Crystal structure of the catalytic domain of the tumor-associated human carbonic anhydrase IX

Vincenzo Alterio[a,1], Mika Hilvo[b,c,1], Anna Di Fiore[a], Claudiu T. Supuran[d,2], Peiwen Pan[b], Seppo Parkkila[b], Andrea Scaloni[e], Jaromir Pastorek[f], Silvia Pastorekova[f], Carlo Pedone[a], Andrea Scozzafava[d], Simona Maria Monti[a,2], and Giuseppina De Simone[a,2]

[a]Istituto di Biostrutture e Bioimmagini-CNR, via Mezzocannone 16, 80134 Naples, Italy; [b]Institute of Medical Technology and School of Medicine, University of Tampere and Tampere University Hospital, FI-33014 University of Tampere, Finland; [c]VTT Technical Research Centre of Finland, Tietotie 2, P.O. Box 1000, FI-02044 VTT, Espoo, Finland; [d]Laboratorio di Chimica Bioinorganica, Università degli Studi di Firenze, Polo Scientifico, Rm. 188, Via della Lastruccia 3, 50019 Sesto Fiorentino (Florence), Italy; [e]Proteomics and Mass Spectrometry Laboratory, ISPAAM-CNR, via Argine 1085, 80147 Naples, Italy; and [f]Institute of Virology, Slovak Academy of Sciences, Dubravska cesta 9, 845 05 Bratislava, Slovak Republic

Communicated by William S. Sly, Saint Louis University School of Medicine, St. Louis, MO, July 24, 2009 (received for review May 25, 2009)

Carbonic anhydrase (CA) IX is a plasma membrane-associated member of the α-CA enzyme family, which is involved in solid tumor acidification. It is a marker of tumor hypoxia and a prognostic factor in several human cancers. An aberrant increase in CA IX expression in chronic hypoxia and during development of various carcinomas contributes to tumorigenesis through at least two mechanisms: pH regulation and cell adhesion control. Here we report the X-ray structure of the catalytic domain of CA IX in complex with a classical, clinically used sulfonamide inhibitor, acetazolamide. The structure reveals a typical α-CA fold, which significantly differs from the other CA isozymes when the protein quaternary structure is considered. Thus, two catalytic domains of CA IX associate to form a dimer, which is stabilized by the formation of an intermolecular disulfide bond. The active site clefts and the PG domains are located on one face of the dimer, while the C-termini are located on the opposite face to facilitate protein anchoring to the cell membrane. A correlation between the three-dimensional structure and the physiological role of the enzyme is here suggested, based on the measurement of the pH profile of the catalytic activity for the physiological reaction, CO_2 hydration to bicarbonate and protons. On the basis of the structural differences observed between CA IX and the other membrane-associated α-CAs, further prospects for the rational drug design of isozyme-specific CA inhibitors are proposed, given that inhibition of this enzyme shows antitumor activity both in vitro and in vivo.

Carbonic anhydrases (CAs) are ubiquitous metallo-enzymes, acting as catalysts in the reversible hydration of CO_2 to HCO_3^- and H^+. All human CAs (hCA*s*) belong to the so-called α-class and differ widely in their cellular localization: CA I, II, III, VII, and XIII reside in the cytosol, CA IV, IX, XII, and XIV are associated with the cell membrane, CA VA, and VB occur in mitochondria, whereas CA VI is secreted into saliva and milk (1). Structural studies on isozymes I, II, III, IV, V, XII, XIII, and XIV (2–10) have revealed that all these isoforms present a considerable degree of three-dimensional similarity and a typical fold characterized by a central antiparallel β-sheet surrounding a Zn^{2+} critical for catalysis. The active site is located in a large, cone-shaped cavity that reaches the center of the protein molecule. The zinc ion is close to the bottom of the cavity and is coordinated by three conserved His residues in a tetrahedral geometry with H_2O or OH^- as the fourth ligand.

CA IX is a very peculiar member of the CA family since it is generally expressed in a limited number of normal tissues (mainly in the gastrointestinal tract), whereas its overexpression is seen on the cell surface of a large number of solid tumors, where it is invariably linked with the hypoxic phenotype, mediated by the transcription factor HIF-1 (11). Moreover, CA IX overexpression is often associated with a poor responsiveness to classical radio- and chemotherapy (12). Since CA IX has also been shown to contribute to pH regulation of tumor cells, cell proliferation, cell adhesion, and malignant cell invasion (13, 14), this protein has been recently recognized as a valuable target for cancer diagnostics and treatment (1).

A cDNA coding for the hCA IX protein was first cloned and investigated by Pastorek et al. (15), and the *CA9* gene was further characterized by the same group (16). On the basis of the sequence similarity, these authors proposed that hCA IX is a multidomain protein consisting of an N-terminal proteoglycan-like (PG) domain, a CA catalytic domain, a transmembrane segment (TM), and an intracytoplasmic (IC) portion (16). The first domain is absent in all other CA isozymes and resembles the keratan sulfate attachment domain of a large aggregating proteoglycan, namely aggrecan; in contrast, the CA catalytic domain presents a significant sequence identity (from 30% to 40%) to other catalytic hCA isozymes. Both CA and PG domains are glycosylated (17). According to recent biochemical reports, the CA IX protein is dimeric, and the dimerization is mediated by the formation of an intermolecular disulfide bond between the same Cys residue located on two CA catalytic domains (17). Despite all of the data available on CA IX, no information is currently available on the three-dimensional structure of its isolated domains or their relative orientation within the whole CA IX three-dimensional structure. This information is fundamental for the rational design of selective CA IX-directed anti-tumor drugs and diagnostic tools.

