scientific reports

Check for updates

OPEN

Lysosomal enzyme binding to the cation-independent mannose 6-phosphate receptor is regulated allosterically by insulin-like growth factor 2

Richard N. Bohnsack[1], Sandeep K. Misra[2], Jianfang Liu[3], Mayumi Ishihara-Aoki[4], Michaela Pereckas[1], Kazuhiro Aoki[4,5], Gang Ren[3], Joshua S. Sharp[2,6] & Nancy M. Dahms[1]✉

The cation-independent mannose 6-phosphate receptor (CI-MPR) is clinically significant in the treatment of patients with lysosomal storage diseases because it functions in the biogenesis of lysosomes by transporting mannose 6-phosphate (M6P)-containing lysosomal enzymes to endosomal compartments. CI-MPR is multifunctional and modulates embryonic growth and fetal size by downregulating circulating levels of the peptide hormone insulin-like growth factor 2 (IGF2). The extracellular region of CI-MPR comprises 15 homologous domains with binding sites for M6P-containing ligands located in domains 3, 5, 9, and 15, whereas IGF2 interacts with residues in domain 11. How a particular ligand affects the receptor's conformation or its ability to bind other ligands remains poorly understood. To address these questions, we purified a soluble form of the receptor from newborn calf serum, carried out glycoproteomics to define the N-glycans at its 19 potential glycosylation sites, probed its ability to bind lysosomal enzymes in the presence and absence of IGF2 using surface plasmon resonance, and assessed its conformation in the presence and absence of IGF2 by negative-staining electron microscopy and hydroxyl radical protein footprinting studies. Together, our findings support the hypothesis that IGF2 acts as an allosteric inhibitor of lysosomal enzyme binding by inducing global conformational changes of CI-MPR.

Lysosomal enzymes constitute a family of ~ 60 different acid hydrolases that degrade proteins, glycans, nucleic acids and other biomolecules. They play essential roles in autophagy by maintaining cellular energy homeostasis and in phagocytic processes by defending against microbial infections[1]. Most lysosomal storage diseases (LSDs) are caused by a mutation in a lysosomal enzyme, resulting in defective catabolism and substrate accumulation. LSDs are progressive diseases, and depending on the severity of the enzyme deficiency, patients can succumb to the disease in early childhood[2]. Of the ~ 70 LSDs, only a handful have FDA-approved treatments. Most of these treatments rely on enzyme replacement therapy (ERT) in which a recombinant enzyme is administered intravenously and becomes endocytosed by cation-independent mannose 6-phosphate receptors (CI-MPRs)[3,4].

CI-MPR participates in biosynthetic and endocytic pathways by trafficking between the trans Golgi network (TGN), cell surface, and endosomal compartments (Fig. 1A). It is best known for its role in lysosome biogenesis, an ancestral property of the receptor found in mollusks, starfish, and nearly all vertebrate species[5]. CI-MPR recognizes newly synthesized, mannose 6-phosphate (M6P)-containing lysosomal enzymes and transports them from the TGN to late endosomal compartments where the acidic pH causes the complex to dissociate. Lysosomal enzymes are packaged into lysosomes and the receptor recycles back to the TGN to repeat the process or moves to the plasma membrane[6]. Lysosomal enzymes that escape this biosynthetic targeting pathway and are secreted can be captured and internalized by cell surface CI-MPRs, a process used clinically in the treatment of LSDs with ERT[3,4].

[1]Department of Biochemistry, Medical College of Wisconsin, 8701 W. Watertown Plank Rd., Milwaukee, WI 53226, USA. [2]Department of BioMolecular Sciences, University of Mississippi, Oxford, MS 38677, USA. [3]The Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [4]Translational Metabolomics Shared Resource, Cancer Center, Medical College of Wisconsin, Milwaukee, WI 53226, USA. [5]Department of Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, Milwaukee, WI 53226, USA. [6]Department of Chemistry and Biochemistry, University of Mississippi, Oxford, MS 38677, USA. ✉email: ndahms@mcw.edu

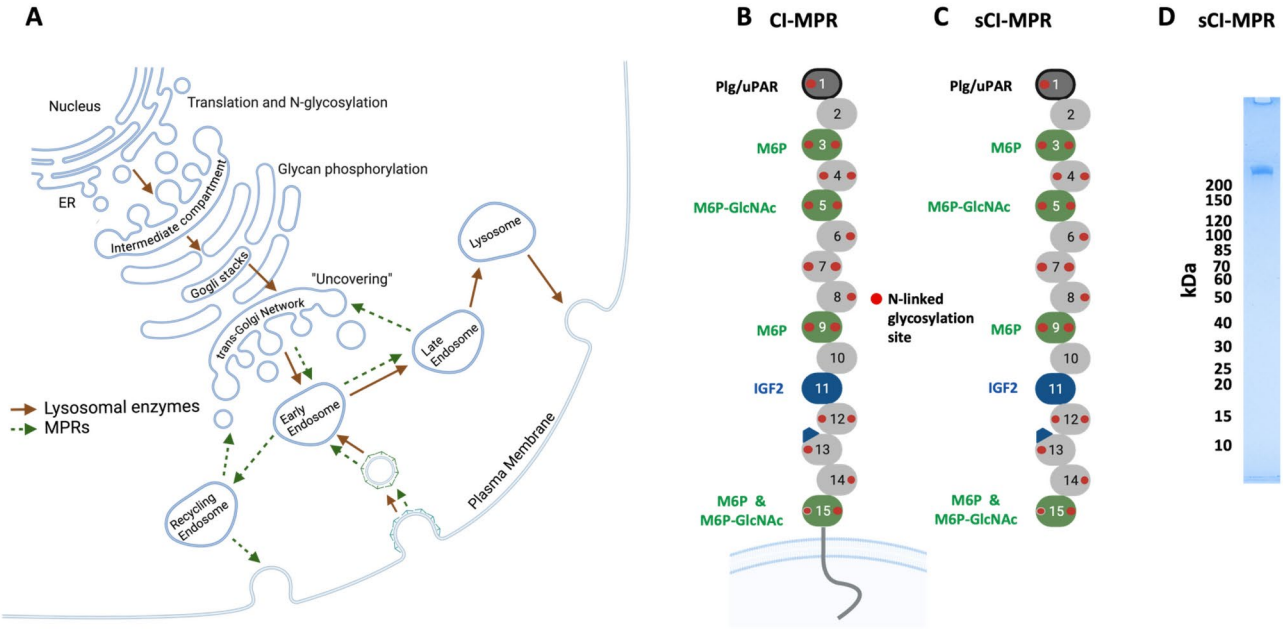

Fig. 1. Purification of sCI-MPR from newborn calf serum. (**A**) Schematic diagram of the trafficking of CI-MPR and lysosomal enzymes between the *trans*-Golgi network, endosomes, and cell surface. CI-MPR and lysosomal enzymes are synthesized in the endoplasmic reticulum (ER) and undergo co-translational N-glycosylation, and their N-glycans are modified in the Golgi. Phosphodiester-containing glycans are converted, or 'uncovered', forming phosphomonoesters. (**B**) Schematic diagram of the ~ 300 kDa full-length CI-MPR, a type I transmembrane protein containing 15 homologous domains in its extracellular region. The carbohydrate recognition domains (CRDs) are highlighted in green, showing domains specific for phosphomonoesters (M6P) and/or phosphodiesters (M6P-GlcNAc). The IGF2 binding site (blue) maps to domain 11 while domain 1 contains residues critical for interacting with plasminogen (Plg) or urokinase-type plasminogen activator receptor (uPAR). The 19 potential N-linked glycosylation sites are shown with red dots. (**C**) Schematic diagram of the soluble form of the receptor, sCI-MPR, found in extracellular fluids following its release at the plasma membrane by a metalloproteinase. (**D**) The soluble receptor was isolated from newborn calf serum by affinity chromatography. Purified sCI-MPR (1 µg) was resolved by SDS-PAGE on a 4–20% gradient polyacrylamide gel and visualized by staining with Coomassie blue G250. The migration of molecular weight markers is indicated. Created with https://www.biorender.com/.

In mammals, where it is also known as the insulin-like growth factor 2 receptor (IGF2R), CI-MPR acquired the ability to bind IGF2, a 7.5 kDa protein that lacks M6P[7–9]. CI-MPR internalizes extracellular IGF2 for delivery and degradation in lysosomes, thus limiting IGF2's proliferative effects mediated by tyrosine kinase receptors, the IGF1 receptor and the type A insulin receptor[10]. The downregulation of circulating levels of IGF2 by CI-MPR modulates embryonic growth and fetal size in mice[11,12]. In humans, evidence supports CI-MPR functioning as a tumor suppressor gene because the loss of heterozygosity of the receptor is linked with several cancers[13,14] and overexpression of CI-MPR in cancer cells reduces the rate of tumor growth[15,16]. Mechanistically, these findings in humans are consistent with CI-MPR's ability to lower extracellular levels of IGF2, thereby limiting the proliferation of tumor cells.

This multifunctional, ubiquitously expressed ~ 300 kDa membrane glycoprotein binds various ligands through its modular extracellular region. Six out of its 15 homologous domains have known ligand-binding properties: residues in domain 1 interact with plasminogen (Plg)[17,18] and urokinase-type plasminogen activator receptor (uPAR)[19], domains 3, 5, 9 and 15 are carbohydrate recognition domains (CRDs) that bind N-glycans bearing a phosphomonoester M6P and/or a phosphodiester M6P-GlcNAc moiety[3,20] and domain 11 houses the primary IGF2 binding site[21] (Fig. 1B). However, how each of CI-MPR's ligands influences the binding of another is not well understood.

Recently, Li and colleagues[22] used cryo-EM to determine the structure of bovine liver CI-MPR at pH 7.4 with IGF2 bound. When we analyzed their cryogenic electron tomography 3D reconstructions of domains 4–14 of CI-MPR bound to IGF2 (PDB 6UM2), we were surprised to find that the M6P binding site in domain 9 is occluded in the IGF2-bound conformation[23]. Due to the absence of structural information at the N- & C-terminus of the IGF2-bound receptor, the possibility remained that lysosomal enzymes could interact with the M6P binding sites located in domains 3 and 15. In the current report, we asked: (1) Whether the interaction of CI-MPR with lysosomal enzymes is inhibited by IGF2, and (2) Whether the binding of IGF2 induces a conformational change in the receptor.

To address these questions, we purified a soluble form of the receptor (sCI-MPR) from newborn calf serum (Fig. 1C,D), carried out glycoproteomics to define the N-glycans at its 19 potential glycosylation sites, probed its

ability to bind lysosomal enzymes in the presence and absence of IGF2 using surface plasmon resonance (SPR), and assessed its conformation in the presence and absence of IGF2 by negative-staining electron microscopy and hydroxyl radical protein footprinting (HRPF) studies. We demonstrate that IGF2, but not the closely related IGF1, allosterically inhibits the ability of sCI-MPR to bind lysosomal enzymes by inducing a global conformational change.

