Glycoconjugate Journal (1994) **11**:507–517

The monosaccharide binding site of lentil lectin: an X-ray and molecular modelling study

REMY LORIS[1]*‡, FLORENCE CASSET[2]*, JULIE BOUCKAERT[1], JURGEN PLETINCKX[1], MINH-HOA DAO-THI[1], FREDDY POORTMANS[3], ANNE IMBERTY[4], SERGE PEREZ[2] and LODE WYNS[1]

[1] *Laboratorium voor Ultrastructuur, Instituut voor Moleculaire Biologie, Vrije Universiteit Brussel, Paardenstraat 65, B-1640 Sint-Genesius-Rode, Belgium*

[2] *Ingénierie Moléculaire, INRA, BP 527, 44026 Nantes cedex 03, France*

[3] *Vlaamse Instelling voor Technologisch Onderzoek—VITO, Boeretang 200, B-2400 Mol, Belgium*

[4] *LSO-CNRS, Faculté des Sciences et Techniques, 2 rue de la Houssinière, 44072 Nantes cedex 03, France*

Received 21 March 1994, revised 17 June 1994

The X-ray crystal structure of lentil lectin in complex with α-D-glucopyranose has been determined by molecular replacement and refined to an *R*-value of 0.20 at 3.0 Å resolution. The glucose interacts with the protein in a manner similar to that found in the mannose complexes of concanavalin A, pea lectin and isolectin I from *Lathyrus ochrus*. The complex is stabilized by a network of hydrogen bonds involving the carbohydrate oxygens O6, O4, O3 and O5. In addition, the α-D-glucopyranose residue makes van der Waals contacts with the protein, involving the phenyl ring of Phe123β. The overall structure of lentil lectin, at this resolution, does not differ significantly from the highly refined structures of the uncomplexed lectin.

Molecular docking studies were performed with mannose and its 2-*O* and 3-*O*-*m*-nitro-benzyl derivatives to explain their high affinity binding. The interactions of the modelled mannose with lentil lectin agree well with those observed experimentally for the protein-carbohydrate complex. The highly flexible Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside and Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside become conformationally restricted upon binding to lentil lectin. For best orientations of the two substrates in the combining site, the loss of entropy is accompanied by the formation of a strong hydrogen bond between the nitro group and one amino acid, Gly97β and Asn125β, respectively, along with the establishment of van der Waals interactions between the benzyl group and the aromatic amino acids Tyr100β and Trp128β.

Keywords: lectin; lentil lectin; carbohydrate specificity; glucose binding; mannose binding

Introduction

The legume lectins are a large family of homologous proteins with unknown physiological function that specifically recognize complex carbohydrate structures. These proteins are abundant in the seeds of most legume plants and relatively easy to purify using affinity chromatography. Therefore, they are considered as model systems for studying the molecular basis of protein–carbohydrate recognition. Their diverse carbohydrate specificities have made them important research tools in glycobiology, immunology, parasitology and medical research [1].

The three-dimensional structures of several legume lectins have been determined by X-ray crystallography [2–5]. Not unexpectedly, their monomeric fold was found to be very similar but the way they associate into dimers or tetramers has been found to be quite diverse and controlled by the presence of covalently bound carbohydrate [6] and point mutations [7]. The recent crystal structures of several legume lectin–monosaccharide complexes [8, 9] and some lectin–oligosaccharide complexes [6, 7, 10–12] have shed some light on the molecular basis of their carbohydrate specificity.

The molecular basis of specific saccharide recognition by proteins has been a much debated topic for several decades [13]. Oligosaccharides are generally flexible entities in solution. It has been suggested that the loss of conformational entropy upon binding to the protein could, at least in part, be compensated for by a favourable increase in entropy due to the liberation of solvent molecules bound to both the saccharide and the protein [14]. Furthermore, polar and nonpolar contacts appear to be equally important

* RL and FC are joint first authors.

‡ To whom correspondence should be addressed.

0282-0080 © 1994 Chapman & Hall

for specificity and affinity. Stacking of aromatic residues against the hydrophobic parts of the sugar is a recurrent theme, as well as the formation of complex hydrogen bonding networks between the protein and saccharide [15].

The structure of lentil lectin was determined recently [5] and refined to high resolution (R. Loris, unpublished). Lentil lectin belongs to the group of Glc/Man-specific lectins, that also includes concanavalin A, LOL I, pea lectin and sainfoin lectin. Despite a similar affinity for mannose and glucose, these lectins do not share the same specificity for oligosaccharides and substituted monosaccharides. The *Viciae* tribe lectins, such as lentil lectin, have a higher affinity for mannoses and glucoses with hydrophobic substituents on the O2 and O3 atoms [16, 17]. In contrast, concanavalin A does not tolerate substitutions on the O3-oxygen of glucose and mannose [18]. These differences in fine specificity have still to be explained. In this study we probe the molecular basis of monosaccharide specificity of lentil lectin by X-ray crystallography and molecular modelling techniques.

Materials and methods

Crystallization and data collection of lentil lectin–glucose complex

Lentil lectin was purified from the seeds of the common lentil (*Lens culinaris*) as previously described [19]. Crystallization conditions were screened using the hanging drop method. Large single crystals (0.2 × 0.3 × 0.5 mm) were obtained when equilibrating 10 µl drops of a 10 mg ml^{-1} solution of lentil lectin in a 100 mM cacodylate buffer (pH 6.5) containing 50 mM glucose and 15% ethanol against 35% ethanol in the same buffer solution. The crystals appeared after 1 week and proved to be very fragile and difficult to handle. Most crystals did not produce measurable diffraction after mounting in glass capillaries and showed visual signs of degradation. Only two crystals could be mounted successfully and were analysed immediately on an Enraf-Nonius FAST area detector. Sharp diffraction spots were observed to a resolution of about 3.0 Å. Autoindexing using the MADNESS [20–21] software yielded a unit cell with dimensions $a = b = 85.71$ Å, c = 165.38 Å, $\alpha = \beta = 90.00°$ and $\gamma = 120.00°$, suggesting trigonal or hexagonal symmetry. Subsequent analysis of the measured data and of low resolution precession photographs taken from one of the crystals after data collection, confirmed the crystals to belong to space group $P6_122$ or $P6_522$. A total of 35751 measured intensities between 10.0 and 3.0 Å resolution were reduced to 6950 unique reflections with a merging *R*-factor of 0.070. The statistics of the data collection are summarized in Table 1.

