Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Binding of *N*-acetylglucosamine oligosaccharides to hen egg-white lysozyme: a powder diffraction study

R. B. Von Dreele

XPS/IPNS Divisions, Argonne National Laboratory, Argonne, IL 60439, USA

Correspondence e-mail: vondreele@anl.gov

Received 31 August 2004
Accepted 12 October 2004

PDB References: NAG_2–HEWL, 1sf4; NAG_3–HEWL, 1sf6; NAG_4–HEWL, 1sf7; NAG_5–HEWL, 1sfb; NAG_6–HEWL, 1sfg.

The binding of *N*-acetylglucosamine oligosaccharides (NAG_n, n = 2–6) to hen egg-white lysozyme (HEWL; EC 3.2.1.17) was investigated by X-ray powder diffraction at room temperature. Each NAG_n examined was found to bind to lysozyme in rapid-precipitation preparations in 1.0 *M* NaCl pH 6.0 buffer. The location of each NAG_n was easily found from difference Fourier maps generated from structure factors extracted during preliminary Rietveld refinements. Full NAG_n–protein structures were subjected to combined Rietveld and stereochemical restraint refinements (R_{wp} = 2.28–2.59%; R_p = 1.81–2.04%; R_F^2 = 3.91–5.80%) and revealed binding modes for NAG_n that depended on the length of the NAG oligosaccharide. The NAG_2 ligand was found in the *BC* sites in the cleft of HEWL, NAG_3 was found to bind in both the *ABC* and *BCD* sites in the ratio 35:65 and NAG_4 and NAG_5 bound to the *ABCD* and *ABCDE* sites, respectively, while NAG_6 only bound to sites *ABCDE*, leaving the *F* site empty with the remaining saccharide ring located in a solvent region adjacent to the *A* site. All protein powder diffraction patterns in this study consisted of extremely sharp Bragg peaks consistent with ∼1 µm crystallites that were devoid of line-broadening defects. Details of the stereochemical restraints used in these refinements and their impact on structural validation are also discussed.

1. Introduction

The central problem of powder diffraction is that the structural information derivable from a powder diffraction pattern is severely limited relative to a single-crystal data set covering the same region of reciprocal space. Scattering from a single crystal is represented in reciprocal space as an array of delta functions, each slightly broadened by crystalline mosaic effects, and the crystal shape overlaid by a smooth experimental background. Single-crystal data analysis is a simple integration of the peak intensity above background and yields values for the individual reflection structure factors, which are readily modeled as the Fourier transform of the unit-cell contents of atom scattering centers. The representation for a powder diffraction experiment on the same material shows instead a set of nested spherical shells of scattering density that are broadened by instrumental and sample effects (James, 1965; Warren, 1990). The number of these shells increases cubically with distance from the reciprocal-space origin corresponding to the reflection density in the single crystal. Consequently, a powder pattern contains all the intensity information from the reciprocal lattice but all the direction information is lost (Von Dreele, 1989). It is measured as a sequence of steps that scan reciprocal space in one dimension

according to Bragg's Law in usually constant size steps in the independent variable (*e.g.* constant $\Delta 2\theta$). This is in contrast to the three-dimensional suite of observations that comprise a single-crystal diffraction data set, in which the number of observations increases cubically with resolution. In the absence of texture, the shells can be modeled by spherical smearing of individual Bragg reflections that are in turn modeled by the same Fourier transform of the unit-cell contents as used in a single-crystal analysis. Consequently, a powder diffraction experiment does not directly measure the magnitudes of the reflection structure factors, but instead measures the results of a sequence of integrals over thin shells in reciprocal space covering parts of many reflection shells. Each reflection shell contributes to a sequence of these integrals depending on its width. There is additionally some background contribution to each of the powder pattern intensities. This suite of intensities forms a smooth and continuous function with respect to the independent variable (*e.g.* 2θ). A Rietveld refinement is a multiparameter curve fitted by least squares to this smooth function (Rietveld, 1969; Post & Bish, 1989) that is modeled *via* a suite of crystallographic and non-crystallographic parameters (atom positions, thermal motions, lattice parameters, profile coefficients, background terms *etc.*). In this form it has been applied to complex oxides, zeolites and small organic molecules (Cheetham & Taylor, 1977; Poojary & Clearfield, 1997; Harris & Tremayne, 1996) from both X-ray and neutron powder diffraction data. For more complex materials, the Rietveld refinement can be augmented by stereochemical restraints and constraints in order to increase the number of data and reduce the number of parameters, respectively. In a recent study (Pagola *et al.*, 2000) of the malaria pigment β-hematin, the porphyrin ring was modeled as a rigid unit and the side chains subjected to restraints on bond lengths and angles. We have extended this approach to powder diffraction studies of proteins by combined refinement with a powder diffraction pattern and all available stereochemical restraints (Von Dreele, 1999, 2001; Von Dreele *et al.*, 2000). The current system of restraints used in this work is outlined in Appendix *A*.

From the above description of a powder diffraction experiment it is clear that it has a different distribution of data points in reciprocal space than does a single-crystal data set for the same material, making comparisons of 'resolution' *via* minimum *d* spacing between the two types of experiment problematic. It is also clear that powder diffraction instrumentation and samples that produce the sharpest lines will provide more information than ones that do not; these are distinguished by referring to them as 'high-resolution' powder diffraction. Moreover, this fundamental difference in the way a powder data set scans reciprocal space means that adoption of reflection 'counts' to powder data is probably not appropriate (McCusker *et al.*, 1999) for evaluating data-to-parameter ratios. Schemes have been proposed for modifying these reflection counts to give a more realistic assessment of the 'information content' of a powder pattern. For example, Altomare *et al.* (1995) suggested a scheme based on the ratio of overlapped Bragg intensity to the total. This can be used to assess the likely success of an *ab inito* structure solution, where a suite of well determined individual structure factors are essential, or a Rietveld refinement based only on a single powder diffraction data set. However, consideration of the least-squares mathematics indicates that the number of measurements used to develop the normal equations and subsequently fill the Hessian matrix is the appropriate scheme. In the case of a Rietveld refinement this is the number of observed profile points (Prince, 1981). Nonetheless, an individual protein powder pattern with its high overlap and limited *d*-spacing coverage does not contain sufficient information for an unrestrained structure refinement, hence our approach of combining stereochemical restraints with the powder diffraction data. Because the results are strongly affected by the restraints on all short-range features of the structure, as will be seen below, this has an impact on the use of these as independent structure validators.

During a Rietveld refinement, the calculated profile intensity is developed as a sum of the contributions from the neighboring reflections. If desired, ratios of these contributions to the total can then be used to apportion the observed intensity among these reflections, which when integrated give values for 'observed' structure factors. These are strongly biased by the model used in the Rietveld refinement but can be used in difference Fourier calculations for *e.g.* structure completion and we recently demonstrated this in our study of the binding of *N*-acetylglucosamine (NAG) to hen egg-white lysozyme (HEWL; Von Dreele, 2001). In this case, NAG was found to bind to HEWL only in the α-anomer form in a different orientation than previously observed from a single-crystal study (Perkins *et al.*, 1978). Rapid precipitation in the preparation of the powder sample probably prevented α/β-anomer conversion, in contrast to the 24 h soaking required for the single-crystal preparation. Subsequent combined Rietveld and stereochemical restraint refinement (Von Dreele, 1999) gave a modest resolution ($\sim$3 Å) description of the structure. To continue the demonstration of the usefulness of powder diffraction for protein–ligand binding, we present the results of binding five oligosaccharides (NAG_n, n = 2–6) to HEWL.

2. Experimental

HEWL (EC 3.2.1.17; Fisher Biotech, 3× crystallized, lot Nos. 995417-12 and 996924-12), *N*,*N*′-diacetylchitobiose (NAG_2, $C_{16}H_{28}N_2O_{11}$, Sigma Chemical Co., lot 21K4059), *N*,*N*′,*N*″-triacetylchitotriose (NAG_3, $C_{24}H_{41}N_3O_{16}$, Sigma Chemical Co., lot 11K4036), tetra-*N*-acetylchitotetraose (NAG_4, $C_{32}H_{54}N_4O_{21}$, Sigma Chemical Co., lot 57H4098), penta-*N*-acetylchitopentaose (NAG_5, $C_{40}H_{67}N_5O_{26}$, Sigma Chemical Co., lot 57H4102), hexa-*N*-acetylchitohexaose (NAG_6, $C_{48}H_{80}N_6O_{31}$, Sigma Chemical Co., lot 57H4099), 0.05 *M* Na_2HPO_4/KH_2PO_4 buffer pH 6.0 (Fisher Scientific) and NaCl ('Certified for Biological Use', Fisher Scientific) were used as received. In a typical sample preparation, a polycrystalline slurry was made by combining 25.7 mg lysozyme (1.8 µmol),

Table 1
Crystallographic and refinement data for *N*-acetyl-D-glucosamine oligomer–HEWL complexes.

