Discovery of a new family of carbonic anhydrases in the malaria pathogen Plasmodium falciparum—The n-carbonic anhydrases

Discovery of a new family of carbonic anhydrases in the malaria pathogen *Plasmodium falciparum* – the η-carbonic anhydrases

Sonia Del Prete,[a] Daniela Vullo,[b] Gillian M. Fisher,[c] Katherine T. Andrews,[c] Sally-Ann Poulsen,[c] Clemente Capasso,[a*] and Claudiu T. Supuran[b,d]*

[a] Istituto di Bioscienze e Biorisorse (IBBR) – CNR, Via P. Castellino 111, 80131 Napoli, Italy.
[b]Università degli Studi di Firenze, Polo Scientifico, Laboratorio di Chimica Bioinorganica, Rm. 188, Via della Lastruccia 3, 50019 Sesto Fiorentino (Florence), Italy.
[c]Eskitis Institute for Drug Discovery, Griffith University, Nathan, Queensland 4111, Australia.
[d]Università degli Studi di Firenze, Dipartimento Neurofarba, Sezione di Scienze Farmaceutiche, Polo Scientifico, Sesto Fiorentino, Firenze, Italy.

•**Abstract**: The genome of the protozoan parasite *Plasmodium falciparum*, the causative agent of the most lethal type of human malaria, contains a single gene annotated as encoding a carbonic anhydrases (CAs, EC 4.2.1.1), thought to belong to the α-class. Here we demonstrate the kinetic properties of *Pf*CA for the CO_2 hydration reaction, as well as an inhibition study of this enzyme with inorganic and complex anions and other molecules known to interact with zinc proteins, including sulfamide, sulfamic acid, and phenylboronic/arsonic acids, detecting several low micromolar inhibitors. A closer examination of the sequence of this and the CA from other *Plasmodium spp*., as well as phylogenetic analysis, revealed that these protozoa encode for a yet undisclosed, new genetic family of CAs termed the η-CA class. The main features of the η-CAs are described in this report.

Keywords: carbonic anhydrase; η-CA-class enzyme; anion; inhibitor; *Plasmodium falciparum*

*Corresponding authors: Phone + 39-081-6132559; Fax +39-081-6132249; e-mail: c.capasso@ibp.cnr.it Phone: +39-055-4573005; Fax: +39-055-4573385; e-mail: claudiu.supuran@unifi.it.

Carbonic anhydrases (CAs, EC 4.2.1.1) are metalloenzymes present in all life kingdoms, with five genetically distinct families described to date in various organisms. [1-3] Most of them are zinc-containing enzymes, but Fe(II) may be present at the active site of the γ-CAs (described so far in Bacteria, Archaea and plants), whereas Cd(II) or Zn(II) ions seem to be equally effective for promoting catalysis in the ζ-CAs (diatoms encode for this class of CAs). [4-6] The metal ion is coordinated by three His residues (in the α-, γ- and δ-class enzymes) or by one His, and two Cys residues (in the β- and ζ-CAs), with the fourth ligand being a water molecule/hydroxide ion.[4-13] The main difference between these α-, γ- and δ-class enzyme families where three His ligands coordinate to zinc, is the spacing between the three His residues in the protein sequence. For example, in all α-CAs investigated so far the His ligands are at positions x, x+2 and x+25 (for example in the human isoform I, hCA I, these are His94, His96 and His119).[1-3] For the γ-CAs the positions of His residues coordinating the metal ion are always x, x + 36 and x + 41, respectively, whereas for the δ-CAs, the zinc ligands are positioned at residues x, x + 3 and x + 112, respectively.[1,4,5]

CAs belonging to various classes have been cloned, purified and characterized from many pathogenic organisms such as bacteria, fungi and worms [14-18] in order to investigate whether inhibition of such enzymes is crucial for their survival or pathogenesis. Indeed, for CAs from in most organisms it has been demonstrated that inhibitors belonging to the sulfonamide class (the most investigated CA inhibitors (CAIs)) [1-4] interfere with the growth, possessing interesting anti-infective properties.[19-22]

Few protozoan parasites have been investigated for the presence and druggability of CAs up until now. The causative agent of human malaria, *Plasmodium falciparum*, was one of the first one to be investigated.[23-25] A truncated form of the *P. falciparum* CA gene was cloned, expressed and purified in 2004 by Krungkrai's group,[23a] who showed that it is an active enzyme, possessing a good esterase activity with 4-nitrophenylacetate as a substrate, and also was inhibited by known sulfonamide-based CA inhibitors such as acetazolamide. The same authors concluded that the enzyme belongs to the α-class of CAs.[23a] Subsequent studies from Krungkrai's and our laboratories showed that different *Plasmodium spp*. encode CAs, all considered to belong to the α-class, and that primary sulfonamides inhibited *in vitro* and *in vivo* growth of Plasmodium parasite.[23b-25] The *P. falciparum* CA, the only Plasmodium CA investigated to date, has been denominated *Pf*CA.[23] Some benzenesulfonamide derivatives showed effective *in vitro* inhibition of the esterase activity of *Pf*CA, and also inhibited the *in vitro* growth of the parasite.[23b-25] Furthermore, some of these sulfonamides were effective as antimalarial agents in mice infected with *P. berghei*, an animal model of human malaria infection, with an efficacy similar to that of the clinically used drug chloroquine.[24] CAIs

were considered to possess ~~anti~~antimalarial activity because their inhibition of the first step of pyrimidine nucleotide biosynthesis in the protozoan parasite, i.e., the CA-mediated carbamoylphosphate biosynthetic pathway.[24] However, this has not been experimentally confirmed and *Pf*CA has never been investigated for its catalytic activity with CO_2 as substrate up until now.

More recently, an α-CA has also been cloned and characterized in another unicellular protozoan, *Trypanosoma cruzi*, the causative agent of Chagas disease.[26] This enzyme, denominated *Tc*CA, had a high catalytic activity for the CO_2 hydration reaction, although it is devoid of the His64 proton shuttle, an amino acid residue involved in the catalytic cycle of α-CAs.[1-3,26] The thiols, another class of CAIs, were the most potent *in vitro* inhibitors of *Tc*CA (K_Is of 21.1-79.0 nM) and some of them also inhibited the epimastigotes growth of two *T. cruzi* strains *in vivo*.[26] Thus, protozoan CA inhibition may be a valid strategy to control infection with protozoans causing diseases such as malaria and Chagas disease.[24,26]

Anions constitute another important class of CAIs.[27-30] As there are no anion inhibition studies of *Pf*CA, here we undertook such an investigation, including in our study the common metal-complexing anions, the halides and pseudohalides, as well as some complex anions and small molecules known to interact with these enzymes, such as sulfamide, sulfamic acid, phenylboronic acid, diethyldithiocarbamate, etc.[31] Up until this study, it should be noted that the catalytic activity of *Pf*CA had only been investigated for the esterase reaction catalyzed by the enzyme, with 4-nitrophenylacetate as a substrate.[24] Here we present the first kinetic study of the CO_2 hydrase properties of this enzyme, together with the anion inhibition data mentioned above. As discussed above, based on amino acid sequence comparisons, *Plasmodium* CA enzymes have until now been classified as members of the α-CA family. A closer look at the amino acid sequence and phylogeny of CAs present in *P. falciparum* and other *Plasmodium spp*., led us to the conclusion that *Pf*CA was erroneously assigned as being an α-class enzyme. Here we propose that *Plasmodia* encode for CAs belonging to a new genetic family that we call the η-CA class.

