This un-edited manuscript has been accepted for publication in Biophysical Journal and is freely available on BioFast at http://www.biophysj.org.
The final copyedited version of the paper may be found at http://www.biophysj.org.

STRUCTURAL FEATURES THAT GOVERN ENZYMATIC ACTIVITY IN CARBONIC ANHYDRASE FROM A LOW TEMPERATURE ADAPTED FISH *CHIONODRACO HAMATUS*

Stefano Marino[1], Kuniko Hayakawa[1,2], Keisuke Hatada[2], Maurizio Benfatto[2], Antonia Rizzello[3], Michele Maffia[3], Luigi Bubacco[1]

1 Department of Biology, University of Padova, Padova, Italy
2 Laboratori Nazionali di Frascati dell'INFN - INFN, c.p. 13, Frascati, Italy
3 Department of Biological and Environmental Science and Technology, University of Salento, Lecce, Italy

Running Title: Icefish carbonic anhydrase

Address correspondence to: Luigi Bubacco, Department of Biology, University of Padova, Viale Ugo bassi 58B, 35121, Padova, Italy. Tel 0039-049-8276346; Fax 0039-049-8276300; Email: luigi.bubacco@unipd.it

The carbonic anhydrase (CA[1]) family of zinc metalloenzymes includes many known isozymes which have different subcellular distributions. The study described here is focused on the identification of the structural features that define low temperature adaptation in a *Chionodraco hamatus* protein, both at an atomic level for the reaction center and for the tertiary structure of the protein. To this aim, a XANES/MXAN analysis of the reaction center was undertaken for both a structurally characterized human CAII (hCAII) and for CA of *C. hamatus* (Ice-CA). Higher structural levels were analyzed by sequence comparison and homology modeling. To establish if the structural insights acquired in fish CAs are general, theoretical models were generated by homology modeling for three temperate climate-adapted fish CAs. The measured structural differences between the two proteins are discussed in terms of the differences in the electrostatic potential between hCAII and Ice-CA. We conclude that modulation of the interaction between the catalytic water molecule and the zinc ion could depend on the effect of the electrostatic potential distribution.

The carbonic anhydrase (CA) family of zinc metalloenzymes includes many known isozymes divided in five distinct classes (α, β, γ, δ and ε) that appear to have evolved independently (1-3). In mammals only the α class isoforms are found with different subcellular distributions: cytoplasmic (CAI, II, III, VII and XIII), plasma membrane associated (CAIV, IX, XII, XIV and XV), mitochondrial (CA VA and VB), or secreted (CAVI) (1,4,5). CAs catalyse the hydration of carbon dioxide to yield a bicarbonate ion and a proton. This reaction occurs through several chemical steps (6-8). The first is a nucleophilic attack of the zinc-bound hydroxide ion to the substrate CO_2 followed by the exchange of the product bicarbonate ion with a water molecule. The metal-bound hydroxide ion is regenerated by transfer of a proton to bulk solvent. The rate constant k_{cat} generally reflects proton transfer between zinc-bound water and an active site "shuttle" residue and the subtle structural differences among the CA isozymes contribute to define a wide range of k_{cat} values (10^3-10^6 s^{-1}) measured for this rate-limiting step. Different residues may function as the intermediary proton shuttle between zinc-bound water and bulk solvent in isozymes.

There are several crystal structures of CAs in the PDB (protein data bank) and for vertebrate enzymes, most of them mammalian CAII (with or without a bound inhibitor). Crystallographic structures are also available for both mammalian CAI (again, with or without a bound inhibitor) and CAIII (the S-glutathiolated form of rat CAIII). All of these structures provide key insights into structure-function relationships for this enzyme (9). The zinc cofactor is placed at the base of the conical active site cleft where it is coordinated to three histidine residues and one water molecule to yield a tetrahedral geometry. At physiological pH, the fourth zinc ligand is a hydroxide molecule which acts as a nucleophile to catalyze the hydrolysis of CO_2 and esters (at a rate near to the limit posed by substrate diffusion) (10).

Other important structural features relate to residues that either protrude into the catalytic cleft or contribute to generate a hydrophobic pocket in which the substrate seems to be held. Among the first, the proton shuttle mentioned above plays a key role in the catalytic process. For instance, His64 is the catalytic proton shuttle in CAII ($k_{cat}=10^6$ s^{-1}) (11,12) and the tautomerization of this residue can mediate the transfers of both protons and water molecules at a neutral pH with high efficiency, requiring no time- or energy-consuming processes (13). The side chain of His200 is the best candidate for a proton shuttle in CAI ($k_{cat}=2x10^5$ s^{-1}) (14) due to its proximity to the zinc ion and the observed pKa value (15). In contrast, it seems that no proton shuttle is present in the active site of CAIII ($k_{cat}=1x10^4$ s^{-1}), in which the product proton is probably transferred directly to the bulk solvent (16).

The active site cleft is split into hydrophobic and hydrophilic regions. Mutagenesis and structural analyses of the hydrophobic pocket adjacent to the zinc-bound hydroxide indicate this region as the site of CO_2 association and that the hydrophobicity modulates the catalytic activity (10,17). The proposed association would place the substrate at about 3 Å from the zinc atom. There is a well-defined hydrophobic pocket adjacent to the zinc-bound hydroxide group. The amino acids involved are, Trp209, Val121, and Leu/Phe198 that define the mouth and the side plus Val143 that defines the bottom. The amino acids that define the hydrophobic pocket are completely conserved in carbonic anhydrase II (with Leu198) and CAIII (with Phe198). Steric effects occurring at the mouth of the pocket have been proposed to justify the lower catalytic activity measured for CAIII; these seem to depend on Phe198 (18).

Among the polar residues of the active site, the hydroxyl group of the Thrl99 side chain forms hydrogen bonds with the zinc-hydroxide group and the Glu106 side chain to form a Zn-OH-/Thr199/Glu106 hydrogen bond network (17,19), which is completely conserved in all animal carbonic anhydrase isozymes. The hydrogen bond between Thr199 and zinc-hydroxide has been postulated to be important to maintain the orientation and reactivity of the zinc-hydroxide group (20), catalyzing proton transfer, and discriminating between protic and aprotic anions (21).

Amino acid substitutions at position 65 of CAII demonstrate that the size of the side chain is critical in the assembly of a bridging solvent network between the zinc-bound solvent and the shuttle residue His64. In particular, the presence in position 65 of both alanine (as in human CAII) or serine (as in almost all other vertebrate CAII) is compatible with a high proton shuttling rate for the adjacent His64 (12,22).

Considering more physiological aspects, carbonic anhydrase plays a crucial role in the excretion of metabolic CO_2 in all vertebrates. In fish, CO_2 produced in the tissues is rapidly hydrated to HCO_3^- by the action of erythrocyte CA. The generated HCO_3^- then diffuses into the blood stream, where it accounts for almost 98 % of the total carbon dioxide stored and transported in the plasma (23,24). At the respiratory epithelia in gills or skin, CA catalyzes the rapid dehydration of HCO_3^- to molecular CO_2, which then diffuses passively into the ventilatory water stream. The CO_2/HCO_3^- system constitutes the most important physiological buffers for acid–base regulation (25).

The Antarctic icefish of the family Channichthyidae (suborder Nothotheniodei), to which *Chionodraco hamatus* belong, are a unique example of adult vertebrates lacking hemoglobin and functionally active erythrocytes, and possessing only a small number of erythrocyte-like cells (26).The absence of hemoglobin does not represent a dramatic limitation to oxygen transport in the icefish. On the other hand, the very limited number of erythrocyte-like cells (and circulating CA) may compromise the CO_2/HCO_3^- equilibrium in the blood (27). However, the elevated solubility of CO_2 in water at low temperatures, and the unusual characteristics of the circulatory system of icefish, contributes to an efficient excretion of CO_2. In teleosts, CA has been found in various tissues. It appears to be present in high concentrations in the gills (28), where it plays an important role in osmoregulation, nitrogen (ammonia) excretion, acid–base balance, and gas exchange (29). Comparisons between activity of *C. hamatus* gill CA (Ice-CA), another icefish *Trematomus bernacchii* (an Antarctic fish with hemoglobin and red blood cells) and the temperate-adapted fish *Anguilla anguilla* showed that CA activity for the fish lacking hemoglobin is high up to a temperature of 30 °C. Above this temperature, there is a dramatic decrease of activity that becomes virtually absent at 37 °C. At this temperature, *A. anguilla* CA is still at its maximum level of activity (30). With these data, we were particularly interested in the identification of the structural features that define low temperature adaptation in the *C. hamatus* protein, both at an atomic level for the reaction center as well as the tertiary structure of the protein. A XANES/MXAN analysis of the reaction center was undertaken for both a structurally characterized human CAII (hCAII) and for Ice-CA (SWISS-PROT accession number P83299). The higher structural levels were analyzed by sequence comparison and homology modeling. To evaluate the general applicability of the structural insights acquired by our study of CAs, we created some theoretical models using homology modeling, generated for three temperate climate-adapted fish CAs.