Molecular replacement and structure refinement

From packing considerations, it was inferred that the crystals most likely contain a single lectin monomer in their asymmetric unit. Therefore, the dimer axis of the complete lectin molecule needs to coincide with a crystallographic two-fold axis, thus limiting the molecular replacement problem to a one-dimensional rotation search followed by the calculation of a single translation vector.

Table 1. Refinement of the lentil lectin–glucose complex.

Unit cell		a = b = 85.71 Å c = 165.38 Å γ = 120.0°
Space group		$P6_522$
Resolution		10.0–3.0 Å
Number of measured reflections		35751
Number of unique reflections		6950
Data completeness		94.5%
R_{symm} [a]		0.070
Crystallographic *R*-factor [b]		0.206
Mean positional error	(From Luzatti Plot)	0.35 Å
Average B-value	Main chain atoms	36.67 Å^2
	Side chain atoms	39.09 Å^2
	Solvent atoms	58.08 Å^2
	Glucose molecule	46.86 Å^2
	Total	38.45 Å^2
RMS deviation on	Bond lengths	0.018 Å
	Bond angles	3.2°
	Peptide plane planarity	6.0°

[a] $R_{symm} = \sum_i \sum_j |I_{ij} - \langle I_j \rangle| \Big/ \sum_{ij} I_{jj}$.

[b] $R = \sum_N (F_o - F_c)/\sum_N F_o$.

The starting model for molecular replacement was the refined structure of orthorhombic lentil lectin [5], from which the phosphate, manganese and calcium ions and all water molecules were removed. All molecular replacement calculations were done on a VAX 3000 server or an Iris Indigo. Rotation functions, initially calculated with MERLOT [22] and ALMN [23], using both a monomer or the complete dimer as search models, did not produce any solution consistent with the lectin dimer axis coinciding with a crystallographic two-fold. When using the ROTFUN function of AMORE [24], a clear solution meeting these criteria was only found when using the complete dimer as search model. When using a monomeric search model, the correct solution was still present in the map, but only in the 60th highest position. The lectin molecule was then oriented in the new unit cell with the dimer axis coinciding with the *a*-axis, following the results from the rotation search. This oriented molecule was then stepped through the cell along the *a*-axis in steps of 0.85 Å, and for each step 10 cycles of rigid body refinement were performed using the FITFUN option [25] of AMORE. For reasons of convenience, these calculations were performed using a lectin dimer in the lower symmetry space groups $P6_1$ and $P6_5$. A clear translation solution was found only in space group $P6_5$. The *R*-factor calculated using the lentil lectin monomer in the real space group $P6_522$ was 34.5% for all data between 8.0 and 3.2 Å resolution.

An electron density map calculated at this stage clearly showed the positions of the calcium and the manganese ions, which were not present in the model, and the presence of electron density corresponding to a glucose molecule in the supposed monosaccharide binding site, thus confirming the correctness of the molecular replacement solution. After introducing these metal ions together with a glucose molecule in its α-anomeric conformation, structural refinement was performed using the program X-PLOR [26] running on a CRAY-YMP supercomputer. After several alternating cycles of refinement and model building, the refinement converged at an *R*-value of 20.6% (27.1% without solvent) with no significant new features present in electron density maps calculated with coefficients F_o–F_c and $2F_o$–F_c.

Identification of possible water molecules was aided by superimposing the high resolution structures of lentil lectin and checking each water site for its presence in the present structure. Neither the present structure nor these high resolution structures have a dense crystal packing. Therefore, most of the water molecules may be considered to be lattice-independent. This led to the identification of 43 crystallographically independent water molecules. The stereochemistry of the final model was examined using PROCHECK [27].

Energy calculations

Modelling studies were performed with the SYBYL molecular modelling package (TRIPOS) of Evans & Sutherland, running on a Silicon Graphics Elan 4000 Indigo. The TRIPOS force field [28] was used, which takes into account the contribution of bond-stretching, angle-bending, torsional and van der Waals energies [29]. The contribution of hydrogen bonding is included in the electrostatic energy term, which can be taken into account or not. Appropriate energy parameters for carbohydrates have recently been developed [30] and are used throughout this study. All the geometry optimizations were performed using the energy minimizer MAXIMIN2 [31]. A procedure consisting of a combination of SIMPLEX and Conjugate Gradient methods was used. The gradient parameter was fixed at 0.5 and the iteration number at 150.

Selection and optimization of amino acids

For reasons of computer time, it was not possible to take into account the entire protein in the energy minimization calculations. Therefore, a region around the binding site was selected, consisting of a sphere of 15 Å radius centred on the P atom of the phosphate anion found in the binding site of the uncomplexed lectin. Thus, 69 amino acids and the Ca^{2+} ion were retained for the following calculations.

Hydrogen atoms were generated in several steps. The hydrogen atoms were first added on carbon and nitrogen atoms of the protein backbone and their positions were optimized with the TRIPOS force field. The second step dealt with the hydrogen atoms of the side chain carbons. Finally, hydrogens linked to oxygen and nitrogen atoms were added. The atomic charges were then calculated by the Pullman method [32] prior to the optimization of all hydrogen atoms. The Ca^{2+} and Mn^{2+} charges were taken to be $2e^-$; a dielectric constant of 4 was used throughout the entire study.

For the docking studies described in the following sections, a 'hot' region is defined which will be allowed to adapt. This region contains 19 amino acids (Ser39β, Ala80β, Asp81β, Gly97β, Gly98β, Gly99β, Tyr100β, Leu101β, Phe123β, Tyr124β, Asn125β, Ala126β, Ala127β, Trp128β, Thr27α, Thr28α, Gly29α, Ala30α, Glu31α), belonging to a sphere having a radius of 8 Å centred about the binding site.

