Crystallization of a fungal lytic polysaccharide monooxygenase expressed from glycoengineered *Pichia pastoris* for X-ray and neutron diffraction

William B. O'Dell,^{a,b} Paul D. Swartz,^a Kevin L. Weiss^b and Flora Meilleur^{a,b*}

Received 24 October 2016

Accepted 21 December 2016

Edited by J. Newman, Bio21 Collaborative Crystallisation Centre, Australia

Keywords: lytic polysaccharide monooxygenases; *Pichia pastoris*; glycoengineering; neutron protein crystallography.

^aDepartment of Molecular and Structural Biochemistry, North Carolina State University, Campus Box 7622, Raleigh, NC 27695, USA, and ^bBiology and Soft Matter Division, Oak Ridge National Laboratory, PO Box 2008, Oak Ridge, TN 37831, USA. *Correspondence e-mail: fmeille@ncsu.edu

Lytic polysaccharide monooxygenases (LPMOs) are carbohydrate-disrupting enzymes secreted by bacteria and fungi that break glycosidic bonds *via* an oxidative mechanism. Fungal LPMOs typically act on cellulose and can enhance the efficiency of cellulose-hydrolyzing enzymes that release soluble sugars for bioethanol production or other industrial uses. The enzyme PMO-2 from *Neurospora crassa* (NcPMO-2) was heterologously expressed in *Pichia pastoris* to facilitate crystallographic studies of the fungal LPMO mechanism. Diffraction resolution and crystal morphology were improved by expressing NcPMO-2 from a glycoengineered strain of *P. pastoris* and by the use of crystal seeding methods, respectively. These improvements resulted in high-resolution (1.20 Å) X-ray diffraction data collection at 100 K and the production of a large NcPMO-2 crystal suitable for room-temperature neutron diffraction data collection to 2.12 Å resolution.

1. Introduction

Lytic polysaccharide monooxygenases (LPMOs) are enzymes that disrupt polysaccharide substrates, including cellulose, chitin and starch, by oxidizing C atoms involved in glycosidic bonds (Vu *et al.*, 2014; Beeson *et al.*, 2015; Hemsworth *et al.*, 2015). Molecular oxygen is activated for insertion into a substrate C–H bond at a mononuclear copper(I)/copper(II) active site contained within an extended planar surface that mediates substrate binding. LPMOs have attracted keen interest in the context of bioethanol production, as cellulose-active LPMOs can increase the efficiency of the conversion of cellulose to soluble sugars by conventional hydrolases (Harris *et al.*, 2014; Johansen, 2016).

X-ray crystallographic studies of both fungal and bacterial LPMOs have revealed that the active-site copper(II) is ligated by a 'histidine brace' formed by the conserved N-terminal histidine and a second conserved histidine. The N-terminal histidine provides a bidentate ligand through the amino-terminus and an imidazole ring N atom, while the second histidine provides a monodentate ligand through an imidazole ring N atom, as shown in Fig. 1. In fungal LPMOs the hydroxy group of a tyrosine residue occupies one axial coordination site; however, this site is unoccupied in bacterial LPMOs (Quinlan *et al.*, 2011; Hemsworth *et al.*, 2013). Various chemical species, including water, chloride ion and superoxide anion, have been observed as equatorial and axial ligands to complete the copper coordination sphere (Li *et al.*, 2012; Gudmundsson *et al.*, 2014; Frandsen *et al.*, 2016). Electron-donating proteins or small molecules can potentiate single-electron reduction of copper(II) to copper(I), which has a high

affinity for molecular oxygen. Upon oxygen binding, the copper ion is oxidized to copper(II) by the formation of a copper-superoxide complex (Beeson *et al.*, 2015). Despite multiple crystallographic, spectroscopic and computational studies, the identity of the activated oxygen species ultimately responsible for H-atom abstraction from the substrate glycosidic C atom remains unknown. Furthermore, all subsequent mechanistic steps still require structural description (Li *et al.*, 2012; Gudmundsson *et al.*, 2014; Kim *et al.*, 2014; Kjaergaard *et al.*,

et al., 2014; Frandsen *et al.*, 2016; Courtade *et al.*, 2016). Additional crystallographic studies of LPMO reactive intermediates are necessary, and direct interrogation of protonation states and H-atom coordinates *via* neutron protein crystallography may ultimately reveal key details of the hydrogen-abstraction reaction.

The structure of the extracellular LPMO enzyme polysaccharide monooxygenase 2 from *Neurospora crassa* (NcPMO-2; UniProt Q1K8B6) has previously been determined by X-ray crystallography of the protein as expressed by the source organism. Crystals of NcPMO-2 purified from the source diffracted to high resolution ($d_{\min} = 1.10 \text{ \AA}$) despite the heterogeneity of the glycan conjugated to the single N-linked site (Asn60; Li *et al.*, 2012). In order to continue investigating the LPMO mechanism using both X-ray and neutron crystallography, we chose to produce NcPMO-2 by heterologous expression from the yeast *Pichia pastoris*. *P. pastoris* is a widely used expression host that tolerates high-cell-density cultivation and typically produces high yields of target proteins (Mattanovich *et al.*, 2012). As a eukaryotic host, *P. pastoris* facilitates targeting pre-proteins for secretion, maturation by signal cleavage and conjugation of N-linked glycans, which are all necessary for the expression of soluble and properly folded NcPMO-2. In addition, *P. pastoris* has been reported to be a suitable host for expressing deuterium (^2H)-labeled proteins, which can be beneficial to neutron protein crystallography studies (Morgan *et al.*, 2000; Pickford & O'Leary, 2004; Sugiki *et al.*, 2012; O'Dell *et al.*, 2016).

