Promiscuity of Carbonic Anhydrase II. Unexpected Ester Hydrolysis of Carbohydrate-Based Sulfamate Inhibitors

Author

Published
2011

Journal Title
Journal of the American Chemical Society

DOI
https://doi.org/10.1021/ja207855c

Copyright Statement
This document is the Accepted Manuscript version of a Published Work that appeared in final form in Journal of the American Chemical Society, copyright 2011 American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see http://dx.doi.org/10.1021/ja207855c.

Downloaded from
http://hdl.handle.net/10072/42683
Promiscuity of Carbonic Anhydrase II. Unexpected Ester Hydrolysis of Carbohydrate-Based Sulframate Inhibitors

Marie Lopez,† Hoan Vu,† Conan K. Wang,† Maarten Wolf,§ Gerrit Groenhof,§ Alessio Innocenti,§ Claudiu T. Supuran † and Sally-Ann Poulsen†,*

† Eskitis Institute for Cell and Molecular Therapies, Griffith University, Nathan, Queensland 4111, Australia
‡ Biomolecular Chemistry Group, Theoretical and Computational Biophysics Department, Max Planck Institute of Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany
§ Polo Scientifico, Laboratorio di Chimica Bioinorganica, Università degli Studi di Firenze, Via della Lastruccia 3, 50019 Sesto Fiorentino, Florence, Italy.

Abstract: Carbonic anhydrases (CAs) are enzymes whose endogenous reaction is the reversible hydration of CO$_2$ to give HCO$_3^-$ and a proton. CA are also known to exhibit weak and promiscuous esterase activity towards activated esters. Here we report a series of findings obtained with a set of carbonic anhydrase inhibitors that showed quite unexpectedly that the compounds were both inhibitors of CO$_2$ hydration and substrates for the esterase activity of CA. The compounds comprised a monosaccharide core with the C-6 primary hydroxyl group derivatized as a sulfamate (for CA recognition). The remaining four sugar hydroxyl groups were acylated. Using protein X-ray crystallography the crystal structures of human CA II in complex with four of the sulfamate inhibitors were obtained. As expected the four structures displayed the canonical CA protein-sulfamate interactions. Unexpectedly a free hydroxyl group was observed at the anomeric centre (C-1) rather than the parent C-1 acyl group. In addition, this hydroxyl group is observed axial to the carbohydrate ring while in the parent structure it is equatorial. A mechanism is proposed that accounts for this inversion of stereochemistry. For three of the inhibitors the acyl group at C-2 or at C-2 and C-3 were also absent with hydroxyl groups observed in their place and retention of stereochemistry. Using electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICR-MS) we observed directly the sequential loss of all four acyl groups from one of the carbohydrate-based sulfamates. For this compound the inhibitor and substrate binding mode were further analyzed using free energy calculations. These calculations suggested that the parent compound binds almost exclusively as a substrate. To conclude, we have demonstrated that acylated carbohydrate-based sulfamates are simultaneously inhibitor and substrate of human CA II. Our results suggest that initially the substrate binding mode dominates, but following hydrolysis the ligand can also bind as a pure inhibitor thereby competing with the substrate binding mode.

Introduction. Carbonic anhydrases (CAs, EC 4.2.1.1) are zinc metalloenzymes that catalyze the reversible hydration of carbon dioxide (CO$_2$) to generate bicarbonate anion (HCO$_3^-$) and a proton (H$^+$). This equilibrium contributes to a range of physiological functions that involve the production, transport and consumption of CO$_2$, H$^+$ and HCO$_3^-$.

The catalytic domain of CAs comprises a tetrahedral Zn$^{2+}$ cation coordinated to the sidechain imidazole of three histidine residues. The fourth Zn$^{2+}$ ligand is the substrate H$_2$O molecule. Coordination to Zn$^{2+}$ lowers the pK$_a$ of the substrate H$_2$O molecule to $\sim$6.8 (the pK$_a$ of bulk water is $\sim$14) and this facilitates formation of bicarbonate hydroxide at physiological pH.$^{1b}$ Hydroxide is the nucleophile that reacts with CO$_2$, the net effect of CA activity is hydration of CO$_2$, Scheme 1a.$^{1b}$

There are growing numbers of promiscuous activities reported for enzymes, with both substrate promiscuity (catalysis of the same chemical reaction for a range of different substrates) and catalytic promiscuity (catalysis of chemically distinct reactions, involving different transition states) reported.$^2$ hCA II (h = human) exhibits weak and promiscuous esterase activity towards activated esters such as p-nitrophenyl acetate ($k_m/K_M \sim 10^3$ M$^{-1}$s$^{-1}$ versus $\sim 10^5$ M$^{-1}$s$^{-1}$ for CO$_2$ hydration).$^3$ The physiological relevance of CAs catalytic esterase activity is unknown, it is however halted in the presence of CA inhibitors, indicating that the zinc-hydroxide mechanism responsible for the CO$_2$ hydration activity by CAs is also responsible for the esterase activity, Scheme 1b.$^4$ CAs have also been reported to catalyze a number of other hydration reactions, again for which the physiological relevance is unknown.$^5$ The hydration substrates are typically simple in structure and reactions include (i) the hydration of cyanate to carbamic acid (Scheme 2a); (ii) the hydration of cyanamide to urea (Scheme 2b); and (iii) the hydration of aldehydes to gem-diols (Scheme 2c).$^6$

