LETTER

doi:10.1038/nature10464

Evolution of a new enzyme for carbon disulphide conversion by an acidothermophilic archaeon

Marjan J. Smeulders[1]*, Thomas R. M. Barends[2]*, Arjan Pol[1], Anna Scherer[2], Marcel H. Zandvoort[1], Anikó Udvarhelyi[2], Ahmad F. Khadem[1], Andreas Menzel[3], John Hermans[1], Robert L. Shoeman[2], Hans J. C. T. Wessels[4], Lambert P. van den Heuvel[4], Lina Russ[1], Ilme Schlichting[2], Mike S. M. Jetten[1] & Huub J. M. Op den Camp[1]

Extremophilic organisms require specialized enzymes for their exotic metabolisms. Acid-loving thermophilic Archaea that live in the mudpots of volcanic solfataras obtain their energy from reduced sulphur compounds such as hydrogen sulphide (H_2S) and carbon disulphide (CS_2)[1,2]. The oxidation of these compounds into sulphuric acid creates the extremely acidic environment that characterizes solfataras. The hyperthermophilic *Acidianus* strain A1-3, which was isolated from the fumarolic, ancient sauna building at the Solfatara volcano (Naples, Italy), was shown to rapidly convert CS_2 into H_2S and carbon dioxide (CO_2), but nothing has been known about the modes of action and the evolution of the enzyme(s) involved. Here we describe the structure, the proposed mechanism and evolution of a CS_2 hydrolase from *Acidianus* A1-3. The enzyme monomer displays a typical β-carbonic anhydrase fold and active site, yet CO_2 is not one of its substrates. Owing to large carboxy- and amino-terminal arms, an unusual hexadecameric catenane oligomer has evolved. This structure results in the blocking of the entrance to the active site that is found in canonical β-carbonic anhydrases and the formation of a single 15-Å-long, highly hydrophobic tunnel that functions as a specificity filter. The tunnel determines the enzyme's substrate specificity for CS_2, which is hydrophobic. The transposon sequences that surround the gene encoding this CS_2 hydrolase point to horizontal gene transfer as a mechanism for its acquisition during evolution. Our results show how the ancient β-carbonic anhydrase, which is central to global carbon metabolism, was transformed by divergent evolution into a crucial enzyme in CS_2 metabolism.

Acidianus species belong to the order Sulfolobales, whose members are common inhabitants of solfataric regions worldwide, and these species are good models of evolution owing to the geographic separation of their habitats[3]. The hyperthermophilic archaeon *Acidianus* A1-3 (refs 4, 5) was isolated from biofilm material from the fumarolic, ancient sauna building at the Solfatara volcano (Naples, Italy) on a mineral medium supplemented with CS_2. It is the first example of a thermophilic CS_2-converting archaeon. Under anaerobic or aerobic conditions, cell extracts of *Acidianus* A1-3 converted CS_2 to H_2S, with COS as the intermediate: $CS_2 + H_2O \rightarrow COS + H_2S$ and $COS + H_2O \rightarrow CO_2 + H_2S$. The enzyme responsible was therefore likely to be a hydrolase, as has been reported for a mesophilic CS_2-converting *Thiomonas* species[6], rather than an oxygenase[7].

We proposed that the hydrolysis of CS_2 to CO_2 by way of COS is similar to the hydration of CO_2 to HCO_3^- by carbonic anhydrases: $CO_2 + H_2O \leftrightarrows HCO_3^- + H^+$. Indeed, conversion of CS_2 by carbonic anhydrase has been explored previously by quantum chemical methods[8–10]. In addition, COS is a known substrate homologue for plant and algal carbonic anhydrases[11,12], and plants may constitute the most important sink of atmospheric COS[13]. Furthermore, the carbonic anhydrase of the flour beetle *Tribolium castaneum* can hydrolyse COS to H_2S[14]. However, CS_2 has not been found to be an effective substrate for carbonic anhydrase, suggesting the presence of a different enzyme for CS_2 conversion.

The CS_2-converting enzyme (subunit mass 24 kDa) was purified from *Acidianus* A1-3 cells and shown to be a CS_2 hydrolase. The Michaelis constant (K_m) for CS_2 is 130 μM, which is 20-fold lower than the K_m for CO_2 of β-carbonic anhydrase (Cab) from *Methanothermobacter thermoautotrophicus* (2.9 mM)[15] and 100-fold lower than the K_m for CO_2 of γ-carbonic anhydrase (Cam) from *Methanosarcina thermophila* (15 mM)[16]. The catalytic rate (k_{cat}) of the CS_2 hydrolase (952 s^{-1}) is also lower than the k_{cat} values of Cab (1.7×10^4 s^{-1}) and Cam (3×10^4 s^{-1}), resulting in a similar overall catalytic efficiency (k_{cat}/K_m) of 2–7.3×10^6 s^{-1} M^{-1} for both the CS_2 hydrolase and the carbonic anhydrases (Supplementary Table 2). The CS_2 hydrolase converted the intermediate COS with a higher affinity (K_m = 22 μM) and k_{cat} (1,800 s^{-1}) than for CS_2 (Supplementary Fig. 1). The catalytic efficiency of the CS_2 hydrolase is almost an order of magnitude higher for COS than for CS_2. The CS_2 hydrolase did not show carbonic anhydrase activity in stopped-flow spectrometry[17] experiments. Inductively coupled plasma mass spectrometry (ICP-MS) analysis showed zinc to be associated with the CS_2 hydrolase, as in most aerobic carbonic anhydrases that have been studied so far. Exceptions are cadmium-containing carbonic anhydrases from diatoms[18] and iron-containing anaerobic carbonic anhydrases from methanogens[13,19]. The gene encoding the CS_2 hydrolase was identified by peptide sequencing and PCR amplification. Sequence analysis showed high homology to β-carbonic anhydrases (Fig. 1 and Supplementary Fig. 2).

The structures of two crystal forms of the CS_2 hydrolase (selenomethionine-labelled heterologously expressed protein and native purified protein) were determined to resolutions of 2.6 Å and 2.4 Å, respectively (Supplementary Table 3). Crystalline CS_2 hydrolase is active, as confirmed by the generation of H_2S from CS_2 in the crystal (Supplementary Fig. 3). The *Acidianus* CS_2 hydrolase displays a β-carbonic anhydrase fold: two monomers form a closely intertwined dimer with a central β-sheet capped by an α-helical domain (Fig. 2a). In both crystal forms, the CS_2 hydrolase dimers form square octamers with a central hole, through interactions of long arms at the N and C termini (Fig. 2a, b). Similar octameric rings have been reported for the β-carbonic anhydrase of *Pisum sativum*[20]. However, in the CS_2 hydrolase, two octameric rings form a hexadecamer by interlocking at right angles to each other, forming a highly unusual catenane structure (Fig. 2c). To our knowledge, the only other examples of natural protein catenanes are the bovine mitochondrial peroxiredoxin III (ref. 21) and the gp5 capsid protein of bacteriophage HK97 (ref. 22). Small-angle X-ray scattering (SAXS) (Fig. 2d) and analytical ultracentrifugation

[1]Department of Microbiology, Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ, Nijmegen, The Netherlands. [2]Department of Biomolecular Mechanisms, Max-Planck Institute for Medical Research, Jahnstrasse 29, D-69120 Heidelberg, Germany. [3]Paul Scherrer Institut, CH-5232 Villigen PSI, Switzerland. [4]Nijmegen Centre for Mitochondrial Disorders, Nijmegen Proteomics Facility, Department of Laboratory Medicine, Radboud University Nijmegen Medical Centre, Geert Grooteplein 10, PO Box 9101, 6500 HB Nijmegen, The Netherlands.
*These authors contributed equally to this work.