This communication reports our successful crystallization, high-resolution X-ray data collection and room-temperature neutron data collection of NcPMO-2 heterologously expressed from *P. pastoris*. This success was enabled by protein expression from the glycoengineered strain *P. pastoris* SuperMan₅ that predominantly produces proteins bearing N-linked glycans with the composition (mannose)₅(N-acetylglucosamine)₂ [(Man)₅(GlcNAc)₂] instead of the high-mannose glycans typically observed in yeast expression products (Vervecken *et al.*, 2004; Jacobs *et al.*, 2009). The combined change of expression host and the use of crystal seeding techniques allowed us to crystallize in a higher symmetry space group, obtain crystals that diffract to higher resolution and improve the crystal morphology compared with our initial attempts.

2. Materials and methods

2.1. Macromolecule production

NcPMO-2 was heterologously expressed in the host *P. pastoris* and purified generally following the methods of Kittl *et al.* (2012). The amino-acid sequence, including the native N-terminal secretion signal peptide, was used as the basis for DNA synthesis of the open reading frame (ORF) with codon optimization for *Pichia* expression (GeneArt, Thermo Fisher Scientific) as shown in Table 1. BstBI and XbaI restriction sites introduced during synthesis were used for subcloning into the pPICZαA expression vector (Thermo

Figure 1
Structure of NcPMO-2 and the 'histidine brace'. (a) Ribbon diagram of NcPMO-2 based on PDB entry 5tkf. Noncrystallographic symmetry (NCS) molecule A is shown. (b) The LPMO 'histidine brace' coordination consists of three nitrogen ligands, here donated by His1 and His84. In fungal LPMOs an additional copper coordination site is occupied by the hydroxy group of a conserved tyrosine residue (Tyr168).

Table 1
Macromolecule-production information.

Source organism	<i>N. crassa</i> OR74A
DNA source	Synthesis
Expression vector	pPICZαA
Expression host	<i>P. pastoris</i> X-33 and <i>P. pastoris</i> SuperMan ₅
Complete amino-acid sequence of the construct produced	HTIFSSLEVNGVNQGLGEGVRVPTYNPIEDVTS- ASIANCGSPNTVASTSKVITVQAGTNVTAIWR- YMLSTTGDSPADVMDSSHKGPTIAYLKKVDNA- ATASGVGNGWFKIQDQMDSSGVWGTERTING- KGRHSIKIPECIAPGGYLLRAEMIALHAASNY- PGAQFYMECAQLNVVGGTGAKTPSTVSFPFGAY- SGSDPGVKISIWPPVTSYTVPGPSVFTC

Fisher Scientific). (Using the BstBI site removes the *Saccharomyces cerevisiae* α-mating factor secretion signal peptide from the vector.) The resulting plasmid was amplified in *Escherichia coli* DH5α using standard techniques and the bleomycin resistance marker. Purified plasmid was linearized by treatment with PmeI, and the linear product was transformed by electroporation into either the *P. pastoris* X-33 wild type or SuperMan₅(HIS⁺) strains (Thermo Fisher Scientific and BioGrammatics, respectively) according to the equipment manufacturer's protocol (Bio-Rad). Transformants were selected for bleomycin resistance using media containing 500 µg ml⁻¹ Zeocin (Thermo Fisher Scientific). Positive colonies were screened for methanol-utilization phenotype and inducible NcPMO-2 expression.

NcPMO-2 expression from the SuperMan₅ strain was performed using a fed-batch approach in a bioreactor system with a 3 l starting volume. The bioreactor vessel was charged with 2700 ml fermentation basal salts medium [26.7 ml l⁻¹ 85% (w/v) H₃PO₄, 0.93 g l⁻¹ CaSO₄·2H₂O, 18.2 g l⁻¹ K₂SO₄, 14.9 g l⁻¹ MgSO₄·7H₂O, 4.13 g l⁻¹ KOH] and steam-sterilized (Stratton *et al.*, 1998). To the cooled medium was added 33 ml of fermentation trace metals solution [6.0 g l⁻¹ CuSO₄·5H₂O, 0.08 g l⁻¹ KI, 3.0 g l⁻¹ MnSO₄·H₂O, 0.2 g l⁻¹ Na₂MoO₄·2H₂O, 0.05 g l⁻¹ CoCl₂·6H₂O, 0.02 g l⁻¹ boric acid, 0.5 g l⁻¹ CaSO₄·2H₂O, 20 g l⁻¹ ZnCl₂, 65 g l⁻¹ FeSO₄·7H₂O, 0.4 g l⁻¹ biotin, 19.2 ml l⁻¹ H₂SO₄ (concentrated); Amresco], 240 ml of 50% (w/v) glycerol containing 12 ml l⁻¹ fermentation trace metals, 300 mg Zeocin and 5 ml Antifoam 204 (Sigma-Aldrich; Tolner *et al.*, 2006). Concentrated NH₄OH was added until the medium equilibrated at pH 5.2. The bioreactor was inoculated at an initial optical density OD₆₀₀ of 4.3 from an overnight culture in buffered minimal glycerol medium. The controller maintained the culture temperature at 30°C, a dissolved oxygen concentration of >30% using agitation and aeration with compressed air, and a pH > 5.0 by the intermittent addition of 6.5% (w/v) NH₄OH. The culture progressed from batch phase upon the depletion of initial glycerol concentration to fed-batch when feeding with 50% (w/v) glycerol containing trace metals was automatically initiated by a spike in the dissolved oxygen concentration. Glycerol feeding continued at a rate of 18 ml h⁻¹ l⁻¹ until the optical density of the culture reached OD₆₀₀ = 110.0. The vessel temperature was then lowered to 25°C, and NcPMO-2 production was induced by the addition of methanol (with 12 ml l⁻¹ fermentation trace metals added) at an initial rate of