In this study, we report the X-ray structure of the catalytic domain of CA IX. Significant differences were identified with respect to the other α-CA isozymes, whose three-dimensional structure has been determined. The molecular bases for dimerization of CA IX in the membrane are provided together with a correlation between the three-dimensional structure and the physiological role of the enzyme. Because of the structural differences observed between CA IX and the other membrane-associated α-CA isoforms, further prospects for the rational drug design of isozyme-specific CA inhibitors are also proposed.

Results

Protein Cloning, Expression, and Purification. Four recombinant hCA IX enzyme forms were produced using the baculovirus-insect cell expression system (17) (Fig. S1). A recombinant product (P1) had both the PG and CA domains of hCA IX; the

Author contributions: M.H., C.T.S., S. Parkkila, A. Scaloni, J.P., S. Pastorekova, C.P., A. Scozzafava, S.M.M., and G.D.S. designed research; V.A., M.H., A.D.F., and P.P. performed research; C.T.S., S.M.M., and G.D.S. analyzed data; and C.T.S. and G.D.S. wrote the paper.

The authors declare no conflict of interest.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 3IAI).

[1]V.A. and M.H. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: claudiu.supuran@unifi.it, marmonti@unina.it, or gdesimon@unina.it.

This article contains supporting information online at www.pnas.org/cgi/content/full/0908301106/DCSupplemental.

Table 1. Crystal parameters, data collection, and refinement statistics for P1 and P4 constructs

Protein	P1	P4
Crystal parameters		
Space group	$P6_3$	$P6_1$
Unit-cell parameters		
a = b (Å)	142.4	144.18
c (Å)	104.9	208.89
Data collection statistics		
Temperature, K	100	100
Resolution range, Å	20.0–3.2	20.00–2.20
Total reflections	93,041	827,770
Unique reflections	19,866	122,974
Completeness, %	99.4 (99.6)	99.3 (95.8)
R_{sym}*	0.097 (0.433)	0.054 (0.231)
Mean $I/\sigma(I)$	14.2 (2.9)	28.5 (5.2)
Refinement statistics		
Resolution range, Å	20.0–3.2	20.00–2.20
Number of reflections used	19,009	119,944
R factor†, %	22.9	15.7
R free†, %	27.3	18.1
rmsd from ideal geometry		
Bond lengths, Å	0.011	0.007
Bond angles, °	1.8	1.3
Protein atoms	7,588	8,116
Water molecules	39	1,662
Inhibitor atoms	—	52
Average B factor, A^2	45.9	24.6
Ramachandran plot		
Most favored, %	84.1	91.8
Additionally allowed, %	15.6	8.2

*$R_{sym} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the i*th* measurement and $<I(hkl)>$ is the weighted mean of all measurements of $I(hkl)$.

†*R* factor = $\Sigma_{hkl}||F_o(hkl)|-|F_c(hkl)||/\Sigma_{hkl}|F_o(hkl)|$; *R* free calculated with 5% of data withheld from refinement. Values in parentheses refer to the highest resolution shell. Values in parentheses refer to the highest resolution shell (3.3–3.2 Å and 2.28–2.20 Å for P1 and P4, respectively).

remaining ones (P2–P4) contained only the CA catalytic domain. In P4, Cys-41 (hCA I numbering system), a residue involved in an intermolecular disulfide bond responsible for CA IX dimerization (17), was mutated into Ser. Proteins were purified to homogeneity by Ni^{2+}-based chromatography; the His-tag was removed using a protease site that was included in the sequences. Proteins were further purified by size exclusion chromatography. Protein purity and homogeneity were assessed by SDS/PAGE in reducing and non-reducing conditions, and by LC-ESI-IT MS analysis. In the latter case, MS spectra confirmed the occurrence of high mannose-type glycans N-linked to all recombinant products (17).

Crystallization, Structure Resolution, and Refinement. Initial crystallization experiments were performed on the P1 recombinant product. Large well-formed crystals were obtained with the hanging-drop vapor diffusion method. Upon diffraction with synchrotron radiation, data were collected to 3.2-Å resolution. The crystals belonged to the space group $P6_3$ with four molecules per asymmetric unit (Table 1). The structure was solved by molecular replacement using the crystallographic structure of mCA XIV as starting model (9). Despite low-resolution data, most of the residue main chains of the catalytic domain were defined in the electron density maps. In contrast, no electron density was observed for the entire PG domain, strongly suggesting the absence of this portion within the crystals. To check this hypothesis, crystals were dissolved and analyzed by SDS/PAGE, revealing a band migrating at the molecular mass of the sole CA domain. This result was confirmed by MALDI-TOF PMF analysis. The absence of the entire PG domain was then ascribed to P1 degradation as a consequence of contaminants present within the purification/crystallization solutions. Additional crystallization experiments performed on the P1 construct were not successful.

On the basis of the observation that the PG domain was not present in the obtained crystals, subsequent crystallization experiments were performed on the other CA IX recombinant products (P2–P4) containing only the catalytic domain either in their unbound or inhibitor-bound form. Acetazolamide (AZM), a strong inhibitor with a nanomolar affinity for hCA IX (17), was used to this purpose. Crystals were obtained only in the case of the P4/AZM complex and they belonged to the space group $P6_1$ with four molecules per asymmetric unit. Upon diffraction with synchrotron radiation, data were collected to 2.20-Å resolution (Table 1). The structure was solved by molecular replacement using the P1 low resolution structure as starting model. In this case, all residues were well defined in the electron density maps, except for the N-terminal region of chains B, C, and D (residues 2 and 3). The structure of the P4 recombinant product was identical to that of P1; hence, we shall describe only its structure in the following sections, referring to P4 as CA IX catalytic domain, unless otherwise stated.