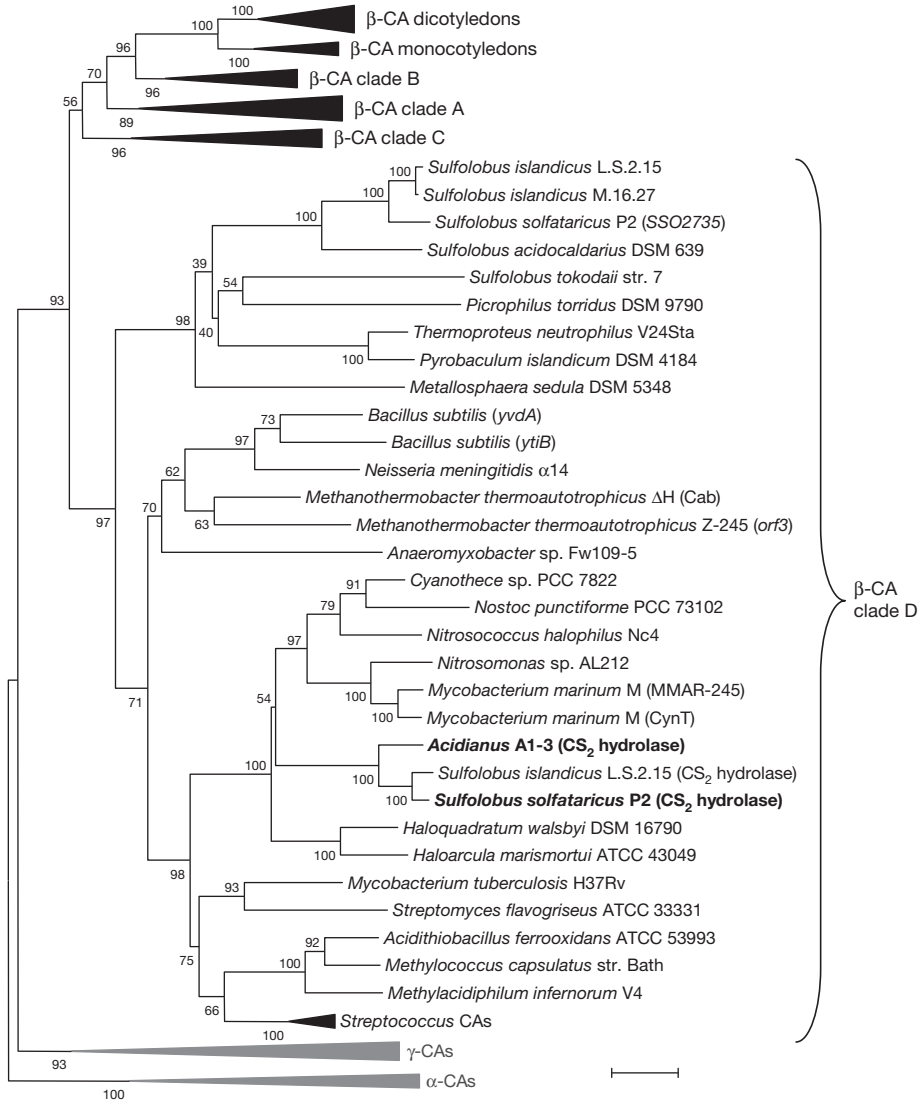

Figure 1 | Neighbour-joining phylogenetic tree showing the evolutionary relationship between carbonic anhydrases (CAs) and the newly discovered CS_2 hydrolases. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of 96 sequences that were retrieved from GenBank and through BLAST, aligned using MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/) and then analysed in MEGA 4.0. Branches corresponding to partitions that were reproduced in less than 50% of bootstrap replicates are collapsed. The bootstrap values are indicated at the branch points. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances that were used to infer the phylogenetic tree (scale bar, 0.2 amino acid substitutions per site). All positions that contain alignment gaps and missing data were eliminated only in pairwise sequence comparisons. The final data set contained 1,382 positions. The eight γ-carbonic anhydrase and six α-carbonic anhydrase sequences were used as outgroups (grey). β-Carbonic anhydrase sequences are clustered in clades (dicotyledons, monocotyledons and clades A–D). CS_2 hydrolase sequences are clustered in clade D. Gene or protein names are indicated in parentheses where relevant. Confirmed CS_2 hydrolases are shown in bold. Accession numbers are listed in the Supplementary Information.

(Supplementary Fig. 4) confirmed the presence of this unusual catenane structure in solution. In addition, both forms of the enzyme were detected and shown to actively convert CS_2 to H_2S on native polyacrylamide gel electrophoresis (PAGE) gels (Supplementary Fig. 5). Considering the large amount of CS_2 hydrolase present in a crude extract of *Acidianus* A1-3 cells growing on CS_2 (Supplementary Fig. 5), as well as the gaseous nature of its intermediate product, COS, tight packing of the enzyme would enable more enzyme to be present in the cell and result in a higher density of active sites for COS to react with before it diffuses away. Moreover, the arms at the N and C termini that form the octamer lie along the sides of the square octameric ring, so interlocking two of these rings may help to keep these arms in place, stabilizing the octameric structure.

The active site of native CS_2 hydrolase contains a zinc ion coordinated by Cys 35, His 88, Cys 91 and a solvent molecule (Fig. 3a). Other solvent molecules were not observed in the active site. Structural comparisons of the active sites of the CS_2 hydrolase and the most closely related β-carbonic anhydrases[23–26], as identified by the Dali program[27], showed zinc-coordinating residues at identical positions (Fig. 3b). This finding suggests that the first step in the catalytic mechanism of the CS_2 hydrolase, involving a hydroxide ion, is analogous to that of β-carbonic anhydrase[28] (Supplementary Fig. 6). Quantum chemical studies of carbonic anhydrase active site models[9,10] suggest that the zinc ion is subsequently coordinated by both the oxygen from the former hydroxide and a sulphur atom from the substrate. The cleavage of the very stable Zn–S bond of the $HS^- - Zn$ complex could be assisted by protonation of the HS^- moiety[9]. The hydrolysis of COS may proceed analogously (Supplementary Fig. 6).

The similarity between the CS_2 hydrolase and carbonic anhydrase active sites raises the question of why the CS_2 hydrolase has a markedly different reaction specificity from carbonic anhydrase, exclusively hydrolysing CS_2 and COS rather than CO_2. The answer may lie in

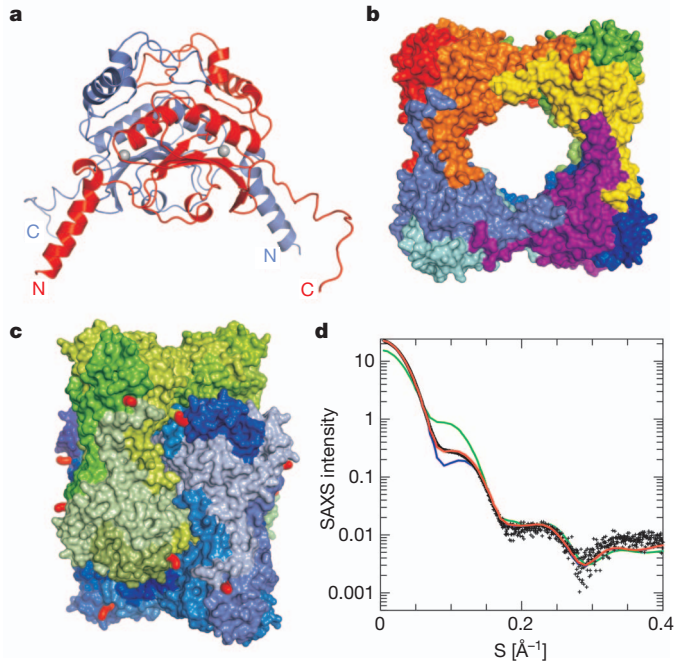

Figure 2 | X-ray crystal structure of the *Acidianus* A1-3 CS_2 hydrolase. **a**, Dimer structure. One monomer is shown in red, and the other in blue. The long arm-like extensions at the N and C termini of each monomer are indicated. The active site zinc ions are indicated as grey spheres. **b**, Four dimers form a square octameric ring through extensive interactions of the N- and C-terminal arms. Monomers are shown in contrasting colours. **c**, The hexadecameric complex is formed by two interlocked octameric rings. One ring is shown with the monomers in various shades of blue, and the other with the monomers in various shades of green. Access routes to the active site tunnels are shown in red. **d**, SAXS results from a solution of CS_2 hydrolase (data points) and fits of the data (curves) assuming 100% hexadecamer (blue), 100% octamer (green) or a mixture of 83% hexadecamer and 17% octamer (red). S corresponds to the momentum transfer in $Å^{-1}$.

