

# Biochemical Characterization of CA IX, One of the Most Active Carbonic Anhydrase Isozymes\*<sup>§</sup>

Received for publication, February 5, 2008, and in revised form, July 25, 2008 Published, JBC Papers in Press, August 13, 2008, DOI 10.1074/jbc.M800938200

Mika Hilvo<sup>†1</sup>, Lina Baranauskiene<sup>§</sup>, Anna Maria Salzano<sup>¶</sup>, Andrea Scaloni<sup>¶</sup>, Daumantas Matulis<sup>§</sup>, Alessio Innocenti<sup>||</sup>, Andrea Scozzafava<sup>||</sup>, Simona Maria Monti<sup>\*\*</sup>, Anna Di Fiore<sup>\*\*</sup>, Giuseppina De Simone<sup>\*\*</sup>, Mikaela Lindfors<sup>‡</sup>, Janne Jänis<sup>‡‡</sup>, Jarkko Valjakka<sup>‡</sup>, Silvia Pastoreková<sup>§§</sup>, Jaromir Pastorek<sup>§§</sup>, Markku S. Kulomaa<sup>‡</sup>, Henri R. Nordlund<sup>‡‡</sup>, Claudiu T. Supuran<sup>||</sup>, and Seppo Parkkila<sup>†¶</sup>

From the <sup>†</sup>Institute of Medical Technology and <sup>¶</sup>School of Medicine, University of Tampere and Tampere University Hospital, FI-33014 Tampere, Finland, the <sup>§</sup>Laboratory of Biothermodynamics and Drug Design, Institute of Biotechnology, LT-02241 Vilnius, Lithuania, the <sup>¶</sup>Proteomics and Mass Spectrometry Laboratory, ISPAAM, National Research Council, 80147 Naples, Italy, the <sup>||</sup>Bioinorganic Chemistry Laboratory, University of Florence, 50019 Sesto Fiorentino (Florence), Italy, the <sup>\*\*</sup>Institute of Biostructures and Bioimages, National Research Council, 80134 Naples, Italy, the <sup>‡‡</sup>Department of Chemistry, University of Joensuu, FI-80101 Joensuu, Finland, the <sup>§§</sup>Centre of Molecular Medicine, Institute of Virology, Slovak Academy of Sciences, 84505 Bratislava, Slovak Republic

Carbonic anhydrase IX (CA IX) is an exceptional member of the CA protein family; in addition to its classical role in pH regulation, it has also been proposed to participate in cell proliferation, cell adhesion, and tumorigenic processes. To characterize the biochemical properties of this membrane protein, two soluble recombinant forms were produced using the baculovirus-insect cell expression system. The recombinant proteins consisted of either the CA IX catalytic domain only (CA form) or the extracellular domain, which included both the proteoglycan and catalytic domains (PG + CA form). The produced proteins lacked the small transmembrane and intracytoplasmic regions of CA IX. Stopped-flow spectrophotometry experiments on both proteins demonstrated that in the excess of certain metal ions the PG + CA form exhibited the highest catalytic activity ever measured for any CA isozyme. Investigations on the oligomerization and stability of the enzymes revealed that both recombinant proteins form dimers that are stabilized by intermolecular disulfide bond(s). Mass spectrometry experiments showed that CA IX contains an intramolecular disulfide bridge (Cys<sup>119</sup>-Cys<sup>299</sup>) and a unique *N*-linked glycosylation site (Asn<sup>309</sup>) that bears high mannose-type glycan structures. Parallel experiments on a recombinant protein obtained by a mammalian cell expression system demonstrated the occurrence of an additional *O*-linked glycosylation site (Thr<sup>78</sup>) and characterized the nature of the oligosaccharide structures. This study provides novel information on the biochemical properties of CA IX and may help characterize the various cellular and pathophysiological processes in which this unique enzyme is involved.

\* This work was supported in part by grants from the Finnish Cultural Foundation, European Union projects Euroxy and DeZnIT, the Lithuanian State Science and Studies Foundation, and the Italian National Research Council (RSTL 862 and MERIT). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>§</sup> The on-line version of this article (available at <http://www.jbc.org>) contains supplemental "Materials and Methods," Table S1, and Figs. S1 and S2.

<sup>†</sup> Deceased July 21, 2008.

<sup>1</sup> To whom correspondence should be addressed: Biokatu 6, FI-33014 University of Tampere, Finland. Tel.: 358-50-5347782; Fax: 358-3-35517710; E-mail: Mika.Hilvo@uta.fi.

Carbonic anhydrases (CAs)<sup>2</sup> form a family of zinc-containing metalloenzymes that catalyze the reversible hydration of carbon dioxide according to the following reaction:  $\text{CO}_2 + \text{H}_2\text{O} \leftrightarrow \text{H}^+ + \text{HCO}_3^-$  (1). Their main function is to maintain an appropriate acid-base balance in organisms; thus, they participate in several physiological processes, such as  $\text{CO}_2$  and  $\text{HCO}_3^-$  transport, bone resorption, production of body fluids, gluconeogenesis, ureagenesis, and lipogenesis (2). The CA family consists of 13 active isozymes in mammals, 12 of which are expressed and function in humans (3). The CA isozymes show various tissue expression patterns, unique kinetic and inhibitory properties, and different subcellular localizations. CAs I, II, III, VII, and XIII reside in the cytosol (2, 4, 5), CAs IV, IX, XII, XIV and XV (XV is not expressed in humans) are associated with the cell membrane (3, 6–10); CA VA and VB localize to mitochondria (11); and CA VI is secreted (12). In general, CAs exhibit a very high catalytic efficiency that approaches the diffusion-controlled limit (1). To date, CA II has been reported to represent the isozyme with the highest catalytic efficiency, with a  $k_{\text{cat}}/K_m = 1.5 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$  (13, 14).

CA IX and XII are tumor-related members of the CA family, and their expression is induced by hypoxia. Of these two, CA IX is the more exceptional because it has been proposed to contribute to cell proliferation, cell adhesion, and malignant cell invasion (15, 16). It is generally expressed in a very limited number of normal tissues (mainly in the gastrointestinal tract), whereas it is highly expressed in several cancers that arise from CA IX-negative tissues, such as renal, lung, cervical, ovarian, esophageal, and breast carcinomas (17). A cDNA coding for the human CA IX protein was first cloned and investigated by Pastorek *et al.* (18), and the CA9 gene was characterized in 1996 by the same group (19). It became evident from this study that CA IX shows significant similarity with the keratan sulfate attach-

<sup>2</sup> The abbreviations used are: CA, carbonic anhydrase; DTT, dithiothreitol; LS, light scattering; MALDI-TOF MS, matrix-assisted laser desorption/ionization-time of flight mass spectrometry; NeuGc, *N*-glycolylneuraminic acid; PSD, post-source decay; SEC, size exclusion chromatography; SES, stepwise elongation of sequence; HPLC, high pressure liquid chromatography; PNGase, peptide:*N*-glycosidase; ah, anti-human.

ment domain of a large human aggregating proteoglycan, aggrecan (19). Mature CA IX consists of the following four domains: an N-terminal proteoglycan-like domain (PG), a CA catalytic domain (CA), a transmembrane helical segment, and a short intracytoplasmic tail (Fig. 1A). However, no studies have proven the occurrence of glycosaminoglycan chains on the PG domain. Recently, CA IX has been suggested to represent a valuable target in cancer diagnostics and treatment. A promising example is Rencarex<sup>®</sup>, an anti-CA IX antibody that is currently in a phase III clinical trial ([www.wilex.com](http://www.wilex.com)) to treat renal cell carcinomas.

Despite all the information on CA IX that has accumulated during the past years, the biochemical characterization of this enzyme has remained incomplete. The main focus of the present study was to fill this gap in knowledge by providing information on the catalytic properties, oligomerization, and post-translational modifications of CA IX. Because it is widely reported that membrane proteins present difficulties for detailed biochemical characterization, soluble CA IX forms were produced using the baculovirus-insect cell expression system, which has not been used previously to investigate CA isoenzymes. These recombinant proteins were subsequently characterized by different biophysical and biochemical methods.

### EXPERIMENTAL PROCEDURES

**Construction of Recombinant Baculoviruses**—Two cDNA constructs were designed to produce the recombinant CA IX forms, both of which lack the small transmembrane and intracytoplasmic domains. Both recombinant proteins had a CA IX signal sequence whose codon usage was optimized for *Spodoptera frugiperda*, eight histidine residues for protein purification and a thrombin cleavage site for tag removal (Fig. 1). The PG + CA form coded for a protein consisting of both the PG and CA domains, encompassing residues 1–377 of the mature CA IX sequence (residues 38–414 of Swiss-Prot entry Q16790). At the N terminus, it contained the removable MAPLCPSPWLPLLIPAPAPGLTVQLLSLLLLMPVHPQRLHHHHHHHHLVPRGS (thrombin cleavage site underlined) sequence (Fig. 1B). The CA form coded only for the CA domain, encompassing residues 101–354 (residues 138–391 of Swiss-Prot entry Q16790), bearing the cleavable MAPLCPSPWLPLLIPAPAPGLTVQLLSLLLLMPVHP signal peptide at the N terminus and a LVPRGSHHHHHHHH sequence at the C terminus (Fig. 1C). cDNA sequences were constructed by the stepwise elongation of sequence-PCR (SES-PCR) (20). The baculoviral genomes encoding CA IX recombinant proteins were generated according to the Bac-To-Bac<sup>®</sup> Baculovirus Expression System instructions (Invitrogen). More detailed description on the SES-PCR and the generation of the baculoviral genomes can be found under supplemental “Materials and Methods.”

