

Chemical Science

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



Analysis of transition state mimicry by tight binding aminothiazoline inhibitors provides insight into catalysis by human O-GlcNAcase

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

N. Cekic,^{a,†} J. E. Heinonen,^{a,†} K. A. Stubbs,^{a,b} C. Roth,^c Y. He,^c A. J. Bennet,^a E. J. McEachern,^{a,†} G. J. Davies,^c D. J. Vocadlo^{a,d,*}

The modification of nucleocytoplasmic proteins with *O*-linked *N*-acetylglucosamine (*O*-GlcNAc) plays diverse roles in multicellular organisms. Inhibitors of *O*-GlcNAc hydrolase (OGA), the enzyme that removes *O*-GlcNAc from proteins, lead to increased *O*-GlcNAc levels in cells and are seeing widespread adoption in the field as a research tool used in cells and *in vivo*. Here we synthesize and study a series of tight binding carbohydrate-based inhibitors of human OGA (hOGA). The most potent of these 2'-aminothiazolines binds with a sub-nanomolar K_i value to hOGA (510 ± 50 pM) and the most selective has greater than 1 800 000-fold selectivity for hOGA over mechanistically related human lysosomal β -hexosaminidase. Structural data of inhibitors in complex with hOGA homologue reveals the basis for variation in binding among these compounds. Using linear free energy analyses, we show binding of these 2'-aminothiazoline inhibitors depends on the pK_a of the aminothiazoline ring system, revealing the protonation state of the inhibitor is a key driver of binding. Using series of inhibitors and synthetic substrates, we show that 2'-aminothiazoline inhibitors are transition state analogues of hOGA that bind to the enzyme up to 1-million fold more tightly than the substrate. These collective data support an oxazoline, rather than a protonated oxazolium ion, intermediate being formed along the reaction pathway. Inhibitors from this series will prove generally useful tools for the study of *O*-GlcNAc. The new insights gained here, into the catalytic mechanism of hOGA and the fundamental drivers of potency and selectivity of OGA inhibitors, should enable tuning of hOGA inhibitors with desirable properties.

Introduction

The modification of serine and threonine residues of nuclear and cytoplasmic proteins with terminal *O*-linked β -*N*-acetylglucosamine (*O*-GlcNAc)¹ has been found on hundreds of proteins.²⁻⁴ *O*-GlcNAc is present in all multi-cellular eukaryotes studied and occurs in a dynamic and reversible manner.⁵ Global *O*-GlcNAc levels have been shown to vary in response to cellular nutrient availability and stress and in some cases has been found to influence protein phosphorylation.⁶ These observations have stimulated interest in the physiological roles played by *O*-GlcNAc and research has implicated this modification in controlling various cellular processes

including, for example, proteosomal degradation of proteins⁷⁻⁹ and transcriptional regulation^{10, 11}. Additionally, a growing body of literature has implicated *O*-GlcNAcylation in chronic diseases such as neurodegeneration¹²⁻¹⁵ and cancer.^{9, 16, 17} Given the mounting potential in targeting protein *O*-GlcNAcylation for therapeutic benefit, there has been rising interest in understanding the molecular basis for inhibition of modulators of the *O*-GlcNAc pathway and the creation of small molecule modulators of this pathway for use in tissues¹⁸⁻²⁰.

The glycosyltransferase uridine diphospho-*N*-acetylglucosamine:peptide β -*N*-acetylglucosaminyl transferase (OGT) installs *O*-GlcNAc residues using uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc) as the sugar substrate donor.^{21, 22} The enzyme responsible for removing *O*-GlcNAc from proteins is *O*-GlcNAcase (OGA),²³ which is a member of glycoside hydrolase family 84 (GH84) of the CAZy classification system.²⁴ Consistent with the reversible nature of protein *O*-GlcNAcylation, inhibitors of these enzymes have been shown to induce time-dependent changes in cellular *O*-GlcNAc levels.^{19, 25-27} Small molecule inhibitors of OGA, in particular, have emerged as commonly used research tools for evaluating the phenotypic effects of increased *O*-GlcNAc levels in cultured cells, as well as *in vivo*.

^a Department of Chemistry, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada

^b School of Chemistry and Biochemistry, The University of Western Australia (M313), 35 Stirling Highway, Crawley WA 6009, Australia

^c York Structural Biology Laboratory, Department of Chemistry, The University of York, YO10 5DD, U.K.,

^d Centres of

^d Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada

[†] These authors contributed equally.

Electronic Supplementary Information (ESI) available: Detailed experimental procedures, NMR spectra and additional figures illustrating K_i graphs, kinetic data, pK_a titration data and structural analysis data. See DOI: 10.1039/x0xx00000x

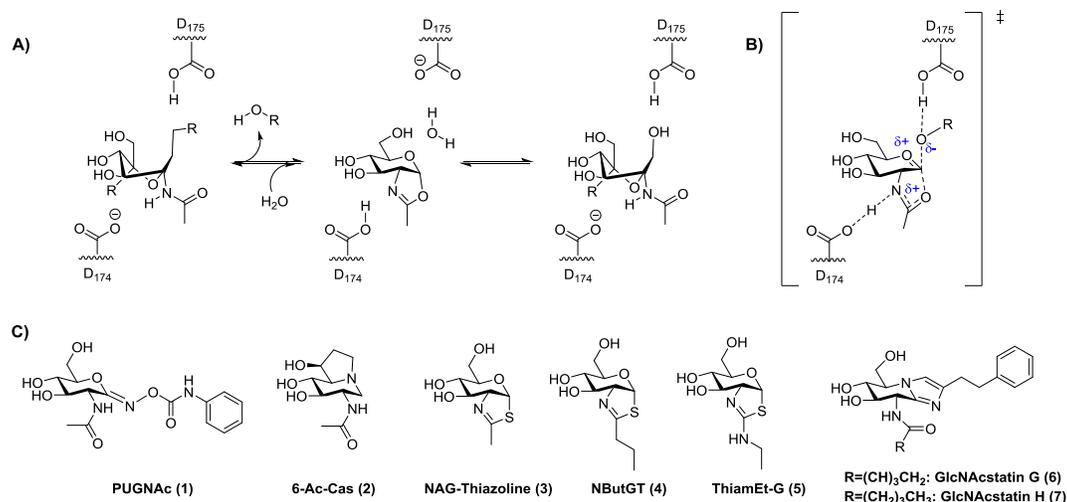


Figure 1. Catalytic mechanism of OGA and lysosomal β -hexosaminidases and some known hOGA inhibitors. **A)** hOGA uses a substrate-assisted catalytic mechanism involving two key catalytic aspartate residues that enable the transient formation of an oxazoline intermediate. R = leaving group. **B)** The proposed transition state (TS) for formation of the oxazoline intermediate. Note that the extent of proton transfer to the stabilizing residue D_{174}/D_{354} is not known. **C)** Some known hOGA inhibitors. Catalytic residues are D_{174}/D_{175} for human OGA and D_{354}/E_{355} for human lysosomal β -hexosaminidase A.