Docking of the Me-α-D-mannopyranoside

The Me-α-D-mannopyranoside in the combining site of lentil was given the same orientation as the one found in the crystal structure of the complex between *Lathyrus ochrus* isolectin I and Me-α-D-mannopyranoside [9]. The carbohydrate moiety atom types and charges were defined as described previously [30]. Optimization of the complex was performed through several cycles of energy minimization. During the energy minimization, not only the ligand was allowed to optimize, but also all the atoms of the 19 amino acids of the 'hot' region and the hydrogen atoms of the whole fragment. The binding energy between the ligand (L) and the protein (P) E_{bind} has been evaluated as:

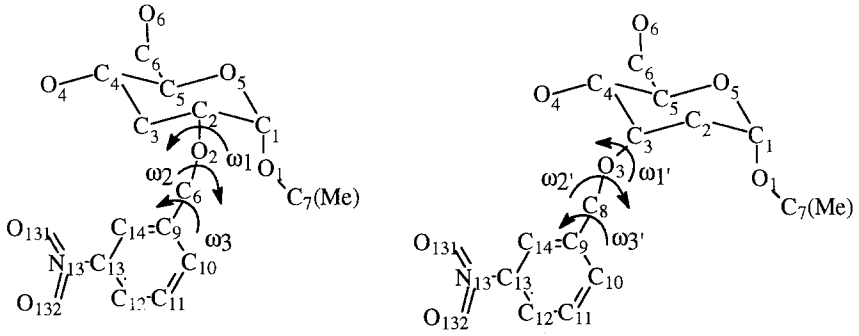
$$E_{bind} = E(P \cdot L) - E(P + L)$$

where $E(P \cdot L)$ represents the potential energy of the complex and $E(P + L)$ represents the potential energy of the protein and the ligand in the same coordinate space, but with no spatial interaction, i.e. apart one from the other [33]. The solvent was not taken into account in these calculations and the computed binding energies should therefore be considered as 'apparent binding energies'.

Conformational search for Me-2-O- or 3-O-(m-nitro-benzyl)-α-D-mannopyranoside

Molecules were built by adding an *m*-nitro-benzyl group on the methyl-α-D-mannopyranoside molecule, either in the isolated state or bound in the lectin. The atoms names and torsion angles of interest are shown in Fig. 1 together with a schematic representation of Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside and Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside. The three torsion angles of interest for the Me-2-*O*- and Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside are:

$\omega 1 = \theta$(C1-C2-O2-C8) $\quad \omega 1' = \theta$(C2-C3-O3-C8)
$\omega 2 = \theta$(C2-O2-C8-C9) $\quad \omega 2' = \theta$(C3-O3-C8-C9)
$\omega 3 = \theta$(O2-C8-C9-C10) $\quad \omega 3' = \theta$(O3-C8-C9-C10)



Me 2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside

Me 3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside

Figure 1. Labelling and torsion angles of Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside and Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside.

The search procedure in the SYBYL molecular modelling package allows a systematic conformational search [34] around several rotatable bonds. Conformational analysis of the *m*-nitro-benzyl group, either on the 2-*O*- position or on the 3-*O*- position of the mannose residue were performed by varying the three torsion angles described above, first for the isolated molecules, and subsequently for the molecules docked to the binding site. The systematic searches were performed with a 5° increment around the $\omega 1$, $\omega 2$ and $\omega 3$ torsions. In such an approach, the hydroxylic hydrogens are not considered and the energy calculation does not take into account the electrostatic contribution. Three-dimensional iso-contour energy potential maps were drawn to illustrate the search results.

The most favourable conformations and orientations of the molecules in the lectin combining site were considered for further optimization. At this stage, the hydroxylic hydrogen atoms were considered and the atomic charges were derived for the *m*-nitro-benzyl group with the MOPAC option using the MNDO Hamiltonian [35]. The energy minimization procedures were conducted as described for the complex with the Me-α-D-mannopyranoside residue.

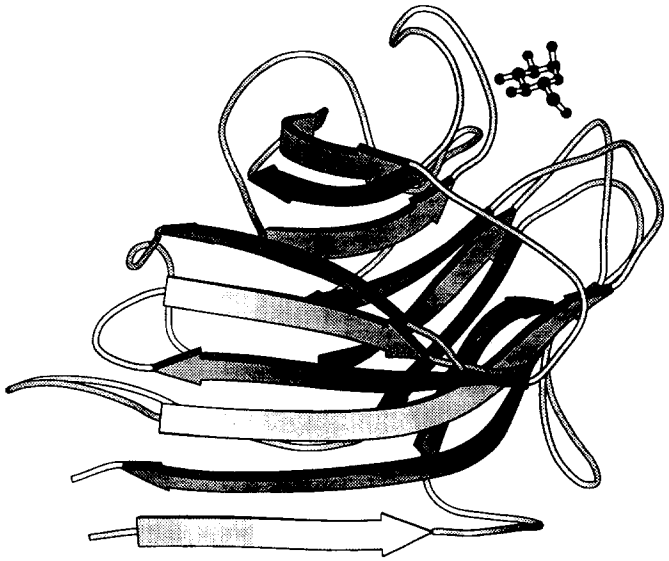


Figure 2. Global view of lentil lectin showing the location of the bound glucose. Figure drawn using MOLSCRIPT [39].

Results

Description of the lentil lectin–glucose complex and comparison with the uncomplexed protein

The refined lentil lectin–glucose complex contains 1777 protein atoms, one manganese and one calcium ion, one glucose molecule and 43 water molecules. The stereochemistry of the model is described in Table 1. A ribbon diagram of the lentil lectin monomer, showing also the position of the glucose binding site, is shown in Fig. 2.

The overall structure of the molecule does not differ significantly from the highly refined structures of the uncomplexed lectin. This is illustrated clearly by the small RMS differences of 0.29 Å and 0.54 Å for the superposition of all backbone and side chain atoms respectively of the refined glucose complex on the starting structure. This agrees well with the mean coordinate error of 0.35 Å calculated from a Luzatti plot. A plot of the main chain and side chain positional RMS differences is given in Fig. 3. The structure has a high mean *B*-value of 38.4 Å^2 and a wider variation of backbone *B*-values than the starting structure. The highest *B*-values are found in the loop around Tyr77β, while in the uncomplexed structures, they are found in the loop around Arg55β. These two loop regions also show the largest RMS differences in backbone coordinates. These differences are most likely due to packing effects and are unrelated to the binding of the glucose molecule.