3.7 ml h⁻¹ l⁻¹. During the 48 h induction period, the methanol feeding rate was gradually increased stepwise to 18 ml h⁻¹ l⁻¹. NcPMO-2 expression from *Pichia* X-33 followed similar procedures, with the reagent amounts adjusted to a starting culture volume of 300 ml. Secreted NcPMO-2 was harvested by centrifuging the culture volume at 1500g and 4°C for 10 min. The supernatant was collected and cleared further by centrifugation at 15 000g and 4°C for 30 min. The supernatant from the second centrifugation was then filtered through 0.22 µm pore-size polyethersulfone (PES) membranes and stored at 4°C until purification.

NcPMO-2 was purified by sequential hydrophobic interaction, anion-exchange and size-exclusion chromatography. Filtered supernatant was warmed to 20°C prior to slow addition of 176 g l⁻¹ ammonium sulfate with stirring. This sample was applied onto a Phenyl Sepharose FF high-substitution column (GE Healthcare Life Sciences, 150 ml column volume) equilibrated with 20 mM sodium acetate pH 5.0 containing 176 g l⁻¹ ammonium sulfate. NcPMO-2 eluted at the end of a 7.5 column-volume linear gradient to 100% 20 mM sodium acetate pH 5.0. Fractions containing NcPMO-2 were pooled, concentrated using 5 kDa molecular-weight cutoff (MWCO) PES centrifugal concentrators and then diafiltered for buffer exchange into 20 mM Tris-HCl pH 8.0. The sample was applied onto a Mono Q column (GE Healthcare Life Sciences, 8 ml column volume) equilibrated with the same buffer. NcPMO-2 eluted in the sample-loading flowthrough, while multiple contaminant proteins bound to the anion-exchange medium and were eluted with 20 mM Tris-HCl pH 8.0, 250 mM NaCl. The flowthrough was concentrated as before and loaded onto a Superdex 75 column (GE Healthcare Life Sciences, 320 ml column volume) equilibrated with 20 mM sodium acetate pH 5.0. Fractions containing NcPMO-2 were collected and concentrated as before. The final yield of purified NcPMO-2 from the SuperMan₅ strain was approximately 50 mg protein per litre of starting culture volume.

2.2. Mass spectrometry

Intact mass analysis of purified NcPMO-2 expressed from both the X-33 and SuperMan₅ strains was performed by MS Bioworks LLC. Samples of NcPMO-2 in 20 mM sodium acetate pH 5.0 were diluted with 0.1% (v/v) formic acid and injected onto a Symmetry C4 column (Waters) equilibrated with 0.1% (v/v) trifluoroacetic acid (TFA). NcPMO-2 eluted during a linear gradient to 100% acetonitrile, 0.1% (v/v) TFA and was analysed using a Q Exactive mass spectrometer (Thermo Scientific) operating with positive-mode electrospray ionization. Data were acquired from 600 to 2500 *m/z* at resolution 17 500 FWHM (400 *m/z*) with ten microscans per spectrum. Mass spectra (Fig. 2) were visualized with XCalibur (Thermo Scientific) and were deconvoluted using MagTran 3.0 (Zhang & Marshall, 1998).

2.3. Crystallization

The crystallization conditions were adapted from those reported by Li *et al.* (2012). NcPMO-2 was crystallized by the

Figure 2

Analysis of NcPMO-2 glycoforms by mass spectrometry. (a, b) Electrospray mass spectrum and deconvolution obtained for NcPMO-2 expressed from *P. pastoris* X-33 showing the presence of (Man)₉(GlcNAc)₂, (Man)₁₀(GlcNAc)₂ and (Man)₁₁(GlcNAc)₂ glycoforms (as water adducts) along with an unidentified species (marked with a star). (c, d) The spectrum and deconvolution for NcPMO-2 expressed from *P. pastoris* SuperMan₅ showing the presence of (Man)₅(GlcNAc)₂, (Man)₇(GlcNAc)₂ and (Man)₁₀(GlcNAc)₂ glycoforms.