The implied target for CA inhibitors is the active site Zn$^{2+}$ cation and zinc binding groups (ZBGs), especially sulfonamides and sulfamates, feature prominently in small molecule drug design against CAs.$^{1a}$ The anion form of these ZBGs, R-SO$_2$NH$^-$ and R-OSO$_2$NH$^-$ respectively, where R is typically an aromatic moiety, anchors them to the active site Zn$^{2+}$, in place of endogenous water/hydroxide. For R-SO$_2$NH$, the sulfonamide NH forms a hydrogen bond with the sidechain oxygen of Thr$_{198}$, while one oxygen atom accepts a hydrogen bond from the backbone NH of Thr$_{198}$, Figure 1b.$^{1b}$ These canonical interactions are shared by sulfamates, and together these inhibitor-CA II interactions are evidenced in approximately 120 X-ray structures of ligands in complex with hCA II in the Protein Data Bank. Binding of R-SO$_2$NH$^-$ and R-OSO$_2$NH$^-$ mimics the transition state structure of CA bound to bicarbonate, where the bicarbonate OH oxygen atom coordinates to the active site Zn$^{2+}$, the bicarbonate OH group also donates a hydrogen bond to the sidechain oxygen of Thr$_{198}$, while the carbonyl oxygen of bicarbonate accepts a hydrogen bond from the backbone NH of Thr$_{198}$, Figure 1a.$^{1b}$ Small molecule CA inhibitors that do not interact with the zinc ion have

Keywords: carbonic anhydrase, esterase, catalytic promiscuity, sulfamate, carbohydrate, bioaffinity mass spectrometry.
Scheme 1. Representation of the catalytic cycle for human CA II catalyzed a) hydration of CO$_2$ to HCO$_3^-$ and H$^+$; and b) ester hydrolysis to carboxylic acid and alkoxide. The zinc bound hydroxide is the active form of the enzyme.

Scheme 2. Other CA catalyzed hydration reactions.

(a) O=C=NH + H$_2$O $\rightarrow$ H$_2$NCOOH

(b) NH=C=NH + H$_2$O $\rightarrow$ H$_2$NCONH$_2$

(c) RCHO + H$_2$O $\rightarrow$ RCH(OH)$_2$

been identified only recently, for example compounds comprising a coumarin scaffold. 

In the current study we present a series of structural observations that were obtained following the analysis of several CA inhibitors in the presence of hCA II using protein X-ray crystallography and bioaffinity mass spectrometry. These CA inhibitors belonged to a novel compound class of carbohydrate-based sulfamates, the carbohydrate hydroxyl groups were either free (-OH) or acylated (-OAcyl). Both biophysical methods provided data that indicated that the sulfamate compounds, in addition to being CA inhibitors were also substrates for the esterase activity of CA, with the ester groups of the sugar hydrolyzed by CA II. For one of these compounds, both the inhibitor and substrate binding mode were further analyzed using free energy calculations. These calculations suggest that the parent compound binds almost exclusively as a substrate. The various hydrolysis products can also bind as a pure inhibitor, competing with the substrate binding mode.

Results and Discussion. The esterase activity reported for hCA II with phenyl acetates shows that activity is sensitive to the chemical nature of both the acyl and the phenolic fragments that comprise the ester substrate. 

Tawfik demonstrated a linear relationship between the pK$_a$ of the phenolic leaving group and hCA II esterase substrate reactivity with a series of substituted phenyl acetates, Table 1. Their study showed the ester substrate reactivity falls as the pK$_a$ of the alcohol leaving group increases, thus activated esters wherein the pK$_a$ of the product alcohol is low, are the best substrates for the esterase activity of hCA II. 4-Nitrophenyl acetate (pK$_a$ of 4-nitrophenol = 7.14) is the most active known hCA II substrate for ester hydrolysis, while 4-methoxyphenyl acetate (pK$_a$ of 4-methoxyphenol = 10.29) is ~ 400-fold less active, Table 1.

The compounds in this study are sulfamates derived from the monosaccharides D-glucose (1a-d), D-galactose (2a-b) and D-mannose (3a-b), Chart 1. The sulfamate moiety is on the C-6 primary hydroxyl group of the monosaccharide, with the remaining four hydroxyl groups either unmodified (1a-3a) or acylated as acetyl (1b-3b), propionyl (1c) or butyryl (1d) ester protecting groups. In a recent contribution we presented the design, synthesis and CA inhibition profile of these compounds. The hCA II inhibition constants (K$_i$) ranged from 11-307 nM as determined using a stopped flow assay that monitors the production of H$^+$ from the CA catalyzed hydration of CO$_2$, Table 2.

Our primary motivation in studying CA inhibitors is therapeutic drug discovery. Given the variable hCA II inhibition profile across this small group of monosaccharide derivatives, differing only by the hydroxyl group acylation profile and the axial/equatorial stereochemistry, we hoped to discern the interactions of these compounds with active site residues of hCA II so as to identify the structural features of these ligands that may be important to directing future medicinal chemistry campaigns against this enzyme class. Protein X-ray structures of native CA with inhibitors can
Table 1. Relationship of CA II catalyzed ester hydrolysis of phenyl acetates with leaving groups of varying pKₐ values.8

<table>
<thead>
<tr>
<th>Substrate</th>
<th>pKₐ</th>
<th>k_cat/K_M × 10^3 M⁻¹ s⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-nitrophenyl acetate</td>
<td>7.14</td>
<td>2050 ± 162</td>
</tr>
<tr>
<td>2,3-difluorophenyl acetate</td>
<td>7.81</td>
<td>1140 ± 62</td>
</tr>
<tr>
<td>2,4-difluorophenyl acetate</td>
<td>8.43</td>
<td>133 ± 8</td>
</tr>
<tr>
<td>4-chlorophenyl acetate</td>
<td>9.38</td>
<td>53 ± 3</td>
</tr>
<tr>
<td>phenyl acetate</td>
<td>10</td>
<td>40 ± 4</td>
</tr>
<tr>
<td>4-methoxyphenyl acetate</td>
<td>10.29</td>
<td>5.3 ± 1</td>
</tr>
</tbody>
</table>

Table 2. Enzyme inhibition data of hCA II with carbohydrate-based sulfamates 1a-d, 2a-b, 3a-b. (h = human).