**Production and Purification of Recombinant CA IX**—The Sf9 insect cells were maintained in HyQ SFX-Insect serum-free cell culture medium (HyClone, Logan, UT) in an orbital shaker at 27 °C (125 rpm). The cells were centrifuged (2000 × g, 20 °C, 5 min) 72 h after infection, and the medium was collected. Purification was performed with the ProBond<sup>®</sup> Purification System

(Invitrogen) under native binding conditions with wash and elution buffers made according to the manufacturer's instructions. The purification procedure per 100 ml of insect cell medium was as follows: 1 liter of native binding buffer and 7 ml of the nickel-chelating resin were added to the insect cell medium, and the His-tagged protein was then allowed to bind to the resin on a magnetic stirrer at 25 °C for 1 h. The resin was washed with 500 ml of washing buffer, and the protein was eluted with elution buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 500 mM NaCl, 250 mM imidazole, pH 8.0).

**Buffer Change, Thrombin Treatment, and Concentration Measurement**—The purified CA form was changed to a buffer of 50 mM Tris-HCl, pH 7.5 (Sigma), using an Amicon Ultra 10-kDa cut-off centrifugal filter device (Millipore). Because the PG domain was unstable in Tris buffer (data not shown), the purified PG + CA form was changed to a buffer of 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 100 mM NaCl, pH 8.0. To remove the His tag, recombinant proteins were treated with 30 μl of resin-coupled thrombin (Thrombin CleanCleave KIT, Sigma) per 1 mg of protein with gentle shaking at 25 °C for 3 h, according to the manufacturer's instructions. The thrombin resin was removed from the protein solution by filtration. The His tag was removed from the protein solution of the PG + CA form by washing the protein using an Amicon Ultra 30-kDa cut-off centrifugal filter device (Millipore). The thrombin treatment was effective for the PG + CA form but failed for the CA form, despite using thrombin from three different manufacturers (Novagen, Sigma, and Amersham Biosciences), a large enzyme excess and long incubation times. Protein concentration was measured at three different dilutions by both the DC Protein Assay (Bio-Rad) and the BCA Protein Assay Reagent Kit (Pierce), and the mean of these results was used as the final value.

**Production of Polyclonal Antibody, SDS-PAGE, and Western Blotting**—Anti-human CA IX (ahCAIX) serum was raised by Innovagen AB (Lund, Sweden) in rabbits against the CA form of CA IX. Purified recombinant CA IX proteins were analyzed by 10% SDS-PAGE under reducing conditions (21) followed either by treatment with the Colloidal Blue Staining Kit (Invitrogen) or Western blotting, which was performed as described previously (22). In Western blots, both the monoclonal M75 (23) and the polyclonal ahCAIX antibodies were used to detect CA IX recombinant proteins; the dilutions were 1:200 and 1:1,000, respectively. The secondary antibodies (Amersham Biosciences) were horseradish peroxidase-labeled anti-mouse IgG for M75 and horseradish peroxidase-labeled anti-rabbit IgG for the ahCA IX antibody, and they were both diluted 1:25,000 in TBST buffer.

**Oligomer Reduction**—To study the covalent oligomerization of CA IX, 7 μg of both recombinant protein samples were incubated at 25 °C for 2 h, with different amounts of dithiothreitol (DTT); these samples were then analyzed by 9.5% SDS-PAGE under non-reducing conditions. The gels were stained using PageBlue<sup>™</sup> dye (Fermentas, Vilnius, Lithuania) according to the manufacturer's recommendations.

**Size Exclusion Chromatography followed by Light Scattering**—The light scattering data were collected using a Superdex 200 10/30 HR size exclusion chromatography (SEC) column (GE Healthcare), connected to high performance liquid chromatog-

raphy system, Alliance 2965 (Waters, Milford, MA), equipped with an autosampler. The elution from SEC was monitored by a photodiode array UV-visible detector 996 (Waters), a differential refractometer (OPTI-Lab or OPTI-rEx Wyatt, Santa Barbara, CA), and a static, multiangle laser light scattering detector (DAWN-EOS, Wyatt). The SEC-UV/LS/RI system was equilibrated in 20 mM HEPES, pH 8.0, 150 mM NaCl buffer at the flow rate of 0.5 ml/min. Two software packages were used for data collection and analysis: the Millennium software (Waters) controlled the HPLC operation and data collection from the multiwavelength UV-visible detector, whereas the ASTRA software (Wyatt) collected data from the refractive index detector, the light scattering detectors, and recorded the UV trace at 280 nm sent from the PDA detector. The peaks from the SEC were collected also for SDS-PAGE analysis under reducing and non-reducing conditions.

**Recombinant CA IX Proteins Produced in the Bacterial and Mammalian Expression Systems**—The recombinant glutathione *S*-transferase-CA construct encoding residues 84–360 of the mature CA IX protein was used in the CA activity assay and has been described previously (24). The production and purification of this glutathione *S*-transferase fusion protein in *Escherichia coli* has already been described (25). A sample of recombinant human CA IX protein produced in murine myeloma cell line NS0 was purchased from R&D Systems (Minneapolis, MN) and used in the mass spectrometric analyses.

**CA Activity Assay**—An Applied Photophysics stopped-flow instrument was used to assay CA-catalyzed CO<sub>2</sub> hydration activity (13). Reaction was measured using 0.2 mM phenol red ( $A_{\text{max}}$  at 557 nm) as an indicator, in 10 mM HEPES, 0.1 M Na<sub>2</sub>SO<sub>4</sub>, pH 7.5, for a period of 10–100 s. To determine the kinetic parameters and inhibition constants, the CO<sub>2</sub> concentration ranged from 1.7 to 17 mM. For the inhibitor assay, at least six replicates of the initial reaction were used to determine the initial velocity. A stock solution of 1 mM acetazolamide in 10–20% (v/v) dimethyl sulfoxide was used to prepare acetazolamide dilutions up to 0.01 nM. To form the enzyme-inhibitor complex, inhibitor and enzyme solutions were preincubated at 25 °C for 15 min, prior to inhibition measurements. Inhibition constants were obtained by non-linear least-squares methods using PRISM 3, whereas kinetic parameters were obtained from Lineweaver-Burk plots, as reported earlier, and represent the mean from at least three different determinations. Assays were performed on the PG + CA and CA forms from the baculovirus-insect cell expression system and on the CA domain of CA IX also from the *E. coli* expression system. The measurements for the proteins produced in the insect cells were also performed in the presence of 50 μM ZnCl<sub>2</sub>, 50 μM MgCl<sub>2</sub>, 50 μM CoCl<sub>2</sub>, or 50 μM MnSO<sub>4</sub>. In the presence of ZnCl<sub>2</sub>, the CO<sub>2</sub> hydration activity was measured also for other CA isozymes, namely I, II, IV, XII, XIV, and XV, and the references for their production can be found in Ref. 26.

**Protein Alkylation**—Protein samples for disulfide assignment were alkylated with 1.1 M iodoacetamide in 0.25 M Tris-HCl, 1.25 mM EDTA, and 6 M guanidinium chloride, pH 7.0, at 25 °C for 1 min in the dark. Samples were separated from excess salts and reagents by passing the reaction mixture through a

PD10 column (Amersham Biosciences), as previously reported (27). Protein samples were collected and lyophilized.

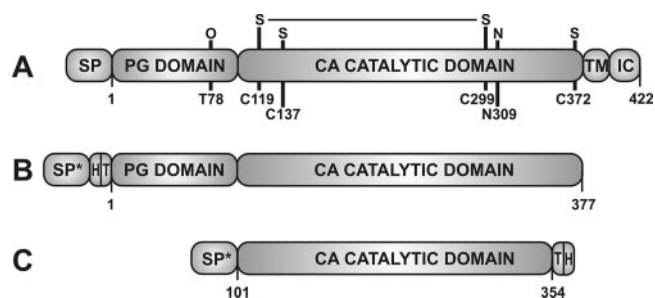
**Protein Digestion and Peptide Separation**—Bands were excised from the gel after SDS-PAGE, triturated, in-gel reduced, *S*-alkylated, and digested with trypsin, as previously reported (28). Gel particles were extracted with 25 mM NH<sub>4</sub>HCO<sub>3</sub>/acetonitrile (1:1 v/v) by sonication, and digests were concentrated. Peptide mixtures were either desalted using μZipTipC<sub>18</sub> pipette tips (Millipore) before direct MALDI-TOF MS analysis or resolved on a capillary Hypersil-Keystone Aquasil C<sub>18</sub> Kappa column (100 × 0.32 mm, 5-μm particle size) (ThermoElectron, Waltham, MA) using a linear gradient from 10 to 60% of acetonitrile in 0.1% formic acid, at a flow rate of 5 μl/min for 65 min. Collected fractions were concentrated and similarly analyzed by MALDI-TOF MS.