vapor-diffusion technique using similar conditions (Table 2) for protein expressed from both the X-33 and SuperMan₅ strains of *P. pastoris*. Hanging drops of 2 µl total volume were set up on siliconized cover slips by mixing equal amounts of protein and reservoir solution (Table 2). Plate-like crystals appeared within a week (Fig. 3). The crystal morphology could be improved by microseeding 600 µl sitting drops composed of 200 µl protein solution and 400 µl 25%(w/v) polyethylene glycol (PEG) 3350, 100 mM HEPES pH 6.0 and equilibrating for 1–2 weeks against a 22%(w/v) PEG 3350, 100 mM HEPES pH 6.0 reservoir. Microcrystals were prepared using PTFE Seed Beads (Hampton Research) and vortexing, and volumes ranging from 1 to 10 µl were added to the sitting drops. (The pH of the drops was not measured before or after equilibration.) Crystals for cryogenic X-ray data collection were harvested and cryoprotected with either reservoir solution

Figure 3

Crystals and X-ray diffraction of NcPMO-2 expressed from *P. pastoris* X-33. (a) Crystals grew as thin plates that often overlaid each other, forming stacks. (b) Representative synchrotron X-ray diffraction shows a highest resolution of ~2.10 Å for these crystals.

Table 2
Crystallization conditions.

PDB code	5tkf	5tkg	5tki
Expression host	<i>P. pastoris</i> X-33	<i>P. pastoris</i> SuperMan ₅	<i>P. pastoris</i> SuperMan ₅
Method	Hanging drop	Sitting drop, microseeding	Sitting drop, microseeding, macroseeding
Plate type	VDX	9-well glass plate	9-well glass plate
Temperature (K)	291	291	291
Protein concentration (mg ml ⁻¹)	12.6	12.5	12.2
Buffer composition of protein solution	20 mM sodium acetate pH 5.0	20 mM sodium acetate pH 5.0	20 mM sodium acetate pH 5.0
Composition of reservoir solution	22%(w/v) PEG 3350, 100 mM HEPES pH 6.4	22%(w/v) PEG 3350, 100 mM HEPES pH 6.0	22%(w/v) PEG 3350, 90 mM HEPES pH 6.0
Volume and ratio of drop	2.0 µl, 1:1 protein:reservoir	600 µl, 1:2 protein:25%(w/v) PEG 3350, 100 mM HEPES pH 6.0	750 µl, 1:2 protein:25%(w/v) PEG 3350, 100 mM HEPES pH 6.0
Volume of reservoir (ml)	1.0	50.0	50.0

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

PDB	5tkf	5tkg	5tki (X-ray)	5tki (neutron)
Diffraction source	22-BM, APS	22-ID, APS	MicroMax-007 HF, Rigaku	CG-4D IMAGINE, HFIR
Wavelength (Å)	1.00	1.00	1.54	2.8–4.5
Temperature (K)	100	100	Ambient	Ambient
Detector	Rayonix MAR 225	Rayonix 300HS	Rigaku R-Axis IV ⁺⁺	ARINAX image plate
Crystal-to-detector distance (mm)	200	110	90	199
Rotation range per image (°)	1	1	1	0†
Total rotation range (°)	180	180	180	120, 90, 50†
No. of exposures	180	180	180	26
Exposure time per image (s)	1	1	120	72000
Space group	<i>P</i> ₁	<i>P</i> ₂ ₁	<i>P</i> ₂ ₁	<i>P</i> ₂ ₁
<i>a</i> , <i>b</i> , <i>c</i> (Å)	43.30, 67.04, 84.15	67.35, 42.21, 69.48	68.12, 42.23, 70.29	68.12, 42.23, 70.29
α , β , γ (°)	97.57, 97.61, 97.43	90.00, 98.96, 90.00	90.00, 98.33, 90.00	90.00, 98.33, 90.00
Resolution range (Å)	35.45–2.10 (2.18–2.10)	31.82–1.20 (1.24–1.20)	36.10–1.50 (1.55–1.50)	32.87–2.12 (2.19–2.12)
Total No. of reflections	102814 (10204)	228097 (21939)	224095 (20772)	31698 (2251)
No. of unique reflections	52131 (5173)	117125 (11360)	58482 (5409)	18141 (1381)
Completeness (%)	97.0 (95.9)	97.0 (94.6)	90.2 (80.3)	79.0 (61.0)
Multiplicity	2.0 (2.0)	1.9 (1.9)	3.8 (3.8)	1.7 (1.6)
$\langle I/\sigma(I) \rangle$	8.8 (3.5)	7.2 (4.2)	13.8 (4.2)	4.9 (3.6)
<i>R</i> _{meas}	0.1232 (0.4136)	0.1034 (0.2138)	0.0633 (0.3963)	0.1806 (0.2691)
Overall <i>B</i> factor from Wilson plot (Å ²)	16.04	8.77	16.04	8.08

† Still quasi-Laue diffraction images were collected with 10° steps in ϕ rotation at three unique crystal orientations that were combined into a single data set. The angular ranges reported are those covered at each crystal orientation.

with 25%(v/v) glycerol added or with low-viscosity cryo-oil (MiTeGen) prior to flash-cooling in liquid nitrogen.