<table>
<thead>
<tr>
<th>Compound</th>
<th>Kᵢ (nM)⁴</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>82</td>
</tr>
<tr>
<td>2a</td>
<td>93</td>
</tr>
<tr>
<td>3a</td>
<td>104</td>
</tr>
<tr>
<td>1b</td>
<td>307</td>
</tr>
<tr>
<td>2b</td>
<td>106</td>
</tr>
<tr>
<td>3b</td>
<td>11.3</td>
</tr>
<tr>
<td>1c</td>
<td>105</td>
</tr>
<tr>
<td>1d</td>
<td>114</td>
</tr>
</tbody>
</table>

⁴Errors in the range of ± 5% of the reported value, from three determinations.
Figure 2. Schematic view of observed ligand interactions from inhibitor-bound hCA II: sulfamate 1b (compound 4), 1c (compound 5), 2b (compound 6) and 3b (compound 7) protein crystal structures. The canonical sulfamate interactions (Figure 1b) are not shown for the purpose of clarity. Hydrogen bond interactions are shown as dotted lines, van der Waals interactions as dashed lines. For compound 1c, the C-1 hydroxyl is observed 95% axial (shown) and 5% equatorial, the water H-bond is conserved in both anomers.

Table 3. Summary of ligand interactions with protein residues in hCA II:ligand complex crystal structures

<table>
<thead>
<tr>
<th>Ligand</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct hydrogen bonds</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>H$_2$O hydrogen bonds</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>vdW interactions</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

*The canonical sulfonamide and sulfamate interactions are not included. *Only bond distances below 3.5 Å were considered. *Only bond distances below 4.5 Å were considered.

Figure 3. Crystal structures of human CA II in complex with sulfamates 1b (compound 4), 1c (compound 5), 2b (compound 6) and 3b (compound 7). Shown are the ligand conformations in the active site. The protein is not shown for clarity. The catalytic zinc ion is colored magenta. The final 2F$_{o}$ – F$_{c}$ density is contoured in blue at 1σ. The figure was prepared with PyMol.°
In addition to the canonical sulfamate interactions, the crystal structures of the four sulfamate ligands reveal that there are several additional interactions between the bound ligands and amino acid residues of hCA II, some of which are highly conserved and involve residues that are involved in zinc binding (His94) or proton transfer (His64). The observed hydrolyzed ligands of 1b, 1c, 2b and 3b are denoted as compounds 4, 5, 6 and 7, respectively. For each ligand, the anomeric hydroxyl group interacts indirectly with the Asn67 and Asn62, while the sidechain of Leu60 provides hydrophobic interactions with Phe130, Leu197, Val121 and/or hydrogen bond with Asn62 and Gln92, while the C-3 alkyl group oxygen of 4 and 5 forms a H-bond with either Glu62 (compounds 4 and 5) or Asp62 (compound 6), and the acyl oxygen of 5 and 7 form a hydrogen bond with Asn62 and Glu62, respectively. The C-4 acyl groups form a network of hydrophobic interactions with Phe236, Leu197, Val121 and/or Pro201 and a hydrogen bond with Glu62 (except compound 7). For compound 7, the distance from the sidechain of Phe236 to the C-4 acyl group CH$_3$ is 2.89 Å, which is considerably smaller in value relative to the other compounds where the respective atoms are separated by at least 3.79 Å. The initial difference density clearly showed the position of the Zn$^{2+}$ ion, surrounding water molecules, and each atom of the bound ligand. Based on the difference density, we were able to confidently identify the composition of the bound ligand and model its conformation in the active site of hCA II. Figure 3 shows the final 2F$_{o}$ – F$_{c}$ density for the bound ligand and Zn$^{2+}$ ion. A summary of ligand interactions with hCA II is presented in Table 3. For the data sets reported here, there was no ambiguity as to the presence of hydrolyzed ligands in the active site cleft. The data collection and refinement statistics of inhibitor-bound hCA II crystal structures are provided in Supporting Information. The majority of publications that describe structures of hCA II annotate residue numbers that are offset by one relative to the amino acid sequence of hCA II. This may be a consequence of the numbering scheme used in the early hCA II crystal structures (PDB accession codes 1CA2, 2CBB) that skip residue number 126. In this study, we have applied a residue numbering scheme that is in accordance with the amino acid sequence of hCA II.

hCA II is a well studied model enzyme and many assays for measuring enzymatic activity, thermodynamic parameters or kinetic parameters for binding of inhibitors or substrates to CA are reported. Assays are typically based on indirect reporters, for example the emission of light (e.g. fluorescence) or heat change (e.g. ITC). There are no assays designed to discern if a compound is both an inhibitor and a substrate of CA. An assay that allows us to directly observe structural features of CA-inhibitor binding interactions is needed to further assess our crystallography-based findings. Electrospray ionization mass spectrometry (ESI-MS) has been used to study proteins and complexes of proteins with naturally occurring substrates, inhibitors, and drugs. It is generally straightforward to adjust the parameters of the ESI-MS measurement so that the signals observed in the mass spectrometer for a protein and the noncovalent complexes thereof reflect these species in solution. Thus if an inhibitor is combined with its target enzyme under native state conditions then the noncovalent complex of [enzyme + inhibitor] may be observed in the ESI mass spectrum. The mass difference ($\Delta$m/z) between the peaks for the unbound [enzyme] and the noncovalent complex [enzyme + inhibitor] can be multiplied by the charge state to give directly the molecular weight of the binding ligand i.e. $\text{MW}_{\text{ligand}} = \Delta m/z \times z$ to provide confirmation of the identity of the bound inhibitor.

We and others have previously demonstrated that electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICR-MS) allows the direct observation of CA in complex with CA inhibitors, and have developed this bioaffinity mass spectrometry methodology (BAMS) to screen for novel CA inhibitors. The FTICR technique allows highly accurate mass measurements. This feature is of great benefit in determining the identity of the unknown binding species in noncovalent protein-ligand complexes. We recently identified a novel class of CA inhibitors belonging to the coumarin chemotype using BAMS. In that study a natural product coumarin, 6-((5-hydroxy-3-methylbutyl)-7-methoxy-2H-chromen-2-one (8), displayed significant CA inhibition. Using protein X-ray crystallography the lactone of 8 was not observed and instead its hydrolysis product, the cis-2-hydroxy-4-(1S,3-methylbutyl)-3-methoxy-cinnamic acid 9 fitted the electron density, Scheme 3. Using BAMS it was confirmed that the CA bound ligand had an increased molecular weight of 18 Da compared to 8, consistent with the mass of the hydrolyzed coumarin 9. A number of follow up studies confirmed that the zinc bound hydroxide anion of the CA enzyme was responsible for the hydrolysis of the lactone ring of 8.