NAG = *N*-acetylglucosamine. Unless otherwise noted, values in parentheses are estimated standard deviations in the values shown. Wavelength calibrations were obtained from the fitted positions of the first six reflections from a NIST SRM1976 alumina plate; r.m.s. position errors were 0.0004° 2θ. Data-collection step count times were determined by the following algorithm: for $2\theta < 5^\circ$, $t = 4$ s; for $2\theta > 5^\circ$, $t = 6.4691 - 0.9877(2\theta) + 0.0988(2\theta)^2$ s. The residuals $R_{wp} = 100 \times [\sum w(I_o - Ic)^2/\sum I_o^2]^{1/2}$, $R_p = 100 \times \sum |I_o - I_c|/\sum I_o$, $R_{wexp} = 100 \times (N_{steps}/\sum wI_o^2)^{1/2}$ and $R_F^2 = 100 \times \sum |F_o^2 - F_c^2|$, where I_o and I_c are the observed and calculated powder diffraction profile intensities in a pattern of N_{steps} and w is the weight associated with each I_o. Those residuals listed 'w/o NAG' are the initial residuals computed using 1ja7 with the NAG removed as the structural model. F_o^2 is the value of the structure factor extracted during the Rietveld refinement and F_o^2 is the calculated structure factor. The correlation coefficient is calculated from $C = \sum(I'_o - \bar{I}'_o)(I'_c - \bar{I}'_c)/[\sum(I'_o - \bar{I}'_o)^2\sum(I'_c - \bar{I}'_c)^2]^{1/2}$, where I'_o and I'_c are the observed and calculated profile intensities minus the background contribution determined during the Rietveld refinement. *PROCHECK* (Laskowski *et al.*, 1993) compares protein stereochemistry with expected values (Engh & Huber, 1991; Morris *et al.*, 1992). *ERRAT* and *ERRAT*2 examine distributions of near atom–atom interactions and compare them with that observed in high-resolution protein structures (Colovos & Yeates, 1993). The Matthews (1976) coefficient includes the contribution from the NAG_n oligomer.

Material	NAG_2–HEWL	NAG_3–HEWL	NAG_4–HEWL	NAG_5–HEWL	NAG_6–HEWL
Crystal data					
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Unit-cell parameters					
a (Å)	78.928 (3)	78.770 (3)	78.933 (3)	79.000 (3)	78.895 (3)
c (Å)	38.178 (2)	38.368 (2)	38.202 (2)	38.137 (2)	38.254 (2)
V (Å^3)	237830 (30)	238060 (30)	238010 (30)	238010 (30)	238110 (30)
Matthews coefficient (Å^3 Da^{-1})	2.163	2.135	2.108	2.080	2.053
Solvent content (%)	42.68	41.93	41.19	40.39	39.62
PDB code	1sf4	1sf6	1sf7	1sfb	1sfg
Sample preparation					
Lysozyme, mg (µmol)	25.7 (1.8)	24.0 (1.7)	24.4 (1.7)	15.3 (1.1)	15.3 (1.1)
NAG_n, mg (µmol)	1.4 (3.3)	1.6 (2.6)	2.1 (2.5)	1.4 (1.4)	1.9 (1.5)
Powder data collection					
2θ range (°)	1.0–13.504	1.0–12.498	1.0–12.498	1.0–12.498	1.0–12.498
Steps	6253	5750	5750	5750	5750
Step time (s)	4.0–11.15	4.0–9.56	4.0–9.56	4.0–9.56	4.0–9.56
I_{max} (σ)	2757 (28)	2878 (26)	2596 (25)	2553 (24)	2332 (23)
I_{min} (σ)	432 (12)	501 (11)	431 (11)	447 (11)	427 (11)
Combined Rietveld and stereochemical restraint refinement results					
N_{ref}	2797	2255	2255	2254	2254
Resolution (Å)	40.13–2.98	40.13–3.22	40.13–3.22	40.13–3.22	40.13–3.22
$N_{restraints}$	5188	5312	5257	5346	5391
N_{obs}	11149	10783	10729	10894	10861
$N_{parameters}$	3111	3196	3195	3237	3279
R_{wp} (%), w/o NAG	3.70	4.52	4.75	4.46	4.58
R_p (%), w/o NAG	2.72	3.27	3.29	3.18	3.26
R_F^2 (%), w/o NAG	10.09	11.64	7.96	6.99	9.35
R_{wp} (%)	2.52	2.28	2.47	2.46	2.59
R_p (%)	1.97	1.81	1.93	1.91	2.04
R_F^2 (%)	4.91	3.91	5.15	4.35	5.80
R_{wexp} (%)	1.56	1.37	1.55	1.48	1.60
Correlation coefficient	0.9815	0.9831	0.9804	0.9782	0.9713
A_s	5.147 (4)	5.354 (5)	5.178 (4)	5.645 (4)	5.619 (5)
B_s	1.301 (7)	1.077 (7)	1.011 (5)	0.713 (4)	0.594 (4)
Powder profile parameters					
X	0.562 (9)	0.538 (13)	0.443 (11)	0.398 (13)	0.363 (15)
X_e	0.115 (16)	0.219 (22)	0.328 (20)	0.314 (23)	0.321 (26)
Validation results from *PROCHECK* and *ERRAT*					
Bond error (Å)	0.025	0.031	0.030	0.027	0.026
Angle error (°)	1.76	2.26	1.98	1.89	1.90
Residues in core region (%)	88.5	85.8	85.8	86.7	85.0
ω torsion-angle error (°)	3.1	2.3	1.9	3.4	3.3
Bad contacts per 100 residues	0	3.1	0.8	0	0
ζ-angle error (°)	2.1	2.4	2.3	2.3	2.1
Hydrogen-bond energy error (kJ mol^{-1})	3.7	3.7	3.3	3.3	3.3
R.m.s. deviations from planes (Å)	0.014	0.011	0.009	0.014	0.014
ERRAT quality factor (%)	90.9	90.9	90.1	89.3	85.1
*ERRAT*2 quality factor (%)	72.7	47.1	59.2	44.6	54.2

1.4 mg NAG_2 (3.3 µmol) and 200 µl 1.0 *M* NaCl pH 6.0 buffer with an agate mortar and pestle. Polycrystalline precipitate formed within a few seconds. The slurry was loaded into a 1.5 mm diameter glass capillary and centrifuged to pack the slurry. Excess mother liquor was removed and the capillary was flame-sealed to prevent subsequent solvent evaporation. Samples were ~8 mm long. As each sample was prepared, collection of X-ray powder diffraction data was immediately begun at room temperature (296 K) on beamline X3b1 at the National Synchrotron Light Source, Brookhaven National Laboratory, equipped with a double Si(111) monochromator and a Ge(111) analyser; $\lambda = 0.70039$ (3) Å; the sample was

spun during data collection to ensure good powder averaging. Steps of $\Delta 2\theta = 0.002^\circ$ were selected along with the data-collection time algorithm indicated in Table 1 to give reasonable counting statistics for all steps within the total scan time limitation imposed by the interval between synchrotron-ring fills. Although this step size provides fewer steps (∼3.5) across the Bragg peaks than recommended by McCusker *et al.* (1999), it is well in line with the experience of Hill & Madsen (1986) in their analysis of the effect of step size on the results of Rietveld refinements. The counting-time algorithm is similar in effect to that proposed by Madsen & Hill (1994) for improving the powder profile counting statistics at the high-angle end of the pattern and compensated for the natural fall-off of scattered intensity with angle and the loss of synchrotron-beam intensity. Each data collection required 10–12 h exposure to the X-ray beam; preliminary experiments showed that radiation damage with this wavelength became discernable only after >24 h exposure. Each powder pattern intensity was corrected for dead-time effects ($\delta t = 1.6$ µs) and scaled to constant monitor counts; weights were determined from counting statistics and proper error propagation through the corrections. Details of the sample preparations and data collections are given in Table 1.

