Glycoconjugate Journal 20, 257–265, 2004

Molecular cloning of goat Mannose 6-phosphate receptors, MPR 300 and 46

Koduru Suresh, Suryanarayana Raju Vegiraju and Siva Kumar Nadimpalli

Protein Biochemistry & Molecular Biology Laboratory, Department of Biochemistry, University of Hyderabad, Hyderabad 500 046, India

Mannose 6-phosphate receptor proteins (MPR 300 and 46) are type 1 transmembrane glycoproteins that mediate transport of lysosomal enzymes to lysosomes. In a recent study we have purified both receptors from goat liver and raised antibodies. An ELISA method was developed that allowed quantification of both receptors in different tissues of goat and chicken and an immuno-affinity method was also developed to purify the receptors. In the present study to understand the structural similarities of the goat receptors to other known receptor proteins, we have prepared cDNA clones for both receptors by RT-PCR approach. A partial cDNA clone (1.368 kb) for the MPR 300 protein, and a full length cDNA clone (0.84 kb) for the MPR 46 protein were obtained. MPR 300 exhibits typical conserved cassette structure in the amino terminal domain similar to other known vertebrate MPR proteins with the conserved cysteine residues and the ligand binding arginine residue in the third domain. MPR 46 exhibits high degree of sequence homology to other known MPR proteins with the conserved cysteine residues, the transmembrane domain and the cytoplasmic tail. mRNA transcript size for both receptors were comparable with that of other vertebrates.
Published in 2004.

Keywords: **Mannose 6-phosphate receptors, goat, cDNA sequencing, RT-PCR, vertebrate**

Introduction

The bovine and human mannose 6-phosphate receptor proteins (MPR 300 and 46) mediate transport of lysosomal enzymes and have been extensively characterized [1]. cDNA sequence for the chicken MPR 300 has been reported [2] and only a partial cDNA clone for chicken MPR 46 is published [3]. Similar proteins have been identified by us in non-mammalian vertebrates such as reptiles, amphibians [4] and in the earliest vertebrate fish [5]. A partial cDNA clone for the fish MPR 300 reported by us suggests that MPR 300 protein exhibits conserved cassette structure in the extracellular domain among all vertebrates [6]. Among the invertebrates, the first report on the presence of MPR 300 in the molluscs came from our studies [7]. Recently the putative MPR 46 in the mollusc was also identified [8].

The extracytoplasmic domain of MPR 300 (mammalian, chicken) is comprised of 15 repetitive units which share significant sequence similarity with each other and also with the single unit that constitutes the extracytoplasmic domain of MPR 46. Further the repetitive domains of known MPR 300 proteins also exhibit 14–38% sequence homology [9]. The long term objective of our laboratory is to understand the structure, function and evolution of MPR proteins. In a recent study our laboratory purified and developed antibodies to the goat MPR proteins and using these antibodies an ELISA method was developed to quantify the receptor proteins from different tissues of goat and chicken [10]. In addition an immuno-affinity method was developed to purify the MPR proteins [11].

To further understand the function of both the goat receptors, it is necessary to obtain their sequence information. Therefore the objective of the present study was to obtain cDNA clones for both receptors. The strategy adopted was to isolate total RNA from goat liver tissue and make RT-PCR experiments using primers designed by multiple sequence alignment of the known MPR 300 and 46 protein sequences. PCR products were cloned and the cDNA sequenced. This approach not only facilitated to know the sequences of these receptors but also enabled us to make a structural comparison of the receptors.

To whom correspondence should be addressed: Siva Kumar Nadimpalli, Protein Biochemistry & Molecular Biology Laboratory, Department of Biochemistry, University of Hyderabad, Hyderabad 500 046, India. Fax: 0091-40-23010120, 23010082; E-mail: drnsk7@yahoo.co.in, nsksl@uohyd.ernet.in

Materials and methods

All reagents and kits were purchased from reputed international firms. Goat liver tissue was obtained from the local slaughter

house under sterile conditions, carried on ice to the laboratory and frozen at −80°C.

Isolation of total RNA from goat liver tissue

Total RNA from goat liver tissue (20 mg) was isolated with the RNeasy kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Purity and integrity of the RNA was analyzed by agarose (1%) gel electrophoresis under denaturing conditions [3].

Reverse transcriptase—Polymerase chain reaction

The following degenerate primer oligonucleotides were used for amplification of cDNA fragments for goat MPR 300: sense primer A (5′-CTGTGCAGTTACACA TGGGAAGC-3′), anti sense primer B (5′-GGCATACTCAGTGATCCACTC-3′), anti sense primer C (5′-GCCATACCACAGCTTCCCAATTYTG-3′). To amplify cDNA fragments for goat MPR 46 the following primers were used: sense primer 1 (5′-GTGGCAGTSAGAGAATCYTGGCAG-3′), sense primer 2 (5′-ATGATGTCCCCCCTCCACAGCTCC-3′) and anti sense primer 3 (5′-TCACATTGGTAAYAARTGRTGATC-3′). Reverse transcriptions (RT) were performed with the Ominiscript kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. One microgram of goat liver total RNA was denatured at 65°C for 5 min and immediately chilled on ice. After addition of 20 pmols of respective anti sense primers, first-strand cDNA synthesis was performed at 37°C for 1 hour. Then, 20% of the first-strand synthesis product was amplified by polymerase chain reaction (PCR) with 1U HotStar Taq polymerase (Qiagen, Hilden, Germany) and 12 pmol of respective forward and reverse primers with following conditions (95°C × 15′; 94°C × 1′; 55°C × 1; 72°C × 1–1.5′ for 30 cycles).

