Journal of Structural Biology 205 (2019) 147–154

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Using neutron crystallography to elucidate the basis of selective inhibition of carbonic anhydrase by saccharin and a derivative

Katarina Koruza[a], Brian P. Mahon[b,f], Matthew P. Blakeley[c], Andreas Ostermann[d], Tobias E. Schrader[e], Robert McKenna[f], Wolfgang Knecht[a], S. Zoë Fisher[a,g,*]

[a] Lund Protein Production Platform (LP3) & Department of Biology, Lund University, 223 62 Lund, Sweden
[b] Department of Molecular Biology, Princeton University, 08544 Princeton, NJ, United States
[c] Large-Scale Structures Group, Institut Laue-Langevin, 38042 Grenoble, France
[d] Heinz Maier-Leibnitz Zentrum (MLZ), Technische Universität München, 85748 Garching, Germany
[e] Forschungszentrum Jülich GmbH, Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), 85747 Garching, Germany
[f] Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, 32610 Gainesville, United States
[g] Scientific Activities Division, European Spallation Source ERIC (ESS), 221 00 Lund, Sweden

ARTICLE INFO

Keywords:
Joint X-ray neutron refinement
Crystallography
Hydrogen atoms
H/D exchange
Ligand protein interactions

ABSTRACT

Up-regulation of carbonic anhydrase IX (CA IX) expression is an indicator of metastasis and associated with poor cancer patient prognosis. CA IX has emerged as a cancer drug target but development of isoform-specific inhibitors is challenging due to other highly conserved CA isoforms. In this study, a CA IX_{mimic} construct was used (CA II with seven point mutations introduced, to mimic CA IX active site) while maintaining CA II solubility that make it amenable to crystallography. The structures of CA IX_{mimic} unbound and in complex with saccharin (SAC) and a saccharin-glucose conjugate (SGC) were determined using joint X-ray and neutron protein crystallography. Previously, SAC and SGC have been shown to display CA isoform inhibitor selectivity in assays and X-ray crystal structures failed to reveal the basis of this selectivity. Joint X-ray and neutron crystallographic studies have shown active site residues, solvent, and H-bonding re-organization upon SAC and SGC binding. These observations highlighted the importance of residues 67 (Asn in CA II, Gln in CA IX) and 130 (Asp in CA II, Arg in CA IX) in selective CA inhibitor targeting.

1. Introduction

Carbonic anhydrases (CAs) are zinc metalloenzymes that catalyze the reversible hydration of carbon dioxide (CO_2) to produce bicarbonate (HCO_3^-) and a proton (H^+) (Silverman & McKenna, 2007). Humans encode 15 α-class CA isoforms (Supuran, 2016). For α-class CAs, in the hydration reaction, CO_2 enters the active site cavity and binds in a hydrophobic pocket, displacing a water molecule (termed Deep Water (DW)) in a side-on orientation and permitting nucleophilic attack by a zinc-bound OH^- (Domsic et al., 2008). As the resultant HCO_3^- exits the active site, a water molecule fills the vacant space and binds to the zinc. For the next cycle of catalysis to occur, the zinc-bound water (ZW) has to be deprotonated to form OH^- with the H^+ transferred out of the active site. This proton transfer occurs via a H-bonded water network to a proton shuttling residue (in most human CAs a histidine), positioned on the edge of the active site. This histidine, in the protonated form, rotates and "flips" conformation towards the bulk solvent, delivering the H^+ to the bulk solvent (Michalczyk et al., 2015; Mikulski et al., 2013).

The membrane anchored CA IX (CA IX) is one of the 15 isoforms expressed in humans and up-regulated by hypoxia in a broad range of tumour cells, while in normal tissues its expression is low to none (Wykoff et al., 2000; Pastorek & Pastorekova, 2015). Hence, the up-regulation of CA IX is an indicator of metastasis and is associated with poor patient survival, as indicated by a meta-analysis of over a hundred independent studies (van Kuijk et al., 2016). The development of CA IX specific compounds is therefore of interest to the medical field, including ligands for use in tumour imaging, cancer detection, cancer staging, and inhibitors as therapeutic drugs to target CA IX-positive cancers (Pastorek & Pastorekova, 2015). CA IX is a multi-domain, glycosylated, membrane protein that functions as a homodimer. It contains an extra-cellular proteoglycan-like domain and a catalytic domain, which are anchored to the cell surface via a trans-membrane domain that connects to an intracellular C-terminal domain (Alterio et al.,

* Corresponding author at: Scientific Activities Division, European Spallation Source ERIC (ESS), 221 00 Lund, Sweden.
E-mail address: zoe.fisher@esss.se (S.Z. Fisher).

https://doi.org/10.1016/j.jsb.2018.12.009
Received 1 October 2018; Received in revised form 30 November 2018; Accepted 21 December 2018
Available online 11 January 2019

2009). The high degree of structural homology and sequence similarities within the active sites of the human CA isoforms makes the design and development of selective CA inhibitors (CAIs) challenging (Aggarwal et al., 2013; Alterio et al., 2012). To facilitate crystallographic studies of ligand binding to CA IX, a mimic to recapitulate the active site environment of CA IX was constructed based on CA II sequence differences. This construct termed CA IX_{mimic} is used in the work reported here. The mutations introduced to CA II to "mimic" the CA IX active site were as follows: Ala65Ser, Asn67Gln, Glu69Thr, Ile91Leu, Phe131Val, Lys170Glu, and Leu204Ala (Pinard et al., 2013).

There are several classes of CAIs (Supuran, 2008; Lomelino et al., 2016). The majority of clinically used CAIs are sulfonamides and their bioisosteres that bind directly to the zinc and inhibit catalysis by displacement of the zinc-bound solvent (McKenna & Supuran, 2014; Supuran, 2008, 2012). Recently, joint X-ray/neutron crystal structures of the clinical sulfonamide-based CAIs, brinzolamide, dorzolamide, and ethoxzolamide, in complex with CA II were determined, and demonstrated that in all three cases the sulfonamide moiety is in the deprotonated, anionic form (Fisher et al., 2012; Kovalevsky et al., 2018). Additionally, these studies showed the ethylamino group of dorzolamide to be protonated, thus having an overall neutral charge (Kovalevsky et al., 2018). The neutron structures also showed a number of H-bonded water molecules retained in the active site, connecting the CAIs to hydrophilic active site amino acids.