The NAG–HEWL structure (PDB code 1ja7) obtained in our previous study (Von Dreele, 2001) was used as the starting model for the analysis of the NAG_n–HEWL powder diffraction data. The NAG atoms were removed and the remaining HEWL structure used for the initial calculations. After preliminary Rietveld refinements of non-structural features (background, lattice parameters and profile coefficients) for each of the NAG_n–HEWL materials, difference Fourier maps were constructed in each case by using structure factors extracted from the respective powder diffraction profiles. The residuals for these initial extractions are shown in Table 1. As noted in the introduction, this extraction process operates during a Rietveld refinement by apportioning the observed diffraction profile above background among the contributing reflections according to the ratio of their calculated intensities to the total. These contributions are then summed to give a set of integrated intensities for all reflections within the range of the powder profile; they are then subjected to the usual corrections for reflection multiplicity, Lorentz and polarization to obtain a set of 'observed' structure factors. The ΔF maps (Fig. 1) prepared from these extracted intensities all showed an extended region of density in the vicinity of the previously identified *A*–*E* saccharide-binding sites for HEWL despite the model-dependent bias arising from the extraction process. This was most apparent for NAG_2, where the map showed extensive density (Fig. 1*a*) only in the vicinity of the ligand-binding site; maps for the other complexes were noisier but a readily identifiable density could be found in the binding site (Figs. 1*b*–1*e*). Although the details of these density features were not clear enough to unambiguously place individual NAG atoms, the extent of this density in most cases was consistent with that expected from the size and orientation of the NAG_n oligomer. However, for NAG_3 the density (Fig. 1*b*) suggested that the four sites *A*–*D* were at least partially occupied and that there might be two alternative locations for the ligand. This was modeled by placing an NAG_4 molecule in sites *A*–*D*. Subsequent refinement of the atom-site fractions for the saccharide rings yielded ∼35% occupancy for the *A* site and the remainder (65%) in the *D* site. Sites *B* and *C* were fully occupied. In the case of the longer oligosaccharides, the maps suggested initial placement of just part of the ligand. Subsequent ΔF maps based on these

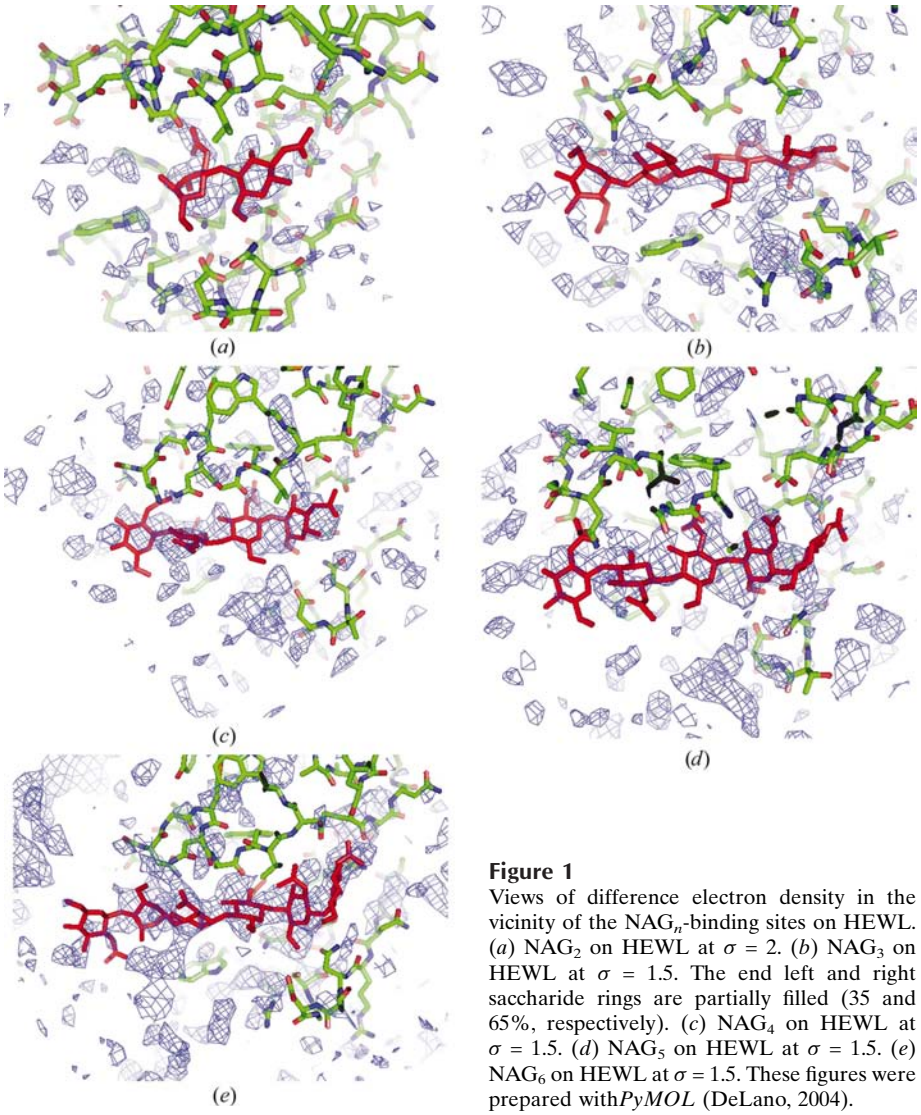

Figure 1
Views of difference electron density in the vicinity of the NAG_n-binding sites on HEWL. (*a*) NAG_2 on HEWL at $\sigma = 2$. (*b*) NAG_3 on HEWL at $\sigma = 1.5$. The end left and right saccharide rings are partially filled (35 and 65%, respectively). (*c*) NAG_4 on HEWL at $\sigma = 1.5$. (*d*) NAG_5 on HEWL at $\sigma = 1.5$. (*e*) NAG_6 on HEWL at $\sigma = 1.5$. These figures were prepared with *PyMOL* (DeLano, 2004).

new partial structures allowed placement of the remainder. Refinement of NAG-group atom fractions was used as an additional check; well placed NAG rings yielded atom fractions that were >90%, while poorly placed ones sometimes yielded group fractions <70%. For NAG_6 the density (Fig. 1*e*) indicated that sites *A*–*E* were occupied, with the remaining ring located in a solvent region adjacent to the *A* site. The *F* site appeared to be unoccupied. It should be noted that the procedure used here differs from our previously reported procedure of subjecting the protein structure to refinement before generating the ΔF maps (Von Dreele, 2001). In the present case, the maps were sufficient to show the extent of the ligand density despite there being no refinement of the protein and maps created after refinement of the protein were differed little in the vicinity of the ligand from the ones shown here.

For each NAG_n–HEWL complex, the ligand was placed by hand to best fit the difference map avoiding intermolecular interferences and the resulting structure was subject to a combined Rietveld and stereochemical restraint least-squares refinement following previously developed procedures (Von Dreele *et al.*, 2000; Von Dreele, 2003) using the *General Structure Analysis System* (*GSAS*; Larson & Von Dreele, 2004) to give the residuals shown in Table 1. Although the disposition of each NAG_n oligosaccharide gives a reasonable fit to the ΔF densities and suitable contacts with the surrounding protein structure, there may be other similar ligand dispositions that satisfy the density and intermolecular contacts and give similar Rietveld fits. The models given here were found by extensive exploration of the two torsion angles about the bridging O atoms in each NAG_n as well as the disposition of the *N*-acetyl and CH_2OH side chains. The Rietveld fit for NAG_2–HEWL is shown in Fig. 2; the others can be found in the supplementary material[1] for this publication. Appendix *A* gives details of the minimization function and some of the restraints used in these analyses. Least-squares protein refinement was achieved by constructing a band-diagonal approximation to the full matrix; band-matrix routines from the *SLATEC* suite (Fong *et al.*, 1993) were adapted for use in *GSAS*. A matrix bandwidth of 100 parameters was chosen for refinement of all NAG_n–HEWL structures studied here. Refinement included the (x, y, z) coordinates of all atoms; all atom displacement factors were fixed at $U = 0.30$ Å^2; the powder diffraction data would not support refinement of individual atom displacement factors or even an overall one for all atoms, as it is essentially identical to the scaling factor. A Babinet's principle modification of all the atom scattering factors according to

$$f = f_0 - A_s \exp(-8\pi^2 B_s \sin^2\theta/\lambda^2)$$

accounted for solvent scattering and facilitated fitting the lowest angle part of the powder diffraction data. The coefficients A_s and B_s were refined as part of the minimization; values are given in Table 1. During all refinements the resulting protein stereochemistry was periodically evaluated with the *PROCHECK* (Laskowski *et al.*, 1993) suite of programs and the structure factors extracted during the Rietveld refinements were evaluated with the *SFCHECK* program (Vaguine *et al.*, 1999). As these results suggested possible modifications to the structures, these were performed graphically using the *SwissPDBViewer* (Guex & Peitsch, 1999)

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: EN5065). Details for accessing these data are described at the back of the journal.