Transformation and isolation of MPR 300 and 46 DNA

The PCR products were analyzed on 1% agarose gel electrophoresis. The single band obtained was excised, gel purified and subjected to TA cloning into pCR 2.1 TOPO vector (Invitrogen) and transformed into OneShot chemically competent cells supplied with the kit. For 46 cloning, in addition to pCR 2.1 TOPO vector, pTZ57R, vector (MBI Fermentas) was also used. The positive clones were selected by blue-white selection and the plasmid DNA isolated using mini prep spin column (Qiagen, Hilden, Germany). The size of the insert was confirmed by restriction digestion of the plasmid DNA with ECoRI for MPR 300 and with KpnI and PstI enzymes for MPR 46.

Northern blot

Fifteen micrograms of total RNA from goat liver tissue was subjected to denaturing agarose (1%) gel electrophoresis [3] and transferred to Hybond N membrane (Amersham). Then, the membranes were cross-linked for 45 s in a UV cross-linker. Membranes were then incubated for 2–3 h at 42°C in hybridization buffer (100ug/ml heat-denatured salmon sperm DNA, 40% formamide, 10% dextran sulfate, 1% Denhardt's solution, 4.8 × SSC buffer, 10mM Tris-HCl buffer pH 7.4 and 1% sodium dodecyl sulfate (SDS).

cDNA fragments of goat MPR 300 (1.368 kb) and MPR 46 (840 bp) obtained in this study were subjected to radiolabeling with the random primer labeling kit (MBI Fermentas) α ^{32}P dCTP (11.1×10^{13} Bq/mmol). ^{32}P labeled fragments were denatured at 98°C for 5 min before dilution with hybridization buffer. Hybridization was performed overnight at 42°C. Membranes were washed for 10 min at room temperature and 30 min at 65°C with 2× SSC buffer containing 0.1% SDS followed by 0.2 × SSC containing 0.1% SDS at 65°C [12]. ^{32}P was detected by exposure to Kodak film (XOMAT AR) overnight at −70°C with an intensifying screen. Alternatively, the membrane was also exposed to phosphorimaging screen and scanned after an overnight exposure.

DNA sequencing

Dideoxy dNTP dye terminator cycle sequencing was performed according to the manufacturer's instructions (Applied Biosystems). Sequence comparisons were performed with the CLUSTALW method [13].

Results

Isolation of partial cDNA encoding goat MPR 300

To get a partial cDNA sequence of the putative MPR 300 protein, we used the single strand cDNA of goat liver tissue prepared, in a PCR experiment employing the following primer pairs (sense primer A and anti sense primer B) that were designed by multiple sequence alignment of the available sequences of the mammalian (human, bovine, rat, mouse) and chicken MPR 300. In PCR a product equal to 800 bp size was amplified covering partial 1st domain, complete 2nd domain and partial 3rd domain of MPR 300. Therefore the fragment obtained was gel purified for TA cloning into pCR2.1 vector. Sequencing with vector derived primers revealed a 800 bp fragment, that at amino acid level exhibited greater then 95–98% similarity to the bovine and other known MPR proteins (data not shown). In order to identify if the ligand binding region (3rd domain in other MPR 300 proteins) is also conserved in goat receptor, we have used the following primer pairs (sense primer A and anti sense primer C). A 1.368 kb fragment was amplified, which was gel purified and cloned as described above. When the plasmid DNA was isolated and digested with EcoRI as described under methods, the insert corresponding to 1.368 kb was released (Figure 1A). In order to determine the mRNA transcript size of the goat MPR 300 protein, the 1.368 kb fragment of goat MPR 300 cDNA, was used as a probe in Northern blotting which gave a 9.3 kb band (Figure 1B). Sequencing of

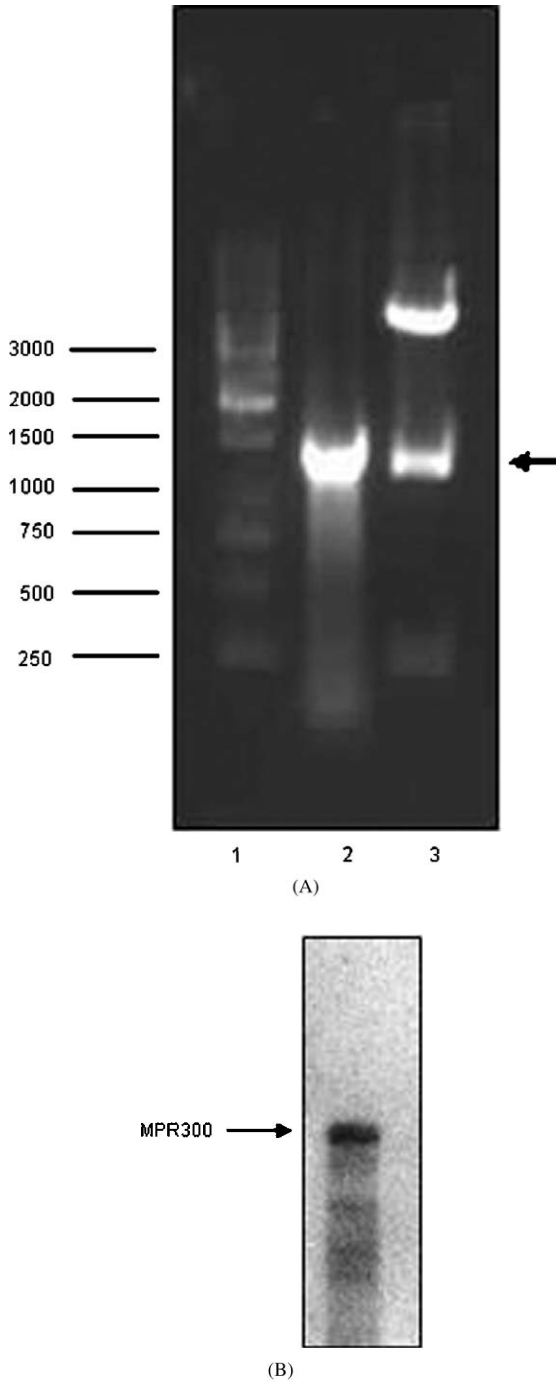

Figure 1. (A) Goat total RNA was used for the first strand synthesis using MPR specific primer (anti sense primer C). For the amplification 2 μl of first strand synthesis product was used along with MPR 300 specific primers (sense primer A and anti sense primer C). Amplified product was subjected to 1% agarose gel electrophoresis. Lane 1, standard DNA ladder, Lane 2, amplified product (1.368 kb), Lane 3, (←) fragment released from the vector after digestion with E.CoRI. (B) Northern Blot Analysis. 15 μg of total RNA isolated from goat liver tissue and subjected to denaturing 1% agarose gel electrophoresis, transferred to hybond-N nylon membrane and hybridized with ^{32}P labeled goat MPR 300 specific cDNA fragment (1.368 kb).