Results

Purification of sCI-MPR from bovine serum

To evaluate the allosteric properties and conformation of CI-MPR, we isolated the mammalian receptor from commercially sourced newborn bovine serum (Thermo Fisher Scientific/Gibco). Soluble forms of the receptor (sCI-MPR, Fig. 1C) are found in the serum of multiple species[24,25] and in the medium of several cultured cell types[26–28], with tumor necrosis factor α convertase (TACE, ADAM-17) identified as mediating the release of CI-MPR's extracellular region from human endothelial cells[28]. We purified sCI-MPR from newborn calf serum by affinity chromatography using resin coupled with a lysosomal enzyme. The purified protein migrated on a SDS polyacrylamide gel as a single, high molecular weight band (Fig. 1D). To probe the sequence of the isolated receptor, tryptic peptides were analyzed by mass spectrometry. The results from three independent purifications show the N-terminus begins with Gly, which is the third amino acid following the predicted cleavage site of the signal sequence (Fig. 2). In contrast, the peptides identified at the C-terminus were more variable, ending 15-, 51-, or 53-residues from the start of the predicted transmembrane region (Fig. 2). This variability may be due, in part, to limitations in detecting the predicted large or small tryptic peptides in this region of the receptor, thus negating our ability to define the exact C-terminus of sCI-MPR. Taken together, these findings indicate that sCI-MPR exists as a stable protein in bovine serum, with an intact N-terminus at domain 1 and at least most of its C-terminal domain 15.

To determine the extent and type of N-glycan modification carried by the bovine serum sCI-MPR, N-linked glycosylation at each glycosite of the receptor was profiled by mass spectrometry (MS). sCI-MPR contains 19

Fig. 2. Analysis of sCI-MPR by mass spectrometry. The affinity-purified sCI-MPR protein isolated from newborn calf serum was digested with trypsin in solution as described in Methods. Results from three separate purifications using two different lots of serum are shown and the identified peptides are highlighted in green, aqua, or yellow in the amino acid sequence of the full-length bovine protein. The 44-residue N-terminal signal sequence is underlined, the triplet sequences identifying potential N-linked glycosylation sites are highlighted in red, the conserved 13-residue sequence at the C-terminus of each of the 15 domains is underlined, and the predicted transmembrane region is boxed.

potential N-linked sites, which are found within the known "N-X-S/T" N-glycosylation consensus sequence[29]. Among the 19 potential N-linked sites, 17 of the 19 sites were identified by a MS workflow (Fig. 3A), which uses a ^{18}O-labeling technique and specific endoglycosidases to introduce mass signatures to distinguish N-glycan types and to quantify each site's occupancy[30]. These analyses demonstrate that most of the sites are either fully or highly glycosylated with complex-type N-glycans, and four of the N-linked sites (N590, N959, N1030 and N1321) carry high-mannose or hybrid-type glycans without core fucosylation as their major glycan components (Fig. 3B). The two remaining sites, N409 and N2094 (indicated with asterisks in Fig. 3B) were not detected from samples treated with EndoF1/PNGase F/^{18}O-water. Instead, they were detected as N-glycopeptides from the tryptic digests, and the analyses indicate that N409 carries high-mannose type glycans while N2094 carries complex-type glycans as major components.

SPR analysis of sCI-MPR's interaction with IGF2 and lysosomal enzymes

We performed SPR experiments to validate that the sCI-MPR purified from bovine serum was functional and capable of binding its ligands, IGF2 and lysosomal enzymes. SPR analyses showed that sCI-MPR binds tightly to immobilized biotinylated IGF2 (K_d = 1.2 nM, Fig. 4A), a value similar to that reported for the interaction of IGF2 with the full-length and soluble forms of the receptor[25,31]. Next, we tested the binding of the receptor to two different lysosomal enzymes: acid α-glucosidase (GAA), a ~ 110 kDa enzyme with 7 N-glycans, and palmitoyl protein thioesterase (PPT1), a ~ 35 kDa enzyme with 3 N-glycans that we showed contains predominantly

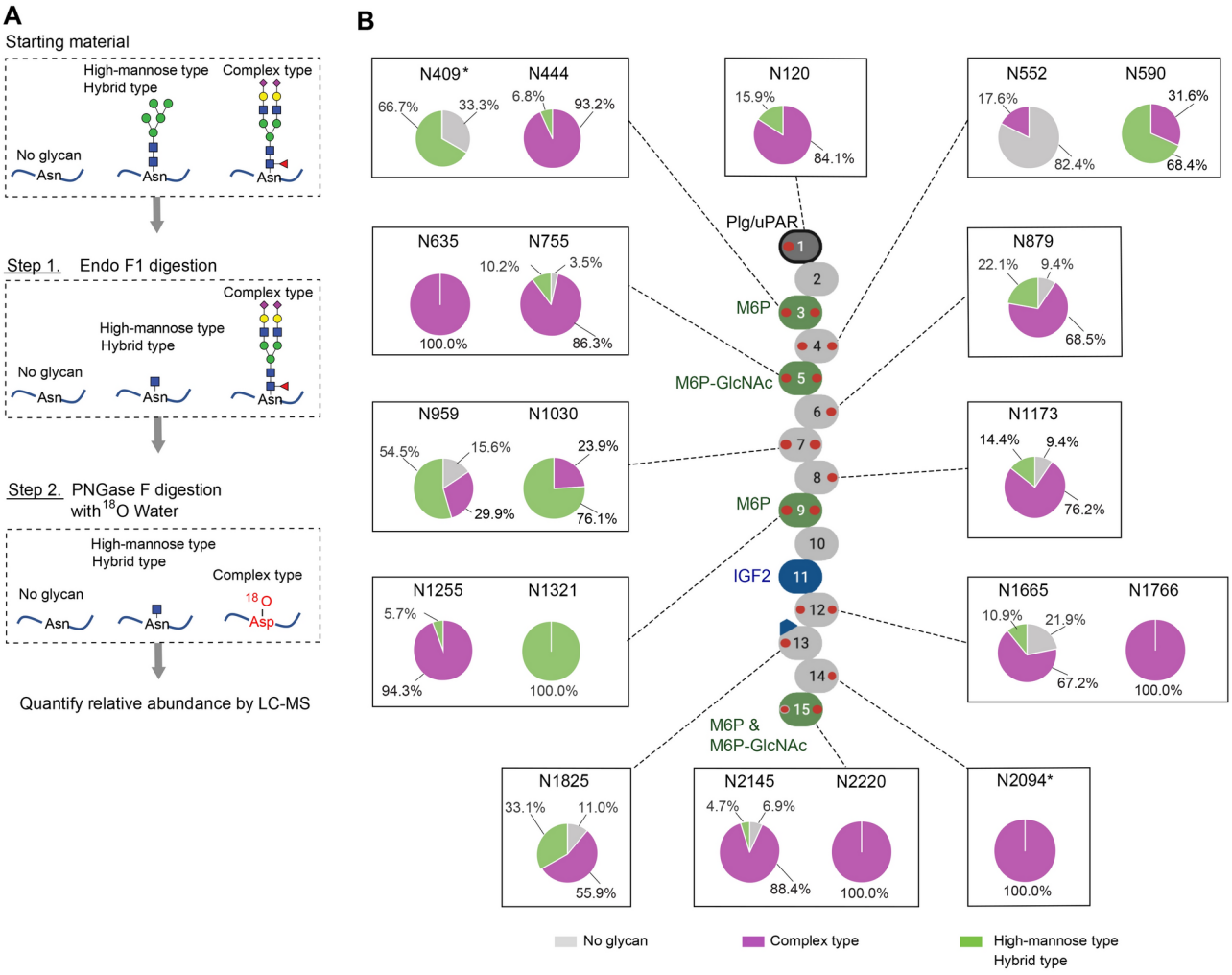

Fig. 3. Site-specific heterogeneity of N-linked glycosylation on sCI-MPR. (**A**) A schematic representation of the mass spectrometry (MS) workflow performed for the site-specific N-glycosylation analysis of sCI-MPR using the method of Cao, et al.[30]. Briefly, tryptic digests from sCI-MPR were further treated sequentially with Endo F1 followed by PNGase F/^{18}O-water to produce unique mass signatures to identify and quantify the types of N-glycans and the occupancy at each of the 19 potential N-linked sites. (**B**) Pie charts illustrate relative percentages of unoccupied (gray) or occupied by high-mannose/hybrid type without core fucosylation (green) or complex-type (purple) N-glycans at each N-linked glycosylation site of sCI-MPR. The MS workflow presented in panel A was used to characterize 17 of the 19 N-linked sites, whereas tryptic digests were used to characterize glycopeptides at the remaining two sites at N409 and N2094 (shown with asterisks).

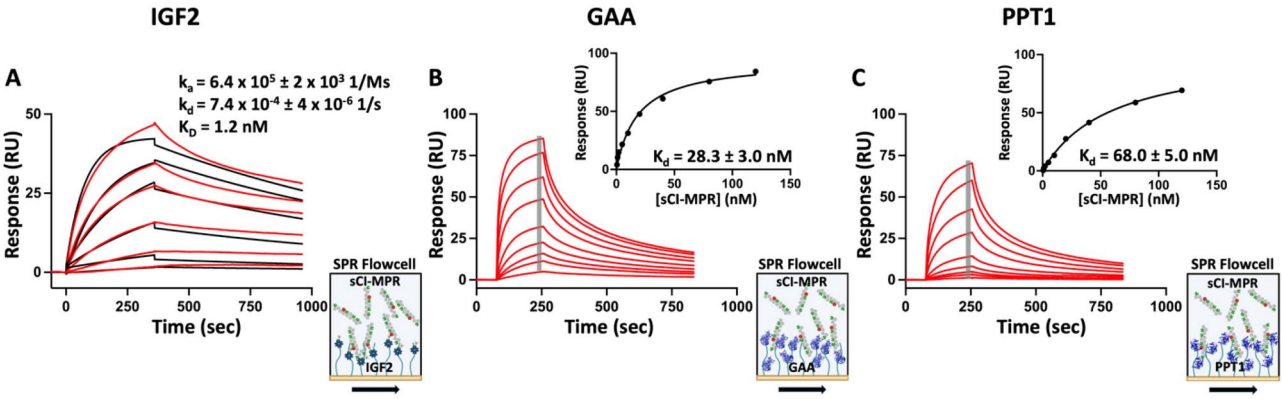

Fig. 4. Binding affinity at pH 7.4 of IGF2 and the lysosomal enzymes GAA and PPT1 to sCI-MPR as assessed by SPR. (**A**) Biotinylated IGF2 (GroPep, Inc.) was immobilized on the surface of a streptavidin-coupled SA sensor chip. Representative sensorgrams (red) of sCI-MPR protein (0.5, 1, 2, 5, 10, and 20 nM) flowed over the IGF2 surface are shown for a single experiment. The resulting curves from three independent experiments were fit to a 1:1 binding model (Biacore S200 evaluation software). The fitted curves are displayed in black, and the resulting kinetic constants are shown ± SE. Human lysosomal enzymes (**B**) GAA and (**C**) PPT1 were immobilized onto the surface of a CM5 sensor chip by amine-coupling. Representative sensorgrams of sCI-MPR (0.2, 0.5, 1, 2, 5, 10, 20, 40, 80, and 120 nM) flowed over each surface are shown. The binding levels (response units, RU) were determined as the average response 4 s before the end of the association phase (vertical gray bar in panels **B**,**C**) of each injection using the Biacore S200 evaluation software (Cytiva, Inc.). The binding levels (RU) for the SPR experiments were plotted against the concentration of receptor (inset graph, panels **B**,**C**) and fitted to a one-site specific binding equation as described in Methods. K_d ± SEM is shown for 10 independent SPR experiments representing four purifications of sCI-MPR from two lots of newborn calf serum. Illustrations created using https://www.biorender.com/.

phosphomonoesters[23]. A deficiency of GAA causes Pompe disease and existing FDA treatments, namely enzyme replacement therapy, significantly extends the lives of these patients who would otherwise succumb to cardiorespiratory failure in early childhood[32]. Mutations in PPT1 lead to a fatal neurodegenerative disease called infantile neuronal ceroid lipofuscinosis for which no FDA approved treatments exist for these patients[33]. Increasing concentrations of sCI-MPR were flowed over the GAA-immobilized surface and the responses were measured for each sample. The response at equilibrium was plotted against receptor concentration and data were fit with a one-site specific binding model, demonstrating that sCI-MPR bound with high affinity to GAA (K_d = 28.3 ± 3.0 nM, Fig. 4B). Similar experiments were performed using sensor chips containing immobilized PPT1. The receptor bound with a slightly poorer affinity to PPT1, with a K_d calculated to be 68.0 ± 5.0 nM (Fig. 4C). In summary, sCI-MPR purified from bovine serum is active and capable of binding IGF2 and phosphomannosyl-containing lysosomal enzymes GAA and PPT1 with high affinity.