Overall Structure and Active Site. The four molecules in the asymmetric unit of the CA IX catalytic domain crystals, called A, B, C, and D, revealed only minor differences: selecting A as the reference, B, C, and D showed an rmsd for the superimposition of the Cα atoms of 0.27 Å, 0.28 Å, and 0.30 Å, respectively. For this reason, the subsequent discussion is based on one arbitrarily chosen monomer, unless otherwise stated.

The CA IX catalytic domain appeared as a compact globular domain, whose roughly ovoid shape is approximately $47 \times 35 \times 42$ $Å^3$ in size. Its structure revealed a fold characteristic of other α-CAs, whose three-dimensional structure has been solved (2–10), in which a 10-stranded antiparallel β-sheet forms the core of the molecule. Two glycans of the pentasaccharide core N-linked to Asn-213 (17) were clearly visible in the electron density maps of all molecules in the asymmetric unit.

As observed in the structure of other α-CA isozymes, the CA IX active site is located in a large conical cavity, which spans from the surface to the center of the protein. The zinc ion is located at the bottom of this cavity. Two distinct regions made of hydrophobic or hydrophilic amino acids delimit the active site. In particular, Leu-91, Val-121, Val-131, Leu-135, Leu-141, Val-143, Leu-198, and Pro-202 define the hydrophobic region, while Asn-62, His-64, Ser-65, Gln-67, Thr-69, and Gln-92 identify the hydrophilic one. A clear electron density for AZM molecule was evident within the active site of each molecule in the asymmetric unit (Fig. 1). The main protein-inhibitor interactions are depicted in Fig. 1.

P1 and P4 Dimeric Arrangement. We recently reported that purified hCA IX is a dimeric protein, and its dimerization is mediated by an intermolecular disulfide bond involving Cys-41 (17). In agreement with these findings, two identical dimers (consisting of the pairs AB and CD), resulting from a Cys-41-mediated intermolecular disulfide bond between two adjacent monomers, were observed in the asymmetric unit of the P1 crystals (Fig. 2*A*). Within each dimer, the two monomers were related by a non-crystallographic 2-fold symmetry axis. The two dimers interacted with each other generating a tetramer that was recognized by the protein interfaces, surfaces and assemblies service PISA (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) (18) as a potential stable oligomeric arrangement in solution. However, further size-exclusion chromatography and dynamic light-scattering experiments, performed at different pH

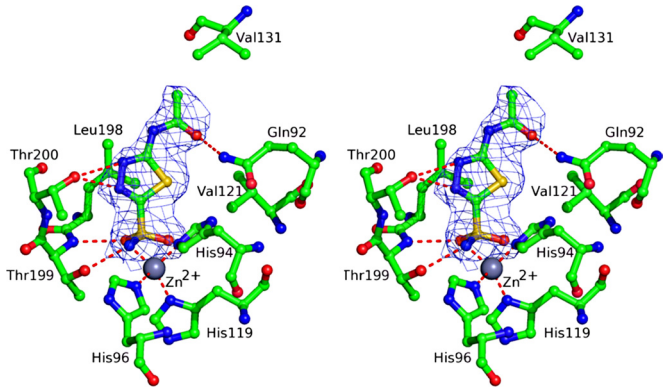

Fig. 1. Stereoview of the active site region of the hCA IX/AZM complex, showing residues participating in AZM recognition. The simulated annealing omit $|2Fo - Fc|$ electron density map of the inhibitor molecule is also shown.

and protein concentrations, did not reveal the presence of a tetrameric structure in any of the tested conditions. Thus, we assume that the observed tetrameric arrangement was induced by the crystal environment, and the physiological oligomeric arrangement of CA IX is dimeric.

Despite the presence of the Cys-41/Ser mutation and the absence of the intermolecular disulfide bond, P4 construct still crystallized with two dimers (consisting of the pairs AB and CD) in the asymmetric unit. These dimers were structurally identical to those observed in the asymmetric unit of the P1 crystals (Fig. 2*B*); by selecting the dimer AB of the P4 crystals as the reference, the rmsd for its superposition with the dimer CD in the same crystals and with the dimer AB and CD of the P1 crystals were 0.27 Å, 0.62 Å, and 0.55 Å, respectively. Interestingly, the sole difference detectable at the dimer interface between P4 and P1 structures was due to the presence of Ser-41 instead of the Cys residue. In P4, the side chains of the two Ser residues did not interact with each other and pointed to opposite directions (Fig. 2*B*). Two hydrogen bonds, involving the Arg-137

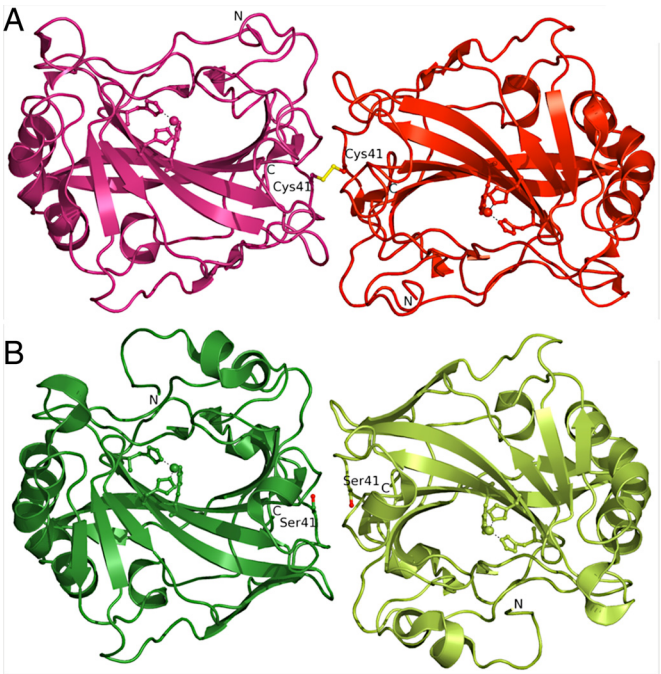

Fig. 2. The dimeric structure observed for P1 (*A*) and P4 (*B*).

side chain and the carbonyl oxygen of Ala-127, and numerous van der Waals interactions, contribute to the stabilization of the dimer, whose interface area extends over 1,590 Å^2.