the 1.368 kb fragment at amino acid level exhibited greater then 95–98% homology to other known MPR 300 proteins (Figure 2). This homology is strikingly similar in the highly conserved cysteine residues, three potential glycosylation sites and the arginine residue which is known to be involved in the ligand binding at 3rd domain of other known MPR 300 proteins (Figure 3). This cDNA clone covers the partial 1st domain, complete 2nd, 3rd and partial 4th domain structures of the receptor. These data allowed us to conclude that the 1.368 kb fragment was derived from goat MPR 300 RNA.

Isolation of full length cDNA clone encoding goat MPR 46

To obtain a partial cDNA clone for the putative goat MPR 46 protein, single strand cDNA prepared by RT reaction was used employing the primer described under methods. In PCR a 760 bp fragment was obtained employing the following primers (sense primer 1 and anti sense primer 3). This was subcloned into TA vector (pCR 2.1, Invitrogen) and the plasmid DNA isolated and sequenced. The results showed extensive homologies to the other known MPR proteins (data not shown). In order to obtain a full length clone of the goat MPR 46 protein, we have used the following primer combinations (sense primer 2 anti sense primer 3). This resulted in obtaining a fragment of 840 bp length which was subcloned into TA vector (pTZ57R, MBI Fermentas). Plasmid DNA was isolated and digested with *Kpn*I and *Pst*I, resulting in release of the 840 bp insert (Figure 4A). In order to determine the mRNA transcript size of the goat MPR 46 protein, the 840 bp fragment of goat MPR 46 cDNA, was used as a probe in Northern blotting which gave a 2.3 kb band (Figure 4B).

From the results of nucleotide and deduced amino acid sequence (Figure 5), it is apparent that the single fragment of goat MPR 46 cDNA contains a 840 bp sequence corresponding to the ORF and the stop codon. The data clearly indicate the presence of a 24 amino acid putative signal sequence, an extracytoplasmic domain, transmembrane domain and cytoplasmic domain that are characteristic of MPR 46 proteins from other known animal species. Structural comparison of the goat sequences to other known receptor sequences is shown in Figure 6. These data allowed us to conclude that the 840 bp fragment was derived from goat MPR 46 RNA.

Discussion

The present study describes for the first time isolation and characterization of cDNA clones for goat receptors another representative species of mammals that has been shown by us to contain the putative Mannose 6-phosphate receptor proteins [10].