EXPERIMENTAL PROCEDURES

Protein purification - Cytoplasmic carbonic anhydrase of the icefish *C. hamatus* gill filaments (Ice-CA) was obtained as previously described (30). Ice-CA was purified by FPLC affinity chromatography on p-aminomethylbenzene-sulphonamide immobilized on cyanogen-bromide-activated agarose gel (30,31). The gel column (1.6 x 20 cm), fitted to a AKTA-Pharmacia FPLC system, was equilibrated with: Tris 25 mM, Na_2SO_4 100 mM, adjusted to pH 8.7 with HCl and rinsed with Tris 25 mM, $NaClO_4$ 300 mM, adjusted to pH 8.7 with HCl; then the enzyme was eluted at 8 ml h^{-1} by CH_3COOH 100 mM, $NaClO_4$ 500 mM, pH 5.6 at 4 °C. Protein elution was monitored by measuring the eluate absorbance at 280 nm, and all fractions containing CA activity, measured by an electrometric method (30), were pooled and concentrated by ultra filtration with YM10 membrane (Amicon Corp, Lexington, U.S.A.), under nitrogen pressure ($7x10^5$ Pa). All purification steps were carried out at 4 °C. Purified human erythrocyte CAII was purchased by SIGMA (St. Louis, MO).

Sample preparation - Powered solid solutions of the protein in saccharose were obtained starting from protein dialyzed against the appropriate buffer for 24 hour at 4° C. Buffers were prepared for a final concentration of 100 mM using secondary and tertiary ammonium bases, such as piperazine and triethanolamine, with glycine and aliphatic carboxylic acids (32). The protein solution, at a concentration of about 10 mg/ml and containing saccharose in a saccharose/protein ratio of 3:1 w/w (corresponding to a concentration of about 0.250 g of protein per gram of final solid-solution), was then rapidly frozen in liquid nitrogen and lyophilized. The solid samples of the CA form were obtained simply by lyophylization of the protein solution in the presence of the cryoprotectants.

The specimens for XAS measurements were prepared by pressing about 50 mg of lyophilized powdered protein solid-solution under 140 atm in a small home-made press having a chamber with lateral dimensions of 2.2 x 0.3 cm; the slides obtained were about 0.2 cm thick.
Solutions of irradiated proteins, used to evaluate the integrity of the samples, were obtained by dissolving the corresponding slide in milliQ water and dialyzing against the appropriate buffer for 48 hours at 4°C.
The XANES (X-ray Absorption Near Edge Structure) measurements were performed at the LURE synchrotron facility, in the D21 (EXAFS II) beamline. The Zn K-edge XANES signals were collected in fluorescence mode, at room temperature, by a seven-element CANBERRA® Ge detector. A Si (311) crystal was used as a monochromator. The energy range was from 9600 to 9800 eV. (Calibration of energy was performed by means of a zinc foil reference). The energy resolution was 0.3 eV, with a counting time of 6 s/point. The absorption spectra were collected using a Perkin-Elmer Lambda 16 double beam spectrophotometer equipped with a thermostated cell holder.
The XANES data analysis was performed by using the MXAN method (33) capable to fit the energy region from the edge up to 200 eV above the threshold in terms of selected structural parameters.
The MXAN method performs the minimization of the residual function R_{sq} defined as:

$$R_{sq} = \sum_{i=1}^{m} \frac{[(y_i^{th} - y_i^{\exp})\varepsilon_i^{-1}]^2}{m},$$

where *m* is the number of experimental data, y_i^{th} and y_i^{exp} are the theoretically computed and experimental measured values of the absorption coefficient, respectively. Our estimate of the experimental error is ε_i, considered constant in the fitting procedure, and in our case, considered equal to about 0,8 % of the experimental edge jump. Details of the MXAN procedure are described in (34-37).
The cluster of atoms used in the fitting procedure includes atoms up to 6.3 Å from the absorber and the muffin-tin radii are chosen according to the Norman criterion with a 5 % overlap. The number of atoms in the cluster used for the calculation was chosen on the basis of a convergence criterion. This cluster was generated starting from the PDB coordinates of the human CA II (PDB reference, 2cba). No core-hole was considered in the final state potential.
Fitting Strategy - The atomic positions are described using a polar coordinate system; the absorbing atom is placed at the origin of the coordinate frame. The atoms of one protein ligand or amino acid residue are considered as a perfectly rigid molecular group and when the Zn-N distance is changed in the fitting, all atoms of the histidine rings move accordingly. In the fit, the bond of the Zn-O of the coordinated water molecule was allowed to vary in both in length and polar angles.
The ionization energy E_o is 9660 eV (see figures 3 and 4), and the spectra were consistently normalized to 1. The contribution of the pre-edge energy region was subtracted using the standard procedure.
Modeling - The first step in choosing the appropriate templates was to use the structure prediction server BioInfoBank Meta Server (http://bioinfo.pl/meta/). The search can be customized choosing the prediction methods that have to be considered by 3D-jury; specifically, the threading method was 'deselect'. Then, the option 'model-to-all' was chosen, which considers the ten best models from each server previously selected to be taken into account (38). For the modeling procedure, the Swiss Model method was used (http://swissmodel.expasy.org//SWISS-MODEL.html) in combination with the DeepView 3.7 (sp5) program (http://swissmodel.expasy.org/spdbv/).

Electrostatic potential distributions were calculated by using numerical solution of the Non-Linear Poisson–Boltzmann equation implemented in APBS software (http://apbs.sourceforge.net/).
All of the CA structures were superimposed and similarly oriented before the calculations. AMBER forcefield was used for partial atomic charges and radii, internal and external dielectric constant values of 2 and 78 were used, and solvent and ionic probe radii of 1.4 and 1.9 Å (with mono-valent salt concentration of 150mM) .
For visualization PyMol (http://pymol.sourceforge.net/) with APBS plugin (APBS_tools.py) was used: APBS calculated values were plotted on a gaussian type surface, as it can be seen in Fig. 2 (according to the coding colour described in the text of Figure 2).
Hex 4.5 (http://www.csd.abdn.ac.uk/hex) was also used for electrostatic potential calculation, which solves the differential Poisson equation for electrostatic potential and which gives as output a value for the potential (in mV) for the tested protein.
To test the models, the SAVS validation server was used (http://nihserver.mbi.ucla.edu/SAVS/).
Sequence analysis - For sequence analysis, Blast (http://www.ncbi.nih.gov/BLAST/), ClustalW (http://www.ebi.ac.uk/clustalw/) and custom designed software were used.

RESULTS

Template structure for the MXAN analysis - The choice of an adequate structural template to be used in the MXAN analysis of the absorption k-edge data for Ice-CA was carried out using the Metap server. The ranking algorithm used was 3D-jury (see Material and Methods section). The structural file resulted was made up of best-confidence predictions (all hits were with a score above 226): 1flj, PDB code for S-gluathiolated *Rattus norvegicus* CAIII (rCAIII), with a resolution of 1.9Å (39); 1v9i, *Bos taurus* CAII (BCAII), with a Gln253>Cys253 mutation and a resolution of 2.5 Å; 2cba**,** *Homo sapiens* CAII (hCAII), with a high resolution of 1.54 Å (17); 12ca, hCAII, with an Ala121>Val121 mutation and a resolution of 2.40 Å (10); 1hcb, hCAI complexed with bicarbonate and a resolution of 1.60 Å (40).
Further criteria used in the selection of the template crystallographic structure were the absence of an exogenous ligand in the active site, wild type sequence, and the level of resolution of the crystal structure. The application of these criteria resulted in two plausible templates: 1flj and 2cba.
In order to choose between these two templates, a sequence analysis was carried out among Ice-CA [SWISS-PROT Protein Data Bank under the accession number P83299, hCAII (that corresponds to the structure 2cba.pdb) and rCAIII (that corresponds to the structure 1flj.pdb) (Figure 1). In addition to the general sequence homology, attention was given to the conserved residues that are believed to be involved in the reaction mechanism, the residues that in the crystal structures are close to the reaction center, and to the fully conserved residues throughout the sequence.
Using the crystal structures 2cba and 1flj and the sequence alignments of these with Ice-CA, it was possible to identify the amino acids in Ice-CA that are likely to be within 10 Å from the zinc atom. The result of the comparison of the selected amino acids indicates that Ice-CA has a higher identity (88%) to hCAII (2cba) than rCAIII (1flj; 76%).
A key residue is His64, which is the proton shuttle of the reaction mechanism in all mammalian CAIIs. The His64 is absent in the corresponding position of rCAIII. Ice-CA has a His residue in the position corresponding to 64 in CAII. Furthermore, this His residue in Ice-

CA is part of a pattern of four highly conserved residues (-GHSF-) present in mammalian CAII.