Structural Comparison with Other Membrane-Associated CAs. As expected on the basis of the high sequence identity observed (Fig. 3), a structural comparison of the CA IX catalytic domain with that of the other membrane-associated α-CAs (2, 8, 9) revealed a substantial degree of three-dimensional similarity. In fact, the rmsd for the superposition of the backbone atoms of the CA IX catalytic domain with the corresponding atoms in hCA IV (8), hCA XII (9) and mCA XIV (2) was calculated as 1.7 Å, 1.4 Å, and 1.5 Å, respectively. The main differences were observed in the N-terminal regions and in the portion 232–238 of these isozymes, where a large sequence deletion is present in CA IX (Fig. 3). Since these portions are located on the protein surface and far away from the active site, the observed differences should not influence either catalytic activity or inhibitor binding. The active site region was almost superimposable between the four enzymes, with only minor differences in the region 125–137, due to one-two amino acid deletions in CA IX with respect to the other isozymes (Fig. 3). Sequence alignment reported in Fig. 3 revealed that the structural similarity in the active site region was associated to a very high conservation of the amino acids delimiting the cavity. Only eight amino acid residues in the active site region differed among the four isozymes, namely residues 67, 69, 91, 131, 132, 135, 136, and 204; interestingly, four of these are localized in the region 125–137. The simultaneous observation of sequence and structural differences in this specific region make it an interesting target to address in the rational drug design of selective CA IX inhibitors.

All crystallographic studies reported so far for mammalian CA isozymes describe monomeric structures (2–8, 10), with the exception of hCA XII for which a dimeric assembling was presented (9). However, a structural comparison revealed that the CA IX dimeric interface was completely different from that reported for hCA XII isozyme (Fig. 4). The observation of an original dimeric structural organization for CA IX was quite surprising, especially in view of its high sequence homology with respect to the other α-CAs. This prompted us to investigate the structural elements responsible for the absence of this oligomeric arrangement in the other CA isozymes. Formation of the CA IX dimer, both in P1 and P4 products, involves a contact area comprising region Ala-39–Ala-43, which also contains Cys-41 involved in the intermolecular disulfide. Measurement of the Cα-Cα distance between the two residues present at position 41 in the P1/P4 dimeric structures revealed a value not higher than 5.3 Å. On the other hand, a structure-based sequence alignment revealed that an Asp residue is always present at position 41 in the other α-CA isozymes with known three-dimensional structure (Fig. 3). Thus the occurrence of this conserved Asp residue and its bulky side chain should destabilize the interactions at the eventual CA dimer interface and would not be compatible with oligomerization.

pH-Dependent Activity of the Various CA IX Products. The pH-dependent profile for the CO_2 hydration reaction catalyzed by P1 and P4 recombinant products was assayed for comparative purposes. Significant differences between the two enzymes were observed (Fig. S2). In fact, the titration curve representing the pH-dependent variation of the k_{cat}/K_M value for the P4 product, consisting uniquely of the CA IX catalytic domain, showed a pK_a value of 7.01. This result was similar to that observed for other enzymes containing uniquely the CA catalytic domain, namely CA I and II, for which a pK_a value of 6.90 and 7.10 was measured, respectively (Fig. S3). In contrast, the same curve for the P1 product, containing both PG and CA domains, showed a pK_a value of 6.49. These results demonstrate that the longer CA IX

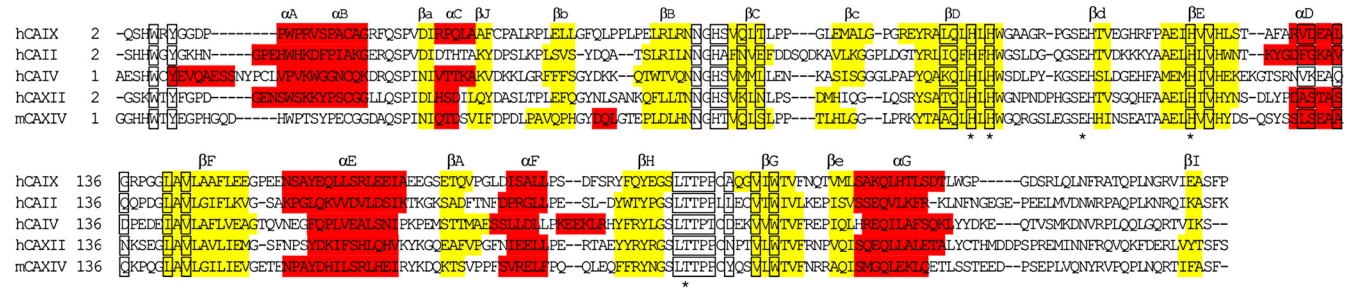

Fig. 3. Structure-based sequence alignment of hCA IX with the membrane-associated CAs whose three-dimensional structure has been solved. hCA II is also included in the alignment for comparison. The helix regions are shown in red and β-strand regions are in yellow. Catalytic triad residues, Thr-199 and Glu-106 are starred, while residues delimiting the active site cavity are boxed.

product, also containing the PG domain, acts as a better catalyst for CO_2 hydration at more acidic pH values, with an optimal activity at pH 6.49. Interestingly this slightly acidic pH value is within a typical pH range of solid and hypoxic tumors, where CA IX is generally expressed (13, 19, 20).