IGF2 inhibits the interaction between sCI-MPR and the lysosomal enzymes GAA and PPT1

To test the hypothesis that IGF2 acts as an allosteric regulator of lysosomal enzyme binding by sCI-MPR, further SPR analyses were conducted. In these experiments, performed at pH 7.4, the receptor was preincubated with increasing concentrations of IGF2 or the control IGF1, which is highly homologous to IGF2 but does not bind CI-MPR[34], before the receptor was flowed over the immobilized lysosomal enzyme. Preincubation with increasing concentrations of IGF2 led to a progressive decrease in response to GAA (Fig. 5A). A 66.3 ± 1.1% ($P < 0.0001$) reduction in maximal receptor binding was observed at the highest concentration of IGF2 (100 nM), whereas no significant decrease in GAA binding was observed with 100 nM IGF1 (Fig. 5A, inset and Fig. 5B). An IC_{50} of 5.5 nM was determined for IGF2 inhibition of GAA binding to sCI-MPR (Fig. 5B). A more striking finding was observed using the PPT1-immobilized surface. Like GAA, preincubation with increasing concentrations of IGF2 resulted in a progressive decrease in response to PPT1 (Fig. 5C). However, a more dramatic reduction (92.6 ± 0.1%, $P < 0.0001$) in maximal receptor binding was observed at the highest IGF2 concentration (100 nM), while no significant decrease in PPT1 binding was observed in the presence of 100 nM IGF1 (Fig. 5C, inset and Fig. 5D). An IC_{50} of 5.5 nM was calculated for IGF2 inhibition of PPT1 binding to sCI-MPR (Fig. 5D). Similar findings were observed when the experiments were conducted at pH 6.8 (Fig. S1), with IC_{50} of 6.7 nM (Fig. S1B) and 8.0 nM (Fig. S1D) calculated for IGF2 inhibition of GAA and PPT1 binding, respectively. Together, these findings demonstrate that IGF2, but not IGF1, allosterically inhibits sCI-MPR binding to the lysosomal enzymes GAA and PPT1.

IGF2 enhances the dissociation of sCI-MPR from GAA and PPT1

To further probe the effect IGF2 has on the receptor's interaction with lysosomal enzymes, a complementary SPR study was performed. In these experiments performed at pH 7.4, the receptor was flowed over the lysosomal enzyme-immobilized surfaces to allow binding. Running buffer alone, without receptor but with increasing concentrations of IGF2 or the control IGF1, was then flowed for 2 min over the sensor chip surface and the

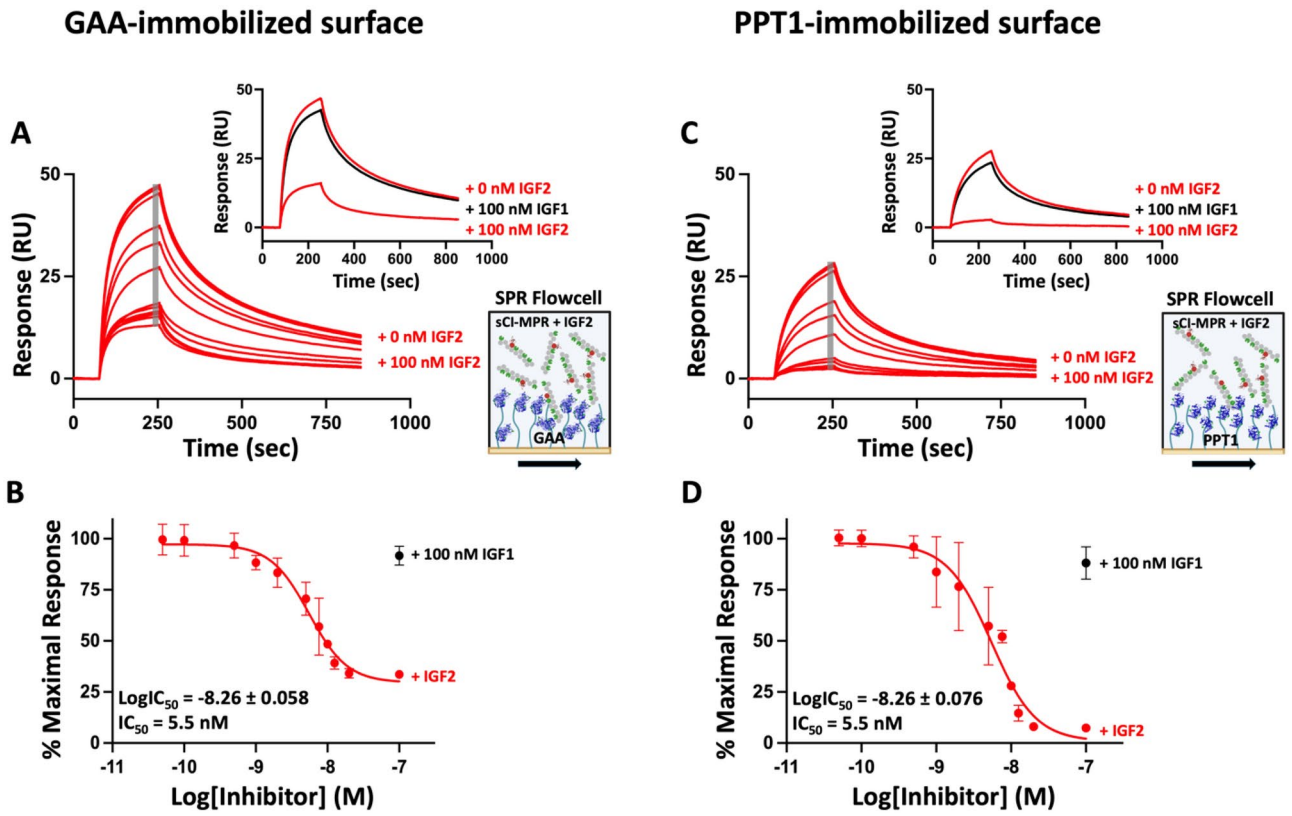

Fig. 5. SPR analysis at pH 7.4 of sCI-MPR incubated with IGF2 prior to interaction with lysosomal enzymes. Samples of sCI-MPR at 20 nM were incubated with IGF2 (0, 0.01, 0.05, 0.1, 0.5, 1, 2, 5, 7.5, 10, 15, 20 and 100 nM) or IGF1 (100 nM) for 1 h at room temperature before flowing the mixture over the GAA- (**A**,**B**) or PPT1- (**C**,**D**) immobilized surfaces. (**A**,**C**) Representative sensorgrams are shown, with insets comparing the sensorgrams of sCI-MPR with IGF2 (0 and 100 nM, red) and IGF1 (100 nM, black). (**B**,**D**) Plotted are the maximal response units (RU, vertical gray bar in panels **A**,**C**) as the mean of the percentage of maximal RU ± SD for three independent SPR experiments, with 0 nM IGF2 set at 100% maximal binding. Data are fit to a 4-parameter dose response curve using GraphPad Prism v 10.2.0. $LogIC_{50}$ values ± SEM are shown. Illustrations created using https://www.biorender.com/.

dissociation of the receptor from the lysosomal enzyme was measured. The response, measured 4 s prior to the end of the 2 min IGF2 injection, showed that increasing concentrations of IGF2 resulted in a progressive enhancement of receptor dissociation from GAA (Fig. 6A). Compared to 0 nM IGF2, a 75.2 ± 3.5% ($P < 0.0001$) reduction in maximal binding of the receptor during the dissociation phase was observed at the highest concentration of IGF2 (100 nM), whereas no significant decrease was observed in the presence of 100 nM IGF1 (Fig. 6A, inset and Fig. 6B). An IC_{50} of 1.1 nM was calculated for IGF2 enhancement of sCI-MPR's dissociation from GAA (Fig. 6B). Again, a more striking finding was observed for PPT1. Like GAA, increasing concentrations of IGF2 resulted in a progressive enhancement of receptor dissociation from PPT1 (Fig. 6D). However, a 94.0 ± 4.4% ($P < 0.0001$) reduction in maximal binding of the receptor during the dissociation phase was observed at the highest concentration of IGF2 (100 nM), whereas no significant decrease was observed in the presence of 100 nM IGF1 (Fig. 6D, inset and Fig. 6E). An IC_{50} of 1.1 nM was calculated for IGF2 enhancement of sCI-MPR's dissociation from PPT1 (Fig. 6E). Similar findings were observed when the experiments were carried out at pH 6.8 (Fig. S2), with IC_{50} of 1.1 nM (Fig. S2B) and 1.4 nM (Fig. S2D) calculated for IGF2 enhancement of the receptor's dissociation from GAA and PPT1, respectively. The rate of dissociation, k_{off}, during the second phase of the injection (IGF2 only) was estimated using a one-site model in BiaEvaluation 4.1. The GAA-immobilized surface showed that IGF2 increased the rate of dissociation at concentrations ≥ 5 nM (Fig. 6C). Similarly, the PPT1-immobilized surfaced showed IGF2 increased the rate of dissociation at concentrations ≥ 7.5 nM (Fig. 6F). In contrast, IGF1 (100 nM) had no effect on the k_{off} rate (Fig. 6C,F). Similar findings were observed when the experiments were performed at pH 6.8 (Fig. S2C,F). In summary, IGF2, but not IGF1, enhances the dissociation of sCI-MPR from the lysosomal enzymes GAA and PPT1.

sCI-MPR assumes a compact conformation upon binding IGF2

To investigate whether sCI-MPR undergoes a conformational change when bound to IGF2, we conducted transmission electron microscopy (TEM) and visualized the individual receptor molecules by negative-staining the sample preparation[35]. sCI-MPR was prepared in the absence (Fig. 7A–C) and presence (Fig. 7D–F) of IGF2.

