The development of a broad-spectrum retaining β-exo-galactosidase activity-based probe.

Kuo, C.L.; Su, Q.; Nieuwendijk, A.M.C.H. van den; Beenakker, T.J.M.; Offen, W.A.; Willems, L.I.; ... ; Aerts, J.M.F.G.

Organic & Biomolecular Chemistry

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d3ob01261a

The development of a broad-spectrum retaining β-exo-galactosidase activity-based probe†

Chi-Lin Kuo,[a] Qin Su,[a] Adrianus M. C. H. van den Nieuwendijk,[b] Thomas J. M. Beenakker,[a] Wendy A. Offen,[c] Lianne I. Willems,[a,c] Rolf. G. Boot,[a] Alexi J. Sarris,[b] André R. A. Marques,[a] Jeroen D. C. Codée,[b] Gijsbert A. van der Marel,[b] Bogdan I. Florea,[b] Gideon J. Davies,[c] Herman S. Overkleeft *[b] and Johannes M. F. G. Aerts*[a]

Received 9th August 2023,
Accepted 12th September 2023

DOI: 10.1039/d3ob01261a

rsc.li/obc

Acid β-galactosidase (GLB1) and galactocerebrosidase (GALC) are retaining exo-β-galactosidases involved in lysosomal glycoconjugate metabolism. Deficiency of GLB1 may result in the lysosomal storage disorders GM1 gangliosidosis, Morquio B syndrome, and galactosialidosis, and deficiency of GALC may result in Krabbe disease. Activity-based protein profiling (ABPP) is a powerful technique to assess the activity of retaining glycosidases in relation to health and disease. This work describes the use of fluorescent and biotin-carrying activity-based probes (ABPs) to assess the activity of both GLB1 and GALC in cell lysates, culture media, and tissue extracts. The reported ABPs, which complement the growing list of retaining glycosidase ABPs based on configurational isomers of cyclophellitol, should assist in fundamental and clinical research on various β-galactosidases, whose inherited deficiencies cause debilitating lysosomal storage disorders.

Introduction

Galactose is incorporated by higher eukaryotes into glycoproteins, glycolipids, and glycosaminoglycans (GAGs). Amongst the possible anomeric galactopyranosidic linkages in glycoconjugates, the predominant ones are the β-galactosides. In humans and other mammals, galactose is transferred from UDP-α-galactose to specific acceptors in the Golgi apparatus by the action of dedicated β-1,3- and β-1,4-galactosyltransferases and is removed from glycoconjugates predominantly by two distinct lysosomal β-galactosidases.[1] Acid β-galactosidase (GLB1, E.C. 3.2.1.23) degrades a variety of β-galactose-containing oligosaccharides, glycoproteins and glycolipids,[2] whereas galactocerebrosidase (GALC, E.C. 3.2.1.46) is much more selective and degrades mainly galactosylceramide.[3]

GLB1 belongs to glycoside hydrolase (GH) family 35.[4] Synthesized as a glycosylated 85 kDa precursor, it is targeted to lysosomes through the mannose-6-phosphate (M6P) dependent pathway, after which proteolytic processing of its C-terminus results in a mature enzyme composed of a 64 kDa subunit associated with a 20 kDa subunit.[5,6] The 64 kDa subunit contains the catalytic glutamates Glu188 and Glu268, while the 20 kDa subunit is required for the stabilization and functioning of the enzyme.[6–8] In lysosomes, GLB1 forms a complex together with cathepsin A (PPCA, *CSTA* gene) and neuraminidase (NEU1), resulting in a stable and efficient catalytic machinery for the degradation of GM1 ganglioside.[9] GALC belongs to the GH59 family.[4] It is synthesized as an 80 kDa precursor containing four N-linked glycans and targeted either directly to the lysosomes *via* the M6P pathway or by re-uptake of secreted proteins.[10,11] In lysosomes, it is processed to a mature enzyme consisting of an N-terminal 50 kDa subunit harbouring catalytic residues Glu198 and Glu274 and a C-terminal 30 kDa subunit.[11–13]