**PNGase F, Dithiothreitol, and Phosphatase Treatment**—Glycopeptides were deglycosylated by reacting with 0.2 unit of PNGase F (Roche) in 50 mM NH<sub>4</sub>HCO<sub>3</sub>, pH 8, at 37 °C for 12 h. Peptides containing disulfide bonds were reduced by reacting with 10 mM DTT in 50 mM NH<sub>4</sub>HCO<sub>3</sub>, pH 8, at 37 °C for 3 h. Potential phosphopeptides were dephosphorylated by reacting with 2 units of shrimp alkaline phosphatase (Roche) in 50 mM NH<sub>4</sub>HCO<sub>3</sub>, pH 8.5, at 37 °C for 1 h. 3 μl of 10% (v/v) formic acid was added to the reaction mixtures, and then they were desalted on μZipTipC<sub>18</sub> pipette tips (Millipore) before MALDI-TOF MS analysis.

**MS Analysis**—Intact recombinant proteins were analyzed with a Surveyor HPLC system connected *on-line* with an LCQ DecaXP Ion Trap mass spectrometer (ThermoElectron) equipped with an OPTON electrospray source operating at a needle voltage of 4.2 kV and a temperature of 320 °C. Analyses were performed using a narrow bore 250 × 2-mm C4 Jupiter column (300 Å pore size) (Phenomenex, Torrance, CA) with a gradient of solvent B (0.05% trifluoroacetic acid in CH<sub>3</sub>CN) to solvent A (0.08% trifluoroacetic acid in H<sub>2</sub>O) that ranged from 30 to 70% over a period of 40 min. Mass spectra were recorded within the *m/z* 400–2000 range in positive mode. Multicharge spectra were deconvoluted using the BioMass program implemented in the Bioworks 3.1 package provided by the manufacturer. Mass calibration was performed by using UltraMark (ThermoElectron) as an internal standard.

Whole protein digests or selected peptide fractions were loaded on the MALDI target together with 2,5-dihydroxybenzoic acid (10 mg/ml in 70% (v/v) acetonitrile, 0.06% trifluoroacetic acid) or 2,4,6-trihydroxyacetophenone (3 mg/ml in 50% (v/v) acetonitrile, 50% 20 mM diammonium citrate) as matrices, using the dried droplet technique. Samples were analyzed with a Voyager-DE PRO spectrometer (Applera, Norwalk, CT) operating with a 337-nm laser (28). Mass spectra were acquired in positive (with 2,5-dihydroxybenzoic acid) and negative (with 2,4,6-trihydroxyacetophenone) polarities, using the instrument either in reflectron or linear mode. Internal mass calibration was performed with peptides from trypsin autolysis or added molecular markers (Applera). Post-source decay (PSD) fragment ion spectra were acquired after isolation of the appropriate precursor, as previously reported (29). In both cases, data were elaborated using the DataExplorer 5.1 software (Applera).

## Biochemical Characterization of CA IX



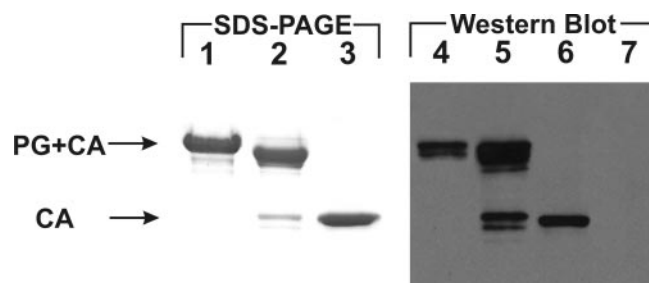
**FIGURE 1. Schematic figure of human CA IX and the recombinant CA IX proteins used in this study.** A, B, and C show native human CA IX, and the PG + CA and CA recombinant forms, respectively. In A, the N- and O-linked glycosylation sites as well as the cysteines are numbered according to the mature CA IX sequence. The corresponding amino acid residues in Swiss-Prot entry Q16790 are Thr<sup>115</sup>, Cys<sup>156</sup>, Cys<sup>174</sup>, Cys<sup>336</sup>, Asn<sup>346</sup>, and Cys<sup>409</sup>. Abbreviations: SP, signal peptide; SP\*, signal peptide whose codon usage has been optimized for *S. frugiperda*; H, polyhistidine tag (His<sub>6</sub>); T, thrombin cleavage site (sequence LVPRGS); TM, transmembrane helix; IC, intracytoplasmic tail.

Observed MALDI-TOF mass values were assigned to peptides, glycopeptides, or disulfide-linked peptides using the GPMW 4.23 software (Lighthouse Data, Odense, Denmark). This software generated a mass/fragment data base output based on protein sequence, protease selectivity, nature of the amino acids susceptible to eventual post-translational modifications, and the molecular mass of the modifying groups. Mass values were matched to protein regions using a 0.02% mass tolerance value. Post-translational modification assignments were always confirmed by additional MS experiments on modified peptides as treated above. All masses are reported as average values. Peptide numbering refers to mature human CA IX.

### RESULTS

**Production of Recombinant CA IX Forms**—Two soluble recombinant forms of human CA IX, both of which were missing the small transmembrane and intracytoplasmic domains at the C terminus, were produced using the baculovirus-insect cell expression system. The PG + CA form consisted of both the PG and CA catalytic domains, which encompass residues 1–377 in the mature CA IX sequence. The CA form included only the CA domain, encompassing residues 101–354 in the mature sequence (Fig. 1). Both recombinant proteins contained the CA IX signal sequence, which targets the protein to the secretory pathway within the cells, and the proteins were secreted into the medium due to the designated truncation at the C terminus. Both forms were purified to homogeneity with a single purification step, as clearly shown by SDS-PAGE analysis (Fig. 2). Only a single or doublet polypeptide band representing each CA IX product were visible on the gel. A thrombin site was included in the recombinant proteins to remove the His tag. Thrombin was active in removing the His tag for the PG + CA form (Fig. 2) but did not work for the CA form (data not shown). Thus, the His tag was removed from the PG + CA form for further analysis, whereas the tag was still present in the CA form.

The nature of both CA IX products was assayed by Western blot using two CA IX-specific antibodies, including M75 (a monoclonal antibody recognizing only the PG domain) and a novel polyclonal antibody (anti-human CA IX, ahCAIX) that recognizes only the CA domain. As expected, the PG + CA



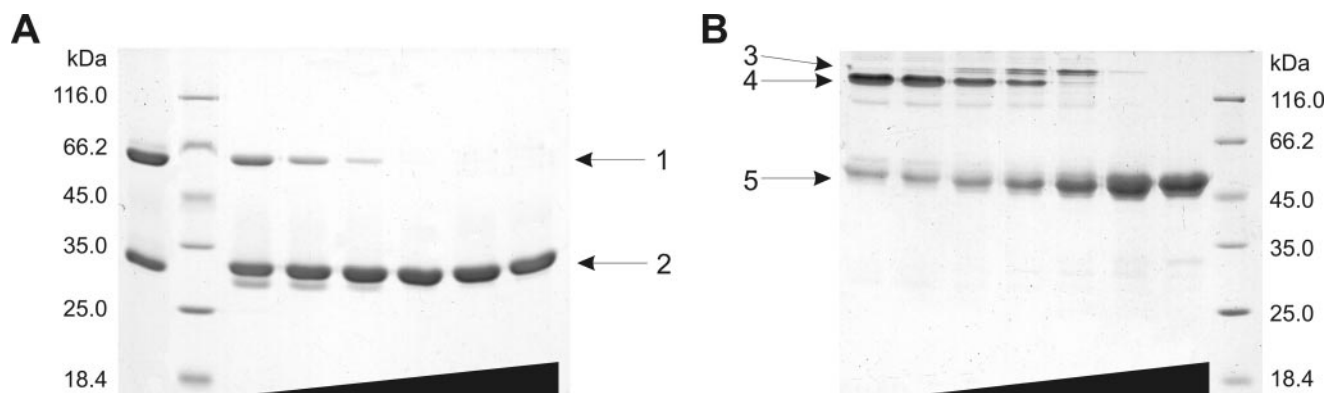
**FIGURE 2. SDS-PAGE and Western blot analysis of the recombinant proteins.** Protein samples applied to the lanes are as follows: 1) 5.5  $\mu$ g of the PG + CA form before His tag removal; 2) 5  $\mu$ g of the PG + CA form after His tag removal; 3) 3.2  $\mu$ g of the CA form before His tag removal; 4) 10 ng of the PG+CA form after His tag removal; 5) 100 ng of the PG+CA form after His tag removal; 6 and 7) 10 ng of the CA form before His tag removal. Lanes 1–3 were visualized by Coomassie staining. Lanes 4–7 were visualized by Western blotting with M75 (detecting the PG domain; lanes 4 and 7) and ahCAIX (detecting the CA domain; lanes 5–6) antibodies.

form was recognized by both antibodies (M75 and ahCAIX), whereas the CA form was only recognized by ahCAIX (Fig. 2). During thrombin treatment of the PG + CA form, a small amount of protein species having an unexpected lower mass (a double band) appeared in the protein fraction. The Western blot showed that the lower mass protein species represented the CA domain, from which the PG domain was missing. This may result from a suboptimal thrombin cleavage site between the PG and CA domains. Another suboptimal thrombin cleavage site at the C terminus is likely to explain the doublet band of the PG + CA form, because the only post-translational modification observed for this protein in the mass spectrometric analysis (as shown later) was an N-linked glycosylation, which was confirmed not to cause the double band, as treatment with PNGase F resulted in a shift of both bands in the Western blot (supplemental Fig. S1).

**Oligomerization State of the Recombinant Proteins**—The oligomerization of the recombinant CA IX proteins was investigated by SDS-PAGE under reducing and non-reducing conditions, SEC, and SEC combined with laser light scattering (SEC/LS).