Interest in inhibitors of OGA has gained increasing attention due to growing recognition of the physiological roles of *O*-GlcNAc. Among the first reported inhibitors of OGA is *O*-(2-acetamido-2-deoxy-D-glucopyranosylidene) amino-N-phenylcarbamate (PUGNAc, Figure 1, **1**)²⁸ (hOGA K_i = 46 nM). This inhibitor, however, has well-described off target effects^{29, 30} including the inhibition of the functionally related lysosomal β -hexosaminidases HEXA and HEXB from family GH20.²⁶ These two lysosomal enzymes cleave β -linked terminal *N*-acetylhexosamine residues from various glycoconjugates including gangliosides. Genetic deficiency of these hexosaminidases results in Tay-Sachs and Sandhoff's disease, which stem from the accumulation of gangliosides within lysosomes. More recently identified inhibitors such as 6-acetamido castanospermine (6-Ac-Cas)²⁹ (hOGA K_i = 300 nM) (Figure 1, **2**), and 1,2-dideoxy-2'-methyl- α -D-glucopyranoso-[2,1-d]- Δ 2'-thiazoline (NAG-thiazoline)²⁶ (hOGA K_i = 70 nM) (Figure 1, **3**) while fairly potent, are also non-selective. Given that gangliosides play varied roles in cellular processes ranging from cell membrane structure to cell signaling, the selectivity of OGA inhibitors has emerged as being important for the creation of useful probe molecules for use *in vivo*.^{29, 31, 32}

Rationally designed OGA inhibitors have been pursued based on knowledge of the catalytic mechanism of hOGA. Detailed mechanistic studies^{26, 33} coupled with structural studies of bacterial homologues of hOGA^{34, 35} have provided clear support for a catalytic mechanism involving substrate-assisted catalysis in which the 2-acetamido group of the substrate serves as a catalytic nucleophile to generate a transient enzyme-bound oxazoline or oxazolinium intermediate (Figure 1A) and stabilizing

an oxocarbenium ion-like transition state (Figure 1B). In this two-step catalytic mechanism, two aspartates (Asp¹⁷⁴ and Asp¹⁷⁵) play key roles as general acid/base catalytic residues.³³ Asp¹⁷⁴ serves to orient and polarize the 2-acetamido group to aid its attack at the anomeric center, accepting a proton during formation of the oxazoline ring. Asp¹⁷⁵ acts as a general acid, donating a proton to the glycosidic oxygen during cleavage of the glycosidic bond (Figure 1A, B). Given the clear resemblance of NAG-thiazoline to the oxazoline intermediate, analogues of this molecule in which the 2'-position of the thiazoline ring is modified showed fair selectivities coupled with moderate nanomolar potencies as exemplified by NButGT (hOGA K_i = 230 nM, $K_i(\text{HEX})/K_i(\text{hOGA})$ = 1,500, Figure 1, **4**)²⁶.

More selective hOGA inhibitors have since been generated including the GlcNAcstatins^{27, 36} such as GlcNAcstatin C (hOGA K_i = 3.2 nM, $K_i(\text{HEX})/K_i(\text{hOGA})$ = 190) and a bioisostere of NButGT, 1,2-dideoxy-2'-ethylamino- α -D-glucopyranoso-[2,1-d]- Δ 2'-thiazoline (ThiamEt-G)(hOGA K_i = 21 nM, $K_i(\text{HEX})/K_i(\text{hOGA})$ = 37,000)¹⁹ (Figure 1C, **5**), which is more synthetically accessible. ThiamEt-G is orally available and increases brain *O*-GlcNAc levels in mammals and has been used by several groups to show chronic OGA inhibition and increased *O*-GlcNAcylation over several months has no apparent deleterious effects and also protects in various mouse models of AD against both tau^{14, 37, 38} and amyloid pathologies.^{39, 40} Given the great interest in OGA inhibitors as research tools, we aimed to explore the basis for the inhibition of hOGA by ThiamEt-G. Here we report on the synthesis and characterization of aminothiazoline inhibitors with

hOGA and human HexB (hHexB), demonstrate the strongly pK_a dependent inhibition of hOGA by such inhibitors, illuminate the molecular basis for observed selectivity and potency using structural biology, and reveal these inhibitors are genuine TS analogues – which reveals new insight into the nature of the catalytic mechanism as well as explaining the picomolar binding of the best representative from this inhibitor family.

Results and Discussion

Early studies showed that varying the 2'-alkyl substituent of NAG-thiazoline resulted in increased selectivity for OGA over the lysosomal hexosaminidases at the slight expense of potency.²⁶ This trend does not hold for PUGNAc analogues which show only modest selectivity.^{41,42} Exploiting a similar approach of increasing steric bulk of the acetamido group, however, cell-penetrant glucoimidazole inhibitors including the hOGA inhibitors GlcNAcstatin G (hOGA = 4.1 nM, $K_i(\text{HEX})/K_i(\text{hOGA}) = 900,000$) (Figure 1C, 6) and GlcNAcstatin H (hOGA = 2.6 nM, $K_i(\text{HEX})/K_i(\text{hOGA}) = 35,000$) (Figure 1C, 7), have been uncovered as some of the most selective hOGA inhibitors to date²⁰. Structural studies have suggested that these OGA inhibitors derive their selectivity from structural differences between the active sites of OGA and the lysosomal β -hexosaminidases in the region that serves to position the 2-acetamido group of the substrate.^{34, 43} We therefore first set out to gain an understanding of the detailed relationships between the size of substituents at the 2'-position of a series of 2'-aminothiazoline inhibitors and the influence on binding of altered electronic properties of the 2'-aminothiazoline system.

Preparation, potency, and selectivity of 2'-alkylaminothiazoline OGA inhibitors

Using three different approaches (Scheme 1) we synthesized a series of 2'-alkylaminothiazoline derivatives. Using the common intermediate hydrochloride salt of 1,3,4,6-*tetra-O*-acetyl-2-amino-2-deoxy- β -D-glucopyranose **8** (Schemes 1A and B), which was conveniently accessed in three steps⁴⁴, we prepared compounds **11a** and **11b** by reacting either *N*-fluorenylmethyloxycarbonyl (Fmoc)-protected isothiocyanate or *N*-allyl isothiocyanate with **8** in the presence of triethylamine to generate the respective thiourea intermediates **9a** and **9b**. Subsequent cyclization with excess SnCl_4 followed by a two step deprotection using catalytic NaOMe in anhydrous methanol, and piperidine catalyzed removal of the Fmoc group, afforded analogue **11a**. Cyclization of thiourea **9b** with excess TFA, followed by a one step deprotection using catalytic K_2CO_3 in anhydrous methanol afforded inhibitor **11b**.

Inhibitors **15a-h** were synthesized using an alternate route from the common isothiocyanate intermediate 1,3,4,6-*tetra-O*-acetyl-2-deoxy-2-isothiocyanato- β -D-glucopyranose (Scheme 1C, **12**), which was prepared from **8** via a biphasic reaction in $\text{H}_2\text{O}/\text{DCM}$ with thiophosgene.⁴⁴ Reaction of isothiocyanate **12**