As mentioned in the introduction, another relevant amino acid for the reactivity of CA is in position 198. This is a phenylalanine in all CAIIIs, but in both mammalian CAII and Ice-CA, it is a leucine which provides further support for the proposed similarity between CAII and Ice-CA.

For the MXAN analysis, the cluster used to generate the simulated absorption spectra includes all atoms within 6.3 Å from the absorption center. Thus, a detailed comparison of all aa positions in this region in 2cba and 1flj was undertaken. An important difference is the presence of a Phe in position 95 in 1flj while a Leu is present in 2cba. This Leu, which is conserved in CAII, is also present in Ice-CA. All atoms of residue 95, which is present in the sequence between the two Zn-coordinating histidine residues (94 and 96), were included in the cluster used for the MXAN analysis, making 2cba the stronger candidate as template. Furthermore, among all of the structures selected by the 3-D jury, 2cba had the highest resolution, confirming it as an ideal starting point for the structural simulations.

Ice-CA Modeling - Starting from the Ice-CA sequence and using the 2cba structure as template (overall 62% of amino acidic identity), the Swiss Model server was used to generate a molecular model of the icefish protein. The refinement of the obtained model was carried out with DeepView 3.7. The resulting model was tested using the SAVS validation server. The obtained model for Ice-CA was a starting point for other computations aimed to compare a few relevant physical and chemical properties of Ice-CA and 2cba.

The first property analyzed was the electrostatic potential distribution using, with undistinguishable results both Hex4.5 and APBS. The two proteins have a significantly different electrostatic potential distribution. Ice-CA has a calculated overall negative potential of -0.22 mV (using Hex4.5). The corresponding value for 2cba (with no water, except the Zn-coordinated one) is +0.70 mV.

A visualization of the electrostatic potential plotted on the Gaussian-type surface is shown in figure 2 (generated using APBS). Different electrostatic potential distributions between the two proteins can be attributed to both a different propensity to the intermolecular interactions and to the properties of the reaction channel that leads the substrate to the zinc atom.

To evaluate the possibility that the observation of a difference in the electrostatic potential distribution between the generated model of Ice-CA and 2cba is more than coincidental, a broader set of high activity carbonic anhydrase (CAII-like) was considered. These were all from fish living in temperate climates: red blood cell (rbc) CA from *Oncorhynchus mykiss* (TCAb, NCBI accession number AAP73748; 41), rbc CAs from *Danio rerio* (ZCA, AAH57412; CAH-Z, AAH65611), cytoplasmic CA from *O. mykiss* (TCAc, AAR99329; 41).

For all of the CAs mentioned above, a molecular model was generated (following the procedure described for icefish CA); the sequence homologies to 2cba were 63% with ZCA, 61% with TCAb, and 63% with TCAc.

For each of the resulting molecular models, the electrostatic potential distribution was calculated. The calculated potential distributions for rbc CAs are similar to that calculated for human CAII, with an overall positive potential (figure 2) in an uneven distribution of neutral, negative and positive potential throughout the molecular surface. The unusual electrostatic potential distribution observed for Ice-CA was also found in the molecular model generated for TCAc. This cytosolic isoform is also found in gills and has a catalytic activity lower than that of the rbc *O. mykiss* CA isoform (41). For comparative purposes, the electrostatic potential distribution was also calculated for the CA of the salt-tolerant unicellular green alga *Dunaliella salina* (pdb code 1y7w; 42).

XANES/MXAN analysis on Ice-CA and hCAII - In figure 3, the normalized experimental k-edge absorption spectra are presented for both the human and the icefish CA. The two spectra

do not show any chemical shift in the time course of the measurements. The most important difference between the two spectra is the feature at about 10 eV that is sharper in the icefish protein (curve A in figure 3). The best fit of the k-edge absorption spectra of the hCAII is presented in figure 4A. The agreement between the experimental data and the calculated spectra is excellent over all spectral regions. The value of the Rsq of the best fit is 4.06. The starting atomic coordinates for the fitting procedure were those obtained from the 2cba PDB file, obtained from the crystal structure of CAII from human erythrocytes. The one-shot calculation of the edge spectra using the crystallographic coordinates gives a Rsq value of 13.77. In Table 1 are reported the bond length values obtained from the fitting procedure for both proteins and, for comparison, also those obtained from the crystal structure of 2cba. The human protein has a greater asymmetric distribution of Zn-N distances which results in a statistically significant difference. Furthermore, the B factors for both the Zn and the atoms of the metal ligands in the 2cba PDB file lead to an expected uncertainty for the bond lengths in the crystallographic data of about 0.25 Å. The reported Zn-O distance for the water molecule coordinated to the Zn atom (H_2O 263 in 2cba) shows an even higher value of the B factors which further increases the uncertainty to 0.32 Å. In figure 4B is presented the best fit of the k-edge absorption spectra of cytosolic CA from *C. hamatus*. The agreement between the experimental data and the calculated spectra is again excellent over all spectral regions, and the value of the Rsq of the best fit is 4.44. The rationale for the choice of the starting coordinates for the MXAN analysis has been in part presented above. Further considerations relate to the result of the simulations of the edge spectrum based only on the coordinates of the crystal structure 2cba (Rsq = 27,36) and 1flj (Rsq = 38,79) which indicate a higher structural similarity of the considered atom cluster for the human enzyme. As mentioned above, a further advantage of the 2cba PDB file is its higher resolution. On these bases, 2cba was used as the template structure for the MXAN fitting procedure for both proteins.

Among the structural parameters explored in the fitting procedure, the most important are the H_2O263–zinc distance (figure 5a) and the theta angle (figure 5b) for the Zn-coordinated water molecule. The H_2O263–zinc distance resulted to be the more relevant structural feature as can be seen in figure 5A. The theta angle for the Zn-coordinated water molecule shows a less pronounced minimum in the Rsq profile for both proteins. (For clarity, only the data for the human protein is reported in figure 5B).

The other determinant structural parameters were the theta angle and the distance from the Zinc atom and Oδ of Thr199. In comparing hCAII and Ice-CA, some important considerations are due on the interplay of zinc-coordinated water and Thr199. In the structural models of the active site obtained in the fitting procedure, a different position is observed for Thr199 (figure 6). The Zn-Oδ (Thr199) distance is shorter in hCAII as is the Zn-O (H_2O263) distance. This is consistent with the observation that Thr199 and the coordinated H_2O are involved in the H-bond network that represents one of the most conserved structural features of cytosolic carbonic anhydrases (CAI, CAII, CAIII).

The zinc-coordinated histidines, that in the fitting procedure were allowed to move, were not effective parameters in the optimization of the fit. The final structural model and the relative bond lengths are presented in figure 6 and in the table (Table 1), respectively. The histidines's ligands are almost superimposable and are not likely to be responsible for the difference between the two proteins analyzed. The first observation is on the comparison between the structural data obtained from the fitting of the human protein and the available crystallographic structure 2cba. As shown in the table, the differences for all considered parameters are within the values of the associated errors. These must be compared to the crystallographic data which are less precise. In comparing the structural models obtained for the human protein and the Ice-CA, a general shortening of bond distances is observed in the reaction center that averages -0.05 Å.