Both recombinant proteins were incubated with various DTT concentrations and subjected to SDS-PAGE under non-reducing conditions (Fig. 3). SDS-PAGE results indicated that both CA IX recombinant proteins existed as oligomeric forms, in addition to monomers. The PG + CA form consisted of ~60% oligomeric and 34% monomeric species; the CA form consisted of about 54% oligomeric and 46% monomeric species (determined by densitometry). Incubating the samples with increasing concentrations of DTT gradually induced the disappearance of the oligomeric species, with a concomitant increase of the monomers, indicating that the oligomers are stabilized by intermolecular disulfide bond(s). For the CA form, the molecular weight standards indicated clearly that the protein is composed of monomers and dimers. For the PG + CA form, however, the size of the monomers was clearly exaggerated, and the size of the oligomers suggested trimerization rather than dimerization.

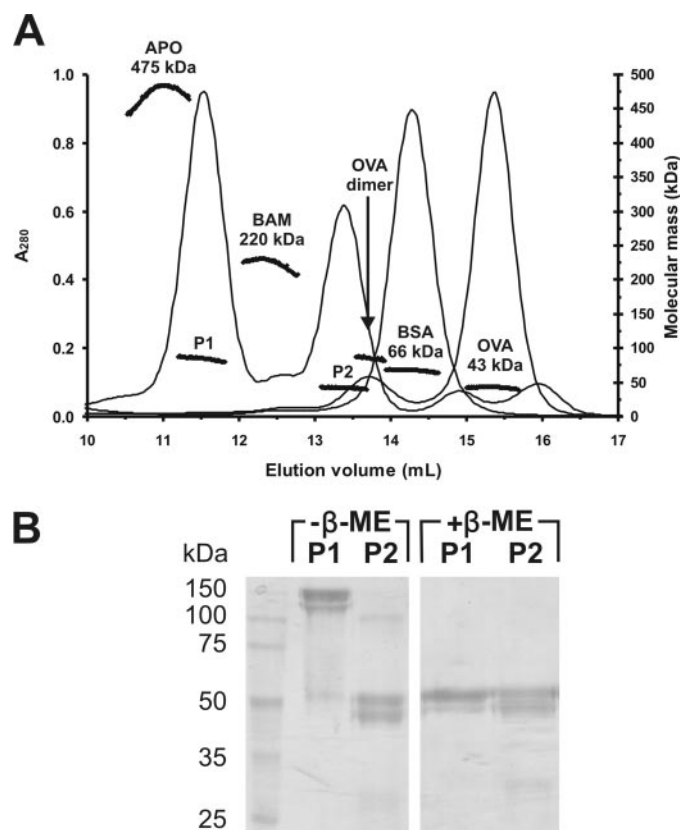
To investigate further the oligomerization of these proteins under native conditions, SEC and SEC/LS analysis were performed. In the SEC experiments the CA form showed two



**FIGURE 3. Effect of DTT on CA IX SDS-PAGE mobility.** A and B show the SDS-PAGE results for the CA and PG + CA forms, respectively. Increments of DTT concentration (lanes from left to right: 0, 0.01, 0.05, 0.1, 0.25, 1.0, and 10.0 mM DTT) are shown by *triangles* in the *bottom* of the gel figures. *Arrows* point to the following CA IX forms: 1) dimer of the CA form; 2) monomer of the CA form; 3) oligomer of the PG + CA form with intramolecular cysteines in a reduced state (-SH-SH); 4) oligomer of the PG + CA form with intramolecular cysteines in an oxidized state (-S-S-); 5) monomer of the PG + CA form.

major peaks, whose molecular mass values (calculated from a calibration curve with protein standards) clearly demonstrated the simultaneous presence of a protein monomer and dimer (data not shown), in accordance with the SDS-PAGE results. For the PG + CA form, the SEC analysis also showed the occurrence of two major peaks (P1 and P2 in Fig. 4A), whose elution position suggested an apparent molecular mass/hydrodynamic radius higher than expected, *i.e.* 334 kDa/5.5 nm and 121 kDa/4.0 nm, respectively (supplemental Fig. S2). These values were much higher than that obtained during mass spectrometry for the monomer (43 kDa, see below), thus suggesting an abnormal SEC migration of this form, probably depending on the presence of the PG domain with a non-globular structure. This hypothesis was confirmed by LS analysis of the eluted SEC fractions, which was performed under native conditions (30). In this case, P1 and P2 showed an average molecular mass value of 86 and 44 kDa, respectively (shown with *dots* in Fig. 4A). These results explicitly indicated also that the PG + CA form existed as a mixture of a protein monomer and dimer. SDS-PAGE analysis of the separated peaks under reducing and non-reducing conditions demonstrated that P1 contained disulfide-bonded oligomers (Fig. 4B), whereas P2 always showed monomeric species. Similarly to the SEC experiments, the molecular masses of the monomeric and dimeric species were exaggerated also in the SDS-PAGE analysis; this phenomenon was particularly evident for the dimer (see also Fig. 3). In summary, the experiments showed that both CA IX forms occurred as a mixture of monomeric and dimeric species.

*In the Presence of Excess Metal Ions CA IX Possesses the Highest Catalytic Activity Measured for a Member of the CA Enzyme Family*—To date, the catalytic activity of CA IX has been measured only for the bacterially expressed recombinant protein consisting of the CA domain (31). In the present study, the CO<sub>2</sub> hydration activities of the recombinant CA IX forms from the baculovirus-insect cell expression system were measured and compared with the previously reported values. Table 1 shows the measured kinetic constants of this study, together with those already published for other mammalian CAs. An excellent concordance was observed between the activity and inhibitory values measured for the CA form when expressed in either the insect cells or *E. coli*; furthermore, the results were consistent



**FIGURE 4. SEC, LS, and SDS-PAGE analysis of the PG + CA form.** A shows the results of the SEC and LS experiments. The x axis indicates the elution position of CA IX and four standard proteins. P1 and P2 refer to the first and second peak of CA IX, respectively. One y axis (*left*) shows  $A_{280}$  measured during the SEC (*curves*); the other (*right*) shows the molecular mass values determined by light scattering (*dots*). In B, collected P1 and P2 fractions were run on SDS-PAGE with and without  $\beta$ -mercaptoethanol. P1 migrated as monomeric species only under reducing conditions, whereas P2 migrated as monomeric species both in the non-reducing and reducing conditions. Abbreviations: APO, apoferritin; BAM,  $\beta$ -amylase from sweet potato; BSA, bovine serum albumin; OVA, ovalbumin.

ent with those reported earlier (31). This suggests that the C-terminal His tag, which could not be removed from the CA form, does not have any effect on the function of the protein *in vitro* because the protein from *E. coli*, from which the glutathione S-transferase tag was removed, showed identical results. In addition

TABLE 1

Kinetic and inhibitory properties of CA IX compared to the other mammalian CA isozymes

The values for isozymes other than CA IX are taken from Ref. 26.

Isoenzyme	CA IX details	$k_{cat}$ $s^{-1}$	$K_m$ $mM$	$k_{cat}/K_m$ $M^{-1} s^{-1}$	$K_I$ (acetazolamide) $nM$	Subcellular localization
hCA I		$2.0 \times 10^5$	4.0	$5.0 \times 10^7$	250	Cytosol
hCA II		$1.4 \times 10^6$	9.3	$1.5 \times 10^8$	12	Cytosol
hCA III		$1.3 \times 10^4$	52.0	$2.5 \times 10^5$	240000	Cytosol
hCA IV		$1.1 \times 10^6$	21.5	$5.1 \times 10^7$	74	GPI-anchored
hCA VA		$2.9 \times 10^5$	10.0	$2.9 \times 10^7$	63	Mitochondria
hCA VB		$9.5 \times 10^5$	9.7	$9.8 \times 10^7$	54	Mitochondria
hCA VI		$3.4 \times 10^5$	6.9	$4.9 \times 10^7$	11	Secreted
hCA VII		$9.5 \times 10^5$	11.4	$8.3 \times 10^7$	2.5	Cytosol
hCA IX	CA, <i>E. coli</i>	$3.8 \times 10^5$	6.9	$5.5 \times 10^7$	25	Transmembrane
hCA IX	CA, Sf9	$3.8 \times 10^5$	7.0	$5.4 \times 10^7$	24	
hCA IX	CA, Sf9, ZnCl <sub>2</sub>	$4.2 \times 10^6$	7.5	$5.6 \times 10^8$	9	
hCA IX	PG + CA, Sf9	$1.1 \times 10^6$	7.5	$1.5 \times 10^8$	16	
hCA IX	PG + CA, Sf9, ZnCl <sub>2</sub>	$2.5 \times 10^7$	7.3	$3.4 \times 10^9$	2325	
hCA IX	PG + CA, Sf9, MgCl <sub>2</sub>	$1.1 \times 10^7$	7.3	$1.5 \times 10^9$	55	
hCA IX	PG + CA, Sf9, CoCl <sub>2</sub>	$1.2 \times 10^7$	7.3	$1.5 \times 10^9$	56	
hCA IX	PG + CA, Sf9, MnSO <sub>4</sub>	$2.5 \times 10^7$	7.3	$3.4 \times 10^9$	41	
hCA XII		$4.2 \times 10^5$	12.0	$3.5 \times 10^7$	5.7	Transmembrane
hCA XIII		$1.5 \times 10^5$	13.8	$1.1 \times 10^7$	16	Cytosol
hCA XIV		$3.1 \times 10^5$	7.9	$3.9 \times 10^7$	41	Transmembrane
mCA XV		$4.7 \times 10^5$	14.2	$3.3 \times 10^7$	72	GPI-anchored