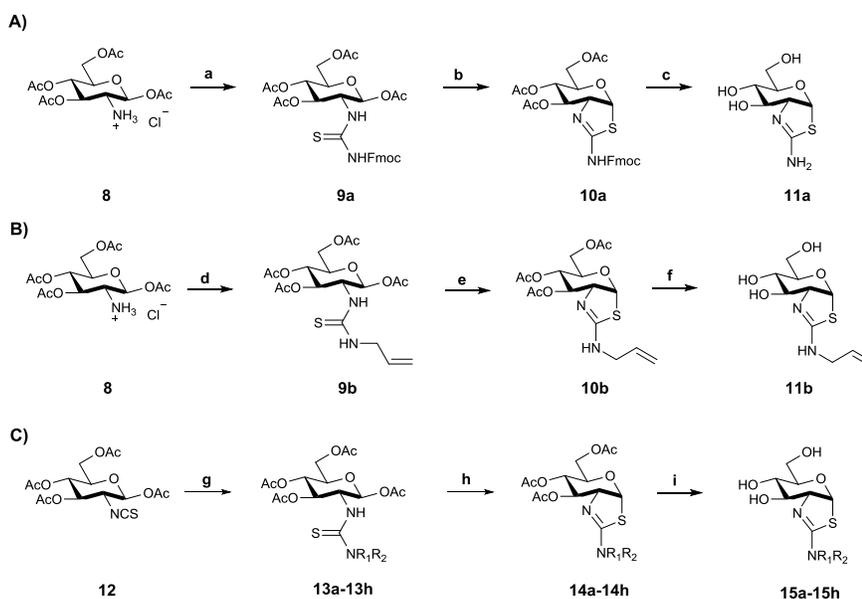
with a series of alkylamines and dialkylamines, or their respective hydrochloride salts, yielded thioureas **13a-h**. Acid catalyzed cyclization of these thiourea-containing compounds using TFA provided protected aminothiazolines **14a-h**, which after deprotection with K_2CO_3 afforded **15a-h**.

We then determined the K_i values for inhibition of hOGA by compounds **11a-b** and **15a-h**. Using Michael-Menten kinetics we obtained the K_i values for the less potent inhibitors **15e-h**, and as expected, double-reciprocal Lineweaver-Burk plots revealing a competitive mode of inhibition (Supplemental Figure S1). For the increasingly tight binding inhibitors we used the non-linear fitting method described by Morrison,⁴⁵ which can be used to determine K_i values which are comparable to the concentration of the enzyme being studied. Refinement of the Morrison approach by Copeland enables defining the K_i value for an inhibitor through the use of relative rates using a quadratic equation (Equation 1).⁴⁶ This approach depends on knowing the initial free enzyme and inhibitor concentrations without the assumption that the free inhibitor concentration is equal to the total inhibitor concentration.

$$v_i = v_0 \left[1 - \frac{([E]_T + [I]_T + K_i^{\text{app}}) - \sqrt{([E]_T + [I]_T + K_i^{\text{app}})^2 - 4[E]_T[I]_T}}{2[I]_T} \right] \quad (1)$$

Further, K_i values can be accurately determined over a wide range of enzyme concentrations using this method,⁴⁷ which enabled us to use hOGA concentrations that permit accurate initial rate determinations. We followed existing guidance⁴⁸ to select inhibitor concentrations for our K_i value determinations and confirmed the accuracy of this method by showing the K_i values for inhibitor **11b** were in reasonable accord when using either the Michaelis-Menten or Morrison method (Supplementary Figure S2).

Using these methods we found the K_i values for inhibition of hOGA by compounds **11a-b** and **15a-h** ranged from the high sub-nanomolar to low nanomolar range (Table 1). Notably, we find that the K_i value for ThiamEt-G (Figure 1, Figure 2) was 10-fold lower ($K_i = 2.1$ nM) than that previously determined¹⁹ using the Michaelis-Menten method ($K_i = 21$ nM). This makes Thiamet-G over 100-fold more potent than the isosteric NButGT ($K_i = 230$ nM). The most tight-binding compound with a K_i of 510 ± 50 pM (Figure 2) is **15a** (ThiamMe-G, Scheme 1), which ranks this compound among the most potent glycoside hydrolase inhibitors known, as well as the most potent selective hOGA inhibitor reported. Interestingly, we note that there is only a slight decrease in potency for hOGA upon increasing the volume of the 2'-aminoalkyl substituent to the point where the alkyl group is a propyl (**15d**, $K_i = 2$ nM). A butyl chain, however, leads to a greater than 100-fold loss of potency (**15e**, $K_i = 350$ nM). Structures of bacterial OGA homologues, in which the active site residues are completely conserved with hOGA,^{34, 35} show a discretely sized



Scheme 1. Synthesis of OGA inhibitors **11a-b** and **15a-h** from **8** and **12**, respectively. A) **11a**: a) 1) NEt_3 , DCM; 2) Fmoc-NCS, pyridine, NEt_3 ; b) SnCl_4 , pyridine, NEt_3 ; c) i) NaOMe, MeOH; ii) AcOH; 2) Piperidine, DMF; B) **11b**: d) allyl-NCS (2 eq), NEt_3 (2 eq), CH_3CN ; e) TFA (7.5 eq), DCM; f) K_2CO_3 , MeOH; C) **15a-h**: g) $\text{NHR}_1\text{R}_2\cdot\text{HCl}$ (1.2 eq), NEt_3 (1.2 eq), CH_3CN ; h) TFA (7.5 eq), DCM; i) K_2CO_3 , MeOH; **13a-15a**: $\text{R}_1 = \text{H}$, $\text{R}_2 = \text{CH}_3$; **13b-15b**: $\text{R}_1 = \text{CH}_3$, $\text{R}_2 = \text{CH}_3$; **13c-15c**: $\text{R}_1 = \text{H}$, $\text{R}_2 = \text{CH}_2\text{CH}_3$; **13d-15d**: $\text{R}_1 = \text{H}$, $\text{R}_2 = (\text{CH}_2)_2\text{CH}_3$; **13e-15e**: $\text{R}_1 = \text{H}$, $\text{R}_2 = (\text{CH}_2)_3\text{CH}_3$; **13f-15f**: $\text{R}_1 = \text{H}$, $\text{R}_2 = (\text{CH}_2)_2\text{F}$; **13g-15g**: $\text{R}_1 = \text{H}$, $\text{R}_2 = \text{CH}_2\text{CHF}_2$; **13h-15h**: $\text{R}_1 = \text{H}$, $\text{R}_2 = \text{CH}_2\text{CF}_3$

pocket having a volume that nicely accommodates the propyl substituent of **15d**.

To assess the structural basis for this stepped decrease in potency observed on going to the 2'-aminobutylthiazoline **15e**, we determined the structure of *Bacteroides thetaiocticron*, a bacterial homolog (*BtGH84*) of hOGA, in complex with the parent 2'-aminothiazoline **11a**, the tighter binding 2'-aminopropylthiazoline **11b**, and the butyl derivative **15e**, which shows greatly diminished binding. In all three structures, the respective inhibitor binds in the active site in a conserved mode with an invariant hydrogen bond pattern for the aminothiazoline moiety (Figure 3). The alkyl chain of the modified aminothiazolines points into the conserved hydrophobic pocket as shown for ThiamEt-G.¹⁹ An inspection of the binding pocket shows no further hydrophobic interactions in the case of the unsubstituted aminothiazoline **11a**, except a possible weak electrostatic interaction with C278 at the bottom of the pocket (Figure 3A). Propylene derivative **11b** matches the size of the pocket, requiring only a minor adjustment of the C278 rotamer and showing favorable hydrophobic interactions with W337, T310, and Y282 (Figure 3B). Even a small further extension of the alkyl chain, as in the butyl derivative **15e**, leads to steric clashes with multiple residues in the cavity, driving the side chain of C278 to adopt a different orientation. Furthermore, we observe the polypeptide main chain surrounding C278 is shifted away from the inhibitor. Thus, the size of the pocket disfavors

substituents longer than three carbon units in chain length, and those inhibitors having such larger groups induce unfavorable conformational changes within the active site (Figures 3C and D).