Taking into account the high resolution of the MXAN analysis, we observe that Zn-nitrogen distances are not equivalent for the coordinated histidine. This observation, at least for the human protein, seems to be in contrast with the crystallographic data. However, as previously observed, the magnitude of the differences in distance that generate this asymmetry are smaller than the error associated to the crystallographic data but larger than the estimated error in the MXAN analysis.
The Zn-O distance for the coordinated water molecule is 0.048 Å, a value that seems to be very significant for the low errors associated to the MXAN distance determination. To further enforce this observation, a series of simulations were made to single out the effect of both the Zn-O distance and the theta angle on the Rsq value. Starting from the best fit structure for the two proteins, we generated individual simulations in which only the Zn-O distance was changed step-wise. A graph is shown in figure 5A, in which the Rsq value is plotted against the Zn-O distance. The individual curves for the two proteins show a well-defined difference in the minima for the explored parameter providing strong support to the claimed difference in distance for the coordinated water molecule in the two proteins. These two minima are not superimposable, demonstrating that this difference can be considered highly significant from a statistical point of view. As will be discussed later, the implications of this difference are of importance in the functional properties of CA. A similar analysis was conducted on both proteins also for the coordinated water and the theta angle parameter. Dependence of the fit on this parameter is present only in the human protein. The results of this analysis are shown in figure 5B. A shallow minimum is observed around 15° (assuming the position derived from 2cba as 0°). However, this parameter has a lower impact on the fit when compared with the Zn-O distance. Superimposition of the best-fit corresponding structures for Ice-CA and hCAII is shown in figure 6.

DISCUSSION

To define the molecular basis of cold adaptation, it is important to characterize the structure of those proteins that, for their physiological roles, are more likely to be modified in the process. Carbonic anhydrase is one of these potential target enzymes for its key role in a number of cellular functions. In this study, as a paradigm of cold adaptation molecular mechanisms, the structural characteristics of a soluble CA isoform isolated from gills of an Antarctic haemoglobinless fish *C. hamatus*, were compared to those of other mesophilic CAs.
Among the CA sequences of teleosts, an average identity of 75% was found. Specifically, the CA from *C. hamatus* was compared with the different isoforms known of *Danio rerio* and *Oncorhynchus mykiss*. Identity ranges from 70% with ZCA and 77% with CAH-Z of *D. rerio* to 79% of gill CA (TCAc, also signed as carbonic anhydrase I like isoform, *O. mykiss* CAI) and 77% of red blood cell CA (TCAb, also signed as carbonic anhydrase II-like isoform, *O. mykiss* CAII) from *O. mykiss*.
The comparison of the Ice-CA sequence with alpha carbonic anhydrases present in the PDB structural database indicates a larger similarity with mammalian CAII than other mammalian isoforms for the amino acids within 10 Å (Figure 1) from the catalytic zinc atom, which are more likely to play a key role in the reaction mechanism.
Some further considerations are due on the residues that define the completely conserved local chemical environment of the zinc atom. The first group of conserved amino acids are involved in proton shuttling (Gly63, His64, Ser65/Ala65, and Phe66, using a numbering that is valid for all mammal CAIIs), where serine is always present in position 65 except in the human enzyme where an alanine is present. The presence of a relatively small amino acid (*i.e.*, Ala or Ser) in proximity of the proton shuttle His64 does not affect the mobility of the latter and, consequentially, the kinetic parameters of the enzyme.

The second group of conserved amino acids defines the hydrophobic pocket (Val121, Val143, Trp209 and Leu198) that are conserved in Ice-CA and in all mammalian CAIIs.
A third class of extremely conserved amino acids in all CAs from vertebrates is that in the second coordination shell of the metal ion in the active site. These are the residues that form a hydrogen bond to the non-coordinating nitrogen of the zinc ligand histidines (Gln92 with His94, Glu117 with His119, and Asn244 with His96) and to the fourth metal ligand, the exchangeable water molecule (Thr199).
However, some amino acids in positions close to the zinc atom are different in icefish CA compared to human CA. These are positions 144 and 146, where Leu and Ile are present in hCAII, and valines are present for icefish in both 144 and 146. It should be mentioned that these positions are not particularly conserved among CAs from vertebrates, although the variability is confined to Leu, Ile and Val. For example, CAs from *Bos taurus* and *Ovis aries* show the same amino acids in this region as icefish, while *D. rerio* CAII has Val144 and Ile146.
The last relevant difference among hCAII and Ice-CA refers to position 245 that is always a Trp in mammalian and a Tyr among fish. This position refers to a difference between fishes and mammals and not to a peculiarity of icefish.
In conclusion, the comparative analysis of the amino acids placed within 10 Å from the zinc atom allows the identification of key residues for the reaction mechanism that represent the criteria in the diversification of the various CA isoforms. However, no indication emerges for a unique sequence feature of icefish CA that can rationalize the low temperature adaptation, simply by referring to the amino acids close to the zinc. The amino acid conservation, as expected, decreases moving away from the zinc atom. Consequentially, it becomes difficult to assign relevance to any of the observed difference between icefish CA and other CAs in the database.
On these premises, the most plausible structural template to be used as starting point for the MXAN analysis is the human carbonic anhydrase II (PDB files 2cba). As mentioned above, the simulation of the k-edge absorption spectra with MXAN allows for the definition of a structural model of the active site where the position of the individual atoms of the protein is defined up to about 6 Å from the zinc atom. The experimental strategy implied the measurement of a known sample of human carbonic anhydrase II to be used as a validation test for the analysis of the structurally unknown icefish CA. The structural metric of the active site of hCAII coming from the MXAN analysis is substantially in agreement with the crystallographic data. The limit in this comparison is caused by the resolution of the crystal structure that permits errors in the atomic positions larger than the observed differences. On the other hand, the accuracy of the MXAN analysis allows discrimination between the human and the icefish proteins, highlighting the structural differences of the active sites.
The Zn-coordinated water molecule is an essential player in the reaction mechanism of CA; consequentially, the definition of the bond length was of key importance. As observed in figure 5A, the proposed difference between hCAII and Ice-CA is substantiated by the presence of two distinct minima when the R_{sq} value is plotted against bond length. A further difference between hCAII and Ice-CA is the angle of the Zn-H_2O bond with the plane, defined by the three coordinating nitrogen atoms of the histidine ligands. As observed in figure 5B, the effect is less pronounced than that observed for the bond length, but a minimum in Rsq value is still observed.
A structural feature that relates to those mentioned above is the position of the δO of Thr199, an atom in the crystal structures that is always reported as hydrogen bonded to the Zn-coordinated water molecule. In the R_{sq} minimization procedure, the movement of this atom, placed at 3.9 Å from the zinc atom, is correlated with the movement of the Zn-coordinated water molecule even when these two atoms are allowed to move independently.

Consequentially, a shorter Zn-δOThr199 distance corresponds also to the shorter Zn-H_2O bond length of the hCAII. It is tempting to propose that all of these structural differences concur to the modulation of the pK value of the coordinated water molecule in the two proteins, but no reliable correlation has been defined between these structural features and the actual pK values that are also likely to depend on other properties, such as the local dielectric constant.

No simple correlation can be made between the observed structural difference in the active site and the amino acid composition for the region around the metal center. The amino acid conservation is complete between hCAII and Ice-CA for the amino acids within a sphere of radius 6 Å centered on the zinc atom. The features that define the observed differences in structure and kinetic properties between these two proteins must be related to something other than what is usually described as governing the molecular mechanism of CA, *i.e.*, Thr199, the hydrophobic pocket and the region around the proton shuttle.

The final step in the structural analysis presented here was the generation of a molecular model of Ice-CA by homology modeling of the template provided by the PDB file of hCAII. For both the generated model of Ice-CA and the crystallographic structure of hCAII, the electrostatic potential was calculated. A striking difference emerged from the comparison: an overall negative potential for Ice-CA (-0.2 mV) and a significant positive potential for hCAII (+0.7 mV).

This difference can be easily visualized when the two proteins are presented with a color coding that refers to the surface potential (figure 2 A, E). The surface potential distribution is substantially different, with a difference in the region of the substrate entrance to the active site but more pronounced in the back side of the protein.

To evaluate the observed differences in surface potential to a larger set of CAs, the molecular modeling was extended to other fish CAs: *Danio rerio* CA (ZCA, sequence identity with hCAII 63%), *O. mykiss* rbc CA (TCAb, sequence identity with hCAII 61%) and *O. mykiss* cytosolic CA (TCAc, sequence identity with hCAII 63%). For all of the molecular models, the surface potentials were calculated with the results reported in figure 2. The two CAs from erythrocytes show an electrostatic potential distribution similar to hCAII with an overall positive potential. On the contrary, the cytosolic CA from *O. mykiss* shows an overall negative electrostatic potential distribution very similar to Ice-CA.

The latter result suggests a diversification of fish CAs based more on cell type rather than on species.