**Scheme 3.** The lactone of the coumarin natural product CA inhibitor is hydrolyzed by hCA II.

![Scheme 3](image)

The BAMS method has proven a valuable biophysical screening method as it allows the direct observation of the inhibitor with the protein and appears ideal to further assess the unexpected crystallographic observations of the sulfamates. We acquired the positive ion mass spectrum of a solution of hCA II (3.4 μM) in 10 mM NH$_4$OAc, pH 7.2 using ESI-FTICR-MS, Figure 4. Peaks corresponding to the 12$^+$ and 10$^+$ charge states of hCA II were observed, this narrow mass distribution is indicative of fully folded (native state) hCA II protein.
**Figure 4.** ESI mass spectra of hCA II obtained under native-state conditions. Peaks corresponding to the $12^+\text{ to }10^+$ charge states of hCA II are observed.

Next we studied a solution of hCA II in the presence of sulfamate inhibitor $1b$. Mass spectra were acquired with this sample following incubation at $16^\circ C$ (the same temperature as used for protein co-crystallization) for 15 min, 1 day, 3 days, 18 days and 21 days. This analysis yielded the series of mass spectra shown in Figure 5b. These spectra show that with increasing incubation time the sequential loss of 42 Da from the hCA II:ligand noncovalent complex is observed. This mass loss is consistent with the sequential loss of acetyl groups from $1b$ to give hydrolyzed ligands $10, 4, 12$ and $1a$, Figure 5a. The 15 min incubation yielded the ESI positive ion mass spectrum of Figure 5b, purple trace (only the $11^+$ charge state is shown for clarity). In this spectrum peaks corresponding to unbound hCA II and hCA II-$1b$ complex are observed. The ESI positive ion mass spectrum of Figure 5b, red trace was obtained following a one day incubation time. In this spectrum peaks corresponding to unbound hCA II and three different noncovalent complexes of hCA II are observed: the latter correspond to [hCA II + $1b$] as well as the loss of one acetate group [hCA II + $10$] and loss of two acetate groups [hCA II + $4$]. Similarly the ESI mass spectra were acquired after three days (Figure 5b, blue trace), 18 days (Figure 5b, green trace) and 21 days (Figure 5b, orange trace). After three days the [hCA II + $4$] complex predominated, with much less of [hCA II + $1b$] and [hCA II + $10$] compared to day 1. The [hCA II + $4$] corresponds to the complex observed by protein X-ray crystallography (above; dataset hCA II:1b_4). After 18 days the [hCA II + $12$] complex predominated, this corresponds to hydrolysis of three acetate groups from the parent sulfamate $1b$. After 21 days the [hCA II + $12$] complex continued to predominate – but also evident was the presence of a complex that corresponds to the hydrolysis product of removal of all four acetates of the ligand $1b$, (i.e. compound $1a$) to give the complex [hCA II + $1a$]. The mass values of the peaks observed and calculation of the mass of the hCA II bound ligands are provided in the Supplementary Data.

We then synthesized compound $10$, the derivative of $1b$ where the C-1 anomeric hydroxyl group is deacetylated to give an alcohol as an anomeric $\alpha/\beta$ mixture in a ratio $\alpha:\beta$ of 6:4. The BAMs studies demonstrate that compound $1b$ is not hydrolyzed following 15 mins of incubation, however after one day both $1b$ and hydrolysis products $10$ and $4$ are bound to hCA II, while the fully deprotected $\alpha$ anomer of $1a$ appears after 21 days. Our CA enzyme inhibition assay includes a pre-incubation period of 15 mins of CA with test inhibitor as this is sufficient time to allow formation of a standard CA-inhibitor complex to allow the inhibition constant ($K_{i}$) to be determined. The BAMs studies demonstrate that the hCA II assay result for $1b$ ($K_{i} = 307$ nM) determined after 15 min pre-incubation is indeed due to inhibition by the $1b$ parent structure. To afford further insight into the crystallography and BAMs observations we have elaborated on the enzyme assays of $1b$ as well as the deacetylated derivative of $1b$, newly synthesized compound $10$. Assays were performed in which hCA II was pre-incubated with $1b$ or $10$ for 15 min, 24 h or 72 h at either 4 °C or 25 °C prior to performing the enzyme inhibition assay, results are given in Table 4. Comparing Table 5, entry 1 with entries 2-4, shows that the inhibition constant for $1b$ decreases with time from $K_{i} = 307$ nM (15 min pre-incubation) to $K_{i} = 100$ nM (24 h pre-incubation at either 4 °C or 25 °C). These results suggest that the hydrolysis products of $1b$ are better inhibitors of hCA II than the parent $1b$ structure. Compound $10$ ($K_{i} = 50$ nM, Table 4, entry 5) and compound $1a$ ($K_{i} = 82$ nM, Table 4, entry 6) are 6-fold and is 3.7-fold more potent than $1b$, respectively when a 15 min pre-incubation period is applied. These measured $K_{i}$ values are consistent with the CA-mediated hydrolysis of $1b$ leading to higher potency hCA II inhibitors.