The importance of GLB1 and GALC activity is illustrated by diseases caused by their deficiency. Deficiency of GLB1 may result in three distinct lysosomal storage disorders, namely GM1 gangliosidosis, Morquio B syndrome, and galactosialidosis. GM1 gangliosidosis is caused by mutations in the *GLB1* gene and is characterized by an elevated cellular level of GM1 ganglioside, causing lysosomal vacuolization in lymphocytes and neuronal cells, which in turn leads to demyelination and neuronal cell death in the peripheral and central nervous systems.[14] Although the molecular mechanisms leading to

[a]*Department of Medical Biochemistry, Leiden Institute of Chemistry, Leiden University, P. O. Box 9502, 2300 RA Leiden, The Netherlands.*
E-mail: j.m.f.g.aerts@lic.leidenuniv.nl
[b]*Bioorganic Synthesis group, Leiden Institute of Chemistry, Leiden University, P. O. Box 9502, 2300 RA Leiden, The Netherlands.*
E-mail: h.s.overkleeft@lic.leidenuniv.nl
[c]*Department of Chemistry, University of York, Heslington, York, YO10 5DD York, UK*
† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3ob01261a

disease pathogenesis are not completely understood, it has been suggested that abnormal accumulation of GM1 ganglioside in the membranes linking the ER and mitochondria can drive Ca^{2+} efflux from the ER to mitochondria, leading to unfolded protein response (UPR) and mitochondrial stress, followed by apoptosis and ultimately neuronal death.[15–17] Different mutations in the *GLB1* gene may lead to Morquio B syndrome, where the primary accumulated substrates are oligosaccharides derived from keratin sulfate and glycoproteins.[2] No central nervous system involvement is observed in these patients, and the major affected tissue is that of the skeletal system. The third disease related to GLB1 deficiency is galactosialidosis and is caused by primary deficiency of cathepsin A which normally forms a protein complex with GLB1 and neuraminidase (NEU1) in lysosomes. Cathepsin A deficiency causes premature degradation of both GLB1 and NEU1, leading to their secondary deficiencies in patients.[9] Deficiency of either GALC or its activator protein, saposin A, forms the basis of Krabbe disease.[18] About 85% of Krabbe disease patients develop the infantile-onset form which presents with developmental delay and severe neurological damage, with death usually within a few years.[19] Later-onset forms have higher residual GALC activity and milder symptoms, and patients may have life spans up to the seventh decade of life. The primarily accumulated substrate is galactosylsphingosine (psychosine) in macrophages (globoid cells) and neural cells (particularly oligodendrocytes and Schwann cells), and this cytotoxic compound is believed to cause demyelination and cell death in both the central and peripheral nervous systems.[20,21] No therapy is yet available for any of these β-galactose-related lysosomal storage disorders besides bone marrow transplantation (BMT) and hematopoietic stem cell transplantation (HCT) for Krabbe disease, which is only effective before the onset of symptoms (which is usually not the case). Several therapeutic approaches are presently studied including substrate reduction therapies (SRT),[22,23] enzyme replacement therapies (ERT),[24,25] pharmacological chaperon therapies (PCT),[26,27] gene therapies[28–30] and lysosomal re-acidification therapy.[31] In addition, the inhibition of lysosomal acid ceramidase which generates the toxic galactosylsphingosine in Krabbe patients is considered as therapeutic intervention.[32]

Mechanistically, both GLB1 and GALC are retaining exo-glycosidases that hydrolyse their substrates following a Koshland double-displacement mechanism (Fig. 1A).[33] A covalent glycosidic bond is formed during the reaction itinerary, which makes both enzymes amenable for activity-based glycosidase profiling.[34] We have previously shown that β-galactose-configured cyclophellitol featuring a fluorophore at C-6 of the cyclophellitol scaffold (galactopyranose numbering) yielded an effective ABP selective for GALC.[35] We also found (Fig. 1B and the ESI†) that both galactose-configured cyclophellitol and cyclophellitol aziridine (compound **1**) led to covalent and irreversible inhibition of the GLB1 homologous retaining β-galactosidase from *Cellvibrio japonicus*, *Cj*GH35A.[36] The latter finding, combined with our previous studies[34] on configurationally diverse, nitrogen-substituted cyclophellitol aziridines as in-class, broad-spectrum retaining glycosidase activity-based probes, instigated the studies presented here.