Molecular cloning of goat MPR 300

The known cDNA sequences of MPR 300 proteins in the coding region are in the range of 7–7.4 kb prolonged to 9.5–11 kb by 3′- and 5′-untranslated regions [3,14–16]. In a recent study

```
1     CTG TGC AGT TAC ACA TGG GAA GCA GCG GAT ACC AAA AAT AAC ATG   45
1      L   C   S   Y   T   W   E   A   A   D   T   K   N   N   M   15
46    CTT TAT AAA ATC AAC ATC TGT GGA AAT ATG GGT ATT GCC CAG TGT   90
16     L   Y   K   I   N   I   C   G   N   M   G   I   A   Q   C   30
91    GGA CCA TCA AGT GCT GTT TGT ATG CAT GAC TTG AAG ACA GAC AGC   135
31     G   P   S   S   A   V   C   M   H   D   L   K   T   D   S   45
136   TTT CAT TCT GTG GGT GAC TCT CTT TTG AAA ACC GCA AGC AGA TCT   180
46     F   H   S   V   G   D   S   L   L   K   T   A   S   R   S   60
181   CTT CTG GAA TTC AAC ACA ACA GTG AAC TGT AAG CAG CAG AAT CAC   225
61     L   L   E   F   N   T   T   V   N   C   K   Q   Q   N   H   75
226   AAA ATT CAG AGT AGC ATC ACC TTC TTA TGT GGG AAA ACC TTG GGA   270
76     K   I   Q   S   S   I   T   F   L   C   G   K   T   L   G   90
271   ACT CCC GAG TTT GTA ACT GCA ACA GAT TGT GTG CAT TAC TTC GAG   315
91     T   P   E   F   V   T   A   T   D   C   V   H   Y   F   E   105
316   TGG AGG ACT ACT GCA GCC TGC AAG AAG GAT ATA TTT AAA GCA AAT   360
106    W   R   T   T   A   A   C   K   K   D   I   F   K   A   N   120
361   AAA GAG GTG CCA TGT TAC GCT TTC GAC AGA GGG CTC AAG AAG CAT   405
121    K   E   V   P   C   Y   A   F   D   R   G   L   K   K   H   135
406   GAT TTA AAC CCA CTG ATC AAG ACC AGC GGT GCT TAC TTG GTG GAC   450
136    D   L   N   P   L   I   K   T   S   G   A   Y   L   V   D   150
451   GAC TCT GAC CCG GAT ACA TCT CTG TTC ATC AAT GTC TGC AGG GAC   495
151    D   S   D   P   D   T   S   L   F   I   N   V   C   R   D   165
496   ATA GAT GCG CTC CGG GCC TCG AGT CCG CGA GTG CGT GTG TGT CCC   540
166    I   D   A   L   R   A   S   S   P   R   V   R   V   C   P   180
541   CCC GGC GCG GCC GCC TGC CTG GTG AGA GGG GAC CGC GCG TTC GAC   585
181    P   G   A   A   A   C   L   V   R   G   D   R   A   F   D   195
586   GTG GGC CGG CCC CAG GAG GGG CTG AAG CTT GTG AGC AAT GAC AGG   630
196    V   G   R   P   Q   E   G   L   K   L   V   S   N   D   R   210
631   CTC GTC TTG AGT TAT GTG AAG GAG GGG GCC GGC CAG CCA GAC TTC   675
211    L   V   L   S   Y   V   K   E   G   A   G   Q   P   D   F   225
676   TGT GAC GGC CAC AGC CCA GCG GTG ACC ATC ACG TTC GTG TGC CCG   720
226    C   D   G   H   S   P   A   V   T   I   T   F   V   C   P   240
721   TCG GAG CGC AGA GAG GGC ACC ATT CCC AAG CTC ACG GCG AAA TCC   765
241    S   E   R   R   E   G   T   I   P   K   L   T   A   K   S   255
766   AAC TGC CGC TTT GAG ATC GAG TGG GTC ACC GAG TAC GCC TGC CAC   810
256    N   C   R   F   E   I   E   W   V   T   E   Y   A   C   H   270
811   AGG GAT TAC CTG GAA AGC CAG AGC TGC TCC CTG AGC AGC GCG CAG   855
271    R   D   Y   L   E   S   Q   S   C   S   L   S   S   A   Q   285
856   CAT GAC GTG GCT GTC GAC CTC CAG CCG CTG AGC CGG GTG GGA GAC   900
286    H   D   V   A   V   D   L   Q   P   L   S   R   V   G   D   300
901   TCC TTG TTC TAC ACT TCG GAG GCA GAT GAG TAT ACA TAT TAT TTG   945
301    S   L   F   Y   T   S   E   A   D   E   Y   T   Y   Y   L   315
946   AGC GTC TGC GGA GGA AGC CAA GTG CCC ATC TGT AAT AAG AAA GAT   990
316    S   V   C   G   G   S   Q   V   P   I   C   N   K   K   D   330
991   GCT GCG GTG TGC CAA GTG AAA AAG GCA GAT TCC ACT CAA GTC AAA   1035
331    A   A   V   C   Q   V   K   K   A   D   S   T   Q   V   K   345
1036  GTG GCC GGG AGA CCC CAG AAC CTG ACC CTC CGG TAC TCG GAT GGA   1080
346    V   A   G   R   P   Q   N   L   T   L   R   Y   S   D   G   360
1081  GAC CTC ACC TTG ATC TAT TTT GGG GGT GAG GAG TGC AGC TCC GGC   1125
361    D   L   T   L   I   Y   F   G   G   E   E   C   S   S   G   375
1126  TTC CAG CGG ATG AGT GTC ATC AAC TTC GAG TGC AAT CAG ACA GCA   1170
376    F   Q   R   M   S   V   I   N   F   E   C   N   Q   T   A   390
1171  GGT AAC AAC GGC AGA GGG GCT CCT GTG TTC ACC GGG GAG GTG GAC   1215
391    G   N   N   G   R   G   A   P   V   F   T   G   E   V   D   405
1216  TGC ACC TAC TTC TTC ACG TGG GAC ACG AAG TAC GCC TGC GTG CAC   1260
406    C   T   Y   F   F   T   W   D   T   K   Y   A   C   V   H   420
1261  GAG AAG GAG GCC CTG CTG TGC AGC GTC TCT GAC GGG AAA CAG CGC   1305
421    E   K   E   A   L   L   C   S   V   S   D   G   K   Q   R   435
1306  TTT GAC CTG TCG GCA CTG GCC CGG CAC TCA GAA CTG GAA CAG AAT   1350
436    F   D   L   S   A   L   A   R   H   S   E   L   E   Q   N   450
1351  TGG GAA GCT GTG GAT GGC   1368
451    W   E   A   V   D   G
```

Figure 2. Nucleotide and deduced amino acid sequence of 1.368 kb partial cDNA clone of goat MPR 300.