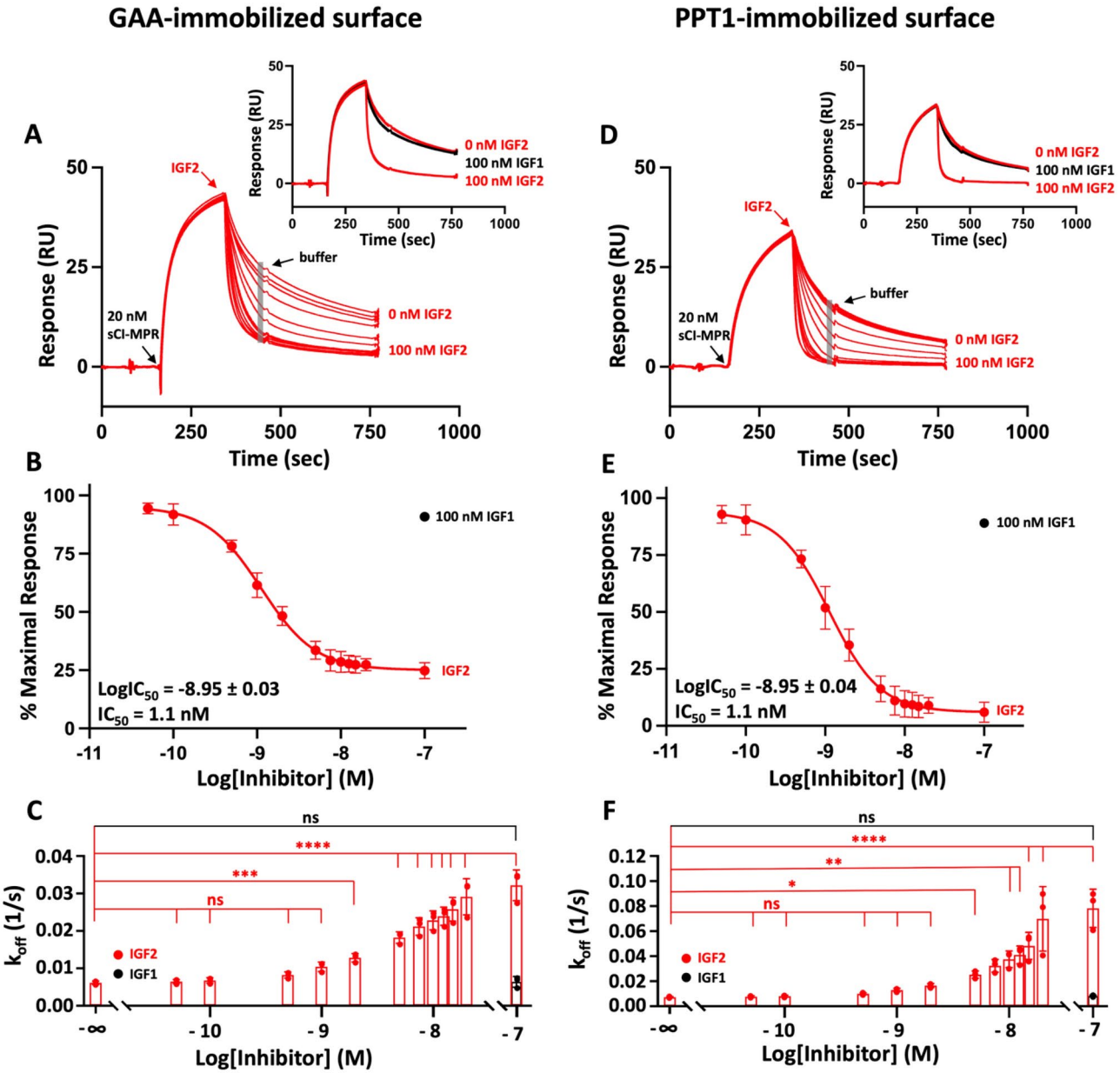

Fig. 6. SPR analysis at pH 7.4 of sCI-MPR dissociation from lysosomal enzymes in the presence and absence of IGF2. Samples of 20 nM sCI-MPR were flowed over the GAA- or PPT1-immobilized sensor chip surface for 3 min followed by a second injection (initiation of dissociation phase) of buffer without receptor but containing increasing concentrations of IGF2 (0, 0.05, 0.1, 0.5, 1, 2, 5, 7.5, 10, 12.5, 15, 20, and 100 nM) or IGF1 (100 nM) for 2 min, and buffer alone was then flowed over the surface for the final 5 min. (**A**,**D**) Representative sensorgrams are shown, with insets comparing the sensorgrams of sCI-MPR with IGF2 (0 and 100 nM, red) and IGF1 (100 nM, black). (**B**,**E**) Plotted are the response units (RU) measured 4 s before the end of the IGF2 injection (vertical gray bar in panels (**A**,**D**)). Percent maximal response was calculated by dividing the response by the response at 0 nM IGF2 multiplied by 100. The normalized responses ± SD for three independent SPR experiments were plotted against the log of the IGF2 concentration. Data are fit to a 4-parameter dose response curve using GraphPad Prism v 10.2.0. $LogIC_{50}$ values ± SEM are shown. (**C**,**F**) Plotted are the k_{off} rates during the 2 min injection of buffer without receptor but containing increasing concentrations of IGF2 (0, 0.05, 0.1, 0.5, 1, 2, 5, 7.5, 10, 12.5, 15, 20, and 100 nM) or IGF1 (100 nM). The k_{off} rates ± SD for three independent SPR experiments were plotted against the log of the IGF2 concentration. Statistical analyses (ordinary one-way ANOVA, GraphPad Prism v 10.2.0) were performed comparing each concentration of IGF2 (red) or IGF1 (black) to 0 nM IGF2. ns, not significant ($P \geq 0.05$), *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$.

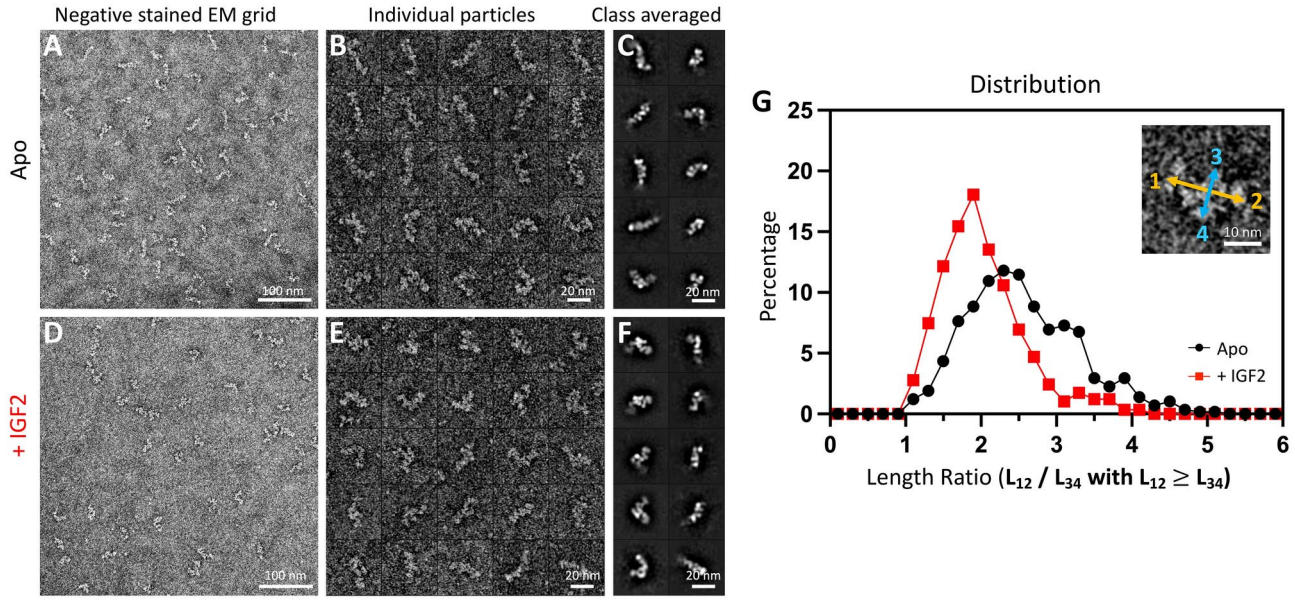

Fig. 7. Negative-staining electron microscopy (EM) of sCI-MPR with and without IGF2. Negative-staining EM studies of sCI-MPR were performed using 1.4 μM of the receptor prepared in the absence and presence of 200 nM IGF2 at pH 6.8. Survey micrograph of sCI-MPR in the absence (**A**) and presence (**D**) of IGF2. Twenty-five representative images of sCI-MPR particles without (**B**) and with (**E**) IGF2. Ten representative reference-free class averaged images of the particles without (**C**) and with (**F**) IGF2. (**G**) The dimensions of 576 individual particles from each condition were measured. The length ratio of each particle was calculated and plotted as a function of the percentage of the total population of particles. This graphical representation highlights the conformational differences of sCI-MPR in the absence and presence of IGF2.

We measured the dimensions of 576 individual particles from each sample and determined the length ratio of each particle (Fig. 7G). In the absence of ligand, the apo form of the receptor displays a striking, broad distribution of size that range from a symmetrical particle (length ratio = 1) to extended conformations (length ratio = 4.5), with two major peaks centered at a length ratio of 2.3 and 3.1 (Fig. 7G). In contrast, the IGF2-bound receptor adopts a more compact conformation, with a single predominant peak centered at a length ratio of 1.9 (Fig. 7G). These findings indicate that the overall conformation of sCI-MPR is significantly influenced by IGF2 binding, with the receptor becoming more compact in size upon binding IGF2 compared to the apo form at pH 6.8.

IGF2 binding induces conformational changes in sCI-MPR distal from the IGF2 binding site

To further investigate the region(s) of the receptor that undergo conformational changes upon IGF2 binding, we conducted HRPF using Fox photolysis system[36]. This technique compares protein topography between two structural states based on changes in the reaction rates of portions of the protein with diffusing hydroxyl radicals. Briefly, the intact receptor protein is reacted with very short-lived hydroxyl radicals, which are generated in situ by flash photolysis of hydrogen peroxide. These radicals diffuse to the surface of the protein and oxidize amino acid side chains forming stable protein oxidation products. The rate of this oxidation reaction is directly proportional to the solvent accessible surface area, with exposure causing an increase in the rate of oxidation and occlusion causing a decrease in the rate of oxidation[37]. We have used HRPF previously to show that a construct encoding domains 1–5 of CI-MPR undergoes conformational changes upon lysosomal enzyme binding or acidification of pH conditions[23]. In the current report, we compared the topography of sCI-MPR at pH 7.4 in the absence and presence of IGF2. Three independent oxidation experiments were performed, with each experiment using triplicate analyses. Due to the large number of peptides in sCI-MPR, a standard t-test would have a large propensity for false positives, and a standard correction for multiple testing would greatly reduce the power of our test. Therefore, we chose to use a two-tailed t-test of each experiment and then select only those peptides that showed significant ($P < 0.05$) changes in oxidation in the same direction in at least two of the three experiments as notable for our analysis.

After oxidation and quenching, the samples were digested with trypsin and chymotrypsin separately to increase sequence coverage, and the modifications were measured by LC–MS/MS (Fig. S3). Peptides showing notable changes in oxidation were mapped onto the cryo-EM structure of the full-length bovine CI-MPR (Fig. 8). These HRPF studies revealed that two peptides of the receptor became protected from hydroxyl radical oxidation upon binding IGF2, and these two peptides (1113–1125 and 1552–1558) map near the IGF2 binding site (Fig. 8). These findings validate the approach because regions near the IGF2 binding site are expected to experience a decrease in solvent accessible surface area in the presence of bound IGF2. Importantly, four peptides (60–72, 233–244, 569–586, and 719–728) that became more exposed upon IGF2 binding are located