tion, we have recently produced another recombinant CA domain of CA IX, which is otherwise identical to the CA form, but has an N-terminal His tag that can be cleaved off; after the His tag removal, this protein showed identical  $k_{cat}$  and  $K_m$  values with the CA form, thus confirming that the C-terminal His tag does not interfere with the catalytic activity.<sup>3</sup> Because it was not known whether the insect cell protein preparations contained enough Zn<sup>2+</sup> to saturate the active sites of all enzyme molecules, measurements were also performed in the presence of 50  $\mu M$  ZnCl<sub>2</sub>. Under this experimental condition, the catalytic efficiency of the CA form became 10 times higher ( $k_{cat}/K_m = 5.6 \times 10^8 M^{-1} s^{-1}$ ) and exceeded that reported for CA II (13, 14). Without addition of ZnCl<sub>2</sub>, the CO<sub>2</sub> hydration activity of the PG + CA form was higher than that measured for the CA form under the same experimental conditions. In this case, addition of 50  $\mu M$  ZnCl<sub>2</sub> increased the activity more than 20-fold ( $k_{cat}/K_m$  was shown to be  $3.4 \times 10^9 M^{-1} s^{-1}$ ), reaching a value that has never been measured for any member of this enzyme family. Surprisingly, this addition also caused a 150-fold increase in the  $K_I$  value of acetazolamide for the PG + CA form (from 16 nM to 2.3  $\mu M$ ).

To investigate further the possible lack of Zn<sup>2+</sup> in the insect cell expression system, CO<sub>2</sub> hydration activity was measured in the presence of ZnCl<sub>2</sub> for another isozyme, namely CA XV, which was produced exactly under the same experimental conditions used for CA IX (26). The experiments demonstrated that the addition of ZnCl<sub>2</sub> to CA XV inhibited the enzyme catalytic activity (Table 2). In control experiments with isozymes I, II, IV, XII, and XIV, the addition of ZnCl<sub>2</sub> to the assay also caused a decrease in the CO<sub>2</sub> hydration activity (Table 2). In summary, these results indicated that the insect cell expression system provided enough Zn<sup>2+</sup> to occupy the active site of the CA molecules, and that the addition of ZnCl<sub>2</sub> did not cause an artificial increase in the CO<sub>2</sub> hydration measured by the CA activity assay. It was concluded that the increase of the catalytic

<sup>3</sup> M. Hilvo, L. Baranauskienė, A. M. Salzano, A. Scaloni, D. Matulis, A. Innocenti, A. Scozzafava, S. M. Monti, A. Di Fiore, G. De Simone, M. M. Lindfors, J. Jänis, J. Valjakka, S. Pastoreková, J. Pastorek, M. S. Kulomaa, H. R. Nordlund, C. T. Supuran, and S. Parkkila, unpublished observations.

TABLE 2

Effect of 50  $\mu M$  ZnCl<sub>2</sub> on the CO<sub>2</sub> hydration activity of various CA isozymes

Activity in the absence of the metal salt was taken as 100%.

Isoenzyme	CA activity <sup>a</sup> %
hCA I	73 ± 6
hCA II	61 ± 4
hCA IV	63 ± 5
hCA IX <sup>b</sup>	2266 ± 131
hCA XII	58 ± 4
hCA XIV	57 ± 3
mCA XV	60 ± 5

<sup>a</sup> Mean ± S.E. (from three different measurements).

<sup>b</sup> The PG + CA form produced in Sf9 cells.

activity resulting from the addition of ZnCl<sub>2</sub> to the enzyme is a unique feature for isozyme IX.

Because both the catalytic activity and inhibition constant were increased following addition of ZnCl<sub>2</sub>, we further investigated the effect of other metal ions on the PG + CA form (Table 1). Addition of MgCl<sub>2</sub>, CoCl<sub>2</sub>, and MnSO<sub>4</sub> also induced an increase in the  $k_{cat}/K_m$  values, with the ratio measured for MnSO<sub>4</sub> being identical to that determined for ZnCl<sub>2</sub>. In contrast to ZnCl<sub>2</sub>, the other metal ions did not exhibit an increase higher than 3-fold for the  $K_I$  of acetazolamide. On the other hand, MgCl<sub>2</sub>, CoCl<sub>2</sub>, and MnSO<sub>4</sub> did not have any effect on the CA form (data not shown).

**CA IX Contains One N-Linked Glycosylation**—To investigate the post-translational modifications of the recombinant CA IX proteins, both monomeric species from SEC were submitted to electrospray ionization-ion trap-mass spectrometry. Both showed the presence of four distinct components that differ by 162 mass units, probably due to different sugar moieties. For the PG + CA form, the determined molecular mass values were  $42486.9 \pm 13.3$ ,  $42656.3 \pm 11.8$ ,  $42812.6 \pm 6.7$ , and  $42969.1 \pm 13.1$  Da. For the CA form, the corresponding values were  $30977.6 \pm 10.1$ ,  $31136.8 \pm 9.8$ ,  $31300.9 \pm 9.7$ , and  $31463.1 \pm 12.0$  Da, respectively.

Glycosylation and the assignment of modified residue(s) were elucidated with protein samples resolved by reducing