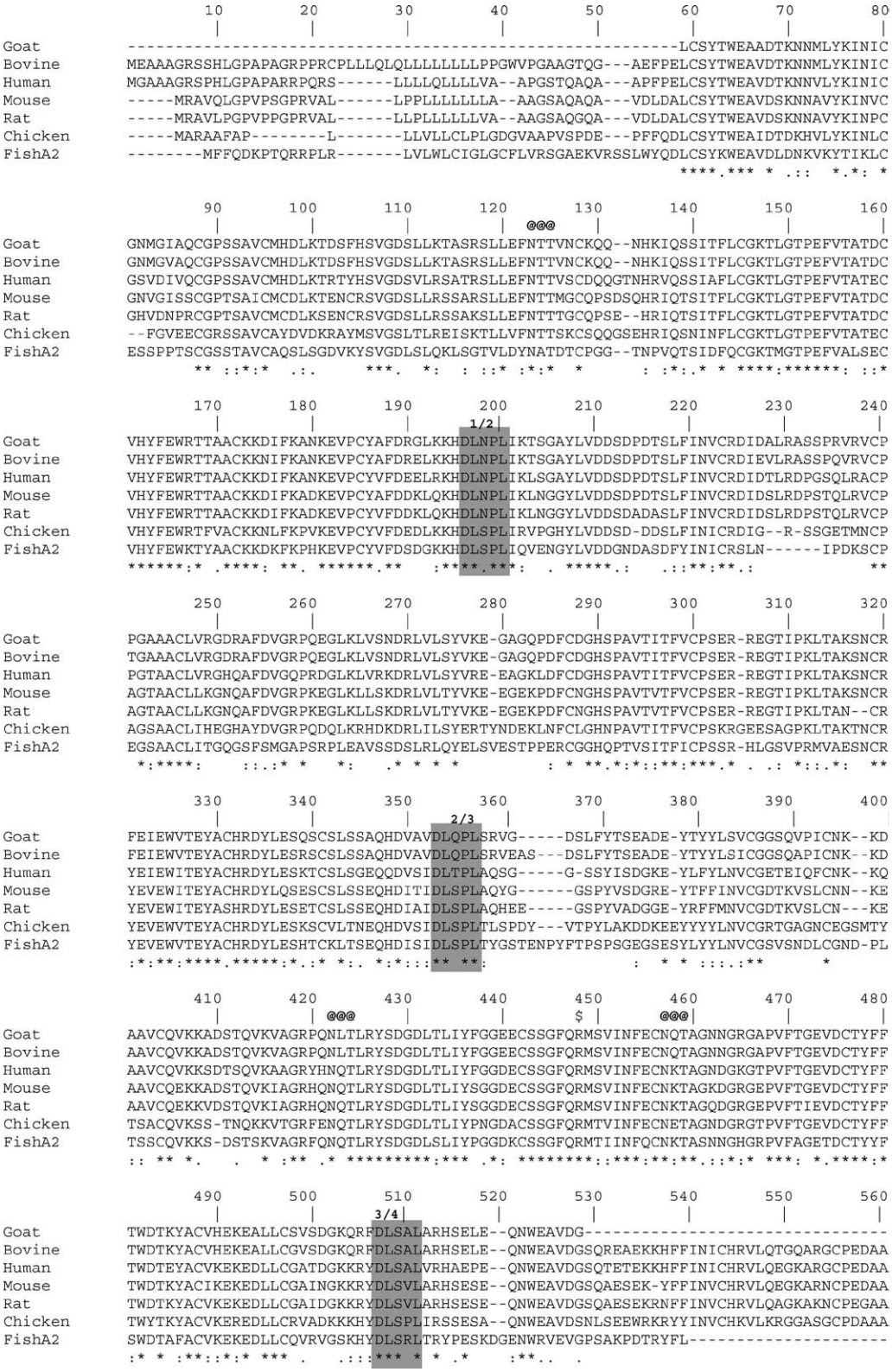

Figure 3. Alignment of the amino acid sequence of 1.368 kb fragment of goat MPR 300. From the partial cDNA sequence of goat MPR 300, domains 1–4 were having conserved motif DLSXL at the carboxyterminal border of each domain. Domains were aligned by the Clustalw method. *: indicates the conserved residues. The region between domain and domain is represented in shade. @@@: indicates potential N-glycosylation sites. $: indicates the arginine residue involved for ligand binding.

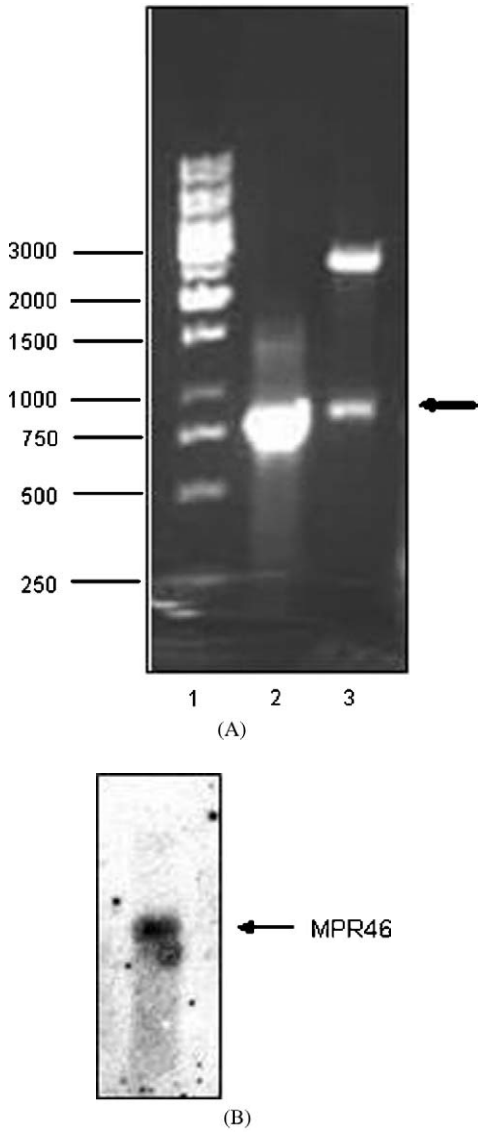

Figure 4. (A) For the PCR amplification of full length clone, 2 μl of first strand synthesis product was used along with MPR 46 specific primers (sense primer 2 and anti sense primer 3). Amplified product was subjected to 1% agarose gel electrophoresis. Lane 1, standard DNA ladder, Lane 2, amplified product (840 bp), Lane 3, (←) fragment released from the vector after digestion with KpnI and PstI. (B) Northern Blot Analysis. 15 μg of total RNA isolated from goat liver tissue and subjected to denaturing 1% agarose gel electrophoresis, transferred to hybond-N nylon membrane and hybridized with ^{32}P labeled goat MPR 46 full length cDNA clone (840 bp).

a partial cDNA clone obtained for the fish MPR 300 protein by us revealed that the receptors in the vertebrates have conserved cassette structures in the extracytoplasmic region [6]. Since the goat receptors have been biochemically characterized in our laboratory [11], we set out to obtain their sequences using a RT-PCR approach. Several lines of evidences support that the partial cDNA clone (1.368 kb) obtained in this study indeed represents the goat MPR 300 protein. First the amino terminal portion of the sequence obtained is divided into domains which show similar length as their chicken and other mammalian counterparts. Second, the carboxyl terminal border of each cassette is defined by the conserved motif DLS(P/R/S)L, which strongly resembles the corresponding sequences found in mammalian and chicken MPRs (DLXXL). Additionally the partial MPR 300 sequence obtained by us for the fish receptor [6] further supports the data obtained for goat MPR 300, that the repetitive domain structure as well as the carboxy terminal of each domain is conserved throughout vertebrates. Third, all the conserved cysteines which have been suggested to form disulfide bonds in bovine MPR 300 [17] and present in all vertebrate species studied so far were also seen in the goat receptor. Fourth, three glycosylation sites were found in the sequences obtained. Fifth, the critical arginine residue involved in ligand binding is located in the third domain of the goat receptor.

