Published in final edited form as:

Protein Expr Purif. 2015 July ; 111: 91–97. doi:10.1016/j.pep.2015.04.002.

Bacterial expression of the phosphodiester-binding site of the cation-independent mannose 6-phosphate receptor for crystallographic and NMR studies

Linda J. Olson, Davin R. Jensen, Brian F. Volkman, Jung-Ja P. Kim, Francis C. Peterson, Rebekah L. Gundry, and Nancy M. Dahms*

Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI 53226, United States

Abstract

The cation-independent mannose 6-phosphate receptor (CI-MPR) is a multifunctional protein that interacts with diverse ligands and plays central roles in autophagy, development, and tumor suppression. By delivering newly synthesized phosphomannosyl-containing acid hydrolases from the Golgi to endosomal compartments, CI-MPR is an essential component in the generation of lysosomes that are critical for the maintenance of cellular homeostasis. The ability of CI-MPR to interact with ~60 different acid hydrolases is facilitated by its large extracellular region, with four out of its 15 domains binding phosphomannosyl residues. Although the glycan specificity of CI-MPR has been elucidated, the molecular basis of carbohydrate binding has not been determined for two out of these four carbohydrate recognition domains (CRD). Here we report expression of CI-MPR's CRD located in domain 5 that preferentially binds phosphodiester-containing glycans. Domain 5 of CI-MPR was expressed in *Escherichia coli* BL21 (DE3) cells as a fusion protein containing an N-terminal histidine tag and the small ubiquitin-like modifier (SUMO) protein. The His₆-SUMO-CRD construct was recovered from inclusion bodies, refolded in buffer to facilitate disulfide bond formation, and subjected to Ni-NTA affinity chromatography and size exclusion chromatography. Surface plasmon resonance analyses demonstrated that the purified protein was active and bound phosphorylated glycans. Characterization by NMR spectroscopy revealed high quality ¹H–¹⁵N HSQC spectra. Additionally, crystallization conditions were identified and a crystallographic data set of the CRD was collected to 1.8 Å resolution. Together, these studies demonstrate the feasibility of producing CI-MPR's CRD suitable for three-dimensional structure determination by NMR spectroscopic and X-ray crystallographic approaches.

Keywords

Mannose 6-phosphate receptor; Lectin; *Escherichia coli* expression; Protein purification; Crystallization; NMR

© 2015 Published by Elsevier Inc.

*Corresponding author at: Department of Biochemistry, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, United States. Tel.: +1 414 955 4698; fax: +1 414 955 6510. ndahms@mcw.edu (N.M. Dahms).

Conflict of interest statement

None declared.

Introduction

The 300 kDa cation-independent mannose 6-phosphate receptor (CI-MPR)¹ supplies the lysosome with ~60 different hydrolytic enzymes that are required for the cellular functioning of phagocytic and autophagic degradative pathways [1,2]. High mannose-type *N*-glycans on newly synthesized acid hydrolases are modified by the Golgi enzymes, UDP-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosamine-1-phosphotransferase (GlcNAc-1-phosphotransferase, EC 2.7.8.17) [3–6] and *N*-acetylglucosamine-1-phosphodiester α -*N*-acetylglucosaminidase (also referred to as uncovering enzyme, EC 3.1.4.45) [7–11], which mark the *N*-glycan chain with one or two phosphomannosyl residues that can be phosphomonoesters (mannose 6-phosphate, M6P) and/or phosphodiester (M6P-GlcNAc) [4]. The large extracellular region of CI-MPR has 15 domains, four of which are carbohydrate recognition domains (CRD) and bind phosphomannosyl residues to allow for the receptor-mediated delivery of newly synthesized acid hydrolases to the lysosome (reviewed in [12,13]; see also [14]) (Fig. 1).

Of the four CRDs, domain 5 binds preferentially to glycans with phosphodiester [15–17]. The importance of this domain in the functioning of CI-MPR comes from studies showing that transgenic mice expressing only M6P-GlcNAc-containing acid hydrolases have a normal phenotype [18]. In contrast, mucopolipidosis II (also known as I-cell disease) patients are deficient in GlcNAc-1-phosphotransferase activity and exhibit aberrant mistargeting and secretion of lysosomal enzymes due to the inability to mark their acid hydrolases with phosphomannosyl residues. The severely compromised lysosomal function exhibited by mucopolipidosis patients results in cardiomegaly, hepatomegaly, and death typically by the age of 10 years [1,19,20]. These studies indicate that domain 5 of CI-MPR by binding phosphodiester, rescues these transgenic mice from the severe ML II-like phenotype observed in mice [21] or patients lacking GlcNAc-1-phosphotransferase activity.