**Table 4.** Time dependent inhibition $K_{i}$ data of hCA II with carbohydrate-based sulfamate $1b$, and its deacetylated analogues $10$ and $1a$.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Compound</th>
<th>Pre-incubation time &amp; temperature</th>
<th>$K_{i}$ (nM)$^{a,b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1b$</td>
<td>15 min, 25 °C</td>
<td>307</td>
</tr>
<tr>
<td>2</td>
<td>$1b$</td>
<td>24 h, 4 °C</td>
<td>113</td>
</tr>
<tr>
<td>3</td>
<td>$1b$</td>
<td>72 h, 4 °C</td>
<td>110</td>
</tr>
<tr>
<td>4</td>
<td>$1b$</td>
<td>72 h, 25 °C</td>
<td>101</td>
</tr>
<tr>
<td>5</td>
<td>$10$</td>
<td>15 min, 25 °C</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>$1a$</td>
<td>15 min, 25 °C</td>
<td>82</td>
</tr>
</tbody>
</table>

$^{a}$Errors in the range of $\pm$ 5% of the reported value, from three determinations. $^{b}$Human (cloned) isozymes.

The stopped-flow together with the BAMs measurements demonstrated that the carbohydrate-based sulfamates are both substrate and inhibitor for hCA II. The crystallography data reveal the binding mode for inhibition of several hydrolyzed compounds, but do not show in what conformation the substrate needs to bind to undergo the hydrolysis reaction. To investigate substrate binding in atomic detail, we performed molecular dynamics (MD) simulations of structural models of the substrate $1b$ in complex with hCA II, generated with docking tools. Of all docked conformations, only those in which the hydroxide oxygen and a carbonyl carbon of $1b$ were in close contact were selected, Figure 6. During the MD simulations, the complexes were stable, with the zinc-bound hydroxide in close proximity to one or more carbonyl carbons of the ester groups.
We further assessed the stability of the modeled substrate-enzyme complex by calculating the non-covalent binding affinity, Table 5. Under the assumption that the docked complexes are representative of reactive conformations, there is a weak tendency for the parent compound 1b to bind as a substrate for hydrolysis, with a binding constant of 1.34 mM. Note, that for a full analysis of the binding affinity the covalent interaction, the second step in the reaction mechanism, should also be accounted for. This step is characterized by the nucleophilic attack of the hydroxide on the carbonyl carbon, but is currently computationally inaccessible. To put the calculated dissociation constant into perspective, we repeated the docking and free energy calculation with an activated and a non-activated ester, p-nitrophenyl acetate and phenyl acetate respectively, for which the esterase activity of hCA II has been measured.\textsuperscript{3b} We find that the noncovalent dissociation constant for the hCA II-1b complex is orders of magnitude lower than for hCA II in complex with the known substrates. This strongly suggests that also 1b can be hydrolyzed by hCA II, even when the catalytic rate for 1b turns out to be lower than in the case of the known substrates.

To estimate the effect of the hydrolysis on the binding as the inhibitor, we computed the difference in binding free energy ($\Delta\Delta G$) of the inhibitor before (compound 1b) and after two hydrolysis steps (compound 4). In these simulations, 1b was assumed to bind in the same conformation as observed in the X-ray structure of the hCA II-compound 4 complex described above (PDB accession code 3T82). Note, that due to the covalent bond between the zinc and the sulfamate we cannot calculate the absolute binding free energy for the inhibitor binding mode and therefore focussed on the $\Delta\Delta G$. The results of these calculations show that 4 binds much stronger than 1b ($\Delta\Delta G = 39$ kJ mol$^{-1}$). This large difference in binding affinity between 1b and 4 does not account for the small differences in the measured inhibition constants in Table 5. Therefore, we can speculate that species 1b does not bind as an inhibitor, but rather as a substrate. This is also supported by the favourable calculated substrate binding affinity and may explain why no X-ray structures were found for the CAII-1b complex.
Table 5. Binding free energies and dissociation constants for ligands in a pure inhibitor binding mode and substrate binding mode.

<table>
<thead>
<tr>
<th>Substrate binding mode</th>
<th>$\Delta G_{\text{bind}}^0$ (kJ mol$^{-1}$)</th>
<th>$k_d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>neutral 1b</td>
<td>-16.5 ± 0.62</td>
<td>1.34×10$^3$ ± 0.13×10$^3$</td>
</tr>
<tr>
<td>phenyl acetate</td>
<td>10.29 ± 0.23</td>
<td>62.9 ± 3.5</td>
</tr>
<tr>
<td>p-nitrophenyl acetate</td>
<td>10.86 ± 0.86</td>
<td>78 ± 15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inhibitor binding mode</th>
<th>$\Delta G_{\text{bind}}$ (kJ mol$^{-1}$)</th>
<th>$k_d/k_d^{1b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1b → 4</td>
<td>-39.25 ± 3.08</td>
<td>6.58×10$^6$</td>
</tr>
</tbody>
</table>

Whitesides recently reviewed the kinetics of association and dissociation of CA inhibitors with a sulfonamide ZBG to hCA II. This analysis showed that ArSO$_2$NH$_2$ compounds fit both a two-state or a three-state model (Scheme 4). In the two-state model, the ArSO$_2$NH$_2$ associates with hCA II and coordinates to the active site Zn$^{2+}$ in one step, while in the three-state model, the ArSO$_2$NH$_2$ first associates with the hCA II to form a non-coordinated complex; that in a second step coordinates to the Zn$^{2+}$. A number of experimental results were described that supported the three-state model. Our findings with the carbohydrate-based sulfamates are also consistent with the three-state model as these compounds necessarily bind to the CA active site as either a substrate (sulfamate not coordinated to Zn$^{2+}$) or as an inhibitor (sulfamate is coordinated to Zn$^{2+}$).

Scheme 4. Two-state (A) and three-state (B) models for the association of arylsulfonamides with CA.