Discussion

A wealth of research data has accumulated during the last decade on the functional aspects of CA IX, even if complete understanding has not yet been obtained of the role of the entire protein and its domains in growth, survival, migration, and invasion of tumor cells. Careful examination of the existing literature seems to attribute a different role to each of the CA IX domains (21). The main player in the growth and survival of tumor cells seems to be the CA IX catalytic domain that converts CO_2, produced in the cytoplasm of hypoxic cells and diffused through the plasma membrane, into bicarbonate and protons contributing to extra-cellular acidosis. The newly generated HCO_3^- ions could then be transported back into the tumor cells or to blood capillaries by HCO_3^- transport proteins. The coupled transport process is probably essential for hypoxic cancer cells to buffer their intracellular pH value to near neutral conditions necessary for their biosynthetic reactions (21). On the other hand, the PG domain function should be related to cell adhesion and intercellular communication (22–24), thus suggesting that it may be the CA IX moiety most probably involved in tumor invasion. It has been proposed that hypoxia can induce tumor invasion via decreased E-cadherin mediated cell-cell adhesion, and CA IX may participate in this process, since it has been shown to reduce this adhesion by interacting with β-catenin (21, 24, 25). Finally, the intracellular tail of CA IX has been shown to participate to signal transduction, since it contains a Tyr residue that, when phosphorylated, can interact with the regulatory subunit of PI3K, which contributes to the activation of the Ser/Thr protein kinase Akt, thereby activating a number of cancer-related signaling cascades, such as the HIF1-mediated one (26).

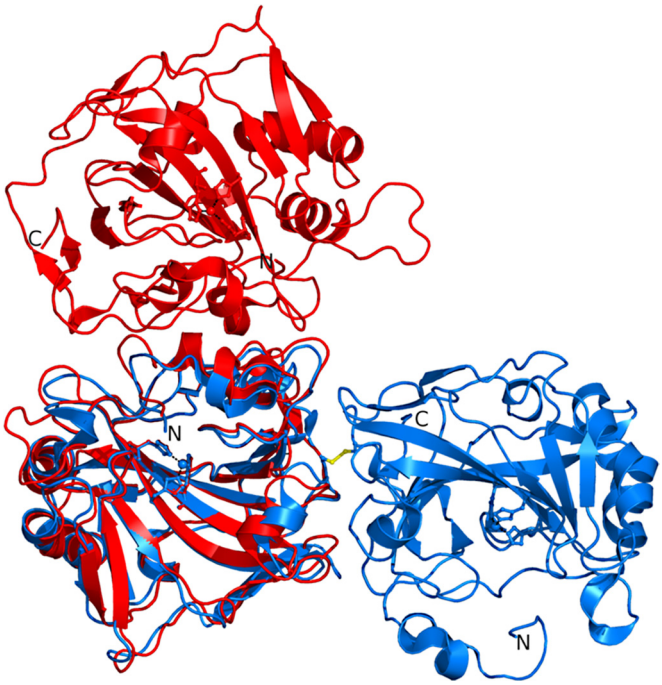

Fig. 4. Dimeric structure of CA IX (blue) and CA XII (red) catalytic domain. To highlight the different dimeric arrangement, monomers A of both dimers have been superposed.

The crystal structure of the catalytic domain of human CA IX, reported in this work, provides structural background to justify the various functions observed for the different CA IX domains. In particular, our crystallographic data confirm the dimeric nature of the enzyme (Fig. 2). The mode of dimer assembly is consistent with the proposed function of the enzyme in tissues where its expression/overexpression has been reported, being both active sites of the dimer clearly exposed to the extracellular medium for efficient CO_2 hydration (Fig. 5). In addition, the N-terminal regions of both monomers are located on the same face of the dimer, while both the C-termini are situated on the opposite face (Fig. 5). This structural organization allows for concomitant positioning of both PG domains, at the entrance to the active site clefts, oriented toward the extracellular milieu to mediate cell interaction, and of both C-terminal transmembrane portions for proper CA IX anchoring to the cell membrane (Fig. 5). More importantly, the position of the PG portion, at the border of the active site, may suggest a further role of this domain in assisting CA domain-mediated catalysis. This hypothesis was further confirmed by our kinetic experiments, which demonstrated the influence of the PG domain on the pH-dependent profile of the k_{cat}/K_M values for CO_2 hydration. In fact, while the CA IX catalytic domain alone presented an optimal pK_a at 7.01, this value changed to 6.49 whenever the PG domain was present in the recombinant product. This presence ensures that CA IX has a better catalytic efficiency at the acidic pH values that the enzyme has to face. These data seem to offer an additional role to the PG domain beyond those already reported (21), since it contributes to the improvement of the CA IX catalytic features at the acidic pH values at which it generally operates, especially in the hypoxic tumors. In this context, it is also worth noting that the upper side of the active site is delimited by three Arg residues, namely Arg-58, Arg-60, and Arg-130, which are not conserved in other members of the α-CA family. Since the PG domain is rich in negatively charged amino acids (eight Asp and 18 Glu residues in a total of 58 residues), it is tempting to