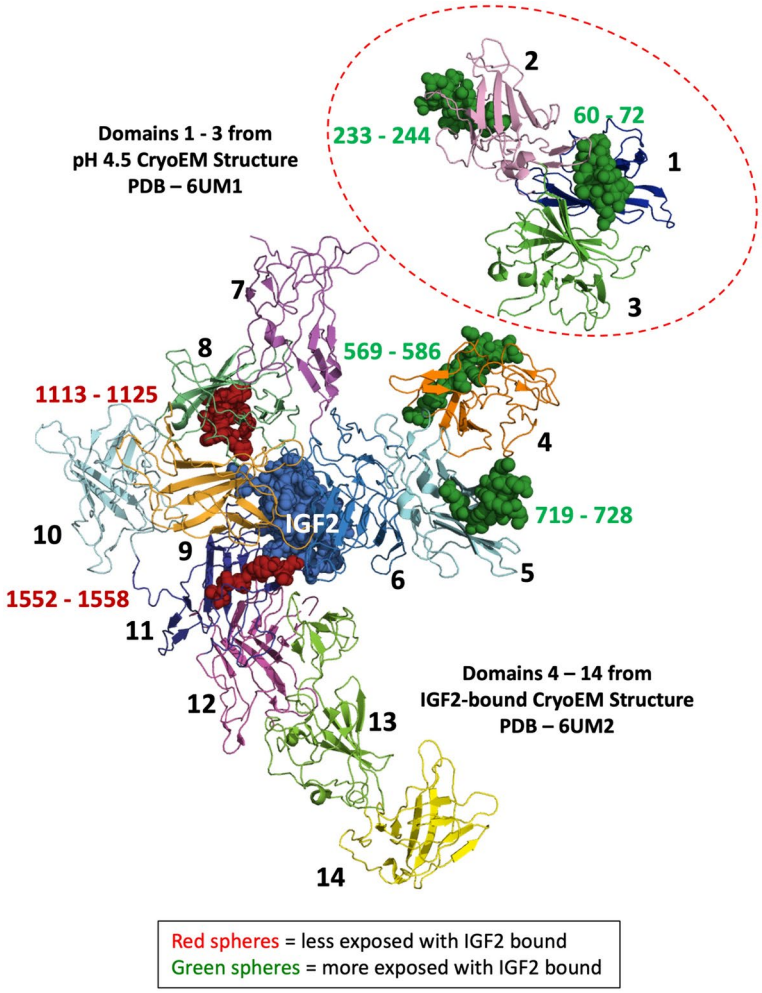

Fig. 8. HRPF analysis of topographical changes of sCI-MPR (1 μM) upon binding IGF2 (1 μM) at pH 7.4. Peptides showing increases in exposure (green spheres) and protection (red spheres) upon IGF2 binding in at least two of the three independent oxidation experiments are mapped onto the cryo-EM structures of CI-MPR. Because there is insufficient resolution of the N-terminal 3 domains of the receptor bound to IGF2 in the cryo-EM structure at pH 7.4 (PDB 6UM2), peptides located in domains 1–3 were mapped onto the receptor in the ligand unbound state using the cryo-EM structure at pH 4.5 (PDB 6UM1). Bound IGF2 is shown in blue spheres. Created using PyMOL Molecular Graphics System, Version 2.5.4 Schrödinger, LLC.

near two (domain 3 and domain 5) of the four CRDs and are distal from the IGF2 binding site in domain 11 (Fig. 8). Together, these data demonstrate sCI-MPR is structurally dynamic and adopts various conformations influenced by IGF2 binding.

Discussion

CI-MPR binds ligands in a pH-dependent manner, interacting optimally at the pH of the Golgi (~pH 6.8), retaining binding ability at the cell surface, and releasing its cargo in the acidic environment of late endosomes[6] (Fig. 1A). The 15 domains that comprise CI-MPR's ~ 2300-residue extracellular region have a similar β-barrel fold and are connected by short linkers[20,22], with the known ligand binding sites mapping to non-adjacent domains (Fig. 1B). Using a recombinant construct expressed in Sf9 insect cells, we previously demonstrated that the N-terminal five domains of CI-MPR exhibit dramatic changes in domain configuration upon carbohydrate binding or exposure to different pH conditions. Unexpectedly, we discovered allosteric interactions between the two CRDs located in domain 3 and domain 5[23]. Data from our study indicate that binding of PPT1's phosphomonoester-containing N-glycans to domain 3 causes a reorientation of domain(s) to prevent domain 5 from interacting with a phosphodiester-containing lysosomal enzyme[23].

Do the conformational changes observed in CI-MPR domains 1–5 produced in Sf9 insect cells reflect the dynamic properties of the native, full-length receptor? Sf9 insect cells produce recombinant proteins with short N-glycan chains of the pauci-mannose type ($Man_3GlcNAc_2$)[38]. These shorter glycan chains may allow greater flexibility compared to the native CI-MPR, which has been reported to possess longer glycans including sialylated complex-type glycans[39,40]. Additionally, when domain 5 is tethered to domain 6 within the context of the full-

length protein, interactions with the receptor's C-terminal region(s) could diminish the overall dynamics and extent of conformational flexibility of CI-MPR in response to changes in ligand occupancy and pH conditions. However, recent studies indicate that the full-length CI-MPR can adopt different conformations[22]. In these experiments, Li et al. described two structures determined by cryo-EM of the CI-MPR purified from bovine liver: (1) an IGF2-bound form at pH 7.4, and (2) a ligand-unbound form at pH 4.5, with CI-MPR at pH 4.5 being more compact than the IGF2-bound structure[22]. The ability to adopt a more compact structure upon acidification appears to be an inherent property of CI-MPR: as measured by size exclusion chromatography we observed recombinant forms of CI-MPR (domains 1–5, 7–15 and 1–15) expressed in Sf9 insect cells each exhibit a decreased Stokes radius at pH 4.5 when compared to the recombinant constructs at pH 6.5[23].

In the current report, we investigated whether IGF2 acts as a negative allosteric regulator of CI-MPR's ability to interact with lysosomal enzymes. This hypothesis came to light during our analysis of the cryogenic electron tomography 3D reconstructions of domains 4–14 of CI-MPR bound to IGF2 (PDB 6UM2)[22] which revealed that the M6P binding site in domain 9 is occluded in the IGF2-bound conformation[23]. However, the possibility remained that lysosomal enzymes could interact with the CRDs located in domain 3 and/or 15, regions not defined in the IGF2-bound cryo-EM structure[22].

To test this hypothesis, we used native sCI-MPR purified from bovine serum. Before performing functional studies, we first characterized the purified receptor using mass spectrometry. We demonstrate sCI-MPR contains an intact N-terminus along with most of its C-terminal domain 15 (Fig. 2). We also reveal that native sCI-MPR is heavily N-glycosylated, with the majority of its 19 potential N-linked sites being fully or highly glycosylated with complex-type N-glycans (Fig. 3B). To our knowledge, this is the first site-specific N-glycan analysis of sCI-MPR from newborn calf serum.

We used two well-characterized lysosomal enzymes as representative examples of CI-MPR's diverse repertoire of acid hydrolases it targets to lysosomal compartments: PPT1, a ~ 35 kDa enzyme with 3 N-glycans containing predominantly phosphomonoesters[23], and GAA, a ~ 110 kDa enzyme containing 7 N-glycans with each GAA containing at least one bis-phosphorylated glycan structure for high-affinity binding to CI-MPR[41]. SPR analyses demonstrate that IGF2, but not IGF1, inhibits sCI-MPR binding to PPT1 and GAA (Fig. 5 and Fig. S1) and enhances the dissociation of the receptor from the lysosomal enzymes (Fig. 6 and Fig. S2). We observed that the effects of IGF2 are more pronounced with PPT1 compared to GAA (92.6% versus 66.3% reduction in maximal binding in preincubation studies, respectively, and 94.0% versus 75.2% reduction in maximal binding during the dissociation phase, respectively).

What could contribute to this differential effect of IGF2 on these two lysosomal enzymes? The CRD(s) where GAA and PPT1 interact with sCI-MPR is unknown. The higher number of phosphorylated glycans and/or the presence of bis-phosphorylated glycans on GAA likely contribute to its higher binding affinity to sCI-MPR compared to PPT1 (K_d = 28.3 nM and 68.0 nM, respectively, Fig. 4) and may increase the chances of GAA binding to two or more of the receptor's CRDs, whereas PPT1 may interact with only a single CRD. Furthermore, the binding of IGF2 to domain 11 may allosterically affect the four CRDs differently. For example, as indicated by the IGF2-bound cryo-EM structure[22], the M6P binding pocket in domain 9 is occluded, predicting a complete blockage of any interaction with a lysosomal enzyme. In contrast, IGF2-mediated allosteric changes of domain 3 could be less dramatic, causing a decrease in carbohydrate binding affinity of tenfold or less. Additionally, a subpopulation of GAA may harbor a higher percentage of phosphodiester-containing glycans than PPT1 to facilitate binding to CI-MPR's CRDs that prefer phosphodiester glycans (domains 5 and 15), thereby potentially allowing the simultaneous binding of IGF2 and phosphodiester-bearing GAA to CI-MPR. To begin to define the mechanism of IGF2's inhibitory effects on lysosomal enzyme binding by CI-MPR, we took two approaches to probe conformational changes in sCI-MPR. First, negative-staining TEM studies comparing the size of the sCI-MPR in the presence and absence of IGF2 at pH 6.8 show that the IGF2-bound structure adopts a more compact configuration. Second, HRPF experiments conducted with sCI-MPR in the presence and absence of IGF2 detected changes in side-chain oxidation in peptides near the IGF2-binding site in domain 11, as expected, and in the distal N-terminal region of the receptor. Together, our results support the hypothesis that the ability of IGF2 to inhibit GAA and PPT1 binding results from global conformational changes in CI-MPR induced by IGF2 binding. Future studies are necessary to interrogate at the functional and structural levels how each of the four CRDs is affected when IGF2 binds CI-MPR.

Our assessment of the overall conformation of sCI-MPR by single particle TEM (Fig. 7) reveals a surprisingly broad size distribution of the ligand-unbound receptor at pH 6.8, which is the pH of the Golgi and the optimal pH for CI-MPR to bind lysosomal enzymes. To our knowledge, this is the first view of the overall conformation of CI-MPR's extracellular region in the apo state and at a neutral pH. Lysosomal enzymes represent a heterogenous cargo that differ in their size and spatial presentation of phosphorylated glycans. Consistent with CI-MPR functioning in the TGN to capture ~ 60 different lysosomal enzymes for delivery to endosomal compartments, it would be advantageous for the receptor to be dynamically flexible to optimize the probability of interacting with similarly dynamic and inherently flexible phosphorylated glycans on lysosomal enzymes.

We recognize the limitations of our study. Although we have shown that IGF2 binding to sCI-MPR inhibits the receptor's ability to bind PPT1 and GAA, we have not determined if this inhibitory effect holds true for the remaining ~ 60 lysosomal enzymes that the receptor delivers to lysosomes. Addressing this issue is challenging because not all of these hydrolytic enzymes are readily accessible for purification in significant quantities, either from native sources or as a recombinantly expressed protein, and importantly they must contain phosphorylated N-glycans in order to interact with CI-MPR. Additionally, we have not examined the function of CI-MPR at the plasma membrane. Therefore, further investigation is needed to determine if the full-length, membrane associated form of CI-MPR when bound to IGF2 is similarly inhibited from binding lysosomal enzymes.

What could be the biological significance and consequences of IGF2 acting as a negative regulator of lysosomal enzyme binding to CI-MPR? CI-MPR bound to IGF2 at the plasma membrane would not be able to function

in the re-uptake of secreted lysosomal enzymes. Similarly, IGF2-bound sCI-MPR found in extracellular fluids could not act as a sink for secreted lysosomal enzymes. Together, these IGF2-mediated effects on the functioning of cell surface and soluble CI-MPRs would result in more hydrolytic enzymes being present in extracellular fluids. IGF2 is synthesized as a pre-pro-protein that is processed in the ER and Golgi, with endoproteolysis occurring in the Golgi to generate the mature IGF2 peptide[42]. Thus, it is possible that IGF2 could bind CI-MPR in the Golgi and/or TGN leading to aberrant secretion of lysosomal enzymes. It is also possible that the cell could respond by elevating levels of CD-MPR, thereby partially compensating for an IGF2-inhibited CI-MPR to mitigate aberrant secretion of lysosomal enzymes. It is also possible that the type of N-glycans (high-mannose, hybrid, complex) and their location in the CI-MPR structure could affect direct IGF2 binding and/or the ability of CI-MPR to adopt specific conformations, thereby modulating allosteric regulation by IGF2 and aberrant lysosomal enzyme secretion. Future experiments are needed to examine the role N-glycans have on CI-MPR function. Some cancer cells have been reported to overexpress and secrete IGF2[43]. Beyond its proliferative effects, IGF2's ability to inhibit CI-MPR's binding of lysosomal enzymes may provide another selective advantage to cancer cells by facilitating a higher level of hydrolytic enzymes in the surrounding extracellular environment, thereby enhancing the invasiveness of the cancer cell by degrading components of the extracellular matrix. Thus, we hypothesize that modulating IGF2 expression is a mechanism a cell can use to regulate the level of hydrolytic enzymes present in its extracellular environment.