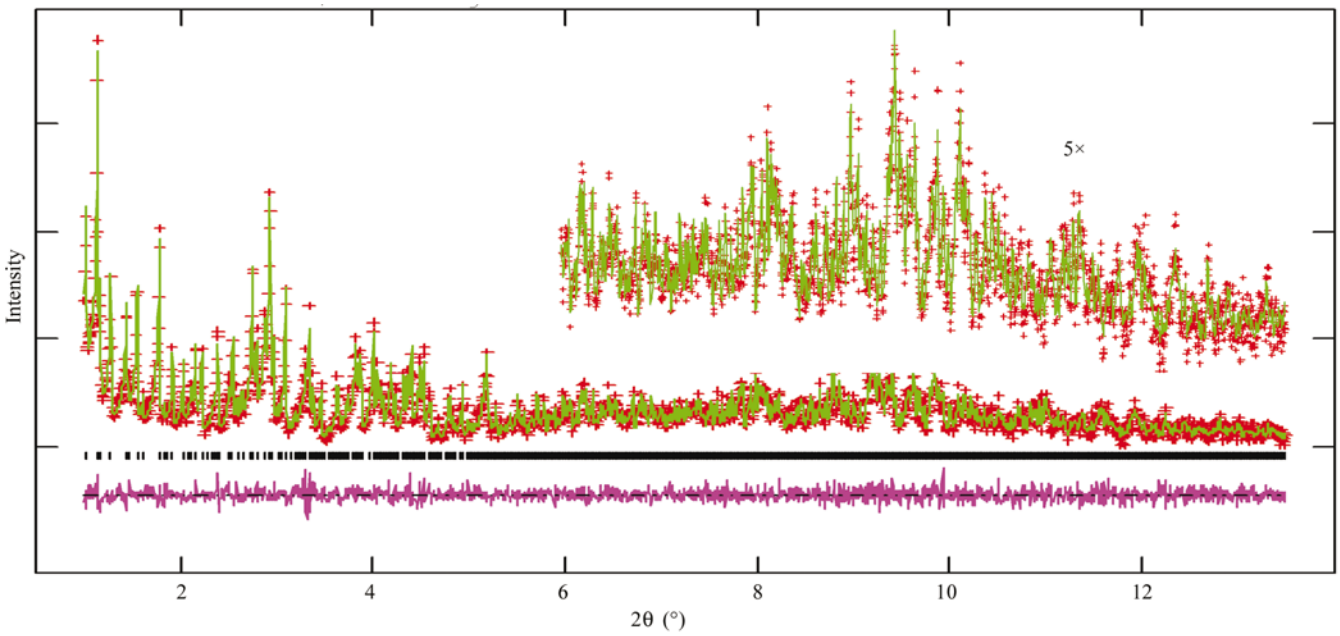

Figure 2
High-resolution X-ray powder diffraction profile from the final Rietveld refinement of NAG_2–HEWL. The observed intensities are shown as red plus signs, the calculated and difference curves as green and magenta lines and the reflection positions as black vertical lines. The background intensity found in the refinement has been subtracted from the observed and calculated intensities for clarity. The inset shows a detail of the fit for the region shown.

Table 2
Root-mean-square displacements in Å of the C^{α} atoms between the structures determined for the NAG_n–HEWL complexes.

	NAG_2–HEWL	NAG_3–HEWL	NAG_4–HEWL	NAG_5–HEWL
NAG_3–HEWL	0.55	—	—	—
NAG_4–HEWL	0.56	0.60	—	—
NAG_5–HEWL	0.69	0.72	0.50	—
NAG_6–HEWL	0.67	0.71	0.44	0.44

Table 3
Root-mean-square displacements in Å of the side-chain atoms between the structures determined for the NAG_n–HEWL complexes.

	NAG_2–HEWL	NAG_3–HEWL	NAG_4–HEWL	NAG_5–HEWL
NAG_3–HEWL	1.07	—	—	—
NAG_4–HEWL	1.17	1.41	—	—
NAG_5–HEWL	1.28	1.48	1.36	—
NAG_6–HEWL	1.32	1.40	1.09	1.24

and *WebLab Viewer Pro* 3.2 (Molecular Simulations Inc.) packages. Modifications included selection of side-chain rotomers and displacement of small groups of atoms to give improvement in protein stereochemistry, occasionally guided by comparison to ΔF maps. As seen in Table 1, inclusion of the ligand and subsequent refinement gave substantially improved residuals. Final evaluation of these structures also included use of the programs *ERRAT* (Colovos & Yeates, 1993) and *ERRAT*2 (Yeates, 2004); the quality factors obtained are listed in Table 1.

3. Discussion

A superposition of the five structures examined here (Fig. 3) suggests that while the protein structure appears to be virtually unchanged (Tables 2 and 3), the NAG oligomers occupy positions *A*–*E* in the HEWL cleft in a variety of ways depending on their length. The greatest variation occurs for the shortest oligomers; as was observed for NAG (Perkins *et al.*, 1978; Von Dreele, 2001), NAG_2 seems to occupy the *CD* sites (Fig. 4) in a different way from the corresponding NAG units of the longer oligomers. As seen in these figures, the longer oligomers occupy many of the sites in more or less the same way. In particular, the saccharide rings in the *C* site are oriented nearly identically for all the longer oligomers; the saccharide rings in the other sites are oriented less identically. Interestingly, the NAG_3 oligomer appears to be disordered and occupies either the *ABC* sites (Fig. 5*a*), identical to previous observations (Cheetham *et al.*, 1992), or the adjacent *BCD* sites (Fig. 5*b*). In the present study, refinement suggests that the latter sites are preferred in the ratio 65:35. This is in contrast to the result reported by Cheetham *et al.* (1992) from single crystals grown over three weeks starting with a 1:1 stoichiometric mixture, where partial hydrolysis of the ligand resulted in only 55% occupancy of the *ABC* sites with no observed occupancy of the *BCD* sites. It may be that the rapid precipitation technique used here trapped NAG_3 ligands in both binding sites and that subsequent hydrolysis of those in the *BCD* sites is somehow inhibited in the solid state; hydrolysis is presumed to occur in the *CD* site (Rupley & Gates, 1967). As expected, the NAG_4 and NAG_5 ligands occupy the *A*–*D* and *A*–*E* sites, respectively. The similarity of their binding can be seen in Figs. 6 and 7.

For NAG_6 (Fig. 8) the ligand appears to occupy the *A*–*E* binding sites on HEWL, thereby leaving the *F* site empty. The

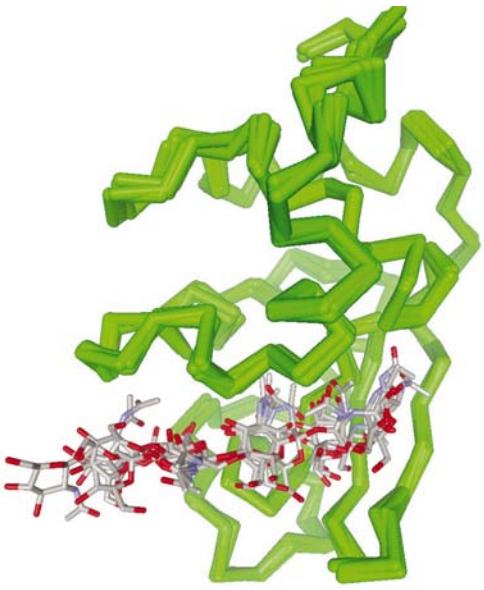

Figure 3
A perspective view of the superposition of the NAG_n–/HEWL (n = 2–6) structures. The backbones of the HEWL structures are in green and the NAG_n ligands are in CPK colors.

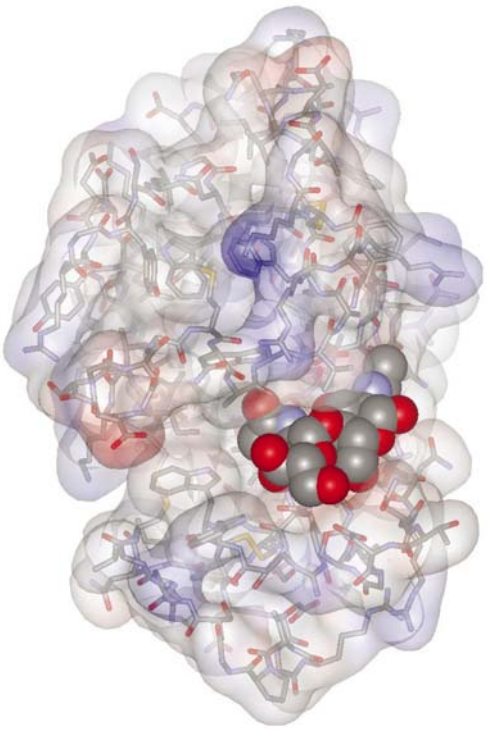

Figure 4
A view of the structure of the HEWL–NAG_2 complex. The ligand occupies lysozyme cleft sites *C* and *D*.