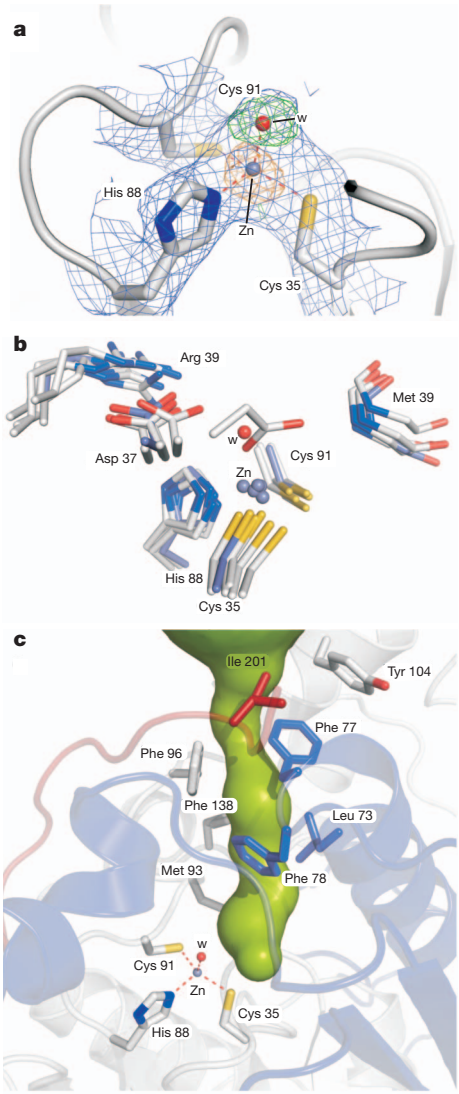

Figure 3 | The CS_2 hydrolase active site and the entrance to the active site. **a**, The active site of the CS_2 hydrolase. The catalytic zinc ion (blue sphere) is coordinated by two cysteine residues, a histidine residue and a solvent molecule (w, red sphere). The interactions between the zinc atom and its ligands are shown as dashed red lines. The final, refined $2mF_o - DF_c$ electron density (blue, 1σ) is shown. The $F_o - DF_c$ electron density (green, 2σ), which was calculated before the incorporation of the solvent molecule into the model, is also shown, as well as the anomalous difference density (orange, 5σ), which confirms the position of the zinc ion. **b**, Superposition of active site residues of the CS_2 hydrolase (pale blue, carbon atoms) and structurally related β-carbonic anhydrases (white, carbon atoms): Cab (clade D) from *M. thermoautotrophicus*[23], Rv3588c (clade C) and Rv1284 (clade D) from *Mycobacterium tuberculosis*[24], Nce103 (clade A) from *Saccharomyces cerevisiae*[25] and CAN2 (clade A) from *Cryptococcus neoformans*[26]. For *Mycobacterium* Rv3588c, which is a type II β-carbonic anhydrase, the equivalent of Asp 37 coordinates the zinc ion, indicating the position of the substrate. By contrast, in the type I β-carbonic anhydrases, this residue forms a salt bridge with an arginine residue, as is the case for the CS_2 hydrolase (Arg 39). The hydrolytic solvent molecule in the CS_2 hydrolase (w) is indicated as a red sphere. **c**, Hydrophobic tunnel (green) providing access to the CS_2 hydrolase active site. The residues that constitute the tunnel wall are indicated, and each protein monomer is individually coloured (red, blue and grey). **a**–**c**, Nitrogen atoms are shown in dark blue, oxygen atoms in red and sulphur atoms in yellow.

the access to the active sites of these enzymes. Superimposing structurally related carbonic anhydrases shows that the same funnel supplies access to the active site (Supplementary Fig. 7). However, in the CS_2 hydrolase, this access route is blocked by the protomer N- and C-terminal arms within the octamer (Supplementary Fig. 7), and the only remaining access to the active site is by way of a very narrow and highly hydrophobic tunnel of 15 Å in length (Fig. 3c), which is lined by residues from three monomers. Thus, the formation of the octamer substantially reduces access to the active site for hydrophilic substrates. Indeed, CS_2 is over an order of magnitude more hydrophobic than CO_2, as judged by octanol:water partitioning coefficients (using the LOGKOW database, http://logkow.cisti.nrc.ca/logkow/). Although we cannot exclude second-sphere effects, we propose that the narrow hydrophobic tunnel is an important factor in determining the substrate specificity of the CS_2 hydrolase.

To test whether active site accessibility determines the reaction specificity, we designed seven CS_2 hydrolase mutants, with the amino acid residues that form the tunnel wall being replaced by smaller or larger residues. Furthermore, the C terminus was truncated (mutant Gly199Stop) (Fig. 4a). The activities of the mutant CS_2 hydrolases were quantified (Fig. 4b) and were visualized on native PAGE activity gels (Supplementary Fig. 5). The mutants with a change in Phe 78 lost most or all of their activity (Fig. 4b). Phe 78 is close to the active site (Fig. 4a), and any change in this region of the protein adversely affected enzyme activity. By contrast, widening the entrance of the tunnel by shortening the C terminus or by replacing large hydrophobic residues with smaller ones (for example, Phe77Ala or Phe96Ser) (Fig. 4a) reproducibly increased the catalytic efficiency of the mutants by up to twofold (Fig. 4b). This finding suggests that the CS_2 conversion rate depends on the properties of the tunnel in modulating the accessibility of the active site to the reactants (CS_2, COS or H_2O) or to product release. Although one might expect the reactivity towards CO_2 to increase when

the tunnel entrance is widened, this was not the case; none of the mutants showed carbonic anhydrase activity when tested using stopped-flow spectrometry. It is possible that it is not CO_2 that poses accessibility problems but its hydrolysis product, HCO_3^-. Furthermore, CO_2 is capable of hydrogen bonding to the surrounding water molecules, whereas CS_2 is not; therefore, the narrow tunnel might discriminate between CO_2–H_2O complexes and CS_2 on the basis of size. This would also explain how the reaction products CO_2 and H_2O could exit the active site through the tunnel.

To investigate possible routes by which the CS_2 hydrolase evolved, we searched for homologous protein sequences and found a 204 amino acid protein (AAK41461) encoded in the genome of *Sulfolobus solfataricus* P2 (ref. 29) (Fig. 1 and Supplementary Fig. 2), one of two clade D β-carbonic anhydrase homologues in this organism. The encoding gene (*SSO1214*) is not present in the genomes of other members of the Sulfolobales that have been sequenced, except in one of the seven sequenced *Sulfolobus islandicus* genomes[3] (strain L.S.2.15) (Fig. 1 and Supplementary Fig. 2). Indeed, we detected CS_2 hydrolase activity in *S. solfataricus* P2 cells but not in cells of its close relative

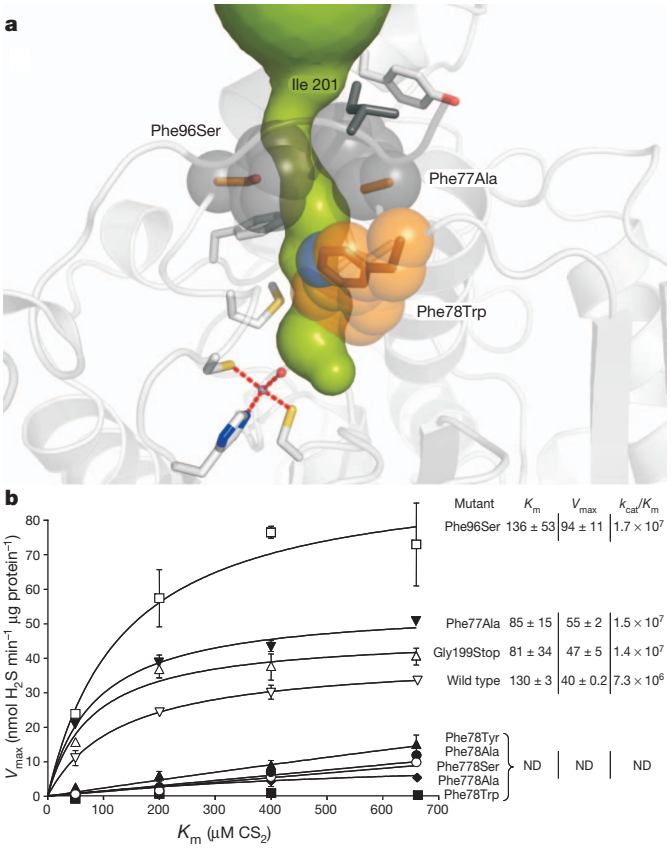

Figure 4 | Effect of mutations in the access tunnel on activity. **a**, The rationale underlying the design of selected mutants. The tunnel is indicated in green. Native residues that were mutated are shown as large grey spheres; some of the mutations are modelled in orange. The large blue sphere is the nitrogen of the Trp residue. Nitrogen atoms are shown in dark blue, oxygen atoms in red and sulphur atoms in yellow. The interactions between the zinc atom (small blue sphere) and its ligands are shown as dashed red lines. Phe 77 and Phe 96 were replaced by smaller residues; Phe 78 was replaced by Trp; and Ile 201 is absent in the Gly199Stop mutant. A stereo version of this panel is provided as Supplementary Fig. 9. **b**, Michaelis–Menten kinetics of the CS_2 hydrolase mutants. The data are presented as the mean ± s.e.m., as calculated from three independent experiments. The K_m, V_{max} and k_{cat}/K_m (in s^{-1} M^{-1}) values were determined by nonlinear regression analysis. Phe778Ser and Phe778Ala correspond to the Phe77Ser plus Phe78Ser and Phe77Ala plus Phe78Ala double mutants, respectively. ND, not determined.