To understand the molecular basis of CI-MPR's ability to target acid hydrolases to the lysosome, structural studies are needed of the receptor bound to its carbohydrate ligand. We determined previously the solution structure of bovine CI-MPR domain 5 and have modeled the location of the ligand in the binding pocket from analyses of NOE data [22]. In this previous study, ¹⁵N-labeled bovine CI-MPR domain 5 was expressed in *Pichia pastoris* yeast as an *N*-glycosylated, secreted glycoprotein. However, our attempts to solve the structure of domain 5 bound to phosphomannosyl residues have been stymied by the poor quality of crystals using protein purified from *P. pastoris* yeast. In addition, challenges have been encountered during the purification of domain 5 from *P. pastoris* yeast involving: (1) removal of the heterogeneous *N*-glycans from the recombinant protein, and (2) removal of mannans that are secreted by *P. pastoris* yeast from the medium containing the secreted recombinant proteins. Because of these issues, an improved expression system and/or purification protocol was needed to generate domain 5 protein suitable for crystallization.

¹Abbreviations used: CI-MPR, cation-independent mannose 6-phosphate receptor; M6P, mannose 6-phosphate; GlcNAc-1-phosphotransferase, UDP-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosamine-1-phosphotransferase; GAA, acid α -glucosidase; Ni-NTA, nickel nitrilotriacetic acid; CRD, carbohydrate recognition domains; MS, mass spectrometry; SUMO, small ubiquitin-like modifier; HSQC, heteronuclear single quantum coherence; G6P, glucose 6-phosphate.

In the current report, we have optimized the expression of domain 5 of the human CI-MPR as a fusion protein with the cleavable small ubiquitin-like modifier protein, SUMO, in *Escherichia coli* to generate non-glycosylated protein (Fig. 1). A simplified purification scheme was developed involving a refolding protocol of the protein derived from *E. coli* inclusion bodies, immobilized metal affinity chromatography, and size exclusion chromatography that resulted in improved yields of ^{15}N -labeled domain 5. High-resolution mass spectrometry (MS) and MS/MS fragmentation confirmed the sequence of the protein and the presence of three disulfide bonds. Surface plasmon resonance studies demonstrated that the refolded protein was functional and bound phosphomannosyl-containing glycans. Analyses of the ^{15}N -labeled domain 5 protein reveal that it exhibits high quality two-dimensional (2D) heteronuclear single quantum coherence (HSQC) spectra suitable for structural studies. Additionally, crystallization conditions were identified that resulted in the collection of a crystallographic data set to 1.8 Å resolution. Taken together, these results demonstrate the feasibility of obtaining the structure of domain 5 in the presence of bound ligand.

Materials and methods

Gene cloning and vector construction

The DNA sequence corresponding to domain 5 of the human CI-MPR (residues 619–761, numbering includes the N-terminal signal sequence and begins with the initiator methionine) was amplified directly from the human clone (GeneBank Accession No. J03528) obtained from the American Type Culture Collection (pGEM-8 ATCC 95661) using standard polymerase chain reaction (PCR) methods. The cDNA was cloned into a pET28a vector (Novagen) modified to express the small ubiquitin-like modifier (SUMO) protein with an N-terminal tag of six histidine residues [23] in which domain 5 was fused to the C-terminus of SUMO (referred to as His₆-SUMO-CRD). DNA sequencing was used to confirm the sequence of the plasmid used for expression in *E. coli* BL21 (DE3) cells.

Bacterial expression of CI-MPR domain 5

The plasmid containing the SUMO-domain 5 fusion construct was transformed into *E. coli* BL21 (DE3) cells. An individual colony was taken from the plate and used to inoculate a 3 ml-culture in Luria–Bertani (LB) medium (1% NaCl, 1% tryptone, 0.5% yeast extract) supplemented with 50 µg/ml kanamycin and grown at 37 °C for approximately 6 h and then transferred to 25 ml of LB medium with kanamycin (50 µg/ml) and allowed to grow overnight. The 25 ml culture was then used to inoculate 1 L of LB medium supplemented with 50 µg/ml kanamycin and allowed to grow at 37 °C until the cell density reached an OD₆₀₀ of 0.8. Expression of the protein was induced by the addition of isopropyl β-D-1-thiogalactopyranoside to the culture (0.5 mM final concentration) and the cells were grown for 18 h at 16 °C. An aliquot of the cells was assayed for protein expression by SDS–PAGE and visualization by Coomassie brilliant blue R-250 (Bio-Rad) dye staining. To generate protein suitable for structural studies by NMR spectroscopy, ^{15}N -labeled protein was produced from M9-based minimal media containing [^{15}N] ammonium chloride (1 g/L, Sigma or Cambridge Isotope Laboratory, MA).