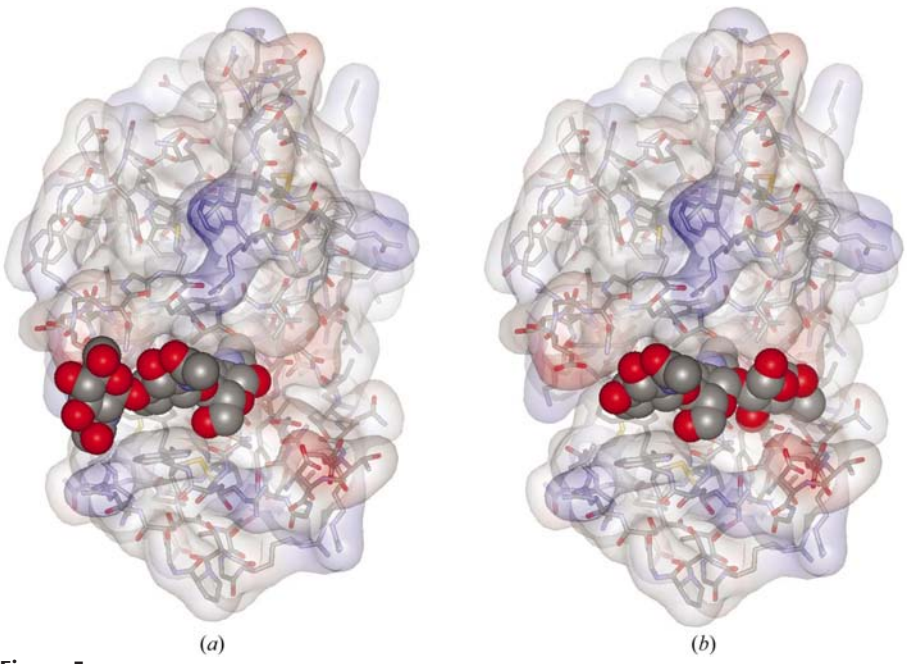

Figure 5
A view of the binding of NAG_3 to HEWL. (*a*) Sites *ABC* with 35% occupancy and (*b*) sites *BCD* with 65% occupancy.

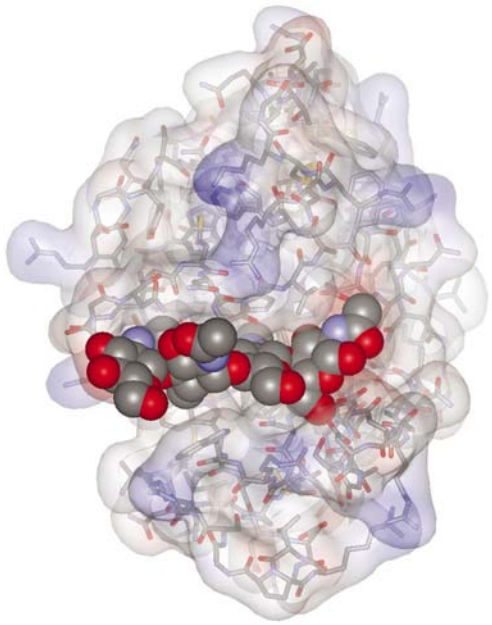

Figure 6
A view of the structure of the HEWL–NAG_4 complex. The ligand occupies lysozyme cleft sites *A*–*D*.

remaining saccharide ring appears to extend out into the solvent region between adjacent protein molecules from the *A*-site end of the ligand. This result conflicts with the long-standing model (Blake *et al.*, 1967) of the hexasaccharide binding to six sites (*A*–*F*) across the cleft of HEWL. Careful examination of the difference map for NAG_6–HEWL in the vicinity of the supposed *F* binding site showed no appreciable density and moreover there is insufficient room for an *N*-acetylglucosamine group in the space available between the neighbouring protein molecules in the NAG_6–HEWL powder structure. Therefore, while it may be that NAG_6 does bind to the six (*A*–*F*) sites on HEWL in solution, the results described here suggest it does not do so in the solid.

As seen in Table 1, the binding of NAG_n oligosaccharides results in significant changes in the unit-cell parameters but less so for the unit-cell volumes. A unit-cell determination of HEWL *via* powder diffraction prepared with the same solvent conditions (1.0 *M* NaCl pH 6.0 buffer) gives a = 79.190 (3), c = 37.925 (2) Å and V = 237 830 (30) Å^3; a second sample prepared the same way gave a = 79.178 (3), c = 37.934 (2) Å and V = 237 810 (30) Å^3. Apparently, the incremental change from increasing the size of the NAG_n oligosaccharide is offset by a reduction in the solvent content. Assuming a volume of 30 Å^3 for a water molecule, each additional NAG group replaces approximately 7.5 water molecules, while the variation in unit-cell size is at most only ∼1 water molecule per HEWL over the entire sequence of materials. As noted earlier by us (Von Dreele, 1999, 2001) Rietveld refinement yields high-precision lattice parameters for these materials; this arises from the sensitivity of the fit to the reflection positions over the entire powder diffraction profile and the sharpness of the diffraction lines (Post & Bish, 1989). However, they are sensitive to systematic effects arising from sample preparation, most notably the solvent conditions. For example, HEWL prepared using the methods described here but with 1.0 *M* NaCl pH 7.0 buffer yields a = 79.020 (3), c = 38.011 (2) Å and V = 237 340 (30) Å^3. Changes are also seen with different NaCl concentrations; our earlier report (Von Dreele, 2001) gave a = 79.132, c = 38.030 Å and V = 238 140 Å^3 for HEWL prepared with 0.5 *M* NaCl pH 6.0 buffer.

The range of r.m.s. distances for the C^α atoms (Table 2) comparing pairs of NAG_n–HEWL structures are 0.44–0.72 Å; a similar study performed for C^α atoms in peanut lectin tetramers (Natchiar *et al.*, 2004) from 2.15–3.5 Å resolution single-crystal data showed r.m.s. distances of 0.13–0.66 Å. A residue-by-residue comparison of r.m.s. values for some of the NAG_n–HEWL pairs shows that most of the differences are concentrated in relatively short segments of the peptide chain. For example, the largest r.m.s. distances for NAG_2/NAG_3–HEWL and NAG_2/NAG_4–HEWL pairs are concentrated in the last ten residues by an essentially uniform ∼1 Å shift in atom positions, while for the NAG_3/NAG_5–HEWL pair it is the Asp18–Gly26 loop that shows a similarly uniform shift. Examination of molecular packing shows that these segments are close to the NAG_n oligosaccharides bound to adjacent

HEWL molecules and they may be forcing these displacements as well as the small lattice-parameter changes noted above. For example, Fig. 9 shows the apparent displacement of the Asp18–Gly26 loop in NAG_5–HEWL compared with its location in NAG_3–HEWL arising from the presence of the NAG segment in the *E* binding site on HEWL. The r.m.s. distances (Table 3) for the side chains reflect different choices of rotomers for their stereochemistry driven by these differences in packing of the NAG_n–HEWL complexes as well as the obvious coupling to the main-chain displacements noted above.

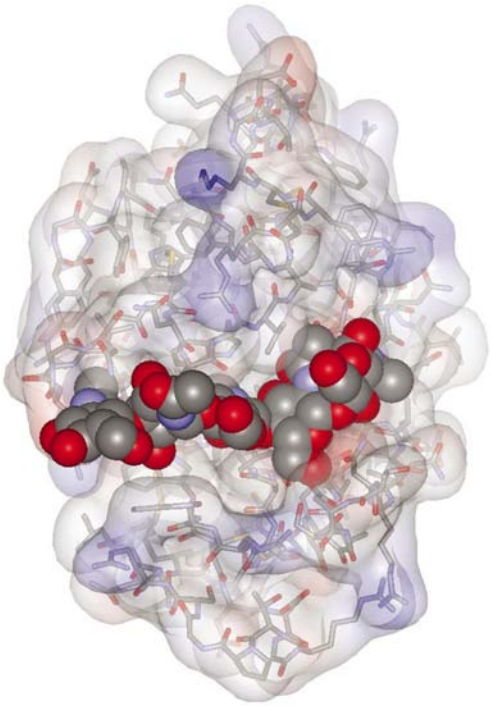

Figure 7
A view of the *A*–*E* binding of NAG_5 to HEWL.