Sulfolobus shibatae or in three *Acidianus* species (*Acidianus brierleyi*, *Acidianus ambivalens* and *Acidianus infernus*). The heterologously expressed *S. solfataricus SSO1214* gene product was shown to have CS_2 hydrolase activity but no carbonic anhydrase activity (Supplementary Fig. 8), suggesting that the *S. solfataricus* AAK41461 protein has been erroneously annotated as a carbonic anhydrase. As is the case for the *Acidianus* CS_2 hydrolase, AAK41461 is a carbonic anhydrase homologue that has evolved divergently to have CS_2 hydrolase activity. Most likely, the true carbonic anhydrase of *S. solfataricus* P2 is the second carbonic anhydrase gene to be identified, *SSO2735*, which encodes the protein AAK42846. This protein clusters with other archaeal β-carbonic-anhydrase-like enzymes, whereas the CS_2 hydrolases cluster mainly with putative bacterial β-carbonic anhydrases (Fig. 1). Inspection of the genomic environment of the CS_2 hydrolase gene *SSO1214* revealed the presence of transposon-associated genes both upstream (*SSO1212*) and downstream (*SSO1216* and *SSO1217*) of *SSO1214*. These transposon-associated genes could also be identified surrounding hypothetical CS_2 hydrolase genes from *S. islandicus* strain L.S.2.15, *Nitrosomonas* sp. AL212, *Mycobacterium marinum* and *Haloquadratum walsbyi*. Furthermore, a Phe-77 Phe-78 motif is present in confirmed CS_2 hydrolases and in the above strains but not in confirmed carbonic anhydrases, strongly suggesting that this motif is a characteristic feature of CS_2 hydrolases (Fig. 1 and Supplementary Fig. 2). In addition, as mentioned above, Phe-78-modified mutants all lose their CS_2 hydrolase activity. Taken together, these findings point to lateral gene transfer as a likely mechanism for the acquisition of the CS_2 hydrolase gene in these organisms, rather than duplication and divergence of their carbonic anhydrase gene. It may also explain why the ability to convert CS_2 is not shared by all Sulfolobales.

Our results show a new example of divergent evolution: CS_2 and CO_2 are non-interchangeable substrates of CS_2 hydrolases and carbonic anhydrases, although the catalytic sites of these enzymes are identical. The difference between the two enzymes lies in the unusual oligomeric structure of CS_2 hydrolase, in which a specificity filter is built into the active site entrance. Thus, a new enzyme emerged owing to the evolution of a new quaternary structure rather than through active site mutations. The omnipresence and convergent evolution of the carbonic anhydrases, as well as the absence of the CS_2 hydrolase in most of the members of the Sulfolobales that have been sequenced, suggests that carbonic anhydrases are the older of the two enzymes. Our data contribute to the ongoing discussion on the evolution and (re)design of enzymes, which has been reviewed recently[30].

METHODS SUMMARY

Acidianus strain A1-3 was cultivated at 70 °C in a chemostat on modified Allen's mineral medium with CS_2 as the main carbon and energy source. Cells from the chemostat effluent were collected by centrifugation and lysed using a French press. CS_2 conversion was examined by gas chromatographic analysis or with an H_2S sensor. Carbonic anhydrase activity was measured by stopped-flow spectrometry at 25 °C. CS_2 hydrolase activity on native PAGE gels was determined using Pb(II) acetate. The *Acidianus* CS_2 hydrolase was purified from cell-free extracts by cation and anion exchange chromatography. The molecular mass of the purified CS_2 hydrolase was determined by matrix-assisted laser desorption/ionization–time of flight mass spectrometry (MALDI-TOF MS), and ICP-MS was used to determine the presence of metal cofactors. The purified CS_2 hydrolase was digested with trypsin, and the dominant peptides were sequenced using tandem mass spectrometry. The amino acid sequences were used to design degenerate PCR primers, and the CS_2 hydrolase gene was amplified with *Acidianus* genomic DNA as a template. *S. solfataricus SSO1214* was cloned after PCR amplification from genomic DNA. Native and mutant proteins were obtained by heterologous expression of synthetic gene constructs and were purified using heat-shock precipitation and gel filtration where applicable. Phylogenetic analyses were carried out using the MEGA 4.0 package. Crystal structures were determined from native and selenomethionine-labelled protein crystals, by using single-wavelength anomalous dispersion methods. SAXS and analytical ultracentrifugation were used to ascertain the oligomeric state of the protein.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 23 March; accepted 19 August 2011.

1. Allard, P., Maiorani, A., Tedesco, D., Cortecci, G. & Turi, B. Isotopic study of the origin of sulfur and carbon in solfatara fumaroles, Campi Flegrei Caldera. *J. Volcanol. Geotherm. Res.* **48,** 139–159 (1991).
2. Vasilakos, C., Maggos, T., Bartzis, J. G. & Papagiannakopoulos, P. Determination of atmospheric sulfur compounds near a volcanic area in Greece. *J. Atmos. Chem.* **52,** 101–116 (2005).
3. Reno, M. L., Held, N. L., Fields, C. J., Burke, P. V. & Whitaker, R. J. Biogeography of the *Sulfolobus islandicus* pan-genome. *Proc. Natl Acad. Sci. USA* **106,** 8605–8610 (2009).
4. Fuchs, T., Huber, H., Burggraf, S. & Stetter, K. O. 16S rDNA-based phylogeny of the archaeal order Sulfolobales and reclassification of *Desulfurolobus ambivalens* as *Acidianus ambivalens* comb nov. *Syst. Appl. Microbiol.* **19,** 56–60 (1996).
5. He, Z., Li, Y., Zhou, P. & Liu, S.-J. Cloning and heterologous expression of a sulfur oxygenase/reductase gene from the thermoacidophilic archaeon *Acidianus* sp. S5 in *Escherichia coli. FEMS Microbiol. Lett.* **193,** 217–221 (2000).
6. Pol, A., van der Drift, C. & Op den Camp, H. J. M. Isolation of a carbon disulfide utilizing *Thiomonas* sp and its application in a biotrickling filter. *Appl. Microbiol. Biotechnol.* **74,** 439–446 (2007).
7. Jordan, S. L. *et al.* Autotrophic growth on carbon disulfide is a property of novel strains of *Paracoccus denitrificans. Arch. Microbiol.* **168,** 225–236 (1997).
8. Sinnecker, S., Brauer, M., Koch, W. & Anders, E. CS_2 fixation by carbonic anhydrase model systems—a new substrate in the catalytic cycle. *Inorg. Chem.* **40,** 1006–1013 (2001).
9. Notni, J., Schenk, S., Protoschill-Krebs, G., Kesselmeier, J. & Anders, E. The missing link in COS metabolism: a model study on the reactivation of carbonic anhydrase from its hydrosulfide analogue. *ChemBioChem* **8,** 530–536 (2007).
10. Schenk, S., Kesselmeier, J. & Anders, E. How does the exchange of one oxygen atom with sulfur affect the catalytic cycle of carbonic anhydrase? *Chem. Eur. J.* **10,** 3091–3105 (2004).
11. Protoschill-Krebs, G., Wilhelm, C. & Kesselmeier, J. Consumption of carbonyl sulphide (COS) by higher plant carbonic anhydrase (CA). *Atmos. Environ.* **30,** 3151–3156 (1996).
12. Blezinger, S., Wilhelm, C. & Kesselmeier, J. Enzymatic consumption of carbonyl sulfide (COS) by marine algae. *Biogeochemistry* **48,** 185–197 (2000).
13. Chin, M. & Davis, D. D. Global sources and sinks of OCS and CS_2 and their distributions. *Glob. Biogeochem. Cycles* **7,** 321–337 (1993).
14. Haritos, V. S. & Dojchinov, G. Carbonic anhydrase metabolism is a key factor in the toxicity of CO_2 and COS but not CS_2 toward the flour beetle *Tribolium castaneum* [Coleoptera: Tenebrionidae]. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* **140,** 139–147 (2005).
15. Smith, K. S. & Ferry, J. G. A plant-type (β-class) carbonic anhydrase in the thermophilic methanoarchaeon *Methanobacterium thermoautotrophicum. J. Bacteriol.* **181,** 6247–6253 (1999).
16. Alber, B. E. *et al.* Kinetic and spectroscopic characterization of the γ-carbonic anhydrase from the methanoarchaeon *Methanosarcina thermophila. Biochemistry* **38,** 13119–13128 (1999).
17. Khalifah, R. G. The carbon dioxide hydration activity of carbonic anhydrase. *J. Biol. Chem.* **246,** 2561–2573 (1971).
18. Lane, T. W. *et al.* A cadmium enzyme from a marine diatom. *Nature* **435,** 42 (2005).
19. MacAuley, S. R. *et al.* The archetype γ-class carbonic anhydrase (Cam) contains iron when synthesized *in vivo. Biochemistry* **48,** 817–819 (2009).
20. Kimber, M. S. & Pai, E. F. The active site architecture of *Pisum sativum* β-carbonic anhydrase is a mirror image of that of α-carbonic anhydrases. *EMBO J.* **19,** 1407–1418 (2000).
21. Cao, Z., Roszak, A. W., Gourlay, L. J., Lindsay, J. G. & Isaacs, N. W. Bovine mitochondrial peroxiredoxin III forms a two-ring catenane. *Structure* **13,** 1661–1664 (2005).
22. Wikoff, W. R. *et al.* Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289,** 2129–2133 (2000).
23. Strop, P., Smith, K. S., Iverson, T. M., Ferry, J. G. & Rees, D. C. Crystal structure of the "cab"-type β class carbonic anhydrase from the archaeon *Methanobacterium thermoautotrophicum. J. Biol. Chem.* **276,** 10299–10305 (2001).
24. Suarez Covarrubias, A. *et al.* Structure and function of carbonic anhydrases from *Mycobacterium tuberculosis. J. Biol. Chem.* **280,** 18782–18789 (2005).
25. Teng, Y. B. *et al.* Structural insights into the substrate tunnel of *Saccharomyces cerevisiae* carbonic anhydrase Nce103. *BMC Struct. Biol.* **9,** 67–76 (2009).
26. Schlicker, C. *et al.* Structure and inhibition of the CO_2-sensing carbonic anhydrase Can2 from the pathogenic fungus *Cryptococcus neoformans. J. Mol. Biol.* **385,** 1207–1220 (2009).
27. Holm, L., Kaarlainen, S., Rosenstrom, P. & Schenkel, A. Searching protein structure databases with DaliLite v.3. *Bioinformatics* **25,** 2780–2781 (2008).
28. Rowlett, R. S. Structure and catalytic mechanism of the β-carbonic anhydrases. *Biochim. Biophys. Acta* **1804,** 362–373 (2010).
29. Smith, K. S., Jakubzick, C., Whittam, T. S. & Ferry, J. G. Carbonic anhydrase is an ancient enzyme widespread in prokaryotes. *Proc. Natl Acad. Sci. USA* **96,** 15184–15189 (1999).
30. Bornberg-Bauer, E., Huylmans, A.-K. & Sikosek, T. How do new proteins arise? *Curr. Opin. Struct. Biol.* **20,** 390–396 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements J. Eygensteyn is acknowledged for ICP-MS analysis. We thank H. Harhangi for advice on cloning, S. Zimmermann for assistance with cloning, A. Meinhart for discussions and support for cloning and crystallography, and D. Ringe and J. Reinstein for discussions. We thank the Dortmund-Heidelberg data collection team, especially W. Blankenfeldt, and the staff of beam lines X10SA and X12SA at the Swiss Light Source of the PSI in Villigen for their help and facilities. We also thank A. Rufer for help with the analytical ultracentrifuge and I. Vetter for support with the crystallographic software. The work was funded by an STW grant (STW_6353) to M.J.S. and M.H.Z. and by the Max-Planck Society.