In this regard, in a recent study on the CA of a green algae *Dunaliella salina* (42), a relation has been proposed between the halo-tolerance of this organism and the observed electrostatic potential distribution that is reported in figure 2F for comparison. The authors emphasize the relation between the overall very negative electrostatic potential and the interaction with anions (particularly halides). More precisely, the low potential in the reaction center associated to the surface potential could rationalize the modulation in the affinity constant of the zinc atom in the active site toward halide binding, together with the other unusual properties of solubility and stability of the algae CA.

The negative electrostatic potential of the branchial Ice-CA and TCAc isoforms could be responsible for their substrate affinity values significantly lower than that reported for the trout rbc CA (30,41). Moreover, a different electrostatic potential in and around the active site of the gill Ice-CA could also explain the observed high catalytic rates at low temperatures (30), with the consequent advantages of a rapid interconversion of CO_2/HCO_3^- in the blood in spite of the lack of a circulating CA in the Antarctic haemoglobinless teleosts (30). According to the structural features already found in other cold adapted enzymes (43,44) as for example subtilisins excreted by Antarctic bacteria (45), an high number of negative charge residues

could improve at low temperatures the protein surface interactions with the solvent, the flexibility of the enzyme structure and a better positioning of the substrates in the active site.
On these bases, the predicted difference in the electrostatic potential between hCAII and Ice-CA represents the best candidate to justify the measured structural differences between these two proteins. The modulation of the interaction between the catalytic water molecule and the zinc atom could, as demonstrated for an analogous interaction of halides, depend on the effect of the electrostatic potential distribution. It should be mentioned, however, that the electrostatic potential distribution itself, being substantially comparable in the "only" two cytosolic fish CA considered, may be not sufficient to account for the cold adaptation of this molecule calling for future studies on other cold-adapted species.

REFERENCES

1. Hewett-Emmett, D., and R. E. Tashian. 1996. Functional diversity, conservation, and convergence in the evolution of the alpha-, beta-, and gamma-carbonic anhydrase gene families. Mol. Phylogenet. Evol. 5:50-77.
2. So, A. K., G. S. Espie, E. B. Williams, J. M. Shively, S. Heinhorst, and G. C. Cannon. 2004. A novel evolutionary lineage of carbonic anhydrase (epsilon class) is a component of the carboxysome shell. J. Bacteriol. 186:626-630.
3. Sawaya, M. R., G. C. Cannon, S. Heinhorst, S. Tanaka, E. B. Williams, T. O. Yeates, and C. A. Kerfeld. 2006. The structure of beta-carbonic anhydrase from the carboxysomal shell reveals a distinct subclass with one active site for the price of two. J. Biol. Chem. 281:7546-7555.
4. Tashian, R. E., D. Hewett-Emmett, N. Carter, and N. C. Bergenhem. 2000. Carbonic anhydrase (CA) -related proteins (CA-RPs), and transmembrane proteins with CA or CA-RPdomains. Exs. 90:105-120.
5. Hewett-Emmett, D. 2000. Evolution and distribution of the carbonic anhydrase gene families. *In* The Carbonic Anhydrases: New Horizons, W. R. Chegwidden, N. D. Carter and Y. H. Edwards, editors. Birkhäuser Verlag, Boston, USA. 29-76,
6. Lindskog. S. 1997. Structure and mechanism of carbonic anhydrase. Pharmacol. Ther. 74:1-20.
7. Christianson, D. W., and C. A. Fierke. 1996. Carbonic anhydrase: evolution of the. zinc binding site by nature and by design. Acc. Chem. Res. 29:331–339.
8. Christianson D. W., and J. D. Cox. 1999. Catalysis by metal-activated hydroxide in zinc and. manganese metalloenzymes. Annu. Rev. Biochem. 68:33-57.
9. Tripp B.C., K. Smith, J.G. Ferry. 2001. Carbonic Anhydrase: New Insights for an Ancient Enzyme. J. Biol. Chem. 276:48615-48618.
10. Nair, S. K., T. L. Calderone, D. W. Christianson, and C. A. Fierke. 1991. Altering the mouth of a hydrophobic pocket. Structure and kinetics of human carbonic anhydrase II mutants at residue Val-121. J. Biol. Chem. 266:17320-17325.
11. Steiner, H., B. H. Jonsson, B and S. Lindskog. 1975. The catalytic mechanism of carbonic anhydrase. Hydrogen-isotope effects on the kinetic parameters of the human C isoenzyme. Eur. J. Biochem. 59:253-259.
12. Tu, C. K., D. N. Silverman, C. Forsman, B. H. Jonsson, and S. Lindskog. 1989. Role of histidine 64 in the catalytic mechanism of human carbonic anhydrase II studied with a site-specific mutant. Biochemistry 28:7913-7918.
13. Shimahara H., T. Yoshida, Y. Shibata, M. Shimizu, Y Kyogoku, F. Sakiyama, T. Nazakawa, S. Tate, S. Ohki, T. Kato, H. Moriyama, K. Kishida, Y. Tano, T. Ohkubo, and Y. Kobayashi. 2007. Tautomerism of histidine 64 associated with proton-transfer in catalysis of carbonic anhydrase. J. Biol. Chem. Doi, 10.1074/jbc.M609679200.
14. Khalifah, R. G. 1971. The carbon dioxide hydration activity of carbonic anhydrase. I. Stop-flow kinetic studies on the native human isoenzymes B and C. J. Biol. Chem. 246:2561-2573.
15. Engstrand, C., B. H. Jonsson, and S. Lindskog. 1995. Catalytic and inhibitor-binding properties of some active-site mutants of human carbonic anhydrase I. Eur. J. Biochem. 229:696-702.
16. Jewell, D. A., C. K. Tu, S. R. Paranawithana, S. M. Tanhauser, P. V. LoGrasso, P. J. Laipis, and D. N. Silverman. 1991. Enhancement of the catalytic properties of human carbonic anhydrase III by site-directed mutagenesis. Biochemistry 30:1484-1490.

17. Hakansson, K., M. Carlsson, L. A. Svensson, and A. Liljas. 1992. Structure of native and apo carbonic anhydrase II and structure of some of its anion-ligand complexes. J. Mol. Biol. 227:1192-1204.
18. Duda, D.M., C. Tu, S.Z. Fisher, H. An, C.Yoshioka, L. Govindasamy, P.J. Laipis, M. Agbandje-McKenna, D.N. Silverman, and R. McKenna. 2005. Human carbonic anhydrase III: structural and kinetic study of catalysis and proton transfer. Biochemistry 44:10046-53.
19. Eriksson, A. E., T. A. Jones, and A. Liljas. 1988. Refined structure of human carbonic anhydrase II at 2.0 A resolution. Proteins 4:274-282.
20. Merz, K. M., Jr. 1990. Insights into the function of the zinc hydroxide-Thr199-Glu106 hydrogen bonding network in carbonic anhydrases. J. Mol. Biol. 214:799-802.
21. Krebs, J. F., J. A. Ippolito, D. W. Christianson, and C. A. Fierke. 1993. Structural and functional importance of a conserved hydrogen bond network in human carbonic anhydrase II. J. Biol. Chem. 268:27458-27466.
22. Scolnick, L. R., and D. W. Christianson. 1996. X-ray crystallographic studies of alanine-65 variants of carbonic anhydrase II reveal the structural basis of compromised proton transfer in catalysis. Biochemistry 35:16429-16434.
23. Perry, S. F., and P. Laurent. 1990. The role of carbonic anhydrase in carbon dioxide excretion, acid-base balance and ionic regulation in aquatic gill breathers. *In* Animal Nutrition and Transport Processes. 2. Transport, Respiration and Excretion, vol. 6 T. J.-P. and B. Lahlou, editors. Basel: Karger. 39-57.
24. Henry, R. P., and E. R. Swenson. 2000. The distribution and physiological significance of carbonic anhydrase in vertebrate gas exchange organs. Respir. Physiol. 121:1-12.
25. Heisler, N. 1984. Acid-base regulation in fishes. *In* Fish Physiology, vol XA, W. S. Hoar and D.J. Randall editors. Academic Press, London. 315-401.
26. MacDonald, J. A., and R. M. G. Wells. 1991. Viscosity of body fluids from Antarctic Notothenioid fish. *In* The Biology of Antarctic Fishes. Springer Verlag, Berlin. 145-162.
27. Feller, G., and C. Gerday. 1997. Psychrophilic enzymes: molecular basis of cold adaptation. Cell. Mol. Life Sci. 53:830-841.
28. Rahim, S. M., J. P. Delaunoy, and P. Laurent. 1988. Identification and immunocytochemical localization of two different carbonic anhydrase isoenzymes in teleostan fish erythrocytes and gill epithelia. Histochemistry 89:451-459
29. Henry, R. P., and T. A. Heming. 1998. Carbonic anhydrase and respiratory gas exchange. *In* Fish Physiology Haemoglobin and Respiration, vol. 17. S. F. Perry and B. L. Tufts, editors. Academic Press, San Diego.
30. Maffia, M., A. Rizzello, R. Acierno, M. Rollo, R. Chiloiro, and C. Storelli. 2001. Carbonic anhydrase activity in tissues of the icefish *Chionodraco hamatus* and of the red-blooded teleosts *Trematomus bernacchii* and *Anguilla anguilla.* J. Exp. Biol. 204:3983-3992.
31. Whitney, P. L. 1974. Affinity chromatography of carbonic anhydrase. Anal. Biochem. 57:467–476.
32. Ascone, I., A. Sabatucci, L. Bubacco, P. Di Muro, and B. Salvato. 2000. Saccharose solid matrix embedded proteins: a new method for sample preparation for X-ray absorption spectroscopy. Eur. Biophys. J. 29:391-397.
33. Benfatto, M., and S. Della Longa. 2001. Geometrical fitting of experimental XANES spectra by a full multiple-scattering procedure. J. Synchrotron. Radiat. 8:1087-1094.
34. Della Longa, S., A. Arcovito, M. Girasole, J. L. Hazemann, and M. Benfatto. 2001. Quantitative Analysis of X-Ray Absorption Near Edge Structure Data by a Full Multiple Scattering Procedure: The Fe-CO Geometry in Photolyzed Carbonmonoxy-Myoglobin Single Crystal. Phys. Rev. Lett. 87:155501-1-155501-4.