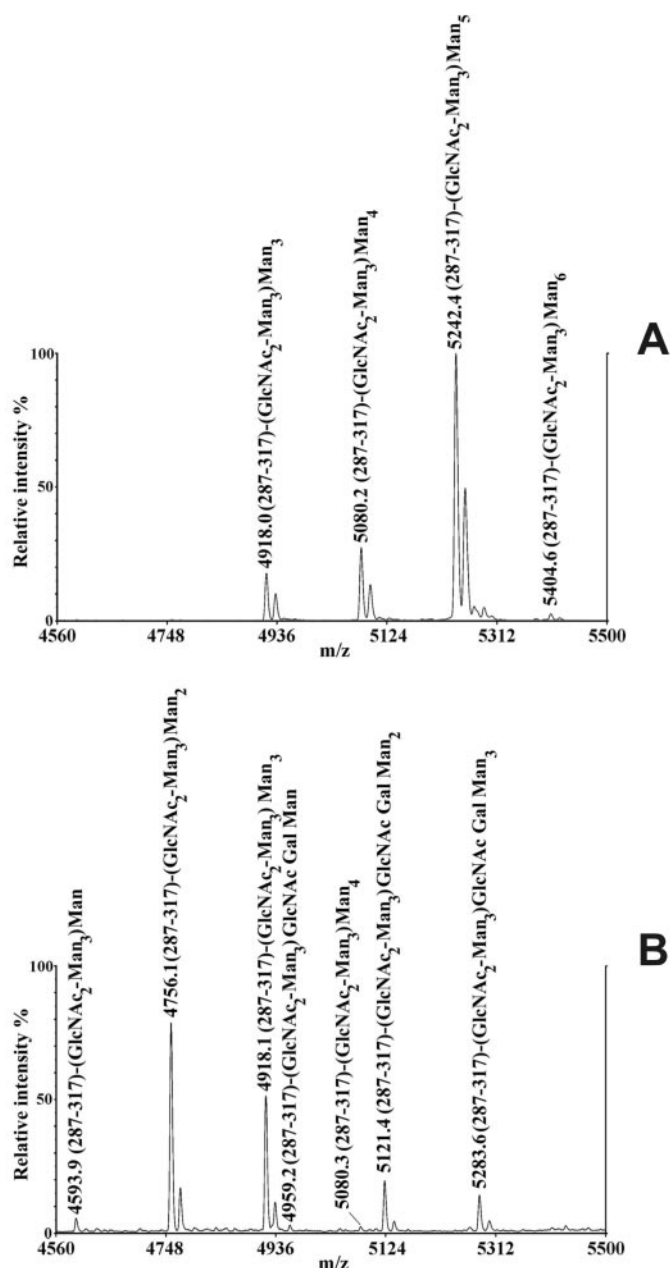


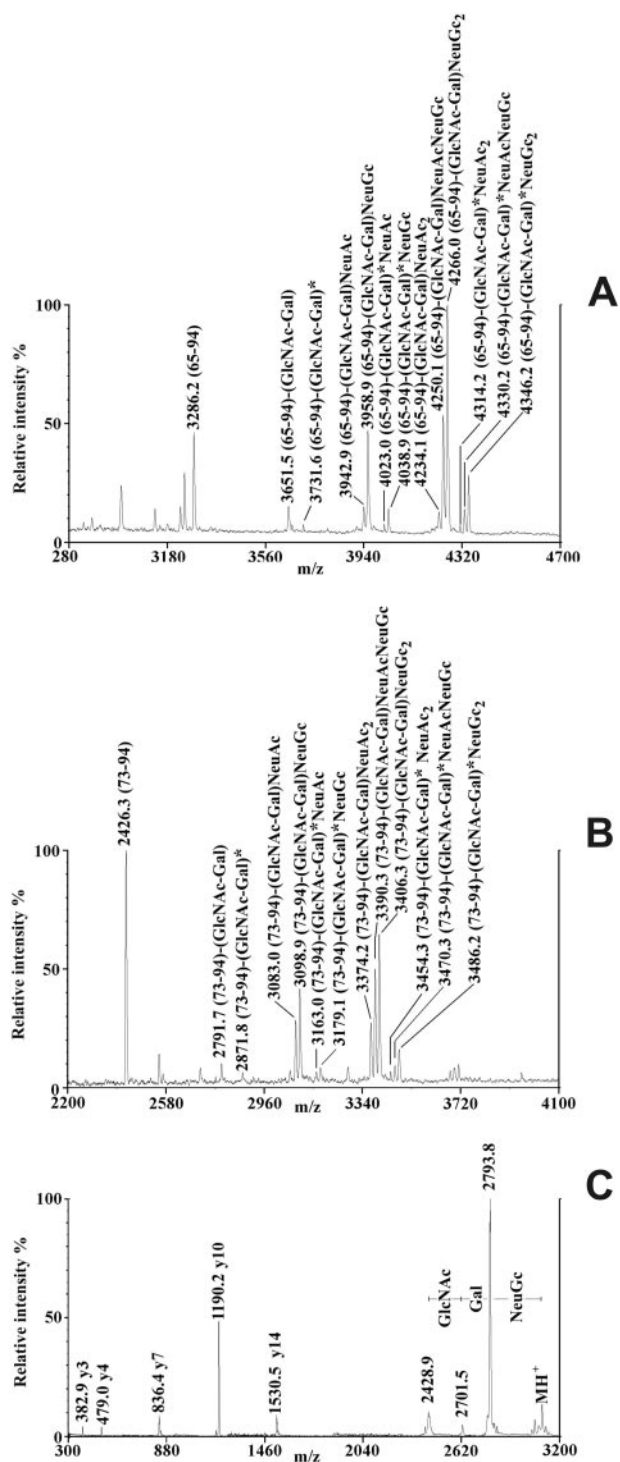
FIGURE 5. MALDI-TOF MS analysis of glycopeptides bearing high mannose and hybrid-type glycans *N*-linked to Asn<sup>309</sup> in recombinant human CA IX. A and B show the results for the protein from the baculovirus-insect cell expression system and from murine cells, respectively.

SDS-PAGE for the PG + CA and CA forms, which migrated at 45 and 29 kDa, respectively. Bands were alkylated with iodoacetamide and digested with trypsin. Peptide digests were either directly analyzed by MALDI-TOF MS in positive polarity or resolved by  $\mu$ LC to isolate glycopeptides, which was then followed by MALDI-TOF MS analysis of the collected fractions. A fraction eluting at 37.8 min, showing multiple peaks in the mass spectrum, was obtained for both recombinant enzyme forms. Fig. 5A illustrates the mass spectrum measured for the PG + CA product, and an almost identical result was obtained for the CA form (data not shown). On the basis of measured masses and known pathways of glycoprotein biosynthesis, these signals were assigned to peptide-(287–317) having a pentasaccharide

core *N*-linked to the specific *N*-glycosylation consensus Asn-Gln-Thr and bearing 3–6 mannose units (theoretical MH<sup>+</sup> values: *m/z* 4918.3, 5080.4, 5242.6, and 5404.7). In both cases, glycopeptides collapsed to a unique component after PNGase treatment, which showed a single MH<sup>+</sup> signal at *m/z* 3540.1). Peptide species before and after PNGase treatment were submitted for PSD analysis, confirming the identity of the glycopeptides and the expected Asn<sup>309</sup> → Asp conversion in the deglycosylated peptide (data not shown). Moreover, no signals associated with the non-glycosylated peptide-(287–317) were detected in any LC fractions from both digests, thus suggesting that both of these proteins were completely modified. Moreover, measured MH<sup>+</sup> signals allowed the analysis of 89.2 and 93.1% of the sequences for the PG + CA and CA forms, respectively, and excluded the occurrence of other post-translational modifications. In conclusion, the analyzed proteins corresponded to the expected thrombin-processed PG + CA form and the unprocessed CA form (from which the signal peptide had been cleaved off by the insect cells), with both containing four high mannose-type glycans *N*-linked to Asn<sup>309</sup> (theoretical mass values 42490.6, 42652.7, 42814.9 and 42977.0 Da (PG + CA form); theoretical mass values 30975.6, 31137.8, 31299.9, and 31462.1 Da (CA form)).

*Human CA IX Produced in Mammalian Cells Shows N- and O-Linked Glycosylations*—To investigate possible differences in glycosylation of human CA IX produced in either insect cells or mammalian cells, we also analyzed a sample of recombinant human enzyme produced in the murine NS0 myeloma cell line. The protein sample was resolved by SDS-PAGE, revealing a unique band that migrated at 47 kDa and was further subjected to MALDI-TOF mass mapping experiments in positive polarity, performed as above. Also in this case, a peptide fraction eluting at almost 37.6 min was observed (Fig. 5B). On the basis of measured masses and known pathways of glycoprotein biosynthesis, these signals were assigned to peptide-(287–317) bearing high mannose and hybrid-type structures *N*-linked to Asn<sup>309</sup> (theoretical MH<sup>+</sup> values: *m/z* 4594.0, 4756.2, 4918.3, 4959.3, 5080.4, 5121.5, and 5283.6). Also in this case, glycopeptides collapsed to a unique component with a MH<sup>+</sup> signal at *m/z* 3540.0 after PNGase treatment. Similar to that reported above, PSD analysis confirmed the glycopeptide nature of signals differing by *m/z* 162 or 203, as well as Asn<sup>309</sup> → Asp conversion in the deglycosylated peptide (data not shown).

By performing MALDI-TOF MS in negative polarity, additional glycopeptides were also detected in the fractions eluting at 29.8 and 33.6 min, and their occurrence was ascertained by parallel MH<sup>-</sup> signals differing by *m/z* 291, 307, and 365 (Fig. 6, A and B). On the basis of measured mass values, these signals were tentatively assigned to peptides-(73–94) and -(65–94) bearing different di-, tri-, and tetrasaccharide chains *O*-linked to a common Thr<sup>78</sup> (theoretical MH<sup>-</sup> values for peptide-(73–94), *m/z* 2791.9, 3083.1, 3099.1, 3374.4, 3390.4, and 3406.4; theoretical MH<sup>-</sup> values for peptide-(65–94), *m/z* 3651.8, 3943.0, 3959.0, 4234.3, 4250.3, and 4266.3). Both *N*-acetylneuraminic acid- (NeuAc) and *N*-glycolylneuraminic acid- (NeuGc)-containing glycans were observed. It was not clear if the complex pattern of peaks measured was generated from a real glycopeptide heterogeneity or the well known partial loss of carbohy-



**FIGURE 6. MALDI-TOF MS analysis of glycopeptides bearing di-, tri-, and tetrasaccharides O-linked to Thr<sup>78</sup> in recombinant human CA IX from murine cells.** Mass spectra in negative mode for glycopeptide-(73–94) and -(65–94) derivatives are reported in A and B, respectively. Asterisks refer to satellite signals with  $\Delta m/z = +80$ , which were assigned to sulfated glycopeptide derivatives. C shows PSD analysis in positive mode of the glycopeptide with  $MH^+$  at  $m/z$  3101.0. On the basis of the occurrence of a sulfated moiety similar to the keratan sulfate unit, the occurrence of a GlcNAc-Gal disaccharide was tentatively proposed for CA IX.

drate groups during MALDI-TOF MS analysis, particularly the loss of NeuAc and NeuGc (32). Satellite signals at  $\Delta m = +80$  Da were also detected in the spectra (Fig. 6, A and B), and these

species were hypothetically associated to additional O-linked glycans bearing a sulfate moiety because they did not disappear following treatment with alkaline phosphatase (theoretical  $MH^-$  values for peptide-(73–94):  $m/z$  2871.9, 3163.2, 3179.2, 3454.4, 3470.4, and 3486.4; theoretical  $MH^-$  values for peptide 65–94:  $m/z$  3731.8, 4023.1, 4039.1, 4314.3, 4330.3, and 4346.3). PSD analysis in positive mode for the molecular species with  $MH^+$  signals at  $m/z$  3101.0 and 3408.5 confirmed the glycopeptide nature by means of detecting fragment ions resulting either from glycan or peptide chain fragmentation, thus confirming the assigning site at Thr<sup>78</sup> (Fig. 6C). Due to the presence of a sulfated moiety similar to the keratan sulfate unit, the occurrence of a GlcNAc-Gal disaccharide was tentatively proposed. In conclusion, a combination of MALDI-TOF mass spectra acquired in positive and negative polarities for recombinant CA IX from murine cells allowed covering 93.8% of the protein sequence.

**CA IX Contains a Single Intramolecular Disulfide Bond**—To investigate the correct pairing of cysteine residues in recombinant CA IX obtained from the baculovirus-insect cell expression system, a sample of the monomeric CA form was alkylated with iodoacetamide under denaturing non-reducing conditions, digested with trypsin, and subjected to MALDI-TOF MS analysis in positive polarity, as reported above. In this case, a peptide fraction eluting at almost 52.7 min showed a panel of signals that differed by 162 mass units, with  $MH^+$  at  $m/z$  5618.9, 5781.6, 5943.2, and 6105.3. On the basis of the glycosylation data reported above, these species were assigned to disulfide-containing peptides-(115–122)-S-S-(287–317) bearing four high mannose-type glycans N-linked to Asn<sup>309</sup> (theoretical  $MH^+$  values:  $m/z$  5619.1, 5781.3, 5943.4, and 6105.5). After PNGase treatment, these glycopeptides collapsed to a unique component with a  $MH^+$  signal at  $m/z$  4240.8. As expected, incubation with DTT caused both the disappearance of the disulfide-containing component and the generation of species with  $MH^+$  signals at  $m/z$  760.9 and 3483.1, which were assigned to reduced peptides-(115–122) and -(287–317), respectively. In the peptide fraction eluting at 10.2 min, MALDI-TOF MS analysis demonstrated the occurrence of the cysteine-containing peptide-(131–141) as a carboxamidomethylated species with a  $MH^+$  signal at  $m/z$  1243.9.