A: Two-state model

B: Three-state model


The C-1 acyl group of the four parent sugar sulfamates is equatorial, while the C-1 hydroxyl group was observed axial to the carbohydrate ring for 1b, 2b and 3b and was observed 95% axial and 5% equatorial for 1c. Hydrolysis of the sugar acyl groups can occur via acyl-oxygen cleavage as in Scheme 1, however this does not account for the inversion of configuration observed at C-1 in the X-ray crystal structures. A mechanism that is consistent with the almost exclusive formation of the axial hydroxyl from the anomeric acyl group is suggested in Scheme 5 (showing inhibitor/substrate 1b as an example). Here hydrolysis occurs via alkyl-oxygen cleavage with formation of the oxocarbenium resonance stabilised cation and departure of the zinc coordinated acyl group. The zinc bound hydroxyl then reacts with the oxocarbenium cation leading to the axial anomer and zinc bound acetate. The acetate is displaced by water and the active CA with zinc bound hydroxyl is regenerated. The formation of a small amount of the equatorial anomer in the case of 1c could be due to competitive hydrolysis via the two mechanisms (Schemes 1 and 5) or anomerization following alkyl-oxygen cleavage (Scheme 5). The sequential hydrolysis of acyl groups around the ring may be due to a proximity effect – after hydrolysis of the most labile acyl group at C-1, the zinc centre is “ready” to bind to the next nearest carbonyl group.

In conclusion, we have demonstrated that per-O-acylated sugar-based sulfamates are both substrate and inhibitor for hCA II. This research grew from the initial observations with protein:ligand structures obtained by X-ray crystallography that unexpectedly revealed the hydrolyzed ligands (compounds 4, 5, 6 and 7) in place of the expected parent ligands 1b, 1c, 2b and 3b, respectively. The stopped-flow together with the BAMS measurements then demonstrated that the carbohydrate-based sulfamates are both substrate and inhibitor for hCA II. To investigate what conformation the substrate needs to bind hCA II to undergo the hydrolysis reaction, we next performed MD simulations of structural models of the substrate 1b in complex with hCA II. Mechanistically the ester hydrolysis is mediated by the zinc-bound hydroxyl and we can speculate that the fully acylated sugar-based sulfamates do not bind as an inhibitor, but rather as a substrate.
Figure 6. Active site structure of hCA II with the ligand inside the pocket. a) Pure inhibitor conformation of compound 4 (X-ray structure CA II:1b,4) and b) Substrate conformation of the parent sulfamate 1b (docking structure). Highlighted are the zinc (grey sphere), hydroxide (red-white spheres), ligand (licorice) and histidine residues (licorice in the background).

Scheme 5. Proposed mechanism for the anomeric acyl group hydrolysis of 1b, 2b, 3b and 1c by alkyl oxygen cleavage (showing inhibitor/substrate 1b as an example).

---

Experimental Section

Protein crystallography. Human recombinant carbonic anhydrase II (CA II) was expressed in bacteria and crystallized as described previously. Ligand-bound crystals were obtained by co-crystallization with 10 mM ligand (added from a 60 mM stock solution in MeOH). Crystals were immersed in cryoprotectant (25% glycerol in mother liquor with 10 mM ligand). X-ray diffraction results of ligand-bound CA II were obtained at the Australian Synchrotron beamline MX1 (equipped with a Quantum ADSC CCD detector). Datasets were indexed with XDS or Mosflm, and scaling, truncation and analysis was performed with programs from the CCP4 suite. The structures were determined using difference Fourier techniques and refined with Phenix. Ligand topology was generated with PRODRG, and manual model building and visual inspection was performed with O and Coot. Coordinates and structure fac-
Mass Spectrometry. Mass spectrometry was performed on an APEX® III 4.7 Tesla FTICR mass spectrometer (Bruker Daltonics, Billerica, MA, USA) fitted with an Apollo™ electrospray ionisation (ESI) source operated in positive ion mode. XMASS NT V6.1.2 mass spectrometry software on a PC platform was used for data acquisition. Broadband excitation was used to analyze a mass range from m/z 50 – 6000 and each spectrum was an average of 64 transients (scans) with 512 K data points acquired in low resolution mode, with an acquisition time of approximately 4 min/sample. Samples were infused into the ESI source at 2 μL/min. Nitrogen was used as both the drying gas (125 °C) and nebulizing gas. Relevant parameters include a 10⁴-fold pressure reduction between source and analyzer regions with an ESI source pressure (10⁻⁶ mbar) and high vacuum analyzer region pressure (6x10⁻¹⁰ mbar). Agilent ESI tuning mix (Santa Clara, California, USA) was used for an external three-point calibration. The hexapole ion accumulation time was 3 s. Samples for MS analysis were prepared as follows: 3.23 μM hCA II and 11.69 μM (3.6 equiv.) ligand 1b in a total volume of 500 μL. Samples were incubated at 16 °C for 15 min, 1 day, 3 days, 18 days and 21 days prior to MS analysis.

Chemical Synthesis, general. All starting reagents were purchased from commercial suppliers. Reactions were monitored by TLC. TLC plates were visualized with UV light and/or orcinol stain (1 g of orcinol monohydrate in a mixture of EtOH:H₂O:H₂SO₄, 72.5:22.5:5 mL). Silicone gel flash chromatography was performed using silicone gel 60 Å (230-400 mesh) from DaviSil. ¹H NMR were recorded on a Varian Unity 500 MHz spectrometer at 30 °C. Chemical shifts for ¹H and ¹³C NMR run in DMSO-d₆ are reported in ppm relative to residual solvent proton (δ = 2.50 ppm) and carbon (δ = 39.5 ppm) signals, respectively. Multiplicity is indicated as follows: s (singlet); d (doublet); t (triplet); m (multiplet); dd (doublet of doublet); ddd (doublet of doublet of doublet); br (broad). Coupling constants are reported in Hertz (Hz). Melting points measured on a Cole Parmer instrument are uncorrected. High-resolution electrospray ionization mass spectra were performed on an Apex III Bruker Daltonics 4.7T Fourier transform mass spectrometer (FTMS) fitted with an Apollo ESI source. Low resolution mass spectra were acquired on an Applied Biosystems Pty Ltd Mariner ESI-TOF mass spectrometer using electrospray as the ionization technique in positive ion and/or negative ion modes as stated. All MS analysis samples were prepared as solutions in methanol. Compounds 1a-3a, 1b-3b, 1c and 1d were synthesized as reported previously by our group.⁹