A crystal obtained by microseeding was used as a macroseed to obtain the crystal used for room-temperature X-ray and neutron data collection. The macroseed was transferred twice through 22%(w/v) PEG 3350, 100 mM HEPES pH 6.0 and placed in an equilibrated 750 µl sitting drop. The crystal reached its final volume within ten weeks of growth. Once grown, the crystal was mounted in a quartz capillary containing artificial mother liquor formulated in deuterium oxide. Vapor exchange over 30 d promoted the replacement of solvent and titratable protein H atoms with D atoms to increase the neutron diffraction intensity and decrease the incoherent neutron scattering background.

2.4. Data collection and processing

X-ray cryocrystallography data were collected using the SER-CAT beamlines of the Advanced Photon Source (APS), Argonne National Laboratory. Room-temperature X-ray data

were collected using a copper rotating-anode home source. All X-ray data were indexed, integrated and scaled using the CCP4 suite (Winn *et al.*, 2011). Quasi-Laue neutron diffraction data were collected using the CG-4D IMAGINE beamline at the High Flux Isotope Reactor (HFIR), Oak Ridge National Laboratory (Meilleur *et al.*, 2013). A total of 26 still diffraction images were collected with 10° ϕ rotation steps using three unique crystal orientations to maximize reciprocal-space coverage in the combined diffraction data set. Quasi-Laue data were indexed and integrated using LAUEGEN prior to wavelength normalization and scaling with LSCALE and SCALA (Campbell *et al.*, 1998; Arzt *et al.*, 1999; Winn *et al.*, 2011). The data-set statistics shown in Table 3 were calculated from scaled but unmerged intensities using the program phenix.table_one in the PHENIX suite (Adams *et al.*, 2010). Data sets 5tkf and 5tkg represent the highest diffraction resolution observed from screening ten or more crystals grown under similar conditions. Refined models were deposited in the Protein Data Bank, and the high-resolution X-ray structure (PDB entry 5tkg) and the joint X-ray/neutron structure

(PDB entry 5tki) are described in detail by O'Dell *et al.* (2016).

2.5. Structure solution and refinement

Initial phases were determined by molecular replacement with *Phaser* as implemented in *PHENIX* using the previous NcPMO-2 structure (PDB entry 4eir) reported by Li *et al.* (2012), which has 99% sequence identity to heterologous NcPMO-2 (McCoy *et al.*, 2007; Adams *et al.*, 2010). Molecular replacement in the *P1* space group for PDB entry 5tkf placed four noncrystallographically symmetric (NCS) molecules of NcPMO-2 within the asymmetric unit with a translation-function Z-score (TFZ score) of 83.8. Molecular replacement

Table 4

Structure solution and refinement.

Values in parentheses are for the outer shell.

PDB code	5tkf	5tkg
Resolution range (Å)	35.45–2.10 (2.18–2.10)	28.73–1.20 (1.24–1.20)
No. of reflections, working set	52110 (5173)	117101 (11354)
No. of reflections, test set	1999 (199)	1997 (202)
Final R_{cryst}	0.1452 (0.1795)	0.1142 (0.1275)
Final R_{free}	0.1813 (0.2193)	0.1399 (0.1650)
Cruickshank DPI	0.142	0.030
No. of non-H atoms		
Protein	6667	3705
Ligand	183	69
Water	1008	921
Total	7858	4699
R.m.s. deviations		
Bonds (Å)	0.003	0.009
Angles (°)	0.73	1.18
Average B factors (Å ²)		
Protein	19.03	11.30
Ligand	42.48	23.56
Water	30.34	26.34
Ramachandran plot		
Most favoured (%)	96.0	97.0
Allowed (%)	3.5	2.9

in space group *P2*₁ for PDB entry 5tkg placed two NCS molecules of NcPMO-2 per asymmetric unit with a TFZ score of 135.3. *Phenix.autobuild* was used for automated structure rebuilding to minimize initial phase bias from the replacement model (Terwilliger *et al.*, 2008). Further structure refinement was performed using *phenix.refine* with manual model building performed in *Coot* (Afonine *et al.*, 2012; Emsley *et al.*, 2010). For model 5tkf, non-H protein atoms were refined in 20 translation–libration–screw (TLS) groups per protein molecule as parametrized by the *TLSMD* server (Painter & Merritt, 2006a,b). For model 5tkg, all non-H atoms were refined with anisotropic atomic displacement parameters. Final models were analyzed using the *Online_DPI* web server to determine the Cruickshank diffraction precision index

Figure 4

Asn60 N-linked glycan observed for NcPMO-2 expressed from *P. pastoris* X-33. (a) A glycan corresponding to $(\text{Man})_3(\text{GlcNAc})_2$ could be modeled into the electron density ($2F_o - F_c$ at $\sigma = 1.0$) adjacent to Asn60 of NCS molecule A. (b) The four NCS molecules in the 5tkf *P1* unit cell are shown as C^α traces as viewed from the *bc* crystal face. The observed glycans are shown as gray sticks, while protein and solvent atoms within 3.0 Å of the glycan are shown as spheres. The N-linked glycan of NCS molecule C (blue) appears to participate in crystal packing with NCS molecule A (red). (c) An enlarged view of the interface between NCS molecule C (blue) and NCS molecule A (orange) shows three possible hydrogen-bonding interactions between Ser48 and Ser50 of NCS molecule A with the glycan from NCS molecule C.