IGF2 is secreted mainly by the placenta during pregnancy, and postnatally the liver is the major source of IGF2 in the circulation. IGF2 is an imprinted gene and multiple mechanisms have been reported to regulate its expression, including *H19/IGF2:IG-DMR* methylation and the direct activation of IGF2's promoters by several transcription factors[44]. Additionally, IGF2 is overexpressed in Wilms' tumors due to loss of heterozygosity or loss of imprinting[45,46]. Thus, different processes can alter the local cellular and/or circulating levels of IGF2. Our findings in the current report may have clinical implications in the treatment of LSDs by ERT. The efficiency of ERT would be lowered in cases where either circulating IGF2 levels or IGF2 secreted by specific tissues or cell types are elevated due to the binding of IGF2 to cell surface CI-MPRs, resulting in the loss of CI-MPR-mediated internalization of the IV-administered recombinant lysosomal enzyme. Studies by Sly and Lebowitz[47] have shown that a lysosomal enzyme with a glycosylation-independent lysosomal tag (GILT) that uses a modified IGF2 with reduced affinity to the IGF1 receptor can be internalized by CI-MPR and is effective in reducing substrate accumulation in mouse models of LSDs[48,49]. Lysosome-targeting chimeras (LYTACs) have been developed for the targeted lysosomal degradation of secreted and membrane proteins, some of which incorporate phosphomannosyl-type moieties for binding to CI-MPRs[50,51]. As in the application to ERT, the use of GILT in the design of LYTACs[52] may be advantageous in situations where IGF2 levels are sufficient to reduce CI-MPR's ability to bind and internalize M6P-tagged molecules. Future studies are needed to examine at the cellular level how IGF2 impacts the functioning of CI-MPR in delivering phosphomannosyl-containing ligands to lysosomes.

Methods

Isolation of soluble cation-independent mannose 6-phosphate receptor (sCI-MPR) from newborn calf serum

sCI-MPR was isolated from newborn calf serum purchased from a commercial vendor (Thermo Fisher Scientific/Gibco) by affinity chromatography using GAA-containing resin. GAA (ATB200, Amicus Therapeutics) was immobilized on NHS-activated agarose (Thermo Fisher Scientific) following the manufacturer's protocol. Briefly, the activated agarose was washed with water, followed by a wash with HBS (10 mM HEPES, pH 7.4, 150 mM NaCl). GAA (1 mg/mL) in HBS was added to the resin and incubated at room temperature for 2 h. The GAA-containing solution was removed, and the resin was suspended in 1 M ethanolamine for 15 min. The resin was washed with HBS and packed into a 5 mL FPLC column. Frozen serum (100 mL) was thawed and mixed with an equal volume of IBS-EDTA (50 mM imidazole pH 6.8, 150 mM NaCl, 5 mM β-glycerol phosphate, and 5 mM EDTA). HALT protease inhibitor cocktail (Thermo Fisher Scientific) was added, and the mixture was passed over a GAA-coupled agarose column. The column was washed with IBS-EDTA, followed by 10 mM glucose 6-phosphate (G6P) in IBS-EDTA. sCI-MPR was eluted with MBS-EDTA (50 mM MES pH 4.6, 150 mM NaCl, 5 mM β-glycerol phosphate, and 5 mM EDTA) followed by immediate neutralization using 1 M imidazole pH 7.5. sCI-MPR was further purified by a second round of affinity chromatography using the GAA-agarose resin, except here the resin was washed sequentially with buffer and 10 mM G6P, and sCI-MPR was eluted with 10 mM mannose 6-phosphate (M6P). Finally, the protein was buffer exchanged on a G-25 column (PD MiniTrap, Cytiva) equilibrated with HBS. The concentration of the purified protein was determined by measuring the absorbance at 280 nm, assuming an extinction coefficient of 1 $OD_{280} = 1$ mg/mL. Protein was stored in HBS at 4 °C.

Expression and purification of PPT1

Recombinant human PPT1 was expressed and purified following the protocol described in Lu et al.[53].

MS peptide analysis

Three preparations of sCI-MPR were isolated from two separate lots of newborn calf serum. The isolated proteins were reduced with TCEP, alkylated with iodoacetamide, digested overnight with trypsin, and desalted using Phoenix columns (PreOmics, Martinsreid Germany). The resulting peptides were analyzed on a Thermo Fisher Scientific Orbitrap Fusion Lumos Mass Spectrometer via two technical replicate injections of 5% of the material each using a data-dependent acquisition (DDA) HCD instrument method with a 90-min gradient on a Thermo Acclaim PepMap C18 75 μm × 50 cm column. MS data were analyzed using Proteome Discoverer version 2.4 (Thermo Fisher Scientific). Protein identifications were filtered to include only those proteins identified by two

or more unique peptides and ranked as high confidence. Mass spectrometry analyses were performed in the Center for Biomedical Mass Spectrometry Research at the Medical College of Wisconsin.

Site-specific N-glycosylation analysis by MS

Site-specific N-glycosylation analysis of the purified sCI-MPR was determined by mass spectrometry using the method of Cao et al.[30]. Briefly, reduced/alkylated/trypsin-digested sCI-MPR was further treated sequentially with EndoF1 followed by PNGase F/^{18}O water to create unique mass signatures to determine the N-glycan type and occupancy of the potential N-linked sites. For direct glycopeptide examination, reduced/alkylated/trypsin-digested sCI-MPR was subjected to LC–MS analysis without further enzymatic treatments.

LC–MS/MS analysis was performed on an Orbitrap-Fusion Lumos equipped with an EASY nanospray source and Ultimate3000 autosampler LC system (Thermo Fisher Scientific). Resuspended tryptic peptides with or without EndoF1/PNGase F/^{18}O-water treatments were chromatographed on a nano-C18 column (Acclaim PepMap, 75 µm × 50 cm, Thermo Fisher Scientific) with a gradient elution of increasing mobile phase B (80% acetonitrile, 0.1% formic acid in distilled H_2O) at a flow rate of 300 nL/min routed directly into the mass spectrometer. A 90-min gradient program was used for separating EndoF1/PNGase F/^{18}O-water-treated peptides, whereas a 180-min gradient was used for separating glycopeptides. Full MS spectra were collected at 60,000 resolution in FT mode, and MS/MS spectra were obtained for each precursor ion by data-dependent scans (top-speed scan, 3 s) utilizing CID, HCD, or ETD activation and subsequent detection in FT mode.

Resulting data were annotated by manual data interpretation following initial processing by Byonic software version 3.2.0 (Protein Metrics). Byonic parameters were set to allow 10 ppm of precursor ion monoisotopic mass tolerance and 20 ppm of fragment ion tolerance. Byonic searches were performed against the bovine CI-MPR sequence with differential modifications of carbamidomethylation at cysteine (+ 57.02146 Da) and oxidation at methionine (+ 15.9949 Da). In addition, ^{18}O-labeled aspartic acid (+ 2.98826 Da) and HexNAc modification of asparagine (+ 203.079373) were allowed for the analysis of Endo-F/PNGase F/^{18}O- data, whereas phosphorylated and non-phosphorylated human/mammalian N-glycosylation was allowed for the direct glycopeptide data analysis.

The proportions of the three glycosylation states at each N-linked site (N + 0, N + 3, and N + 203 modifications) in the Endo-F/PNGase F/^{18}O- data were calculated based on the intensity peak area of corresponding peptide ions over all identified charge states. N-glycan prevalence at each N-linked site in the direct glycopeptide analysis data were calculated based on spectral counts of each N-glycopeptide ion. Mass spectrometry analyses for glycopeptide characterization were performed by the Translational Metabolomics Shared Resource at the Medical College of Wisconsin Cancer Center.

Binding affinity and inhibition studies using SPR

SPR studies were conducted at 25 °C using a Biacore S200 instrument (Cytiva). Biotinylated IGF2 (GroPep, Inc.) was captured on the surface of a streptavidin-coupled SA sensor chip (Cytiva, Inc.). Human lysosomal enzymes GAA (Amicus Therapeutics) and PPT1 were immobilized onto the surface of a CM5 sensor chip (Cytiva, Inc.) by amine-coupling as described previously[54]. sCI-MPR was diluted in running buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM $CaCl_2$, 5 mM $MnCl_2$, 5 mM β-glycerol phosphate, and 0.05% surfactant polysorbate 20 (P20) and passed over the surfaces at a flow rate of 40 µL/min. All SPR experiments were carried out in triplicate. Sensorgrams were double-referenced to control for refractive index changes by subtracting the response on the reference surface and the response for buffer-only injections as described[55]. Measurement of IGF2 binding kinetics used a 5-min association phase and a 10-min dissociation phase, and the resulting sensorgrams were fit to a 1:1 binding model (S200 Evaluation Software, Cytiva). Measurement of lysosomal enzyme binding used a 3-min association phase followed by a 10-min dissociation phase. The response levels were determined as the average response 4 s before the end of the association phase of each injection (Biacore S200 Evaluation Software, Cytiva), plotted against the sCI-MPR concentrations, and the resulting curves were fit to a one-site specific binding equation (GraphPad Prism v 10.2.0). In some experiments, sCI-MPR (20 nM) was incubated with IGF2 (rhIGF2, Shenandoah recombinant proteins, FUJIFILM Irvine Scientific) or IGF1 (rhIGF1, Cell Sciences) for 1 h at room before flowing over the GAA or PPT1 surface for 3 min, followed by a 5-min dissociation phase. The response levels were determined as the average response 4 s before the end of the association phase (Biacore S200 Evaluation Software, Cytiva) and normalized as a percentage of the response at 0 nM IGF2. The response levels were plotted against the log of the corresponding IGF2 or IGF1 concentration and fit to a 4-parameter dose response curve (GraphPad Prism v 10.2.0). The dual injection function of the Biacore S200 instrument was used to evaluate IGF2's effect on the dissociation of the receptor from lysosomal enzymes. Samples of sCI-MPR (20 nM) were flowed over the GAA or PPT surfaces for 3 min, followed by a second injection of running buffer without receptor but containing IGF2 or IGF1 for 2 min, and finally buffer alone for an additional 5 min. The response level at the end of the IGF2/IGF1 second injection phase was determined as the average response for the 4 s before the end of the IGF2/IGF1 second injection (Biacore S200 Evaluation Software, Cytiva). Responses were normalized as the percentage of response at 0 nM IGF2 and plotted against the log of the corresponding IGF2 or IGF1 concentration. Data were fit to a 4-parameter dose response curve (GraphPad Prism v 10.2.0). The rate of dissociation (k_{off}) during the IGF2 or IGF1 injection phase of the sensorgrams was estimated using a one-site model in BiaEvaluation v 4.1 (Cytiva). The off rates (k_{off}) were plotted against the log of the inhibitor concentration and analyzed by a one-way ANOVA (GraphPad Prism v 10.2.0) in which the Alpha value is 0.05 and n = 3 for each concentration of inhibitor. Dunnett's multiple comparison test was used to compare the mean of the control (0 inhibitor) to the mean at each inhibitor concentration individually and the test method assumes a normal Gaussian distribution.