Figs. 1 -3 here

The *Pf*CA full length enzyme contains 600 amino acid residues (PlasmoDB: PF3D7_1140000), in contrast to hCA I and II which have 260 and 259 residues, respectively.[25] An alignment of the amino acid sequences of a truncated[23a] *Pf*CA sequence~~,~~ with the two human α-CAs, hCA I and II, is shown in Fig. 1, in order to identify some features of the protozoan enzyme previously assigned to the α-class.[23b] The ~~sequence of~~ truncated *Pf*CA sequence can be aligned to that of hCA I and II, but only if gaps are added to ~~in~~ the amino acid sequence regions in which the zinc coordinating histidines are located: six gaps must be placed in the *Pf*CA sequence (after residue 96 – the hCA I numbering system is used throughout the paper), and five gaps in the hCA

I/II sequences (after Leu118). However, we consider this ~~as~~ to be a "forced" alignment that has erroneously led to the assignment of *Pf*CA as belonging to the α-class. Indeed, the sequence alignment of Fig. 1 demonstrates that other features of α-CAs are not present in the *Pf*CA sequence, such as the proton shuttling residue in position 64 (His in most, but not all α-CAs, but Gln in *Pf*CA), or the Thr199 residue. The dyad Glu106 – Thr199 is conserved in all α-CAs investigated so far, being involved in the orientation of the substrate for the nucleophilic attack by the zinc hydroxide species of the enzyme.[1-4] It may be observed that the side chain of Lys present in position 199 in the *Pf*CA sequence in the forced alignment would be too bulky for assuring the correct hydrogen bonding network with Glu106, when CO_2 is bound in the hydrophobic pocket of α-CAs.[2] All these features of *Pf*CA which are not typical of an α-class enzyme, prompted us to compare the sequences of other *Plasmodium spp* CAs. An alignment of CA sequences available in the Plasmodium genome database, as well as *Plasmodium reichenowi* and *Plasmodium vinckei* sequences, is shown in Fig. 2.[32] These alignment data indicate that the CAs present in these protozoa are not α-CAs, and that they belong to a yet undescribed, new CA genetic family, for which we propose the name η-CA. The main features of this new enzyme class are as follows:

(i) the predicted metal ion coordinating residues are His94, His96 and His118 (again the hCA I numbering system is used for allowing us to better describe the differences between the α- and η-CAs). Thus, the metal ion coordination pattern (x, x+2, x+25 in α-CAs) is x, x+2, x+24, for the η-CAs. For the first four *Plasmodium* species shown in the alignment in Fig. 2, which include three *P. falciparum* lines and the chimpanzee malaria parasite *P. reichenowi* which is phylogenetically related to the human parasite *P. falciparum*, the same putative zinc coordination pattern is observed whereas for the other ~~last~~ four sequences, which are all murine malaria parasite species, His94 is replaced by an Asn, which cannot coordinate Zn(II). However, in the murine Plasmodia four sequences there are two other His residues, in position 91 and 92 (conserved in all these four CAs), which in principle could coordinate Zn(II). However, this may represent another genetic CA family (proposed to be denominated θ-CA class). Alternatively, similar to the CA-related proteins (CARPs),[33] these four proteins may lack one of the Zn(II) ligands, being catalytically inactive η-class enzymes. The hypothesis as to whether there is a seventh CA genetic family (considering the η-CA class as the sixth one) may be confirmed or refuted only when one of these enzymes is cloned and the presence or absence of CO_2 hydrase activity verified;

(ii) the η-CAs do not have His in position 64 (a major difference compared to α-CAs). In all sequences of Fig. 2, Phe is present in that position;

(iii) the Glu106 – Thr199 dyad, present in all α-CAs, is also absent in the η-CAs, with Ser being present in both these positions (Fig. 2);

(iv) the η-CAs discovered so far have a much longer amino acid sequence compared to the α-CAs, which typically are 250-280 amino acid residues long enzymes. It may be observed that the η-class enzymes have > 400 amino acid residues in their sequence. However it should be noted that on average Plasmodium gene sequences, excluding introns, are larger than those of other organisms.[34]

These conclusions were also reinforced when the phylogenetic analysis shown in Fig. 3 was performed. Enzymes from all CA families and all types of prokaryotes and eukaryotes have been included in the analysis (Fig. 3). It may be observed that the enzymes belonging to the δ-, α- and η-CA classes clustered together on the upper part of the tree. There are two main branches, one encompassing all the δ-CAs, and the second one with α- and η-CAs. The prokaryotic α-CAs are more distantly related to the eukaryotic enzymes as well as to the η-class enzymes. Among these, CAs from the human or chimpanzee infecting Plasmodium species that have the three His ligands (positions 94, 96 and 118) all cluster in the upper part of the branch, whereas the murine Plasmodium CAs without His94 cluster in the lower part of the phylogenetic tree. Thus, the η-CA class seems to be the most recent genetic family of such enzymes, being genetically closer to the eukaryotic α-CAs. The γ-CAs are also phylogenetically related to these three CA genetic families, their branch being the closest to the one that led to δ-, α- and η-CAs. It should be recalled that all these four genetic families have the metal ion coordinated by three His residues. On the other hand, on the lower branches of the tree are found the β- and ζ-CAs, which have the Zn(II) ion coordinated by two Cys and one His residue. The β-CAs from various organisms are, as expected, more related to each other than with the ζ-class enzymes which are on the lowest branch of the tree.

Tables 1 and 2 here

To begin to understand the enzymatic characteristics of *Pf*CA, we next measured the CO_2 hydrase activity of this recombinant enzyme by a stopped-flow assay.[28] Data in Table 1, where enzymes belonging to all six CA families are included, show that *Pf*CA has a good CO_2 hydrase activity at pH 7.5, with a k_{cat} of 1.4 x 10^5 s^{-1} and a k_{cat}/K_M of 5.4 x 10^6 M^{-1} x s^{-1}. Furthermore, this activity is inhibited by the clinically used sulfonamide acetazolamide (as for all other classes of CAs), with an inhibition constant of 170 nM. Thus, *Pf*CA is among the least effective catalysts shown in Table 1, being around 27.5 – 28 times less efficient compared to hCA II (α-class) or ZnCA1-R1 (ζ-class CA), which are among the enzymes with the highest turnover numbers described so far.[1,6] The η-class enzyme shows however a significant catalytic action, as its kinetic parameters are of the same order of magnitude as those of FbiCA (β-CA from the plant *Flaveria bidentis*) or TweCA (δ-CA from the diatom *Thalassiosira weissflogii*).

We next investigated the anion inhibition profile of recombinant *Pf*CA with simple and complex anions, as well as small molecules known to inhibit other CA families (Table 2). For

comparison, we have also included the anion inhibition data of the two human enzymes hCA I and II as well as the protozoan *T. cruzi* enzyme TcCA, all of which belong to the α-CAs and as reported earlier by us.[35] The following should be noted regarding the inhibition data of Table 2:

(i) Perchlorate, tetrafluoroborate and hydrogensulfite were not inhibitors of *Pf*CA, a behavior observed with most other α- and β-CAs investigated so far for perchlorate and tetrafluoroborate, but not for hydrogensulfite, which is an effective (submillimolar) TcCA inhibitor.[27]

(ii) A number of the investigated anions showed weak potency of inhibition against *Pf*CA, with inhibition constants ranging between 2.24 and 9.76 mM (Table 2). They include the halides (fluoride, chloride, bromide and iodide), azide, bicarbonate, nitrite, tellurate, perosmate, divanadate, tetraborate, perrhenate, perruthenate, peroxydisulfate, trithiocarbonate, triflate, sulfate, fluorosulfonate and iminodisulfonate. It is notable that for the halides the least efficient inhibitor was chloride, and the inhibition power increased with the increase atomic weight of the halogen, from chloride to iodide. Fluoride was slightly more inhibitory than bromide. However there is a very important difference in the inhibition profile by halides of the two protozoan enzymes, with TcCA-being much more sensitive to these anions than *Pf*CA.