Purification of CI-MPR domain 5

Bacterial cells (5 g wet weight) were collected by centrifugation of 1 L of culture medium at 5000 g for 10 min and resuspended in 15 ml lysis buffer (20 mM Tris pH 7.8, 300 mM NaCl) at 4 °C. After sonication (3 × 15 s, with 10 s between pulses), the cell lysate was centrifuged at 12,500 g for 25 min at 4 °C. The pellet containing the inclusion bodies was solubilized in 18 ml buffer AD (6 M guanidine-HCl, 20 mM Tris pH 7.8, 300 mM NaCl, 0.1% (v/v) 2-mercaptoethanol) at 4 °C followed by centrifugation at 15,000 g for 20 min at 4 °C to remove insoluble material. The sample was passed over a Ni-NTA agarose (Clontech) column, washed, and then eluted with 40 ml buffer BD (6 M guanidine-HCl, 100 mM sodium acetate pH 4.5, 300 mM NaCl). Refolding of the His₆-SUMO-CRD fusion protein (100 mg at 2.5 mg/ml) was performed by diluting the Ni-NTA eluate dropwise into 480 ml (i.e., 12 times the volume of the Ni-NTA eluate) refolding buffer (100 mM Tris pH 8.0, 10 mM cysteine and 0.5 mM cysteine) at 4 °C and then stirring for 2 h at 22 °C. Following concentration by filtration using Amicon stirred cells (10 kDa MWCO membranes), the protein was dialyzed against Ni-NTA binding buffer (20 mM Tris, 300 mM NaCl pH 7.8) for 18 h at 4 °C. The SUMO-specific peptidase, Ulp1 (1.86 mg, ~1/45 the weight to the target fusion protein, which is now at 86 mg following concentration and dialysis), was added to remove the N-terminal His₆-tag and SUMO portion of the fusion protein. After incubation for 30 min at 4 °C, the sample was centrifuged at 12,500 g for 20 min at 4 °C to remove insoluble material. To separate the His₆ tag-SUMO protein that was cleaved from CI-MPR domain 5, the cleared supernatant was passed over a Ni-NTA agarose column, washed with Ni-NTA binding buffer containing 20 mM imidazole, and then eluted with Ni-NTA binding buffer containing 400 mM imidazole. The fractions from the Ni-NTA column that contained domain 5 (45 mg) were concentrated by filtration using Amicon stirred cells prior to size exclusion chromatography using a Sephacryl S-300 column (10 × 300 mm) or Superdex 75 column (1.6 × 60 cm, GE Healthcare) to isolate a single species. Starting with 100 mg of fusion protein isolated from the Ni-NTA affinity chromatography step, 14 mg of Ulp1-cleaved domain 5 protein with >95% purity was obtained after size exclusion chromatography and used for crystallization. Protein for NMR spectroscopic analyses was buffer exchanged into a solution containing 10 mM deuterated imidazole (pH 6.5) and 150 mM NaCl. The concentration of the purified protein was determined using the Bradford assay (Bio-Rad) with bovine serum albumin as the standard.

Mass spectrometry of purified CI-MPR domain 5

Purified CI-MPR domain 5 protein (100 µg) was desalted using a MicroSpin C₁₈ column (Harvard Apparatus) and eluted in 70% (v/v) acetonitrile containing 0.1% (v/v) formic acid. To allow for confirmation of disulfide bonds, protein (100 µl at 62 µM) was also reduced with 10 mM Tris(2-carboxyethyl)phosphine-HCl for 1 h at 60 °C in buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, 6 M guanidine HCl prior to desalting. Samples were analyzed by direct infusion using a *Q Exactive*TM Hybrid Quadrupole-Orbitrap mass spectrometer (ThermoFisher Scientific). All analyses were performed with resolving power set to 140,000. For top-down analyses, tandem mass spectrometry (MS/MS) was performed using higher-energy collisional dissociation (HCD) with collision energy set to 25 and each scan included four microscans. Data were deconvoluted and deisotoped using *Xtract* and

*Xcalibur*TM in QualBrowser (ThermoFisher Scientific) and MS/MS data were analyzed using ProSight Lite [24].

Biosensor studies

Surface plasmon resonance analyses were performed using a BIAcore 3000 instrument (BIAcore, GE Healthcare) as described previously [14]. Briefly, the lysosomal enzyme acid α -glucosidase (GAA) with defined *N*-glycans, either phosphomonoesters or phosphodiester [16], was immobilized on CM5 sensor chips. Purified CI-MPR domain 5 in buffer containing 50 mM imidazole, pH 6.5, 150 mM NaCl, 5 mM MgCl₂, 5 mM MnCl₂, and 5 mM CaCl₂ supplemented with 0.005% (v/v) surfactant polysorbate 20 (GE Healthcare) was injected in a volume of 80 μ l over the coupled and reference flow cells at a flow rate of 40 μ l/min. After 2 min, the complexes were allowed to dissociate for 2 min in buffer lacking domain 5. Representative sensorgrams are shown from three replicate experiments. In other experiments, purified CI-MPR domain 5 (40 μ M) was injected in the presence of M6P (1 or 5 mM), or glucose 6-phosphate (G6P, 1 or 5 mM). G6P was used as a control in these inhibition studies because in contrast to M6P ($K_D = \sim 10$ μ M), G6P binds with low affinity ($K_D = 10$ –80 mM) to the full-length CI-MPR [25].

¹H–¹⁵N HSQC of CI-MPR domain 5

NMR spectra of purified ¹⁵N-labeled domain 5 at a concentration of 100 μ M were acquired at 25 °C on a Bruker 600-MHz spectrometer equipped with a triple resonance CryoProbeTM as described previously [22]. Spectral data were processed using NMRPipe [26] and visualized using XEASY [27].