Negative-staining EM

Negative-staining EM studies of the receptor were conducted using 1.4 μM sCI-MPR prepared in the absence and presence of 200 nM IGF2 at pH 6.8, following methods similar to those used by Olson et al.[23] Samples were prepared using optimized negative-staining electron microscopy (OpNS-EM)[35,56,57]. Briefly, a 4 μL aliquot of the sample was placed on an ultra-thin carbon continuous film grid (CF-200-Cu-UL, Electron Microscopy Sciences, Hatfield, PA) that had been glow-discharged for 15 s. After a one-minute incubation, excess solution was removed from the grid. The grid was then stained with three drops of 1% (w/v) uranyl formate (UF). Excess solution was removed with filter paper by blotting from the rear side, and the grid was then immediately dried with nitrogen.

The EM specimens were loaded into a Zeiss Libra 120 Plus TEM (Carl Zeiss NTS, Oberkochen, Germany), equipped with an LaB6 gun operating at 120 kV, an in-column energy filter, and a 4 k × 4 k Gatan UltraScan 4000 charge-coupled device (CCD) camera. Micrographs were acquired under near Scherzer defocus at a magnification of 80,000 × (corresponding to 1.48 Å/pixel). For 2D reference-free class averaging analysis, particles were selected from the micrographs that were low-pass filtered to 15 Å after removal of X-ray speckles, and then subjected to reference-free class averaging using the program EMAN[58].The dimensions of 576 individual particles from each sample were measured from two perpendicular directions of a particle to catalog the long distance and short distance, and the length ratio of each particle was determined.

Hydroxyl radical protein footprinting

A final concentration of 1 μM sCI-MPR was incubated in 10 mM Tris pH 7.4, 150 mM NaCl buffer in the presence or absence of 1 μM IGF2. HRPF was performed as described previously[59]. Briefly, 20 μL of protein sample mixture containing 1 mM adenine and 100 mM hydrogen peroxide was irradiated using Fox Photolysis System (GenNext Technologies) with the lamp set at 700 V. The flow rate was adjusted to 15 μL/min to ensure a 15% exclusion volume between irradiated segments. After the illumination, each replicate was collected in a microcentrifuge tube containing 25 μL of quench mixture that contained 0.5 μg/μL H-Met-NH2 and 100 mM DMTU to eliminate secondary oxidants[59]. The adenine hydroxyl radical dosimetry readings were measured inline at 265 nm to ensure that all the samples were exposed to equivalent effective doses of hydroxyl radical[60,61]. All oxidations were performed in triplicate for statistical analysis.

After photolysis and quenching, the samples were divided equally to digest with trypsin and chymotrypsin separately. To digest with trypsin, 50 mM Tris, pH 8.0 containing 1 mM $CaCl_2$, 5 mM DTT was added to the protein samples and to digest with chymotrypsin, 150 mM Tris, pH 8.0 was used. The samples were incubated at 95 °C for 15 min to denature and reduce the protein. The sample was cooled on ice, and a 1:20 ratio of protease to protein was added. The solution was incubated with rotation at 37 °C for 12 h for trypsin and 25 °C for digestion with chymotrypsin. Sample digestion was stopped by adding 0.1% formic acid, and the samples were analyzed on a Dionex Ultimate 3000 nano-LC system coupled to an Orbitrap Fusion (Thermo Fisher Scientific). Samples were trapped on a 300 μm id X5 mm PepMap 100, 5 μm C18 trapping cartridge (Thermo Fisher Scientific), then back-eluted onto an Acclaim PepMap 100 C18 nanocolumn (75 μm × 15 cm, 2 μm, Thermo Fisher Scientific). Peptides were separated by chromatography using a binary gradient of solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in acetonitrile) at a 0.30 μL/min flow rate. The peptides were eluted with a gradient consisting of 2–10% solvent B over 4 min, increasing to 32% B over 25 min, ramped to 95% B over 4 min, held for 4 min, and then returned to 2% B over 2 min and held for 8 min. Peptides were eluted directly into the nanospray source of an Orbitrap Fusion Tribrid mass spectrometer using a conductive nanospray emitter obtained from Thermo Fisher Scientific. All data were collected in positive ion mode. Collision-induced dissociation CID was used to fragment peptides, with an isolation width of 2 m/z units. The spray voltage was set to 2300 V, and the heated capillary temperature was set to 300 °C. Full MS scans were acquired from m/z 350 to 2000, followed by eight subsequent MS2 CID scans on the top eight most abundant peptide ions. Peptides from sCI-MPR tryptic digests were identified using ByOnic version v4.4.2 (Protein Metrics). The search parameters included all possible major oxidation modifications as variable modifications, and the enzyme specificity was set to cleave the protein after arginine and lysine residues for trypsin digested samples and tyrosine, phenylalanine, tryptophan, and leucine for chymotrypsin treated samples, excluding cleavage before proline. Average oxidation events per peptide were calculated using FoxWare v1.1.8f. (GenNext Technologies), with manual auditing of selected peptides to ensure proper automated calculations. No manual changes of any automated results were performed.

Statistical analyses

Statistical comparisons were performed using GraphPad Prism or Microsoft Excel and are described in the legends. Statistical significance levels were set at $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[62] partner repository with the dataset identifier PXD052326 for the hydroxy radical protein footprinting studies and dataset identifier PXD052183 for the LC-MS/MS analysis of the CI-MPR isolated from bovine serum. Mass spectrometry glycopeptide data are available at GlycoPost[63] with the dataset identifier GPST000429.

Received: 20 May 2024; Accepted: 3 October 2024
Published online: 06 November 2024

References

1. Settembre, C. & Perera, R. M. Lysosomes as coordinators of cellular catabolism, metabolic signalling and organ physiology. *Nat. Rev. Mol. Cell Biol.* **25**, 223–245. https://doi.org/10.1038/s41580-023-00676-x (2024).
2. Platt, F. M., d'Azzo, A., Davidson, B. L., Neufeld, E. F. & Tifft, C. J. Lysosomal storage diseases. *Nat. Rev. Dis. Primers* **4**, 27. https://doi.org/10.1038/s41572-018-0025-4 (2018).
3. Dahms, N., Braulke, T. & Varki, A. in *Essentials of Glycobiology* (eds A. Varki *et al.*) pp. 443–454 (2022).
4. Freeze, H. H., Steet, R., Suzuki, T., Kinoshita, T. & Schnaar, R. L. in *Essentials of Glycobiology* (eds A. Varki *et al.*) pp. 583–598 (2022).
5. Nadimpalli, S. K. & Amancha, P. K. Evolution of mannose 6-phosphate receptors (MPR300 and 46): Lysosomal enzyme sorting proteins. *Curr. Protein Pept. Sci.* **11**, 68–90. https://doi.org/10.2174/138920310790274644 (2010).
6. Ghosh, P., Dahms, N. M. & Kornfeld, S. Mannose 6-phosphate receptors: New twists in the tale. *Nat. Rev. Mol. Cell Biol.* **4**, 202–212. https://doi.org/10.1038/nrm1050 (2003).
7. Morgan, D. O. et al. Insulin-like growth factor II receptor as a multifunctional binding protein. *Nature* **329**, 301–307. https://doi.org/10.1038/329301a0 (1987).
8. MacDonald, R. G. et al. A single receptor binds both insulin-like growth factor II and mannose-6-phosphate. *Science* **239**, 1134–1137. https://doi.org/10.1126/science.2964083 (1988).
9. Williams, C. et al. An exon splice enhancer primes IGF2:IGF2R binding site structure and function evolution. *Science* **338**, 1209–1213. https://doi.org/10.1126/science.1228633 (2012).
10. Surinya, K. H. et al. An investigation of the ligand binding properties and negative cooperativity of soluble insulin-like growth factor receptors. *J. Biol. Chem.* **283**, 5355–5363. https://doi.org/10.1074/jbc.M707054200 (2008).
11. Wang, Z. Q., Fung, M. R., Barlow, D. P. & Wagner, E. F. Regulation of embryonic growth and lysosomal targeting by the imprinted Igf2/Mpr gene. *Nature* **372**, 464–467. https://doi.org/10.1038/372464a0 (1994).
12. Lau, M. M. et al. Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. *Genes Dev.* **8**, 2953–2963. https://doi.org/10.1101/gad.8.24.2953 (1994).
13. De Souza, A. T., Hankins, G. R., Washington, M. K., Orton, T. C. & Jirtle, R. L. M6P/IGF2R gene is mutated in human hepatocellular carcinomas with loss of heterozygosity. *Nat. Genet.* **11**, 447–449. https://doi.org/10.1038/ng1295-447 (1995).
14. Martin-Kleiner, I. & Gall Troselj, K. Mannose-6-phosphate/insulin-like growth factor 2 receptor (M6P/IGF2R) in carcinogenesis. *Cancer Lett.* **289**, 11–22. https://doi.org/10.1016/j.canlet.2009.06.036 (2010).
15. O'Gorman, D. B., Weiss, J., Hettiaratchi, A., Firth, S. M. & Scott, C. D. Insulin-like growth factor-II/mannose 6-phosphate receptor overexpression reduces growth of choriocarcinoma cells in vitro and in vivo. *Endocrinology* **143**, 4287–4294. https://doi.org/10.1210/en.2002-220548 (2002).
16. Lee, J. S., Weiss, J., Martin, J. L. & Scott, C. D. Increased expression of the mannose 6-phosphate/insulin-like growth factor-II receptor in breast cancer cells alters tumorigenic properties in vitro and in vivo. *Int. J. Cancer* **107**, 564–570. https://doi.org/10.1002/ijc.11453 (2003).
17. Bohnsack, R. N., Patel, M., Olson, L. J., Twining, S. S. & Dahms, N. M. Residues essential for plasminogen binding by the cation-independent mannose 6-phosphate receptor. *Biochemistry* **49**, 635–644. https://doi.org/10.1021/bi901779p (2010).
18. Leksa, V. et al. The N terminus of mannose 6-phosphate/insulin-like growth factor 2 receptor in regulation of fibrinolysis and cell migration. *J. Biol. Chem.* **277**, 40575–40582. https://doi.org/10.1074/jbc.M207979200 (2002).
19. Nykjaer, A. et al. Mannose 6-phosphate/insulin-like growth factor-II receptor targets the urokinase receptor to lysosomes via a novel binding interaction. *J. Cell Biol.* **141**, 815–828. https://doi.org/10.1083/jcb.141.3.815 (1998).
20. Kim, J. J., Olson, L. J. & Dahms, N. M. Carbohydrate recognition by the mannose-6-phosphate receptors. *Curr. Opin. Struct. Biol.* **19**, 534–542. https://doi.org/10.1016/j.sbi.2009.09.002 (2009).
21. Schmidt, B., Kiecke-Siemsen, C., Waheed, A., Braulke, T. & von Figura, K. Localization of the insulin-like growth factor II binding site to amino acids 1508–1566 in repeat 11 of the mannose 6-phosphate/insulin-like growth factor II receptor. *J. Biol. Chem.* **270**, 14975–14982. https://doi.org/10.1074/jbc.270.25.14975 (1995).
22. Wang, R., Qi, X., Schmiege, P., Coutavas, E. & Li, X. Marked structural rearrangement of mannose 6-phosphate/IGF2 receptor at different pH environments. *Sci Adv* **6**, eaaz1466. https://doi.org/10.1126/sciadv.aaz1466 (2020).
23. Olson, L. J. et al. Allosteric regulation of lysosomal enzyme recognition by the cation-independent mannose 6-phosphate receptor. *Commun. Biol.* **3**, 498. https://doi.org/10.1038/s42003-020-01211-w (2020).
24. Causin, C. et al. Mannose 6-phosphate/insulin-like growth factor II-binding proteins in human serum and urine. Their relation to the mannose 6-phosphate/insulin-like growth factor II receptor. *Biochem. J.* **252**, 795–799. https://doi.org/10.1042/bj2520795 (1988).
25. Valenzano, K. J., Remmler, J. & Lobel, P. Soluble insulin-like growth factor II/mannose 6-phosphate receptor carries multiple high molecular weight forms of insulin-like growth factor II in fetal bovine serum. *J. Biol. Chem.* **270**, 16441–16448. https://doi.org/10.1074/jbc.270.27.16441 (1995).
26. Bobek, G., Scott, C. D. & Baxter, R. C. Secretion of soluble insulin-like growth factor-II/mannose 6-phosphate receptor by rat tissues in culture. *Endocrinology* **128**, 2204–2206. https://doi.org/10.1210/endo-128-4-2204 (1991).
27. Confort, C., Rochefort, H. & Vignon, F. Insulin-like growth factors (IGFs) stimulate the release of alpha 1-antichymotrypsin and soluble IGF-II/mannose 6-phosphate receptor from MCF7 breast cancer cells. *Endocrinology* **136**, 3759–3766. https://doi.org/10.1210/endo.136.9.7649082 (1995).
28. Leksa, V. et al. Soluble M6P/IGF2R released by TACE controls angiogenesis via blocking plasminogen activation. *Circ. Res.* **108**, 676–685. https://doi.org/10.1161/CIRCRESAHA.110.234732 (2011).
29. Mellquist, J. L., Kasturi, L., Spitalnik, S. L. & Shakin-Eshleman, S. H. The amino acid following an asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency. *Biochemistry* **37**, 6833–6837. https://doi.org/10.1021/bi972217k (1998).
30. Cao, L. et al. Global site-specific N-glycosylation analysis of HIV envelope glycoprotein. *Nat. Commun.* **8**, 14954. https://doi.org/10.1038/ncomms14954 (2017).
31. Tong, P. Y., Tollefsen, S. E. & Kornfeld, S. The cation-independent mannose 6-phosphate receptor binds insulin-like growth factor II. *J. Biol. Chem.* **263**, 2585–2588 (1988).
32. Meena, N. K. & Raben, N. Pompe disease: new developments in an old lysosomal storage disorder. *Biomolecules*https://doi.org/10.3390/biom10091339 (2020).
33. Rosenberg, J. B., Chen, A., Kaminsky, S. M., Crystal, R. G. & Sondhi, D. Advances in the treatment of neuronal ceroid lipofuscinosis. *Expert Opin. Orphan Drugs* **7**, 473–500. https://doi.org/10.1080/21678707.2019.1684258 (2019).
34. Frago, S. et al. Functional evolution of IGF2:IGF2R domain 11 binding generates novel structural interactions and a specific IGF2 antagonist. *Proc. Natl. Acad. Sci. USA* **113**, E2766-2775. https://doi.org/10.1073/pnas.1513023113 (2016).
35. Rames, M., Yu, Y. & Ren, G. Optimized negative staining: a high-throughput protocol for examining small and asymmetric protein structure by electron microscopy. *J. Vis. Exp.* https://doi.org/10.3791/51087 (2014).
36. Sharp, J. S. et al. Flash oxidation (FOX) system: a novel laser-free fast photochemical oxidation protein footprinting platform. *J. Am. Soc. Mass Spectrom* **32**, 1601–1609. https://doi.org/10.1021/jasms.0c00471 (2021).
37. Shi, L. & Gross, M. L. Fast photochemical oxidation of proteins coupled with mass spectrometry. *Protein Pept. Lett.* **26**, 27–34. https://doi.org/10.2174/0929866526666181128124554 (2019).