We next set out to assess the selectivity of this series of inhibitors for hOGA over the lysosomal β -hexosaminidases, which are comprised of combinations of α and β subunits that are products of the highly homologous HEXA and HEXB genes. Using purified human hexosaminidase B (hHexB) we determined the approximate K_i values for compounds **11a** to **15h** using Dixon plot analysis and found remarkably high inhibitor selectivities ranging from 1100- to 1 850 000-fold preference for hOGA (Table 1). We confirmed these Dixon plot analyses for the two most potent compounds, **11a** and **15a**, by determining full K_i values for their inhibition of hHexB. Remarkably, inhibitors **11a** and **15a** still retain 1100 and 3400-fold selectivity for hOGA despite their similarity in size to NAG-thiazoline, which itself demonstrated no selectivity²⁶. Accordingly, the presence of the 2'-amino substituent, on its own, confers at least 1000-fold selectivity for hOGA over hHexB.

We also noted that the selectivity ratio for hOGA increases as the 2'-aminoalkyl chain length increases to the three-carbon propyl (**15d**) and propylene (**11b**) derivatives, but this trend reverses once the chain length increases further, as seen for the 2'-aminobutyl analogue (**15e**). Since the active site pocket for hHexB is more constrained in the vicinity of the acetamido group, these observations are consistent with structural observations of

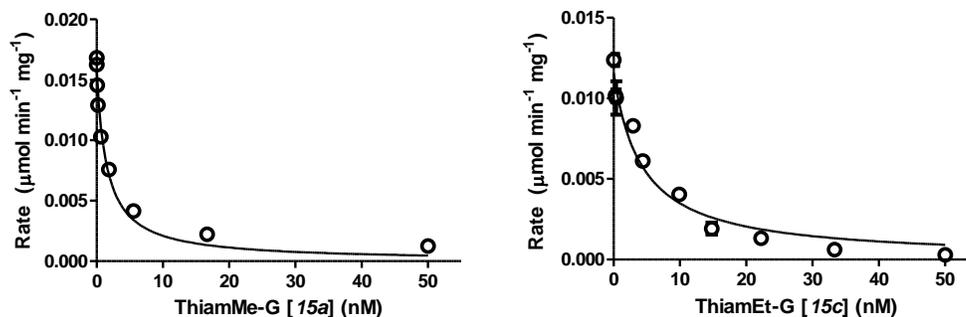


Figure 2. Representative Morrison data and fitted curves used to obtain K_i values for the tight binding hOGA inhibitors, shown for ThiamMe-G (**15a**) and ThiamEt-G (**15c**). Data was obtained in triplicate and error bars represent the standard error of the mean (S.E.M.).

both bacterial hOGA homologues³⁴ and hHexB⁴³. This suggests that once the 2'-substituent passes the volume that can be accommodated in the active site of hOGA, increases in bulk are even slightly more deleterious for hOGA as compared to hHexB.

Preparation of 2'-alkylaminothiazoline OGA inhibitors for evaluating electronic effects in OGA inhibition

Notably, ThiamEt-G ($K_i = 2.1$ nM) binds over 100-fold more tightly than NButGT ($K_i = 230$ nM) and is 25-fold more selective

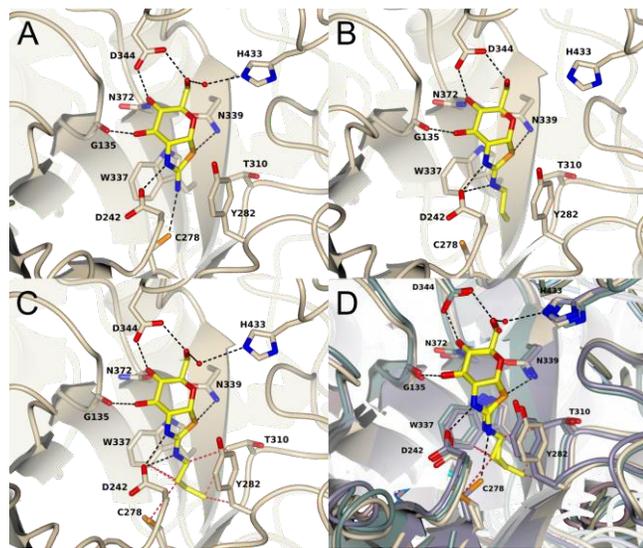


Figure 3. Cartoon plots of BtGH84 in complex with **11a** (A), **11b** (B) and **15e** (C). The inhibitors are shown in stick representation. The side chains of the residues forming the active site are shown in stick representation. Hydrogen bonds between the protein and the inhibitor are shown as black dashed lines. Close contacts ($d < 3.4$ Å) of butylaminothiazoline (C) are shown as red dashed lines. D shows an overlay of all three structures based on the bound aminothiazoline. The respective protein environment is coloured in darkblue for aminothiazoline containing structure, for propyleneaminothiazoline in bluegreen and butylaminothiazoline in beige. A change of the rotamer of C278 and a shift of the corresponding β -strand is observed in response to the steric repulsion with the butyl substituent.

for hOGA. Detailed mechanistic studies in combination with pH-rate profiles of wild-type and mutant hOGA revealed the key catalytic residue Asp¹⁷⁴ in the OGA catalytic site (Asp²⁴² in BtGH84) acts as a general base to assist the attack of the substrate 2-acetamido group onto the anomeric center. The kinetic pK_a of this residue was determined to be 5.2 so that at physiological pH this residue is expected to be in its carboxylate form³³ and therefore suitably ionized in the resting enzyme to facilitate catalysis. Given that aminothiazolines are known to be more basic than thiazolines, it was speculated that installation of the 2'-alkylamino group would increase the basicity of ThiamEt-G as compared to NButGT and thereby contribute to its enhanced potency relative to inhibitors bearing 2'-alkyl groups at physiological pH. Structural data of ThiamEt-G bound within the active site of BtGH84 is consistent with this proposal, revealing that both the endo and exocyclic amines engaged Asp242 of BtOGA.¹⁹

We set out to evaluate this proposal and assess the importance of the inhibitor pK_a on potency by studying a series of 2'-aminoethylthiazoline inhibitors with increasing fluorine substitution at the terminal methyl group (**15f-h**). Evaluation of the potency of these compounds revealed a progressive increase in K_i value upon increasing substitution with fluorine [$K_i = 2.1$ (CH₃, **15b**) 15 (CH₂F, **15f**), 60 (CHF₂, **15g**) and 1000 (CF₃, **15h**) nM] (Table 1). To clearly understand the relationship between inhibitor basicity and potency we used ¹³C NMR titration to determine the pK_a values of conjugate acids of these inhibitors (**15b**, **15f-h**). NMR methods are highly accurate and can be used to determine relative pK_a values between one or more compounds having an unknown pK_a and a reference compound having a well established pK_a value.⁴⁹

Practically, two advantages of this relative measurement approach is that it does not require repeated pH measurements throughout the titration and it can be used to evaluate small changes in pK_a values. In this way, a non-linear plot of the difference between NMR resonance frequencies for a compound of interest and those for a standard, such as 3-nitrophenol, can be used to determine the ratio (R) of the acid dissociation