New classes of CAIs chemical scaffolds are being explored and some exhibit higher selectivity for CA IX over other isoforms (Lomelino et al., 2016; Supuran, 2017). Previous studies have highlighted the potential of saccharin (SAC) and SAC-based derivatives for selective inhibition towards CA IX (Fig. 1A) (Moeker et al., 2014a; D'Ascenzio et al., 2017). SAC is a cyclic secondary sulfonamide, which utilizes both the sulfonamide and lactam carbonyl group to interact with the zinc. It shows 60-fold specificity towards CA IX (K_i = 100 nM) in comparison to the ubiquitously distributed CA II (K_i = 6000 nM) (Köhler et al., 2007; Mahon et al., 2015a). Based on X-ray crystal structures of SAC binding to CA II and CA IX_{mimic}, it was not possible to ascertain which interactions conferred isoform selectivity (Fig. 1A). It was postulated that the hydrophobic face of the active site affects how SAC enters the active site with a bulky Phe131 in CA II (Val131 in CA IX) causing steric clashes and reducing binding affinity (Moeker et al., 2014a, Mahon et al., 2015a). A SAC-derivative, produced by addition of a triazole linker and sugar to the SAC (SAC-glucose conjugate (SGC)), was reported to be a better inhibitor of CA IX (K_i = 50 nM, Fig. 1A). Modification of SAC in this way led to 2-fold improvement in binding selectivity to CA IX (SGC K_i = 50 nM and SAC K_i = 100 nM) while decreasing its inhibition against CA II almost 10-fold (SGC $K_i \geq$ 50000 nM and SAC K_i = 6000 nM) (Moeker et al., 2014b). In addition to SGC binding to the metal ion, the inhibitor also interacts with regions around the active site, contributing to isoform-specific inhibition. For designing new or improved isoform-specific inhibitors, it is crucial to have more detailed structural information about ligand binding interactions. Such information includes how ligand binding influences H-bonds and water displacement and re-organization. Neutron protein crystallography offers the only unambiguous approach for locating hydrogen atoms (H) and can contribute knowledge that feeds back into the drug design process, leading to ligands with improved characteristics.

Approximately half of the atoms in proteins are H atoms and they play a crucial role in protein folding, enzyme catalysis, water coordination, and ligand binding interactions. However, it is very challenging to determine H positions using X-ray scattering alone, due to the small atomic Z number (H has 1 electron). In neutron scattering, deuterium (D) has a strong positive coherent scattering length (6.67 fm) in comparison with H (−3.74 fm), and a significantly smaller incoherent component: ~2.05 barn vs. 80.3 barn for D and H, respectively. To benefit from these intrinsic scattering properties and maximize the signal-to-noise ratio in crystallography, H is often exchanged with D in proteins to increase the signal-to-noise ratio, reduce background, and to obtain better nuclear scattering length density maps from neutron protein crystallography experiments (Blakeley et al., 2015).

Based on the neutron's unique sensitivity to isotopes, it is possible to "see" numerous features involved with ligand binding to proteins that X-rays alone cannot reveal. These include the charged state of ligands, protonation and H-bonding capacity of amino acid residues, water networks, H-bonds, and how water molecules may mediate ligand binding, and the possibility to discern solvent species (H^+ vs OH^- vs

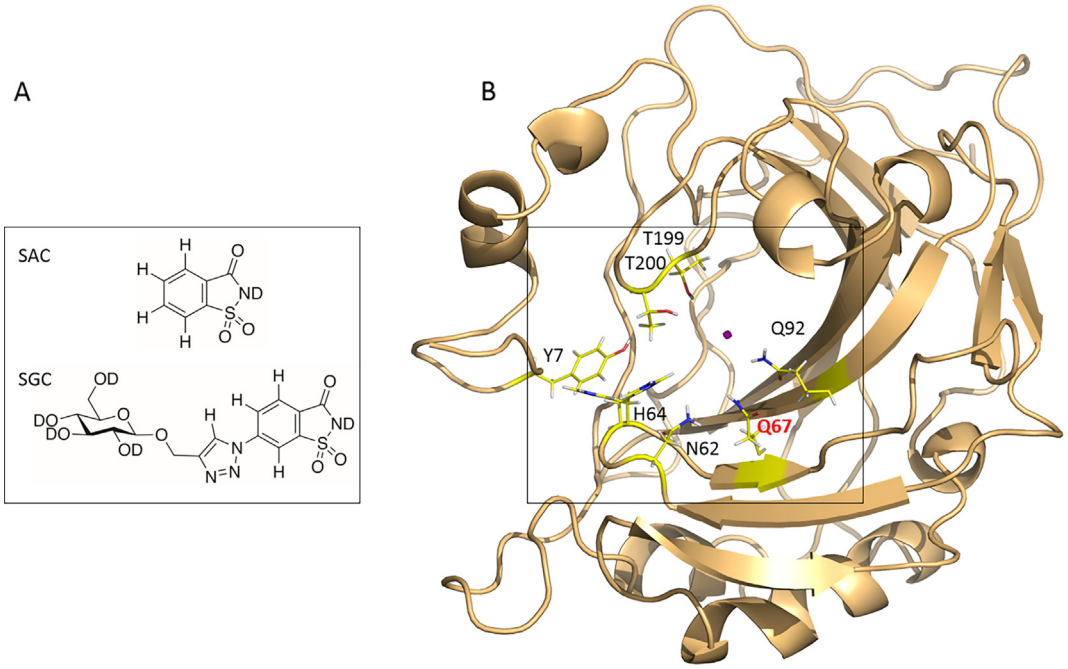

Fig. 1. (A) Structure formulae of SAC and SGC. H and exchanged D atoms are indicated, (B) Ribbon diagram of CA IX with active site residues (stick) and zinc (pink sphere) shown. Mutated residue Q67 is labeled in red, other residues that are conserved in CA II are labeled in black.

H_2O vs H_3O^+) (Weber et al., 2013; Kovalevsky et al., 2011; Ikeda et al., 2017). The SAC and SGC used in this study have both exchangeable and non-exchangeable H atoms (Fig. 1A). For neutron studies it is expected that –CH groups may show signal cancellation due to the negative troughs created by H atoms, while exchanged –OD, –SD or –ND groups should show positive neutron scattering due to the scattering length of D.

Previous work using X-rays alone did not reveal the level of molecular detail needed to understand the different inhibition constants for SAC and SGC to CA II and CA IX_{mimic} (Mahon et al., 2015a). For the structural studies presented in this article a CA IX_{mimic} construct was used. The original construct was designed by Genis *et al.* and further modified by Pinard *et al.* (Genis et al., 2009; Pinard et al., 2013). Briefly, CA IX_{mimic} was produced via site-directed mutagenesis of the active site of CA II to generate a structural and functional analogue to the native CA IX active site (Fig. 1B, Fig. S1). To investigate SAC and SGC binding the joint X-ray/neutron crystal structures for three different forms were determined: CA IX_{mimic} (no ligand), CA IX_{mimic}:SAC, and CA IX_{mimic}:SGC. The structures reveal the details of inhibitor binding to the protein, associated water displacement, and a detailed view of the re-organization of H-bonds in the CA IX active site in response to inhibitor binding. Using this approach provided insight as to how these inhibitors exhibit CA IX selective inhibition.