38. Shi, X. & Jarvis, D. L. Protein N-glycosylation in the baculovirus-insect cell system. *Curr. Drug Targets* **8**, 1116–1125. https://doi.org/10.2174/138945007782151360 (2007).
39. Goldberg, D. E., Gabel, C. A. & Kornfeld, S. Studies of the biosynthesis of the mannose 6-phosphate receptor in receptor-positive and -deficient cell lines. *J. Cell Biol.* **97**, 1700–1706. https://doi.org/10.1083/jcb.97.6.1700 (1983).
40. Chen, M. et al. Comparative site-specific N-glycoproteome analysis reveals aberrant N-glycosylation and gives insights into mannose-6-phosphate pathway in cancer. *Commun. Biol* **6**, 48. https://doi.org/10.1038/s42003-023-04439-4 (2023).
41. Do, H. V., Khanna, R. & Gotschall, R. Challenges in treating Pompe disease: An industry perspective. *Ann. Transl. Med.* **7**, 291. https://doi.org/10.21037/atm.2019.04.15 (2019).
42. Duguay, S. J. et al. Post-translational processing of the insulin-like growth factor-2 precursor. Analysis of O-glycosylation and endoproteolysis. *J. Biol. Chem.* **273**, 18443–18451. https://doi.org/10.1074/jbc.273.29.18443 (1998).
43. Livingstone, C. IGF2 and cancer. *Endocr. Relat. Cancer* **20**, R321-339. https://doi.org/10.1530/ERC-13-0231 (2013).
44. Selenou, C., Brioude, F., Giabicani, E., Sobrier, M. L. & Netchine, I. IGF2: Development, genetic and epigenetic abnormalities. *Cells* https://doi.org/10.3390/cells11121886 (2022).
45. Ogawa, O. et al. Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. *Nature* **362**, 749–751. https://doi.org/10.1038/362749a0 (1993).
46. Taniguchi, T., Sullivan, M. J., Ogawa, O. & Reeve, A. E. Epigenetic changes encompassing the IGF2/H19 locus associated with relaxation of IGF2 imprinting and silencing of H19 in Wilms tumor. *Proc. Natl. Acad. Sci. USA* **92**, 2159–2163. https://doi.org/10.1073/pnas.92.6.2159 (1995).
47. LeBowitz, J. H. et al. Glycosylation-independent targeting enhances enzyme delivery to lysosomes and decreases storage in mucopolysaccharidosis type VII mice. *Proc. Natl. Acad. Sci. USA* **101**, 3083–3088. https://doi.org/10.1073/pnas.0308728100 (2004).
48. Maga, J. A. et al. Glycosylation-independent lysosomal targeting of acid alpha-glucosidase enhances muscle glycogen clearance in pompe mice. *J. Biol. Chem.* **288**, 1428–1438. https://doi.org/10.1074/jbc.M112.438663 (2013).
49. Catalano, F. et al. Tagged IDS causes efficient and engraftment-independent prevention of brain pathology during lentiviral gene therapy for Mucopolysaccharidosis type II. *Mol. Ther. Methods Clin. Dev.* **31**, 101149. https://doi.org/10.1016/j.omtm.2023.101149 (2023).
50. Banik, S. M. et al. Lysosome-targeting chimaeras for degradation of extracellular proteins. *Nature* **584**, 291–297. https://doi.org/10.1038/s41586-020-2545-9 (2020).
51. Stevens, C. M. et al. Development of oligomeric mannose-6-phosphonate conjugates for targeted protein degradation. *ACS Med. Chem. Lett.* **14**, 719–726. https://doi.org/10.1021/acsmedchemlett.2c00479 (2023).
52. Mikitiuk, M. et al. IGF2 peptide-based LYTACs for targeted degradation of extracellular and transmembrane proteins. *Molecules* https://doi.org/10.3390/molecules28227519 (2023).
53. Lu, J. Y., Hu, J. & Hofmann, S. L. Human recombinant palmitoyl-protein thioesterase-1 (PPT1) for preclinical evaluation of enzyme replacement therapy for infantile neuronal ceroid lipofuscinosis. *Mol. Genet. Metab.* **99**, 374–378. https://doi.org/10.1016/j.ymgme.2009.12.002 (2010).
54. Chavez, C. A. et al. Domain 5 of the cation-independent mannose 6-phosphate receptor preferentially binds phosphodiesters (mannose 6-phosphate N-acetylglucosamine ester). *Biochemistry* **46**, 12604–12617. https://doi.org/10.1021/bi7011806 (2007).
55. Myszka, D. G. Kinetic, equilibrium, and thermodynamic analysis of macromolecular interactions with BIACORE. *Methods Enzymol.* **323**, 325–340. https://doi.org/10.1016/s0076-6879(00)23372-7 (2000).
56. Zhang, L. et al. An optimized negative-staining protocol of electron microscopy for apoE4 POPC lipoprotein. *J. Lipid Res.* **51**, 1228–1236. https://doi.org/10.1194/jlr.D002493 (2010).
57. Liu, J. et al. Optimized negative-staining protocol for lipid-protein interactions investigated by electron microscopy. *Methods Mol. Biol.* **2003**, 163–173. https://doi.org/10.1007/978-1-4939-9512-7_8 (2019).
58. Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97. https://doi.org/10.1006/jsbi.1999.4174 (1999).
59. Shami, A. A., Misra, S. K., Jones, L. M. & Sharp, J. S. Dimethylthiourea as a quencher in hydroxyl radical protein footprinting experiments. *J. Am. Soc. Mass Spectrom* **34**, 2864–2867. https://doi.org/10.1021/jasms.3c00323 (2023).
60. Sharp, J. S., Misra, S. K., Persoff, J. J., Egan, R. W. & Weinberger, S. R. Real time normalization of fast photochemical oxidation of proteins experiments by inline adenine radical dosimetry. *Anal. Chem.* **90**, 12625–12630. https://doi.org/10.1021/acs.analchem.8b02787 (2018).
61. Misra, S. K., Orlando, R., Weinberger, S. R. & Sharp, J. S. Compensated hydroxyl radical protein footprinting measures buffer and excipient effects on conformation and aggregation in an adalimumab biosimilar. *AAPS J.* **21**, 87. https://doi.org/10.1208/s12248-019-0358-2 (2019).
62. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552. https://doi.org/10.1093/nar/gkab1038 (2022).
63. Watanabe, Y., Aoki-Kinoshita, K. F., Ishihama, Y. & Okuda, S. GlycoPOST realizes FAIR principles for glycomics mass spectrometry data. *Nucleic Acids Res.* **49**, D1523–D1528. https://doi.org/10.1093/nar/gkaa1012 (2020).

Acknowledgements

We thank Dr. Sandra L. Hofmann for providing the CHO cell line expressing recombinant human PPT1 and Dr. Linda J. Olson for purifying PPT1. We thank Amicus Therapeutics for generously providing IGF2 and the lysosomal enzyme GAA. This research was funded by NIH grants R01DK042667 (to N.M.D.), R01HL115153, R01GM104427 and R01MH077303 (to G.R) and R01GM127267 (to J.S.S.). The work at the molecular foundry, LBNL was supported by the Office of Science, Office of Basic Energy Sciences of the United States Department of Energy (contract no. DE-AC02-05CH11231). Research reported in this publication using the Biacore S200 instrument was supported by the Office of The Director, NIH under Award Number S10OD028640. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Author contributions

R.N.B., K.A., R.G., J.S.S., N.M.D. designed the study. R.N.B., S.K.M., J.L., M.I.-A., M.P., and K.A. performed the experiments. All authors analyzed data. R.N.B., M.I.-A., and J.L. prepared figures and initial drafts of legends and methods. N.M.D. wrote the manuscript. All authors revised and edited the manuscript.

Declarations

Competing interests

N.M.D. is a member of the scientific advisory board of M6P Therapeutics. J.S.S. discloses a significant interest in GenNext Technologies, Inc., a company commercializing technologies for protein higher-order structure analysis. R.N.B., S.K.M., J.L, M.I.-A., M.P., K.A., and G.R. declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-75300-9.

Correspondence and requests for materials should be addressed to N.M.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.