35. Benfatto, M., S. Della Longa, and C. R. Natoli. 2003. The MXAN procedure: a new method for analysing the XANES spectra of metallononeproteins to obtain structural quantitative information. J. Synchrotron Radiat. 10:51-57.
36. Frank, P., M. Benfatto, R. K. Szilagyi, P. D'Angelo, S. Della Longa, and K. O. Hodgson. 2005. The solution structure of $[Cu(aq)]^{2+}$ and its implications for rack-induced bonding in blue copper protein active sites. Inorganic chemistry 44:1922-1933.
37. Sarangi, R., M. Benfatto, K. Hayakawa, L. Bubacco, E. I. Solomon, K. O. Hodgson, and B. Hedman. 2005. MXAN analysis of the XANES energy region of a mononuclear copper complex: applications to bioinorganic systems. Inorganic chemistry 44:9652-9659.
38. Ginalski, K., A. Elofsson, D. Fischer, and L. Rychlewski. 2003. 3D-Jury: a simple approach to improve protein structure predictions. Bioinformatics 19:1015-1018.
39. Mallis, R. J., B. W. Poland, T. K. Chatterjee, R. A. Fisher, S. Darmawan, R. B. Honzatko, and J. A. Thomas. 2000. Crystal structure of S-glutathiolated carbonic anhydrase III. FEBS Lett 482:237-241.
40. Kumar, V., and K. K. Kannan. 1994. Enzyme-substrate interactions. Structure of human carbonic anhydrase I complexed with bicarbonate. J. Mol. Biol. 241:226-232.
41. Esbaugh, A. J., S. F. Perry, M. Bayaa, T. Georgalis, J. Nickerson, B. L. Tufts, and K. M. Gilmour. 2005. Cytoplasmic carbonic anhydrase isozymes in rainbow trout *Oncorhynchus mykiss*: comparative physiology and molecular evolution. J. Exp. Biol. 208:1951-1961.
42. Premkumar, L., H. M. Greenblatt, U. K. Bageshwar, T. Savchenko, I. Gokhman, J. L. Sussman, and A. Zamir. 2005. Three-dimensional structure of a halotolerant algal carbonic anhydrase predicts halotolerance of a mammalian homolog. Proc. Natl. Acad. Sci. U S A. 102:7493-7498.
43. Davail, S., G. Feller, E. Narinx, and C. Gerday. 1994. Cold adaptation of proteins. Purification, characterization, and sequence of the heat-labile subtilisin from the antarctic psychrophile Bacillus TA41. J. Biol. Chem., 269:17448–17453.
44. Feller, G., F. Payan, F. Theys, M. Qian, R. Haser, and C. Gerday. 1994. Stability and structural analysis of alpha-amylase from the antarctic psychrophile *Alteromonas haloplanctis* A23. Eur. J. Biochem., 222:441–447.
45. Narinx, E., E. Baise, and C. Gerday. 1997. Subtilisin from psychrophilic antarctic bacteria: characterization and site-directed mutagenesis of residues possibly involved in the adaptation to cold. Protein Eng.10:1271-1279.

[1] The abbreviations used are: CA, carbonic anhydrase; XANES, X-ray Absorption Near Edge Spectroscopy; MXAN, Minuit Xanes; aa, amino acid; FPLC, Fast Protein Liquid chromatography; XAS, X-ray Absorption Spectroscopy; rbc, red blood cells

Acknowledgements
This work is in the framework of the Italian National Programme for Antarctic Research. The authors also wish to thank Dr. Isabella Ascone for the excellent technical support at the LURE facility. We wish to thank Prof. Federico Fogolari for the helpful discussions on the manuscript.

Table 1:

Bond lengths obtained for the best fit procedures for the experimental data. The crystallographic bond lengths extracted from the crystal structure of human carbonic anhydrase II (2cba) are presented for comparison.

	Ice-CA	***hCAII***	***2cba***
d (Nε His94-Zn)	2,044 (±0,0001)*	1,996 (±0,035)	2,10 (±0,25)**
d (Nε His96-Zn)	2,076 (±0,033)	2,137 (±0,034)	2,12 (±0,25) **
d (Nδ His119-Zn)	2,084 (±0,046)	2,045 (±0,036)	2,11 (±0,25) **
d (HOH263-Zn)	2,044 (±0,0001)*	1,996 (±0.020)	2,05 (±0,32) **
d (Oγ Thr199-Zn)	4,060 (±0,059)	3,913 (±0.038)	3,93 (±0,25) **

* In considering the reported error value of the statistical error (evaluated by the MIGRAD subroutine of the MINUIT program (F. JAMES, CERN, Geneva), it should be mentioned that the systematic error is on the order of 1%, as inferred for previous work on model systems.

** The reported uncertainty in the distances represents the root mean square deviation calculated from the individual B values reported in the PDB file.

FIGURE LEGENDS

Figure 1: Alignment of *C. hamatus* CA, *H. sapiens* CAII, *R. norvegicus* CAIII.
Label on the sequence: *, refers to the positions near the reaction center that are (typically) different in mammalian CAII and CAIII. Underlined are the amino acids placed (with one or more of their atoms) within 10 Å distance from the zinc atom.

Figure 2: Electrostatic potential distribution mapped on the Gaussian-type surface calculated by APBS and visualized with PyMol for ZCA (A), hCAII (B), TCAb (C), TCAc (D), Ice-CA (E) and dCAII (F). The structures are graphically depicted looking down the active site cleft (first row) and in a 180° rotated view (second row). The potentials are coloured to range from -1.5 kT per proton charge (red) to +1.5 kT per proton charge (blue).

Figure 3: Zn K-edge XANES experimental spectra for Ice-CA *C. hamatus* (A) and hCAII *H. sapiens* (B).

Figure 4: Comparison of the best-fit calculation (continuous line) and the experimental data (crosses) of hCAII (Panel A) and Ice-CA (Panel B).

Figure 5: Panel A) Rsq profile as function of the bond length of Zn-O of the H_2O 263. Line a) hCAII, line b) Ice-CA. Panel B) Rsq profile as function of the theta angle between the z-axis and the Zn-O bond. For simplicity, the 0° value corresponds to the orientation of the Zn-O bond in the crystal structure of 2cba.

Figure 6: Comparison between the structural models of the reaction centers of Ice-CA (blue) and hCAII (red) obtained in the MXAN analysis. Zinc atom (violet) and coordinated waters are shown in ball and stick mode, while Thr199 and coordinated histidines are shown in wire frame mode. In addition to the three coodinating histidine residues on the left side of the figure, Thr199 is also shown .