2. Experimental section

2.1. CA IX_{mimic} expression and purification

The CA IX_{mimic} coding region was codon optimized for expression in *E. coli* and cloned into pET-26b (+) (Novagen) using NdeI/XhoI by Genscript. Active site mutations of CA II in the CA IX_{mimic} include Ala65Ser, Asn67Gln, Glu69Thr, Ile91Leu, Phe131Val, Lys170Glu, and Leu204Ala (Pinard et al., 2013) as marked on (Fig. S1) and give an active site that is approximately ~95% homologous to wild type CA IX. Inhibitor interactions that occur in CA IX_{mimic} should be very similar to the interaction in wild type CA IX (Mahon et al., 2015b). CA IX_{mimic} was expressed in *E. coli* BL21(DE3) cells grown in LB Broth, Miller (Difco) at 37 °C in the presence of kanamycin (50 μg/mL). Upon reaching an OD_{600} of 1, IPTG was added to the culture at 1 mM final concentration to induce protein expression in the presence of 1 mM $ZnSO_4$. After 3 h the cells were harvested by centrifugation at 6000 × *g* and cell pellets were frozen at −20 °C until lysis. The pellets were thawed and diluted in wash buffer 1 (0.2 M Na_2SO_4, 100 mM Tris, pH 9.0) before passed dropwise twice through the French press cell at 20 000 psi. Cell lysate was clarified by centrifugation at 14000 × *g* for 60 min at 4 °C prior to purification by affinity chromatography using p-amino-methyl-benzenesulfonamide (Sigma Aldrich). Cell lysate was loaded on the column and non-specifically bound proteins were washed off by 10 column volumes (CV) of wash buffer 1. The column was then washed with wash buffer 2 (0.2 M Na_2SO_4, 100 mM Tris, pH 7.2). Bound protein was eluted from the column with elution buffer (50 mM Tris pH 7.8, 0.4 M NaN_3). The eluted protein was concentrated with Amicon Ultra Centrifugal Filter Units (Merck) with weight cut-off 10 kDa filters and subsequently purified by size exclusion chromatography HiLoad 26/600 Superdex 75 (GE Healthcare Life Sciences) in 50 mM Tris pH 8.5. Fractions eluted from affinity chromatography and size exclusion chromatography were analyzed by sodium dodecyl sulfate − polyacrylamide gel electrophoresis (SDS − PAGE) to assess protein purity and homogeneity (Fig. S2A). Elution fractions containing target protein were concentrated to 20 mg/mL by using Amicon Ultra Centrifugal Filter Units (Merck) with molecular weight cut-off 10 kDa.

2.2. Crystallization

Crystallization of purified CA IX_{mimic} was carried out in a sitting drop in siliconized glass plate Sandwich Box setup (Hampton Research). This technique allows to set up larger crystallization drops, in this case 100–250 μL by mixing 1:1 ratios of reservoir solution (1.4 M sodium citrate, 0.1 M Tris pH 8.5) and 20 mg/mL CA IX_{mimic} and 5 μL of additive (No.2 from Detergent Screen #HR2-410, Hampton Research), against 25 mL of reservoir solution. Crystal trays were sealed and incubated at 20 °C to undergo vapour diffusion over a period varying from 1 to 3 months. Crystals larger than 1 mm^3 typically took ~6 weeks to grow (Fig. S2B-C) and were then used to prepare complexes with ligands (Fig. S2D-E).

SAC (Sigma product No 240931) stock solution was prepared by dissolving the powder in reservoir solution to the final concentration of 20 mM. The ligand solution was then injected in 5–10 μL batches into a drop containing large CA IX_{mimic} crystals to the final concentration of 20 mM. Crystals were incubated for several weeks at 20 °C to allow soaking of the ligand into the crystal and then mounted in a quartz capillary with D_2O mother liquor for X-ray and neutron data collection SAC-glucose-conjugate-inhibitor ([1-(1,1,3-trioxo-2H-1λ^6,2-benzothiazol-6-yl)-1,2,3-triazol-4-yl]-methyl-β-D-glucopyranoside) was synthesized by Red Glead Discovery (Lund, Sweden) according to a published procedure (Moeker et al., 2014b). We refer to this inhibitor as saccharin-glucose-conjugate (SGC). Despite numerous efforts, simple soaking used for SAC did not work for SGC. Instead we used the dry co-crystallization method as described by Gelin et al. (Gelin et al., 2015). It is challenging to move large drops together with crystals and it did result in some mechanical damage (Fig. S2D-E), but the crystals survived well enough and still gave very good diffraction (Table S3).

A stock solution of 20 mM inhibitor SGC was prepared in water and 150 μL of stock solution was dried overnight at room temperature in a well of a siliconized glass plate in the Sandwich Box setup (Hampton Research). Large CA IX_{mimic} crystals were selected by visual inspection and moved together with the crystallization drop (~100–150 μL) in the dry well against fresh 25 mL of 1.4 M Na-Citrate in reservoir solution. Crystals were incubated for several weeks at 20 °C prior X-ray and neutron data collection.

Large single crystals (0.7–1.8 mm^3) were mounted in a 3 mm diameter quartz capillaries as described before (Fisher et al., 2012) and allowed to undergo hydrogen/deuterium-exchange of labile H atoms inside the capillary (Fig. S2C). According to a report by Bennett and colleagues, up to 80% of the amide backbone exchanges in four weeks (Bennett et al., 2008). All described crystals were allowed to exchange for at least 6 weeks before transportation to X-ray and neutron facilities. Smaller crystals (0.05–0.10 mm^3) for room temperature X-ray diffraction experiments were prepared in exactly the same way as the neutron samples.

Neutron and X-ray diffraction data sets of H/D exchanged large crystals were collected at room temperature, to resolutions ranging from 1.2 to 2.0 Å. All data collection and model refinement statistics are given for unbound CA IX_{mimic} (Table S1), CA IX_{mimic}:SAC (Table S2), and CA IX_{mimic}:SGC (Table S3).

2.3. X-ray and neutron data collection and processing

X-ray diffraction data from the CA IX_{mimic} crystal were collected at room temperature (RT) on the I911-2 Cassiopeia beamline at MAX lab (Lund, Sweden). The CA IX_{mimic}:SAC complex data were collected at room temperature on the I911-3 beamline at MAX lab. Diffraction data for the CA IX_{mimic}:SGC complex were collected at RT on the BM-30/FIP beamline at ESRF (Grenoble, France). X-ray data collection, data set statistics are shown in Table S1-3.

RT neutron diffraction data on the CA IX_{mimic} crystal were collected using the LADI-III instrument at ILL (Blakeley et al., 2015). As LADI-III is a quasi-Laue instrument, only 28 images were collected. The CA IX_{mimic}:SAC and CA IX_{mimic}:SGC data were collected on the Biodiff beamline at the FRM-II reactor source (Garching, Germany). Neutron data collection and data set statistics are shown in Table S1-3. All X-ray

data sets were processed and scaled using XDS and XSCALE (Kabsch, 2010). Neutron data collected at Biodiff were processed with either HKL2000 (v.705b) (Otwinowski & Minor, 1997). Neutron data collected at LADI-III were processed using LAUEGEN (Campbell, 1995). LSCALE was used to determine the wavelength-normalization curve using intensities of symmetry-equivalent reflections observed at different wavelengths and the data were merged in SCALA (Evans, 2006).

2.4. Joint X-ray and neutron refinement

The models were initially refined against the X-ray data alone in phenix.refine (Adams et al., 2011) using the previously determined CA IX_{mimic} crystal structure as the starting model (PDB entry 4riv). All waters, metals, and ligands were removed prior to starting iterative rounds of refinement. Once the R-factors converged and all metals, solvent, and ligands were properly placed, H and D atoms were generated in the (X-ray only) refined model.