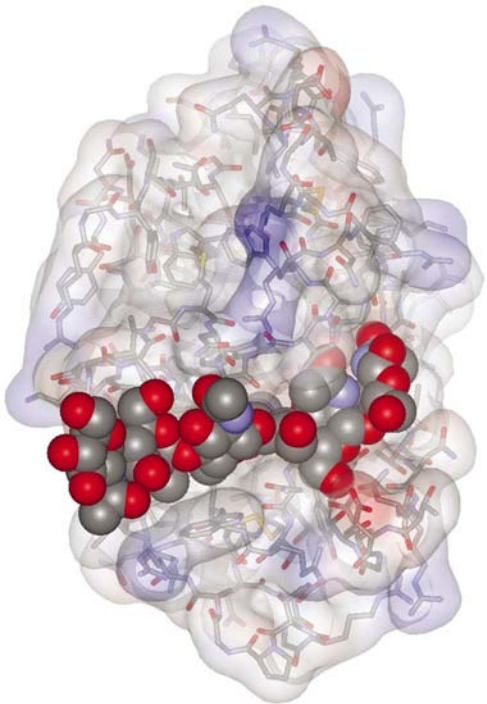

Figure 8
A view of the *A*–*E* binding of NAG_6 to HEWL.

Examination of the validation results from *PROCHECK* and *ERRAT* (Table 1) illustrates the impact of the restraints on these refinements. Most of the values from *PROCHECK* are within their expected ranges for 2 Å resolution single-crystal studies even though the powder data extends to only ∼3 Å and the effective resolution may be closer to 4 Å (Von Dreele, 1999). The latter resolution value is more consistent with the *ERRAT* quality factors, which would be more sensitive to structural errors not subject to the restraints and thus may represent more appropriate cross-validators for these results. Clearly, the *PROCHECK* results are compromised as cross-validators since they also form the basis for the suite of restraints used in these refinements (Dodson, 1998). The values listed instead show the quality of the stereochemical fit and should be considered in conjunction with the *R* values given for the Rietveld component of the refinements.

Our previous X-ray powder diffraction studies (Von Dreele, 1999, 2001; Von Dreele *et al.*, 2000) all yielded material with extremely sharp diffraction peaks and little or no sample broadening. The apparent crystallite sizes in these cases were of the order of 1 µm and the microstrain broadening effects were less than 0.1%. As in the latter study, polycrystalline materials have been prepared from amorphous lyophilized protein and relatively high-salt buffered solvents and as before we obtain polycrystalline material that gives extremely sharp diffraction patterns. Interpretation (Larson & Von Dreele, 2004) of the *X* and X_e coefficients (Table 1) obtained from the Rietveld refinements gives a remarkably uniform effective crystallite size of 0.8 µm for all the NAG_n–HEWL complexes.

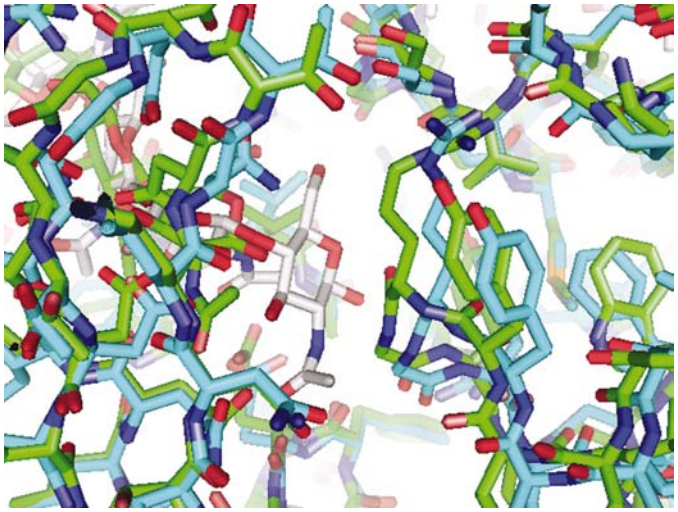

Figure 9
View of the interface between the molecules at x, y, z (left) and $1/2 - y$, $1/2 + x$, $3/4 + z$ (right) of NAG_3–HEWL (C atoms in green; NAG_3 not shown) and NAG_5–HEWL (C atoms in light blue and NAG_5 C atoms in white). The NAG group (center) found in the *E* site of NAG_5–HEWL (left) appears to displace the Asp18–Gly26 loop (blue on right) from its location as found in NAG_3–HEWL (green on right).

Oddly, the slight anisotropy in the crystallite size broadening indicates that all of the complexes show a slightly plate-like habit with an apparent thickness of 0.5 μm; given the lattice parameters one might have expected a more needle-like habit, although large crystals grown for single-crystal studies are usually not needles. No microstrain broadening was observed for any of the complexes. Evidently, all of the polycrystalline lysozyme samples prepared here as precipitate appear to be essentially free of lattice defects and are quite homogeneous with no variability in lattice parameters. Even the NAG_3–HEWL material shows no microstrain effects despite the apparent disorder in ligand binding.

Although the quality of the ΔF maps obtained here from the highly biased structure factors extracted from the powder patterns during Rietveld refinement is not high, it is evident that there is sufficient information to give features in these maps that allow one to at least recognize the location of the NAG_n ligands on HEWL. The relative scattering powers of the ligand (∼200–600 e^-) to protein (∼7000 e^-) in this case are more favorable than for a recent study on the sodium salt of *p*-hydroxybenzoic acid (Dinnebier *et al.*, 1999), where a single H atom (1 e^-) was located in a ΔF map generated from structure factors extracted during a Rietveld refinement of the partial structure (81 e^-). It can be anticipated that improvements in powder data collection by, for example, high-resolution area detectors will improve the quality of the maps over what has been presented here.

Table 4
Weight factors for combined Rietveld and stereochemical refinement of HEW lysozyme–NAG_n complexes.

f_I	f_a	f_d	f_t	f_p	f_v	f_h	f_x	f_R
1.00	1.50	0.30	1.00	0.50	0.50	0.50	0.50	0.20–0.50

Table 5
Pseudopotential coefficients for torsion-angle restraints.

Torsion	A_1	φ_1	B_1	A_2	φ_2	B_2	A_3	φ_3	B_3
Arg-χ_1	−1.96	61.13	0.80	−2.66	181.10	0.58	−2.82	290.89	0.47
Arg-χ_2	−2.30	66.75	0.61	−2.80	177.25	0.48	−2.53	287.29	0.53
Arg-χ_3	−2.63	89.56	0.26	−2.60	180.06	0.49	−2.72	266.30	0.39
Asn-χ_1	−2.44	58.95	1.24	−2.60	187.67	0.83	−2.87	285.22	0.54
Asp-χ_1	−2.34	56.44	1.05	−2.70	186.38	0.85	−2.83	284.31	0.65
Cys-χ_1	−2.12	58.74	2.10	−2.68	179.30	1.46	−2.94	290.59	0.98
Gln-χ_1	−1.77	56.69	0.75	−2.43	182.25	0.54	−2.77	288.81	0.55
Glu-χ_1	−1.89	54.05	0.60	−2.48	180.15	0.53	−2.63	291.54	0.36
His-χ_1	−2.25	58.73	1.54	−2.76	181.58	1.27	−2.93	290.20	0.91
Ile-χ_1	−1.88	62.39	0.58	−1.85	187.47	0.81	−2.99	291.26	1.04
Leu-χ_1	−1.14	63.12	1.45	−2.58	184.80	0.71	−2.91	280.20	0.51
Lys-χ_1	−1.94	57.26	0.64	−2.51	178.70	0.44	−2.85	288.81	0.42
Lys-χ_2	−1.86	58.77	0.58	−2.83	176.38	0.21	−1.91	291.75	0.67
Lys-χ_3	−1.80	63.04	0.63	−2.86	175.22	0.34	−1.88	288.81	0.58
Met-χ_1	−1.78	57.33	1.10	−2.31	184.06	0.60	−2.88	286.14	0.97
Met-χ_2	−2.79	71.23	0.44	−2.38	176.71	0.69	−2.83	282.30	0.34
Phe-χ_1	−2.51	59.62	2.38	−2.90	179.38	1.34	−2.85	289.38	0.70
Ser-χ_1	−2.88	59.79	0.55	−2.53	177.84	0.52	−2.66	289.41	0.50
Thr-χ_1	−2.67	56.95	0.68	−1.77	182.41	0.66	−2.69	295.55	0.76
Trp-χ_1	−2.37	59.64	1.96	−2.91	175.69	2.01	−2.96	288.12	1.15
Tyr-χ_1	−2.19	61.57	1.14	−2.84	177.36	1.22	−3.12	288.94	1.26
Val-χ_1	−2.04	83.41	0.05	−2.19	172.02	2.24	−1.30	292.04	1.94
ω	−1.22	0.00	4.43	−2.86	180.00	1.85	0	0	0