(iii) More effective *Pf*CA inhibitors detected here were cyanate, thiocyanate, cyanide, bicarbonate, carbonate, nitrate, hydrogensulfide, stannate, selenite, selenocyanide and diethylditiocarbamate, which showed inhibition constants in the range of 0.55 – 0.95 mM (Table 2). Some interesting correlations between the nature of the anion and its inhibitory activity against *Pf*CA are now apparent. For example, for the pseudohalides, such as cyanate, thiocyanate, cyanide, are generally inhibitory against many metalloenzymes, and this is also the case for *Pf*CA.[1-6] It is well known that these anions readily complex cations in solution or within the enzyme active sites. In fact many of them efficiently inhibit α-CAs such as hCA I/II or TcCA (Table 2). A very interesting case is constituted by bicarbonate/carbonate, which are rather inhibitory against *Pf*CA (and TcCA) and much less so against the human isoforms hCA I and II. The opposite was observed for trithiocarbonate, which is an effective, micromolar inhibitor of hCA I, II and TcCA and a quite weak *Pf*CA inhibitor. Nitrate on the other hand is also a rather effective inhibitor of the protozoan enzymes but not of the human ones. Diethyldithiocarbamate, which contains the same zinc-binding function as trithiocarbonate, was on the other hand a much more effective *Pf*CA inhibitor.

(iv) The best *Pf*CA inhibitors detected here were sulfamide ($H_2NSO_2NH_2$), sulfamic acid (H_2NSO_3H), phenylboronic acid (Ph-$B(OH)_2$) and phenylarsonic acid (Ph-AsO_3H_2), which had inhibition constants in the low micromolar range (K_Is of 6-9 μM). All four compounds are much more effective η-CA inhibitors than α-CA inhibitors, for which millimolar affinities of these inhibitors were measured, Table 2.

(v) The inhibition profile of the protozoan enzyme *Pf*CA is very different from that of the human enzymes (hCA I and II) and also the *T. cruzi* TcCA, which is probably due to the fact that the active sites of the η- and α-class CAs are probably quite different.

As for most anions, it is probable that these inhibitors bind to the metal ion from the η-CA active site, in tetrahedral or trigonal-bipyramidal geometries of the Zn(II).

In conclusion, we demonstrated that *Pf*CA, an enzyme considered earlier to belong to the α-CA class, has significant CO_2 hydrase activity, with a k_{cat} of 1.4 x 10^5 s^{-1} and a k_{cat}/K_M of 5.4 x 10^6 M^{-1} x s^{-1}. The sequence of this and annotated CAs from other *Plasmodium spp*., as well as their phylogenetic analysis, allowed us to conclude that these protozoa encode for a yet undisclosed, new genetic family of CAs, which was called the η-CA class. The main features of the η-CAs are also delineated here. We also evaluated a series of inorganic simple/complex anions and other small molecules known to bind to metalloenzymes (sulfamide, sulfamic acid, phenylboronic/arsonic acids), for the inhibition of *Pf*CA, detecting several low micromolar inhibitors however with a differing inhibition profile to a-CAs.

Acknowledgments: This research was financed by an FP7 EU project (Dynano) to CTS, the Australian Research Council (FT0991213 to KTA, FT10100185 to S-AP) and the Australian National Health and Medical Research Council (PhD Scholarship to GF).

References and Notes

1. Alterio, V.; Di Fiore, A.; D'Ambrosio, K.; Supuran, C. T.; De Simone, G. *Chem. Rev.* **2012,** *112*, 4421.
2. Domsic, J. F.; Avvaru, B. S.; Kim, C. U.; Gruner, S. M.; Agbandje-McKenna, M.; Silverman, D. N.; McKenna, R. *J. Biol. Chem.* **2008,** *283*, 30766.
3. Supuran, C. T. *Nature Rev. Drug Discov.* **2008,** *7*, 168.
4. a) Supuran, C. T. *Front. Pharmacol.* **2011,** *2*, 34; b) Del Prete, S.; Vullo, D.; Scozzafava, A.; Capasso, C.; Supuran, C.T. *Bioorg Med. Chem.* **2014**, *22*, 531; c) Vullo, D.; Del Prete, S.; Osman, S.M.; De Luca, V.; Scozzafava, A.; AlOthman, Z.; Supuran, C.T.; Capasso, C. *Bioorg Med. Chem. Lett.* **2014**, *24*, 275.
5. a) Ferry, J.F. *Biochim. Biophys. Acta* **2010**, *1804*, 374; b) Smith, K.S.; Jakubzick, C.; Whittam, T.S.; Ferry, J.G. *Proc Natl Acad Sci U S A.* **1999**, *96*, 15184; c) Zimmerman, S.A.; Tomb, J.F.; Ferry, J.G. *J. Bacteriol*. **2010**, *192*, 1353; d) Zimmerman, S.A.; Ferry, J.G.; Supuran, C.T. *Curr. Top. Med. Chem*. **2007**, *7*, 901 ; e) Tripp, B. C.; Bell, C. B., 3rd; Cruz, F.; Krebs, C.; Ferry, J. G. *J. Biol. Chem.* **2004,** *279*, 6683.
6. a) Xu, Y.; Feng, L.; Jeffrey, P. D.; Shi, Y.; Morel, F. M. *Nature* **2008,** *452*, 56; b) Alterio, V.; Langella, E.; Viparelli, F.; Vullo, D.; Ascione, G.; Dathan, N.A.; Morel, F.M.; Supuran, C.T.; De Simone, G.; Monti, S.M. *Biochimie* **2012**, *94*, 1232; c) Viparelli, F.; Monti, S.M.; De Simone, G.; Innocenti, A.; Scozzafava, A.; Xu, Y.; Morel, F.M.; Supuran, C.T. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4745.
7. a) Alterio, V.; Hilvo, M.; Di Fiore, A.; Supuran, C. T.; Pan, P.; Parkkila, S.; Scaloni, A.; Pastorek, J.; Pastorekova, S.; Pedone, C.; Scozzafava, A.; Monti, S. M.; De Simone, G. *Proc. Natl. Acad. Sci.* USA **2009,** *106*, 16233; b) Vullo, D.; De Luca, V.; Scozzafava, A.; Carginale, V.; Rossi, M.; Supuran, C.T.; Capasso, C. *Bioorg. Med. Chem. Lett*. **2012**, *22*, 6324.
8. a) Neri, D.; Supuran, C. T. *Nature Rev. Drug Discov.* **2011,** *10*, 767; b) Supuran, C. T. *J. Enzyme Inhib. Med. Chem*. **2012**, *27*, 759.
9. Schlicker, C.; Hall, R. A.; Vullo, D.; Middelhaufe, S.; Gertz, M.; Supuran, C. T.; Muhlschlegel, F. A.; Steegborn, C. *J. Mol. Biol.* **2009,** *385*, 1207-1214.
10. a) Monti, S.M.; De Simone, G.; Dathan, N.A.; Ludwig, M.; Vullo, D.; Scozzafava, A.; Capasso, C.; Supuran, C.T. *Bioorg. Med. Chem. Lett*. **2013**, *23*, 1626; b) Suarez Covarrubias, A.; Larsson, A. M.; Hogbom, M.; Lindberg, J.; Bergfors, T.; Bjorkelid, C.; Mowbray, S. L.; Unge, T.; Jones, T. A. *J. Biol. Chem.* **2005,** *280*, 18782.
11. a) Tu, C.; Tripp, B. C.; Ferry, J. G.; Silverman, D. N. *J. Am. Chem. Soc.* **2001,** *123*, 5861; b) Del Prete, S.; Vullo, D.; De Luca, V.; Carginale, V.; Scozzafava, A.; Supuran, C.T.; Capasso, *Bioorg. Med. Chem. Lett*. **2013**, *23*, 4067.
12. Covarrubias, A. S.; Bergfors, T.; Jones, T. A.; Hogbom, M. *J. Biol. Chem.* **2006,** *281*, 4993.
13. Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2010,** *20*, 3467.
14. Burghout, P.; Vullo, D.; Scozzafava, A.; Hermans, P. W.; Supuran, C. T. *Med. Chem.* **2011,** *19*, 243.
15. Joseph, P.; Ouahrani-Bettache, S.; Montero, J. L.; Nishimori, I.; Minakuchi, T.; Vullo, D.; Scozzafava, A.; Winum, J. Y.; Kohler, S.; Supuran, C. T. *Med. Chem.* **2011,** *19*, 1172.
16. Joseph, P.; Turtaut, F.; Ouahrani-Bettache, S.; Montero, J. L.; Nishimori, I.; Minakuchi, T.; Vullo, D.; Scozzafava, A.; Kohler, S.; Winum, J. Y.; Supuran, C. T. *J. Med. Chem.* **2010,** *53*, 2277.
17. Nishimori, I.; Minakuchi, T.; Morimoto, K.; Sano, S.; Onishi, S.; Takeuchi, H.; Vullo, D.; Scozzafava, A.; Supuran, C. T. *J. Med. Chem.* **2006,** *49*, 2117.
18. Nishimori, I.; Minakuchi, T.; Vullo, D.; Scozzafava, A.; Innocenti, A.; Supuran, C. T. *J. Med. Chem.* **2009,** *52*, 3116.
19. a) Supuran, C. T. *Curr. Pharm. Des.* **2010,** *16*, 3233; b) Carta, F.; Aggarwal, M.; Maresca, A.; Scozzafava, A.; McKenna, R.; Masini, E.; Supuran, C. T. *J. Med. Chem.* **2012,** *55*, 1721.