For bovine, human and chicken MPR 300 proteins, domains 3 and 9 have been shown to contain two independent binding sites for mannose 6-phosphate containing ligands. Binding of mannose 6-phosphate in both cases depends on a critical arginine residue (R435 in domain 3 and R1334 in domain 9) [17]. Further, recent partial cDNA analysis of the fish MPR 300 also reveals the presence of the critical arginine residue in domain 3 [6] that is important for ligand binding. The sequence obtained for goat MPR 300 reveals the presence of a corresponding arginine residue (R435) in domain 3, which is surrounded by the highly conserved sequence motif CSSGFQRM (S/T)(V/I)INF (E/Q)C. Taken together these data allow us to conclude that the cDNA clone indeed represents goat MPR 300 protein. In addition to the critical arginine residue, conserved in domain 3, other residues such as glutamine, serine, glutamic acid and tyrosine identified for the bovine MPR 300 protein to be important in carbohydrate binding [1] are also conserved in the goat MPR 300 protein.

Molecular cloning of goat MPR 46

The known MPR 46 cDNA sequences are in the range of 830–840 bp [1]. In the present study, first an attempt was made to obtain a partial cDNA clone for the goat MPR 46 spanning the extracytoplasmic region of the receptor. Initially, using the total RNA from the liver tissue, first strand cDNA synthesis was accomplished using the anti sense primer 3. This RT product was then amplified by PCR (sense primer 2 and anti sense primer 3) and the product (760 bp) cloned into the vector. From the sequence results obtained for the partial cDNA clone, it is clear that the fragment (760 bp) represents the MPR 46 cDNA.

By using the sense primer 2 and anti sense primer 3 in the PCR reaction a fragment of 840 bp was amplified. From the sequence results obtained, it is clear that the cDNA isolated is the putative cDNA for MPR 46.

```
1     ATG ATG TCC CCC CTC CAC AGC TCC TGG AGG ACT GGC CTG CTC CTG   45
1      M   M   S   P   L   H   S   S   W ▼R   T   G   L   L   L    15
46    CTG CTG CTC TTC TCC ATG GCA GTC AGA GAA TCT TGG CAG ACT GAA   90
16     L   L   L   F   S   M   A   V   R   E   S   W   Q   T   E    30
91    GAG AAA ACA TGC GAC CTG GTG GGA GAA AAG GGT AAA GAA TCA GAG   135
31     E   K   T   C   D   L   V   G   E   K   G   K   E   S   E    45
136   AAA GAG TTG GCT CTC CTG AAG AGG CTG ACA CCG CTA TTT AAC AAA   180
46     K   E   L   A   L   L   K   R   L   T   P   L   F   N   K    60
181   AGC TTT GAG AGC ACC GTG GGC CAG AGC CCA GAT ATG TAC AGC TAT   225
61     S   F   E   S   T   V   G   Q   S   P   D   M   Y   S   Y    75
226   GTG TTC CGG GTG TGC CGA GAA GCT GGC AAC CAC TCC TCT GGG GCA   270
76     V   F   R   V   C   R   E   A   G   N   H   S   S   G   A    90
271   GGC CTC GTG CAG ATC AAC AAA AGT AAC GGG AAG GAG ACA GTA GTT   315
91     G   L   V   Q   I   N   K   S   N   G   K   E   T   V   V   105
316   GGG AGA TTC AAC GAG ACT CAG ATC TTC AAT GGA AGT AAT TGG ATC   360
106    G   R   F   N   E   T   Q   I   F   N   G   S   N   W   I   120
361   ATG CTG ATC TAT AAA GGG GGT GAT GAA TAT GAC AAC CAC TGT GGC   405
121    M   L   I   Y   K   G   G   D   E   Y   D   N   H   C   G   135
406   AGG GAG CAG CGG AGG GCA GTG GTG ATG ATC TCC TGC AAT CGA CAC   450
136    R   E   Q   R   R   A   V   V   M   I   S   C   N   R   H   150
451   ACT CTA GCG GAC AAT TTT AAC CCT GTG TCT GAG GAG CGA GGC AAA   495
151    T   L   A   D   N   F   N   P   V   S   E   E   R   G   K   165
496   GTC CAA GAT TGT TTC TAC CTC TTT GAG ATG GAC AGC AGC CTG GCG   540
166    V   Q   D   C   F   Y   L   F   E   M   D   S   S   L   A   180
541   TGT TCC CCA GAG ATC TCC CAT CTT AGC GTG GGT TCT ATC TTA CTT   585
181    C   S   P   E   I   S   H   L   S   V   G   S   I   L   L   195
586   GTC ACG TTT GCA TCA CTG GTC GCA GTC TAT ATC ATC GGG GGG TTC   630
196    V   T   F   A   S   L   V   A   V   Y   I   I   G   G   F   210
631   CTG TAC CAG CGA CTG GTG GTC GGA GCC AAA GGA ATG GAG CAG TTT   675
211    L   Y   Q   R   L   V   V   G   A   K   G   M   E   Q   F   225
676   CCT CAC TTG GCC TTC TGG CAG GAT CTT GGA AAC CTG GTA GCA GAT   720
226    P   H   L   A   F   W   Q   D   L   G   N   L   V   A   D   240
721   GGC TGT GAC TTT GTA TGC CGC TCT AAA CCC CGA AAT GTG CCT GCT   765
241    G   C   D   F   V   C   R   S   K   P   R   N   V   P   A   255
766   GCC TAT CGT GGT GTG GGG GAT GAT CAG CTG GGG GAG GAG TCA GAA   810
256    A   Y   R   G   V   G   D   D   Q   L   G   E   E   S   E   270
811   GAA AGG GAT GAC CAC TTA TTA CCA ATG TGA   840
271    E   R   D   D   H   L   L   P   M   *
```

Figure 5. Nucleotide and deduced amino acid sequence of goat MPR 46 full length cDNA clone. (↓) Putative signal sequence cleavage site.