APPENDIX *A*
Combined Rietveld and stereochemical restraint refinement

Although a Rietveld refinement will extract the maximum information available in a powder pattern, this is not sufficient for refinement of a protein crystal structure, so we have included all available prior stereochemical information (Sussman, 1984; Dodson, 1998) as a suite of restraints. The minimization function is an extension of that previously reported (Von Dreele, 1999)

$$\begin{aligned} M_r = {} & f_I \sum_{i=1}^{N_I} w_i (I_{oi} - I_{ci})^2 + f_a \sum_{i=1}^{N_a} w_i (\alpha_{oi} - \alpha_{ci})^2 \\ & + f_d \sum_{i=1}^{N_d} w_i (d_{oi} - d_{ci})^2 + f_t \sum_{i=1}^{N_t} w_i (-T_{ci}^2)^2 \\ & + f_p \sum_{i=1}^{N_p} w_i (-p_{ci})^2 + f_v \sum_{i=1}^{N_v} w_i (v_{oi} - v_{ci})^4 \\ & + f_h \sum_{i=1}^{N_h} w_i (h_{oi} - h_{ci})^2 + f_x \sum_{i=1}^{N_x} w_i (x_{oi} - x_{ci})^2 \\ & + f_R \sum_{i=1}^{N_R} w_i (-R_{ci}^2)^2, \end{aligned}$$

where the first term is the traditional Rietveld refinement term with I_{oi} as the observed powder profile intensity, I_{ci} is that calculated from the model and w_i is the weight as determined from counting statistics in I_{oi}. The subsequent terms are stereochemical restraints for bond angles, interatomic distances, torsion-angle pseudopotentials, atom displacements from planar groups (*e.g.* phenyl), van der Waals contacts (*i.e.* 'bump' restraints), hydrogen bonds, chiral volumes and coupled torsion-angle pseudopotentials, respectively. The observed value for each restraint was developed from published values [*e.g.* Engh & Huber (1991) for bond lengths and angles]. For each individual term, a weight w_i is chosen from the distribution of these values used to develop each restraint. For example, the observed bond-length term used for a Gly $C^\alpha-C$ bond has d_{oi} = 1.516 (10) Å and $w_i = 1/\sigma_i^2$ = 100. Each class of terms in the minimization function has an associated weight factor *f*, which is used to adjust its contribution. All refinements in this work used the weight factors in Table 4; the f_R for χ_1/χ_2 pseudopotential terms was 0.20 and that for φ/ψ terms was 0.50. These values were established in a somewhat *ad hoc* fashion by noting the distribution of the relevant stereochemical parameters as measured by *PROCHECK* after refinement of a protein structure. If a distribution is judged to be too 'tight' then the associated weight factor was reduced or *vice versa*; the refinement was then repeated. Lebedev *et al.* (2003) has suggested a similar but more mathematical basis to this process by estimating these weights based on a marginal likelihood; this has yet to be applied to combined Rietveld and stereochemical refinements. In the present case, the weight factors are scaled to that of the powder diffraction data as it is, from the nature of the experiment, based firmly on counting statistics. This differs

Table 6
Coefficients for coupled torsion pseudopotentials.

Each has three to seven terms depending on the complexity of the distributions; 'MC-φ/ψ' is the pseudopotential terms for the main chain excluding Gly and Pro.

Torsion pair	A_i	δ_i or χ_1	ψ_i or χ_2	B_i	C_i	D_i
MC-φ/ψ	−3.23	−93.55	−25.72	0.14	0.21	0.06
	−3.53	−114.97	144.46	0.13	0.15	0.02
	−2.59	56.79	48.61	0.86	0.31	0.13
	−1.85	54.16	−151.75	1.97	1.28	0.08
Pro-φ/ψ	−3.40	−70.21	142.33	3.05	1.64	1.13
	−3.26	−66.39	−30.88	4.46	2.52	2.36
	−1.26	−82.10	63.68	18.50	0.90	−0.64
Gly-φ/ψ	−1.42	−173.72	−187.36	0.26	0.94	0.22
	−1.58	76.29	−173.82	2.55	0.89	1.21
	−2.36	−73.67	−29.92	3.05	1.75	1.79
	−2.88	79.77	4.48	1.98	1.11	0.97
	−1.50	−84.89	−190.30	3.00	1.03	1.25
Asn-χ_1/χ_2	−1.88	60.40	2.26	5.58	0.31	0.52
	−1.23	189.07	27.19	1.94	0.50	0.02
	−2.49	288.82	128.27	2.67	0.38	0.23
	−3.14	286.35	314.71	2.83	0.42	0.29
	−1.19	181.94	279.76	5.67	0.04	0.03
Asp-χ_1/χ_2	−1.89	58.96	−13.62	5.07	0.27	−0.04
	−2.12	183.96	5.41	2.45	0.21	0.14
	−3.08	286.69	328.58	2.88	0.56	0.58
Gln-χ_1/χ_2	−1.76	60.07	177.53	2.58	2.74	0.34
	−2.04	176.63	62.12	2.91	3.90	0.48
	−2.51	179.19	175.40	2.62	1.92	−0.02
	−3.05	288.53	175.50	2.07	1.70	−0.16
	−2.30	294.45	289.02	3.36	2.11	0.67
	−0.76	290.09	74.90	3.76	1.73	0.23
Glu-χ_1/χ_2	−1.27	58.44	177.21	2.24	1.36	0.05
	−1.14	180.79	60.46	1.84	1.97	−0.24
	−2.44	180.02	174.34	1.34	1.42	−0.28
	−1.66	294.53	73.70	2.39	2.78	−0.16
	−3.02	288.23	175.33	1.95	1.34	−0.21
	−2.36	288.71	289.98	2.55	2.02	0.86
	−1.22	59.36	272.26	7.46	7.42	−6.00
His-χ_1/χ_2	−1.25	60.06	89.67	13.26	1.60	−2.24
	−1.43	61.01	273.83	2.71	2.16	0.91
	−2.72	176.93	65.24	4.53	1.94	0.37
	−2.27	182.55	269.53	4.50	2.42	−1.58
	−2.19	291.84	92.21	3.53	0.91	−0.40
	−2.04	285.66	168.59	9.08	1.70	−0.58
	−3.02	291.32	279.04	3.02	0.97	−0.31
Ile-χ_1/χ_2	−2.45	58.75	166.08	5.01	2.61	−0.08
	−1.88	184.24	160.64	3.73	3.68	−1.25
	−3.05	291.91	164.28	2.93	1.08	−0.32
	−2.17	297.51	296.37	3.95	2.22	0.56
	−1.20	190.40	63.49	15.58	3.30	0.18
Leu-χ_1/χ_2	−2.73	178.53	60.99	1.85	2.28	0.72
	−3.25	290.21	172.12	1.66	1.63	−0.01
	−1.44	258.79	30.13	2.35	0.72	−0.66
	−1.27	186.43	153.69	5.11	2.75	−1.13
	−1.10	210.74	210.68	3.23	3.08	2.01
Met-χ_1/χ_2	−2.91	292.36	292.22	4.37	2.50	0.01
	−2.27	179.22	175.34	2.78	2.44	−0.10
	−1.28	63.01	178.18	4.60	2.35	−1.11
	−1.69	179.66	66.57	10.55	5.31	−4.12
	−2.98	287.15	175.30	2.73	1.92	−0.37
	−0.34	179.90	271.63	3.47	4.12	1.52
Trp-χ_1/χ_2	−1.47	52.67	85.71	4.33	9.46	−2.14
	−2.67	60.77	265.57	8.00	11.86	6.43
	−2.60	175.89	79.17	12.15	2.84	−3.15
	−1.94	174.99	251.61	3.59	2.46	−0.96
	−2.81	289.17	96.32	3.62	1.43	−0.30
	−1.78	283.28	264.33	11.77	3.74	−1.77
	−1.41	288.72	348.83	6.30	1.84	1.14

from Lebedev *et al.* (2003), who argue that the interatomic distance and angle restraints should be the basis for weighting choices. In fact, we find the weighting factors for distances and angles to be quite different from unity and each other.