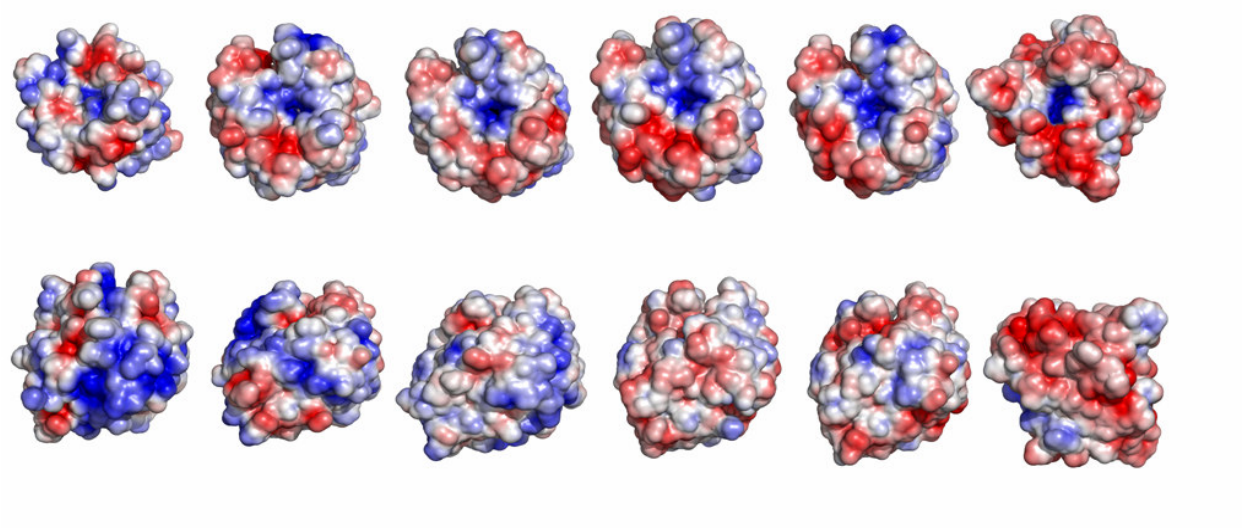
Figure_Marino_1

```
Ice-CA  .AHAWGYGPTDGPDKWVSNFPIADGPRQSPIDILPGGASYDSGLKPLSLKYDPS
hCAII   MSHHWGYGKHNGPEHWHKDFPIAKGERQSPVDIDTHTAKYDPSLKPLSVSYDQA
rCAIII  .AKEWGYASHNGPEHWHELYPIAKGDNQSPIELHTKDIRHDPSLQPWSVSYDPG

Ice-CA  NCLEILNNGHSFQVTFADDSDSSTLKEGPISGVYRLKQFHFHWGASNDKGSEHT
hCAII   TSLRILNNGHAFNVEFDDSQDKAVLKGGPLDGTYRLIQFHFHWGSLDGQGSEHT
rCAIII  SAKTILNNGKTCRVVFDDTFDRSMLRGGPLSGPYRLRQFHLHWGSSDDHGSEHT
                 *                                 *
Ice-CA  VAGTKYPAELHLVHWNTKYPSFGEAASKPDGLAVVGVFLKIGDANASLQKVLDA
hCAII   VDKKKYAAELHLVHWNTKYGDFGKAVQQPDGLAVLGIFLKVGSAKPGLQKVVDV
rCAIII  VDGVKYAAELHLVHWNPKYNTFGEALKQPDGIAVVGIFLKIGREKGEFQILLDA

Ice-CA  FNDIRAKGKQTSFADFDPSTLLPGCLDYWTYDGSLTTPPLLESVTWIVCKEPIS
hCAII   LDSIKTKGKSADFTNFDPRGLLPESLDYWTYPGSLTTPPLLECVTWIVLKEPIS
rCAIII  LDKIKTKGKEAPFNHFDPSCLFPACRDYWTYHGSFTTPPCEECIVWLLLKEPMT
                                          *
Ice-CA  VSCEQMAKFRSLLFSAEGEPECCMVDNYRPPQPLKGRHVRASFQ
hCAII   VSSEQVLKFRKLNFNGEGEPEELMVDNWRPAQPLKNRQIKASFK
rCAIII  VSSDQMAKLRSLFASAENEPPVPLVGNWRPPQPIKGRVVRASFK
```