The crystals have a fairly high solvent content (65%). This is reflected in the crystal packing, which is dominated by large solvent channels, 40 Å in diameter and spanning the whole length of the crystal, around the crystallographic six-fold axis. This high solvent content together with the

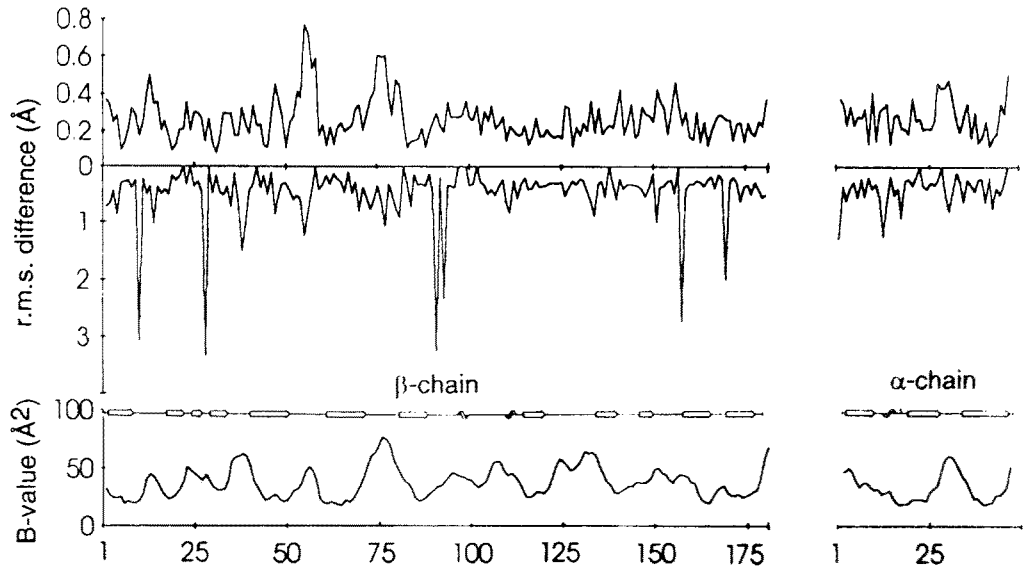


Figure 3. The per residue averaged RMS differences of the main chain (a) and side chain atom (b) positions between the lentil lectin–glucose complex and the uncomplexed lectin molecule that was used as the starting model for the refinement are shown. They are aligned with the per residue averaged main chain B-values of the lentil lectin–glucose complex (c) and the secondary structure assignments as based on the main chain hydrogen bonding pattern, determined using the program DSSP [40].

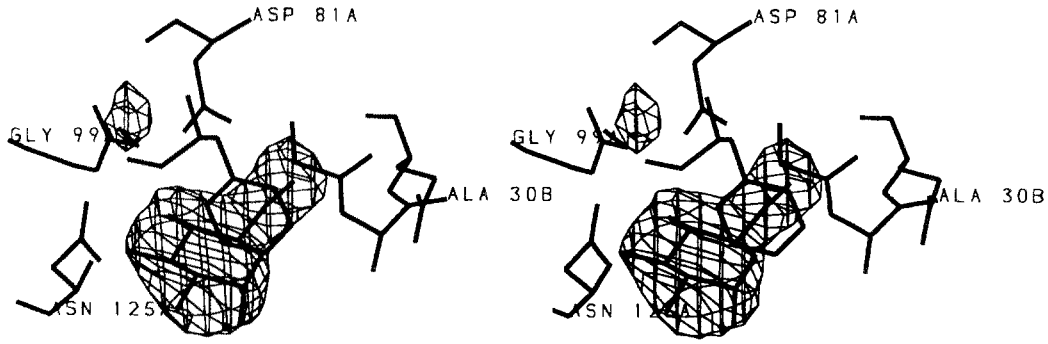


Figure 4. Final difference electron density of the region around the bound glucose. The map was calculated after taking the glucose out of the final model and subsequent 40 steps of energy-restrained X-ray refinement with X-PLOR. The final model is superimposed. The small spherical blob of electron density near Gly99β corresponds to a water molecule.

large thermal motions in the crystal are probably responsble for the relatively high R-value of the refined structure as compared to the refined structures of the uncomplexed lectin [5]. At the resolution that can be obtained with these crystals, it is not possible to determine an accurate solvent model. The electron densities of the protein molecule and the bound glucose, however, are satisfactory and could be interpreted easily.

Description of the glucose binding site

The electron density around the glucose molecule and the hydrogen bonding scheme between the glucose molecule and the protein are shown in Fig. 4 and Tables 2 and 3. The mode of binding observed here is similar to the one found in the crystal structure of LOL I complexed with glucose. The essential features of the hydrogen bond network that stabilizes the complex involve (a) an Asn residue as a donor to O4 of mannose, (b) an Asp residue as an acceptor from the same O4, (c) the O3 of mannose is an acceptor of an NH of the backbone. Further stabilization of the complex is provided by extensive van der Waals interactions between the glucose and the phenyl ring of Phe123. No significant conformational adjustments of the protein seem to occur upon glucose binding, when compared to the uncomplexed binding site, at least at this resolution. The only binding partner of the glucose molecule that does move to some degree is Asp81β, but as a whole the recognition seems to occur via a lock-and-key model.

Docking of Me-α-D-mannopyranoside and comparison with glucose binding

The results of the optimization of the Me-α-D-manno-pyranoside in the binding site of lentil lectin is represented in Fig. 5. The hydrogen bonds and the binding energies are

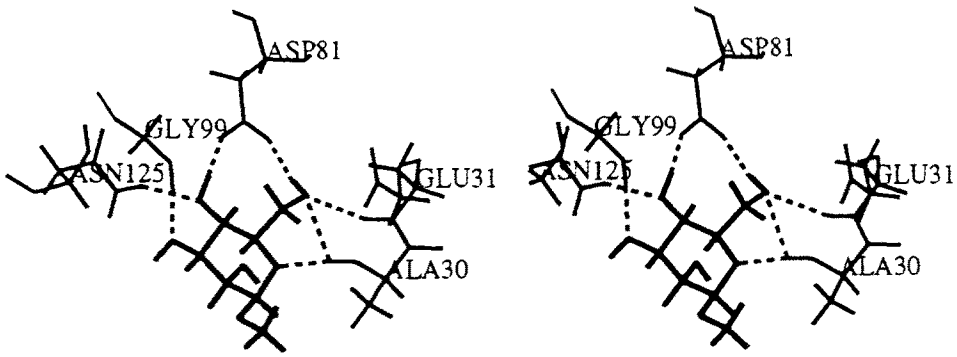


Figure 5. Stereoscopic representation of Me-α-D-mannopyranoside modelled in the binding site of lentil lectin. For the sake of clarity, only amino acids involved in intermolecular hydrogen bonding are displayed. Hydrogen bonds are represented by dotted lines.