The most important observations and conclusions that can be drawn from these results are as follows. First, the entire cytoplasmic tail of the goat 46 protein is identical to other known MPR 46 proteins [3]. Second, the transmembrane domain is highly conserved among all receptors. Third, the cysteine residues that aid in disulfide pairing in the amino terminal domain of the receptor are highly conserved among all receptors. Fourth the ligand binding residue arginine in the extracytoplasmic domain is also conserved in goat protein. Further the potential N-glycosylation sites (five) have been found to be conserved in the goat receptor. Additionally from the crystal structure of bovine MPR 46, it is evident that several other residues are important and are found in the binding pocket. These include the cysteine 134 and 169 that are known to form the disulfide bonds. In addition to histidine 133 and arginine 139, aspartic acid 131 and glutamic acid 129 are known to interact with Mn^{+2}. Additionally several amino acids have been implicated to play a crucial role in forming hydrogen bonds with the hydroxyl groups of mannose [18]. The sequences obtained for goat MPR 46 (Figure 6) exhibit high degree of homology to the bovine receptor in the sugar binding pocket. The present study provides the first evidence to show that the goat MPR 46 protein sequences are homologous to other known MPR 46 sequences and to the partial chicken MPR 46 sequence [3], suggesting that this receptor is highly conserved.

From literature it is also known that the mammalian MPR 300 and 46 sequences exhibit homology with each other in the extracytoplasmic region [9]. The extracytoplasmic domain sequence of goat MPR 46 also exhibits 26–28% homology with bovine, human and mouse CI-MPR proteins. From Table 1, it is evident that both MPR 300 and 46 receptor functional regions are evolutionarily conserved among different species in the vertebrates. Only a partial sequence of the fish MPR 300 protein is published [6]. Recent biochemical and immunological studies revealed the presence of both putative receptors in the invertebrate mollusc [8]. It would be interesting to obtain complete sequence information of the fish and mollusc receptors, to establish the evolutionary pattern of both the receptors in the animal kingdom. Subsequent analysis of their functions in non-mammalian vertebrates and invertebrates, should provide useful information on structure-function relationships of these interesting proteins. Current work in our laboratory is focused on these lines.

Table 1. Similarity in various regions of the goat MPR 300 and MPR 46 (*italics*) to other known MPR proteins

Species	*Size of mature protein (kDa)*	*cDNA clone size (kb)*	*Repetitive domain sequence/* transmembrane domain	*Conservation of cystein residues*	*Ligand binding region sequences (in 3rd domain for MPR 300)*	*Potential N-glycosylation sites*	*Size of RNA transcript (kb)*
Bovine	300	7.497	DL(N/Q)(P/A)L	Yes	CSSCHQ**R**MSV	Conserved	~9.5
	46	*0.840*	*ILLVTLASLVAVYIIGGFLY*	*Yes*	*DN**H**CGREQ**R**RA*	*Conserved*	*~2.3*
Human	300	7.473	DL(N/T/S)(P/A)L	Yes	CSSGFQ**R**MSV	Conserved	~9.4
	46	*0.837*	*ILLVTFASLVAVYVVGGFLY*	*Yes*	*DN**H**CGKEQ**R**RA*	*Conserved*	*~2.3*
Mouse	300	7.446	DL(N/S)(P/V)L	Yes	CSSGFQ**R**MSV	Conserved	~10.0
	46	*0.837*	*ILLVIFASLVAVYIIGGFLY*	*Yes*	*DN**H**CGKEQ**R**RA*	*Conserved*	*~2.3*
Goat	300	Partial	DL(N/Q/S)(P/A)L	Yes	CSSGFQ**R**MSV	Conserved	~9.3
	46	*0.840*	*ILLVTFASLVAVYIIGGFLY*	*Yes*	*DN**H**CGREQ**R**RA*	*Conserved*	*~2.3*
Chicken	300	7.410	DLSPL	Yes	CSSGFQ**R**MTV	Conserved	~11.0
	46	*Partial*	*ILLITFSALVTVYIVGGFLY*	*Yes*	–	*Conserved*	*~3.3*
Fish	300	Partial	DLS(P/R)L	Yes	CSSGFQ**R**MTI	Conserved	~13.5
	46	–	–	–	–	–	–