The next phase was joint X-ray and neutron refinement in phenix.refine and involved several rounds of maximum-likelihood-based refinement of individual coordinates, individual B-factors and occupancies of H/D atoms against both neutron and X-ray experimental data sets. After every round of joint refinement the model was visually inspected and manipulated in Coot using a series of F_o-F_c and $2F_o$-F_c positive electron density maps and both positive and negative F_o-F_c and $2F_o$-F_c nuclear scattering length density maps to guide modeling of solvent and protein H/D atoms (Emsley & Cowtan, 2004). Extra care was taken with waters as the D atom positions could be ambiguous and the maps were often challenging to interpret. Water oxygen atoms were initially placed based on omit density in the electron density maps and subsequently refined until the R-factors converged. In preparation for joint refinement, all waters are generated as D_2O molecules. During joint refinement the water molecule positions were manually inspected, aided by successive omit map generation. Water molecules (as D_2O) were manually oriented as phenix.refine can place them in random orientations (i.e. not in chemically possible positions). After modeling into omit F_o-F_c maps, new $2F_o$-F_c maps are then intermittently calculated to check that the placement was sensible. The same approach was used in a consistent way for all three structures reported here. Final model statistics are shown in Table S1-3. Figures were generated with Pymol (Schrödinger, 2015).

3. Results & discussion

3.1. Crystallography

All three structures were determined from joint refinement against room temperature X-ray and neutron diffraction data and a brief summary is shown in Table 1. Model coordinates and experimental data were deposited with the Protein Data Bank under accession codes PDB: 6fji for CA IX_{mimic} structure; 6fjj for CA IX_{mimic}:SAC; 6gcy for CA IX_{mimic}:SGC. The space group for all six crystals was $P2_1$ with mean unit cell dimensions $a = 42.6$, $b = 41.9$, $c = 72.8$ Å, $\beta = 104.1$. The structures were highly homologous with least squares superposition for all atom positions across the three structures having an rmsd of ~0.8.

3.1.1. Unbound CA IX_{mimic}

Fig. 2 shows the active site of the joint X-ray/neutron (X/N) structure of unbound CA IX_{mimic} with amino acid side chains, solvent, and H-bonds indicated. Within the active site of unbound CA IX_{mimic} we observed an organized H-bonded water network that was homologous to that previously seen in CA II neutron and CA IX_{mimic} X-ray crystal structures (Genis et al., 2009; Fisher et al., 2011). The DW is present, as was observed in CA II (Domsic et al., 2008). The network of ordered waters extends from the catalytic ZW to the proton shuttling residue, His64. A network of H-bonds can connect the ZW to His64 and represent four possible pathways for protons to leave the active site: ZW-W1-W2, ZW-W1-W2-W3a, ZW-W1-W4-W3b, or ZW-W1-W2-W3b (Fig. 2). The structure reveals most of the positions of D atoms on the water molecules, allowing detailed mapping of thirteen observed H-bonds between waters and water to side-chains. Five hydrophilic amino acid residues are involved in coordinating the water network: Tyr7, Gln67, Gln92, Thr199, and Thr200. The least ordered waters are W3b, W4, and W5 with poorly defined nuclear scattering length density and high crystallographic B-factors (~34 $Å^2$ vs. ~14 $Å^2$ for the other active site waters).

There are some key similarities and differences in how the waters are H-bonded to each other and the CA IX_{mimic} compared to the pH 7.8 neutron structure of CA II (Fisher et al., 2011). Of note, DW, ZW, W1 are in the same orientations as their counterparts in CA II. ZW acts as H-bond donor to OG1 of Thr199. Whereas, W1, while in the same general orientation as observed in CA II, is now a H-bond acceptor from Thr200, the opposite to CA II (Fig. 2). Also, W2, W3a, and W3b are in different orientations and thus make different H-bonds compared to CA II.

Glutamine at position 67 (Gln67) (position 67 is an asparagine residue (Asn67) in CA II) is observed as a H-bond donor to W3b (Fig. 2), and unlike CA II, there is no H-bond between W3b and Asn62 (Fisher et al., 2011). Kinetic measurements of k_{cat}/K_M ($M^{-1}s^{-1}$) show about a 2-fold difference between CA II and the native CA IX domain, at $1.2 \times 10^{-8}\,M^{-1}s^{-1}$ and $0.55 \times 10^{-8}\,M^{-1}s^{-1}$, respectively. The proton transfer rate constant k_B (s^{-1}) for CA II and CA IX is also about 2-fold different at $0.8 \times 10^{-6}\,s^{-1}$ and $1.4 \times 10^{-6}\,s^{-1}$, respectively (Mahon et al., 2016). The measured differences in kinetic parameters could be due to the presence of Gln67 (Asn67 in CA II) and the associated alternative organization of the H-bond donors/acceptors around the Zn center, specifically W1.

His64, located on the edge of the active site, is the proton shuttling residue during catalysis and observed in numerous X-ray crystal structures to be alternating between two conformations (Tu et al., 1989; Nair & Christianson, 1991). Likewise, in CA IX_{mimic} His64 is observed in both an "in" and "out" conformation, with the "in" conformation slightly favored with occupancy refined to ~0.6 "in" vs. 0.4 "out". Due to the dual conformation, it is not possible to unambiguously orient the imidazole ring around $\chi 2$ angle or to observe the protonation state of the "in" or the "out" form.

3.1.2. CA IX_{mimic}:SAC complex

SAC binds to CA IX_{mimic} primarily by coordination of the deprotonated N atom to the active site zinc, similar to the sulfonamide-based CAIs, acetazolamide and dorzolamide, to CA II in neutron structures

Table 1
Summary of crystal structures presented in this work. For complete crystallographic statistics see Supplemental Information (Tables S1–S3).

	CA IX_{mimic}	CA IX_{mimic}:SAC	CA IX_{mimic}:SGC.
PDB ID	6fji	6fjj	6gcy
Resolution X/N	29.00–1.60 (1.70–1.60) 40.00–2.00 (2.11–2.00)	35.00–1.20 (1.23–1.20) 26.90–2.00 (2.07–2.00)	36.00–1.30 (1.40–1.30) 29.00 – 2.00 (2.06 – 2.00)
Source X/N	MAX lab (Lund, SE) ILL (Grenoble, FR)	MAX lab (Lund, SE) FRM-II (Garching, DE)	ESRF (Grenoble, FR) FRM-II (Garching, DE)
R_{cryst}/R_{free} X	12.9/16.1	13.5/16.5	14.8/15.5
R_{cryst}/R_{free} N	18.5/22.2	20.4/21.7	19.1/21.4