20. a) Supuran, C. T.; Scozzafava, A.; Casini, A. *Med. Res. Rev.* **2003,** *23*, 146; b) Del Prete, S.; Isik, S.; Vullo, D.; De Luca, V.; Carginale, V.; Scozzafava, A.; Supuran, C. T.; Capasso, C. *J. Med. Chem.* **2012,** *55*, 10742; c) Capasso, C.; Supuran, C.T. *Expert Opin. Ther. Pat.* **2013**, *23*, 693.
21. Vullo, D.; Nishimori, I.; Scozzafava, A.; Kohler, S.; Winum, J. Y.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2010,** *20*, 2178.
22. Winum, J. Y.; Kohler, S.; Supuran, C. T. *Curr. Pharm. Des.* **2010,** *16*, 3310-3316.
23. a) Reungprapavut, S.; Krungkrai, S. R.; Krungkrai, J. *J. Enzyme Inhib.* **2004**, *19*, 249; b) Krungkrai, J.; Krungkrai, S. R.; Supuran, C. T. *Curr. Top. Med. Chem.* **2007,** *7*, 909.
24. Krungkrai, J.; Krungkrai, S. R.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2008,** *18*, 5466.
25. a) Krungkrai, J.; Supuran, C. T. *Curr. Pharm. Des.* **2008,** *14*, 631. b) Andrews, K.T.; Fisher, G.M.; Sumanadasa, S.D.M.; Sinner-Adams, T.; Moeker, J.; Lopez, M.; Poulsen, S.-A. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 455; c) Fisher, G.M.; Tanpure, R.P.; Douchez, A.; Andrews, K.T.; Poulsen, S.-A. *Chem. Biol. Drug Des.* **2014**, doi: 10.1111/cbdd.12335
26. Pan, P.; Vermelho, A. B.; Capaci Rodrigues, G.; Scozzafava, A.; Tolvanen, M. E.; Parkkila, S.; Capasso, C.; Supuran, C. T. *J. Med. Chem.* **2013**, *56*, 1761.
27. De Simone, G.; Supuran, C. T. *J. Inorg. Biochem.* **2012,** *111*, 117.
28. Khalifah, R. G. *J. Biol. Chem.* **1971,** *246*, 2561. An Applied Photophysics stopped-flow instrument has been used for assaying the CA catalyzed CO_2 hydration activity. Phenol red (at a concentration of 0.2 mM) has been used as indicator, working at the absorbance maximum of 557 nm, with 10 – 20 mM Hepes (pH 7.5) as buffer, and 20 mM Na_2SO_4 or 20 mM $NaBF_4$ for maintaining constant the ionic strength, following the initial rates of the CA-catalyzed CO_2 hydration reaction for a period of 10-100 s. The CO_2 concentrations ranged from 1.7 to 17 mM for the determination of the kinetic parameters and inhibition constants. For each inhibitor at least six traces of the initial 5-10% of the reaction have been used for determining the initial velocity. The uncatalyzed rates were determined in the same manner and subtracted from the total observed rates. Stock solutions of inhibitor (10 mM) were prepared in distilled-deionized water and dilutions up to 0.01 μM were done thereafter with distilled-deionized water. Inhibitor and enzyme solutions were preincubated together for 15 min at room temperature prior to assay, in order to allow for the formation of the E-I complex. The inhibition constants were obtained by non-linear least-squares methods using the Cheng-Prusoff equation whereas the kinetic parameters for the uninhibited enzymes from Lineweaver-Burk plots, as reported earlier,[29-31] and represent the mean from at least three different determinations.
29. a) Innocenti, A.; Scozzafava, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2009,** *19*, 1855; b) Innocenti, A.; Scozzafava, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2010,** *20*, 1548.
30. a) Kolayli, S.; Karahalil, F.; Sahin, H.; Dincer, B.; Supuran, C. T. *J. Enzyme Inhib. Med. Chem.* **2011,** *26*, 895; b)Maresca, A.; Scozzafava, A.; Kohler, S.; Winum, J. Y.; Supuran, C. T. *J. Inorg. Biochem.* **2012,** *110*, 36; c)Ozensoy, O.; Arslan, M.; Supuran, C. T. *J. Enzyme Inhib. Med. Chem.* **2011,** *26*, 749; d) Temperini, C.; Scozzafava, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2010,** *20*, 474.
31. a) Carta, F.; Aggarwal, M.; Maresca, A.; Scozzafava, A.; McKenna, R.; Supuran, C. T. *Chem. Commun.* **2012,** *48*, 1868; b) Maresca, A.; Carta, F.; Vullo, D.; Supuran, C. T. *J. Enzyme Inhib. Med. Chem.* **2013,** *28*, 407.
32. Aurrecoechea C, Brestelli J, Brunk BP, Dommer J, Fischer S, Gajria B, Gao X, Gingle A, Grant G, Harb OS, Heiges M, Innamorato F, Iodice J, Kissinger JC, Kraemer E, Li W, Miller JA, Nayak V, Pennington C, Pinney DF, Roos DS, Ross C, Stoeckert CJ Jr, Treatman C, Wang H. *Nucleic Acids Res.* 2009, 37(Database issue):D539-43. Epub 2008 Oct 28.
33. Nishimori, I.; Vullo, D.; Minakuchi, T.; Scozzzafava, A.; Capasso, C.; Supuran, C.T. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 256.

34. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B. *Nature* **2002** 3;419(6906), 498.
35. Pan, P.; Vermelho, A.B.; Scozzafava, A.; Parkkila, S.; Capasso, C.; Supuran, C.T. *Bioorg. Med. Chem*. **2013**, *21*, 4472.