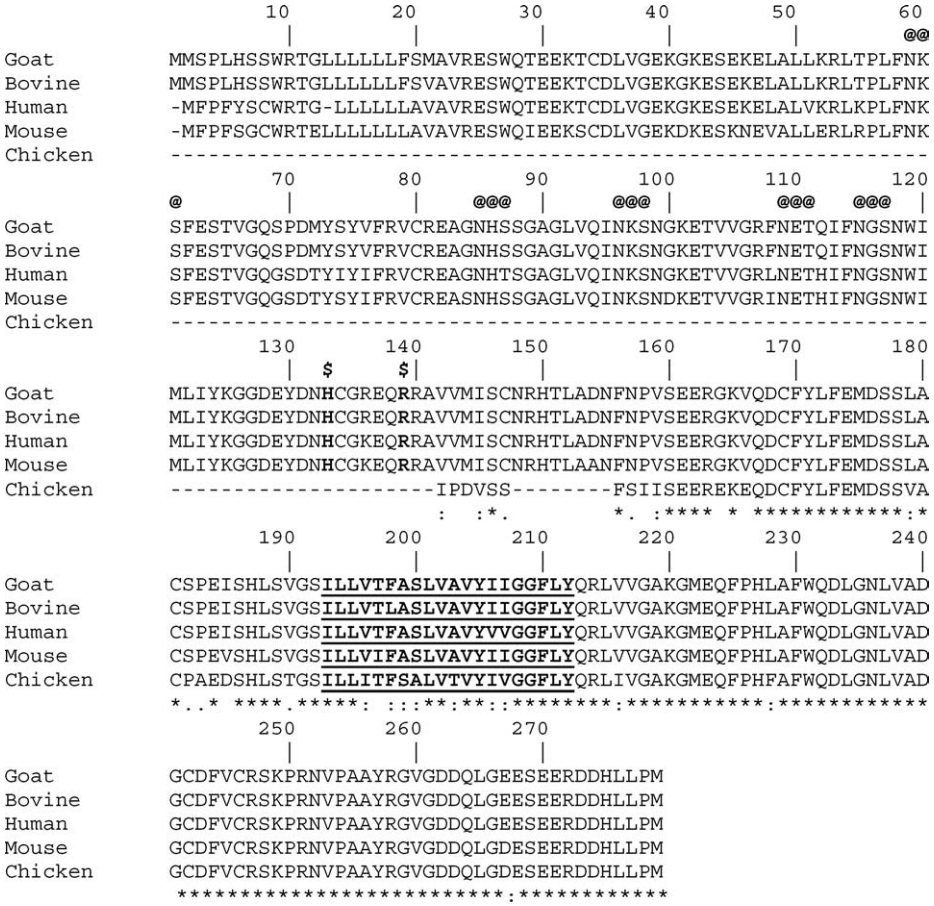

Figure 6. Alignment of the amino acid sequence of full length goat MPR 46. Amino acids were aligned by clustalw method. *: indicates the conserved residues. @@@: indicates potential N-glycosylation sites. $: indicates the arginine and histidine residues involved in ligand binding. ___: indicates the transmembrane domain.

Acknowledgments

The authors thank Department of Science and Technology, New Delhi, India, for a research support to carry out this work (SP/SO/D21/97). The authors also acknowledge the financial assistance provided by VWS, Germany (I/78913) that also helped in carrying out these studies.

References

1 Dahms NM, Hancock MK, P-type lectins, *Biochim Biophys Acta* **1572**, 317–40 (2002).
2 Zhou M, Ma Z, Sly WS, Cloning and expression of the cDNA of chicken cation-independent mannose 6-phosphate receptor, *Proc Natl Acad Sci USA* **92**, 9762–6 (1995).
3 Matzner U, Annette HR, von Figura K, Pohlmann R, Expression of mannose 6-phosphate receptors in chicken, *Dev Dynamics* **207**, 11–24 (1996).
4 Siva Kumar N, Hille-Rehfeld A, von Figura K, Mannose 6-phosphate receptor proteins from reptiles and amphibians: Evidence for the presence of MPR 300 and MPR 46, *Comp Biochem and Physiol* **118B**, 805–9 (1997).
5 Siva Kumar N, Udaya Lakshmi Y, Hille-Rehfeld A, von Figura K, Mannose 6-phosphate receptors (MPR 300 and MPR 46) from a teleostean fish (trout), *Comp Biochem and Physiol* **123B**, 261–5 (1999).
6 Udaya Lakshmi Y, Siva Kumar N, Schu P, von Figura K, Rehfeld AH, Conserved cassette structure of vertebrate Mr 300 kDa mannose 6-phosphate receptors: Partial cDNA sequence of fish MPR 300, *Comp Biochem Physiol* **127**, 433–41 (2000).
7 Udaya Lakshmi Y, Radha Y, Hille-Rehfeld A, von Figura K, Siva Kumar N, Identification of the putative mannose 6-phoshate receptor protein (MPR 300) in the invertebrate *unio, Biosci Reports* **15**, 403–9 (1999).
8 Siva Kumar N, von Figura K, Identification of the putative mannose 6-phosphate receptor (MPR 46) protein from the invertebrate mollusk, *Biosci Rep* **22**, 513–21 (2002).
9 Hille-Rehfeld A, Mannose 6-phosphate receptors in sorting and transport of lysosomal enzymes, *Biochim Biophys Acta* **1241**, 177–94 (1995).
10 Suresh K, Ramanadham M, Siva Kumar N, An ELISA method to quantify mannose 6-phosphate receptors, *J Biochem Biophys Meth* **52**, 111–9 (2002).
11 Suresh K, Siva Kumar N, An immno-affinity method for the purification of mannose 6-phosphate receptor proteins, *J Biochem Biophys Meth* **57**, 237–45 (2003).
12 Sambrook J, Fritsch EF, Maniatis T, *Molecular Cloning: A Laboratory Manual*, 2nd ed. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989).
13 Combet C, Blanchet C, Geourjon C, Deleage G, *Trends Bioch Sci* **25**, 147–50 (2000).
14 Lobel P, Dahms NM, Kornfeld S, Cloning and sequence analysis of the cation independent mannose 6-phosphate receptor. *J Biol Chem* **263**, 2563–70 (1988).
15 Oshima A, Nolan CM, Kyle J, Grubb HJ, Sly WS. The human cation independent mannose 6-phosphate receptor. Cloning and sequence of the full length cDNA and expression of functional receptor in COS cells, *J Biol Chem* **263**, 2553–62 (1988).
16 Szebenyi G, Rotwein P, The mouse insulin like growth factor II/cation independent mannose 6-phosphate (IGF-II/MPR) receptor gene. Molecular cloning and genome organization, *Genomics* **19**, 120–9 (1994).
17 Dahms NM, Rose PA, Molkentin JD, Zhang Y, Brzycki MA, The bovine mannose 6-phosphate/insulin-like growth factor II receptor, *J Biol Chem* **268**, 5457–63 (1993).
18 Roberts LD, Weix DJ, Dahms NM, Kim JP, Molecular basis of lysosomal enzyme recognition: Three-dimensional structure of the cation-dependent mannose 6-phosphate receptor, *Cell* **93**, 639–48.

Received 14 October 2003; revised 25 November 2003;
accepted 21 January 2004