Crystallization of CI-MPR domain 5

Purified domain 5 protein at 6.6 mg/ml in 20 mM Tris pH 7.5/150 mM NaCl was screened for crystallization conditions using the Joint Center for Structural Genomics screening kit (Molecular Dimensions JCSG⁺ screen) in a 96-well Art Robbins Intelli crystallization plate in the presence of 100 mM M6P and 10 mM MnCl₂. The robotic Crystal Phoenix liquid handling system (Art Robbins Instruments) was used to set up equal volume hanging drops (0.15 μ l each of protein and well solution) and the plates were incubated at 19 °C. Crystals were obtained in buffer containing 0.2 M MgCl₂, 0.1 M Bis–Tris pH 5.5, 25% (v/v) polyethylene glycol 3350 after ~6 weeks. The crystal was quickly passed through a cryoprotection solution containing 0.2 M MgCl₂, 0.1 M Bis–Tris pH 5.5, 28% (v/v) polyethylene glycol 3350, and 20% (v/v) glycerol before flash-freezing in liquid nitrogen. A native data set was collected using a Rigaku R-Axis IV⁺⁺ image plate detector system and MicroMax 007 generator equipped with Osmic confocal mirrors and X-treme crystal cooling system operating at –175 °C.

Results and discussion

Expression and purification of CI-MPR domain 5 in *E. coli*

To obtain protein that contains solely the native sequence of the CI-MPR for subsequent structural studies, a plasmid was constructed that encodes domain 5 of human CI-MPR (residues 619–761) joined to the C-terminus of a His₆-tagged small ubiquitin-like modifier

(SUMO) protein. The presence of the hexahistidine tag permits purification of the His₆-SUMO-CRD fusion protein by immobilized metal affinity chromatography whereas the inclusion of SUMO in the fusion protein takes advantage of the specificity of the SUMO-specific peptidase, Ulp1. Ulp1 peptidase recognizes natively folded SUMO and its cleavage removes the SUMO protein such that no additional residues are associated with the adjacent protein domain in the fusion construct [28,29]. Therefore, cleavage by Ulp1 can serve to indirectly monitor the native structure of SUMO during refolding procedures and its mode of cleavage results in a native sequence of the downstream fusion partner.

The plasmid encoding His₆-SUMO-CRD was used to transform *E. coli* BL21 (DE3) cells. Colonies were screened for protein expression by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) following overnight growth at 16 °C in medium containing isopropyl β-D-1-thiogalactopyranoside. The results show that the His₆-SUMO-CRD fusion protein of the expected size of ~35 kDa was expressed predominantly in the insoluble fraction (Fig. 2). The fusion protein was recovered from inclusion bodies (1 L culture) using 6 M guanidine–HCl and purified by affinity chromatography using Ni-NTA agarose. Because CI-MPR domain 5 contains six cysteine residues that form three disulfide bonds, refolding conditions were optimized using buffers containing a mixture of cysteine and cystine to facilitate disulfide bond formation. Following dialysis, the fusion protein (33 mg) was cleaved by Ulp1 peptidase and reapplied to a Ni-NTA column to remove the N-terminal His₆-tagged SUMO protein from CI-MPR domain 5 (Fig. 3). The unbound fractions from the Ni-NTA column containing the cleaved domain 5, which migrated at the expected size of 16 kDa, were subjected to preparative size-exclusion chromatography on Sephacryl S-300 and the purity of the fractions was analyzed by SDS–PAGE (Fig. 4).

Peak fractions (lanes 5–8) consistent with a monomeric form of the protein and containing greater than 90% pure domain 5 were pooled (14 mg per liter of bacterial culture; an additional 7 mg of domain 5 was present in fractions exhibiting ~75–90% purity) and subjected to mass spectrometry. High-resolution mass spectrometry analyses of non-reduced and reduced protein demonstrate the intact mass of CI-MPR domain 5 is consistent with the proposed amino acid sequence shown in Fig. 5. In the non-reduced condition, the mass is consistent with all six cysteine residues being oxidized and involved in disulfide bonds, whereas the mass of the protein under reducing conditions is consistent with the presence of six cysteine residues in their reduced form. MS/MS fragmentation confirms the sequence of CI-MPR domain 5, and as expected, the reduced form gives rise to more complete fragmentation within the mass spectrometer resulting in higher sequence coverage, protein characterization score (PCS), and more significant *P*-score (Fig. 5B). Together, these data are consistent with Ulp1 peptidase cleaving at the predicted site and releasing the CI-MPR domain 5 that contains only its native sequence. Furthermore, these data demonstrate that expression of SUMO-MPR fusion protein in BL21 (DE3) *E. coli* cells represents a robust expression system for the generation of sufficient amounts of CI-MPR domain 5 for structural studies.