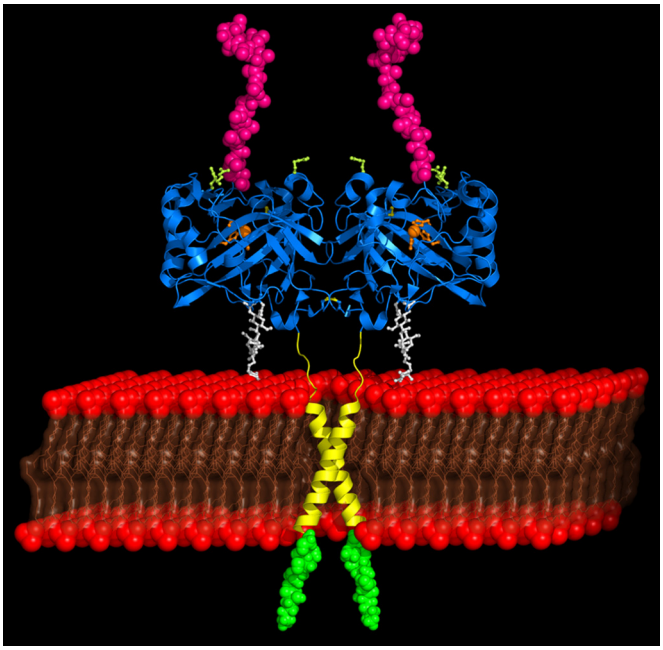

Fig. 5. Proposed model showing the structural arrangement of the full-length CA IX dimer on the cellular membrane. The X-ray structure of the dimeric catalytic domain is reported in cyan, with Arg-58, Arg-60, and Arg-130 in green and the glycan moieties in white. The hypothetical arrangements of the PG domains, the transmembrane helices and the cytoplasmic portions are schematically reported in magenta, yellow, and green, respectively.

speculate that the PG domain could interact with positively charged residues that delimit the active site, sterically controlling substrate accession or participating in the proton-transfer reaction (1). Further experiments are currently under way to confirm these hypotheses.

The crystal structure resolution of the CA IX catalytic domain also provides important results for the structure-based drug design of selective CA IX inhibitors with potential as anticancer drugs. Owing to the extracellular location of isozymes CA IV, IX, XII, and XIV, an initial step in the design of selective inhibitors can be obtained with the development of membrane-impermeant inhibitors, which inhibit membrane-associated CAs without interacting with the cytosolic or mitochondrial isoforms. This selectivity has been recently obtained through the design of positively charged sulfonamides that, being permanently charged, are restricted to the extracellular space (1, 27, 28). Structural comparison with the other membrane-associated CA isozymes reported herein, offers some additional insights for the design of more selective CA IX inhibitors. In particular, since the region 125–137 differs both in structure and sequence in all these isozymes, it represents a "hot spot" to be considered in structure-based drug design.

The results presented in this work provide structural background for the elucidation of the complex enzymatic machinery associated to CA IX, a multidomain membrane protein consisting of extra-, trans-, and intra-cellular portions, which act together to generate different biochemical functions (21). Interestingly, a similar PG-CA domain combination has also been reported in extracellular parts of some human receptor protein Tyr phosphatases (RPTPβ/γ) and a related rat phosphacan, which contain inactive CA domains and are expressed in the form of proteoglycans (29). These proteins also contain a cytoplasmic phosphorylation-based signaling module, function as receptor molecules and, through the binding of neuronal adhesion molecule contactin/tenascin, mediate adhesive communication between astroglial cells and gonadotrophin-releasing hormone neurons (30, 31). In some cases, RPTPs modulate their activity through a disulfide-based dimerization process involving the same Cys residue present in two monomers (32–34). We believe that the occurrence of a number of structural/functional analogies with some RPTPs as well as the recognition of CA IX as a valuable target for cancer diagnostics and treatment will promote further crystallographic studies to completely elucidate the whole CA IX structure. These investigations will definitively allow the rational drug design of isozyme-specific inhibitors for the different CA IX functions.

Materials and Methods

Production and Purification of Recombinant CA IX Enzymes. The recombinant CA IX enzymes used for the crystallization studies were produced using the Bac-To-Bac Baculovirus Expression System (Invitrogen) (17). Four CA IX proteins, whose sequences are reported in Fig. S1, were constructed. All of the enzymes had a CA IX signal sequence, whose codon usage was optimized for *Spodoptera frugiperda*, eight His residues for protein purification and a protease cleavage site (thrombin or PreScission protease, GE Healthcare) to remove the His-tag. The cDNA sequences were constructed by stepwise elongation sequence-PCR similarly to what recently described (17).

The production and purification of the recombinant enzymes as well as the thrombin treatment for P1 were performed as previously described (17). Before protease treatment of P2-P4 products, they were changed to a buffer of 50 mM Tris-HCl, pH 7.5. To remove the His-tag from P2, the protein was treated with 170 μL resin-coupled thrombin (Thrombin CleanCleave KIT, Sigma-Aldrich) per 1 mg protein, with gentle mixing at 25 °C for 1 h. The thrombin and the His-tag were removed from the solution as previously described (17). For P3 and P4, the His-tag was removed by incubating proteins with 2 U GST-tagged PreScission protease per 1 mg protein at 4 °C for 20 h. PreScission protease was removed using Bulk GST Purification module (GE Healthcare), and the His-tag was removed from the solution as previously described (17). As final purification step the proteins were loaded on a Superdex 200 10/30 (P1) or Superdex 75 10/30 (GE Healthcare) (P2-P4) column equilibrated with 50 mM phosphate, 100 mM NaCl, pH 8.0 (P1), or 30 mM Tris-HCl, 100 mM NaCl, pH 7.5 (P2–P4). Highly purified proteins were pooled, concentrated to 5 mg/mL and buffer exchanged to water before crystallization experiments. Purity was assessed by SDS/PAGE and LC-ESI-IT MS as reported previously (10).