Author Contributions The research was conceived by A.P., M.S.M.J. and H.J.M.O. M.H.Z. performed the sampling, enrichment and isolation. M.J.S., M.H.Z. and A.P. performed the physiological experiments. M.J.S., J.H., A.F.K., L.P.V. and H.J.C.T.W. performed the purification and protein and gene sequencing. A.F.K. and H.J.M.O. performed the MALDI-TOF MS. H.J.M.O. and M.J.S. performed the alignments and phylogenetic analyses. M.J.S. and L.R. performed the site-directed mutant studies, including the activity measurements. A.S. and T.R.M.B. grew the crystals; T.R.M.B. determined the crystal structures and suggested the amino acid residues to be mutated; I.S., A.U. and A.M. performed the SAXS experiments and analyses; and R.L.S. performed the analytical ultracentrifugation experiments . M.J.S., T.R.M.B., I.S., M.S.M.J. and H.J.M.O. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information The sequence for the CS_2 hydrolase from *Acidianus* A1-3 has been deposited in GenBank under the accession number HM805096. Atomic coordinates and structure factor amplitudes have been deposited in the Protein Data Bank under accession numbers 3TEO and 3TEN. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.S.M.J. (m.jetten@science.ru.nl).

METHODS

Strains, media and culture conditions. *Acidianus* strain A1-3 was isolated from the fumarolic, ancient sauna building at the Solfatara volcano (Naples, Italy). Enrichment and pure cultures of this strain were grown in a chemostat at 70 °C on modified Allen's mineral medium[31] supplemented with 0.1 g l^{-1} yeast extract and Vishniac trace elements[32] with an additional 4.5 mg l^{-1} $Na_2B_4O_7 \cdot 10H_2O$ and 0.03 mg l^{-1} $VOSO_4 \cdot 5H_2O$. The medium was acidified with 0.125% (v/v) sulphuric acid, resulting in a pH of approximately 1. The main carbon and energy source was supplied as 80 μM CS_2 gas mixed with air, bubbled through the reactor with a gas flow of 20 ml min^{-1}. The growth yield owing to the consumption of CS_2 was calculated to be 80% of the total growth yield under these conditions. A pure culture of *Acidianus* A1-3 was obtained by plating an enrichment culture onto Gelrite plates containing mineral medium with 2% elemental sulphur (S^0) as the electron donor, 0.1 g l^{-1} yeast extract and 0.1% (v/v) H_2SO_4. Plates were incubated at 70 °C with 5% (v/v) CO_2 gas as the carbon source. Single colonies were subcultured twice until purity. After a pure culture was obtained with growth on CO_2 and S^0, the capacity of the microorganism to utilize CS_2 was reconfirmed.

A. ambivalens 3772, *A. brierleyi* 6334 and *A. infernus* 3191 were obtained from DSMZ and grown at 80 °C, 70 °C and 88 °C, respectively. They were grown in medium 358 (http://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium358.pdf), pH 2.5, supplemented with 1% (w/v) S^0 (*A. ambivalens* and *A. infernus*) or 5% (w/v) S^0 (*A. brierleyi*), 5% CO_2 and 0.02% (w/v) yeast extract (*A. ambivalens*) or 1% (w/v) yeast extract (*A. brierleyi* and *A. infernus*). *S. solfataricus* P2 and *S. shibatae* (courtesy of J. van der Oost) were grown at 80 °C in modified Allen's mineral medium with either sterile S^0 plus 5% CO_2 or 16.7 mM glucose. Strains were grown in serum bottles closed with grey butyl rubber stoppers. When cultures had grown for approximately 1 week, CS_2 was added to a head-space concentration of 20 μM, and the conversion of CS_2 to COS and H_2S was examined using a Chrompack CP9001 gas chromatograph, as described previously[33].

Cell-free extracts. Chemostat effluent was collected and kept at 4 °C for a maximum of 4 days before cells were collected by centrifugation (10,000*g*, 10 min, 4 °C). The pellet was washed in 10 mM phosphate buffer, pH 7, until the pH was between 6 and 7. The cells were lysed by at least four passes through a French press (American Instrument Corporation) at 1,260 psi. Cell-free extract was obtained by centrifuging the suspension at 48,000*g* for 30 min at 4 °C, to remove unlysed cells and cell debris. Glycerol was added to 5% (v/v), and extracts were stored at −20 °C.

Mechanism of CS_2 conversion. The conversion of CS_2 into H_2S and COS was followed in aerobic and anaerobic 120 ml stopper bottles containing 5 ml 20 mM phosphate buffer, pH 7. For the anaerobic experiments, the bottles were flushed with N_2 for 30 min. In addition, the syringes were flushed with N_2 before use. The CS_2 gas stock bottles were prepared by adding 20 μl liquid CS_2 to 120 ml bottles, with the liquid evaporating inside the bottles to form CS_2 gas stock bottles. CS_2 gas from the aerobic and anaerobic stock bottles was injected into the buffer-containing aerobic and anaerobic 120 ml bottles, to achieve a head-space concentration of 8.5 μM. The bottles were incubated at 60 °C with shaking at 100 r.p.m. Pre-warmed, diluted *Acidianus* cell-free extract (13 μg protein in 100 μl) was injected into the bottles, and the conversion of CS_2 into its products, COS and H_2S, was followed by gas chromatography.