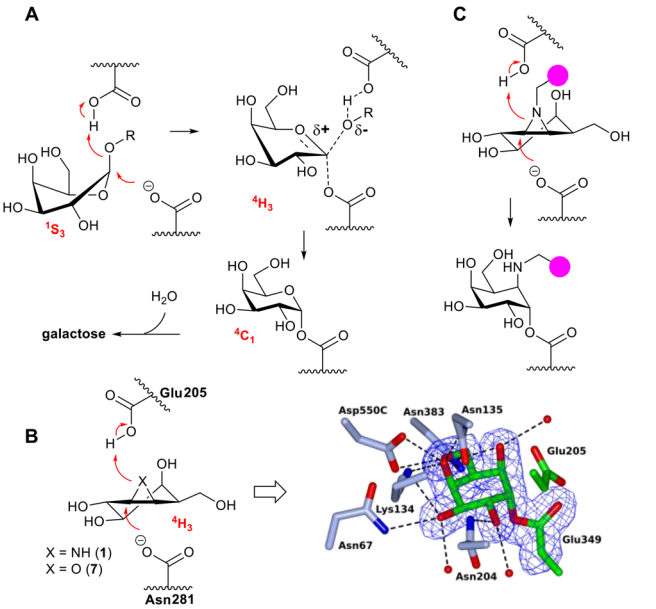
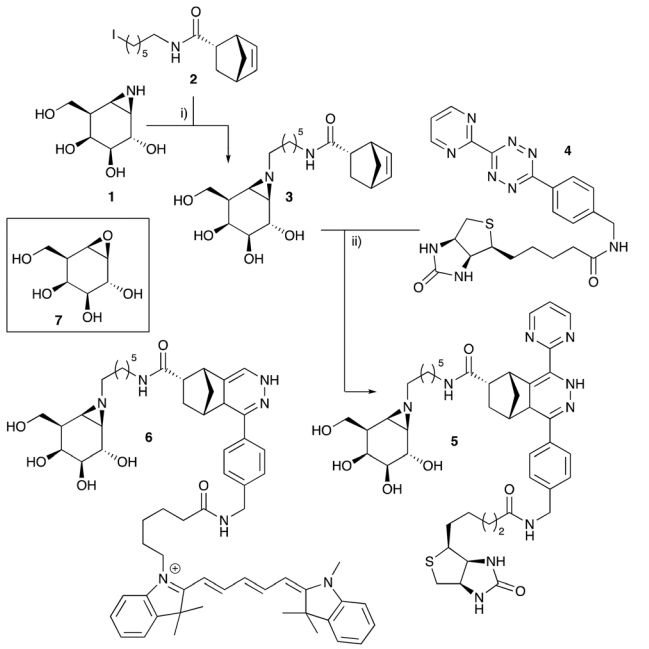


Fig. 1 (A) GALC and GLB1 are retaining β-galactosidases that process their substrates in a Koshland two-step double displacement mechanism, with the substrate β-galactoside undergoing a $^1S_3$–$^4H_3$–$^4C_1$ reaction itinerary in the first half of the catalytic cycle leading to the covalent enzyme–substrate intermediate. (B) Structure of the *Cellvibrio japonicus* retaining β-galactosidase *Cj*GH35A, which reacts with β-galactose-configured cyclophellitol (**7**) (galactose-configured cyclophellitol aziridine reacts in the same manner[36]). β-*Gal*-Cyclophellitol adopts a $^4H_3$-configuration emulating that of the transition state of *Cj*GH35A-mediated substrate processing. Nucleophilic opening of the epoxide by the *Cj*GH35A Asn281 nucleophile then leads to a $^4C_1$-bound enzyme–inhibitor complex which by nature of the more stable (ester as opposed to acylal that would emerge from β-galactoside processing) enzyme–inhibitor linkage is stable over time. (C) Substituted β-galactose-configured cyclophellitol aziridines as β-galactosidase activity-based probes, the subject of the here-presented studies (the purple bulb represents biotin or a fluorophore).

Here we present the synthesis of *N*-alkylated β-galactose-configured cyclophellitol aziridines. We show that the introduction of a fluorophore for in-gel activity-based protein profiling and biotin for chemical proteomics, namely Cy5-ABP **6** and biotin-ABP **5**, allows inhibition of both GLB1 and GALC and facilitates the identification of both β-galactosidases from complex biological samples.