Synthesis of 2,3,4-tri-O-acetyl-6-sulfamoyl-a,b-d-glucopyranose (10)²⁵ Hydrazine acetate (24 mg, 0.27 mmol, 1.3 equiv.) was added to a solution of compound 1b (91 mg, 0.21 mmol, 1.0 equiv.) in DMF (1 mL) at 75 °C. The reaction stirred for 20 min at 75 °C the diluted in EtOAc (20 mL), washed with brine (×2) and the aqueous fractions back extracted with EtOAc (×2). The organic fractions were then combined, dried over MgSO₄, filtrated and the solvent evaporated leaving a crude residue that was purified by flash chromatography (1:1 EtOAC/hexane) to afford the title compound 10 (66 mg, 0.17 mmol, 81% yield, a/b 6/4) as a white solid. m.p. = 147 °C. Rₓ = 0.28 (3:2 EtOAC/hexane). ¹H NMR (500 MHz, DMSO-d₆): δ = 7.53 (br s, 2H, NH₂); 7.36 (d, J = 6.0 Hz, 0.4H, OH-1); 7.33 (d, J = 4.5 Hz, 0.6H, OH-1); 5.36 (t, J = 10.0 Hz, 0.6H, H-3); 5.25 (t, J = 10.0 Hz, 0.4H, H-3); 5.25 (t, J = 4.0 Hz, 0.6H, H-1); 4.90 (t, J = 10.0 Hz, 0.6H, H-4); 4.89 (t, J = 7.0 Hz, 0.4H, H-2); 4.83 (t, J = 9.5 Hz, 0.4H, H-4); 4.73 (dd, J = 10.0, 3.5 Hz, 0.6H, H-2); 4.71 (dd, J = 10.0, 8.0 Hz, 0.4H, H-2); 4.19 (m, 0.6H, H-5); 4.05-4.01 (m, 2.4H, H-5); 6.6a-6.6a; 2.02, 2.01, 1.99, 1.97, 1.94 (5 × s, 9H, OCOC₃H₇), assignments were confirmed by ¹H-¹H gCOSY. ¹³C NMR (152 MHz, DMSO-d₆): δ = 170.2, 170.1, 169.7 (OCOC₃H₇); 94.1 (C-1); 89.3 (C-1); 72.7 (C-2); 72.6 (C-3); 71.0 (C-2); 70.6 (C-5); 69.7, 69.7 (C-3); 68.8 (C-4); 68.7 (C-5); 67.5 (C-6); 66.6 (C-6); 20.7, 20.6, 20.5 (OCOC₃H₇), assignments were confirmed by ¹H-¹³C HSQC. LRMS (ESI): m/z = 408 [M + Na]+. HRMS: Calculated for C₁₂H₁₈N₄O₁₃SNa 408.0571, Found 408.0577.

Carbonic anhydrase catalytic inhibition assays. An SX.18MV-R Applied Photophysics stopped-flow instrument has been used for assays the CA II CO₂ hydration activity. Phenol red (at a concentration of 0.2 mM) was used as indicator, working at the absorbance maximum of 557 nm, with 10 mM Hepes (pH 7.5) as buffer, 0.1 M NaClO₄ (for maintaining constant the ionic strength, this anion is not inhibitory), following the CA-catalyzed CO₂ hydration reaction for a period of 10 – 100 s. Saturated CO₂ solutions in water at 20 °C were used as substrate. Stock solutions of inhibitors were prepared at a concentration of 10 – 50 mM (in the assay buffer) and dilutions up to 1 nM done with the assay buffer mentioned above. Inhibitor and enzyme solutions were preincubated together for the required time at room temperature prior to assay, in order to allow for the formation of the E-I complex and time for the acyl group hydration. The inhibition constants were obtained by non-linear least-squares methods using PRISM 3. The curve-fitting algorithm allowed us to obtain the IC₅₀ values, working at the lowest concentration of substrate of 1.7 mM), from which Kᵢ values were calculated by using the Cheng-Prusoff equation. The catalytic activity (in the absence of inhibitors) of these enzymes was calculated from Lineweaver-Burk plots and represent the mean from at least three different determinations. The hCA II enzyme concentration was 10.3 nM.

Free-energy calculations
Simulation setup. All simulations were performed using the Gromacs software package. The md integrator was used. All bonds were constrained using the LINCS algorithm, allowing a time step of 2 fs. The temperature was connected to a heat bath of 300 K via the v-rescale thermostat. The pressure was maintained constant at 1 atm using the Parrinello-Rahman barostat with τp set to 1.0 and the compressibility to 4.5×10⁻⁵. Van der Waals interactions were switched off between 0.8 and 0.9 and a dispersion correction term was included. The electrostatic interactions were treated using PME²⁸ with a real space cut-off of 1.0 nm. The simulation box was generated by creating a dodecahedron box around the protein and/or ligand, with a minimum distance of 1.0 nm between the box and the solute. The rest of the box was filled with TIP3P²⁹ water molecules and Na⁺ and Cl⁻ atoms were added to obtain a neutral box of approximately 0.1 M. The resulting system was energy minimized.
by steepest descent and simulated for 1 ns with position restraints on the protein and/or ligand.

We have used the Amber force fields for our simulations, with the Amber99sb parameter set for the protein. The parameters for the active site were taken from Suárez et al. To describe the carbohydrate ligands we used the carbohydrate optimized Glycam06 parameter set. The missing parameters for the sulfamate were taken from the Gaff parameter set. To describe the phenyl acetate, p-nitrophenyl acetate we used the Gaff parameter set. For all ligands, the derivation of the charge set was done following the Glycam06 procedure. Glycam06 was preferred over Gaff, because all atom charge fitting leads to overfitting of the electrostatic potential. Removing the aliphatic hydrogens from the fit will reduce the chance of overfitting without large effects for the charge distribution of the molecule. The charges we used for the ligands and the active site are provided as supporting information.