Crystallization and Data Collection. The crystals of the P1 recombinant product were obtained at 20 °C using the hanging-drop vapor diffusion method. Drops of 2 μL, prepared mixing 1 μL protein solution with an equal volume of precipitant solution (0.4 M $NH_4H_2PO_4$, 0.1 M sodium citrate, pH 5.0), were suspended over wells containing 1.0 mL precipitant. In about two weeks crystals of $0.2 \times 0.1 \times 0.1$ mm^3 were grown. P1 diffraction data were collected at 3.20-Å resolution from a single crystal at the temperature of 100 K, using a MAR CCD detector, at the synchrotron source Elettra in Trieste. The crystals were transferred to a cryoprotectant solution containing 35% (vol/vol) glycerol, in addition to the precipitant buffer, before X-ray data diffraction collection. Data were processed using the HKL crystallographic data resolution package (Denzo/Scalepack) (35) (Table 1). The crystals belonged to space group $P6_3$ with four molecules in the asymmetric unit according to a solvent content of 53.4%.

P4/AZM complex was prepared by adding a 10-M excess of inhibitor to a 5 mg/mL protein solution in water. The crystals were grown using the hanging drop vapor diffusion technique at 20 °C by using a precipitant solution whose formulation was similar to that of P1 enzyme (1.0 M $NH_4H_2PO_4$, 0.1 M sodium acetate, pH 4.0). Crystals appeared in the drops within 3 days and grew in about 2 weeks to maximum dimensions of $0.4 \times 0.2 \times 0.2$ mm^3.

P4/AZM X-ray diffraction data were collected and processed as for P1, with the exception that the cryoprotectant solution contained 25% glycerol. The crystals belonged to the hexagonal space group $P6_1$ with four complex molecules in the asymmetric unit and showed an unusual solvent content of 77.2%, which corresponded to a Matthews coefficient of 5.4 Å^3 Da^{-1}. Data reduction statistics for both P1 and P4 are recorded in Table 1.

Structure Resolution. The P1 crystal structure was solved by molecular replacement using the program AMoRe (36) and the crystallographic structure of mCA XIV (PDB code 1RJ5) (9) as starting model. P1 structure refinement was performed with CNS (37) and model building was performed with O (38). The four molecules in the asymmetric unit were refined by using NCS-restraints and thermal B-factors were initially fixed at 20.0 Å^2. Many cycles of positional and temperature-factor refinement alternated with manual rebuilding were necessary to reduce the crystallographic Rfactor and Rfree values to 0.229 and 0.273, respectively.

The structure of P4/AZM complex was solved by molecular replacement em-

ploying the refined P1 structure as search model. After initial refinement limited to the enzyme structure, a model for the inhibitor, clearly visible in the electron density maps, was built, and introduced into the atomic coordinate set for further refinement. The first cycles of the refinement were carried out with 4-fold NCS-restraints. After Rfactor and Rfree reached 0.287 and 0.297, respectively, the NCS-restraints were removed and further cycles of manual rebuilding and positional and temperature-factor refinement were necessary to reduce the crystallographic Rfactor and Rfree values to 0.157 and 0.181, respectively. Refinement statistics for both P1 and P4 are summarized in Table 1.

CA Activity Assay. The CA-activity was measured as previously described (17). Assays were performed on P1 and P4 recombinant CA IX forms. For obtaining the pH dependency, the following buffer/indicator pairs were used: pH 6.0, Mes, methyl red, working at 433 nM; pH 6.2 and 6.5, Mes, chlorophenol red, working at 574 nM; pH 6.8, HEPES, 4-nitrophenol, 400 nM; pH 7.0, HEPES, neutral red, working at 540 nM; pH 7.4, HEPES, phenol red, 557 nM (this is the buffer-indicator pair in which most measurements of CA catalytic and inhibition are usually performed, due to the physiological pH value of 7.4) (39).

ACKNOWLEDGMENTS. This work was supported by a 6th Framework Program (FP) of the European Union (DeZnIT project) grant. We thank the Sincrotrone Trieste Consiglio Nazionale delle Richerche/Elettra for giving us the opportunity to collect data at the Crystallographic Beamline.