Table 2. Comparison of the donor–acceptor distances for the possible hydrogen bonds in the lentil lectin binding site of the experimental D-glucose complex and the modelled Me-α-D-mannopyranoside complex.

Hydrogen bonds	D-*glucose*	*Me-α-D-mannopyranoside*
Ala30α NH → O5	3.1 Å	2.9 Å
Ala30α NH → O6	3.0 Å	3.1 Å
O4H → Asp81β OD2	2.8 Å	2.6 Å
O6H → Asp81β OD1	2.8 Å	3.2 Å
Gly99β NH → O3	3.0 Å	3.0 Å
Gly99β NH → O4	3.2 Å	3.8 Å
Glu31α NH → O6	3.0 Å	3.5 Å
Asn125β ND2(H) → O4	3.1 Å	3.0 Å

summarized in Tables 2 and 3. The position of Me-α-D-mannopyranoside in the binding site is similar to what has been observed in the crystal structure of lentil lectin with glucose described above and in some other lectin–mannose complexes: *Lathyrus ochrus* lectin with Me-α-D-mannopyranoside [9], pea lectin with a trimannoside [12] and ConA with Me-α-D-mannopyranoside [8]. In a previous modelling study, it was established that several orientations of a mannose in the binding site of ConA are possible, but that the one described here is the one with the lowest binding energy [30].

*Conformational analysis of Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside and Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside in the isolated state and the bound state*

The three ω torsion angles represented on Fig. 1 have been varied by 5° steps each, therefore creating 373 248 starting conformers to be tested for an *m*-nitro-benzyl group in position 2-*O*- or 3-*O*- of mannose. Only the conformations without severe steric conflict and within an energy window of 20 kcal mol^{-1} above the energy minimum have been retained. In the isolated state, both molecules are highly flexible since 140 585 conformers and 136 222 conformers are retained for the *m*-nitro-benzyl group on position 2-*O*- and 3-*O*- of Me-α-D-mannopyranoside, respectively. The situation is drastically different when the mannopyranose moiety is buried in the binding site of lentil lectin. The number of conformational states available for the *m*-nitro-benzyl group undergoes a significant reduction. Within a 20 kcal mol^{-1} energy window, there are only 14 024 possible conformations on position 2-*O*- and 9020 for position 3-*O*-. Thus, only 10% of possible conformations still exist for the Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside molecule when bound and 6% for the Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside. The larger reduction of flexibility for the 3-*O*- position can be explained since O3 is more buried in the binding site than O2.

The locations of the low energy conformations can be visualized in a cube where the three axes represent the three rotatable bonds (Fig. 6). Several low energy families are found in both cubes. The pseudo two-fold axis of symmetry for the rotation about $\omega 3$ reflects the symmetry of the phenyl group, only broken by the *m*-nitro substituent. For further refinement, the lowest energy conformation in each of the four main low energy regions of the Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside molecule was selected and similarly one conformation in each of the six low energy regions of the Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside molecule.

For both molecules, the two conformers with the lowest energy (termed A and B for the Me-2-*O*- case and A′ and B′ for the Me-3-*O*- case) are displayed in Fig. 7 and their geometric characteristics are listed in Table 3. The mannose residue always remains in the same position in the binding site and seven hydrogen bonds between the mannose moiety and the amino acids always exist. For both molecules, the orientations with the lowest energy, namely A and A′, allow for the formation of a strong hydrogen bond between the nitro group and one amino acid. The nitro group of Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside hydrogen bonds to the main chain NH group of Gly97β, while in the Me-3-*O*- case a hydrogen bond is formed between the nitro

Table 3. Description of the hydrogen bonds between the lentil lectin binding site and D-glucose, Me-α-D-mannopyranoside, Me-2-*O*- and Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside.

Name	*Torsion angle* (°)	*Relative* E_{bind} (*kcal mol*$^{-1}$)	*Strong hydrogen bond*	*Weak hydrogen bond*
Glucose		ND	Gly99 NH → O3 Asn125 ND2(H) → O4 O6H → Asp81 OD1	Ala30 NH → O6 Glu31 NH → O6 O4H → Asp81 OD2 Ala30 NH → O5 Gly99 NH → O4
Me-α-D-mannopyranoside		ND	Ala30 NH → O5 O4H → Asp81 OD2 Gly99 NH → O3 Asn125 ND2(H) → O4	Ala30 NH → O6 Glu31 NH → O6 O6H → Asp81 OD1
Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside				
A	ω1 = 293.2° ω2 = 204.8° ω3 = 235.5°	0	Ala30 NH → O5 O4H → Asp81 OD2 O6H → Asp81 OD1 Gly99 NH → O3 Asn125 ND2(H) → O4 Gly97 NH → O132 (nitro)	Ala30 NH → O6 Glu31 NH → O6
B	ω1 = 261.6° ω2 = 61.6° ω3 = 263.4°	3.6	Ala30 NH → O5 O4H → Asp81 OD2 O6H → Asp81 OD1 Gly99 NH → O3 Asn125 ND2(H) → O4	Ala30 NH → O6 Glu31 NH → O6
Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside				
A′	ω1′ = 308.3° ω2′ = 303.8° ω3′ = 123.9°	2.7	Ala30 NH → O5 O4H → Asp81 OD2 O6H → Asp81 OD1 Gly99 NH → O3 Asn125 ND2(H) → O4 Asn125 ND2(H) → O131 (nitro)	Ala30 NH → O6 Glu31 NH → O6
B′	ω1′ = 257.8° ω2′ = 62.1° ω3′ = 249.4°	5.5	Ala30 NH → O5 O4H → Asp81 OD2 O6H → Asp81 OD1 Gly99 NH → O3 Asn125 ND2(H) → O4	Ala30 NH → O6 Glu31 NH → O6

Definition of strong hydrogen bond: 2.5 Å < dist(D-A) < 3.1 Å and 120° < ang(D-H-A).
Definition of weak hydrogen bond: 2.5 Å < dist(D-A) < 3.5 Å and 105° < ang(D-H-A) < 120° where: dist(D-A) = distance between the donor and the acceptor and ang(D-H-A) = angle between donor, hydrogen and acceptor. In the case of D-glucose, hydrogen positions were calculated using X-PLOR.
ND, not determined.

group and the side chain ND2(H) of Asn125β. In the case of Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside the energy is further lowered by significant van der Waals interactions between the benzyl group and the aromatic amino acids Tyr100β and Trp128β. This situation is more pronounced in the A′ case. In the Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside structures, the benzyl group is pointing towards the external part of the binding site and the interactions between the aglycon and the protein are less intensive, although in the A conformation, van der Waals interactions between the nitrobenzyl group and the side chain of Tyr100β are clearly present.