Table 1. Kinetic parameters for the CO_2 hydration reaction catalysed by various CAs belonging to the various families.[1-7] The α-class CAs were the human cytosolic isozymes hCA I and II and the bacterial SazCA (from *Sulfurihydrogenibium azorense*).[5d,6] The β-class includes the fungal enzyme Can2 from *Cryptococcus neoformans* [4b] and FbiCA1 from the plant *Flaveria bidentis*.[18] The γ-class enzyme was PgiCA from the anaerobic bacterium *Porphyromonas gingivalis*,[19] whereas the δ and ζ-class enzymes [2,9] (the last with zinc and cadmium at the active site) were from the diatom *Thalassiosira weissflogii.* Inhibition data with the clinically used sulfonamide acetazolamide (5-acetamido-1,3,4-thiadiazole-2-sulfonamide) are also provided. All data were obtained in the author's laboratories.

Isozyme	Class	Organism	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}.s^{-1}$)	K_I (acetazolamide) (nM)
hCA I[a]	α	human	$2.0x10^5$	$5.0x10^7$	250
hCA II[a]	α	human	$1.4x10^6$	$1.5x10^8$	12
SazCA[b]	α	bacterium	$4.4x10^6$	$3.5x10^8$	0.9
Can2[c]	β	fungus	$3.9x10^5$	$4.3x10^7$	10.5
FbiCA1[d]	β	plant	$1.2x10^5$	$7.5x10^6$	27
PgiCA[e]	γ	bacterium	$4.1x10^5$	$5.4x10^7$	324
CdCA1-R1[f]	ζ	diatom	$1.5x10^6$	$1.4 x10^8$	82
ZnCA1-R1[f]	ζ	diatom	$1.4x10^6$	$1.6 x10^8$	58
TweCA[g]	δ	diatom	$1.3x10^5$	$3.3 x10^7$	83
*Pf*CA[h]	η	protozoa	$1.4 x10^5$	$5.4x10^6$	170

[a] Data from ref.[1]

[b] Data from ref.[7b]

[c] Data from ref.[9]

[d] Data from ref.[10a]

[e] Data from ref.[11b]

[f] Data from ref.[6]

[g] Data from ref.[4b]

[h]This work.

Table 2: Inhibition constants of anion inhibitors against α-CAs from mammals (hCA I, and II, human isoforms, and the protozoan enzyme from *T. cruzi*, TcCA), and the η-CA *Pf*CA from *P. falciparum*, for the CO_2 hydration reaction, at 20 °C and pH 7.5.[28]

Inhibitor[§]	K_I [mM][#]			
	hCA I[a]	hCA II[a]	TcCA[b]	*Pf*CA[c]
F^-	> 300	>300	0.90	5.78
Cl^-	6	200	0.81	9.76
Br^-	4	63	0.73	6.05
I^-	0.3	26	0.044	2.74
CNO^-	0.0007	0.03	0.080	0.81
SCN^-	0.2	1.60	0.084	0.95
CN^-	0.0005	0.02	1.01	0.76
N_3^-	0.0012	1.51	1.12	5.60
HCO_3^-	12	85	0.58	0.78
CO_3^{2-}	15	73	0.69	0.90
NO_3^-	7	35	0.77	0.66
NO_2^-	8.4	63	0.70	2.46
HS^-	0.0006	0.04	0.078	0.68
HSO_3^-	18	89	0.91	>100
SnO_3^{2-}	0.57	0.83	0.72	0.73
SeO_4^{2-}	118	112	0.66	0.90
TeO_4^{2-}	0.66	0.92	0.71	3.10
OsO_5^{2-}	0.92	0.95	0.80	4.93
$P_2O_7^{4-}$	25.77	48.50	16.9	9.53
$V_2O_7^{4-}$	0.54	0.57	0.68	5.70
$B_4O_7^{2-}$	0.64	0.95	24.3	2.24
ReO_4^-	0.11	0.75	0.65	3.47
RuO_4^-	0.101	0.69	0.65	4.51
$S_2O_8^{2-}$	0.107	0.084	0.84	3.84
$SeCN^-$	0.085	0.086	0.78	0.87
CS_3^{2-}	0.0087	0.0088	0.093	4.36
$Et_2NCS_2^-$	0.00079	0.0031	0.005	0.55
$CF_3SO_3^-$	nt	nt	nt	7.75

(Table 2, continued)

SO_4^{2-}	63	>200	6.9	9.2
ClO_4^-	>200	>200	>100	>100
BF_4^-	>200	>200	>100	>100
FSO_3^-	0.79	0.46	15.8	9.45
$NH(SO_3)_2^{2-}$	0.31	0.76	7.1	5.96
$H_2NSO_2NH_2$	0.31	1.13	0.12	0.008
H_2NSO_3H	0.021	0.39	10.6	0.009
Ph-$B(OH)_2$	58.6	23.1	0.86	0.007
Ph-AsO_3H_2	31.7	49.2	0.62	0.006

[§]As sodium salt, except sulfamide, phenylboronic acid and phenylarsonic acid; [#]Errors were in the range of 3-5 % of the reported values, from three different assays, by a CO_2 hydrase assay method; [a] From ref. [27]; [b] From ref. [33]; [c] This work.

Table 3

CA class, organism (belonging to the Bacteria, Archaea and Eukarya kingdoms), accession numbers and cryptonyms of the amino acid sequences used to construct the alignment and the phylogenetic tree.

CA class	Organism	Accession number	Cryptonim
Alpha (α)	*Helicobacter pylori J99*	NP_223829..1	HpylCA_alpha
	Homo sapiens, isoform II	AAH11949.1	HumCAII_alpha
	Homo sapiens, isoform I	NP_001158302.1	HumCAI_alpha
	Sulfurihydrogenibium yellowstonense YO3AOP1	ACD66216.1	SspCA_alpha
	Streptococcus salivarius	EIC81445.1	SsalCA_alpha
	Neisseria gonorrhoeae	CAA72038.1	NgonCA_alpha
Beta (β)	*Schizosaccharomyces pombe*	CAA21790	SpoCA_beta
	Brucella suis 1330	NP_699962.1	BsuCA_beta
	Burkholderia thailandensis	ZP_02386321	BthCA_beta
	Coccomyxa sp.	AAC33484.1	CspCA_beta
	Chlamydomonas reinhardtii	XP_001699151.1	CreCA_beta
	Acinetobacter baumannii	YP_002326524	AbaCA_beta
	Porphyromonas gingivalis	YP_001929649.1	PgiCA_beta
	Myroides injenensis	ZP_10784819	MinCA_beta
	Zea mays	NP_001147846.1	ZmaCA_beta
	Vigna radiata	AAD27876	VraCA_beta
	Flaveria bidentis, isoform I	AAA86939.2	FbiCA_beta
	Arabidopsis thaliana	AAA50156	AthCA_beta
	Helicobacter pylori	BAF34127.1	HpyCA_beta
	Legionella pneumophila	YP_003619232	LpnCA_beta
	Escherichia coli	ACI70660	EcoCa_beta
	Methanobacterium thermoautotrophicum	GI:13786688	Cab_beta
	Saccharomyces cerevisiae	GAA26059	SceCA_beta
	Dekkera bruxellensis	EIF49256	DbrCA_beta
Gamma (γ)	*Pseudomonas sp.*	ZP_10427314.1	PseCA_gamma
	Burkholderia gladioli	YP_004359911.1	BglCA_gamma
	Methanosarcina thermophila	ACQ57353.1	CAM_gamma
	Chlamydomonas reinhardtii	XP_001703237.1	CreCA_gamma
	Arabidopsis thaliana	NP_564091.1	AthCA_gamma
	Porphyromonas gingivalis	YP_001929649.1	PgiCA_gamma
Delta (δ)	*Thalassiosira weissflogii*	AAV39532.1	TweCA_delta
	Thalassiosira pseudonana	XP_002287620.1	TpsCA_delta
	Emiliania huxleyi	ABG37687.1	EhuCA_delta
	Bathycoccus prasinos	CCO20234.1	Bpr_delta
	Lingulodinium polyedrum	ABS87870.1	LpoCA_delta