Inhibitor binding free energy difference. The binding free energy of the ligand in the inhibitor binding mode cannot be calculated directly, because the sulfamate nitrogen is chemically bonded to the zinc of the hCA II active site. The energy of a chemical bond cannot be assessed in a classical MD simulation and consequently the associated binding free energy cannot be calculated. Therefore we calculated the free energy difference between ligand 1b and compound 4 in water and in the protein. These are related to the binding free energies through the thermodynamic cycle displayed in Figure 7.

$$\Delta G_{\text{hyd,prot}} = \Delta G_{\text{bind,prot}} - \Delta G_{\text{hyd,water}}$$

Figure 7. Free energy cycle used to calculate the difference in binding free energy between compound 4 and 1b (\(\Delta G_{\text{bind}} = \Delta G_{\text{4,bind}} - \Delta G_{\text{1b,bind}}\)).

Then \(\Delta G_{\text{bind}} = \Delta G_{\text{4,bind}} - \Delta G_{\text{1b,bind}} = \Delta G_{\text{hyd,prot}} - \Delta G_{\text{hyd,water}}\). The difference in binding free energy \(\Delta G_{\text{bind}}\) is then related to the ratio of the dissociation constants through

$$\frac{k'_{4}}{k'_{1b}} = e^{-\Delta G_{\text{bind}}/RT}.$$ The free-energy differences \(\Delta G_{\text{hyd,prot}}\) and \(\Delta G_{\text{hyd,water}}\) were computed using the Crooks gaussian intersection method (CGI) and the error is estimated as described.

For the short simulations to calculate the work to change 1b into 4 and vice versa the van der Waals switching parameters were changed to 1.0 and 1.1 nm and the PME real space cut-off to 1.2 nm. For \(\Delta G_{\text{hyd,water}}\) morphing one state into the other was done within 50 ps. This was repeated 100 times with different starting conformations obtained from at least 10 ns equilibrium trajectory. For \(\Delta G_{\text{hyd,prot}}\) the same procedure was followed with morphing times increased to 1 ns to obtain reasonable accuracy.

Ester hydrolysis by hCA II. We used Flexx to dock neutral 1b, phenyl acetate and p-nitrophenyl acetate into the active site of hCA II. The initial protein structure for the docking was the X-ray structure PDB ID 3T82 (CA II:1b_4). After adding hydrogens to this structure we substituted the ligand N-H (ligated to Zn\(^{2+}\)) with O-H and removed all other atoms of the ligand. We defined the binding site as a sphere with a radius of 8 Å around the oxygen of the hydroxide ion. For every ligand we created the maximum number of docking poses (2000). From the created database of docking poses we visually inspected the structures. The four highest ranked structures with a unique close contact interaction between the hydroxide ion and the carboxyl carbon of the ester were selected for further investigation. Then, we equilibrated these complexes for 2 ns including position restraints on the ligand Zn\(^{2+}\), OH and the three histidine residues coordinated to the Zn\(^{2+}\) to allow for side chain equilibration. Finally, a short production run was done without position restraints for 10 ns.

When the ligand is considered as a substrate for ester hydrolysis there are no bonded interactions between the ligand and the active site. Consequently, we can calculate the binding free energy directly. Therefore we use discrete thermo-dynamic integration (DTI) and Bennet Acceptance Ration (BAR) to extract the ligand binding free energy. First, the ligands coulomb interactions are removed in five steps \(\lambda = \{0,0.25,0.5,0.75,1\}\). Then, the van der Waals interactions between the ligand and protein/solvent are removed in 16 steps with \(\lambda = \{0,0.05,0.1,0.2,0.3,0.4,0.5,0.6,0.65,0.7,0.75,0.8,0.85,0.9,0.95,1\}\). The spacing between the \(\lambda\)-points is chosen to have sufficient overlap between two neighbouring points. This procedure was performed for the ligand in water and in the protein to obtain the binding free energy. In order to maintain the average bonded conformation of the ligand in the protein we introduced position and orientation restraints acting on the ligand with respect to the protein. These restraints were introduced in three steps \(\lambda' = \{0,0.5,1\}\). Furthermore, a standard state correction was applied. For these simulations the md integrator was replaced by the sd integrator with the reference temperature set to 300 K. This ensured appropriate sampling of the decoupled state. The van der Waals switching parameters were changed to 1.0 and 1.1 nm and the PME real space cut-off to 1.2 nm.

Supporting Information. Data collection and refinement statistics of inhibitor-bound hCA II crystal structures; ESI-FTICR-MS calculated mass values of the peaks observed in the mass spectra; and the charges used for the ligands and the active site in MD. This material is available free of charge via the Internet at http://pubs.acs.org.

* Corresponding author. S.-A.P. Telephone: +61 7 3735 7825; Fax: +61 7 3735 7656; e-mail: s.poulsen@griffith.edu.au.

Acknowledgment. We thank Professor Emeritus Ian D Jenkins for helpful discussions. cDNA for hCA II was kindly provided by Prof Carol Fierke. This work was financed by the Australian Research Council (Grant number DP1110071 to S.-A.P.); an EU grant of the 7th framework programme (METOXIA project to C.T.S.). M.G.W. thanks the Humboldt Foundation for financial support. C.K.W. is supported by a NHMRC postdoctoral training Fellowship. We gratefully acknowledge the award of beam time by the Australian Synchrotron.
ESI-FTICR-MS, electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry; CA, carbonic anhydrase; ZBG, zinc binding group; PDB, Protein Data Bank; SAR, structure-activity relationships; BAMS, bioaffinity mass spectrometry; MD, molecular dynamics

References


JACS Authors note: Authors are required to submit a graphic entry for the Table of Contents (TOC) that, in conjunction with the manuscript title, should give the reader a representative idea of one of the following: A key structure, reaction, equation, concept, or theorem, etc., that is discussed in the manuscript. The TOC graphic may be **no wider than 9.0 cm** and **no taller than 3.5 cm** as the graphic will be reproduced at 100% of the submission size. A surrounding margin will be added to this width and height during *Journal* production.

Insert Table of Contents artwork here