1. Supuran CT (2008) Carbonic anhydrases: Novel therapeutic applications for inhibitors and activators. *Nat Rev Drug Discov* 7:168–181.
2. Whittington DA, et al. (2001) Crystal structure of the dimeric extracellular domain of human carbonic anhydrase XII, a bitopic membrane protein overexpressed in certain cancer tumor cells. *Proc Natl Acad Sci USA* 98:9545–9550.
3. Boriack-Sjodin PA, Heck RW, Laipis PJ, Silverman DN, Christianson DW (1995) Structure determination of murine mitochondrial carbonic anhydrase V at 2.45 Å resolution: Implications for catalytic proton transfer and inhibitor design. *Proc Natl Acad Sci USA* 92:10949–10953.
4. Kannan KK, Ramanadham M, Jones TA (1984) Structure, refinement, and function of carbonic anhydrase isozymes: Refinement of human carbonic anhydrase I. *Ann NY Acad Sci* 429:49–60.
5. Eriksson AE, Jones TA, Liljas A (1988) Refined structure of human carbonic anhydrase II at 2.0 Å resolution. *Proteins* 4:274–282.
6. Eriksson AE, Liljas A (1993) Refined structure of bovine carbonic anhydrase III at 2.0 Å resolution. *Proteins* 16:29–42.
7. Duda DM, et al. (2005) Human carbonic anhydrase III: Structural and kinetic study of catalysis and proton transfer. *Biochemistry* 44:10046–10053.
8. Stams T, et al. (1996) Crystal structure of the secretory form of membrane-associated human carbonic anhydrase IV at 2.8 Å resolution. *Proc Natl Acad Sci USA* 93:13589–13594.
9. Whittington DA, et al. (2004) Expression, assay, and structure of the extracellular domain of murine carbonic anhydrase XIV: Implications for selective inhibition of membrane-associated isozymes. *J Biol Chem* 279:7223–7228.
10. Di Fiore A, et al. (2008) Crystal structure of human carbonic anhydrase XIII and its complex with the inhibitor acetazolamide. *Proteins* 74:164–175.
11. Wykoff CC, et al., (2000) Hypoxia-inducible expression of tumor-associated carbonic anhydrases. *Cancer Res* 60:7075–7083.
12. Potter C, Harris AL (2004) Hypoxia inducible carbonic anhydrase IX, marker of tumour hypoxia, survival pathway and therapy target. *Cell Cycle* 3:164–167.
13. Pastoreková S, Pastorek J (2004) in *Carbonic Anhydrase*: *Its Inhibitors and Activators*, eds Supuran CT, Scozzafava A, Conway J (CRC Press, Boca Raton, FL), pp 255–281.
14. Swietach P, et al. (2008) Tumor-associated carbonic anhydrase 9 spatially coordinates intracellular pH in three-dimensional multicellular growths. *J Biol Chem* 283:20473–20483.
15. Pastorek J, et al. (1994) Cloning and characterization of MN, a human tumor-associated protein with a domain homologous to carbonic anhydrase and a putative helix-loop-helix DNA binding segment. *Oncogene* 9:2877–2888.
16. Opavský R, et al. (1996) Human MN/CA9 gene, a novel member of the carbonic anhydrase family: Structure and exon to protein domain relationships. *Genomics* 33:480–487.
17. Hilvo M, et al. (2008) Biochemical characterization of CA IX, one of the most active carbonic anhydrase isozymes. *J Biol Chem* 283:27799–27809.
18. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.
19. Svastová E, et al. (2004) Hypoxia activates the capacity of tumor-associated carbonic anhydrase IX to acidify extracellular pH. *FEBS Lett* 577:439–445.
20. Chiche J, et al. (2009) Hypoxia-inducible carbonic anhydrase IX and XII promote tumor cell growth by counteracting acidosis through the regulation of the intracellular pH. *Cancer Res* 69:358–368.
21. Pastorekova S, Zatovicova M, Pastorek J (2008) Cancer-associated carbonic anhydrases and their inhibition. *Curr Pharm Des* 14:685–698.
22. Závada J, et al. (2000) Human tumour-associated cell adhesion protein MN/CA IX: identification of M75 epitope and of the region mediating cell adhesion. *Br J Cancer* 82:1808–1813.
23. Zavadova Z, Zavada J (2005) Carbonic anhydrase IX (CA IX) mediates tumor cell interactions with microenvironment. *Oncol Rep* 13:977–982.
24. Svastová E, et al. (2003) Carbonic anhydrase IX reduces E-cadherin-mediated adhesion of MDCK cells via interaction with beta-catenin. *Exp Cell Res* 290:332–345.
25. Beavon IRG (1999) Regulation of E-cadherin: Does hypoxia initiate the metastatic cascade? *J Clin Pathol Mol Pathol* 52:179–188.
26. Dorai T, Sawczuk IS, Pastorek J, Wiernik PH, Dutcher JP (2005) The role of carbonic anhydrase IX overexpression in kidney cancer. *Eur J Cancer* 18:2935–2947.
27. Pastorekova S, et al. (2004) Carbonic anhydrase inhibitors: The first selective, membrane-impermeant inhibitors targeting the tumor-associated isozyme IX. *Bioorg Med Chem Lett* 14:869–873.
28. Scozzafava A, Briganti F, Ilies MA, Supuran CT (2000) Carbonic anhydrase inhibitors: Synthesis of membrane-impermeant low molecular weight sulfonamides possessing in vivo selectivity for the membrane-bound versus cytosolic isozymes. *J Med Chem* 43:292–300.
29. Barnea G, et al. (1994) Receptor tyrosine phosphatase beta is expressed in the form of proteoglycan and binds to the extracellular matrix protein tenascin. *J Biol Chem* 269:14349–14352.
30. Peles E, et al. (1995) The carbonic anhydrase domain of receptor tyrosine phosphatase beta is a functional ligand for the axonal cell recognition molecule contactin. *Cell* 82:251–260.
31. Parent AS, et al. (2007) A contactin-receptor-like protein tyrosine phosphatase beta complex mediates adhesive communication between astroglial cells and gonadotrophin-releasing hormone neurones. *J Neuroendocrinol* 19:847–859.
32. van der Wijk T, Overvoorde J, den Hertog J (2004) H_2O_2-induced intermolecular disulfide bond formation between receptor protein-tyrosine phosphatases. *J Biol Chem* 279:44355–44361.
33. Groen A, Overvoorde J, van der Wijk T, den Hertog J (2008) Redox regulation of dimerization of the receptor protein-tyrosine phosphatases RPTPalpha, LAR, RPTPmu and CD45. *FEBS J* 275:2597–2604.
34. Tabernero L, Aricescu AR, Jones EY, Szedlacsek SE (2008) Protein tyrosine phosphatases: Structure-function relationships. *FEBS J* 275:867–882.
35. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
36. Navaza J (1994) AMoRe: An automated package for molecular replacement. *Acta Crystallogr Sect A* 50:157–163.
37. Brünger AT, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr Sect D* 54:905–921.
38. Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr Sect A* 47:110–119.
39. Behravan G, Jonsson BH, Lindskog S (1990) Fine tuning of the catalytic properties of carbonic anhydrase. Studies of a Thr200 à His variant of human isoenzyme II. *Eur J Biochem* 190:351–357.