Zeta (ζ)	*Thalassiosira weissflogii*	AAX08632.1	TweCA_zeta,CDCA1
	Micromonas pusilla	XP_003063214.1	MpuCA_zeta
	Thalassiosira oceanica	EJK51395.1	TocCA_zeta
	Thalassiosira pseudonana	XP_002295227.1	TpseCA_zeta
	Micromonas sp.	XP_002504722.1	MicCA_zeta
Eta (η)	*Plasmodium falciparum* 3D7	PF3D7_1140000	Pf3D7_eta
	Plasmodium reichenowi	CDO65199.1	Preich_eta
	Plasmodium yoelii yoelii 17X	PY17X_0910400	Py17X_eta
	Plasmodium yoelii yoelii 17XNL	PY00744	Py17XNL_eta
	Plasmodium vinckei vinckei	KEG02328.1	Pvvinc_eta
	Plasmodium vinckei petteri	EUD73019.1	Pvpet_eta
	Plasmodium berghei ANKA	PBANKA_090900	PbANKA_eta
	Plasmodium chabaudi chabaudi	PCHAS_071030	Pcchab_eta
	P. falciparum CA(residues 211 to 445[24, 25]	-	PfCA_eta

Fig. 1: “Forced” alignment of the amino acid sequences of the truncated *P. falciparum* CA (*Pf*CA), which contains three additional C-terminal aminno acids (GNK) not present in the native sequence, with the human α-CA isoforms I and II (hCA I and hCA II). In order to align the Zn(II) ion ligands (in red) of the three enzymes, six “missing” residues have to be inserted in the protozoan enzyme sequence, and five in the mammalian enzymes sequences. It may be seen however that the other residues crucial for the catalytic mechanism of the α-CAs (i.e., the proton shuttle residue His64 shown in blue) and the gatekeeper residues (Glu106 and Thr199, in orange) are not conserved in the protozoan enzyme (except for Glu106). The hCA I numbering system has been used.[29-34]. Amino acid sequence cryptonym is indicated in Table 3.

```
Pf3D7_eta      MKLLYLLYPILLFYNVNVFINYKKSRLMLEMIDKYNTHFVQTTKPYYEFNVTNLTNSKKK
PfIT_eta       MKLLYLLYPILLFYNVNVFINYKKSRLMLEMIDKYNTHFVQTTKPYYEFNVTNLTNSKKK
Preich_eta     MKLLYLLYPILLFYNVNVFINYKKSRLMLEMIEKYNTHFVQTTKPYYEFNVTNLSNSKKK
Py17X_eta      -MKHIIFLSIVLCFCDNVMYNNYVERMLFELPNNITDDLNSDPIVEYKIKEKKNDNIDIN
Py17XNL_eta    -MKHIIFLSIVLCFCDNVMYNNYVERMLFELPNNITDDLNSDPIVEYKIKEKKNDNIDIN
Pvvinc_eta     -MKHIIFLSIVFCFCDNVVYNNYVGRILFEFPDNIIHDLSSGPIVEYEIKEHKDDNPDIN
Pvpet_eta      -MKHIIFLSIVFCFCDNVVYNNYVARILFELPDNIIYDLNSSPIVEYEIKEHKDDNPDIN
PbANKA_eta     -MKHIIFLSIVLCFCDNVMYNNYVERILFELPNNITDDLNSEPMVEYEVKEKKDDNIDIN
Pcchab_eta     -MKHIIFLSIVFCFCDHVVSNNYVGRILFELPDNIMDDLSSGPIVEYEIKEHKDDNPDIN
                    :: .*:: :  :*. *    *:::*: ::   .: . .   *:.:  :  * . :

KKKKKRENHLIGSGENMQKKDEKNIKDFHIN----------------------------
KKKKKRENHLIGSGENMQKKDEKNIKDFHIN----------------------------
KKKKKRGNHLIGSGENMQKKDEKHIKDFHIN----------------------------
KDVRHWDIEINEHKDDPNIQRNIEWHDNNDGNGNNSGNNSGNNNGNNSGNNNDDNNDNDY
KDVRHWDIEINEHKDDPNIQRNIEWHDNNDGNGNNSGNNSGNNNGNNSGNNNDDNNDNDY
KEVNHWNIEINEHKDNPNIQRNNEGNDDSNNN---------------------------
KEVRHWNIEINEHKDNPNIQRSNEGNDDSNNN---------------------------
KDVRHWNIEINEHKDNPNIQRNPEGNDNNHNNENND-----------------------
KDARHWNIEINGDKDNPNIQRSNEGNDDNNDN----------------------------
*. .:   .:    :: : : . . :*   .

---------------------DYEIDGKTIHNKENKDSFKMNKNKLNDNEELFYMDNILS
---------------------DYEIDGKTIHNKENKDSFKMNKNKLNDNEELFYMDNILS
---------------------DYEIVGKTIHNKENKDAFKMNKNKLNDNEELFYMDSILS
GNDKNWEYNSNYNDEEFERQNENERDEFSLKNEVEKNSEERKERAFDESNEYADFENMND
GNDKNWEYNSNYNDEEFERQNENERDEFSLKNEVEKNSEERKERAFDESNEYADFENMND
-----WQYHSNYNDEQSESQNENERNEFSLKNEMEKKPEERKDTQFDKYNEYDDFENMN-
-----RQYHSNYNDEQSESQNENERNEFSLKNEAEKNTEERKDTQFDKYNEYDNFENMN-
----NWEYHSNYNDEKFESQNENERNGFSLKNEVEKNPEERKDTPFDEYNEYANFENMN-
-----WQYHSNYNDKQSESQNENERNEFSLKNEMEKKPEEIKDTQFDKYNEYDDFENVD-
                     : *    :::*: :*.. : :.  ::. :*   ::.:

------YKPNKKKLFTYSFSENEGNSEKEETLYNFKNMKNIN------------------
------YKPNKKKLFTYSFSENEGNSEKEETLYNFKNMKNIN------------------
------YKPNKKKLFTYSFSENEGNSEKEETLYNFENRKNIN------------------
LENMNNIEKEKKKYFEDMQSKYVEDNTSDGNKEYMGEMKNQQNEYEQNEHQKNEYEQNEH
LENMNNIEKEKKKYFEDMQSKYVEDNTSDGNKEYMGEMKNQQNEYEQNEHQKNEYEQNEH
----NNFEENKKKSFEAMQSEDMEDKKGEENKEYAGWAEDKKREDNRDYIDRMDDKNSAG
----NNFEENKKKHFEAMQSEDMEDKKRMDNKEYVDWIEDKQREDNRDYANGMEDKSSAG
----NNFEKNKKKYFEDMQSTYMEDKKNVDNKEYMDEVKNKKIEYQKN------------
----NNFEENKRKHFEDMQSEDMEDKKRADNKEYADWIEDKKRSDNRGYTGGMEDKNSAG
       : :*:* *    *    :.    .       :: :

------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
QQNEYEQNEHQQNEYEQ------------------NEHQQNEYEQN--------------
QQNEYEQNEHQQNEYEQ------------------NEHQQNEYEQNEHQQ----------
NREYIDGMEYKKEAENKDYIGWVEDKNSAGNRDYIDGMEDKKREENKDYAGWVEDKNSAG
NREYIDGMEDKKGEENK---------------DYIDGMEDKKGEENKDYIGWVEDKNIAG
------------------------------------------------------------
NRSYIDGVEDKNSASNK---------------EYISWMDDKNSASNKGYAGWTDDKNGAS


------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
NREYIDRAEDKKGEENKEYGGWVKDKNNRAEDKKGEENKEYGGWVKDKNSAENIEYIDGM
NREYIG------------------------------------------------------GM
------------------------------------------------------------
N-------------------------------------------------------------


-------------------------SVQNNINKTFLYNKLKNVDYYEHGYNWDIGQCKTG