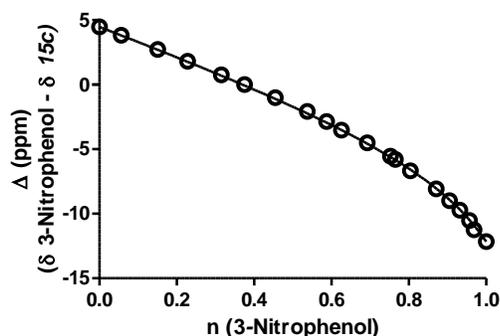


Figure 4. A representative example for ^{13}C -NMR determination of the pK_a value for ThiamEt-G (**15c**) with chemical shifts of the 2'-C of the thiazoline ring resonances (Δ ppm) for compounds **15c** as a function of the fractional protonation (n) of the reference compound, 3-nitrophenol (pK_a of 8.42). The solid line is the best non-linear least squares fit to equation S1 (supplementary information).

constants between these two materials, a typical plot for the difference in ^{13}C chemical shifts is shown in Figure 4 (see supplementary methods for a full discussion). To assess the extent to which inhibitor potency depends on its pK_a value, we plotted the pK_a value of inhibitors (**15c**, **15f-h** and NButGT, **4**) with the corresponding $\log K_i$, which are both free energy terms. The resulting linear free energy relationship (LFER) shows a linear correlation ($R^2 = 0.9876$) with a slope of -1.12 ± 0.09 (Figure 5), which is consistent with the pK_a value of inhibitors mostly dominating the effect of binding as compared to steric effects associated with increasing fluorine substitution. Notably, we also find that NButGT, which is isosteric to ThiamEt-G (**15c**), matches reasonably well within this correlation, supporting the electronic effects dominating this correlation. While the affinities of each protonation state of these inhibitors for hOGA (Figure 5) cannot be readily determined because the enzyme itself has various ionization states, these data collectively suggest that the pK_a value of the inhibitor, either by favoring the protonated inhibitor form or by optimizing hydrogen bonding strength, plays a key role in binding of these 2'-aminothiazoline inhibitors.

To assess the extent to which inhibitor potency depends on its pK_a value, we plotted the pK_a value of inhibitors (**15c**, **15f-h** and NButGT, **4**) with the corresponding $\log K_i$, which are both free energy terms. The resulting linear free energy relationship (LFER) shows a linear correlation ($R^2 = 0.9876$) with a slope of -1.12 ± 0.09 (Figure 5), which is consistent with the pK_a value of inhibitors mostly dominating the effect of binding as compared to steric effects associated with increasing fluorine substitution. Notably, we also find that NButGT, which is isosteric to ThiamEt-G (**15c**), matches reasonably well within this correlation, supporting the electronic effects dominating this correlation. These data collectively suggest that the pK_a value of the inhibitor, either by favoring the protonated inhibitor form or by optimizing

hydrogen bonding strength, plays a key role in binding of these 2'-aminothiazoline inhibitors.

Assessment of 2'-aminothiazoline inhibitors as transition state analogs.

The tight binding of these 2'-aminothiazoline inhibitors prompted us to consider their potency in the context of their size. One widely used parameter to understand the efficiency of binding as a function of molecular weight is to consider the ligand efficiency (LE) of a ligand. This measure provides the binding affinity of the compound as a measure of the number of heavy atoms. We calculate a remarkably high LE of 0.88 kcal/mol/heavy atom for compound **15a**. Such a LE is comparable to some of the highest ever observed LEs observed for compounds in the size range of between 10-50 heavy atoms⁵⁰ and suggests to us that these compounds could well be TS analogues, as had been observed for the related thiazoline inhibitors⁵¹. Tight-binding inhibitors that bear resemblance to enzyme substrates or intermediates have often been considered to be TS analogues simply by virtue of their potency. However, because enzymes are thought to catalyze reactions by tightly binding the TS, for genuine TS analogues changes in free energies of binding of a series of TS analogues ($\log K_i$) should parallel changes in the free energies of a series of related transition states TS ($\log k_{\text{cat}}/K_m$). Bartlett has formalized these concepts and methods to quantitatively assess whether compounds are TS analogues using LFERs.^{52, 53} Using this method, genuine TS analogues yield plots of $\log K_m/k_{\text{cat}}$ values, for a series of substrates having defined structural differences, versus $\log K_i$ values, for a series of inhibitors having the analogous structural changes, which show linear correlations having a slope of unity.

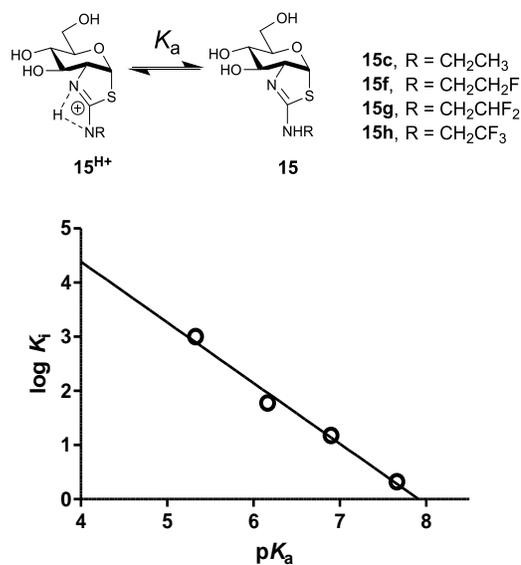


Figure 5. Acid dissociation of the 2'-aminothiazolium protonation state of 2'-aminothiazoline inhibitors **15c** and **15f-h** and the linear free energy relationship (LFER) analysis between the pK_a and $\log K_i$ of these inhibitors shows a negative correlation of -1.12 ± 0.09 .

Table 1. K_i selectivity ratios of inhibitors **11a** to **15h** for hOGA over hHexB and the pK_a values for **15c**, **15f–h**.

Inhibitor	hOGA K_i (nM) ^[a]	hHexB K_i (μ M) ^[b]	(hHexB/hOGA) ^[c]	pK_a ^[e]	Fraction protonated at pH 7.4
11a : $R_1 = R_2 = H$	4.7 ± 0.3	5.0 ± 0.6 ^[d]	1100		
15a : $R_1 = H, R_2 = CH_3$	0.51 ± 0.05	1.7 ± 0.19 ^[d]	3300		
15b : $R_1 = R_2 = CH_3$	2.4 ± 0.2	13.0 ± 3.8	5400		
15c : $R_1 = H, R_2 = CH_2CH_3$	2.1 ± 0.3	740 ± 60 ^[d]	350000	7.68	0.66
11b : $R_1 = H, R_2 = CH_2CHCH_2$	3.2 ± 0.4	2850 ± 570	950000		
15d : $R_1 = H, R_2 = (CH_2)_2CH_3$	2.0 ± 0.2	3700 ± 670	1850000		
15e : $R_1 = H, R_2 = (CH_2)_3CH_3$	350 ± 90 ^[d]	4800 ± 763	13700		
15f : $R_1 = H, R_2 = (CH_2)_2F$	15 ± 5 ^[d]	180 ± 44	12000	6.92	0.25
15g : $R_1 = H, R_2 = CH_2CHF_2$	60 ± 10 ^[d]	150 ± 50	2500	6.18	0.06
15h : $R_1 = H, R_2 = CH_2CF_3$	1000 ± 200 ^[d]	4200 ± 1525	4200	5.33	0.01

^[a]Determined using the Morrison K_i fit if the values are below 5-mM. ^[b]Determined using Dixon plot analysis. ^[c]Selectivity ratios representing the favored selectivity for hOGA compared to hHexB. ^[d]Determined using Michaelis-Menten inhibition analysis. ^[e]see supplementary information for full details, pK_a (NButGT) = 4.65.