Discussion

Although the binding of mannose by the Glc/Man specific legume lectins has been investigated in detail by crystallographic techniques, only in the case of LOL I has a refined complex with glucose been described [9]. The lentil lectin–glucose complex presented in this paper confirms the

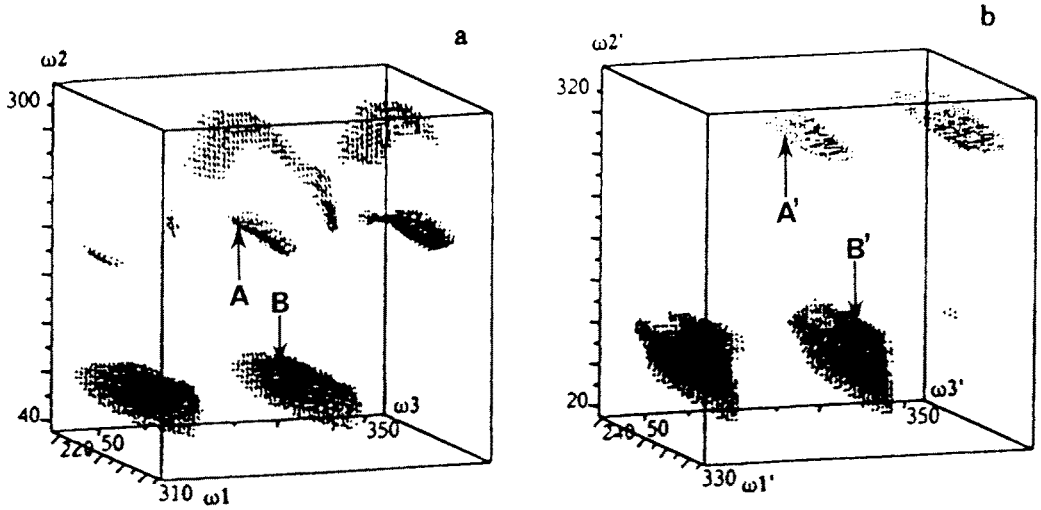


Figure 6. Three-dimensional iso-contour energy maps of (a) Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside and (b) Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside in the binding site as a function of ω1, ω2, ω3 and ω1′, ω2′, ω3′, respectively. The dots represent all the conformers with energy less than 10 kcal mol^{-1} above the minimum. The location of the best conformers A, B and A′, B′ are indicated by arrows.

mode of binding of glucose to the Glc/Man specific group of lectins and confirms our modelling studies with Me-α-D-mannopyranoside.

In the crystal structures of LOL I with mannose and glucose, the loop around Ala30α undergoes a significant shift of about 1 Å upon binding the monosaccharide. No such structural adjustment is observed for lentil lectin. This can be explained by the presence of a phosphate ion in the binding site of the uncomplexed lentil lectin structures (R. Loris, unpublished). This phosphate is connected by a bridging water molecule to the main chain NH of Ala30α and thus pulls this loop closer to the monosaccharide binding site. The conformational rigidity of the lentil lectin molecule and the lack of structural adjustment upon glucose binding that is apparent from our crystal structure are essential to the success of the modelling studies presented here.

It was our aim to understand the molecular basis of the higher affinity for 2-*O*- and 3-*O*-substituted monosaccharides by lentil lectin and the reasons why such substituents do not have any significant effect on the affinity of monosaccharides to ConA. The present calculations yield 'apparent' binding energies for the binding of Me-2-*O*- and Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside to lentil lectin which are in the range of -30 to -35 kcal mol^{-1}. They can be compared to a value of about -25 kcal mol^{-1} calculated for the Me-α-D-mannopyranoside residue. Such a decrease in the binding energy is due to favourable interactions between the *m*-nitro-benzyl group and the adjacent amino acids and therefore explains the increase in affinity experimentally observed [17]. However, we do not observe much difference between the calculated binding energies of the Me-2-*O*- and Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranosides despite the observed lower inhibition concentration measured for the Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside [17]. There are two reasons that can explain the discrepancy between observed affinities and calculated energies of binding.

First, the entropy term has to be taken into account when comparing affinity. When binding a flexible molecule, the entropy barrier $T\Delta S^\circ$ due to the loss of flexibility can have a value of several kcal mol^{-1} [36]. The calculations of entropy terms in binding studies present several difficulties. Second, it should be noted that no water molecules were explicitly included in our modelling studies as there is at the moment no adequate procedure developed to treat the influence of solvent during docking simulations. Nevertheless, from the high resolution structures of legume lectins with larger oligosaccharides [7, 10–11], the importance of bridging waters in the recognition of carbohydrates by legume lectins has been demonstrated [37]. In order to calculate accurately the relative free energy of binding, molecular dynamics or Monte Carlo methods have to be considered (see review by Kollman [38]). Such extensive simulations were not envisaged in the present study, where the modelling part aimed at a qualitative description of the interactions.