(DPI; Cruickshank, 1999; Kumar *et al.*, 2015). Refinement statistics for both models are shown in Table 4.

3. Results and discussion

Mass-spectrometric measurements of the protein expressed from *P. pastoris* X-33 (Figs. 2*a* and 2*b*) gave intact masses corresponding to water adducts of NcPMO-2 bearing three glycan chains: (Man)₉(GlcNAc)₂ (predicted 25 159.8 Da, observed 25 159.0 Da), (Man)₁₀(GlcNAc)₂ (predicted 25 321.9 Da, observed 25 321.0) and (Man)₁₁(GlcNAc)₂ (predicted 25 484.13, observed 25 482.0). Mass-spectrometric measurements of NcPMO-2 expressed from *P. pastoris* SuperMan₅ (Figs. 2*c* and 2*d*) gave the expected (Man)₅(GlcNAc)₂ species (predicted 24 993.3 Da, observed 24 492.0 Da) along with the presence of (Man)₇(GlcNAc)₂

(predicted 24 817.6 Da, observed 24 817.0 Da) and (Man)₁₀(GlcNAc)₂ (predicted 25 304.0 Da, observed 25 302.0 Da) species. As reported by Vervecken *et al.* (2004), a (Man)₅(GlcNAc)₂ glycan accounts for 90–95% of all N-linked glycosylation of target proteins expressed by this strain.

Initial crystallization attempts were conducted using NcPMO-2 heterologously expressed from the wild-type *P. pastoris* X-33 strain. Crystals could be obtained with a morphology of either thin or stacked plates, as shown in Fig. 3. Synchrotron X-ray diffraction at 100 K could only be observed to a highest resolution of 2.10 Å and revealed a triclinic space group (Table 3; PDB entry 5tkf). Examination of the electron-density maps revealed densities corresponding to N-linked glycans attached to Asn60 of each of the four protein molecules in the asymmetric unit. Different lengths of glycans ranging up to approximately seven sugar units could be modeled into the electron densities (Fig. 4*a*). Increases in the *B* factor of each subsequent sugar residue from the Asn60 linkage suggest that additional sugar residues could not be

Figure 5
Crystals and X-ray diffraction of NcPMO-2 expressed from *P. pastoris* SuperMan₅. (a) Spontaneously nucleating crystals grew as thin plates that typically formed stacks or clusters. (b) Microseeds prepared from spontaneous crystals and added to pre-equilibrated drops nucleated primarily single crystals with increased growth in all three dimensions. (c) Representative synchrotron X-ray diffraction shows reflections exceeding a resolution of 1.20 Å for the seeded crystals.

Figure 6
Asn60 N-linked glycan observed for NcPMO-2 expressed from *P. pastoris* SuperMan₅. (a) A glycan corresponding to (GlcNAc)₂ could be modeled into the electron density ($2F_o - F_c$ at $\sigma = 1.0$) adjacent to Asn60 of NCS molecule A. (b) The truncated N-linked glycans of NcPMO-2 (sticks) in the P2₁ unit cell of PDB entry 5tkf do not participate in intermolecular crystal contacts, as all atoms within 3.0 Å of either glycan (spheres) correspond to the NCS molecule to which the glycan is conjugated.

observed owing to disorder. Notably, the glycans for NCS molecules *A*, *B* and *D* each face towards solvent channels within the crystal and only show interactions with the parent protein chain, as shown in Fig. 4(b); however, the (Man)₂(GlcNAc)₂ glycan observed for NCS molecule *C* participates in crystal contacts with Ser48 and Ser50 of NCS molecule *A* via two water molecules (Fig. 4c). This contact was not observed in the previous structure of *NcPMO-2* and disrupts the *P*2₁ symmetry of the homodimer of NCS molecules *A* and *B*, yielding a triclinic space group.

As this packing contact between NCS molecules *A* and *C* is mediated by the core GlcNAc residues of the N-linked glycan, we initially attempted complete removal of the glycan from purified *NcPMO-2* using peptide:N-glycosidase F (PNGase F) to cleave the Asn60–GlcNAc amide bond. Enzymatic glycan cleavage yielded protein with very poor solubility, as did the expression of an *NcPMO-2* construct bearing an N60D mutation that removed the single N-glycosylation site. As these trials showed the N-linked glycan to be necessary for the solubility of *NcPMO-2*, reducing the length of the conjugated glycan was tried. Prior to attempting treatment of *NcPMO-2*