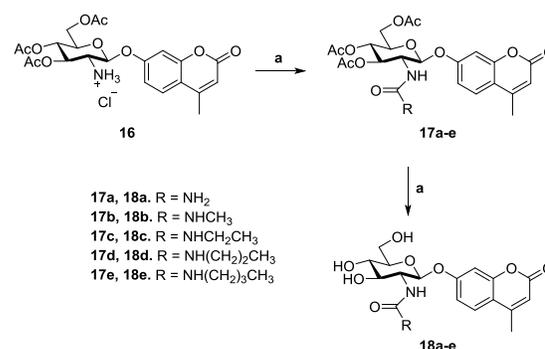
Furthermore, $\log K_m$ versus $\log K_i$ are not correlated for TS analogues but do correlate for substrate analogues.⁵²

Previous studies showed that NAG-thiazoline analogues are TS analogues despite their obvious resemblance to the oxazoline intermediate, perhaps due to the longer C-S bonds altering the thiazoline ring to resemble a late TS.⁵¹ Given the greater than 100-fold increase in potency we observe for the 2'-aminothiazoline inhibitors over their thiazoline counterparts and their structural resemblance to the oxazoline intermediate found along the reaction coordinate of hOGA, we wanted to assess whether incorporation of the 2'-amino group benefitted binding through serendipitous interactions, or whether the presence of the charge included in this class of inhibitors also made them TS analogues. We therefore turned to using the Bartlett LFER approach. With the series of inhibitors in hand we synthesized a series of fluorogenic 4-methylumbelliferyl 2-deoxy-2-urea- β -D-glucopyranoside substrates (**18a–e**, Scheme 2) bearing *N*-alkyl substituents on the terminal urea nitrogen that correspond to those alkyl groups present on the series of 2'-aminothiazoline inhibitors (**11a**, **15a**, **15c–e**). We started from 4-methylumbelliferyl 2-amino-2-deoxy- β -D-glucopyranoside hydrochloride (**16**) as a common intermediate⁵⁴. Per-*O*-acetylated urea substrates **17a–e** were prepared by reacting **16** with the appropriate alkylisocyanate in the presence of triethylamine, followed by Zemplen de-*O*-acetylation. With this series of substrates (**18a–e**) in hand we determined the k_{cat}/K_m values governing their hOGA catalyzed hydrolysis (Table 2). Plotting these data to assess TS analogy, we observe (Figure 6) a clear correlation ($R^2 = 0.9950$) with a slope of 1.08 ± 0.04 between $\log K_i$ values for the inhibitors versus the $\log K_m/k_{cat}$ values for the corresponding series of urea substrates (Figure 6A). In contrast, we find no correlation between $\log K_i$ values for

the inhibitors and $\log K_m$ values for the series of substrates (Figure 6B). These results indicate the 2'-aminothiazoline inhibitors are TS analogues for the hOGA catalyzed hydrolysis of urea substrates.

We recognized that our observations showed 2'-aminothiazoline inhibitors are TS analogues for the hOGA-catalyzed hydrolysis of unnatural urea substrates (**18a–e**), however, we were curious as to whether these inhibitors would also be TS analogues for the hOGA-catalyzed hydrolysis of the more natural *N*-acyl substrates. To address this question, we examined the correlation between $\log K_i$ values for the same series of 2'-aminothiazoline inhibitors and the series of *N*-acyl substrates (Figure 6C) having the analogous structural changes for which k_{cat}/K_m values are reported (Figure 6D).⁵¹ We find a fair correlation ($R^2 = 0.9768$) with a slope of 2.3 ± 0.3 . For this analysis we excluded the 2-aminothiazoline (**11a**) because of its unexpectedly poor inhibition of hOGA (Table 1).

Scheme 2. Synthesis of 4-methylumbelliferyl fluorogenic substrates **18a–e**.^a



^a a) R-NCO, NEt_3 , CH_3CN ; b) i) NaOMe, MeOH; ii) Dowex 50- H^+



Chemical Science

ARTICLE

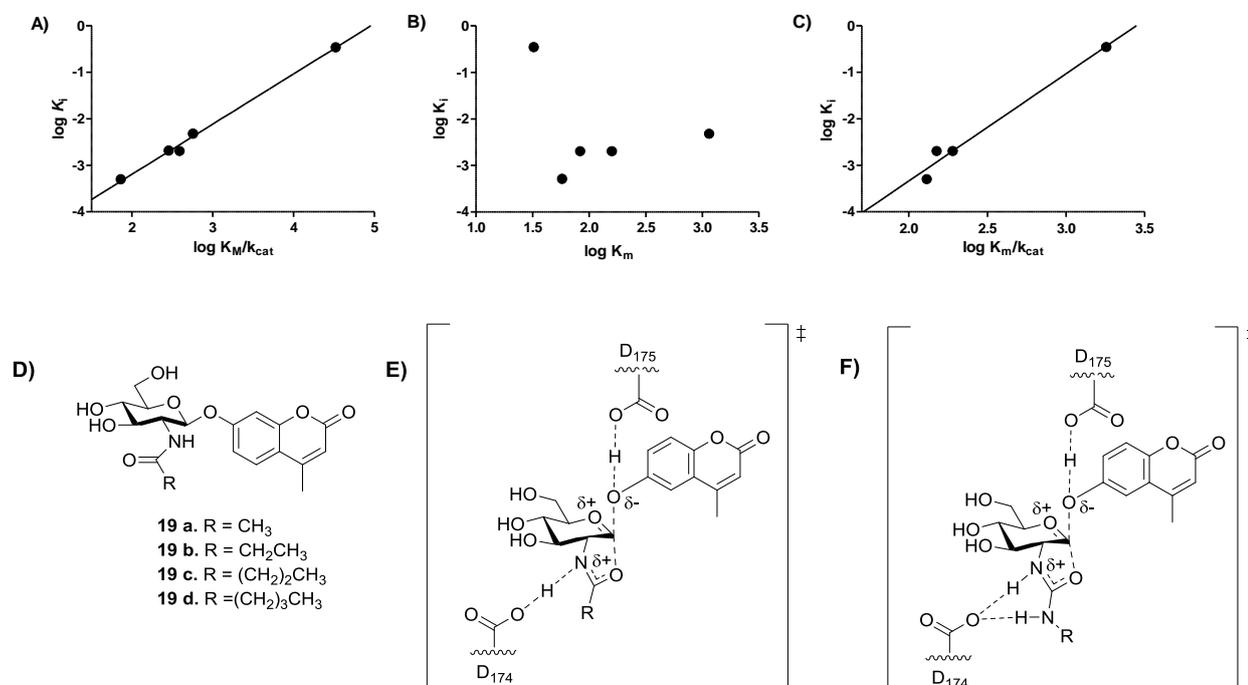


Figure 6. Analysis of transition state analogy for hOGA 2'-aminothiazoline inhibitors (**11a**, **15a**, **15c-e**) and urea (**18a-e**) and acyl (**19a-e**) substrates. A) The correlation between $\log K_m/k_{cat}$ of substrates **18a-e** and $\log K_i$ of inhibitors **11a**, **15a** and **15c-e** ($m = 1.08 \pm 0.04$, $R^2 = 0.9950$). B) The correlation between the $\log K_m$ of substrates **18a-e** and $\log K_i$ of inhibitors **11a**, **15a**, **15c-e**. C) The correlation between the $\log K_m/k_{cat}$ of substrates **19a-d** and $\log K_i$ of inhibitors **15a** and **15c-e** ($m = 2.31 \pm 0.25$, $R^2 = 0.9768$). D) Substrates **19a-d** synthesized by Whitworth *et al*⁵¹. E) Transition state for the hOGA catalyzed hydrolysis of substrates **19a-d** likely involves transfer of the amidic proton. F) Transition state for the hOGA catalyzed hydrolysis of substrates **18a-e** likely involves no transfer of the amidic proton.