-------------------------SVQNNINKTFLYNKLKNVDYYEHGYNWDIGQCKTG
-------------------------SVQNNINKTFLYNKLKNVDYYEHGYNWDIGQCKTG
```

```
--------------EHQQYEQNEQNEEGNIKNGMIQNNENLSFNYAKHGMDWNVGICKNG
--------NEYEQNEHQQYEQNEQNEEGNIKNGMIQNNENLSFNYAKHGMDWNVGICKNG
EYKKEAENKEYGGWMGDKQIGSHRKKENNIKSDTTEHNDNLSFEYSKQGMDWAAGICKNG
EYKKEAENKEYAGWMNDKQIEDHRNEENNIKSDTTQHNDNLSFDYSKQGMDWAAGVCKNG
-------------------------EENNIKNGIIQYNDNLSFDYSKHGMDWNVGICKNG
--------KEYAGWRDDKQIEDHRNEENNTKSDTTQDNDNLSFDYSKQGVNWDVGVCKNG
                         .  *  .     *.  ..:* ::* :*  * **.*

KYQSPVDLPMKDLKERELKNISDVYLNLFDD-DNYAWNNYNKPWMKGDFFYYYEYFIKKI
KYQSPVDLPMKDLKERELKNISDVYLNLFDD-DNYAWNNYNKPWMKGDFFYYYEYFIKKI
KYQSPVDLPMKDLKERELKNISDVYLNLFDD-DNYAWNNYNKPWMKGDFFYYYECFIKKI
KYQSPVDLHMHTLKERELKNLSDFYLNAFYDNDEYSWNNFNRPWFKGDIFYYYENLINKI
KYQSPVDLHMHTLKERELKNLSDFYLNAFYDNDEYSWNNFNRPWFKGDIFYYYENLINKI
KYQSPVDLHMHTLKERELKNLSDFYLNAFYDNDENSWNNYNRPWFKGDIFYYYENLVNKI
KYQSPVDLHMHTLKERELKNLSDFYLNAFYDNDEYSWNNYNRPWFKGDIFYYYENLINKI
KYQSPVDLHMHTLKERELKNLSDFYLNAFYDNDEYSWNNYNKPWFKGDIFYYYENLINKI
KYQSPVDLHMHTLKERELKNLSDFYLNAFYDNDEYSWNNYNRPWFKGDIFYYYENLINKI
******** *: ********:**.*** * * *: :***:*:**:***:***** :::**

         64                                94 96
VINRQNNIFQIKAARDGIIPFGVLFTTEQPAMFYADQIHFHAPSEHTFQGSGNRREIEMQ
VINRQNNIFQIKAARDGIIPFGVLFTTEQPAMFYADQIHFHAPSEHTFQGSGNRREIEMQ
VINRQNNIFQIKAARDGIIPFGVLFTTEEPAMFYADQIHFHAPSEHTFQGSGNRREIEMQ
IINRQNNMFKIKASNNEIIPFGVLFTTDEPTIFYSHHINFHSPSEHTFEGSGNRRHIEMQ
IINRQNNMFKIKASNNEIIPFGVLFTTDEPTIFYSHHINFHSPSEHTFEGSGNRRHIEMQ
IINRQNNMFKIKASNNDIIPFGVLFTTDEPAIFYSHHINFHSPSEHTFEGSGNRRHIEMQ
IINRQNNMFKIKATNNEIIPFGVLFTTDEPAIFYSHHINFHSPSEHTFEGSGNRRHIEMQ
IINRQNNMFKIKASNNEIIPFGVLFTTDEPAIFYSHHINFHSPSEHTFEGSGNRRHIEMQ
IINRQNNMFKIKASNNEIIPFGVLFTTDEPAIFYSHHINFHSPSEHTFEGSGNRRHIEMQ
:******:*:***:.: **********::*::**:.:*:**:******:******.****

  118
IFHSTNYFYDIQDDKSKYKKKYGLHIYNNLKKNSKETSKKDSSRYHSYLMSFLMNSLSNE
IFHSTNYFYDIQDDKSKYKKKYGLHIYNNLKKNSKETSKKDSSRYHSYLMSFLMNSLSNE
IFHSTNYFYDIQDDKSKYKKKYGLHIYNNLKKNSKETSKKDSSRYHSYLMSYLMNSLSNE
IYHSTNEIYDYDENK----------WNGVFEKKNYKKKNNETNIQHSYILTFLMNSLSNP
IYHSTNEIYDYDENK----------WNGVFEKKNYKKKNNETNIQHSYILTFLMNSLSNP
IYHSTNEIYDYDESK----------WNGIFGKKKNQKKNNETNIKHSYILTFLRNSLSNP
IYHSTNEIYDYDESK----------WNGIFGKKKNQKKNNETNIKHSYILTFLRNSLSNP
IYHSTNEIYDYDENK----------WNGVFGKKTYKKKNNETNIQHSYILTFLMNSLSNP
IYHSTNEIYDYDESK----------WNGILGKKKNQKKNNETNIQHSYILTFLRNSLSNP
*:**** :** ::.*            . : *:. :..:::::.  ***::::* *****

QLQNKYNKKKRIKKMKN-------------------QYEVISITFTSAEINASTINAFKK
QLQNKYNKKKRIKKMKN-------------------QYEVISITFTSAEINASTINAFKK
QLQNKYNKKKRIKKMKN-------------------QYEVISITFTSAEINASTINAFKK
HLGQQYTKNKKRNKRSKSLYNSIRLDENGKNTKRENQYQVISITFSSAEIDKSTINNFKK
HLGQQYTKNKKRNKRSKSLYNSIRLDENGKNTKRENQYQVISITFSSAEIDKSTINNFKK
HLGHQNPKNKKRNKRSK-SYNNTQLGRNGKNTKKLNQYQVISITFSSAEINKSTINNFKK
RLGHQNPKNKKRNKRSK-SYNNTQLGRNGKNTKRLNQYQVISITFSSAEINKSTINNFKK
HLSQQYTKNKKRNKRSK-SYNSIRMGRNDKNTKRESQYQVISITFSSAEIDKSTINNFKK
HLGHQNTKNKKRNKRSK-SYNNIQLGRNGKNTKRINQYQVISITFSSAEIDNSTINNFKK
:* ::  *:*: :* .:                  **:******:****: **** ***

LPSEKFLRTIINVSSAVHVGSDPTLVELKDALNLDALMMMLNIEDMQFLSYQGSSTLPLC
LPSEKFLRTIINVSSAVHVGSDPTLVELKDALNLDALMMMLNIEDMQFLSYQGSSTLPLC
LPSEKFLRTIINVSSAVHVGSDPTLVELKEALNLDALMMMLNIEDMQFLSYQGSSTLPLC
LPSEKFLKTILEASQNVPVGSDPKLVNLKKPLNLNSLLMMLNMKSMEFFAYHGSSTSPDC
LPSEKFLKTILEASQNVPVGSDPKLVNLKKPLNLNSLLMMLNMKSMEFFAYHGSSTSPDC
LPSEKFLKTILEGTQNVPVGSDPTLVDLKAPLNLNSVLMMLNMKSMEFFAYHGSSTSPDC
LPSEKFLKTILEGTQNVPVGSDPTLVDLKAPLNLNSVLMMLNMKSMEFFAYHGSSTSPDC
LPSEKFLKTILEGSQNVPVGSGPKLVNLKEPLNLNSLLMMLNMKSMEFFAYHGSSTSPGC
LPSEKFLKTILEGTQNIPVGSDPTLVDLKVPLNLNSVLMMLNMKSMEFFAYHGSSTTPDC
*******:**:: :. : ***.*.**:** .***::::****::.*:*::*:**** * *

DENVSWKVAKQPLPVSTETILNFYYLLKKHTPNYSGSDNDNYRSLQNVEDNTRHYRKFSL
DENVSWKVAKQPLPVSTETILNFYYLLKKHTPNYSGSDNDNYRSLQNVEDNTRHYRKFSL
DENVSWKVAKQPLPVSTETILNFYYLLKKHTPNYSGSDNDNYRSLQNVEDNTRHYRKFSL
NENVHWKVAKKSLPISTETMLKFYNMLKKTTPDYNASDNDNFRALQNVQGNIHNYGRVYL
NENVHWKVAKKSLPISTETMLKFYNMLKKTTPDYNASDNDNFRALQNVQGNIHNYGRVYL
```

```
NENVHWKVAKKSLPISTETMLKFYNMLKKTTPDYNGSDNDNFRALQNVQGNIHNYGRVYL
NENVHWKVAKKSLPISTETMLKFYNMLKKTTPDYNGSDNDNFRALQNVQGNIHNYGRVYL
NENVHWKVAKKSLPISTETMLKFYNMLKKTTPYYNASDNDNFRALQNVQGNIHNYGRVYL
SENVHWKVAKKSLPISTETMLKFYNMLKKTTPDYNSSDNDNFRALQNVQGNVHNYGRVYL
.*** *****:.**:****:*:** :*** ** *..*****:*:****:.* ::* :. *

VQVFPIQVLISSAISNIEDKKVINIIKDISPKNMSFTYYSKWDIYFILFIFYNIVLFLF
VQVFPIQVLISSAISNIEDKKVINIIKDISPKNMSFSYYSKWDIYFILFIFYNIVLFLF
VQVFPIQVLISSAISNVEDKEVINIIKDISPKNMSFSYYSKWDIYFILFIFYNIVLFLF
IQGFPVQLLISSALTTSEDKNVIENIKLAYSKSSGNYIYFNLIFLLLIFMFLQNY----
IQGFPVQLLISSALTTSEDKNVIENIKLAYSKSSGNYIYFNLIFLLLIFMFLQNY----
IQGFPVQLLISSILTTSEDKTVIENIKQAYSKSNGNYICFNFIFLLLIFIFLQNY----
IQGFPVQLLISSILTTSEDKTVIENIKQAYSKSNGNYICFNFIFLLLIFMFLQNY----
IQGFPVQLLISSALMTSDDKNVIENIKLAYSKSSANYIYFNFIFLLLIFMFLQNY----
IQGFPVQLLISSILTTSEDKTVIENIKQAYSKSNGNYIYFNFIFLLLIFMYLQNY----
:* **:*:**** : . :** **: **   .*. .     :  : :::*:: :
```