Results and discussion

The synthesis of biotin-ABP **5** is depicted in Scheme 1. In our previous cyclophellitol aziridine-based ABP designs[34] we connected the inhibitor and reporter parts through a copper(I)-catalysed [2 + 3] alkyne–azide cycloaddition, with the azide connected to the aziridine nitrogen through a six–seven carbon alkyl chain and the alkyne to the reporter (biotin, fluorophore).



Scheme 1 (i) DiPEA, DMF, 90 °C, 14%. (ii) MeOH, CH_2Cl_2, 39%.

For reasons that are yet elusive to us this procedure proved abortive for the β-galactose-configured cyclophellitol aziridine, which degraded under the copper-catalysed click conditions. We therefore resorted to inverse electron demand Diels–Alder (IEDDA) as a metal-free, mild ligation chemistry to connect the mechanism-based inhibitor to reporters. Thus, β-D-*galacto*-cyclophellitol aziridine (**1**), which we prepared as reported previously[37] following a strategy adapted from that reported by Llebaria and colleagues,[38] was alkylated with iodide (**2**), prepared in two steps from 6-aminohexanol and featuring a norbonene as the IEDDA dienophile, to give norbonene species **3**.[39,40] The ensuing IEDDA ligation with biotin-tetrazine (**4**) afforded an ABP (**5**) in 39% yield based on aziridine (**2**) after HPLC purification. A fluorescent ABP (**6**) was prepared by IEDDA ligation of **3** with the appropriate Cy5-tetrazine as published previously.[40] It should be noted that both ABPs were obtained as a mixture of stereoisomers because of the enantiomeric mixture of norbonenes (**3** or **6**) we used, as well as the asymmetric tetrazine used. The ABPs were used as a mixture based on the thought that the targeted β-galactosidases, like most exoglycosidases, are very particular to the configuration of the sugar (like) substrate (here, β-*galacto*-cyclophellitol aziridine, which is enantiomerically pure), but less so to the aglycon.

ABPs **5** and **6** were next tested for their inhibitory potency towards GLB1 and GALC in comparison with free aziridine **1**, norbonene intermediate **3**, and D-*galacto*-cyclophellitol **7**. For this purpose, a 4-methylumbelliferyl β-D-galactopyranoside (4-MU-β-D-Gal) substrate assay was developed first, with a focus on pH optimum.[41] GLB1 in human fibroblast lysates and mouse GALC (collected culture medium of HEK293T cells overexpressing this protein) were used for this purpose and 11 μM $AgNO_3$ was added to selectively inhibit GLB1 over GALC when aiming for GALC activity assessment. The β-galactosidase activity in both materials had a pH optimum of 4.5, but at pH 6.0 the culture medium retained 60% activity, in contrast to the 20% remaining β-galactosidase activity in the fibroblast lysates (Fig. S1A†). In the presence of $AgNO_3$, no activity remained in the fibroblast lysates, while the culture medium retained full β-galactosidase activity (Fig. S1B†), indicating that the fibroblast lysates contained only GLB1 and that the culture medium contained only GALC. The inhibitory potency (IC_{50}) of **1**, **3**, and **5**–**7** towards GLB1 and GALC as measured in the above-described 4-MU-β-D-Gal assays is depicted in Table 1. Compounds **1** and **5** proved to be low nanomolar inhibitors of both GLB1 and GALC, while compound **3** is revealed as the weakest inhibitor, although still with an IC_{50} below 100 nM for both enzymes. Importantly, both biotin-ABP **5** and Cy5-ABP **6** inhibit both enzymes potently, inviting their assessment in activity-based protein profiling assays to visualize and identify retaining β-galactosidases in complex biological samples.