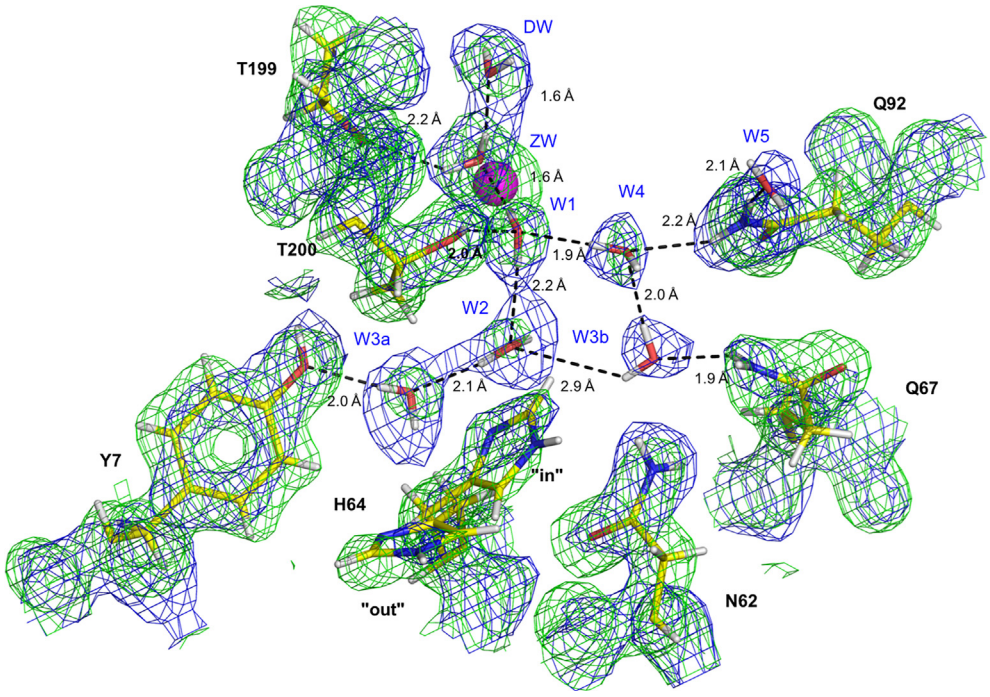

Fig. 2. Stick representation of unbound CA IX_{mimic} active site. Observed H-bonds are shown as black dashed lines. The $2F_o$-F_c nuclear density map is contoured at 1.5σ and is shown in blue mesh. The $2F_o$-F_c electron density map is contoured at 1.0 σ for all residues and 1.5 σ for water molecules and is in green mesh. Residues and solvent are as labeled, Zn^{2+} is represented as a magenta sphere.

(Fig. 3A) (Fisher et al., 2012; Kovalevsky et al., 2018). The –NH group in SAC has a pK_a of ~1.3 and the observed anionic form is expected at the crystallization pH of 8.3 (Pearson, 2001). The O8 of SAC is an H-bond acceptor from Thr200 and forms a H-bond network involving Tyr7, W2, and W3a (Fig. 3A-B). In the unbound CA IX_{mimic} structure, Thr199 is an H-bond acceptor from ZW with an O…O distance of 2.8 Å and an H-bond donor to Glu106. SAC binds as an anion and the cyclic sulfonamide N to Thr199 OG1 distance is 3.3 Å. Thr199 is engaged with Glu106 as H-bond donor. The side chain of Gln67 has flipped and there was a compensatory rearrangement of W3b to maintain the H-bond. The loss of W4 also facilitates the re-orientation of W3b that is now only weakly coordinated by Gln67 and Gln92, with corresponding weak density (Fig. 3). The modeling of W3b is further complicated as there is no electron density for it and we were only guided by the weak nuclear density. The aromatic ring of SAC packs against a hydrophobic patch in the CA IX_{mimic} active site composed of Val121, Val 141, and Leu198 (Fig. S3).

SAC binding displaces four ordered water molecules: DW, ZW, W1, and W4 (Figs. 2 and 3). The C1, C7, N9, and S10 atoms of the SAC molecule overlap closely with the oxygen atoms of the four displaced water molecules. The binding of SAC results in a net loss of six H-bonds between the water and protein while seven H-bonds are retained “as is” or through some re-organization of water to accommodate SAC. Water molecules W3a, W2, W3b, and W5 are still in place, with W3b displaying the most disorder. His64 retains its mobility with about 0.5 occupancy for both the “in” and “out” conformations. This is unusual

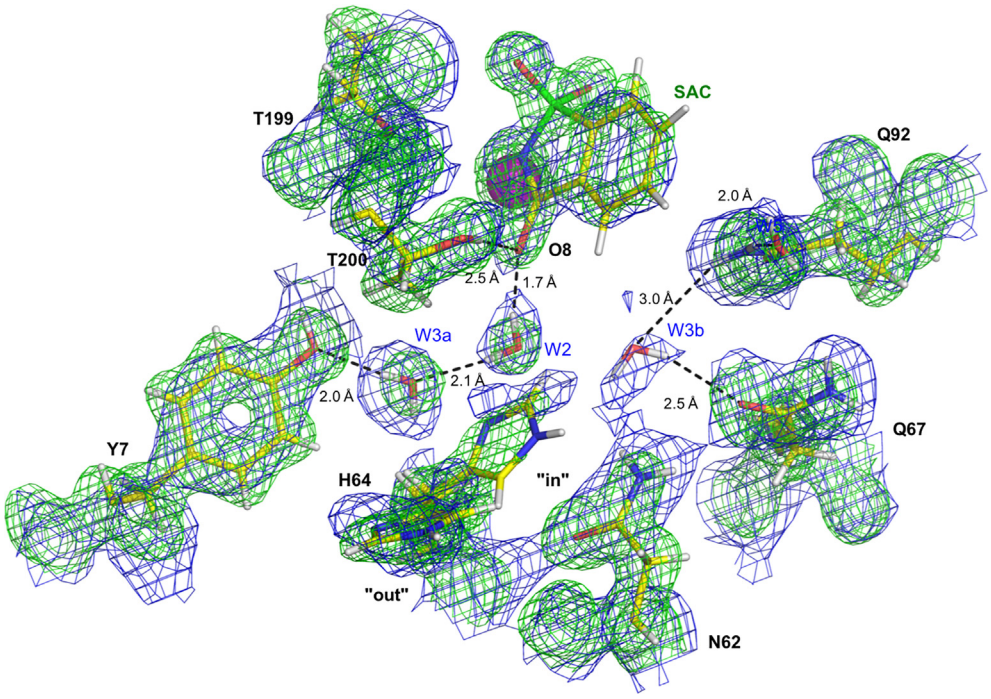

Fig. 3. Stick representation of CA IX_{mimic}:SAC complex. Observed H-bonds are shown as black dashed lines. (A) The $2F_o$-F_c nuclear density map is contoured at 1.5σ and is shown in blue mesh. The $2F_o$-F_c electron density map is contoured at 1.5σ and is in green mesh. (B) Close-up of SAC interactions to Thr200. The $2F_o$-F_c nuclear density map is contoured at 1.0 σ for His and 1.5 σ for water and ligand molecules and is in blue mesh. Residues and solvent are as labeled, Zn^{2+} is represented as a magenta sphere.