Previous modelling calculations on the binding site of ConA using the same force field [30] have proved to predict correctly the crystallographic mode of binding of monosaccharides. Therefore, our results are believed to be reliable with regard to the essential features of the interaction. It is thus likely that the interactions between lentil lectin and the 2-*O*- and 3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranosides are further stabilized by solvent bridges that may also modulate the difference in fine specificity between 2-*O*- and 3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranosides observed in solution. Lentil lectin belongs, together with pea lectin and LOL I, to a group of highly homologous

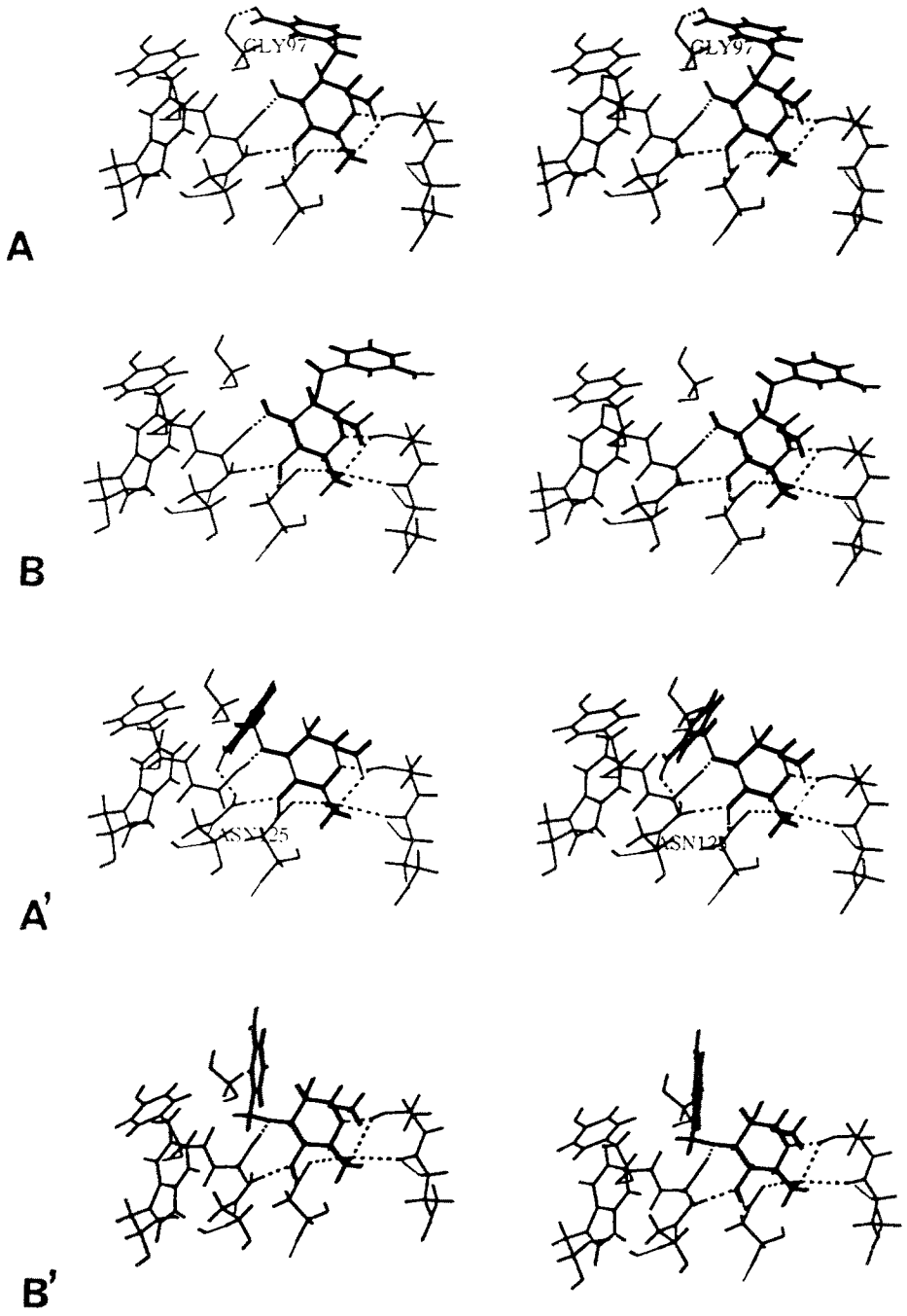


Figure 7. Stereoscopic representation of the two lowest energy conformations of Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside (A and B) and Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside (A′ and B′) in the binding site of lentil lectin. For the sake of clarity, only the amino acids involved in the hydrogen bonds and in the stacking of the aromatic group are displayed. Hydrogen bonds are represented as dotted lines.

Glc/Man specific lectins that all possess a high affinity for 2-*O*- and 3-*O*- substituted mannoses and glucoses. The other well-studied Glc/Man specific lectin, concanavalin A, does not share the same preference for hydrophobic groups on O3 of glucose and mannose. It was suggested [9, 17] that substituents on O3 and O2 of glucose and mannose would interact primarily with the side chains Tyr100β and Trp128β of lentil lectin (and of the other *Viciae* tribe lectins with a similar specificity). In ConA, the two equivalent residues are Leu229 and Ile17 respectively. From our modelling studies, it was found that Tyr100β and Trp128β of lentil lectin are indeed involved in nonpolar interactions with the nitrobenzyl group of Me-3-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside and that in the case of Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside little or no interaction can be seen at all with these side chains. The binding of these

ligands is further stabilized by hydrogen bonds that could also be produced in a complex with ConA.

A complete answer to the problem of the different specificity of ConA and lentil lectin was provided upon inspection of the coordinates of ConA superimposed upon the modelled structures. In the case of Me-3-*O*-(*m*-nitrobenzyl)-α-D-mannopyranoside, it was immediately clear that the A′ conformation was totally inaccessible because of strong steric clashes with the side chain of Arg228 of ConA (Gly99*β* in lentil lectin) and further unfavourable interactions between the nitro group and the carboxyl group of Asp16. In contrast, the sterical problems produced by the B′ conformation were less severe, but this conformation lacks, of course, the stabilizing hydrogen bond with the backbone ND2(H) Asn14, the ConA equivalent of Asn125*β* in lentil lectin. This then explains the extremely poor inhibitory capacities of 3-*O*-substituted mannoses and glucoses for ConA: a number of large side chains, not present in the *Viciae* group of lectins, strongly restricts the conformational space available to an otherwise highly flexible molecule while this drastic loss in entropy is not compensated by any significant other interaction.

In the case of Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside, the possible stabilizing interaction of the nitro group in the A conformation with the NH group of Thr226, the ConA equivalent of Gly97*β*, is also impaired by steric conflicts, i.e. with the main chain carbonyl of Gly224 and the side chain of Thr226. Also, the B conformation is unfavourable in this case as possible clashes can be observed with Ser169 and Leu99. These unfavourable interactions are, however, not as pronounced as for Me-3-*O*-(*m*-nitrobenzyl)-α-D-mannopyranoside and could be accounted for by some relatively small movements of both the relevant side chains and the nitrobenzyl group. This then explains why Me-2-*O*-(*m*-nitro-benzyl)-α-D-mannopyranoside is still an inhibitor of dextran precipitation by ConA, although not as strongly as for lentil lectin. These conclusions should, of course, be confirmed by further crystallographic studies.