Figure 2: Alignment of *Plasmodium* putative CA amino acid sequences. which demonstrate the existence of a new CA family, denominated η-CA class. The presumed Zn(II) ligands are at positions x, x+2 and x+24. Amino acid sequence cryptonym is indicated in Table 3.

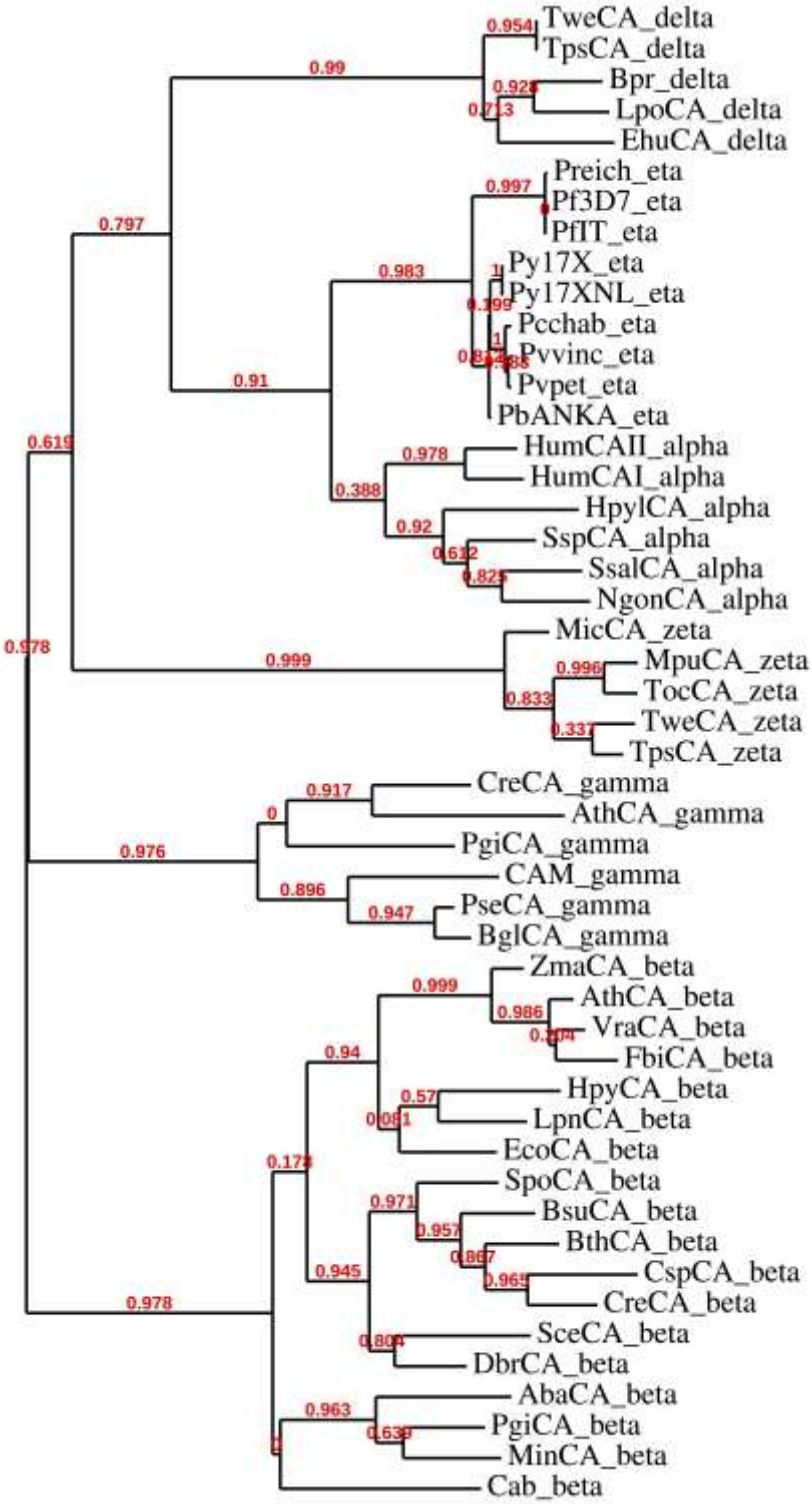

Fig. 3: Phylogenetic tree leading to the discovery of the η-CA genetic family in *Plasmodia* and constructed using amino acid sequences of the α-, β-, γ-, δ-, ζ- and η-CAs from selected prokaryotic and eukaryotic species. The tree was created using the program PhyML 3.0. Branch support values have been reported at branch points. Class, organisms, accession numbers and cryptonyms of the sequences used in the alignment have been indicated in Table 3.