Notably, previous studies have shown k_{cat} and k_{cat}/K_m values for the OGA-catalyzed hydrolysis of aryl 2-acetamido-2-deoxyglucopyranosides, including **19a**, vary in according to the pK_a value of the phenolic leaving group.⁵⁵ Additionally, such substrates bearing different sized N-acyl groups, including **19a-d**⁵¹ coupled with the large $\alpha\text{-D(V)}$ -KIE ($k_H/k_D = 1.14 \pm 0.02$) value observed for the OGA-catalyzed hydrolysis of *p*-nitrophenol 2-acetamido-2-deoxyglucopyranoside⁵⁵ all support k_{cat} and k_{cat}/K_m values reflect a chemical step being rate limiting for human OGA, which supports the validity of this TS analogy study.

The steep slope observed for this series of 2-acyl substrates in correlation with the aminothiazoline inhibitors is surprising. However, slopes other than unity in Bartlett LFER plots are preceded though not often rationalized.⁵⁶ Here, we interpret this steep slope as indicating that the TS for the hOGA catalyzed hydrolysis of *N*-acyl substrates bears less positive charge in the forming oxazoline ring system. This may arise because the amide proton is in flight in the TS, as compared to the *N*-urea substrates (Figure 6E and F), which are expected to be more basic and

therefore likely lead to the formation of 2'-amino-oxazolinium ion intermediates that retain their proton. Accordingly, the TS leading to such aminooxazolinium ion intermediates is expected to have more positive charge than the corresponding transition state leading to the oxazoline intermediate (Figure 6E and F). In keeping with this proposal, it is notable that site-directed deletion of the side chain of Asp¹⁷⁴, which is the catalytic general acid/base catalytic residue that interacts with the acetamido group of the substrate, leads to a similar drop of between 150 to 750-fold in second order rate constant³³ as seen on going from *N*-acyl to *N*-urea substrates (250-fold). Accordingly, these data indicate that 2'-aminothiazoline inhibitors are TS analogues for hOGA, by virtue of both their shape and general charge distribution. However, the steep slope observed for the LFER between $\log K_m/k_{cat}$ values observed for *N*-acyl substrates and the $\log K_i$ values seen for the 2'-aminothiazoline inhibitors suggests that this feature lends improved binding over the corresponding partial charge that likely develops for the TS found for the hOGA catalyzed processing of natural *N*-acyl-containing substrates. These data support a catalytic

Chemical Science Accepted Manuscript

mechanism in which residue D174 of hOGA acts as a general acid/base catalytic residue rather than simply stabilizing an oxazolinium ion intermediate as proposed for GH20 β -hexosaminidases⁵⁷.

Conclusion

In summary, we describe a series of aminothiazoline inhibitors for human OGA having picomolar and low nanomolar K_i values. The great potency of this inhibitor family is in large part attributable to their pK_a values since a clear correlation was observed between the pK_a and $\log K_i$ of a series of these compounds. Structures of these inhibitors in complex with BtOGA reveal the molecular basis for the trends in observed inhibitor potencies and selectivities. Using quantitative methods we find that these 2'-aminothiazoline inhibitors are tight-binding TS analogues for hOGA. These inhibitors benefit from their formal positive charge at physiological pH, harnessing favorable interactions that are only partly realized within the transition state for the natural 2-acyl-containing substrates. These observations should permit the design of more potent and selective inhibitors, not only in this class of inhibitor but also using other inhibitor scaffolds. Finally, the great potencies and selectivities of these inhibitors reveal a series of useful tool compounds that can be used to manipulate hOGA activity *in vivo*.

Acknowledgements

The authors thank Diamond Light Source for beamtime (proposals mx-1221, mx-7864 and mx-9948 and), and the staff of beamlines I02, I03 and I24 for assistance with crystal testing and data collection. We thank Dr. Michael Tropak and Prof. Donald Mahuran, Sick Children's Hospital, for the samples of human hexosaminidase B. This research was supported by funding from Canadian Institutes of Health Research for support (MOP-123341) (DJV), Biotechnology and Biological Sciences Research Council (BB/K003836/1) (GJD and CR), and the Natural Sciences and Engineering Research Council (NSERC, DG-121348) of Canada (AJB). DJV also thanks Brain Canada, Genome British Columbia, and the Michael Smith Foundation for Health Research for support of this research and the Canada Research Chairs program for support as a Tier I Canada Research Chair in Chemical Glycobiology.