Similarly to that reported above, the tryptic digest of the PG + CA monomeric form showed the occurrence of peptides with  $MH^+$  at  $m/z$  5618.8, 5781.8, 5943.3, and 6105.5, which were assigned to disulfide-containing peptides-(115–122)-S-S-(287–317) bearing four N-linked high mannose-type glycans. These glycopeptides again collapsed to a unique component following PNGase F treatment. Together, these results demonstrated that both recombinant CA IX forms exhibited a S-S bridge linking Cys<sup>119</sup> to Cys<sup>299</sup>.

## DISCUSSION

A series of novel biochemical features, which differentiate human CA IX from the other members of the CA family, were evident following characterization of two recombinant protein forms that were produced by the baculovirus-insect cell expression system. The insect cell expression system has never been reported for the production of recombinant CA proteins, and it

allowed us to obtain enzymes with a higher similarity to the mammalian native counterpart than that produced in bacteria.

The oligomerization of the recombinant enzymes was investigated by several experimental approaches, *i.e.* size exclusion chromatography, size exclusion chromatography combined with laser light scattering, and SDS-PAGE under reducing and non-reducing conditions. These experiments demonstrated that both recombinant CA IX forms consisted of a mixture of monomeric and disulfide-linked dimeric species. The results were clearly interpretable for the CA form; in contrast, the use of a specific technique, namely laser light scattering (30), considered as preferential for the absolute determination of the size of a protein and its oligomers under native conditions, was essential to clarify the multimeric state of the PG + CA form, consisting of the whole CA IX extracellular portion. It allowed us to correctly interpret the results from SEC and SDS-PAGE experiments, both showing an abnormal migration of the recombinant product, because of the occurrence of the PG domain, which probably possesses non-globular structure. Our experiments provided an alternative interpretation to previous SDS-PAGE-based estimations of the oligomeric state of CA IX, reporting the protein as able to form a disulfide-linked trimer (23, 33).

Because our MS experiments demonstrated that both recombinant CA IX forms have an intramolecular S-S bridge linking Cys<sup>119</sup> to Cys<sup>299</sup>, which is also conserved in other CA isozymes, we can deduce that the CA form may have only one cysteine residue (Cys<sup>137</sup>) able to form the disulfide-stabilized dimer. On the other hand, the availability of two cysteine residues (Cys<sup>137</sup> and Cys<sup>372</sup>) in the PG + CA form, capable of forming the intermolecular disulfide bond(s), leaves open the question on the number and nature of the intermolecular disulfides in the dimeric form, but not in an eventual trimeric species. Although our results do not support the previous concept of trimerization of CA IX, a possibility remains that CA IX may need the transmembrane helix and/or the intracytoplasmic tail to form the oligomeric trimer. Further studies with the full-length protein are needed to elucidate whether the C-terminal part of this protein may affect the CA IX oligomerization status.

Our previous conceptions regarding CA IX were dramatically changed by the results addressing the CO<sub>2</sub> hydration activity of the recombinant protein forms. The results revealed a catalytic activity that was considerably higher than that previously reported for this enzyme. Initially, an excellent correlation was observed between the data obtained for the catalytic domain produced in both the baculovirus-insect cell and the bacterial expression systems; however, addition of ZnCl<sub>2</sub> to the insect cell-derived enzyme increased its catalytic efficiency by approximately 1 order of magnitude. The PG + CA form showed even more surprising results; without metal additives, it showed the same  $k_{\text{cat}}/K_m$  value as that measured for CA II, which has been considered to be the most active CA isozyme thus far studied. The only major difference between the PG + CA and CA forms was the presence of the PG domain in the PG + CA form; our results suggest that this domain contributes to an increase in the CO<sub>2</sub> hydration activity of the CA domain by an unknown mechanism. Addition of ZnCl<sub>2</sub> increased the catalytic activity of the PG + CA form by 23-fold, reaching a

level ( $k_{\text{cat}}/K_m = 3.4 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ ) never measured for any member of the CA isozyme family. Additional ZnCl<sub>2</sub> also increased the  $K_i$  value for acetazolamide considerably. Comparative experiments with other isozymes revealed that the increased catalytic activity in the presence of ZnCl<sub>2</sub> was not due to lack of Zn<sup>2+</sup> in the insect cell expression system nor reflect an artificial acceleration of the CO<sub>2</sub> hydration activity in the assay. Further experiments performed in the presence of other metals demonstrated a common effect of divalent cations on the catalytic activity of the PG + CA form. In contrast to the addition of ZnCl<sub>2</sub>, MgCl<sub>2</sub>, CoCl<sub>2</sub>, and MnSO<sub>4</sub> showed no effect on the  $k_{\text{cat}}/K_m$  and  $K_i$  values of the CA form, and the presence of these cations caused only a slight increase in the inhibition constant of acetazolamide for the PG + CA form.

Together, the data suggest that the metal ions may bind some negatively charged molecular moieties either on the proteoglycan or catalytic domain. Zn<sup>2+</sup> ion seems to be mainly directed toward the sites within the catalytic domain because addition of ZnCl<sub>2</sub> to the protein solution increased the catalytic activity of both recombinant protein forms. On the other hand, the other metal ions may bind the PG domain, which contains many negatively charged amino acid repeats, because their effects were only observed with the PG + CA form. The metal ions enhance catalytic activity by possibly relieving the electrostatic repulsions that occur especially in the PG domain, thereby stabilizing the whole enzyme. Only in the case of the PG + CA form, the subtle structural rearrangement induced by the Zn<sup>2+</sup> ion may also contribute to a PG domain-dependent selective masking or distortion of the binding pocket for acetazolamide, which results in the observed inhibition constant increase. It is worthwhile to note that both the catalytic and PG domains of native human CA IX reside in the interstitial fluid, which contains a considerable amount of free ions, *e.g.* Na<sup>+</sup> in the order of 100 mM (34). The PG domain of the PG + CA form of CA IX has been observed to be stable mainly in a phosphate-based buffer, which contained 100 mM NaCl.<sup>3</sup> However, experiments *in vivo* have indicated that native CA IX protein on the cell surface is highly stable (35). The results in the present study lead to the speculation that the free ions in the interstitial fluid may result in the very high stability and catalytic activity of CA IX in human tissues.

Mass spectrometric analysis revealed information on the post-translational modifications present in human CA IX. A common *N*-linked glycosylation site (Asn<sup>309</sup>) within the CA domain, which was modified by high mannose-type glycans, was ascertained in both recombinant CA IX forms from the baculovirus-insect cell expression system as well as in that from the murine cells. In the latter case, *N*-linked hybrid-type structures were also detected. Our finding of the high mannose-type glycans is in good agreement with previous preliminary data on glycosylation of the native protein (23). In addition, the presence of *O*-linked glycosylation at Thr<sup>78</sup> was also demonstrated in the recombinant product from mammalian cells. This residue is present within a region that has significant sequence similarity with the keratan sulfate attachment domain of a large aggregating proteoglycan, aggrecan (19). Different di-, tri-, and tetrasaccharide NeuAc- or NeuGc-containing *O*-linked glycans with, and without, a

sulfate moiety were detected. Some of these oligosaccharides highly resemble the keratan sulfate unit that has already been described in the proteoglycan domain of other proteins involved in cell adhesion processes and tumor progression (36). In CD44, a protein implicated in cell motility, tumor metastasis, and lymphocyte activation, keratan sulfate modifications modulate adhesion to the extracellular matrix component, hyaluronan (37). Increased sialylation, defective *N*-linked glycosylation, and substitution of the CD44 glycosaminoglycans with keratan sulfate structures all seem to occur in neuroblastoma, myeloid, and melanoma cells (38–41). Similarly, changes in the glycosylation of membrane mucins have been studied in many carcinomas by analyzing both the composition and density of the *O*-glycans added and the profile of glycosyltransferases and sulfotransferases expressed (42, 43). In breast cancer tumors and cell lines, a shift has been observed in protein modifications toward the addition of shorter *O*-glycans (43–45). These findings correlated well with the measurements of higher expression levels of the glycosyltransferases that are responsible for the sialylation of core-1 disaccharides and the termination of *O*-glycan chain extension (46). Although a mechanism is not well elucidated, sulfotransferases may also control the addition of *O*-glycans to form specific molecular signatures (47, 48), and several of these enzymes have different expression and activity in cancer and inflammation (49). On this basis, our results suggest that a more detailed characterization of the post-translational processes modifying final protein structure is necessary to ascertain the eventual occurrence of complex keratan sulfate polymeric structures in CA IX and to investigate the roles of *N*- and *O*-linked glycans and their eventual changes during malignant transformation.

Even though the exact biological roles of CA IX still remain an open question, the present results describe novel biochemical properties of CA IX, which provide important further information on its function. Only a detailed structural characterization of the full-length enzyme in a membrane-like environment by high resolution mass spectrometric, x-ray diffraction, or cryo-electron microscopic techniques as well as biological assays on cellular adhesion and proliferation using specific recombinant CA IX forms, will allow the structure-function relationship of this protein to be advanced. Future studies will hopefully deepen our knowledge on the function of different domains present within CA IX, contributing to the understanding of the various pathophysiological processes in which this unique enzyme is involved.