Two types of pseudopotentials are used for some of the restraints in our combined refinements; these avoid the problem of constructing a restraint to an uncertain target noted by Dodson (1998). Torsion angles (*e.g.* side-chain χ_1) are restrained by a monovariant Gaussian distribution with typically three terms

$$T_c = A_0 + \sum_{i=1}^{3} A_i \exp[-B_i(\varphi_{oi} - \varphi_c)^2]$$

that reflect the three possible values for these torsions. We have determined the coefficients A_i, φ_i and B_i by fitting to the log-normalized distributions of torsion angles obtained from high-resolution protein structures and used by *PROCHECK* (Laskowski *et al.*, 1993) for stereochemical evaluation; the A_0 bias is set so that the function has a minimum of zero. These are tabulated in Table 5 for various torsion angles. In the refinements reported here each torsion restraint was assigned a weight of $w_i = 4$.

Coupled torsion angles (φ/ψ or χ_1/χ_2) are restrained by a bivariant Gaussian distribution with three to seven terms; these reflect stereochemically allowed and forbidden values for the two torsions. For example, this restraint for φ/ψ torsion angles drives the refinement toward the most favored regions of the Ramachandran plot (Fig. 10; Table 6). The function is

$$R_c = A_0 + \sum_{i=1}^{n} A_i \exp[-B_i(\varphi_{oi} - \varphi_c)^2 - C_i(\varphi_{oi} - \varphi_c)^2 - D_i(\varphi_{oi-\varphi_c})(\varphi_{oi} - \varphi_c)].$$

The six coefficients for each term have been determined by fitting to a log-normalized distribution of the paired torsions found in a suite of high-resolution protein structures and used by *PROCHECK* (Laskowski *et al.*, 1993) for stereochemical evaluation; the A_0 bias is set so that the function has a minimum of zero. In the refinements reported here each paired torsion restraint was assigned a weight of $w_i = 100$. This is an extension of our earlier description (Von Dreele, 1999) of this type of restraint that was based on the scoring originally

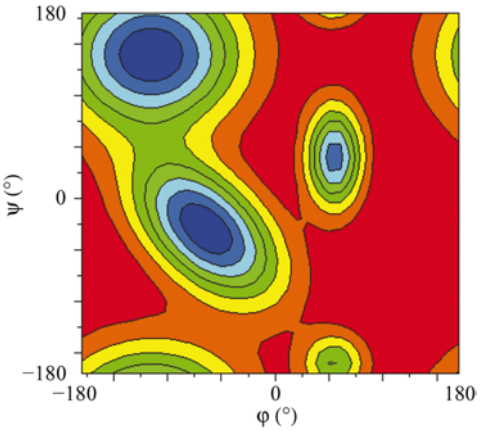

Figure 10
Ramachandran φ/ψ pseudopotential restraint surface; the most favored regions are blue and forbidden regions are shown in red as calculated from the MC − φ/ψ terms given in Table 6.

given by Ramachandran *et al.* (1963) for the regions of this plot.

The author thanks P. Stephens and A. Light for assistance in collecting the powder diffraction data. He also thanks the referees and journal co-editor for their efforts during the preparation of this manuscript. The US DOE/OS/BES under contracts W-7405-ENG-36 and W-31-109-ENG-38 supported this work and the SUNY X3 beamline was supported by US DOE Grant No. DE-FG02-86ER45231.

References

Altomare, A., Casarano, G., Giacovazzo, C., Guagliardi, A. & Moliterna, G. G. (1995). *J. Appl. Cryst.* **28**, 738–744.
Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1967). *Proc. R. Soc. London Ser. B*, **167**, 378–388.
Cheetham, J. C., Artymiuk, P. J. & Phillips, D. C. (1992). *J. Mol. Biol.* **224**, 613–628.
Cheetham, A. K. & Taylor, J. C. (1977). *J. Solid State Chem.* **21**, 253–275.
Colovos, C. & Yeates, T. O. (1993). *Protein Sci.* **2**, 1511–1519.
DeLano, W. L. (2004). *The PyMOL Molecular Graphics System.* DeLano Scientific LLC, San Carlos, CA, USA. http://www.pymol.org/.
Dinnebier, R. E., Von Dreele, R. B., Stephens, P. W., Jelonek, S. & Sieler, J. (1999). *J. Appl. Cryst.* **32**, 761–769.
Dodson, E. (1998). *Acta Cryst.* D**54**, 1109–1118.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Fong, K., Jefferson, T., Suyehiro, T. & Walton L. (1993). *Guide to the SLATEC Common Mathematical Library*. http://www.netlib.org/slatec.
Guex, N. & Peitsch, M. C. (1999). *SwissPDBViewer.* Glaxo Wellcome Experimental Research. http://www.expasy.ch/spdbv.
Harris, K. D. M. & Tremayne, M. (1996). *Chem. Mater.* **8**, 2554–2570.
Hill, R. J. & Madsen, I. C. (1986). *J. Appl. Cryst.* **19**, 10–18.
James, R. W. (1965). *The Optical Principles of The Diffraction of X-rays*, pp. 517–518. Ithaca: Cornell University Press.
Larson, A. C. & Von Dreele, R. B. (2004). *General Structure Analysis System (GSAS)*, Los Alamos National Laboratory Report LAUR 86–748. http://www.ccp14.ac.uk/ccp/ccp14/ftp-mirror/gsas/public/gsas.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lebedev, A. A., Tickle, I. J., Laskowski, R. A. & Moss, D. S. (2003). *Acta Cryst.* D**59**, 1557–1566.
McCusker, L. B., Von Dreele, R. B., Cox, D. E., Löuer, D. & Scardi, P. (1999). *J. Appl. Cryst.* **32**, 36–50.
Madsen, I. C. & Hill, R. J. (1994). *J. Appl. Cryst.* **27**, 385–392.
Matthews, B. W. (1976). *J. Mol. Biol.* **33**, 491–497.
Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992). *Proteins*, **12**, 345–364.
Natchiar, S. K., Jeyaprakash, A. A., Ramya, T. N. C., Thomas, C. J., Suguna, K., Surolia, A. & Vijayan, M. (2004). *Acta Cryst.* D**60**, 211–219.
Pagola, S., Stephens, P. W., Bohle, D. S., Kosar, A. D. & Madsen, S. K. (2000). *Nature (London)*, **404**, 307–310.
Perkins, S. J., Johnson, L. N., Machin, P. A. & Phillips, D. C. (1978). *Biochem. J.* **173**, 607–616.
Poojary, D. M. & Clearfield, A. (1997). *Acc. Chem. Res.* **30**, 414–422.
Post, J. E. & Bish, D. L. (1989). *Modern Powder Diffraction. Reviews in Mineralogy*, Vol. 20, edited by D. L. Bish & J. E. Post, pp. 278–308. Washington: Mineralogical Society of America.
Prince, E. (1981). *J. Appl. Cryst.* **14**, 157–159.
Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). *J. Mol. Biol.* **7**, 95–99.
Rietveld, H. M. (1969). *J. Appl. Cryst.* **2**, 65–71.
Rupley, J. A. & Gates, V. (1967). *Proc. Natl Acad. Sci. USA*, **57**, 496–510.
Sussman, J. L. (1984). *Methods and Applications in Crystallographic Computing*, edited by S. R. Hall & T. Ashida, pp. 206–237. Oxford University Press.
Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.
Von Dreele, R. B. (1989). *Modern Powder Diffraction. Reviews in Mineralogy*, Vol. 20, edited by D. L. Bish & J. E. Post, pp. 334–369. Washington: Mineralogical Society of America.
Von Dreele, R. B. (1999). *J. Appl. Cryst.* **32**, 1084–1089.
Von Dreele, R. B. (2001). *Acta Cryst.* D**57**, 1836–1842.
Von Dreele, R. B. (2003). *Methods Enzymol.* **368**C, 254–267.
Von Dreele, R. B., Stephens, P. W., Blessing, R. H. & Smith, G. W. (2000). *Acta Cryst.* D**56**, 1549–1553.
Warren, B. E. (1990). *X-Ray Diffraction*, pp. 47–49. New York: Dover.
Yeates, T. O. (2004). Private communication. http://www.doe-mbi.ucla.edu/Services/ERRATv2.