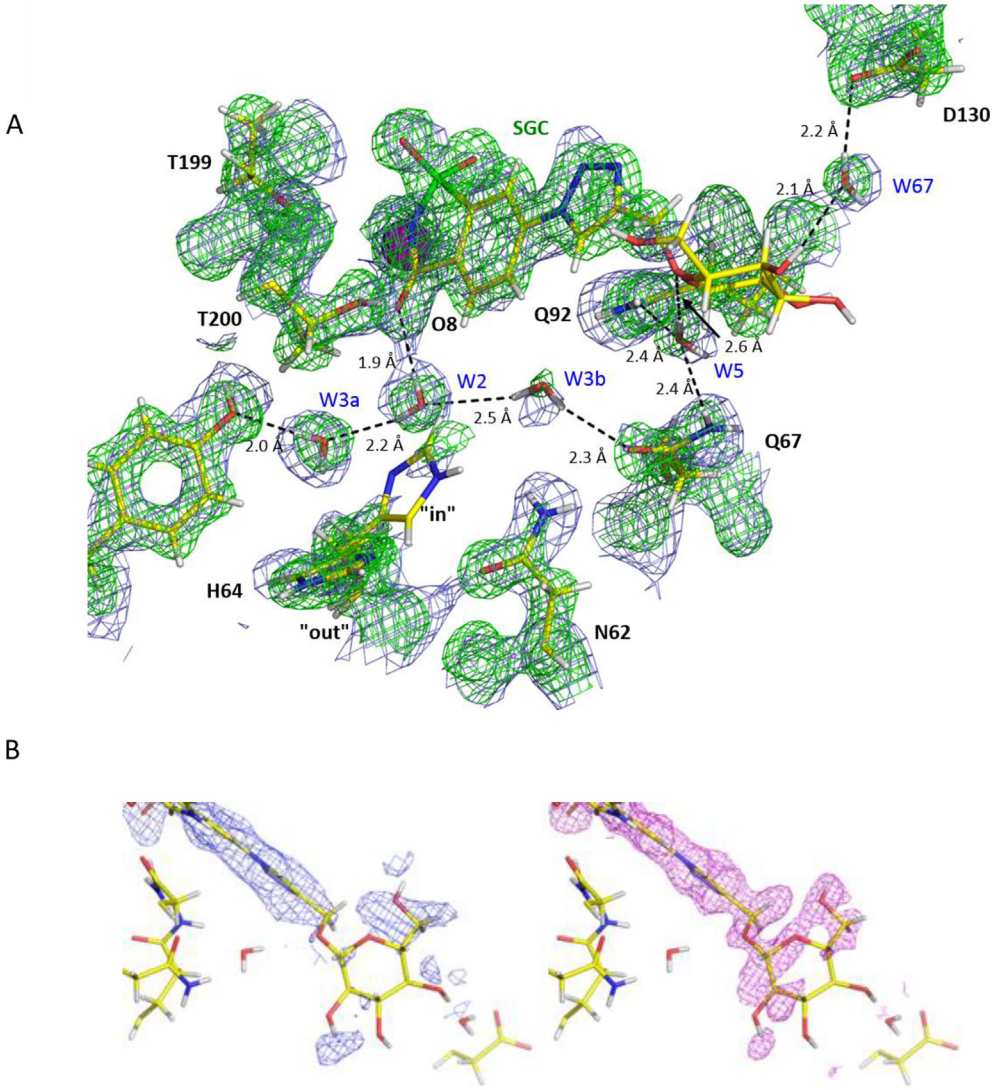

Fig. 4. Stick representation of CA IX_{mimic}:SGC complex. Observed H-bonds are shown as black dashed lines. (A) The $2F_o$-F_c nuclear density map is contoured at 1.5σ and is shown in blue mesh. The $2F_o$-F_c electron density map is contoured at 1.5σ and is in green mesh. (B) Electron and nuclear density maps for the glucose moiety of SGC. The $2F_o$-F_c nuclear density map is shown in blue and is contoured at 0.75 σ. The $2F_o$-F_c electron density map is shown in magenta and is contoured at 0.75 σ. Residues and solvent are as labeled, Zn^{2+} is represented as a magenta sphere.

compared to other inhibited CA structures where His64 is fully in the "out" conformation (Fisher et al., 2012; Kovalevsky et al., 2018). Hence, while SAC binding displaces the catalytic ZW and inhibits activity, parts of the active site have the same organization as without SAC.

3.1.3. CA IX_{mimic}:SGC complex

The SAC moiety of SGC binds in a slightly different orientation to that observed for SAC alone, but still with the deprotonated N coordinating to the zinc (Fig. 4A, Fig. S4). SGC makes three indirect interactions for a total of ten H-bonds to amino acid residues that are present in the SGC binding site. Similar to SAC binding, O8 is H-bonded to Tyr7 via W2 and W3a. But unlike the SAC complex, for SGC there is no H-bond between O8 and Thr200. Instead, Thr200 makes a water-mediated H-bond to the backbone of neighboring residue Pro201. This is most probably due to the difference in position of the SAC-moiety of SGC compared to SAC, that increases the distance between H-bond donor (O28 of SGC) and acceptor (OG1 of Thr200), from 3.3 to 3.5 Å, respectively.

The higher X-ray resolution (1.3 Å) structure compared to previous studies (1.65 Å, PDB ID 4RUI) does not reveal ordered density for the glucose group and therefore its flexibility makes it challenging to model (Fig. 4B; Mahon et al., 2015b). At 0.75σ the $2F_o$-F_c electron density map provides partial density for half of the glucose moiety, guiding placement (Fig. 4B). The residues and water molecules around the glucose group have well-defined nuclear and electron density peaks, somewhat limiting the area where the glucose moiety may be located. The triazole linker does not induce any new H-bonds, but based on how the glucose is modeled, O12 of the glucose group is in a position to make interactions with Gln67 through H-bonds with W5 (Fig. 4A). There are also a putative H-bonds between O19, W67, and Asp130 (Fig. 4A). Similar to the SAC structure, DW, ZW, W1, and W4 are displaced by SGC, while W2, W3a, W3b, and W5 remain in the active site. W3b and W5 have changed orientations and W5 has moved ~2.5 Å compared to its position in the unbound CA IX_{mimic} structure (Fig. 5B). His64 is still seen in two conformations but now with 0.3 occupancy for the "in" conformation. This reduced "in" occupancy could be due to the presence of the bulkier SGC molecule or slight water rearrangements. These factors together could also be affecting the pK_a of His64, causing the protonated form (i.e. "out" conformation) to be slightly more favoured. As in the unbound and in the CA IX_{mimic}:SAC complex, the data is very weak for this mobile His residue and it was modeled as neutral in both "in"

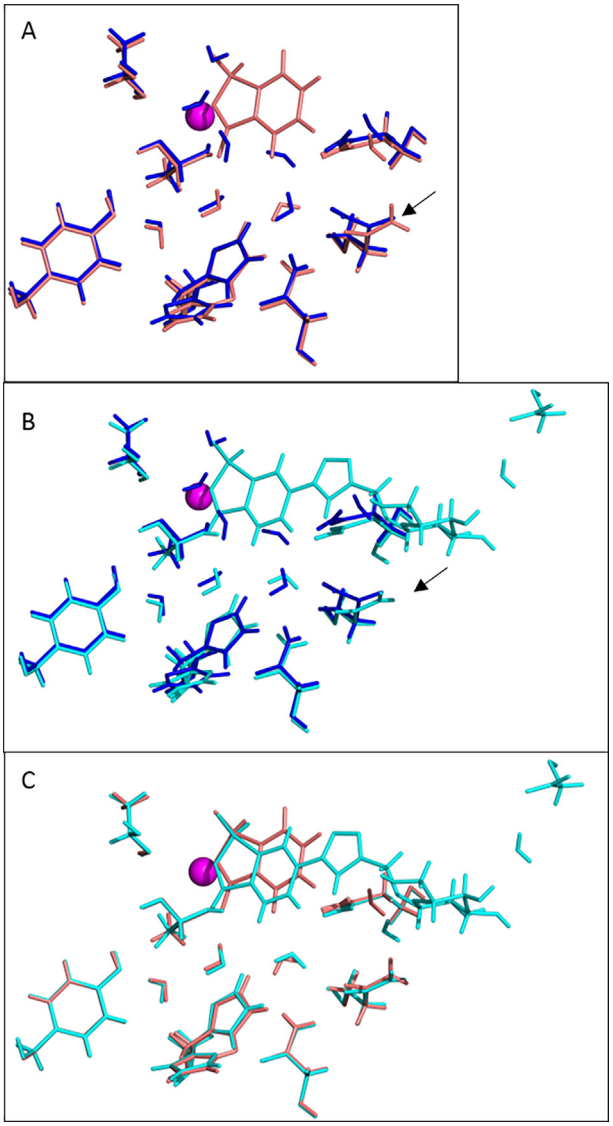

Fig. 5. Overlays of the active site of unbound CA IX_{mimic} and in complex with SAC and SGC show both similarities and changes that occur in the crystal structures upon ligand binding. (A) Unbound CA IX_{mimic} in blue and CA IX_{mimic}:SAC in pink, (B) Unbound CA IX_{mimic} in blue, CA IX_{mimic}:SGC in cyan, (C) Overlay of the SAC (pink) and SGC (cyan) complexes. The black arrows indicate the flipping of the Gln67 side chain.

and "out" conformations.