Labeling of mouse kidney extracts, containing both GALC and GLB1, with Cy5-ABP **6** was performed in samples pre-incubated with a broad-spectrum β-glucosidase suicide inhibitor (see for its structure the ESI†).[42] Pre-incubation with this compound was done because we expected from previous studies that β-galactose-configured cyclophellitols are cross-reactive (although at higher concentrations) towards the lysosomal retaining β-glucosidase GBA1 (lysosomal glucosylceramidase) and, to a lesser extent, GBA2 (cytosolic glucosylceramidase). Optimal labeling of both GLB1 and GALC occurred at 1 μM Cy5-ABP **6** (Fig. 2A), with 30 min incubation time (Fig. 2B), and at pH from 4.0 to 5.0 (Fig. 2C). Labeling was partially abrogated by pre-incubating the samples with 4-MU-β-D-Gal, consistent with enzyme active site labeling (Fig. 2D).

As the last experiment, the ability of biotin-ABP (**5**) to identify GLB1 and GALC, as well as other (off) targets, was assessed by treatment of mouse kidney extracts with either **5** at 20 μM final concentration or a vehicle at pH 4.5 for 1.5 hours. Probe-reacted proteins, now endowed with biotin, were enriched by streptavidin pull-down and next on-bead-digested by addition of trypsin. The digested peptides were eluted from the beads and analysed by LC-MS/MS separation and fragmentation. Comparison of the data obtained from probe **5**-treated and vehicle-treated samples yielded six proteins which were identified in the probe-treated samples exclusively. These proteins are listed in Table 2 and include, besides the expected GLB1

Table 1 Apparent IC_{50} values (nM) of compounds towards the retaining β-exogalactosidases GLB1 and GALC. Error range = ± SD, *n* = 2 biological replicates

Compound	GLB1	GALC
1	2.55 ± 0.59	5.57 ± 0.36
3	57.8 ± 3.05	98.6 ± 20.8
5	4.79 ± 4.09	9.95 ± 5.73
6	14.6 ± 0.98	61.0 ± 6.89
7	21.7	39.1

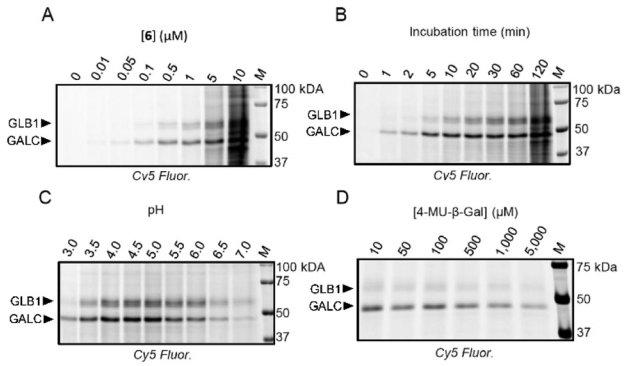


Fig. 2 Labeling of mouse kidney extracts with Cy5-ABP (**6**). (A) Concentration-dependent labeling. (B) Time-dependent labeling. (C) pH-dependent labeling. (D) ABPP using 4-MU-β-D-Gal as the competitor.

Table 2 List of (putative) glycosidases from mouse kidney extracts treated with ABP **5** and identified by LC-MSMS. Counts: number of tryptic peptides identified per protein. MANBA: β-mannosidase; BGAL: β-galactosidase; GLBL2: β-galactosidase-1-like protein 2; GLB1L: β-galactosidase-1-like protein 1; GALC: galactosylceramidase; GBA1: lysosomal acid glucosyl-ceramidase

Accession no.	Entry	Peptide counts	Sequence coverage	Score
Q8K2I4	MANBA_MOUSE	32	39.1%	323
P23780	BGAL_MOUSE	19	37.2%	230
Q3UPY5	GLBL2_MOUSE	14	25.9%	121
Q8CV60	GLB1L_MOUSE	21	41.6%	118
P54818	GALC_MOUSE	11	24.3%	95
P17439	GBA1_MOUSE	13	30.9%	54