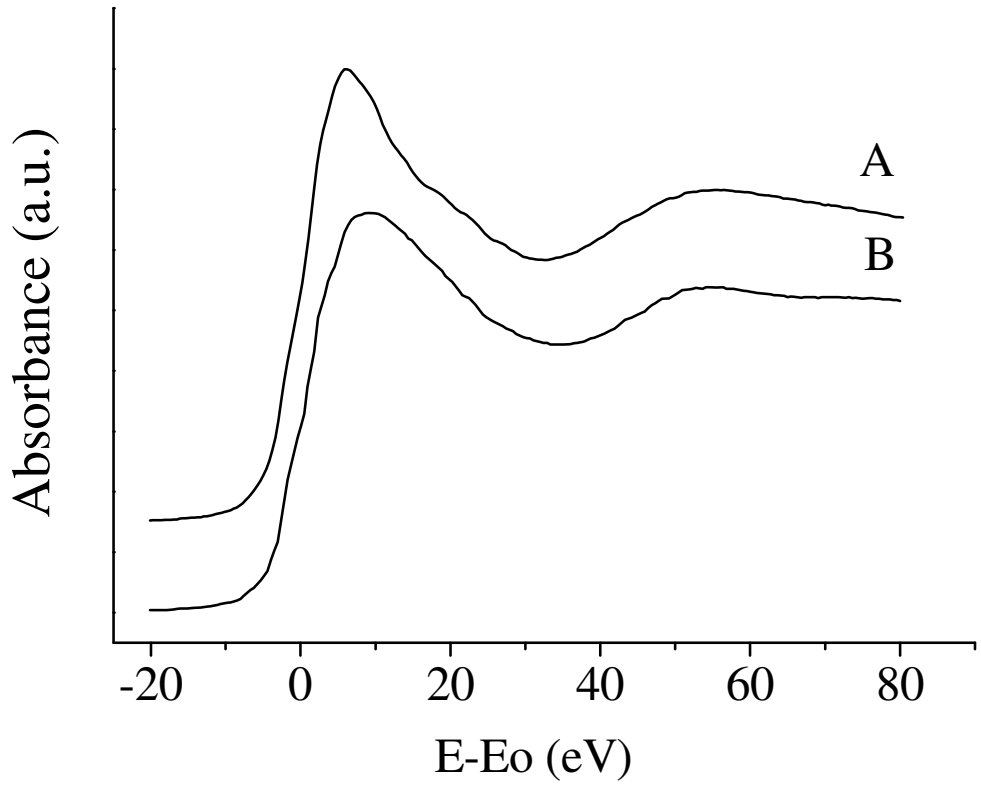
Figure_Marino_2



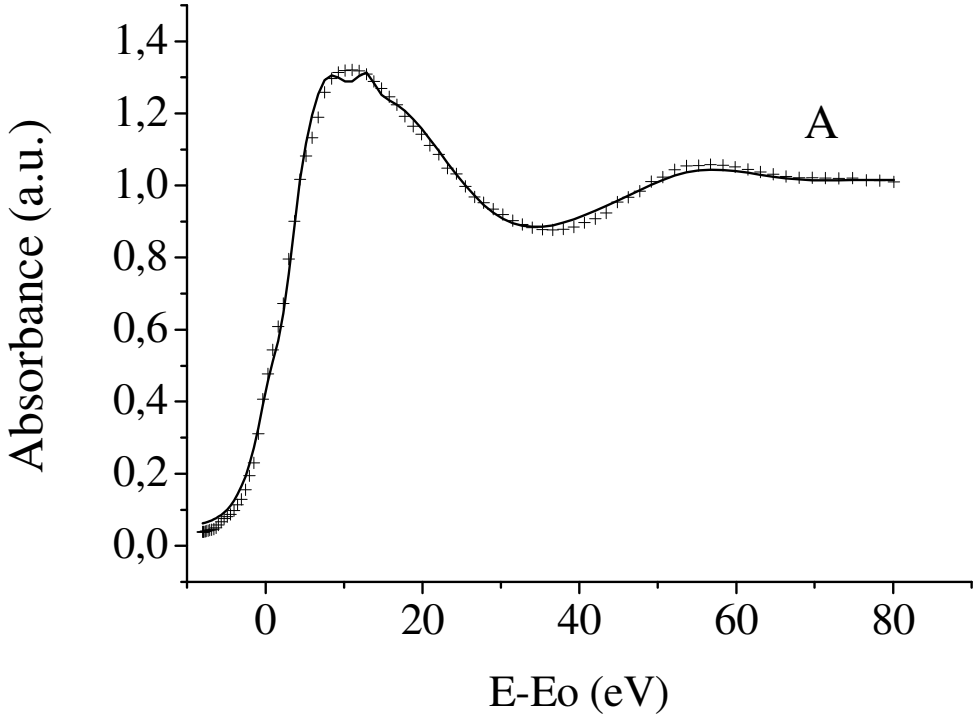
A B C D E F

Figure 2

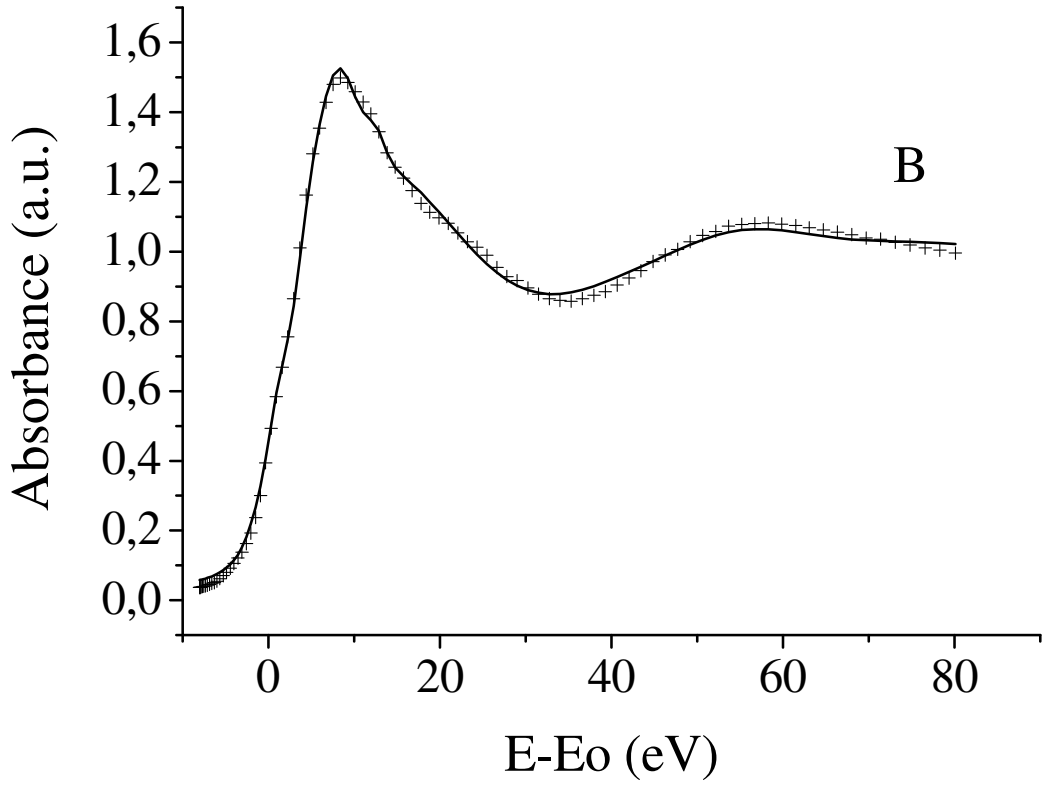
Figure_Marino_3



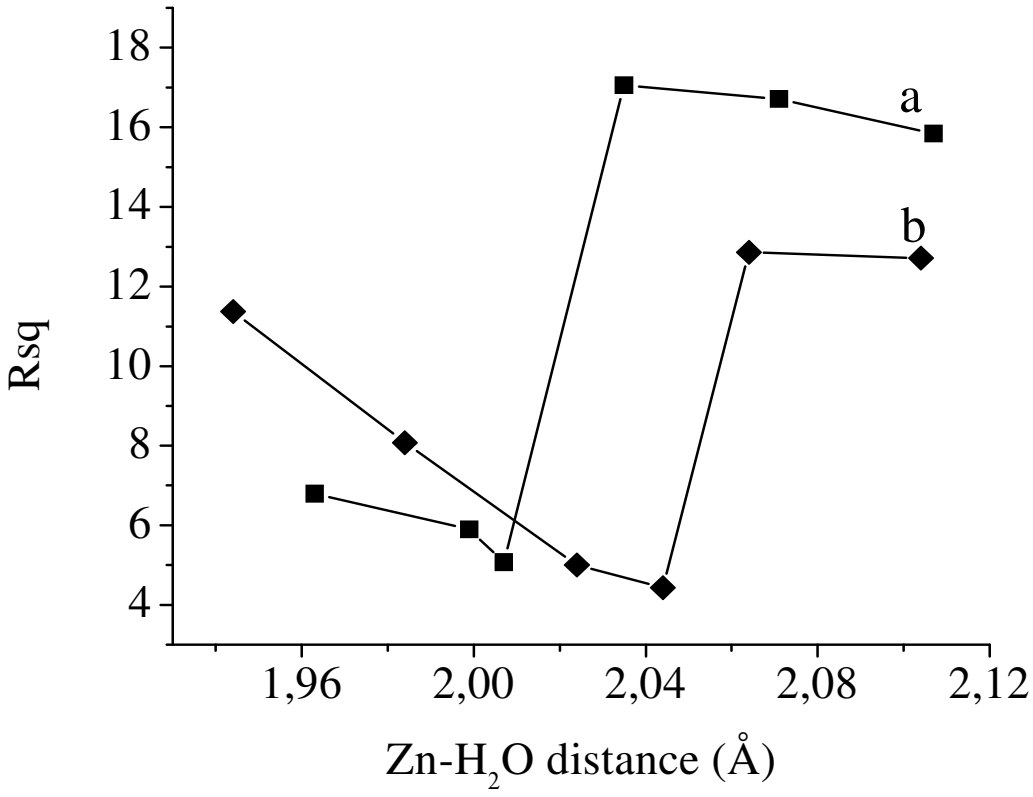
Figure_Marino_ 4A



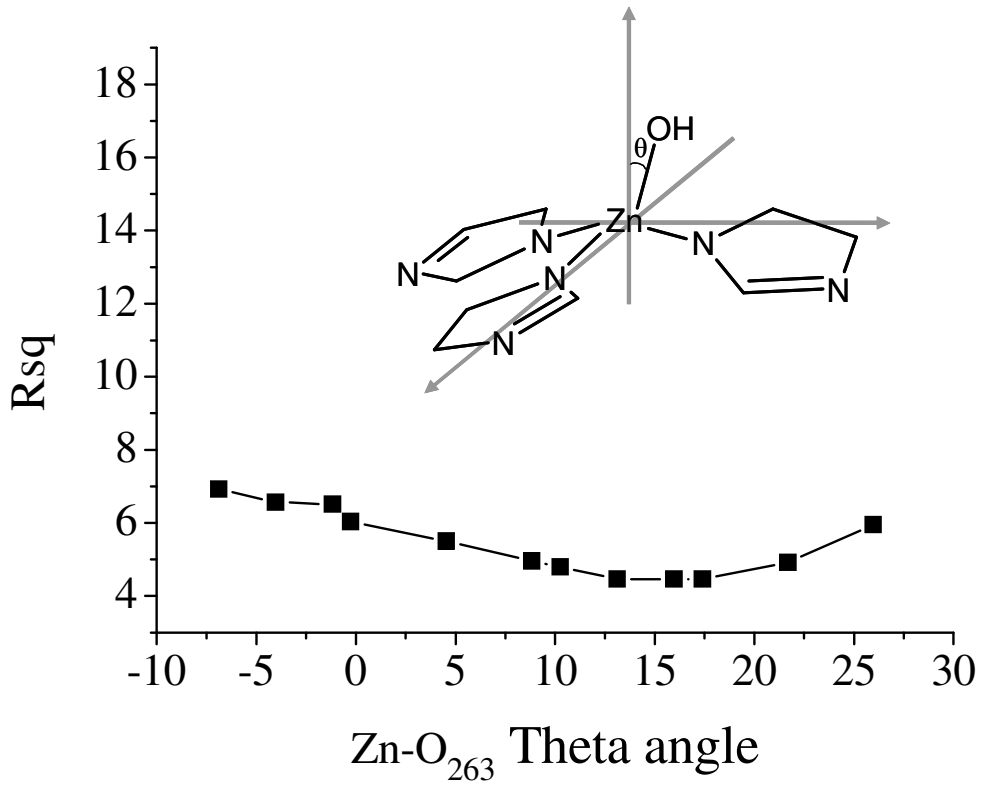
Figure_Marino_4B:



Figure_Marino_5A



Figure_Marino_5B



Figure_Marino_6: