Organic & Biomolecular Chemistry

Cite this: *Org. Biomol. Chem.*, 2012, **10**, 3019

www.rsc.org/obc

PAPER

Combining weak affinity chromatography, NMR spectroscopy and molecular simulations in carbohydrate–lysozyme interaction studies†

Jens Landström,[a] Maria Bergström,[b] Christoffer Hamark,[a] Sten Ohlson[b] and Göran Widmalm*[a]

Received 9th December 2011, Accepted 1st February 2012
DOI: 10.1039/c2ob07066a

By examining the interactions between the protein hen egg-white lysozyme (HEWL) and commercially available and chemically synthesized carbohydrate ligands using a combination of weak affinity chromatography (WAC), NMR spectroscopy and molecular simulations, we report on new affinity data as well as a detailed binding model for the HEWL protein. The equilibrium dissociation constants of the ligands were obtained by WAC but also by NMR spectroscopy, which agreed well. The structures of two HEWL–disaccharide complexes in solution were deduced by NMR spectroscopy using ^{1}H saturation transfer difference (STD) effects and transferred ^{1}H,^{1}H-NOESY experiments, relaxation-matrix calculations, molecular docking and molecular dynamics simulations. In solution the two disaccharides β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-OMe and β-D-Glc*p*NAc-(1→4)-β-D-Glc*p*NAc-OMe bind to the B and C sites of HEWL in a *syn*-conformation at the glycosidic linkage between the two sugar residues. Intermolecular hydrogen bonding and CH/π-interactions form the basis of the protein–ligand complexes in a way characteristic of carbohydrate–protein interactions. Molecular dynamics simulations with explicit water molecules of both the *apo*-form of the protein and a ligand–protein complex showed structural change compared to a crystal structure of the protein. The flexibility of HEWL as indicated by a residue-based root-mean-square deviation analysis indicated similarities overall, with some residue specific differences, *inter alia*, for Arg61 that is situated prior to a flexible loop. The Arg61 flexibility was notably larger in the ligand-complexed form of HEWL. *N*,*N*′-Diacetylchitobiose has previously been observed to bind to HEWL at the B and C sites in water solution based on ^{1}H NMR chemical shift changes in the protein whereas the disaccharide binds at either the B and C sites or the C and D sites in different crystal complexes. The present study thus highlights that protein–ligand complexes may vary notably between the solution and solid states, underscoring the importance of targeting the pertinent binding site(s) for inhibition of protein activity and the advantages of combining different techniques in a screening process.

Introduction

Studies of biomolecular interactions are of great importance in order to understand biological phenomena and for guidance in the quest for new drugs. A variety of different biophysical methods are available for research involving molecular interactions. The different methods have their own advantages and disadvantages, making combinations of techniques an attractive approach to extract reliable data in an effective manner.

Weak affinity chromatography (WAC) was introduced about twenty years ago, combining high performance liquid chromatography with affinity chromatography.[1] It was shown that an efficient separation system could be obtained by immobilizing low affinity ligands ($K_D > 10^{-5}$ M) onto microparticulate porous silica. The WAC technique has been shown to be a valuable tool especially in characterizing protein–carbohydrate interactions since these often are of low affinity.[2–4] Normally the protein (for example an antibody, lectin or toxin) is immobilized and small amounts of carbohydrate solutions are injected into the column. From chromatography theory the retention volume of an injected substance should be directly related to the affinity and the number of binding sites under standardized conditions.[5] The K_D value of the interaction can therefore be easily extracted from the retention in combination with the number of binding sites, or by comparing the retention in relation to a compound of known affinity. The advantages with the WAC technique include low consumption of analyte (injected substance) and short analysis time, since K_D might be obtained from a single injection at low concentration. Another benefit is that racemates, sample

[a]*Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden. E-mail: gw@organ.su.se*
[b]*School of Natural Sciences, Linnaeus University, S-391 82 Kalmar, Sweden*

†Electronic supplementary information (ESI) available. Experimental details, characterizations and additional results. See DOI: 10.1039/c2ob07066a

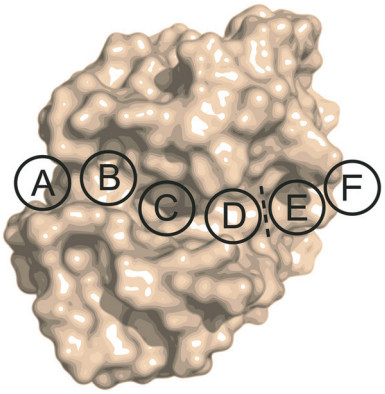


Fig. 1 The binding cleft of HEWL is divided into six subsites depicted as A–F. Each subsite can accommodate one monosaccharide unit. The hydrolysis site is situated between subsites D and E. The HEWL structure is from ref. 23.

mixtures and impure substances can be analyzed, and that the K_D of the individual components in the mixture can be determined concomitantly.

During the past decade, NMR spectroscopy has become an important tool in molecular interaction studies.[6] A major advantage is that the molecules can be studied label-free in solution. NMR spectroscopy gives valuable dynamic as well as structural information of both the target protein and the ligand on an atomic level. In order to predict, understand and interpret experimental results it is often necessary to implement computational methods. Molecular docking simulations are being used to predict the structure of intermolecular complexes[7,8] and are often used to screen libraries of compounds *in silico*. Docking gives a static picture of the binding event and if information on dynamic properties is required, more computationally demanding molecular dynamics (MD) simulations can be carried out.[9,10] MD simulations give information about the intra- and intermolecular motions of molecules over a time period. Results from docking and MD simulations can be translated to NMR spectroscopy data by relaxation-matrix calculations[11,12] and compared to experimental results. By combining these two experimental methods, WAC and ligand detected NMR spectroscopy, with molecular docking, relaxation-matrix calculations and molecular dynamics simulations, the high throughput characteristics of WAC can be complemented with qualitative and quantitative information on an atomic level.

Hen egg-white lysozyme (HEWL) is being used as a model protein in a great number of structural[13–18] and mechanistic[19,20] studies. HEWL is known to catalyze the hydrolysis of the glycosidic bond of the β-(1→4)-linkage between *N*-acetyl-muramic acid (MurNAc) and *N*-acetyl-D-glucosamine (GlcNAc) in peptidoglycans that constitute the cell wall of Gram-positive bacteria. It is also known to hydrolyze the β-(1→4)-linkages in chitin.[21] HEWL has a large binding cleft that can be divided into six subsites, A–F, each one can contain one monosaccharide unit and the hydrolysis is known to take place between subsites D and E (see Fig. 1).[19] Furthermore, HEWL binds chitin polymers of different lengths and it has been reported that GlcNAc binds to subsites C and E,[22] *N*,*N*′-diacetylchitobiose in the CD site[23] and chitotriose in the BCD as well as in the ABC sites.[23]

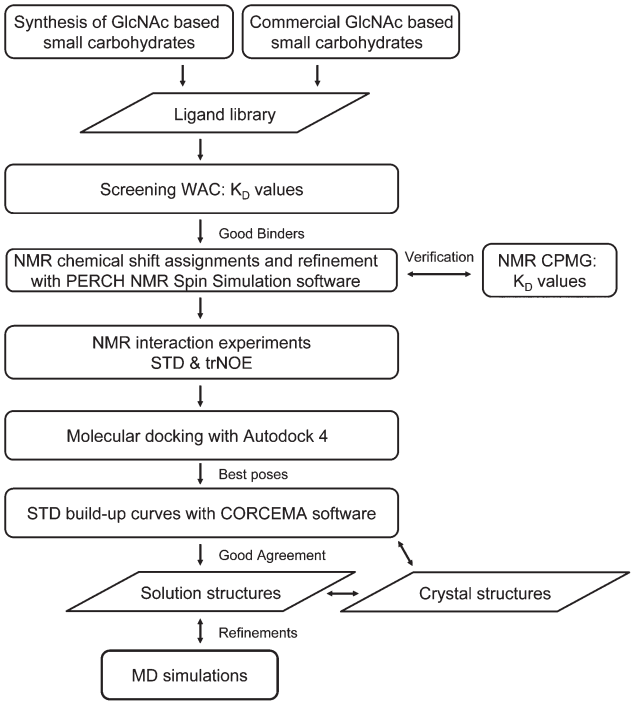


Fig. 2 Flowchart for the generation of HEWL–ligand structure complexes in solution.

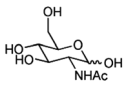
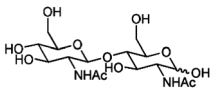
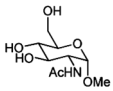
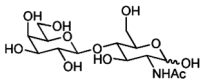
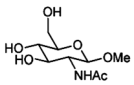
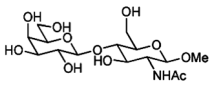
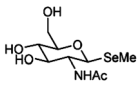
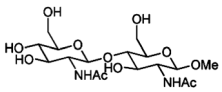
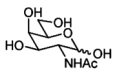
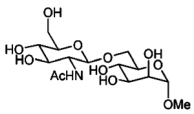
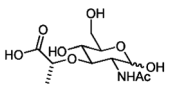
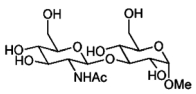
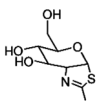
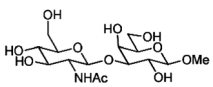
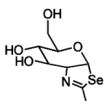
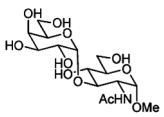
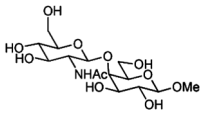
We have chosen HEWL as a model system to investigate protein–ligand interactions with a combination of weak affinity chromatography (WAC), NMR spectroscopy and computational methods. This study provides the scientific community with efficient quantitative and qualitative methodology for ligand–protein interaction studies.

Results and discussion

Protocol

A process of ligand screening, in combination with binding and structure analysis, is outlined in Fig. 2. In the present study the interactions between HEWL and small carbohydrates were analyzed. The ligands were obtained either by chemical synthesis or from commercial sources and together they make up a library of small carbohydrate compounds (Table 1). The screening of the ligand-binding to HEWL first employed the high-throughput technique weak affinity chromatography which resulted in K_D values for the ligands. For the subsequent NMR experiments they were characterized with respect to, in particular, ^{1}H NMR chemical shifts. By subsequent NMR titration experiments using a T_2 relaxation experiment it was possible to obtain K_D values from an orthogonal technique. The interactions at an atomic level were then investigated by STD (saturation transfer difference) and trNOE (transfer nuclear Overhauser effect) NMR experiments. *In silico* molecular docking was able to propose binding poses which were possible to evaluate by comparison to theoretically calculated STD build-up curves. Finally, the protein–ligand solution structures may then be compared to crystal structures and refined using MD simulations.

Table 1 Schematic structures of compounds studied by weak affinity chromatography and NMR spectroscopy (**11**) with corresponding K_D values

Structure	No.	K_D/mM	Structure	No.	K_D/mM
	1	$\alpha = 27$ $\beta = 47$		9	0.27
	2	46		10	0.64
	3	43		11	0.62 0.4[a]
	4	42		12	0.29
	5	$\alpha = 31$ $\beta = 90$		13	>100
	6	61		14	38
	7	38		15	>100
	8	24		16	>100
				17	>100

[a] Measured by NMR spectroscopy.

Synthesis

The thio-analogue of the oxazoline derivative of *N*-acetylglucosamine,[24] also known as thiazoline (**7**), is an inhibitor of NAGases[25] and hence a potential inhibitor of HEWL. To test this hypothesis, thiazoline and its seleno-analogue **8**, were synthesized (Scheme S1†). Synthesis of selenazoline **8** has not been previously published, so a new synthetic procedure based on the synthesis of **7** was developed. The only difference between the syntheses of **7** and **8** was that Woollins' reagent[26] was used in the synthesis of **8** instead of Lawesson's reagent.[27] The disaccharide methyl glycosides, **11** and **12**, were synthesized according to procedures that are described in Schemes S2 and S3.†

Frontal chromatography

The number of active binding sites on the HEWL silica was assessed by frontal chromatography at pH 5.5. The analysis was performed with newly dissolved monosaccharide (**1**, α-anomer) and disaccharide (**10**, anomeric mixture) in order to determine if these saccharides interacted differently. Analysis of α-D-Glc*p*NAc is possible as mutarotation of monosaccharides is a rather slow process at ambient temperature.[28,29] Break-through curves in the frontal analysis were determined in less than ten minutes after the sample was dissolved and the derived values were therefore judged to be a good approximation of α-D-Glc*p*NAc binding. It was found that the concentration of binding sites in the column was 4.8 mM when determined with the monosaccharide (**1**) and 3.0 mM with the disaccharide (**10**). These values were compared with the immobilized amount of HEWL that were estimated to 133 mg g^{-1}, corresponding to a HEWL concentration of 4.7 mM (silica density = 0.5 g mL^{-1}). Previous immobilizations of other proteins have shown that about 50% of the binding sites are inactivated after immobilization, probably due to the random orientations in coupling.[2,3] This implies that the expected number of sites on the column is about 2.3 mM (assuming one site per HEWL), a value that is close to what was obtained with **10**, and it seems reasonable to assume a 1 : 1 interaction between **10** and HEWL. This result contrasts with the situation with monosaccharide **1** where an interaction model with two independent sites on HEWL is most likely, as the concentration of binding sites in this analysis was

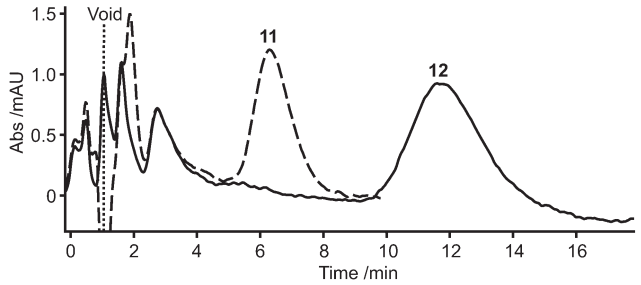


Fig. 3 Elution profiles of disaccharides **11** (dashed line) and **12** (solid line) on a WAC HEWL column (3 × 0.21 cm). See Materials and methods section for chromatography conditions.

2.1 times higher than the calculated concentration of HEWL. The frontal analysis therefore indicates that disaccharides and monosaccharides interact differently with HEWL. A number of studies have shown that *N*-acetylglucosamine-containing oligosaccharides may bind in several positions close to subsite C in the binding cleft.[13,23,30] X-ray crystallography studies have found **1** mainly in site C[31,32] but other studies, performed in solution with NMR, has indicated several binding possibilities.[22] The WAC results imply that the monosaccharide is small enough to interact with at least two of these sites at the same time.

Weak affinity chromatography (WAC)

Lysozyme has previously been reported to have an optimal binding between pH 4 and 6.[33] This was confirmed by determining the retention of **1** and **5** with mobile phases buffered to pH 3.0, 5.5 and 7.0 (0.1 M sodium phosphate, data not shown), using the 50 × 0.46 cm HEWL column. Both compounds had a maximal retention at pH 5.5 and all library compounds were therefore analyzed using this mobile phase. The 50 × 0.46 cm HEWL column was used for all analytes except for **9–12** that were run on the 3 × 0.21 cm HEWL column because of their higher affinity. The WAC analysis of **11** and **12** is shown in Fig. 3. The affinity was calculated using the relation $K_D = B_{max}/\Delta v$ where K_D is the equilibrium dissociation constant, Δv is the retention volume and B_{max} is equal to the number of binding sites present in the column.[3] All monosaccharides as well as **7** and **8** were assumed to recognize the same number of sites as monosaccharide **1**, which was calculated from the frontal analysis to be 40.2 μmol regarding the 50 × 0.46 cm column. All disaccharides were assumed to be similar to **10**, resulting in 24.6 μmol and 0.31 μmol binding sites for the 50 × 0.46 cm and 3 × 0.21 cm HEWL columns, respectively. K_D values of all compounds are presented in Table 1.

The K_D values of the α- and β-anomeric forms of **1** were determined concomitantly by analyzing the anomeric mixture at equilibrium. This is possible since the chromatographic separation is faster than the anomeric re-equilibration.[28,29] Repeated analysis at 15 minute intervals of newly dissolved α-D-Glc*p*NAc (the α-anomeric form was determined by ^{1}H NMR) on the 50 × 0.46 cm HEWL column confirmed that the anomeric equilibrium was reached after approximately one hour (data not shown). It was found that α-D-Glc*p*NAc interacts with HEWL with higher affinity compared to the β-anomeric form. The analysis of **1** was compared with the α- and β-methyl glycosides of GlcNAc (**2** and **3**), which exhibited K_D values in the same range as β-D-Glc*p*NAc. A reason for this could be that the anomeric hydroxyl proton in the α-anomeric form of **1** is involved in hydrogen bonding with the protein.[31] The affinities of the other monosaccharides and the thio- and selenoanalogues of the oxazoline derivatives of *N*-acetylglucosamine (**7** and **8**) were in the same range as **1** with K_D values ranging from 24 to 90 mM. The group of *N*-acetylglucosamine-containing disaccharides, on the other hand, had a high variation in HEWL affinity. The suggested reason for the difference in behavior is that the small monosaccharides, which all contain the key *N*-acetyl group, interact in a similar way with HEWL, while the interaction with larger *N*-acetyl-containing saccharides are more selective. Compounds **9–12** being β-(1→4)-linked disaccharides, having at least one *N*-acetyl-D-glucosamine residue, binds with <1 mM affinity whereas disaccharides **13–17**, which also have a D-Glc*p*NAc residue but where anomeric or ring-carbon configurations as well as linkage positions are different, do not represent compounds with any significant binding affinity, *i.e.*, $K_D \geq 38$ mM. The low affinity of the latter compounds are presumably due to steric clashes in the binding site, thereby preventing any tighter binding even though the monosaccharide D-Glc*p*NAc residue of these compounds binds to site C in HEWL.

The concentrations of the injected saccharide in weak affinity chromatography should be less than the K_D of the interaction in order to not overload the column. Overriding this prerequisite will induce the compound to elute earlier resulting in an erroneous (higher) K_D value. Due to detection limitations the concentration of saccharides **9** and **12** was analyzed at a concentration close to the obtained K_D value (0.3 mM), which means that the affinity could be slightly underestimated. The analysis error due to a nonlinear behavior is however likely to be small because of the sample dilution in the column (band-broadening effects). The obtained K_D values of **1** as well as **9** do also agree well with previous studies performed at similar conditions using NMR and microcalorimetry.[22,34–36] The reported K_D values are ≈30 mM for *N*-acetyl-D-glucosamine (**1**) and ≈0.3 mM for *N*,*N*′-diacetylchitobiose (**9**) (*cf.* Table 1).

The thio- and selenoanalogues **7** and **8** of the oxazoline derivatives of *N*-acetylglucosamine were bound only weakly to HEWL (Table 1) indicating that they do not represent the hydrolysis reaction intermediate[25,37] thereby supporting the mechanism with a covalently linked intermediate, proposed by Vocadlo *et al.*[20] and confirmed by QM/MM simulations performed by Bowman *et al.*[38] instead of participation by the acetamido group.

NMR spectroscopy

NMR assignments of ^{1}H and ^{13}C chemical shifts at 343 K and ^{1}H chemical shifts at 278 K were performed for **11** and **12**. ^{1}H chemical shifts and coupling constants were refined using the NMR spin simulation software PERCH and are summarized in Table S1† for **11** and Table S2† for **12**. Subsequently, the protein–ligand dissociation constant, K_D, was measured for **11** using ligand observed transverse relaxation rates, R_2.[39] The HEWL concentration was kept constant and T_2 CPMG

experiments were performed on **11** at eight different concentrations. By plotting the inverse of the difference between R_2 of the ligand in exchange with lysozyme and R_2 of the free ligand *versus* ligand concentrations a K_D of 0.4 mM was extracted (ESI†). This value is in very good agreement with that measured by WAC (0.6 mM). The small difference between the techniques may be due to immobilization effects and minor pH differences because the NMR experiments were run in D_2O solution instead of water.

Saturation transfer difference (STD) experiments[40] in which binding epitopes on the ligand may be identified were performed on HEWL in complex with **11** and **12**, using saturation times of 0.5, 1, 2 and 4 s and two different saturation frequencies, 0.5 and 7.0 ppm. The STD NMR effects for **11** and **12** in complex with HEWL at different saturation times and saturation frequencies are compiled in Tables S4 and S5.† We also carried out control STD experiments without protein present and did not observe any STD effects. In Fig. 4, conspicuous STD effects are observed for the methyl resonance of the *N*-acetyl group at the reducing end residue of both disaccharides when a 0.5 s irradiation is set at 0.5 ppm (Fig. 5a and 5b), indicating that the protons are in close proximity to methyl groups in the protein. The methyl resonance of the *N*-acetyl group at the reducing end residue and the $H6_{pro\text{-}S}'$ in the terminal residue of both disaccharides give notable STD effects when irradiation is set at 7 ppm (Fig. 5c and 5d), indicating that these protons are in close contact with aromatic residues, *e.g.* Trp, in the protein complex (Fig. 6a and 6b; poses obtained from docking, *vide infra*). These results may be compared to a crystal structure of the HEWL–**9** complex,[23] where the methyl group of the *N*-acetyl group in the terminal residue is in close contact with methyl and aromatic protons in the protein (Fig. 6c). A published crystal structure of HEWL in which **10** as a glycoside is covalently attached as an ester to Asp52 *via* a three-carbon linker, pdbid: 1UC0,[41] displays a disaccharide binding mode (Fig. 6d) similar to that obtained herein based on experimental NMR data in solution. Disaccharides **11** and **12** show very similar STD patterns, indicating that they have similar binding modes.

In solution a 1D $^1H,^1H$-NOESY experiment at 298 K and 700 MHz of β-D-Glc*p*NAc-(1→4)-β-D-Glc*p*NAc-OMe (**12**) with selective excitation of the resonance from H1′ and a mixing time of 300 ms revealed a prominent NOE to H4 across the glycosidic linkage, consistent with a *syn* conformation.[42] This result is in complete agreement with the transient NOE experiment performed on **12** at 293 K and 400 MHz which showed a strong NOE between H1′ and H4, besides an intra-residue NOE between H1′ and H5′.[43] An MD simulation of **12** with explicit water molecules as solvent at 290 K[43] had average torsion angles $\phi_H \approx 59°$ and $\psi_H \approx 15°$ over the 50 ps duration where the molecule was confined to a single conformational state at the glycosidic linkage, *i.e.*, a *syn* conformation.

Further information on ligand-bound conformation comes from trNOE experiments[44,45] of both complexes that show strong negative cross-peaks between H1′ and H4, indicating a *syn* conformation also in the bound state. This is exemplified for **12** at 289 K and 700 MHz showing a stronger inter-residue NOE from H1′ to H4 than the intra-residue NOEs to H3′ and H5′ (Fig. 7). At this temperature and magnetic field strength the NOE was close to zero for a ligand preparation of **12** without protein. Thus, in the tr-NOESY experiments the contribution from the ligand in solution should be negligible compared to that originating from the protein–ligand complex. This trNOE result is fully consistent with a *syn* conformation of **12** when bound to HEWL. A weak cross-peak between the methyl resonance of the *N*-acetyl group of the reducing end sugar residue and the methyl

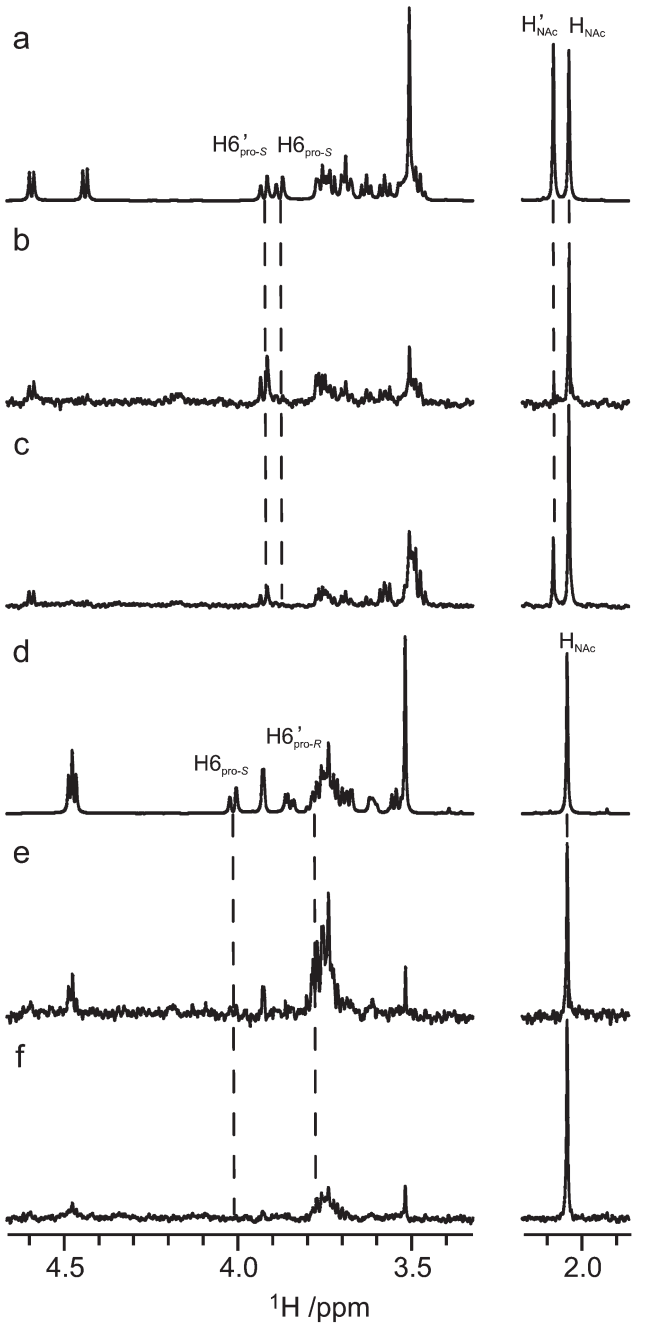


Fig. 4 STD NMR and ^1H NMR reference spectra of HEWL and **12** (a–c) and **11** (d–f) showing two different irradiation (0.5 s) positions. (a) ^1H reference spectrum of **12** and HEWL in a ratio of 50 : 1; (b) STD spectrum of **12** and HEWL with irradiation at 7 ppm; (c) STD spectrum of **12** and HEWL with irradiation at 0.5 ppm; (d) ^1H reference spectrum of **11** and HEWL in a ratio of 30 : 1; (e) STD spectrum of **11** and HEWL with irradiation at 7 ppm; (f) STD spectrum of **11** and HEWL with irradiation at 0.5 ppm.

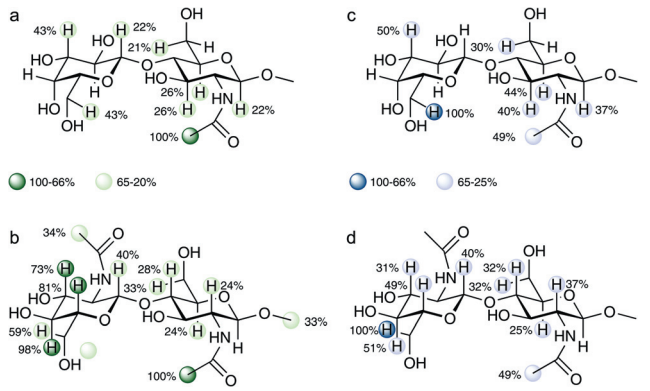


Fig. 5 Normalized levels of saturation (%) from ^{1}H STD NMR spectroscopy of HEWL in complex with (a) **11** and (b) **12** (on-resonance irradiation δ = 0.5 ppm, off-resonance irradiation δ = 60 ppm), (c) **11** and (d) **12** (on resonance irradiation δ = 7.0 ppm, off-resonance irradiation δ = 60 ppm). Protons included in the figures indicate detected STD effects.

region of the protein could also be observed in the 2D-NOESY spectra.[46] Similar results were obtained by 2D-NOESY experiments for the complex between compound **11** and HEWL.

Hydrolysis of the methyl glycosides was not detected by ^{1}H NMR spectroscopy, which could have been the case if binding had occurred at site D. Furthermore, hydrolysis of the glycosidic linkage between the sugar residues in either of the disaccharides was not detected, indicating that binding at sites D and E is of limited importance. Moreover, a ^{1}H NMR spectrum of an anomeric mixture of **1** and HEWL showed a difference in the chemical shifts of the methyl resonances of the *N*-acetyl groups of the α- and β-anomeric forms compared to the chemical shifts in absence of HEWL (data not shown). This indicates different binding modes for α- and β-anomeric forms of **1**, which is in agreement with WAC data (see Table 1) and previously published results.[22,31]

Molecular docking and simulation

TrNOE results indicate that the major conformation in the bound state is *syn* for both of the disaccharides, **11** and **12**. Molecular docking of disaccharides **11** and **12** into the crystal structure 1SF4 were performed with the ϕ-torsion angles restrained at ϕ_H = 50° using Autodock 4.[47] The top ranked structure clusters from docking were compared with NMR data and theoretical STD build-up curves, calculated for the complexes using CORCEMA-ST.[12] Fig. 8 shows the docking result for **12** (in blue) that was in best agreement with experimental data. This structure was ranked 6th place by the Autodock scoring function. The ranking of docked ligands still poses significant problems in the evaluation of the resulting poses due to insufficiently developed scoring functions,[7] but in some cases illuminating results on the carbohydrate–protein interactions have been obtained for the most favorable poses,[48] stressing that Autodock is still an attractive tool for carbohydrates.[7] We emphasize that in the present study the selection of reasonable poses is based on docking in conjunction with experimental NMR data. The docked structure for **11**, which is in best agreement with experimental data (ranked 7th place by Autodock), shows a similar binding mode as **12**. Both disaccharides populate the B and C subsites. These results contrast the crystal structure of HEWL with bound *N*,*N*′-diacetylchitobiose[23] in which the disaccharide is present in subsites C and D (shown in red in Fig. 8). Theoretical STD build-up curves, calculated using CORCEMA-ST,[12] of both the proposed docked structure of **12** and the crystal structure with bound *N*,*N*′-diacetylchitobiose are presented in Fig. 9. The docked conformation shows much better agreement with experimental STD NMR data (*cf.* Fig. 4) compared to the crystal structure, giving credence to the interpretations presented herein.

The docked structures of disaccharides **11** and **12** in the BC binding site of HEWL are shown in Fig. 10. The characteristic interactions between the ligands with hydrogen bonding and CH/π-stacking[49–52] are present to different extents. The π-stacking interaction for benzene dimers[53] has been described as parallel-stacked or T-shaped and such arrangements were observed in the crystal structures of molecules that were synthesis intermediates in the present study, *viz.*, ethyl 3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-1-thio-β-D-glucopyranoside[54] having an intramolecular parallel-stacked arrangement and ethyl 4,6-*O*-benzylidene-2-deoxy-2-phthalimido-1-thio-β-D-glucopyranoside[55] showing an intermolecular T-shape-like arrangement. The disaccharide interactions with HEWL differ slightly where the tighter binder compound **12** (*cf.* Table 1) shows a closer stacking (Fig. 10a) than compound **11** (Fig. 10b) which has a lower binding affinity. However, in the latter complex a hydrogen bond is present between HO4′ and the carboxylate group of Asp101 (Fig. 6b and 10b). A corresponding hydrogen bond from hydroxyl groups in **12** that are proximate to Asp101 was not present, neither from HO6′ nor from HO4′. The interaction strength of a classical hydrogen bond and a CH/π-stacking interaction is of similar magnitude.[56,57] The CH/π-stacking interaction between the terminal sugars in the two complexes were further analyzed with respect to distances and angles. In order to compare the two structure complexes, centroids for the H1′/H3′/H5′ protons and centroids for each ring of the indole part of Trp62 were defined as well as a plane based on the three sugar protons and another plane for the indole part. The distance separation in the HEWL-**12** complex (Fig. 11a) is significantly shorter than for the HEWL-**11** complex (Fig. 11b) and whereas the planes are almost parallel in the first case (16°) the angle between the planes is large (43°) in the second case. The geometrical arrangements for both hydrogen bonds and CH/π-stacking interactions are important descriptors to characterize interaction energies for complexes. The former are highly directional whereas the latter show only weak directionality.[58,59] The distance between protons and carbon atoms in the π-plane are on the order of 2.5–3.0 Å for geometrically optimized structures with a favorable CH/π-stacking.[49–52] The somewhat tighter binding of **12** compared to **11** may thus be reasonable to attribute to a stronger CH/π-stacking interaction with Trp62 in HEWL.

To study structure further, as well as dynamics, MD simulations with explicit solvent were performed on HEWL, HEWL-**11** complex from docking, HEWL-**12** complex from docking and HEWL–*N*,*N*′-diacetylchitobiose complex from a crystal structure (1SF4). A CHARMM22 force field with CMAP[60] was used to describe the protein; the carbohydrates were described by PARM22/SU01.[61] The durations of the simulations were

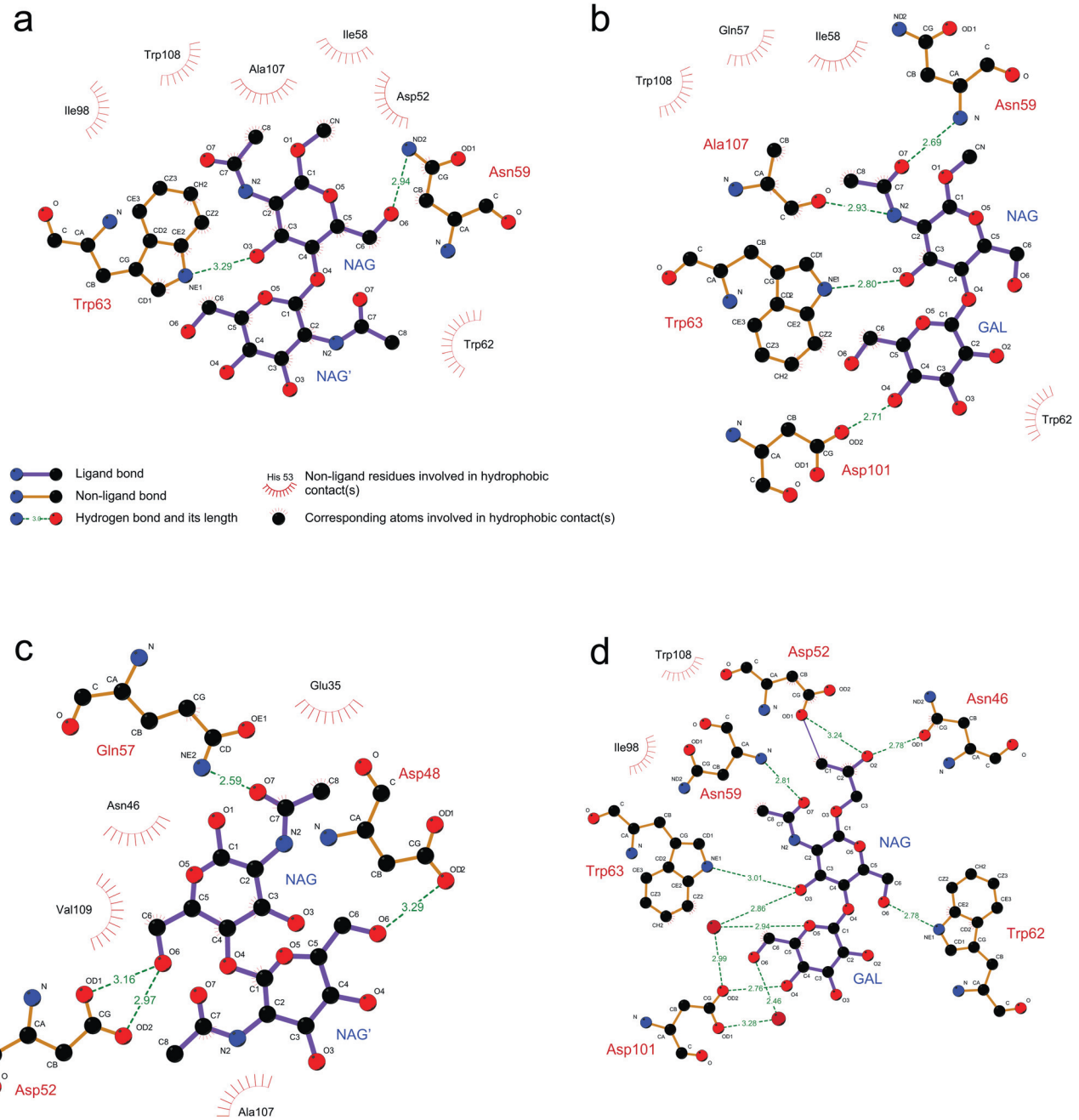


Fig. 6 Possible hydrogen bonding and hydrophobic interactions between HEWL and disaccharide ligands **12** (a), **11** (b), *N*,*N*′-diacetylchitobiose from crystal structure 1SF4 (c) and β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→1)-Gro-3-Asp52 from crystal structure 1UC0 (d). The numerical values are the distances in Å between the donor and acceptor atoms responsible for the interaction.

175 ns for HEWL and the HEWL-**11** complex and 50 ns for the HEWL-**12** complex from docking and the HEWL–*N*,*N*′-diacetylchitobiose complex from a crystal structure. The conformational preference of the glycosidic linkage of **11** in complex with HEWL from the 175 ns simulation is shown in Fig. 12, where the ligand exists in two conformations during the simulation. The second conformation appears after approximately 50 ns and is then stable for the remaining part of the simulation (Table S3†). The MD simulation then suggests that the conformation after 50 ns of simulation is close to the conformation of the disaccharide in complex with HEWL in solution. The glycosidic torsion angles of the docked HEWL-**12** complex adopt one conformation, close to the starting structure, during the 50 ns simulation. Table S3† shows the torsion angle averages of **12** where the ψ_H torsion angle shows a higher flexibility compared to the ϕ_H torsion angle.

The complex of HEWL and *N*,*N*′-diacetylchitobiose (**9**) from crystal structure 1SF4 dissociates during the MD simulation by losing contact between the reducing end residue and the protein binding site D, making the reducing end to protrude into

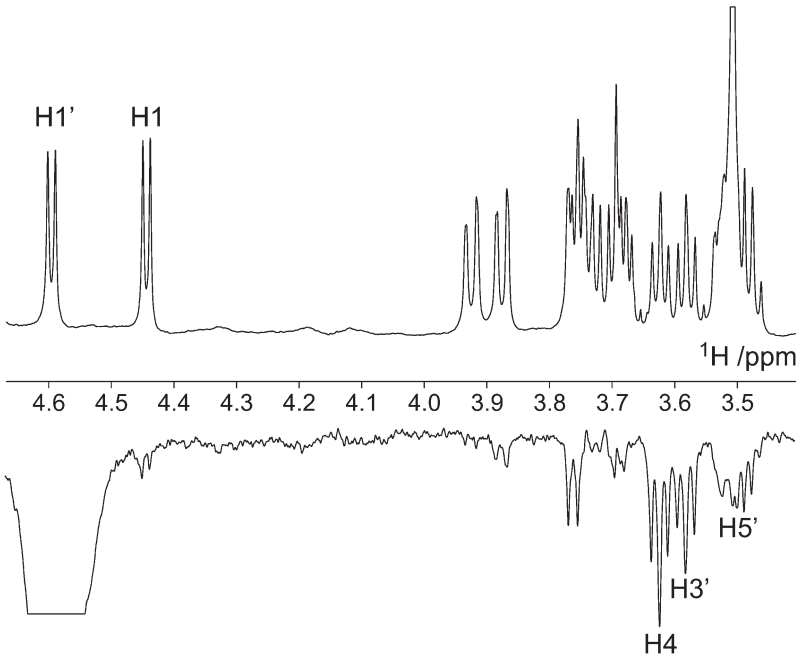


Fig. 7 Part of the ^{1}H NMR spectrum of **12** (1 mM) in the presence of HEWL (25 μM) at 289 K (top); the corresponding 1D DPFGSE ^{1}H,^{1}H-NOESY spectrum with selective excitation of the resonance from H1′ and a mixing time of 300 ms (bottom). Pertinent resonances are annotated.

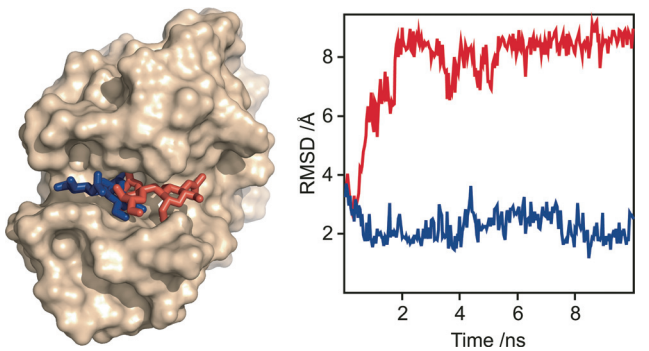


Fig. 8 (left) Complex between HEWL and β-D-Glc*p*NAc-(1→4)-β-DGlc*p*NAc-OMe from docking (blue) and HEWL and *N*,*N*′-diacetylchitobiose from crystal structure 1SF4 (red); (right) time dependence of the RMSD of the atoms in the disaccharides in complex with HEWL, referenced from docking (blue) and crystal structure (red), during 10 ns production runs of the MD simulations.

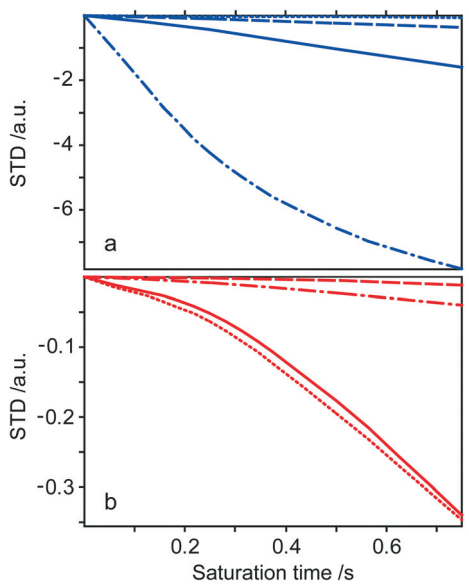


Fig. 9 Theoretical CORCEMA-ST STD build-up curves of **12** in complex with HEWL from proposed structure from docking (a, in blue) and crystal structure 1SF4 (b, in red); irradiation was set at the aromatic protons in the protein. The build-up curves describe the methyl protons of the *N*-acetyl group at the reducing end residue (solid), $H6_{pro\text{-}S}$ in the reducing end residue (dashed), methyl protons of the *N*-acetyl group at the terminal residue (dots) and $H6_{pro\text{-}S}$ in the terminal residue (dashed-dots).

solution. In the right panel of Fig. 8, the time dependency of the RMSD of the ligand compared to the starting structure illustrates that the crystal structure complex (red) breaks early in the simulation, but the docked complex (blue) remains intact. The fact that the crystal structure complex, which has an *anti*-ϕ conformation at the glycosidic linkage of the ligand,[23] does not survive the MD simulation is a strong indication that the crystal structure complex with the disaccharide is not likely to be present in solution.

From the MD simulation it is also evident that the overall backbone structure of HEWL in solution changes compared to the crystal structure (Fig. 13). This occurs both for the *apo*-form and somewhat more for the HEWL-**11** complex where the fluctuations after 50 ns (*vide infra*) take place around new structural equilibria. As an experimental evaluation of the simulated

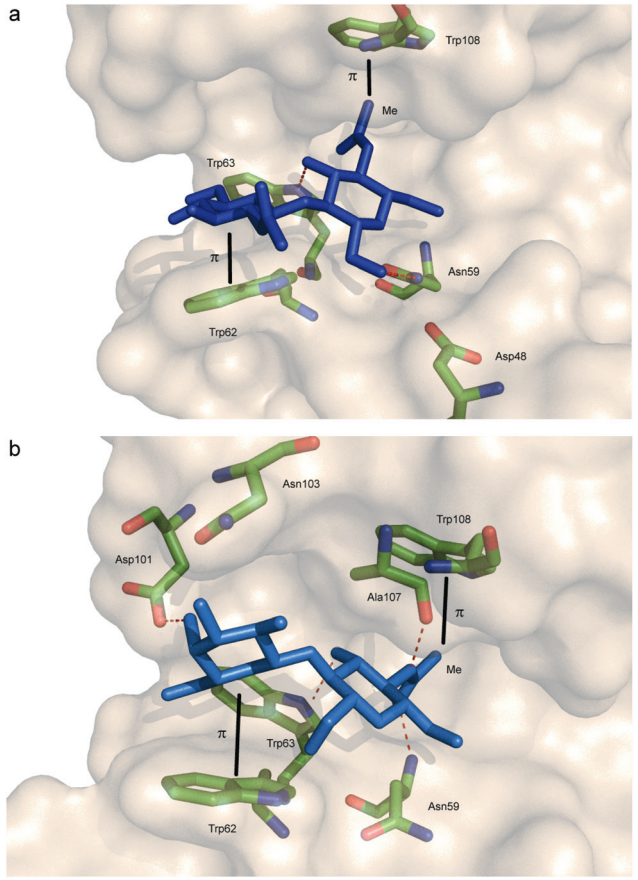


Fig. 10 Intermolecular interactions at sites B and C for **12** (a) and **11** (b) docked into complex with HEWL.

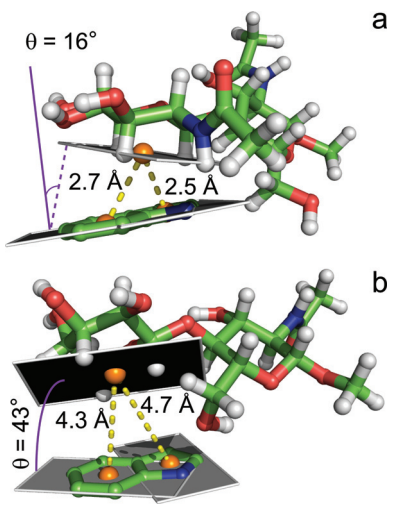


Fig. 11 The intermolecular stacking interactions of the indole part of Trp62 in HEWL with the terminal sugars in **12** (a) and **11** (b). The distances between the centroids are given in Å and the angle between the planes, given in degrees, is denoted by θ.

HEWL structures, a previously published NMR structure ensemble of HEWL (pdbid: 1E8L)[18] was used. The NMR result

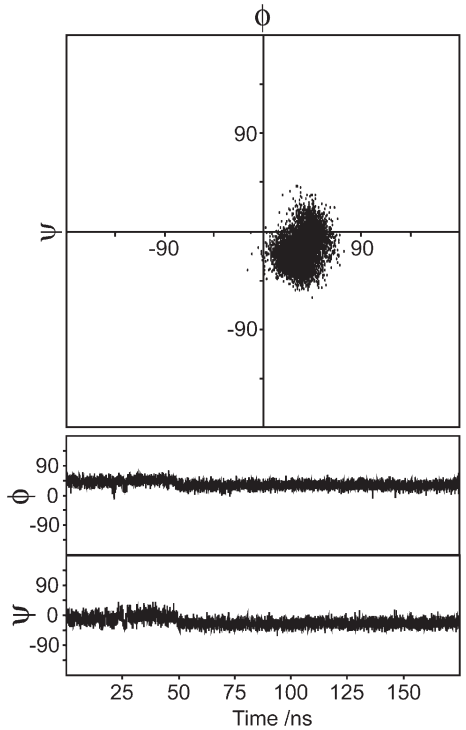


Fig. 12 Scatter plot (top) and time dependence (bottom) of the glycosidic torsion angles of disaccharide **11** in complex with HEWL from the 175 ns MD simulation.

consists of 50 structures that fulfil the NMR data and the RMSD of this data can be used as a model of the dynamics of the protein. The backbone RMSDs, referenced for the MD to the average trajectory structure during 50–175 ns of the simulations, for every residue of the HEWL-**11** complex (black) and the *apo*-form of HEWL (brick red) from simulations as well as the NMR structure (green) have been plotted in the lower panel of Fig. 13. Experimental data from NMR is in good agreement with MD simulations, which supports the reliability of the simulations. Conspicuously higher RMSDs are observed for Arg61 and Arg68 in the HEWL-**11** complex. The three structural models are presented in Fig. 14.

It may be noted that the higher flexibility in the loop region, *e.g.*, amino acids 68–70 (Fig. 14), observed in the MD simulations as well as from the RMSD-analysis of the 50 NMR structures (Fig. 13) determined by Schwalbe *et al.*[18] agrees well with the ^{15}N NMR relaxation data previously reported by Buck *et al.*,[17] which show main-chain values of $S^2 \approx 0.7$ whereas most residues have $S^2 \geq 0.8$ and some helices have $S^2 \approx 0.9$. A compilation of *B*-factors in crystal structures of the *apo*-form of the protein showed that they were anti-correlated to the generalized order parameters, *e.g.*, larger for residues in the region around Pro70. The accessible surface area was likewise larger in this region. However, for many of the amino acid residues involved in direct or close contact with the ligands in the present study (Fig. 6 and 10) the S^2 values from the NMR relaxation study are high in the *apo*-form of HEWL with one exception, *viz.*, Asn103 which has $S^2 \approx 0.5$, attributed to its high surface accessibility.[17]

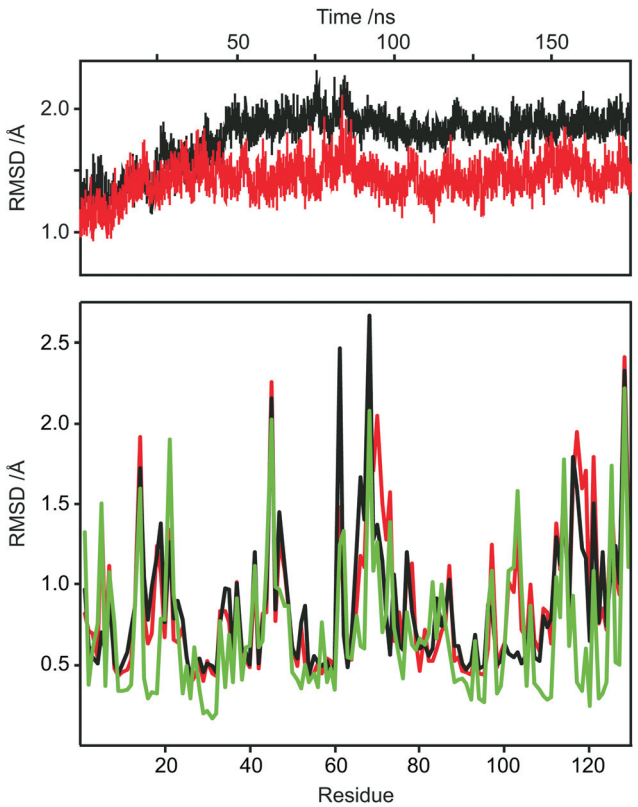


Fig. 13 (top) Time dependence of the RMSD of the protein backbone of HEWL in complex with **11** (black) and in the *apo*-form (brick red) during the 175 ns MD simulation, referenced to crystal structure 1SF4; (bottom) RMSD per amino acid residue for the protein backbone of HEWL in complex with **11** (black) and in the *apo*-form (brick red) during the production run between 50 and 175 ns of the MD simulation, referenced to the trajectory average structure as well as 50 NMR structures[18] (green), referenced to its average structure.

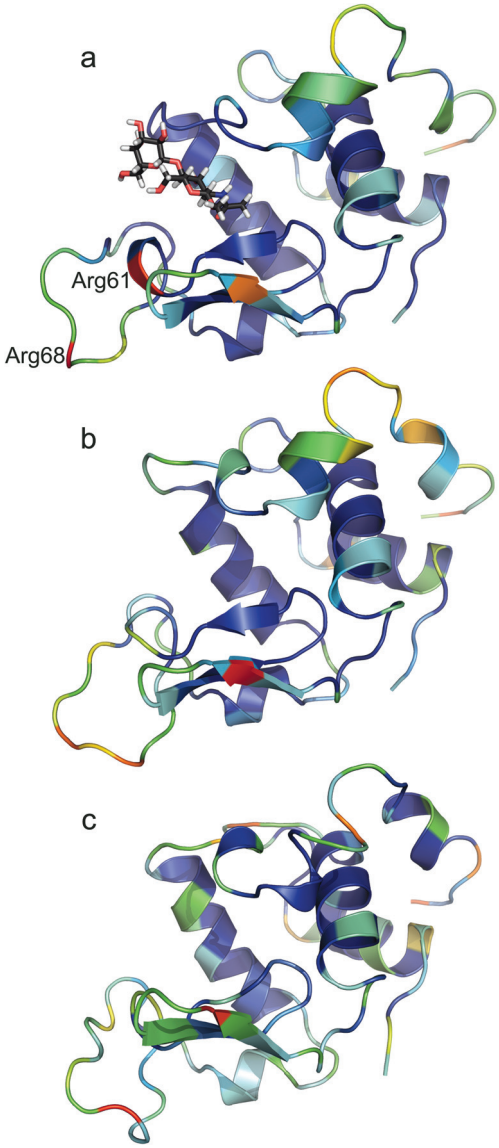


Fig. 14 HEWL colored by RMSD relative to its pertinent average structure, where blue represents low RMSD and red represents high RMSD; (a) HEWL in complex with **11** between 50 and 175 ns of the MD simulation, (b) HEWL between 50 and 175 ns of the MD simulation and (c) HEWL from the 50 experimental NMR structures (pdbid: 1E8L).

Conclusions

We have described a combination of techniques to study the interaction between HEWL and carbohydrate ligands that give both qualitative and quantitative information and the same procedures can be applied to many types of protein–ligand systems. The high through-put qualities of WAC have been demonstrated and it complemented structural information obtained from NMR spectroscopy, docking and MD simulations. The methodology does not require isotopically enriched samples nor is a large molecular mass of the protein a limitation. The dissociation constants differ for ligands that bind by two orders of magnitude with $K_D \approx 30$ mM for monosaccharides related to GlcNAc whereas $K_D \approx 0.3$ mM for GlcNAc-containing disaccharides. Two new solution structures of protein–ligand complexes have been proposed and this shows that results obtained from a crystal structure, disaccharide binding at C and D sites discussed herein, are not always present for solution complexes where the corresponding disaccharide binding occurs at sites B and C. The occurrence of binding in this region is fully consistent with the results of Dobson and co-workers[30] who analyzed ^{1}H NMR chemical shifts changes in HEWL upon addition *N*,*N*′-diacetylchitobiose (Fig. 15). The proposed structure poses did not show the highest ranking from the docking simulation. This may be due to conformational changes of the protein upon complex formation, which were not taken into consideration in the docking procedure and show that a combination of docking and MD simulations may be needed in order to obtain good agreement with experimental data. Furthermore, the results of this study underscore the importance of targeting the appropriate binding site(s), *e.g.*, when small molecule fragment-based libraries are employed[62] in the process of generating ligands with high affinity for protein receptors.

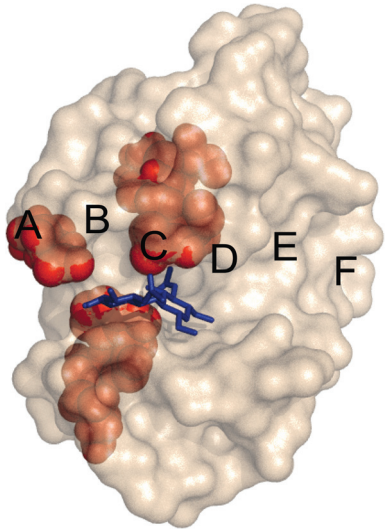


Fig. 15 HEWL with amino acid residues colored in red based on chemical shift changes upon formation of an *N*,*N*′-diacetylchitobiose complex in a study by Dobson and co-workers.[30] $\Delta\delta_H$ limit for coloring: 0.1 ppm. Due to signal overlap the authors were not able to detect chemical shift changes of a number of proton resonances in the B–C sites. The docked structure of compound **12** is shown in dark blue colour. The HEWL structure is from ref. 23.

Material and methods

Synthesis

Organic synthesis procedures and characterization of compounds are presented in the ESI.†

Preparation of HEWL columns

HEWL silica was prepared by coupling the protein to aldehyde functionalized silica, essentially as been described previously.[3] In short, spherical and porous silica Kromasil® silica (EKA chemicals AB, Bohus, Sweden) with a diameter of 10 μm and pore size of 100 Å was derivatized with (3-glycidyloxypropyl)-trimethoxysilane to obtain diol silica which was converted to aldehyde silica by periodate oxidation. Aldehyde silica (6.81 g) was mixed with HEWL (1.35 g in 70 mL of 0.1 M sodium phosphate, pH 7.0) and sodium cyanoborohydride (0.7 g). The coupling was allowed to proceed at 4 °C for 48 h. The coupling efficiency was determined by measuring the absorbance at 280 nm in the supernatant and 133 mg HEWL was found to be immobilized per gram of silica. The HEWL silica was packed into a 50 × 0.46 cm stainless steel column (Hichrom, Berkshire, UK) and a 3 × 0.21 cm PEEK column (Poros, Applied Biosystems, Carlsbad, CA, USA), using an air-driven liquid pump (Haskel, Burbank, CA, USA) and an ordinary HPLC pump (Varian, Walnut Creek, CA, USA), respectively.

Frontal chromatography

Frontal chromatography[3,63] was performed in order to determine the number of binding sites on the HEWL-immobilized silica. A Varian HPLC system (Walnut Creek, CA, USA) equipped with a 5012 pump, a Rheodyne manual injector with a 1 mL injection loop and a 9050 single UV wavelength detector, were used in the experiment. The mobile phase was 0.1 M sodium phosphate buffer pH 5.5 and the temperature was 23 ± 1 °C. Two different compounds were used in separate sets of experiments: newly dissolved GlcNAc (**1**, α-anomeric form) and LacNAc (**10**, anomeric mixture). Each compound was dissolved in 0.1 M sodium phosphate buffer pH 5.5 at several concentrations and injected until the column was saturated; the break-through curve (front) was detected at 220 nm. The concentrations of **1** were varied from 4 to 320 mM and concentrations of **10** were varied from 1.5 to 5.0 mM. A negative spike from the injection of pure water was used to determine the void volume of the system. Chromatography data were evaluated with EZChrom software version 6.8 (Scientific Software, San Ramon, CA, USA) and the midpoint of each front was determined from the first derivative of the curve (peak apex). The number of binding sites on the column and the dissociation constant (K_D) of the analyzed compound were obtained from the experimental data by non-linear regression analysis using GraphPad Prism 5 (SanDiego, CA, USA) applying a one-site binding model as been described previously;[3] see the ESI† for additional information.

Weak affinity chromatography (zonal chromatography)

Zonal chromatography experiments were performed on an Agilent 1100 series HPLC system controlled by ChemStation chromatography data system (Agilent Technologies, Santa Clara, CA, USA). The mobile phase was 0.1 M sodium phosphate buffer pH 5.5, and the temperature was 23 ± 1 °C. The flow rate was 0.1 mL min^{-1} when using the 3 × 0.21 cm column and 0.5 mL min^{-1} when using the 50 × 0.46 cm column. Compounds (see Table 1) were injected individually and the retention was recorded. An average from three separate injections was used in the calculations. The 3 × 0.21 cm HEWL column was used for compounds with moderate affinity (K_D values in the range of 0.1–5 mM) and the 50 × 0.46 cm column was used for compounds with very low affinities (K_D values above 5 mM). Sample concentration was typically between 25–100 μg mL^{-1} (corresponding to a concentration of about 0.1 mM, depending on the molecular weight of each compound).

NMR spectroscopy

For the proton R_2 relaxation rate measurements to obtain K_D a Bruker Advance 500 MHz spectrometer equipped with a 5 mm PFG triple-resonance CryoProbe was used. The experiments were carried out at 294 K in D_2O solutions with 20 μM TSP (sodium 3-trimethylsilyl-(2,2,3,3-$^2H_4$)-propanoate), 25 mM sodium phosphate buffer, pH 5.5, and 50 mM NaCl. Disaccharide **11** was titrated into a solution of 30 μM HEWL (kept constant) and R_2 measurements were performed at eight concentrations (0.2–13.4 mM) with the CPMG spin-echo pulse sequence. The methyl resonance at 2.04 ppm was observed using 12 delay times (10 ms–10 s). A decaying curve, based on the peak heights, was generated for each concentration. These curves were fitted to an exponentially decaying function to give relaxation rate R_2 values that for each concentration were plotted

and the K_D value was extracted according to eqn (1)[39] (plots are displayed in the ESI†):

$$[L]_0 = \frac{[P]_0 \Delta R_2^{max}}{\Delta R_2^{obs}} - K_D, \quad (1)$$

where $[L]_0$ and $[P]_0$ are the total concentrations of ligand and protein, respectively, ΔR_2^{obs} is the observed change in spin–spin relaxation rate constant for a given concentration relative to that of the free ligand, and ΔR_2^{max} is the limiting case for bound and free states. For processing of the K_D data Excel and Matlab programs were used. The FID from each experiment was acquired with 32–96 scans and the spectral width used was 8 kHz which together with the 16 k data points corresponds to an acquisition time of 1 s. The waiting period between scans was set to 16 s.

NMR assignments of disaccharides **11** and **12**, STD and trNOE NMR experiments were carried out on a Bruker Avance III 700 MHz NMR spectrometer with a 5 mm TCI Z-gradient high resolution CryoProbe. The STD and trNOE NMR experiments were carried out using Norell S-3-HT-7 3 mm tubes at 278 K with an oligosaccharide concentration of 1 mM and a protein concentration of 20 μM. Reference samples with free ligand (26 mM, 5 mm tube) were analyzed with STD, 1D and 2D NOESY NMR experiments. The buffer was 25 mM sodium phosphate, 50 mM sodium chloride pH 5.5. The buffer was lyophilized from H_2O and subsequently exchanged with D_2O.

The STD experiments were recorded using the standard pulse sequence[40] with a 60 ms 5 kHz spin-lock pulse to reduce the background protein resonances, excitation sculpting[64] with a 224 Hz square π-pulse to suppress residual HDO and purge pulses to remove unwanted magnetization. Saturation of the protein NMR signals of the enzyme was performed using a train of selective 65 Hz Gaussian pulses with duration of 50 ms, adding up to total saturation times of 0.5, 1.0, 2.0 and 4.0 s. The on-resonance frequency was set to 0.5 and 7.0 ppm and off-resonance irradiation was applied at 60 ppm. STD NMR spectra were acquired with a total of 2048, 1536, 1024 and 1024 transients for the four different saturation times respectively, in addition to 32 dummy scans. Spectra were recorded with a spectral width of 11 kHz and 18 k data points corresponding to an acquisition time of 0.8 s followed by a relaxation delay of 4.2 s.

For the trNOE experiments standard 1D and phase-sensitive 2D ^{1}H,^{1}H-NOESY pulse sequences were used together with excitation sculpting, with a 224 Hz square π-pulse, to suppress residual HDO and purge pulses to remove unwanted magnetization. In order to determine suitable experimental conditions 1D NOESY experiments were performed on the free ligand **12** at different temperatures and it was found that the NOE was close to zero at 289 K (at 16.4 T). The 1D experiments were performed with 32 k points with a sweep width of 11 kHz giving an acquisition time of 2.1 s. Selective excitation of the H1′ resonance was enabled using a 30 ms long r-SNOB shaped pulse.[65] A mixing time of 300 ms was used with a relaxation delay of 2.9 s, giving a recycle time between scans of 5.3 s. Ten thousand scans were acquired and 64 dummy scans were used to reach a steady state. The 2D experiments were performed with 4096 points in the direct dimension, 320 increments, with a sweep width of 11 kHz in both dimensions, and a relaxation delay of 1.1 s. Eighty scans were acquired per increment and 32 dummy scans were used.

Disaccharides **11** and **12** were dissolved in D_2O to a concentration of 15 and 26 mM respectively, and assigned using standard ^{1}H,^{1}H-homonuclear and ^{1}H,^{13}C-heteronuclear 2D experiments at 343 K and 278 K. Proton chemical shifts and coupling constants were refined using the NMR spin simulation[66] software PERCH (PERCH Solutions Ltd, Kuopio, Finland). Assignments of the NMR resonances of the protected compounds utilized a limited number of 1D and 2D NMR experiments but were complemented with chemical shift predictions made by PERCH. The ^{1}H and ^{13}C chemical shift assignments and J_{HH} are presented in the ESI.† The data from the NMR experiments were processed and analyzed with Bruker TOPSPIN 2.1 software.

Computer simulations

Molecular dynamics (MD) simulations used NAMD[67,68] (parallel version, 2.6) employing a CHARMM22 force field with CMAP[60] for the protein. The carbohydrates were described by the PARM22/SU01 force field.[61] Initial coordinates for the protein were taken from the crystal structure 1SF4[23] and topology files, TIP3P solvent boxes and ions were generated in VMD.[69] Pdb and psf files of the oligosaccharides were created using VEGA ZZ[70] and assigned CHARMM partial charges. The MD simulations were carried out with multiple-time-stepping of 2 fs, 2 fs, and 6 fs were used as the inner, middle and outer time steps in the NPT ensemble (P = 1 atm, 294 K) with a cutoff distance for non-bonded interactions set at 12 Å and periodic boundary conditions giving box sizes of approximately 50 Å × 44 Å × 51 Å. The smooth particle-mesh Ewald (SPME) method was used to calculate the full electrostatic interactions. The temperature and pressure were kept constant using a Langevin thermostat and a Langevin barostat, respectively. All bonds to hydrogen atoms were kept rigid. After energy minimization, heating and 1 ns of equilibration, production runs were carried out during 175 ns for HEWL and for the complex between disaccharide **11** and HEWL. For the crystal structure and docked complex of disaccharide **12** and HEWL, simulations were carried out for 50 ns. Data were saved every 1000 time steps for analysis. Trajectory analyses were carried out with VEGA ZZ and VMD.

For molecular docking simulations the crystal structure of the complex between HEWL and *N*,*N*′-diacetylchitobiose (pdbid: 1SF4),[23] with the disaccharide removed, was used as the receptor with AutoDock 4.0.[47] Coordinates for oligosaccharides **11** and **12** were obtained by the Glycam Biomolecule Builder (www.glycam.com). The ϕ-torsions were restrained in an *exo*-anomeric conformation at ϕ_H = 50° and all other torsions were flexible. The grid dimensions were 50 × 50 × 70 points, with points separated by 0.375 Å. The grids were chosen to be centered on binding site C and sufficiently large to cover the whole pocket. In total, 150 runs with the Lamarckian genetic algorithm were performed and the maximum number of free energy evaluations was set to 5 × 10^6. Other parameters were set to default values. In the complexes obtained by docking, the centroids used for CH/π-stacking analyses were based on the H1′/H3′/H5′ protons

in the terminal sugar residues and 4/6/7a and 2/3/7a atoms in the indole part of Trp62. The pertinent planes were based the H1′/H3′/H5′ protons in the terminal sugar residues and the 2/5/6 atoms in the indole part of Trp62.

Theoretical STD build-up curves were calculated from the docking results and the crystal structure 1SF4 using a modified version of CORCEMA-ST.[12,71] K_D values were taken from the WAC results. We have assumed the diffusion limited on-rate as 10^8 s^{-1} M^{-1}, a generalized order parameter S^2 of 0.85 and a uniform leakage relaxation of 0.25 s^{-1}. The ligand correlation time (τ_L) and the methyl group internal correlation time (τ_m) were chosen as 0.12 ns and 5 ps, respectively, since this resulted in the best agreement with experimental T_1 values. The protein correlation time (τ_P) was calculated to 12 ns using Stokes' law.[72] The conformation of the ligand was assumed to be the same in both the free and the bound state. The SHIFTX software[73] was used to calculate the ^{1}H chemical shifts of the protein and the protons resonating between 0.3 and 0.7 ppm were assumed to be saturated when irradiation was set at 0.5 ppm. All aromatic protons were assumed to be saturated when irradiation was set at 7 ppm.

Acknowledgements

This work was supported by grants from the Swedish Research Council, The Knut and Alice Wallenberg Foundation, and Magn. Bergvalls Stiftelse. Computing resources were kindly provided by the Center for Parallel Computers (PDC), Stockholm, Sweden. We are also grateful for the contribution from Emma Fång regarding the preparation of the HEWL WAC columns.

References

1 D. Zopf and S. Ohlson, Weak-affinity chromatography, *Nature*, 1990, **346**, 87–88.
2 S. Ohlson, M. Bergström, P. Påhlsson and A. Lundblad, Use of monoclonal antibodies for weak affinity chromatography, *J. Chromatogr., A*, 1997, **758**, 199–208.
3 M. Bergström, S. Liu, K. L. Kiick and S. Ohlson, Cholera toxin inhibitors studied with high-performance liquid affinity chromatography: a robust method to evaluate receptor–ligand interactions, *Chem. Biol. Drug Des.*, 2009, **73**, 132–141.
4 P. Cheshev, L. Morelli, M. Marchesi, Č. Podlipnik, M. Bergström and A. Bernardi, Synthesis and affinity evaluation of a small library of bidentate cholera toxin lignads: towards nonhydrolyzable ganglioside mimics, *Chem.–Eur. J.*, 2010, **16**, 1951–1967.
5 S. Ohlson and D. Zopf, Weak affinity chromatography, in *Handbook of Affinity Chromatography*, ed. T. Kline, Marcel Dekker, New York, NY, 1993, pp. 299–314.
6 M. Pellecchia, I. Bertini, D. Cowburn, C. Dalvit, E. Giralt, W. Jahnke, T. L. James, S. W. Homans, H. Kessler, C. Luchinat, B. Meyer, H. Oschkinat, J. Peng, H. Schwalbe and G. Siegal, Perspectives on NMR in drug discovery: a technique comes of age, *Nat. Rev. Drug Discovery*, 2008, **7**, 738–745.
7 D. Neumann and O. Kohlbacher, Structural glycomics – Molcular details of protein–carbohydrate interactions and their prediction. In *Proceedings of the International Beilstein Symposium in Glyco-Bioinformatics*, ed. M. G. Hicks and C. Kettner, Beilstein-Institut, Frankfurt/Main, 2009, pp. 101–122.
8 I. Halperin, B. Ma, H. Wolfson and R. Nussinov, Principles of docking: an overview of search algorithms and a guide to scoring functions, *Proteins*, 2002, **47**, 409–443.
9 S. A. Adcock and J. A. McCammon, Molecular dynamics: survey of methods for simulating the activity of proteins, *Chem. Rev.*, 2006, **106**, 1589–1615.
10 H. Alonso, A. A. Bliznyuk and J. E. Gready, Combining docking and molecular dynamics simulations in drug design, *Med. Res. Rev.*, 2006, **26**, 531–568.
11 B. A. Borgias, M. Gochin, D. J. Kerwood and T. L. James, Relaxation matrix analysis of 2D NMR data, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1990, **22**, 83–100.
12 V. Jayalakshmi and N. Rama Krishna, Complete relaxation and conformational exchange matrix (CORCEMA) analysis of intermolecular saturation transfer effects in reversibly forming ligand-receptor complexes, *J. Magn. Reson.*, 2002, **155**, 106–118.
13 C. C. Blake, D. F. Koenig, G. A. Mair, A. C. North, D. C. Phillips and V. R. Sarma, Structure of hen egg-white lysozyme: a three-timensional Fourier synthesis at 2 Å resolution, *Nature*, 1965, **206**, 757–761.
14 L. N. Johnson and D. C. Phillips, Structure of some crystalline lysozyme-inhibitor complexes determined by X-ray analysis at 6 Å resolution, *Nature*, 1965, **206**, 761–763.
15 K. Kurachi, L. C. Sieker and L. H. Jensen, Structures of triclinic mono- and di-*N*-acetylglucosamine: lysozyme complexes – A crystallographic study, *J. Mol. Biol.*, 1976, **101**, 11–24.
16 C. B. Post, B. R. Brooks, M. Karplus, C. M. Dobson, P. J. Artymiuk, J. C. Cheetham and D. C. Phillips, Molecular dynamics simulations of native and substrate-bound lysozyme. A study of the average structures and atomic fluctuations, *J. Mol. Biol.*, 1986, **190**, 455–479.
17 M. Buck, J. Boyd, C. Redfield, D. A. MacKenzie, D. J. Jeenes, D. B. Archer and C. M. Dobson, Structural determinants of protein dynamics: analysis of ^{15}N NMR relaxation measurements for main-chain and side-chain nuclei of hen eggwhite lysozyme, *Biochemistry*, 1995, **34**, 4041–4055.
18 H. Schwalbe, S. B. Grimshaw, A. Spencer, M. Buck, J. Boyd, C. M. Dobson, C. Redfield and L. J. Smith, A refined solution structure of hen lysozyme determined using residual dipolar coupling data, *Protein Sci.*, 2001, **10**, 677–688.
19 D. M. Chipman and N. Sharon, Mechanism of lysozyme action, *Science*, 1969, **165**, 454–465.
20 D. J. Vocadlo, G. J. Davies, R. Laine and S. G. Withers, Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate, *Nature*, 2001, **412**, 835–838.
21 L. R. Berger and R. S. Weiser, The β-glucosaminidase activity of egg-white lysozyme, *Biochim. Biophys. Acta*, 1957, **26**, 517–521.
22 S. J. Perkins, L. N. Johnson, D. C. Phillips and R. A. Dwek, The binding of monosaccharide inhibitors to hen egg-white lysozyme by proton magnetic resonance at 270 MHz and analysis by ring-current calculations, *Biochem. J.*, 1981, **193**, 553–572.
23 R. B. Von Dreele, Binding of *N*-acetylglucosamine oligosaccharides to hen egg-white lysozyme: a powder diffraction study, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2005, **61**, 22–32.
24 M. Noguchi, T. Tanaka, H. Gyakushi, A. Kobayashi and S.-I. Shoda, Efficient synthesis of sugar oxazolines from unprotected *N*-acetyl-2-amino sugars by using chloroformamidinium reagent in water, *J. Org. Chem.*, 2009, **74**, 2210–2212.
25 S. Knapp, D. Vocadlo, Z. N. Gao, B. Kirk, J. P. Lou and S. G. Withers, NAG-thiazoline, an *N*-acetyl-β-hexosaminidase inhibitor that implicates acetamido participation, *J. Am. Chem. Soc.*, 1996, **118**, 6804–6805.
26 P. Bhattacharyya and J. D. Woollins, Selenocarbonyl synthesis using Woollins reagent, *Tetrahedron Lett.*, 2001, **42**, 5949–5951.
27 T. Ozturk, E. Ertas and O. Mert, Use of Lawesson's reagent in organic syntheses, *Chem. Rev.*, 2007, **107**, 5210–5278.
28 G. Livingstone, F. Franks and L. J. Aspinall, The effects of aqueous solvent structure on the mutarotation kinetics of glucose, *J. Solution Chem.*, 1977, **6**, 203–216.
29 P. W. Wertz, J. C. Garver and L. Anderson, Anatomy of complex mutarotation. Kinetics of tautomerization of α-D-galactopyranose and β-D-galactopyranose in water, *J. Am. Chem. Soc.*, 1981, **103**, 3916–3922.
30 K. J. Lumb, J. C. Cheetham and C. M. Dobson, ^{1}H nuclear magnetic resonance studies of hen lysozyme-*N*-acetylglucosamine oligosaccharide complexes in solution. Application of chemical shifts for the comparison of conformational changes in solution and in the crystal, *J. Mol. Biol.*, 1994, **235**, 1072–1087.
31 C. C. F. Blake, L. N. Johnson, G. A. Mair, A. C. T. North, D. C. Phillips and V. R. Sarma, Crystallographic studies of the activity of hen egg-white lysozyme, *Proc. R. Soc. London, Ser. B*, 1967, **167**, 378–388.
32 R. B. Von Dreele, Binding of *N*-acetylglucosamine to chicken egg lysozyme: a powder diffraction study, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2001, **57**, 1836–1842.

33 S. K. Banerjee and J. A. Rupley, Temperature and pH dependence of the binding of oligosaccharides to lysozyme, *J. Biol. Chem.*, 1973, **248**, 2117–2124.

34 F. W. Dahlquist and M. A. Raftery, A nuclear magnetic resonance study of association equilibria and enzyme-bound environments of *N*-acetyl-D-glucosamine anomers and lysozyme, *Biochemistry*, 1968, **7**, 3269–3277.

35 B. D. Sykes and C. Parravano, A nuclear magnetic resonance study of the inhibition of lysozyme by *N*-acetyl-D-glucosamine and di-*N*-acetyl-D-glucosamine, *J. Biol. Chem.*, 1969, **244**, 3900–3904.

36 C. Bjurulf and I. Wadsö, Thermochemistry of lysozyme-inhibitor binding, *Eur. J. Biochem.*, 1972, **31**, 95–102.

37 Y. He, M. S. Macauley, K. A. Stubbs, D. J. Vocadlo and G. J. Davies, Visualizing the reaction coordinate of an O-GlcNAc hydrolase, *J. Am. Chem. Soc.*, 2010, **132**, 1807–1809.

38 A. L. Bowman, I. M. Grant and A. J. Mulholland, QM/MM simulations predict a covalent intermediate in the hen egg white lysozyme reaction with its natural substrate, *Chem. Commun.*, 2008, 4425–4427.

39 L. Fielding, NMR methods for the determination of protein–ligand dissociation constants, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2007, **51**, 219–242.

40 M. Mayer and B. Meyer, Characterization of ligand binding by saturation transfer difference NMR spectroscopy, *Angew. Chem., Int. Ed.*, 1999, **38**, 1784–1788.

41 M. Muraki and K. Harata, X-Ray structural analysis of the ligand-recognition mechanism in the dual-affinity labeling of c-type lysozyme with 2′,3′-epoxypropyl β-glycoside of *N*-acetyllactosamine, *J. Mol. Recognit.*, 2003, **16**, 72–82.

42 G. Widmalm, General NMR spectroscopy of carbohydrates and conformational analysis in solution, in *Comprehensive Glycoscience*, ed. J. P. Kamerling, Elsevier, Oxford, 1st edn, 2007, pp. 101–132.

43 M. Aida, Y. Sugawara, S. Oikawa and K. Umemoto, Structural fluctuation of methyl *N*,*N*′-diacetyl-β-D-chitobioside in vacuo and in aqueous solution: molecular dynamics simulations and proton NMR spectroscopy, *Int. J. Biol. Macromol.*, 1995, **17**, 227–235.

44 G. M. Clore and A. M. Gronenborn, Theory and applications of the transferred nuclear Overhauser effect to the study of the conformations of small ligands bound to proteins, *J. Magn. Reson.*, 1982, **48**, 402–417.

45 C. B. Post, Exchange-transferred NOE spectroscopy and bound ligand structure determination, *Curr. Opin. Struct. Biol.*, 2003, **13**, 581–588.

46 T. Scherf and J. Anglister, A $T_{1\rho}$-filtered two-dimensional transferred NOE spectrum for studying antibody interactions with peptide antigens, *Biophys. J.*, 1993, **64**, 754–761.

47 G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *J. Comput. Chem.*, 1998, **19**, 1639–1662.

48 J. Landström, E.-L. Nordmark, R. Eklund, A. Weintraub, R. Seckler and G. Widmalm, Interaction of a Salmonella enteritidis O-antigen octasaccharide with the phage P22 tailspike protein by NMR spectroscopy and docking studies, *Glycoconjugate J.*, 2008, **25**, 137–143.

49 V. Spiwok, P. Lipovová, T. Skálová, E. Buchtelová, J. Hašek and B. Králova, Role of CH/π interactions in substrate binding by *Escherichia coli* β-galactosidase, *Carbohydr. Res.*, 2004, **339**, 2275–2280.

50 M. del Carmen Fernández-Alonso, F. J. Cañada, J. Jiménez-Barbero and G. Cuevas, Molecular recognition of saccharides by proteins. Insights on the origin of carbohydrate–aromatic interactions, *J. Am. Chem. Soc.*, 2005, **127**, 7379–7386.

51 V. Spiwok, P. Lipovová, T. Skálová, E. Vondráčková, J. Dohnálek, J. Hašek and B. Králova, Modelling of carbohydrate–aromatic interactions: *ab initio* energetics and force field performance, *J. Comput.-Aided Mol. Des.*, 2006, **19**, 887–901.

52 S. Kozmon, R. Matuška, V. Spiwok and J. Koča, Three-dimensional potential energy surface of selected carbohydrates' CH/π dispersion interactions calculated by high-level quantum mechanical methods, *Chem.–Eur. J.*, 2011, **17**, 5680–5690.

53 G. B. McGaughey, M. Gagné and A. K. Rappé, π-Stacking interactions, *J. Biol. Chem.*, 1998, **273**, 15458–15463.

54 C. Hamark, J. Landström, L. Eriksson and G. Widmalm, Ethyl 3,6-di-*O*-benzyl-2-deoxy-*N*-phthalimido-1-thio-β-D-glucopyranoside, *Acta Crystallogr., Sect. E: Struct. Rep. Online*, 2010, **66**, o3250–o3251.

55 C. Hamark, J. Landström, L. Eriksson and G. Widmalm, Ethyl 4,6-*O*-benzylidene-2-deoxy-*N*-phthalimido-1-thio-β-D-glucopyranoside, *Acta Crystallogr., Sect. E: Struct. Rep. Online*, 2010, **66**, o3249.

56 M. S. Sujatha, Y. U. Sasidhar and P. V. Balaji, Insights into the role of the aromatic residue in galactose-binding sites: MP2/6–311G++** study on galactose- and glucose-aromatic residue analogue complexes, *Biochemistry*, 2005, **44**, 8554–8562.

57 R. Sharma, J. P. McNamara, R. K. Raju, M. A. Vincent, I. H. Hillier and C. A. Morgado, The interaction of carbohydrates and amino acids with aromatic systems studied by density functional and semi-empirical molecular orbital calculations with dispersion corrections, *Phys. Chem. Chem. Phys.*, 2008, **10**, 2767–2774.

58 S. Tsuzuki and A. Fujii, Nature and physical origin of CH/π-interaction: significant difference from convetional hydrogen bonds, *Phys. Chem. Chem. Phys.*, 2008, **10**, 2584–2594.

59 S. Tsuzuki, T. Uchimaru and M. Mikami, Magnitude and nature of carbohydrate–aromatic interactions: *ab initio* calculations of fucose–benzene complex, *J. Phys. Chem. B*, 2009, **113**, 5617–5621.

60 A. D. Mackerell Jr., M. Feig and C. L. Brooks III, Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, *J. Comput. Chem.*, 2004, **25**, 1400–1415.

61 R. Eklund and G. Widmalm, Molecular dynamics simulations of an oligosaccharide using a force field modified for carbohydrates, *Carbohydr. Res.*, 2003, **338**, 393–398.

62 C. Rademacher, J. Landström, N. Sindhuwinata, M. M. Palcic, G. Widmalm and T. Peters, NMR-based exploration of the acceptor binding site of human blood Group B galactosyltransferase with molecular fragments, *Glycoconjugate J.*, 2010, **27**, 349–358.

63 K.-I. Kasai, Y. Oda, M. Nishikata and S.-I. Ishii, Frontal affinity-chromatography: theory for its application to studies on specific interactions of biomolecules, *J. Chromatogr., Biomed. Appl.*, 1986, **376**, 33–47.

64 T. L. Hwang and A. J. Shaka, Water suppression that works. Excitation sculpting using arbitrary waveforms and pulsed field gradients, *J. Magn. Reson., Ser. A*, 1995, **112**, 275–279.

65 Ē. Kupče, J. Boyd and I. D. Campbell, Short selective pulses for biochemical applications, *J. Magn. Reson., Ser. B*, 1995, **106**, 300–303.

66 R. Laatikainen, M. Niemitz, U. Weber, J. Sundelin, T. Hassinen and J. Vepsäläinen, General strategies for total-lineshape-type spectral analysis of NMR spectra using integral-transform iterator, *J. Magn. Reson., Ser. A*, 1996, **120**, 1–10.

67 L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan and K. Schulten, NAMD2: greater scalability for parallel molecular dynamics, *J. Comput. Phys.*, 1999, **151**, 283–312.

68 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, Scalable molecular dynamics with NAMD, *J. Comput. Chem.*, 2005, **26**, 1781–1802.

69 W. Humphrey, A. Dalke and K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graphics*, 1996, **14**, 33–38, 27–28.

70 A. Pedretti, L. Villa and G. Vistoli, VEGA: a versatile program to convert, handle and visualize molecular structure on Windows-based PCs, *J. Mol. Graphics Modell.*, 2002, **21**, 47–49.

71 S. Kemper, M. K. Patel, J. C. Errey, B. G. Davis, J. A. Jones and T. D. W. Claridge, Group epitope mapping considering relaxation of the ligand (GEM-CRL): including longitudinal relaxation rates in the analysis of saturation transfer difference (STD) experiments, *J. Magn. Reson.*, 2010, **203**, 1–10.

72 J. Cavanagh, W. J. Fairbrother, A. G. Palmer III and N. J. Skelton, Classical NMR spectroscopy, in *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, Boston, 1st edn, 1996, p. 17.

73 S. Neal, A. M. Nip, H. Zhang and D. S. Wishart, Rapid and accurate calculation of protein ^{1}H, ^{13}C and ^{15}N chemical shifts, *J. Biomol. NMR*, 2003, **26**, 215–240.