The coordinates of the lentil lectin–glucose complex have been submitted to the Protein Data Bank and are available as entry 1LEM.

Acknowledgements

This work has been supported by a grant from the V.L.A.B. project of the Flemish government. R. Loris is a research associate of the N.F.W.O. Julie Bouckaert and Jurgen Pletinckx acknowledge the I.W.O.N.L. for financial support.

We thank Maria Vanderveken for excellent technical assistance. The authors are indebted to the following colleagues that made available crystallographic coordinates prior to publication: Dr Y. Bourne and Dr C. Cambillau for the *Lathyrus ochrus* lectin structures, Dr L. Delbaere for the coordinates of lectin IV from *Griffonia simplicifolia* and Dr Naismith and Dr J Helliwell for the refined structure of cadmium-substituted concanavalin A.

References

1. Van Driessche E (1988) In *Advances in Lectin Research* (Franz H, ed.) Vol. 1, pp. 73–134. Berlin: VEB Verlag Volk und Gesundheit.
2. Hardman KD, Agarwal RC, Freiser MJ (1982) *J Mol Biol* **157**:69–89.
3. Einspahr H, Parks EH, Suguna K, Subramanian E, Suddath FL (1986) *J Biol Chem* **261**:16518–27.
4. Bourne Y, Abergel C, Cambillau C, Frey M, Rougé P, Fontecilla-Camps JC (1990b) *J Mol Biol* **214**:571–84.
5. Loris R, Steyaert J, Maes D, Lisgarten J, Pickersgill R, Wyns L (1993) *Biochemistry* **32**:8772–81.
6. Shaanan B, Lis H, Sharon N (1991) *Science* **254**:862–66.
7. Delbaere LTJ, Vandonselaar M, Prasad L, Quail JW, Wilson KS, Dauter Z (1993) *J Mol Biol* **230**:950–65.
8. Derewenda Z, Yariv J, Helliwell JR, Kalb AJ, Dodson EJ, Papiz MZ, Wan T, Campbell J (1989) *Embo J* **8**:2189–93.
9. Bourne Y, Roussel S, Frey M, Rougé P, Fontecilla-Camps JC, Cambillau C (1990a) *Proteins* **8**:365–76.
10. Bourne Y, Rougé P, Cambillau C (1990c) *J Biol Chem* **265**:18161–65.
11. Bourne Y, Rougé P, Cambillau C (1992) *J Biol Chem* **267**:197–203.
12. Rini JM, Hardman KD, Einspahr H, Suddath FL, Carver JP (1993) *J Biol Chem* **268**:10126–32.
13. Imberty A, Bourne Y, Cambillau C, Rougé P, Pérez S (1993) *Adv Biophys Chem* **3**:71–118.
14. Lemieux RU (1989) *Chem Soc Rev* **18**:347–74.
15. Quiocho FA (1989) *Pure Appl Chem* **61**:1293–306.
16. Allen AK, Desai NN, Neuberger A (1976) *Biochem J* **155**:127–35.
17. Loganathan D, Osborne SE, Glick GD, Goldstein IJ (1992) *Arch Biochem Biophys* **299**:268–74.
18. Poretz RD, Goldstein IJ (1970) *Biochemistry* **9**:2890–96.
19. Loris R, Lisgarten J, Maes D, Pickersgill R, Körber F, Reynolds C, Wyns L (1992) *J Mol Biol* **223**:579–81.
20. Messerschmidt A, Pflugrath JW (1987) *J Appl Cryst* **20**:306–15.
21. Pflugrath JW, Messerschmidt A (1989) *MADNESS Manual of FAST Diffractometer*. Delft, The Netherlands: Enraf-Nonius.
22. Fitzgerald PMD (1988) *J Appl Cryst* **21**:273–78.
23. CCP4 (1979) The SERC (UK) *Collaborative Computing Project No. 4:* A Suite of Programs for Protein Crystallography, distributed from Daresbury Laboratory, Warrington WA 4AD, UK
24. Navaza J (1990) *Acta Cryst* **A46**:619–20.
25. Castellano EE, Oliva G, Navaza J (1992) *J Appl Cryst* **25**:281–84.
26. Brünger AT (1990) *X-PLOR Version 2.1: A System for Crystallography and NMR*. New Haven, CT: Yale University.
27. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) *J Appl Cryst* **24**:946–50.
28. White DNJ, Guy MHP (1975) *J Chem Soc Perkin Trans* **2**:43–6.

29. Clark M, Cramer RD, Van Opdenbosch N (1989) *J Comp Chem* **10**:982–1012.
30. Imberty A, Hardman KD, Carver JP, Pérez S (1991) *Glycobiology* **1**:631–42.
31. Press WH, Flannery BP, Teukolsky SA and Vetterling WT (1986) In *Numerical Recipes, the Art of Scientific Computing.* Cambridge: Cambridge University Press.
32. Berthod H, Pullan A (1965) *J Chem Phys* **62**:942–46.
33. Haneef I (1990) *J Mol Graph* **8**:45–51.
34. Mayer D, Naylor CB, Motoc I, Marshall GR (1987) *J Comput-Aided Mol Design* **1**:3–16.
35. Dewar MJS, Thiel W (1977) *J Am Chem Soc* **99**:4899–907.
36. Carver JP, Michnick SW, Imberty A, Cumming DA (1989) In *Computer Modelling of Carbohydrate Molecules* (Brady JW, French AD eds), ACS Series 430, pp. 266–80. Washington DC: Americal Chemical Society.
37. Bourne Y, Cambillau C (1993) In *Water and Biological Macromolecules* (Westho E ed.), *Topic Molec Struct Biol* **17**:321–37.
38. Kollman P (1993) *Chem Rev* **93**:2395–417.
39. Kraulis PJ (1993) *J Appl Cryst* **24**:946–50.
40. Kabsch W, Sander S (1983) *Biopolymers* **22**:2577–637.