3.2. SAC and SGC isoform selectivity

The coordination of sulfonamide N to the zinc and the binding of SAC, alone and in SGC, against the hydrophobic face (Val121, Val141, and Leu198) of the active site is similar for the two ligands. Ligand and active site desolvation are known drivers for favorable binding interactions and here we observe the loss of four water molecules for both SAC and SGC, and the associated reorganization of existing H-bonds and formation of new ones. There is a network of seven H-bonds that mediate the binding of SAC compared to as many as ten H-bond involved with SGC binding. Due to the uncertainty of the glucose moiety position it could be fewer interactions mediating overall SGC binding.

The binding interactions between SAC and CA II and native CA IX should be the same as Thr200 and Tyr7, the residues involved with direct and water-mediated H-bonding, are conserved in both isoforms. Similarly, the hydrophobic patch where the aromatic ring of SAC sits is composed of Val121, Val141, and Leu198 and these residues are also conserved in CA II and native (or mimic) CA IX (Fig. 5C; Fig. S3). The only different amino acid between CA II and CA IX_{mimic} in the SAC binding area is Gln67. The side chain of Gln67 moves in response to SAC binding, leading to remodeling of H-bonds and water orientations (Fig. 5A-B). Hence, the presence of a Gln at position 67 most likely plays a defining role in explaining the difference in measured K_i for SAC between CA II and native CA IX.

In contrast to the SAC binding site, for SGC binding, there are some notable differences between CA II and native CA IX in the residues involved. As noted before, the glucose moiety is not well ordered as is apparent in the $2F_o$-F_c electron and nuclear density maps shown in Fig. 4B. Based on the modeling of SGC, the glucose group can be situated close to Leu91 (Ile91 in CA II) and Val131 (Phe131 in CA II) (Fig. S4). The presence of these two bulky hydrophobic residues in CA II would cause steric clashes with SGC and could explain why SGC binds so poorly to CA II compared to native CA IX. Asp130 is Arg130 in native CA IX and an Arg in that position could make a direct H-bond to SGC, possibly leading to improved binding to CA IX. Furthermore, in the active site of CA IX_{mimic}, residue Gln67 is responsible for an additional water-mediated H-bond to SGC. The collective effects of the additional H-bonds and less bulky hydrophobic residues in native CA IX that are different to those in CA II, could underlie the observed isoform discrimination by SGC.

Although SAC and SGC share the same chemical SAC group, it is very interesting to observe that they have different interactions with Thr200, probably due to a different position of SAC compared to SAC-moiety in SGC. This could be due to the presence of the additional triazole linker and glucose group and interactions that glucose may make to the protein. These interactions could cause the slight shift of the SAC moiety observed compared to SAC alone (Fig. 5C). Taken together, these interactions explain the modest 2-fold difference in K_i of 100–50 nM for SAC and SGC towards CA IX.

The selectivity of SAC for native CA IX over CA II is difficult to interpret as only Gln67 and some water positions are different between the isoforms. Perhaps this difference is strong enough for the 60-fold difference in K_i or there are some other effects not observable in static crystal structures. From the bound SGC structure there are several H-bond interactions involving CA IX_{mimic} residues (Gln67, Asp130) in addition to hydrophobic interactions (Leu91, Val131). These interactions observed in the CA IX_{mimic}:SGC complex could explain the observed selectivity between CA isoforms.

4. Conclusions

This study described the jointly refined, room temperature X-ray and neutron crystal structures of CA IX_{mimic} alone and in complex with two isoform selective CAIs, SAC and SGC. The differences in how SAC and SGC bind to CA IX_{mimic} reported here contribute to our understanding for the observed differences in K_i to native CA IX. This approach also can give clues as to the basis for isoform specificity of these compounds towards native CA IX over CA II.

The structures analyses based on high resolution X-ray and medium resolution neutron data provided insights to how SGC can be improved to capitalize on the observed H-bonds. In the same way the Arg at position 130 in native CA IX may be able to directly H-bond to SGC, it is possible to modify SGC so it can directly H-bond to Gln67 and displace W5. The triazole linker is situated very close to the hydrophobic pocket that interacts with the SAC and glucose moieties. Perhaps the linker can be methylated or replaced with a more hydrophobic group to encourage hydrophobic interactions along that side of the active site. The glucose disorder indicates that there is enough room to add bulkier groups or more sugars that can fill this area. Unfortunately, the glucose disorder renders the conclusions speculative and further work is necessary to explore the application of neutrons in drug design for this target. These data and analyses do indicate a proof-of-concept that neutrons can

contribute unique information on inhibitor binding and can give clues as to which specific interactions drive binding and selectivity. In addition, having experimental coordinates for the H atoms in a protein structure can be useful to computational chemistry and improve drug docking studies.

Acknowledgements

The authors would like to thank Esko Oksanen for useful discussions and Anna Leung for help with chemical drawings. The authors would also like to thank the Institut Laue-Langevin for provision of beamtime on LADI-III, and MAX lab and ESRF beamline scientists for expert assistance. We thank Red Glead Discovery for the custom synthesis of the SGC. We also need to thank the Lund Protein Production Platform staff for providing technical support for many of the experiments. The authors declare no competing interests.

Funding information

The authors thank the Integrated Infrastructure Initiative No. 262348 European Soft Matter Infrastructure, ESMI for providing beam time at the instrument BIODIFF at the Heinz Maier-Leibnitz Zentrum (MLZ). We thank Lund University, the Royal Physiographic Society of Lund, Interreg/MAX4ESSFUN, and The Crafoord Foundation for financial support. This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 654000.