CI-MPR domain 5 binds phosphomannosyl residues

Surface plasmon resonance analyses were performed to determine whether the CI-MPR domain 5 protein isolated and refolded from *E. coli* inclusion bodies attained its native conformation and was competent to bind carbohydrates. The lysosomal enzyme acid α -glucosidase (GAA) with defined high mannose-type *N*-glycans containing either phosphomonoesters (M6P) or phosphodiester (GlcNAc-M6P) [16], was immobilized on the sensor surface and increasing concentrations of the purified CI-MPR domain 5 protein was passed over the immobilized GAA. The representative sensorgrams show that CI-MPR domain 5 binds with a higher affinity to GAA bearing phosphodiester glycans (GAA monoester $K_D = 292 \pm 66 \mu\text{M}$, GAA diester $K_D = 72 \pm 7 \mu\text{M}$) (Fig. 6, *left panels*). These values are similar to those we have previously obtained of bovine CI-MPR domain 5 containing a C-terminal hexahistidine tag purified from *P. pastoris* yeast before and after enzymatic deglycosylation using endoglycosidase H (GAA monoester $K_D = 438 \pm 63 \mu\text{M}$ versus GAA diester $K_D = 18 \pm 2 \mu\text{M}$ [16]; GAA monoester $K_D = 145 \pm 33 \mu\text{M}$ versus GAA diester $K_D = 58 \pm 9 \mu\text{M}$ [22]). The addition of G6P (1 mM or 5 mM) did not appreciably inhibit the interaction of CI-MPR domain 5 with either coupled surface (Fig. 6, *right panels*). In contrast, the addition of 1 mM M6P had a minimal inhibitory effect whereas 5 mM M6P resulted in a ~15–25% inhibition of CI-MPR domain 5 binding to both GAA monoester- and GAA diester-coupled surfaces (Fig. 6, *right panels*). These findings are consistent with our previously published inhibition studies of the bovine CI-MPR domain 5 purified from *P. pastoris* which demonstrated that M6P was a 2.4–2.6-fold better inhibitor than G6P [16]. Furthermore, the lack of complete inhibition by 5 mM M6P (Fig. 6, *right panels*) is consistent with the higher concentrations of M6P required to achieve 50% inhibition of the interaction between bovine CI-MPR domain 5 purified from *P. pastoris* and GAA (GAA monoester $K_i = 18 \pm 2 \text{ mM}$, GAA diester $K_i = 14 \pm 2 \text{ mM}$) [16]. These quantitative binding studies demonstrate that human CI-MPR domain 5 can be refolded from *E. coli* inclusion bodies to produce a functional protein capable of binding phosphomannosyl-containing glycans.

Two-dimensional NMR spectroscopy

To further assess the folded status of CI-MPR domain 5 and to obtain protein suitable for structure determination in solution, transformed BL21 (DE3) cells were grown in M9-based minimal medium supplemented with [^{15}N] ammonium chloride. ^{15}N -labeled domain 5 protein was refolded and purified as described above. Size exclusion chromatography on a Superdex 75 column revealed a significant fraction of the ^{15}N -labeled domain 5 protein eluted as higher molecular weight species (Fig. 7). For structural characterizations, peak fractions consistent with a monomeric form of the protein and containing greater than 90% pure domain 5 were pooled from the size exclusion column (Fig. 7). The yield of purified ^{15}N -labeled domain 5 from 1 L of bacterial culture was 1.3 mg. Importantly, as shown in Fig. 8, a high quality 2D ^1H – ^{15}N HSQC spectrum was obtained of domain 5 in the absence of ligand.

These results demonstrate that a sufficient amount of ^{15}N -labeled human CI-MPR domain 5 can be produced in *E. coli* and refolded from inclusion bodies, and is suitable for structure determination by NMR spectroscopy. Although this protocol results in a lower yield of ^{15}N -

labeled CI-MPR domain 5 than we had obtained from *P. pastoris* yeast [22], this method is significantly less time intensive and less costly. The purification protocol of domain 5 from BL21 (DE3) bacterial cells requires fewer steps than our published protocol of domain 5 expressed in *P. pastoris* yeast to achieve a high level of purity. We found that domain 5, when produced as a secreted, *N*-glycosylated protein in *P. pastoris* yeast, displays poor NMR 2D HSQC spectral quality that is dramatically improved following enzymatic deglycosylation using endoglycosidase H coupled with multiple chromatography steps using Concanavalin A Sepharose to remove the enzymatically released glycans and contaminating mannans that are secreted by *P. pastoris* yeast [22].

Crystallization of CI-MPR domain 5

To determine whether bacterially expressed CI-MPR domain 5 would have improved crystallization properties compared to that expressed in *P. pastoris* yeast, purified CI-MPR domain 5 at a concentration of 6.6 mg/ml was screened in 96-well plates incubated at 19 °C for crystallization conditions using the Joint Center for Structural Genomics screening kit (Molecular Dimensions JCSG⁺ screen). The screens were performed in the presence and absence of the carbohydrate ligand, M6P. Crystals were obtained in buffer containing 0.2 M MgCl₂, 0.1 M Bis-Tris pH 5.5, 25% polyethylene glycol 3350, 100 mM M6P and 10 mM MnCl₂ after 6 weeks (Fig. 9, inset). A single crystal fragment was removed from the cluster shown in Fig. 9. The crystal was found to diffract at high resolution (1.77 Å) (Fig. 9) and a data set was collected. The overall R_{merge} value of 0.053 (50.00–1.77 Å) and 0.379 in the outermost shell (1.80–1.77 Å) (redundancy 7.1 overall, 6.2 outermost shell) serves to demonstrate the overall high quality of the data obtained. Initial analyses of the data set reveal two molecules per asymmetric unit and a P2₁2₁2 space group with unit cell dimensions $a = 32.5$ Å, $b = 70.1$ Å, and $c = 113.9$ Å. These results demonstrate that the human CI-MPR domain 5 protein produced in *E. coli* and refolded from inclusion bodies is superior to that of the protein expressed as a secreted protein in *P. pastoris* yeast with respect to the attainment of crystals that diffract to high resolution.