Enzyme purification. The *Acidianus* CS_2 hydrolase was purified from cell-free extracts by cation exchange chromatography, followed by anion exchange chromatography. The chemicals used for chromatography were of ultrapure grade. The *Acidianus* cell-free extract (150 mg protein) was loaded onto a 14 × 2 cm column packed with Carboxymethyl Sepharose Fast Flow beads (GE Healthcare), which was equilibrated with 50 mM acetate buffer, pH 5. The proteins were separated at room temperature, using fast protein liquid chromatography (Hitachi) with a linear gradient from 0 to 1 M NaCl over 10 column volumes and a flow rate of 2 ml min^{-1}. The fractions were collected and assayed for CS_2 hydrolase activity (see below). The active fractions were pooled and concentrated twofold, and the buffer was exchanged to 20 mM Bis-Tris propane-HCl, pH 6.5, by filtration in an ultrafiltration cell with a 43-mm diameter PM30 membrane (Amicon) at room temperature and with stirring at 100 r.p.m. The concentrated and washed fractions were filtered through a 0.45-μm syringe filter unit (Millex-HV, Durapore PVDF) and loaded onto an anionic DEAE Sepharose Fast Flow 5PW column (75 × 7.5 mm, TSK) equilibrated with 20 mM Bis-Tris propane-HCl, pH 6.5, using high-performance liquid chromatography (with Agilent 1100 series equipment). The proteins were separated at room temperature, using a gradient from 0 to 0.5 M NaCl over 25 min (4.5 column volumes) with a flow rate of 0.6 ml min^{-1}. The fractions containing the purified enzyme were stored on ice or, if not used within 1 week, were frozen at −20 °C after the addition of glycerol to 5% (v/v).

MALDI-TOF MS. The molecular mass of the purified CS_2 hydrolase was determined by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). The purified protein fraction (4.7 pmol in 0.5 μl) was mixed with an equal volume of matrix solution, consisting of 10 g l^{-1} 3,5-dimethoxy-4-hydroxycinnamic acid in 0.05% (v/v) trifluoroacetic acid and 50% (v/v) acetonitrile. The resultant mixture was spotted onto a stainless steel target plate, dried and analysed in linear mode on a Biflex III MALDI-TOF mass spectrometer (Bruker), and then calibrated with a ProteoMass Protein MALDI-MS Calibration Kit (Sigma).

Determination of metal-associated cofactors by ICP-MS. ICP-MS was used to determine the presence of metal cofactors associated with 17.6 μg purified CS_2 hydrolase from the anion exchange purification step (in 20 mM Bis-Tris propane-HCl, pH 6.5, and 230 mM NaCl). The negative controls consisted of protein fractions without CS_2 hydrolase activity, buffer and deionized water. Carbonic anhydrase from bovine erythrocytes (20 μg; SERVA) served as a positive control. Samples were prepared and analysed on an ELEMENT 2 ICP MS (Thermo Scientific) according to the manufacturer's protocol. Height point calibration curves were created by dilution of a standard multielement solution of zinc and copper (at a final concentration of 0, 10, 20 and 50 p.p.b.; Merck) in 0.6% nitric acid solution. Quantification was performed using PlasmaLab software (Thermo).

Protein gel electrophoresis. SDS–PAGE (12% polyacrylamide, pH 8.3) and native PAGE (8% polyacrylamide, pH 8.3) were performed using a Mini-PROTEAN 3 Cell (Bio-Rad) at room temperature. Fermentas SM0431 molecular weight standards were used for SDS–PAGE. For native PAGE, various mixtures of bovine erythrocyte carbonic anhydrase (SERVA), blue dextran, apoferritin, β-amylase, β-lactoglobulin, thyroglobulin and albumin (Sigma) were used as indicators of molecular weight. Proteins were visualized with Coomassie Brilliant Blue G-250. The CS_2 hydrolase activity on native PAGE gels was determined by incubating the gels in an airtight jar at 60 °C in 20 mM Tris, pH 6.9, containing 1 mM Pb(II) acetate, in the presence of about 2 mM CS_2 gas (achieved by injecting 15 ml CS_2-saturated air from an airtight 1 l bottle containing 50 ml CS_2). The gels were incubated until brown bands, consisting of lead sulphide precipitate, appeared.

CS_2 hydrolase activity measurements. The CS_2 hydrolase activity in the purification fractions was monitored qualitatively by adding samples to microtitre wells containing 100 μl 10 mM phosphate buffer or 20 mM HEPES buffer, pH 7, and 10 mM Pb(II) acetate. The Pb(II) acetate precipitated over time, but this did not negatively influence the activity tests. The microtitre plate was incubated for 5–30 min at 50–70 °C in an airtight jar in the presence of CS_2 as described above. In the wells containing the active fractions, the H_2S produced by the CS_2 hydrolase formed a clearly visible brown lead sulphide precipitate. The CS_2 hydrolase activity was quantified by gas chromatography, as well as with an H_2S microsensor. For gas chromatography, appropriate amounts of the enzyme were diluted in 10 mM phosphate buffer, pH 7, in 120 ml bottles sealed with grey butyl rubber stoppers, and 5 ml CS_2-saturated air was injected to a final concentration of about 700 μM CS_2. The bottles were incubated at 50 °C with vigorous shaking to ensure good transfer of gas into the buffer. The formation of CS_2 hydrolysis products (COS and H_2S) was measured by gas chromatography for about 45 min, and the initial production rates were used to calculate enzyme activity. Alternatively, an H_2S Clark-type microsensor (Unisense) was used. A 6.6 mM CS_2 stock was prepared by mixing 200 μl with 500 ml distilled H_2O in a 500 ml serum bottle by vigorous shaking for at least 30 min. Mutant and native enzyme extracts were diluted in 20 mM HEPES, pH 7. Depending on the activity of the enzyme, 0.5–6 μg protein was added to a 1 ml vessel containing 1 ml 20 mM HEPES, pH 7, and was stirred at 500 r.p.m. in a 50 °C water bath. CS_2 was added from the stock bottle to concentrations of 50–660 μM CS_2, and the formation of H_2S was followed for approximately 30 s. The initial H_2S production rates were calculated, and the Michaelis–Menten parameters (K_m and V_{max}) were determined by nonlinear regression using Prism 5 (GraphPad). COS activity measurements were performed by gas chromatography as described above, using 600 μl enzyme solution and 20 mM HEPES buffer, pH 7.

Peptide sequencing. The purified CS_2 hydrolase was reduced and alkylated before digestion using dithiothreitol and iodoacetamide, followed by trypsin digestion, desalting and concentration using stop-and-go elution tips[34]. Mass spectrometric *de novo* sequencing of peptides was performed on a 7 Tesla hybrid linear ion-trap Fourier transform ion-cyclotron-resonance mass spectrometer (LTQ-FT; Thermo Scientific), coupled with an 1100 series nanoflow liquid chromatograph (Agilent). Chromatographic separations were performed using a 15-cm fused silica emitter (PicoTip emitter; tip, 8 ± 1 μm; internal diameter, 100 μm; FS360-100-8-N-5-C15; New Objective) packed in-house with reverse phase ReproSil-Pur C18AQ 3 μm resin (Dr. Maisch). The peptides were loaded directly onto the analytical column over 20 min at a flow rate of 600 nl min^{-1} in 3% (v/v) acetonitrile and 0.5% (v/v) acetic acid. The peptides were eluted from the column using a 60-min linear gradient of 10% acetonitrile and 0.5% acetic acid to 40% acetonitrile and 0.5% acetic acid at a flow rate of 300 nl min^{-1}.

The mass spectrometer was operated in positive ion mode and was programmed to analyse the two most abundant ions from each precursor scan by collision-induced dissociation experiments. Precursor scans by the ion-cyclotron-resonance cell were set to cover a range of 350–2,000 m/z at a resolution power of 5×10^4 FWHM at $m/z = 400$. Each peptide ion selected for collision-induced dissociation analysis was fragmented twice in the linear ion trap (30% normalized collision energy, 30 ms activation time, activation $q = 0.250$, AGC target 3×10^4 ions, and 4 Th isolation width), and fragment ions were analysed in the ion trap and in the ion-cyclotron-resonance cell.