References

1. C. R. Torres and G. W. Hart, *J Biol Chem*, 1984, **259**, 3308-3317.
2. N. Khidekel, S. B. Ficarro, E. C. Peters and L. C. Hsieh-Wilson, *Proc Natl Acad Sci, USA* 2004, **101**, 13132-13137.
3. S. A. Whelan and G. W. Hart, *Circ Res*, 2003, **93**, 1047-1058.
4. J. C. Trinidad, D. T. Barkan, B. F. Gullledge, A. Thalhammer, A. Sali, R. Schoepfer and A. L. Burlingame, *Mol Cell Proteom*, 2012, **11**, 215-229.
5. G. W. Hart, *Front Endocrin*, 2014, **5**, 183.
6. Z. Wang, M. Gucek and G. W. Hart, *Proc Natl Acad Sci, USA* 2008, **105**, 13793-13798.
7. H. B. Ruan, Y. Nie and X. Yang, *Mol Cell Proteom*, 2013, **12**, 3489-3497.
8. Y. Zhu, T. W. Liu, Z. Madden, S. A. Yuzwa, K. Murray, S. Cecioni, N. Zachara and D. J. Vocadlo, *J Mol Cell Biol*, 2016, **8**, 2-16.
9. C. M. Ferrer, T. P. Lynch, V. L. Sodi, J. N. Falcone, L. P. Schwab, D. L. Peacock, D. J. Vocadlo, T. N. Seagroves and M. J. Reginato, *Mol Cell*, 2014, **54**, 820-831.
10. S. Ozcan, S. S. Andrali and J. E. Cantrell, *Biochim Biophys Acta*, 2010, **1799**, 353-364.
11. S. M. Ranuncolo, S. Ghosh, J. A. Hanover, G. W. Hart and B. A. Lewis, *J Biol Chem*, 2012, **287**, 23549-23561.
12. F. Liu, J. Shi, H. Tanimukai, J. Gu, J. Gu, I. Grundke-Iqbal, K. Iqbal and C. X. Gong, *Brain*, 2009, **132**, 1820-1832.
13. F. Liu, K. Iqbal, I. Grundke-Iqbal, G. W. Hart and C. X. Gong, *Proc Natl Acad Sci, USA* 2004, **101**, 10804-10809.
14. S. A. Yuzwa, X. Shan, M. S. Macauley, T. Clark, Y. Skorobogatko, K. Vosseller and D. J. Vocadlo, *Nat Chem Biol*, 2012, **8**, 393-399.
15. S. A. Yuzwa and D. J. Vocadlo, *Chem Soc Rev*, 2014, **43**, 6839-6858.
16. Z. Ma and K. Vosseller, *Amino Acids*, 2013, **45**, 719-733.
17. W. Yi, P. M. Clark, D. E. Mason, M. C. Keenan, C. Hill, W. A. Goddard, 3rd, E. C. Peters, E. M. Driggers and L. C. Hsieh-Wilson, *Science*, 2012, **337**, 975-980.
18. R. F. Ortiz-Meoz, J. Jiang, M. B. Lazarus, M. Orman, J. Janetzko, C. Fan, D. Y. Duveau, Z. W. Tan, C. J. Thomas and S. Walker, *ACS Chem Biol*, 2015, **10**, 1392-1397.
19. S. A. Yuzwa, M. S. Macauley, J. E. Heinonen, X. Shan, R. J. Dennis, Y. He, G. E. Whitworth, K. A. Stubbs, E. J. McEachern, G. J. Davies and D. J. Vocadlo, *Nat Chem Biol*, 2008, **4**, 483-490.
20. H. C. Dorfmüller, V. S. Borodkin, M. Schimpl, X. Zheng, R. Kime, K. D. Read and D. M. van Aalten, *Chem Biol*, 2010, **17**, 1250-1255.
21. W. A. Lubas and J. A. Hanover, *J Biol Chem*, 2000, **275**, 10983-10988.
22. L. K. Kreppel, M. A. Blomberg and G. W. Hart, *J Biol Chem*, 1997, **272**, 9308-9315.
23. Y. Gao, L. Wells, F. I. Comer, G. J. Parker and G. W. Hart, *J Biol Chem*, 2001, **276**, 9838-9845.
24. V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho and B. Henrissat, *Nucl Acids Res*, 2014, **42**, D490-495.
25. R. S. Haltiwanger, K. Grove and G. A. Philipsberg, *J Biol Chem*, 1998, **273**, 3611-3617.
26. M. S. Macauley, G. E. Whitworth, A. W. Debowski, D. Chin and D. J. Vocadlo, *J Biol Chem*, 2005, **280**, 25313-25322.
27. H. C. Dorfmüller, V. S. Borodkin, M. Schimpl, S. M. Shepherd, N. A. Shpiro and D. M. van Aalten, *J Am Chem Soc*, 2006, **128**, 16484-16485.
28. D. Beer, J.-L. Maloisel, D. M. Rast and A. Vasella, *Helv Chim Acta*, 1990, **73**, 1918-1922.
29. M. S. Macauley, Y. He, T. M. Gloster, K. A. Stubbs, G. J. Davies and D. J. Vocadlo, *Chem Biol*, 2010, **17**, 937-948.
30. A. Mehdy, W. Morelle, C. Rosnoblet, D. Legrand, T. Lefebvre, S. Duvet and F. Foulquier, *J Biochem*, 2012, **151**, 439-446.
31. M. S. Macauley, X. Shan, S. A. Yuzwa, T. M. Gloster and D. J. Vocadlo, *Chem Biol*, 2010, **17**, 949-958.
32. M. S. Macauley, A. K. Bubb, C. Martinez-Fleites, G. J. Davies and D. J. Vocadlo, *J Biol Chem*, 2008, **283**, 34687-34695.

33. N. Cetinbas, M. S. Macauley, K. A. Stubbs, R. Drapala and D. J. Vocadlo, *Biochemistry*, 2006, **45**, 3835-3844.
34. R. J. Dennis, E. J. Taylor, M. S. Macauley, K. A. Stubbs, J. P. Turkenburg, S. J. Hart, G. N. Black, D. J. Vocadlo and G. J. Davies, *Nat Struct Mol Biol*, 2006, **13**, 365-371.
35. F. V. Rao, H. C. Dorfmüller, F. Villa, M. Allwood, I. M. Eggleston and D. M. van Aalten, *EMBO J*, 2006, **25**, 1569-1578.
36. H. C. Dorfmüller, V. S. Borodkin, M. Schimpl and D. M. van Aalten, *Biochem J*, 2009, **420**, 221-227.
37. P. Borghgraef, C. Menuet, C. Theunis, J. V. Louis, H. Devijver, H. Maurin, C. Smet-Nocca, G. Lippens, G. Hilaire, H. Gijssen, D. Moechars and F. Van Leuven, *PLoS one*, 2013, **8**, e84442.
38. D. L. Graham, A. J. Gray, J. A. Joyce, D. Yu, J. O'Moore, G. A. Carlson, M. S. Shearman, T. L. Dellovade and H. Hering, *Neuropharm*, 2014, **79**, 307-313.
39. S. A. Yuzwa, X. Shan, B. A. Jones, G. Zhao, M. L. Woodward, X. Li, Y. Zhu, E. J. McEachern, M. A. Silverman, N. V. Watson, C. X. Gong and D. J. Vocadlo, *Mol Neurodegen*, 2014, **9**, 42.
40. C. Kim, D. W. Nam, S. Y. Park, H. Song, H. S. Hong, J. H. Boo, E. S. Jung, Y. Kim, J. Y. Baek, K. S. Kim, J. W. Cho and I. Mook-Jung, *Neurobiol Aging*, 2013, **34**, 275-285.
41. K. A. Stubbs, N. Zhang and D. J. Vocadlo, *Org Biomol Chem*, 2006, **4**, 839-845.
42. E. J. Kim, M. Perreira, C. J. Thomas and J. A. Hanover, *J Am Chem Soc*, 2006, **128**, 4234-4235.
43. B. L. Mark, D. J. Mahuran, M. M. Cherney, D. Zhao, S. Knapp and M. N. G. James, *J Mol Biol*, 2003, **327**, 1093-1109.
44. M. Avalos Gonzalez, J. Fuentes Mota, I. M. Gomez Monterrey, J. L. Jimenez Requejo, J. C. Palacios Albarran and M. C. Ortiz Mellet, *Carb Res*, 1986, **154**, 49-62.
45. J. F. Morrison, *Biochim Biophys Acta*, 1969, **185**, 269-286.
46. R. A. Copeland, *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists*, John Wiley and Sons Inc., Hoboken, New Jersey, 2005.
47. P. Kuzmic, K. C. Elrod, L. M. Cregar, S. Sideris, R. Rai and J. W. Janc, *Analytical biochemistry*, 2000, **286**, 45-50.
48. D. J. Murphy, *Anal Biochem*, 2004, **327**, 61-67.
49. L. C. Perrin and M. A. Fabian, *Anal Chem*, 1996, **68**, 2127-2134.
50. I. D. Kuntz, K. Chen, K. A. Sharp and P. A. Kollman, *Proc Natl Acad Sci, USA* 1999, **96**, 9997-10002.
51. G. E. Whitworth, M. S. Macauley, K. A. Stubbs, R. J. Dennis, E. J. Taylor, G. J. Davies, I. R. Greig and D. J. Vocadlo, *J Am Chem Soc*, 2007, **129**, 635-644.
52. P. A. Bartlett and C. K. Marlowe, *Biochemistry*, 1983, **22**, 4618-4624.
53. M. M. Mader and P. A. Bartlett, *Chem Rev*, 1997, **97**, 1281-1301.
54. K.-R. Roeser and G. Legler, *Biochim Biophys Acta*, 1981, **657**, 321-333.
55. M. S. Macauley, K. A. Stubbs, D. J. Vocadlo, *J Am Chem Soc*, 2006, **127**, 17202-17203.
56. R. Mosi, H. Sham, J. C. Uitdehaag, R. Ruitenkamp, B. W. Dijkstra and S. G. Withers, *Biochemistry*, 1998, **37**, 17192-17198.
57. I. R. Greig, F. Zahariev and S. G. Withers, *J Am Chem Soc*, 2008, **130**, 17620-17628.