and GALC, also β-mannosidase (MANBA) and lysosomal glucosylceramidase (GBA1). MANBA is responsible for the cleavage of the core Man-β-1,4-GlcNAc linkage during lysosomal turnover of *N*-glycoproteins.[43] We previously found that MANBA also reacts with retaining β-glucosidase targeting probes[44] and it appears that the MANBA active site is flexible and can accommodate, besides mannopyranose-configured ligands, also glucopyranose and galactopyranose ones. GBA1, whose physiological role is to degrade glucosylceramidase and thus is the penultimate enzyme in the lysosomal turnover of glycosphingolipids, has previously been shown to be cross-reactive towards β-galactopyranose-configured ligands – substrates, competitive inhibitors and mechanism-based irreversible inhibitors. Its enrichment with galactose-configured ABP **5** is therefore not surprising and matches the outcome of our gel-based experiments. The other two unique targets are GLB1-like proteins 1 and 2 (GLB1L and GLBL2), both of which share high homology with BGAL. Both proteins are classified in the GH35 family; however, no physiological role has yet been ascribed to either of these proteins. The fact that they react with our probes suggests that GLB1L and GLBL2 are indeed *bona fide* β-galactosidases, which may help in further uncovering their biological function.

Conclusion

This work reveals that the cyclophellitol aziridine scaffold, which in our hands proved useful to visualise and identify a range of retaining exo- and endoglycosidases, provides also a suitable starting point for the design of activity-based probes for the retaining β-galactosidases GLB1 and GALC. Matching the configuration of the cyclophellitol scaffold to that of the GLB1/GALC sugar substrates yields inhibitors and probes that potently inhibit the target galactosidases and also have sufficient selectivity to identify these enzymes from complex biological samples. Target selectivity is not complete as the inhibitors/ABPs also target the retaining β-glucosidases GBA1 and GBA2, but preincubation with complementary β-glucose-configured cyclophellitol inhibitors/probes allows for discrimination between the two retaining GH families. In our studies on differently configured retaining glycosidase ABPs we grafted biotin/fluorophores onto the cyclophellitol aziridine scaffold by using a copper(I)-catalysed azide–alkyne [2 + 3] click ligation, with the azide installed onto the aziridine and the alkyne on the reporter. This approach proved abortive here because the β-galactose-configured cyclophellitol aziridine proved unstable under the click ligation conditions. Reverting to inverse electron demand Diels–Alder (IEDDA) ligation conditions did yield the desired probes that, although obtained as diastereomeric mixtures with respect to the Diels–Alder adduct parts of the molecules, behaved otherwise as expected in biological sample labeling and pull-down. In all, the ABPs presented here add to the growing list of retaining glycosidase ABPs which can be used to visualize and identify retaining β-galactosidases irrespective of their biological source. Specific to the work presented here, we put forward our ABPs for monitoring GLB1 and/or GALC in relation to the inherited lysosomal storage disorders related to genetic mutations in these enzymes and also for monitoring the efficacy of drug treatment for these, in the context of GM1 gangliosidosis, Morquio B syndrome, galactosialidosis and Krabbe disease.

Experimental

Chemicals for biological assays

Chemicals were obtained from Sigma-Aldrich, if not otherwise indicated. Protein concentration was measured using a PierceTM BCA assay kit (Thermo Fisher).

Synthesis

Synthesis of β-Gal-aziridine ABP 5. To a solution of aziridine **3**[40] (8.0 mg, 20.3 μmol) in methanol (0.5 mL) was added a solution of biotin-tetrazine **4**[45] (10.2 mg, 20.7 μmol) in DCM/MeOH (1 : 1, 2.5 mL). The resulting red/purple solution was stirred at room temperature while slowly fading to become almost colourless. LCMS analysis after 3 hours shows the target conjugate **5** as a mixture of isomers with trace amounts of both starting materials **3** and **4** present. The mixture was left stirring overnight after which no further conversion was

observed by LCMS. The solvents were evaporated and ^{1}H NMR analysis showed the presence of *circa* 10 mol% of tetrazine **4** co-eluting with the target compound on HPLC. The crude material was dissolved in MeOH (1 mL) and treated with 0.40 mL of a methanolic solution (10 mg mL^{-1}) of (bicyclo [2.2.1]hept-5-en-2-yl)methyl *tert*-butyldimethylsilyl ether[44] to remove unreacted tetrazine before HPLC purification of aziridine probe **5**. HPLC purification afforded two fractions, each containing two isomers, of probe **5** in a total yield of 39%.