expressed from wild-type *P. pastoris* with exoglycosidases or endoglycosidases, we tested the crystallization behaviour of *NcPMO-2* expressed from the glycoengineered *P. pastoris* SuperMan₅. Crystallization with the largely truncated *NcPMO-2* glycoforms obtained from this strain under similar conditions also gave crystals with a thin or stacked plate morphology. However, these crystals diffracted synchrotron X-rays to 1.20 Å resolution at 100 K in space group *P*2₁ (Fig. 5). This diffraction resolution and space group matches those observed from crystals of natively expressed protein (Li *et al.*, 2012). Microseeds prepared from these crystals and added to pre-equilibrated vapor-diffusion drops yielded single crystals with equal diffraction resolution (Table 3; PDB entry 5tkg). Examination of the electron-density maps showed density corresponding to (GlcNAc)₂ and (Man)₁(GlcNAc)₂ species attached to Asn60 of the two molecules in the asymmetric unit, as shown in Fig. 6(a). (The inability to model the entire glycan is again presumed to result from disorder of the sugar residues farthest from the Asn60 linkage.) In addition, the glycans are not observed to participate in crystal contacts as was the case in the *NcPMO-2* structure of Li *et al.* (2012).

Having improved the diffraction resolution and crystal morphology of heterologous *NcPMO-2*, we attempted to grow crystals with sufficient volume for neutron protein crystallography. Single crystals grown by microseeding were transferred into larger pre-equilibrated vapor-diffusion drops to ultimately obtain an *NcPMO-2* single crystal measuring ~0.35 mm³ in total volume (Fig. 7). This crystal was used for successful room-temperature neutron (Fig. 8) and X-ray diffraction data collection to resolutions of 2.12 and 1.50 Å, respectively (Table 3; PDB entry 5tki).

Glycosylation constitutes a paradoxical situation for protein crystallization. While glycosylation can be required for protein stability or solubility, the chemical and/or conformational heterogeneity of glycan chains may prevent crystallization or lead to crystals with limited diffraction resolution (Davis & Crispin, 2011). Here, we describe how the same target protein, *NcPMO-2*, expressed from wild-type and glycoengineered *P. pastoris* showed an improvement in X-ray diffraction resolution from 2.10 to 1.20 Å and improvement in crystallographic symmetry under similar crystallization conditions. In addition, successful neutron data collection was ultimately possible with this protein. We attribute the improved crystal symmetry and diffraction resolution to the elimination of an intermolecular protein–glycan crystal contact. Conformational changes in the N-linked glycan caused by truncating to a (Man)₅(GlcNAc)₂ structure were likely to have allowed the change in crystal packing, as the core GlcNAc residues involved in the contact remained in the truncated form. It should also be noted that expression from glycoengineered *P. pastoris* produced amounts of glycan-modified protein sufficient for growing large crystals and precluded the need for potentially expensive and time-consuming post-expression treatment with glycan-modifying enzymes. These results demonstrate the utility and potential for improved crystallization behavior of heterologously expressed glycoproteins when expressed from eukaryotic hosts with engineered

Figure 7
NcPMO-2 crystal used for room-temperature X-ray and neutron data collection as seen through the quartz capillary mount. The crystal volume is approximately 0.35 mm³. Many of the fissures in the crystal appeared during room-temperature X-ray data collection, which followed the neutron data collection.

Figure 8
Neutron Laue diffraction recorded from ¹H/²H vapor-exchanged *NcPMO-2*. A 1 h exposure of the ~0.35 mm³ *NcPMO-2* crystal to neutrons with wavelengths 2.8 ≤ λ ≤ 10 Å produced the above pattern. For data collection, the incident wavelengths were restricted to 2.8 ≤ λ ≤ 4.5 Å to minimize spatial and harmonic overlaps.

glycosylation pathways producing shortened and more chemically homogeneous N-linked glycans.