Conclusions

In this study we have successfully expressed and purified from *E. coli* inclusion bodies human CI-MPR domain 5 that contains a phosphodiester-binding site. High-resolution mass spectrometry confirmed the presence of the expected three disulfide bonds. NMR spectroscopy demonstrated the protein was folded as evidenced by the well resolved 2D HSQC spectra. These high quality spectra indicated the protein was suitable for structural studies. Quantitative binding studies using surface plasmon resonance confirmed that the protein had folded into a native conformation as it was active and bound phosphomannosyl-containing lysosomal enzymes. Furthermore, we successfully found a crystallization condition in the presence of ligand, M6P. The crystal diffracted to high resolution and a data set has been collected. We are now poised to continue our structural studies of the CI-MPR to further define how this receptor recognizes a diverse repertoire of lysosomal enzymes. Our findings also provide a starting point for structural studies of the remaining two carbohydrate-binding sites of the CI-MPR located in domain 9 and domain 15. Our efficient bacterial expression and purification strategy, which circumvents the issues of producing a

glycan binding protein in the mannan-rich environment of *P. pastoris* yeast, allows for the production of truncated forms of the CI-MPR that contain multiple disulfide bonds in sufficient quantity and quality to be useful for structural studies. Furthermore, this method of expression and purification can be applied to other MRH domain-containing proteins found in the secretory pathway.

Acknowledgments

Funding

This work was supported in part by the National Institutes of Health Grants R01DK042667 to N.M.D. and R01AI0580723 to B.F.V.

We thank Ulrike Galasinski for technical assistance. The BIAcore 3000 instrument was purchased through a grant from the Advancing a Healthier Wisconsin Program.

References

1. Kornfeld S.; Sly, WS. I cell disease and pseudo-Hurler polydystrophy: disorders of lysosomal enzyme phosphorylation and localization. In: Scriver, CR.; Beaudet, AL.; Sly, WS.; Valle, D., editors. *Metabolic and Molecular Bases of Inherited Diseases*. McGraw Hill; New York: 2001. p. 3469-3482.
2. Braulke T, Bonifacino JS. Sorting of lysosomal proteins. *Biochim Biophys Acta*. 2009; 1793:605–614. [PubMed: 19046998]
3. Hasilik A, Klein U, Waheed A, Strecker G, von Figura K. Phosphorylated oligosaccharides in lysosomal enzymes: identification of alpha-N-acetylglucosamine(1)phospho(6)mannose diester groups. *Proc Natl Acad Sci USA*. 1980; 77:7074–7078. [PubMed: 6938953]
4. Varki A, Kornfeld S. Structural studies of phosphorylated high mannose-type oligosaccharides. *J Biol Chem*. 1980; 255:10847–10858. [PubMed: 7430158]
5. Bao M, Booth JL, Elmendorf BJ, Canfield WM. Bovine UDP-N-acetylglucosamine:lysosomal-enzyme N-acetylglucosamine-1-phosphotransferase. I. Purification and subunit structure. *J Biol Chem*. 1996; 271:31437–31445. [PubMed: 8940155]
6. Kudo M, Bao M, D'Souza A, Ying F, Pan H, Roe BA, Canfield WM. The alpha- and beta-subunits of the human UDP-N-acetylglucosamine:lysosomal enzyme phosphotransferase are encoded by a single cDNA. *J Biol Chem*. 2005; 280:36141–36149. [PubMed: 16120602]
7. Waheed A, Hasilik A, von Figura K. Processing of the phosphorylated recognition marker in lysosomal enzymes. Characterization and partial purification of a microsomal alpha-N-acetylglucosaminyl phosphodiesterase. *J Biol Chem*. 1981; 256:5717–5721. [PubMed: 6263889]
8. Varki A, Sherman W, Kornfeld S. Demonstration of the enzymatic mechanisms of alpha-N-acetyl-D-glucosamine-1-phosphodiester N-acetylglucosaminidase (formerly called alpha-N-acetylglucosaminylphosphodiesterase) and lysosomal alpha-N-acetylglucosaminidase. *Arch Biochem Biophys*. 1983; 222:145–149. [PubMed: 6301379]
9. Kornfeld R, Bao M, Brewer K, Noll C, Canfield W. Molecular cloning and functional expression of two splice forms of human N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase. *J Biol Chem*. 1999; 274:32778–32785. [PubMed: 10551838]
10. Rohrer J, Kornfeld R. Lysosomal hydrolase mannose 6-phosphate uncovering enzyme resides in the trans-Golgi network. *Mol Biol Cell*. 2001; 12:1623–1631. [PubMed: 11408573]
11. Do H, Lee WS, Ghosh P, Hollowell T, Canfield W, Kornfeld S. Human mannose 6-phosphate-uncovering enzyme is synthesized as a proenzyme that is activated by the endoprotease furin. *J Biol Chem*. 2002; 277:29737–29744. [PubMed: 12058031]
12. Kim JJ, Olson LJ, Dahms NM. Carbohydrate recognition by the mannose-6-phosphate receptors. *Curr Opin Struct Biol*. 2009; 19:534–542. [PubMed: 19801188]
13. Brown J, Jones EY, Forbes BE. Keeping IGF-II under control: lessons from the IGF-II-IGF2R crystal structure. *Trends Biochem Sci*. 2009; 34:612–619. [PubMed: 19796953]