**Acknowledgments**—We thank Dr. Ewa Folta-Stogniew (Biophysics Facility, W. M. Keck Foundation Biotechnology Resource Laboratory, Yale University, New Haven, CT) for performing the SEC/LS analysis and Aulikki Lehmus and Jukka Lehtonen for skillful technical assistance.

### REFERENCES

- Supuran, C. T. (2004) in *Carbonic Anhydrase: Its Inhibitors and Activators* (Supuran, C. T., Scozzafava, A., and Conway, J., eds) pp. 1–23, CRC Press, Boca Raton, FL
- Sly, W. S., and Hu, P. Y. (1995) *Annu. Rev. Biochem.* **64**, 375–401
- Hilvo, M., Tolvanen, M., Clark, A., Shen, B., Shah, G. N., Waheed, A., Halmi, P., Hänninen, M., Hämäläinen, J. M., Vihinen, M., Sly, W. S., and Parkkila, S. (2005) *Biochem. J.* **392**, 83–92
- Montgomery, J. C., Venta, P. J., Eddy, R. L., Fukushima, Y. S., Shows, T. B., and Tashian, R. E. (1991) *Genomics* **11**, 835–848
- Lehtonen, J., Shen, B., Vihinen, M., Casini, A., Scozzafava, A., Supuran, C. T., Parkkila, A. K., Saarnio, J., Kivelä, A. J., Waheed, A., Sly, W. S., and Parkkila, S. (2004) *J. Biol. Chem.* **279**, 2719–2727
- Zhu, X. L., and Sly, W. S. (1990) *J. Biol. Chem.* **265**, 8795–8801
- Pastoreková, S., Parkkila, S., Parkkila, A. K., Opavský, R., Zelník, V., Saarnio, J., and Pastorek, J. (1997) *Gastroenterology* **112**, 398–408
- Türeci, Ö., Sahin, U., Vollmar, E., Siemer, S., Göttert, E., Seitz, G., Parkkila, A. K., Shah, G. N., Grubb, J. H., Pfreundschuh, M., and Sly, W. S. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 7608–7613
- Fujikawa-Adachi, K., Nishimori, I., Taguchi, T., and Onishi, S. (1999) *Genomics* **61**, 74–81
- Whittington, D. A., Grubb, J. H., Waheed, A., Shah, G. N., Sly, W. S., and Christianson, D. W. (2004) *J. Biol. Chem.* **279**, 7223–7228
- Fujikawa-Adachi, K., Nishimori, I., Taguchi, T., and Onishi, S. (1999) *J. Biol. Chem.* **274**, 21228–21233
- Murakami, H., and Sly, W. S. (1987) *J. Biol. Chem.* **262**, 1382–1388
- Khalifah, R. G. (1971) *J. Biol. Chem.* **246**, 2561–2573
- Nishimori, I., Innocenti, A., Vullo, D., Scozzafava, A., and Supuran, C. T. (2007) *Bioorg. Med. Chem. Lett.* **17**, 1037–1042
- Pastoreková, S., and Pastorek, J. (2004) in *Carbonic Anhydrase: Its Inhibitors and Activators* (Supuran, C. T., Scozzafava, A., and Conway, J., eds) pp. 255–281, CRC Press, Boca Raton, FL
- Ortova Gut, M. O., Parkkila, S., Vernerová, Z., Rohde, E., Závada, J., Höcker, M., Pastorek, J., Karttunen, T., Gibadulinová, A., Zavadová, Z., Knobeloch, K. P., Wiedenmann, B., Svoboda, J., Horak, I., and Pastoreková, S. (2002) *Gastroenterology* **123**, 1889–1903
- Pastoreková, S., and Závada, J. (2004) *Cancer Therapy* **2**, 245–262
- Pastorek, J., Pastoreková, S., Callebaut, I., Mornon, J. P., Zelník, V., Opavský, R., Zát'ovicová, M., Liao, S., Portetelle, D., Stanbridge, E. J., Závada, J., Burny, A., and Kettmann, R. (1994) *Oncogene* **9**, 2877–2888
- Opavský, R., Pastoreková, S., Zelník, V., Gibadulinová, A., Stanbridge, E. J., Závada, J., Kettmann, R., and Pastorek, J. (1996) *Genomics* **33**, 480–487
- Majumder, K. (1992) *Gene (Amst.)* **116**, 115–116
- Laemmli, U. K. (1970) *Nature* **227**, 680–685
- Hilvo, M., Rafajová, M., Pastoreková, S., Pastorek, J., and Parkkila, S. (2004) *J. Histochem. Cytochem.* **52**, 1313–1322
- Pastoreková, S., Závadová, Z., Kostál, M., Babusiková, O., and Závada, J. (1992) *Virology* **187**, 620–626
- Zát'ovicová, M., Tarábková, K., Svastová, E., Gibadulinová, A., Mucha, V., Jakubíčková, L., Biesová, Z., Rafajová, M., Ortova Gut, M., Parkkila, S., Parkkila, A. K., Waheed, A., Sly, W. S., Horak, I., Pastorek, J., and Pastoreková, S. (2003) *J. Immunol. Methods* **282**, 117–134
- Özensoy, Ö., Puccetti, L., Fasolis, G., Arslan, A., Scozzafava, A., and Supuran, C. T. (2005) *Bioorg. Med. Chem. Lett.* **15**, 4862–4866
- Hilvo, M., Innocenti, A., Monti, S. M., De Simone, G., Supuran, C. T., and Parkkila, S. (2008) *Curr. Pharm. Des.* **14**, 672–678
- Grillo, C., D'Ambrosio, C., Consalvi, V., Chiaraluce, R., Scaloni, A., Macerioni, M., Eufemi, M., and Altieri, F. (2007) *J. Biol. Chem.* **282**, 10299–10310
- Salzano, A. M., Arena, S., Renzone, G., D'Ambrosio, C., Rullo, R., Bruschi, M., Ledda, L., Maglione, G., Candiano, G., Ferrara, L., and Scaloni, A. (2007) *Proteomics* **7**, 1420–1433
- Bernardini, G., Renzone, G., Comanducci, M., Mini, R., Arena, S., D'Ambrosio, C., Bambini, S., Trabalzini, L., Grandi, G., Martelli, P., Achtman, M., Scaloni, A., Ratti, G., and Santucci, A. (2004) *Proteomics* **4**, 2893–2926
- Folta-Stogniew, E., and Williams, K. R. (1999) *J. Biomol. Tech.* **10**, 51–63
- Wingo, T., Tu, C., Laipis, P. J., and Silverman, D. N. (2001) *Biochem. Biophys. Res. Commun.* **288**, 666–669
- Holland, J. W., Deeth, H. C., and Alewood, P. F. (2004) *Proteomics* **4**, 743–752
- Závada, J., Zavadová, Z., Pastoreková, S., Ciampor, F., Pastorek, J., and Zelník, V. (1993) *Int. J. Cancer* **54**, 268–274
- Fogh-Andersen, N., Altura, B. M., Altura, B. T., and Siggaard-Andersen,

- O. (1995) *Clin. Chem.* **41**, 1522–1525
35. Rafajová, M., Zatovicová, M., Kettmann, R., Pastorek, J., and Pastoreková, S. (2004) *Int. J. Oncol.* **24**, 995–1004
36. Funderburgh, J. L. (2000) *Glycobiology* **10**, 951–958
37. Takahashi, K., Stamenkovic, I., Cutler, M., Dasgupta, A., and Tanabe, K. K. (1996) *J. Biol. Chem.* **271**, 9490–9496
38. Maiti, A., Maki, G., and Johnson, P. (1998) *Science* **282**, 941–943
39. Delcommenne, M., Kannagi, R., and Johnson, P. (2002) *Glycobiology* **12**, 613–622
40. Gasbarri, A., Del Prete, F., Girnita, L., Martegani, M. P., Natali, P. G., and Bartolazzi, A. (2003) *Melanoma Res.* **13**, 325–337
41. Gross, N., Balmas, K., and Beretta Brognara, C. (2001) *Med. Pediatr. Oncol.* **36**, 139–141
42. Hakomori, S. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 10231–10233
43. Brockhausen, I. (2006) *EMBO Rep.* **7**, 599–604
44. Brockhausen, I., Yang, J. M., Burchell, J., Whitehouse, C., and Taylor-Papadimitriou, J. (1995) *Eur. J. Biochem.* **233**, 607–617
45. Lloyd, K. O., Burchell, J., Kudryashov, V., Yin, B. W., and Taylor-Papadimitriou, J. (1996) *J. Biol. Chem.* **271**, 33325–33334
46. Sewell, R., Bäckström, M., Dalziel, M., Gschmeissner, S., Karlsson, H., Noll, T., Gätgens, J., Clausen, H., Hansson, G. C., Burchell, J., and Taylor-Papadimitriou, J. (2006) *J. Biol. Chem.* **281**, 3586–3594
47. Honke, K., and Taniguchi, N. (2002) *Med. Res. Rev.* **22**, 637–654
48. Brockhausen, I. (2003) *Biochem. Soc. Trans.* **31**, 318–325
49. Chandrasekaran, E. V., Xue, J., Neelamegham, S., and Matta, K. L. (2006) *Carbohydr. Res.* **341**, 983–994