Acknowledgements

Protein expression and purification experiments were conducted at the Center for Structural Molecular Biology, a DOE Office of Biological and Environmental Research User Facility. Diffraction data were collected on the Southeast Regional Collaborative Access Team 22-BM and 22-ID beamlines at the Advanced Photon Source and on the CG-4D IMAGINE beamline at the High Flux Isotope Reactor, DOE Office of Science User Facilities. Student support was provided by the ORNL Graduate Opportunities (GO!) Program and by NSF Integrative Graduate Education and Research Training program award 1069091. The authors acknowledge M. B. Goshe for useful discussions during manuscript preparation. FM acknowledges support from USDA National Institute of Food and Agriculture Hatch Award 211001.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* **D68**, 352–367.
- Arzt, S., Campbell, J. W., Harding, M. M., Hao, Q. & Helliwell, J. R. (1999). *J. Appl. Cryst.* **32**, 554–562.
- Beeson, W. T., Vu, V. V., Span, E. A., Phillips, C. M. & Marletta, M. A. (2015). *Annu. Rev. Biochem.* **84**, 923–946.
- Campbell, J. W., Hao, Q., Harding, M. M., Nguti, N. D. & Wilkinson, C. (1998). *J. Appl. Cryst.* **31**, 496–502.
- Courtade, G., Wimmer, R., Røhr, A. K., Preims, M., Felice, A. K. G., Dimarogona, M., Vaaje-Kolstad, G., Sørli, M., Sandgren, M., Ludwig, R., Eijssink, V. G. H. & Achmann, F. L. (2016). *Proc. Natl Acad. Sci. USA*, **113**, 5922–5927.
- Cruickshank, D. W. J. (1999). *Acta Cryst.* **D55**, 583–601.
- Davis, S. J. & Crispin, M. (2011). *Functional and Structural Proteomics of Glycoproteins*, edited by R. J. Owens & J. E. Nettleship, pp. 127–158. Dordrecht: Springer.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Frandsen, K. E. *et al.* (2016). *Nature Chem. Biol.* **12**, 298–303.
- Gudmundsson, M., Kim, S., Wu, M., Ishida, T., Momeni, M. H., Vaaje-Kolstad, G., Lundberg, D., Royant, A., Stahlberg, J., Eijssink, V. G. H., Beckham, G. T. & Sandgren, M. (2014). *J. Biol. Chem.* **289**, 18782–18792.
- Harris, P. V., Xu, F., Kreel, N. E., Kang, C. & Fukuyama, S. (2014). *Curr. Opin. Chem. Biol.* **19**, 162–170.
- Hemsworth, G. R., Johnston, E. M., Davies, G. J. & Walton, P. H. (2015). *Trends Biotechnol.* **33**, 747–761.
- Hemsworth, G. R., Taylor, E. J., Kim, R. Q., Gregory, R. C., Lewis, S. J., Turkenburg, J. P., Parkin, A., Davies, G. J. & Walton, P. H. (2013). *J. Am. Chem. Soc.* **135**, 6069–6077.
- Jacobs, P. P., Geysens, S., Vervecken, W., Contreras, R. & Callewaert, N. (2009). *Nature Protoc.* **4**, 58–70.
- Johansen, K. S. (2016). *Biochem. Soc. Trans.* **44**, 143–149.
- Kim, S., Stahlberg, J., Sandgren, M., Paton, R. S. & Beckham, G. T. (2014). *Proc. Natl Acad. Sci. USA*, **111**, 149–154.
- Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D. & Ludwig, R. (2012). *Biotechnol. Biofuels*, **5**, 79.
- Kjaergaard, C. H., Qayyum, M. F., Wong, S. D., Xu, F., Hemsworth, G. R., Walton, D. J., Young, N. A., Davies, G. J., Walton, P. H., Johansen, K. S., Hodgson, K. O., Hedman, B. & Solomon, E. I. (2014). *Proc. Natl Acad. Sci. USA*, **111**, 8797–8802.
- Kumar, K. S. D., Gurusaran, M., Satheesh, S. N., Radha, P., Pavithra, S., Thulaa Tharshan, K. P. S., Helliwell, J. R. & Sekar, K. (2015). *J. Appl. Cryst.* **48**, 939–942.
- Li, X., Beeson, W. T. IV, Phillips, C., Marletta, M. & Cate, J. D. (2012). *Structure*, **20**, 1051–1061.
- Mattanovich, D., Branduardi, P., Dato, L., Gasser, B., Sauer, M. & Porro, D. (2012). *Methods Mol. Biol.* **824**, 329–358.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Meilleur, F., Munshi, P., Robertson, L., Stoica, A. D., Crow, L., Kovalevsky, A., Koritsanszky, T., Chakoumakos, B. C., Blessing, R. & Myles, D. A. A. (2013). *Acta Cryst.* **D69**, 2157–2160.
- Morgan, W. D., Kragt, A. & Feeney, J. (2000). *J. Biomol. NMR*, **17**, 337–347.
- O'Dell, W. B., Agarwal, P. K. & Meilleur, F. (2016). *Angew. Chem. Int. Ed.*, <https://doi.org/10.1002/anie.201610502>.
- O'Dell, W. B., Bodenheimer, A. M. & Meilleur, F. (2016). *Arch. Biochem. Biophys.* **602**, 48–60.
- Painter, J. & Merritt, E. A. (2006a). *Acta Cryst.* **D62**, 439–450.
- Painter, J. & Merritt, E. A. (2006b). *J. Appl. Cryst.* **39**, 109–111.
- Pickford, A. R. & O'Leary, J. M. (2004). *Methods Mol. Biol.* **278**, 17–33.
- Quinlan, R. J. *et al.* (2011). *Proc. Natl Acad. Sci. USA*, **108**, 15079–15084.
- Stratton, J., Chiruvolu, V. & Meagher, M. (1998). *Methods Mol. Biol.*, **103**, 107–120.
- Sugiki, T., Ichikawa, O., Miyazawa-Onami, M., Shimada, I. & Takahashi, H. (2012). *Methods Mol. Biol.* **831**, 19–36.
- Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L.-W., Read, R. J. & Adams, P. D. (2008). *Acta Cryst.* **D64**, 61–69.
- Tolner, B., Smith, L., Begent, R. H. J. & Chester, K. A. (2006). *Nature Protoc.* **1**, 1006–1021.
- Vervecken, W., Kaigorodov, V., Callewaert, N., Geysens, S., De Vusser, K. & Contreras, R. (2004). *Appl. Environ. Microbiol.* **70**, 2639–2646.
- Vu, V. V., Beeson, W. T., Span, E. A., Farquhar, E. R. & Marletta, M. A. (2014). *Proc. Natl Acad. Sci. USA*, **111**, 13822–13827.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Zhang, Z. & Marshall, A. G. (1998). *J. Am. Soc. Mass Spectrom.* **9**, 225–233.