Appendix A. Supplementary data

PDB references: The atomic coordinates and structure factors for the reported crystal structures have been deposited in the Protein Data Bank under accession codes PDB: 6fji for CA IX_{mimic} structure; 6fjj for CA IX_{mimic}:SAC; 6gcy for CA IX_{mimic}:SGC. Authors will release the atomic coordinates and experimental data upon article publication.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsb.2018.12.009.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Echols, N., Headd, J.J., Hung, L.W., Jain, S., Kapral, G.J., Grosse Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R.D., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2011. Methods 55, 94–106.
Aggarwal, M., Kondeti, B., McKenna, R., 2013. Bioorgan. Med. Chem. 21, 1526–1533.
Alterio, V., Di Fiore, A., D'Ambrosio, K., Supuran, C.T., De Simone, G., 2012. Chem. Rev. 112, 4421–4468.
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., 2009. Proc. Natl. Acad. Sci. U.S.A. 106, 16233–16238.
Bennett, B.C., Gardberg, A.S., Blair, M.D., Dealwis, C.G., 2008. Acta Crystallogr. D 64, 764–783.
Blakeley, M.P., Hasnain, S.S., Antonyuk, S.V., 2015. IUCrJ 2, 464–474.
Campbell, J., 1995. J Appl. Crystal. 28, 228–236.
Domsic, J.F., Avvaru, B.S., Kim, C.U., Gruner, S.M., Agbandje-McKenna, M., Silverman, D.N., McKenna, R., 2008. J. Biol. Chem. 283, 30766–30771.
D'Ascenzio, M., Guglielmi, P., Carradori, S., Secci, D., Florio, R., Mollica, A., Ceruso, M., Akdemir, A., Sobolev, A.P., Supuran, C.T., 2017. J. Enzyme Inhib. Med. Chem. 32, 51–59.
Emsley, P., Cowtan, K., 2004. Acta Crystallogr. D 60, 2126–2132.
Evans, P.R., 2006. Acta Crystallogr. D 62, 72–82.
Fisher, S.Z., Aggarwal, M., Kovalevsky, A.Y., Silverman, D.N., McKenna, R., 2012. J. Am. Chem. Soc. 134, 14726–14729.
Fisher, Z., Kovalevsky, A.Y., Mustyakimov, M., Silverman, D.N., McKenna, R., Langan, P., 2011. Biochemistry 50, 9421–9423.
Gelin, M., Delfosse, V., Allemand, F., Hoh, F., Sallaz-Damaz, Y., Pirocchi, M., Bourguet, W., Ferrer, J.L., Labesse, G., Guichou, J.F., 2015. Acta Crystallogr. D 71, 1777–1787.
Genis, C., Sippel, K.H., Case, N., Cao, W., Avvaru, B.S., Tartaglia, L.J., Govindasamy, L., Tu, C., Agbandje-McKenna, M., Silverman, D.N., Rosser, C.J., McKenna, R., 2009. Biochemistry 48, 1322–1331.
Ikeda, T., Saito, K., Hasegawa, R., Ishikita, H., 2017. Angew. Chem. Int. Ed. Engl. 56, 9151–9154.
Kabsch, W., 2010. Acta Crystallogr. D 66, 133–144.
Kovalevsky, A., Aggarwal, M., Velazquez, H., Cuneo, M.J., Blakeley, M.P., Weiss, K.L., Smith, J.C., Fisher, S.Z., McKenna, R., 2018. Structure 26, 383–390.e383.
Kovalevsky, A.Y., Hanson, B.L., Mason, S.A., Yoshida, T., Fisher, S.Z., Mustyakimov, M., Forsyth, V.T., Blakeley, M.P., Keen, D.A., Langan, P., 2011. Angew. Chem. Int. Ed. Engl. 7520–7523.
Köhler, K., Hillebrecht, A., Schulze Wischeler, J., Innocenti, A., Heine, A., Supuran, C.T., Klebe, G., 2007. Angew. Chem. Int. Ed. Engl. 46, 7697–7699.
Lomelino, C.L., Supuran, C.T., McKenna, R., 2016. Int. J. Mol. Sci. 17.
Mahon, B.P., Bhatt, A., Socorro, L., Driscoll, J.M., Okoh, C., Lomelino, C.L., Mboge, M.Y., Kurian, J.J., Tu, C., Agbandje-McKenna, M., Frost, S.C., McKenna, R., 2016. Biochemistry 55, 4642–4653.
Mahon, B.P., Hendon, A.M., Driscoll, J.M., Rankin, G.M., Poulsen, S.-A., Supuran, C.T., McKenna, R., 2015a. Bioorgan. Med. Chem. 23, 849–854.
Mahon, B.P., Lomelino, C.L., Ladwig, J., Rankin, G.M., Driscoll, J.M., Salguero, A.L., Pinard, M.A., Vullo, D., Supuran, C.T., Poulsen, S.-A., McKenna, R., 2015b. J. Med. Chem. 58, 6630–6638.
McKenna, R., Supuran, C.T., 2014. In: Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications. Springer Netherlands, Dordrecht, pp. 291–323.
Michalczyk, R., Unkefer, C.J., Bacik, J.P., Schrader, T.E., Ostermann, A., Kovalevsky, A.Y., McKenna, R., Fisher, S.Z., 2015. Proc. Natl. Acad. Sci. U.S.A. 112, 5673–5678.
Mikulski, R., West, D., Sippel, K.H., Avvaru, B.S., Aggarwal, M., Tu, C., McKenna, R., Silverman, D.N., 2013. Biochemistry 52, 125–131.
Moeker, J., Mahon, B.P., Bornaghi, L.F., Vullo, D., Supuran, C.T., McKenna, R., Poulsen, S.-A., 2014a. J. Med. Chem. 57, 8635–8645.
Moeker, J., Peat, T.S., Bornaghi, L.F., Vullo, D., Supuran, C.T., Poulsen, S.-A., 2014b. J. Med. Chem. 57, 3522–3531.
Nair, S.K., Christianson, D.W., 1991. J. Am. Chem. Soc. 113, 9455–9458.
Otwinowski, Z., Minor, W., 1997. Methods Enzymol. Academic Press, pp. 307–326.
Pastorek, J., Pastorekova, S., 2015. Semin. Cancer. Biol. 31, 52–64.
Pearson, R.L., 2001. In: Alternative Sweeteners, 3rd ed. Marcel Dekker Inc., pp. 147–166.
Pinard, M.A., Boone, C.D., Rife, B.D., Supuran, C.T., McKenna, R., 2013. Bioorg. Med. Chem. 21, 7210–7215.
Schrödinger, L. (2015). The PyMOL Molecular Graphics System, Version 1.8.
Silverman, D.N., McKenna, R., 2007. Acc. Chem. Res. 40, 669–675.
Supuran, C.T., 2008. Nat. Rev. Drug Discov. 7, 168–181.
Supuran, C.T., 2012. J. Enzyme Inhib. Med. Chem. 27, 759–772.
Supuran, C.T., 2016. J. Biochem. 473, 2023–2032.
Supuran, C.T., 2017. Expert Opin. Drug Discov. 12, 61–88.
Tu, C., Silverman, D.N., Forsman, C., Jonsson, B.H., Lindskog, S., 1989. Biochemistry 28, 7913–7918.
van Kuijk, S.J., Yaromina, A., Houben, R., Niemans, R., Lambin, P., Dubois, L.J., 2016. Front. Oncol. 6, 69.
Weber, I.T., Waltman, M.J., Mustyakimov, M., Blakeley, M.P., Keen, D.A., Ghosh, A.K., Langan, P., Kovalevsky, A.Y., 2013. J. Med. Chem. 56, 5631–5635.
Wykoff, C.C., Beasley, N.J.P., Watson, P.H., Turner, K.J., Pastorek, J., Sibtain, A., Wilson, G.D., Turley, H., Talks, K.L., Maxwell, P.H., Pugh, C.W., Ratcliffe, P.J., Harris, A.L., 2000. Cancer Res. 60, 7075–7083.