14. Olson LJ, Castonguay AC, Lasanajak Y, Peterson FC, Cummings RD, Smith DF, Dahms NM. Identification of a fourth mannose 6-phosphate binding site in the cation-independent mannose 6-phosphate receptor. *Glycobiology*. 2015
15. Reddy ST, Chai W, Childs RA, Page JD, Feizi T, Dahms NM. Identification of a low affinity mannose 6-phosphate-binding site in domain 5 of the cation-independent mannose 6-phosphate receptor. *J Biol Chem*. 2004; 279:38658–38667. [PubMed: 15252023]
16. Chavez CA, Bohnsack RN, Kudo M, Gotschall RR, Canfield WM, Dahms NM. Domain 5 of the cation-independent mannose 6-phosphate receptor preferentially binds phosphodiester (mannose 6-phosphate N-acetylglucosamine ester). *Biochemistry*. 2007; 46:12604–12617. [PubMed: 17927214]
17. Bohnsack RN, Song X, Olson LJ, Kudo M, Gotschall RR, Canfield WM, Cummings RD, Smith DF, Dahms NM. Cation-independent mannose 6-phosphate receptor: a composite of distinct phosphomannosyl binding sites. *J Biol Chem*. 2009; 284:35215–35226. [PubMed: 19840944]
18. Boonen M, Vogel P, Platt KA, Dahms N, Kornfeld S. Mice lacking mannose 6-phosphate uncovering enzyme activity have a milder phenotype than mice deficient for N-acetylglucosamine-1-phosphotransferase activity. *Mol Biol Cell*. 2009; 20:4381–4389. [PubMed: 19710420]
19. Tiede S, Storch S, Lubke T, Henrissat B, Bargal R, Raas-Rothschild A, Bräulke T. Mucopolipidosis II is caused by mutations in GNPTA encoding the alpha/beta GlcNAc-1-phosphotransferase. *Nat Med*. 2005; 11:1109–1112. [PubMed: 16200072]
20. Kudo M, Brem MS, Canfield WM. Mucopolipidosis II (I-cell disease) and mucopolipidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha/beta-subunits precursor gene. *Am J Hum Genet*. 2006; 78:451–463. [PubMed: 16465621]
21. Gelfman CM, Vogel P, Issa TM, Turner CA, Lee WS, Kornfeld S, Rice DS. Mice lacking alpha/beta subunits of GlcNAc-1-phosphotransferase exhibit growth retardation, retinal degeneration, and secretory cell lesions. *Invest Ophthalmol Vis Sci*. 2007; 48:5221–5228. [PubMed: 17962477]
22. Olson LJ, Peterson FC, Castonguay A, Bohnsack RN, Kudo M, Gotschall RR, Canfield WM, Volkman BF, Dahms NM. Structural basis for recognition of phosphodiester-containing lysosomal enzymes by the cation-independent mannose 6-phosphate receptor. *Proc Natl Acad Sci USA*. 2010; 107:12493–12498. [PubMed: 20615935]
23. Takekoshi T, Ziarek JJ, Volkman BF, Hwang ST. A locked, dimeric CXCL12 variant effectively inhibits pulmonary metastasis of CXCR4-expressing melanoma cells due to enhanced serum stability. *Mol Cancer Ther*. 2012; 11:2516–2525. [PubMed: 22869557]
24. Fellers RT, Greer JB, Early BP, Yu X, LeDuc RD, Kelleher NL, Thomas PM. ProSight Lite: graphical software to analyze top-down mass spectrometry data. *Proteomics*. 2014
25. Tong PY, Kornfeld S. Ligand interactions of the cation-dependent mannose 6-phosphate receptor. Comparison with the cation-independent mannose 6-phosphate receptor. *J Biol Chem*. 1989; 264:7970–7975. [PubMed: 2542255]
26. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Vax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. 1995; 6:277–293. [PubMed: 8520220]
27. Bartels C, Xia TH, Billeter M, Guntert P, Wuthrich K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J Biomol NMR*. 1995; 6:1–10. [PubMed: 22911575]
28. Malakhov MP, Mattern MR, Malakhova OA, Drinker M, Weeks SD, Butt TR. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J Struct Funct Genomics*. 2004; 5:75–86. [PubMed: 15263846]
29. Weeks SD, Drinker M, Loll PJ. Ligation independent cloning vectors for expression of SUMO fusions. *Protein Expr Purif*. 2007; 53:40–50. [PubMed: 17251035]

Fig. 1. Schematic diagram of human CI-MPR and His₆-SUMO-CRD fusion protein. CI-MPR is a type I integral membrane proteins and the location of the single transmembrane domain is shown by a *vertical hatched bar*. CI-MPR contains 15 contiguous MRH domains, with domain 13 containing a 48-residue fibronectin type II (FnII) insert (gray).