The fragmentation spectra from the precursor ions of m/z $[M + 2H]^{2+} = 518.2834$, $[M + 2H]^{2+} = 669.8513$, $[M + 2H]^{2+} = 733.8698$ and $[M + 2H]^{2+} = 947.5529$ were further analysed by automated *de novo* analysis using PEAKS Studio 4.2 (Bioinformatics Solutions). The settings used for *de novo* sequence interpretation included trypsin specificity with a maximum of one missed cleavage and carboxamidomethyl on cysteine residues as a fixed modification, and they allowed the oxidation of methionine as a variable modification. Mass tolerances were set to 20 p.p.m. for precursor ions and 0.6 Da for fragment ions measured in the linear ion trap or to 20 p.p.m. for fragment ions analysed in the ion-cyclotron-resonance cell. Automated *de novo* sequencing that was validated and corrected by manual interpretation resulted in the following partial sequences from the respective $[M + 2H]^{2+}$ ions: LEDYALR ($m/z = 518.2834$), VSEYLDSE ($m/z = 669.8513$), TGYVYEVETHR ($m/z = 733.8698$) and KPTEVQLDP ($m/z = 947.5529$). The peptide sequences were compared with databases at the National Center for Biotechnology Information using BLASTP. The closest homologous protein was the annotated carbonic anhydrase AAK41461 from *S. solfataricus* P2.

Cloning the *Acidianus* A1-3 CS_2 hydrolase gene. Genomic *Acidianus* DNA was isolated[35], and an internal fragment of the *Acidianus* CS_2 hydrolase gene was amplified by degenerate PCR, using the amino acid sequence obtained by the *de novo* sequencing of the four most abundant peptides, in combination with the aligned homologous gene from *S. solfataricus* P2 and the *A. brierleyi* codon usage, to design degenerate PCR primers (Supplementary Table 1). PCR was performed using the green GoTaq Flexi kit (Promega) with 200 pmol of each of the degenerate primers 1bF and 525bR, using a hot start reaction in a PCR machine (Biometra) with a gradient of annealing temperatures between 39 °C and 46 °C. Two main products, 500 base pairs (bp) and 300 bp, were obtained, and these were excised from the gel and purified using the QIAEX II Gel Extraction Kit (Qiagen). The cleaned 1bF-525bR PCR products were used as template for nested PCR with the primers 40bF and 333bR (Supplementary Table 1) at an annealing temperature of 44 °C. From both the 300-bp and the 500-bp fragments, nested PCR products of the expected 310 bp were obtained. The purified PCR products were cloned into pGEM-T Easy (Promega), and ligated plasmids were transformed into *Escherichia coli* Top 10 (Invitrogen) by heat shock. Purified plasmid DNA was sequenced with the m13F primer at the sequencing facility of the Anthropogenetics Department of the Radboud University Nijmegen Medical Centre. The sequences were analysed using Vector NTI software.

The C- and N-termini of the *Acidianus* A1-3 CS_2 hydrolase gene were cloned by inverse PCR: 10 μg genomic *Acidianus* DNA was digested in three different reactions with either EcoRI, RsaI or Sau3AI. The EcoRI digestion was incubated for 2 h before a 15-min inactivation at 80 °C. From the RsaI and Sau3AI digests, 10 μl samples were taken during incubation and were inactivated at 80 °C every 5 min up to 25 min to prevent overdigestion, and the five inactivated samples were pooled. The digested genomic DNA fragments were ligated to form circular pieces of DNA using T4 DNA ligase (Fermentas). To prevent the ligation of multiple fragments at high DNA concentrations, three ligations were performed per digest, with decreasing DNA concentrations (2, 1 and 0.4 ng μl^{-1}). Ligations were performed at 4 °C for 72 h, and the three ligations per digest were pooled. Inverse PCR was performed using the primers F4inv and R3inv (Supplementary Table 1) with Phusion High-Fidelity DNA polymerase (Finnzymes), in the supplied buffer HF according to the manufacturer's instructions. With each of the three different ligated digests, a dominant PCR product was obtained with the sizes as follows: 2.5 kilobases (kb) (EcoRI), 5 kb (RsaI) and 1.3 kb (Sau3AI). The products were cloned into pJET1/blunt (Fermentas) and sequenced using the pJET sequencing primers provided with the GeneJET PCR Cloning Kit (Fermentas). All three PCR products yielded sequence data corresponding to the *Acidianus* CS_2 hydrolase gene (the *S. solfataricus SSO1214* homologue) and the promoter region upstream, as well as the gene downstream of the CS_2 hydrolase. The amino acid sequence derived from the complete 615-bp CS_2 hydrolase gene was found to be 75% identical to the *S. solfataricus SSO1214* gene product AAK41461. It has a predicted isoelectric point of 5.92 and a molecular mass of 23,576 Da.

Cloning *S. solfataricus* P2 *SSO1214*. DNA encoding the annotated carbonic anhydrase of *S. solfataricus* P2 was cloned into pET-30a(+) (Novagen) after PCR amplification using the primers S.solf P2 CS_2 hydrolase F1, F2, R1 and R2 (Supplementary Table 1) in all four combinations as described above, using 120 ng genomic *S. solfataricus* P2 DNA and an annealing temperature of 59.3 °C. PCR products were cloned into pGEM-T Easy (Promega) according to the manufacturer's instructions and were checked by sequencing with the primer m13F (Supplementary Table 1); they were then cloned into pET-30a(+) and rechecked by sequencing with the primer T7rev (Supplementary Table 1).

Phylogenetic analysis. In total, 96 carbonic anhydrase or carbonic-anhydrase-like protein sequences were retrieved from GenBank and via BLAST. These sequences were aligned using MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/), and phylogenetic analysis was performed in MEGA 4.0 (ref. 36), using the neighbour-joining method[37], and displayed in a bootstrap consensus tree inferred from 500 replicates[38]. Branches corresponding to partitions that were reproduced in less than 50% of bootstrap replicates were collapsed. The evolutionary distances were computed using the Poisson correction method[39] and are shown in units of the number of amino acid substitutions per site. Positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (pairwise deletion option). There are a total of 1,382 positions in the final data set.

Stopped-flow spectrometry. Carbonic anhydrase activity was measured by stopped-flow spectrometry at 25 °C, using the changing pH indicator method[16], on an RX 2000 stopped-flow system (Applied Photophysics) coupled to an HP-8453 spectrophotometer. Enzyme solution containing the pH indicator phenol red (0.2 mM phenol red ($A_{557\ nm}$; Merck) in 50 mM HEPES and 50 mM $NaSO_4$ (ionic strength 0.15), pH 7.5) was mixed 1:1 with CO_2-saturated water (32.9 mM) in the reaction chamber to a total volume of 200 μl. Measurements were taken every 1 s.

Expression construct. A synthetic gene encoding the *Acidianus* CS_2 hydrolase—which was codon optimized for *E. coli* and had a G+C content of 47%, a stop codon, and a 5′ NdeI restriction site and a 3′ NotI restriction site—was obtained from GenScript and cloned into pET24b using the NdeI and NotI restriction sites. The resultant plasmid, pET24bCS_2hydsynth, was used to transform *E. coli* BL21(DE-3) cells.

Site-directed mutagenesis. Single and double base substitutions and a deletion mutant of the synthetic *Acidianus* CS_2 hydrolase gene in pET24bCS_2hydsynth were made using a QuikChange Site-Directed Mutagenesis Kit (Stratagene). Eight primers were designed with the required mutation in the centre of the primer (Supplementary Table 1). The whole plasmid was amplified using the overlapping primer sets (125 ng each), 15 ng methylated template DNA, 5 mM dNTPs and 2.5 U PfuUltra High-Fidelity DNA Polymerase (Stratagene), according to the manufacturer's instructions, with an annealing temperature of 55 °C. DpnI (10 U) was added to the PCR products to degrade the methylated template DNA at 37 °C for 2–16 h. Treated PCR products (2 μl) were transformed into *E. coli* Top10 by heat shock, and transformants were selected on LB plates containing 30 μg ml^{-1} kanamycin. Plasmid DNA was isolated from up to six clones (FlexiPrep Kit; Amersham Biosciences) and was checked by sequencing using the primer T7rev. Plasmids containing the correct mutation(s) were transformed into *E. coli* BL21 for expression studies.

