data for the chloracetamido sugars derived from GlcNAc and Man. This material is available free of charge via the Internet at http://pubs.acs.org.

JA800335M

(34) Sheldrick, G. M. *SHELXL97, Program for Crystal Structure Refinement*; University of Gottingen: 1997.