Fig. 2.

SDS-PAGE analysis of His₆-SUMO-CRD fusion protein ("SUMO-Dom5") expressed in BL21 (DE3) *E. coli* cells. Following 18 h induction with 0.5 mM isopropyl β-D-1-thiogalactopyranoside, a portion of the total (T) and soluble (S) fraction of BL21 (DE3) *E. coli* cells were evaluated by SDS-PAGE on a 10% acrylamide gel run under non-reducing conditions. Proteins were visualized by Coomassie staining.

Fig. 3. SDS-PAGE analysis of Ni-NTA resin purification of CI-MPR domain 5. Following incubation with the SUMO-specific protease Ulp1, the sample was subjected to Ni-NTA affinity chromatography. The Ni-NTA column was washed with buffer (fractions 1–5) and eluted with buffer containing 400 mM imidazole (fraction 6). Aliquots of each fraction were analyzed on a 10% acrylamide gel run under non-reducing conditions and the proteins were visualized by Coomassie staining.

Fig. 4. Purification of CI-MPR domain 5 by gel filtration. The run-through fractions from the Ni-NTA affinity column were subjected to size exclusion chromatography using a Sephacryl S-300 column (10 × 300 mm) run at a flow rate of 0.4 ml/min in 20 mM Tris pH 7.5 and 150 mM NaCl. Fractions (0.8 ml) comprising the major peak (mAU₂₈₀) were assayed for purity by SDS-PAGE and the proteins were visualized by Coomassie staining.

Fig. 5.

Mass spectrometry analysis of purified CI-MPR domain 5. Data are shown for non-reduced (A) and reduced (B) samples. In each panel, isotopically resolved spectra of intact purified CI-MPR domain 5 are shown. Inset: deconvoluted, deisotoped spectra for each condition labeled with neutral [M] monoisotopic masses. Bottom: sequence coverage of CI-MPR domain 5 from top-down MS/MS analyses of the 16053.53 and 16059.50 ions. Protein characterization score (PCS), *P*-score and sequence coverage image were generated using Prosight Lite [24].

Fig. 6.

Surface plasmon resonance analyses of the interaction between CI-MPR domain 5 and lysosomal enzyme GAA with defined glycans. Increasing concentrations (1, 5, 10, 20, 40, 60, 80, and 100 μM) of the domain 5 protein were flowed over the reference surface or GAA monoester or GAA diester coupled surfaces for 2 min followed by buffer for 2 min. Representative sensorgrams are shown. The response to the reference surface, which ranges from 0 to 30 response units (RU) with 30 RU being observed only at the highest concentration of domain 5 protein used, has been subtracted from the representative sensorgrams. Insets depict the response at equilibrium (R_{eq}) plotted against receptor concentration and data are fitted by non-linear regression to a one-site saturation-binding model (SigmaPlot version 10.0, Systat Software, Inc.) to calculate the K_D under steady-state conditions [22]. The *right panels* show representative sensorgrams in which domain 5 was flowed over either the reference surface or GAA monoester- or diester-coupled surface in the absence or presence of 1 or 5 mM M6P or G6P.

Fig. 7.

Purification of ^{15}N -labeled CI-MPR domain 5 by Ni-NTA affinity chromatography and gel filtration. Following incubation with the SUMO-specific protease Ulp1, the sample was subjected to Ni-NTA affinity chromatography. The Ni-NTA column was washed with buffer (fractions 1 and 2) and eluted with buffer containing 400 mM imidazole (fraction 3). Aliquots of each fraction were analyzed on a 10% acrylamide gel run under non-reducing conditions and the proteins were visualized by Coomassie staining (*left panel*). The run-through fractions from the Ni-NTA affinity column were subjected to size exclusion chromatography using a Superdex 75 column (1.6×60 cm) run at a flow rate of 0.75 ml/min in 20 mM Tris pH 7.5 and 150 mM NaCl. Fractions (1.5 ml) comprising the major peak

(mAU₂₈₀) were assayed for purity by SDS–PAGE and the proteins were visualized by Coomassie staining (*right panel*). Pooled fractions beginning with lane 3 (*right panel*) were used for NMR analyses.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 8.

Two dimensional ^1H - ^{15}N HSQC spectrum of 100 μM domain 5 in buffer containing 10 mM deuterated imidazole (pH 6.5) and 150 mM NaCl is shown.

Fig. 9. Crystallization of CI-MPR domain 5. The crystal (inset) and resulting diffraction pattern are shown.