Protein expression. *E. coli* BL21(DE-3) cells that had been transformed with pET24bCS_2hydsynth were grown and induced in 5–6 l LB medium for native protein production or in minimal medium supplemented with L-selenomethionine (SeMet) for SeMet-labelled protein production. Cells were collected by centrifugation at 5,500*g* for 15 min at 4 °C and resuspended in PBS. A Complete, EDTA-Free Protease Inhibitor Cocktail Tablet (Roche), 1 mM $MgCl_2$ and a spatula tip of solid DNase I were added, after which the cells were lysed by two passes through a fluidizer at 100 psi (Microfluidics). The resultant suspension was incubated in a water bath at 70 °C for 30 min and then cooled on ice. The mixture was then centrifuged at 73,000*g* for 40 min at 4 °C, and the supernatant (80 ml) was collected. This supernatant was dialysed overnight against 2 l gel filtration buffer (25 mM Bis-Tris propane and 200 mM NaCl, pH 6.5, adjusted with HCl), after which the solution was concentrated to 5 ml by ultrafiltration (Amicon, 30 kDa molecular weight cutoff; Millipore) and submitted to gel filtration on a 300 ml Superdex 200 HR column. The fractions containing pure recombinant CS_2 hydrolase were identified using 15% SDS–PAGE, pooled and concentrated, as described above, to an A_{280} of ~5.

Alternatively, *E. coli* BL21 cells containing synthetic native or mutant enzyme constructs (see below) were grown and induced in 50 ml cultures, collected as described above and lysed for 30 min at 37 °C with 2 mg ml^{-1} lysozyme followed by three rounds of bead beating (1 min at 30 Hz; Retsch) with intermittent cooling on ice. The lysates were incubated at 70 °C for 10 min to denature *E. coli* proteins and centrifuged at 13,000*g* for 10 min to remove precipitated proteins.

Crystallization and crystal handling. Plate- and box-like single crystals were obtained using the hanging-drop vapour diffusion method, mixing 2 μl protein solution with 1 μl reservoir solution containing 25% PEG 1000, 200 mM Li_2SO_4, 100 mM LiCl and 50 mM phosphate/citrate buffer, pH 4.2, followed by equilibration against this reservoir solution at 12 °C for 4 days (monoclinic crystals) or 20 °C (orthorhombic crystals). When SeMet-labelled protein was used, 5 mM

dithioerythritol was added to the set-ups. Crystals were flash-cooled in liquid nitrogen, after cryoprotection by soaking in reservoir solution containing 30% PEG 1000.

Structure determination. A 2.7 Å resolution single-wavelength anomalous dispersion data set was collected from an SeMet-labelled crystal at the X10SA beam line of the Swiss Light Source at a wavelength of 0.9790 Å (Supplementary Table 3). All data were processed using XDS[40], and care was taken to exclude data not conforming to Wilson statistics. Using SHELXD[41], a plausible solution with 60 selenium sites was found only in approximately 1 of 500 trials. The best solution was used as input into autoSHARP[42] for phase determination (phasing power for anomalous data 0.838, and $FOM_{acentric}$ 0.33 before solvent flattening), which resulted in an excellent electron density map, into which one monomer of the protein was built manually using Coot[43]. This monomer was then copied 16 times to complete the model of the asymmetrical unit. The structure was then refined using CNS[44] and REFMAC5 (ref. 45) software against the single-wavelength anomalous dispersion data set and finally against a 2.4 Å resolution data set collected at the same beam line at a wavelength of 0.98089 Å from a crystal soaked in potassium thiocyanate, using a test set defined in thin resolution shells because of the high non-crystallographic symmetry. The structure of the orthorhombic crystal form was solved by molecular replacement using PHASER[46] and refined as described above against data collected at X10SA at a wavelength of 0.97178 Å from a crystal of native protein purified from *Acidianus* that had been soaked in uranyl acetate. For the orthorhombic crystal form, eight-fold non-crystallographic symmetry restraints were applied during refinement. Using the PISA server[47], the hexadecameric structure of two interlocking octameric rings was detected both in the asymmetrical unit of the *C*2 crystals from the SeMet-labelled protein and in the primitive orthorhombic crystals of the native protein. For the monoclinic structure, the number of residues in the allowed, additional allowed and generously allowed regions of the Ramachandran plot were 89.0, 11.0 and 0.1%, respectively. For the orthorhombic structure, these values were 90.4%, 9.6% and 0.0%, respectively. No residues were in the disallowed regions of the Ramachandran plot in either structure.

Analytical ultracentrifugation. The oligomeric state of CS_2 hydrolase in solution was investigated by analytical ultracentrifugation at three concentrations ($A_{280}{}^{1\ cm} = 0.25$, 0.5 and 0.75, corresponding to 0.25, 0.5 and 0.75 mg ml^{-1}) in 25 mM HEPES/NaOH, 25 mM KCl, pH 7.5, with and without 250 mM NaCl. The protein solution was centrifuged at 130,000*g* in a ProteomeLab XL-I ultracentrifuge (Beckman) at 20 °C using 2-sector cells with a 1.2-cm optical path length. Absorption data were analysed using SEDFIT[48]. For each combination of protein and salt concentration, two species were detected: a small one, with a sedimentation coefficient of 7.8–8.5S, corresponding to 20–23% of the protein; and a large one, with a sedimentation coefficient of 13.0–13.9S, corresponding to 74–78% of the protein. For comparison, the program HYDROPRO[49] was used to estimate the sedimentation coefficients of the hexadecamer and octamer from the crystal structure. The calculated values were 14.9S and 8.8S, respectively.

SAXS. SAXS was performed at the X12SA beam line, cSAXS, at the Swiss Light Source, with the protein at $A_{280}{}^{1\ cm} = 7.0$ and 3.5, corresponding to 7 and 3.5 mg min^{-1} in 25 mM HEPES/NaOH, 25 mM KCl, pH 7.5. Solutions were added to 2-mm diameter quartz capillaries and kept at 10 °C during measurement. X-rays were used at 12.4 keV, and 200 measurements of 0.25 s each were recorded over 10 positions along the length of the capillary, which was mounted at a detector distance of 200 mm. Background measurements with the buffer only were taken using the identical capillaries, positions and measurement protocol. Data were used to a maximum momentum transfer of 0.4 Å^{-1}. Scattering curves calculated for the octameric structure by using CRYSOL[50] did not agree with the measured data, whereas calculated scattering curves for the hexadecamer were in better agreement with the measurements, although there remained a notable discrepancy at 0.08 Å^{-1}. However, fitting a linear combination of the theoretical curves for the octamer and hexadecamer using the program OLIGOMER[51] resulted in excellent correspondence after the fitting procedure converged to volume fractions of 0.8–0.9 for the hexadecamer and 0.1–0.2 for the octamer, depending on the sample.

31. Allen, M. B. Studies with *Cyanidium caldarium*, an anomalously pigmented chlorophyte. *Arch. Mikrobiol.* **32,** 270–277 (1959).
32. Vishniac, W. & Santer, M. Thiobacilli. *Bacteriol. Rev.* **21,** 195–213 (1957).
33. Derikx, P. J. L., Op den Camp, H. J. M., van der Drift, C., Van Griensven, L. J. L. D. & Vogels, G. D. Odorous sulphur-compounds emitted during production of compost used as a substrate in mushroom cultivation. *Appl. Environ. Microbiol.* **56,** 176–180 (1990).
34. Wessels, H. J., Gloerich, J., van der Biezen, E., Jetten, M. S. M. & Kartal, B. Liquid chromatography–mass spectrometry-based proteomics of *Nitrosomonas*. *Methods Enzymol.* **486,** 465–482 (2011).
35. Kowalchuk, G. A., de Bruijn, F. J., Head, I. M., Akkermans, A. D. & van Elsas, J. D. (eds) *Molecular Microbial Ecology Manual* 2nd edn, Vol. 1 (Kluwer Academic, 2004).
36. Tamura, K., Dudley, J., Nei, M. & Kumar, S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24,** 1596–1599 (2007).
37. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4,** 406–425 (1987).
38. Felsenstein, J. Confidence imits on phylogenies: an approach using the bootstrap. *Evolution* **39,** 783–791 (1985).
39. Zuckerkandl, E. & Pauling, L. Molecules as documents of evolutionary history. *J. Theor. Biol.* **8,** 357–366 (1965).
40. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26,** 795–800 (1993).
41. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D* **58,** 1772–1779 (2002).
42. Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364,** 215–230 (2007).
43. Emsley, P. & Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
44. Brünger, A. T. *et al.* Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
45. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53,** 240–255 (1997).
46. McCoy, A., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D* **61,** 458–464 (2005).
47. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372,** 774–797 (2007).
48. Schuck, P. Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78,** 1606–1619 (2000).
49. García de la Torre, J., Huertas, M. L. & Carrasco, B. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.* **78,** 719–730 (2000).
50. Svergun, D. I., Barberato, C. & Koch, M. CRYSOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28,** 768–773 (1995).
51. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. & Svergun, D. I. *PRIMUS*: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36,** 1277–1282 (2003).