Fraction 1 (2.77 mg after lyophilization, major isomers reported (1 : 1 ratio)): ^{1}H NMR (600 MHz, MeOD) δ 8.84 (dd, J = 4.9, 1.0 Hz, 2H, Ar H-21, H-23), 7.43 (t, J = 4.9 Hz, 1H, Ar, H-22), 7.40–7.35 (m, 2H, Ar H-28, H-30), 7.32 (dd, J = 8.4, 1.8 Hz, 2H, Ar H-29, H-31), 4.52–4.47 (m, 1H, H-41), 4.46–4.34 (m, 2H, CH_2-33), 4.28 (ddd, J = 12.0, 7.9, 4.5 Hz, 1H, H-40), 3.91 (d, J = 7.9 Hz, 0.5H, H-2), 3.87 (d, J = 7.9 Hz, 0.5H, H-2), 3.82 (ddd, J = 10.4, 7.0, 1.7 Hz, 1H, H-6), 3.79–3.75 (m, 1H, H-6′), 3.73 (t, J = 2.3 Hz, 1H, H-4), 3.53–3.36 (m, 1H, H-13), 3.27 (d, J = 3.6 Hz, 1H, H-18), 3.25–3.15 (m, 3H, H-3, H-13′, H-38), 3.10 (dt, J = 4.3, 1.4 Hz, 1H, H-16), 2.99–2.90 (m, 2H, H-14, H-39), 2.88 (d, J = 2.4 Hz, 1H, H-15), 2.71 (t, J = 12.0 Hz, 1H, H-39′), 2.50–2.39 (m, 1H, H-8), 2.28 (td, J = 7.3, 3.9 Hz, 2H, CH_2-34), 2.22–2.13 (m, 1H, H-8′), 2.02–1.96 (m, 2H, H-1, H-5), 1.92–1.86 (m, 2H, CH_2-19 (obscured by solvent)), 1.79 (t, J = 5.7 Hz, 1H, H-7), 1.77–1.53 (m, 8H, 4 × CH_2), 1.50–1.39 (m, 7H, H-20, 3 × CH_2), 1.01 (d, J = 8.5 Hz, 1H, H-20′). ^{13}C NMR (151 MHz, MeOD) δ 176.66, 175.95 & 175.89 (C-44, C-43), 166.12 (C-42), 163.88 (C-24), 158.27 (C-21, 23), 140.57 & 140.53 (C-25), 136.73 & 136.68 (C-26), 134.46 & 134.44 (C-32), 133.47 (C-27), 129.17 (C-28), 129.13 (C-30), 128.65 (C29, 31), 121.20 (C-22), 114.52 (C-17), 78.44 & 78.42 (C-3), 72.53 & 72.46 (C-2), 72.32 (C-4), 63.38 (C-40), 62.65 (C-6), 61.64 & 61.62 (C-41), 60.72 & 60.68 (C-8), 57.14 & 57.04 (C-38), 48.21 (C-14), 47.91 (C-16), 44.96 & 44.94 (C-7), 44.78 (C-20), 43.79 & 43.76 (C-33), 43.02 & 43.00 (C-15), 42.50 (C-1), 41.79 (C-5), 41.10 (C-39), 40.81 (C-18), 40.78 (C-13), 36.77 & 36.60 (C-34), 33.30 (C-19), 30.55 & 30.53, 30.48, 29.80 & 29.71, 29.50 & 29.44, 28.22 & 28.18, 28.07, 26.96 & 26.85 (7 × CH_2, C-9–C-12 and C-35–C-37). HRMS calculated for $[C_{44}H_{59}N_9O_7S + H]^+$ 858.4331; found 858.4331.

Fraction 2 (4.09 mg after lyophilization, major isomers reported (1 : 1 ratio)): ^{1}H NMR (600 MHz, MeOD) δ 8.91 (dd, J = 4.8, 1.2 Hz, 2H, Ar H-21, H-23), 7.49 (d, J = 8.2 Hz, 2H, Ar H-28, H-30), 7.45 (tdd, J = 4.8, 2.0, 0.7 Hz, 1H, Ar, H-22), 7.36 (d, J = 8.2 Hz, 2H, Ar H-29, H-31), 4.49 (ddd, J = 7.9, 5.0, 0.9 Hz, 1H, H-41), 4.46–4.36 (m, 2H, H-33), 4.28 (dt, J = 7.9, 4.0 Hz, 1H, H-40), 4.17 (d, J = 3.1 Hz, 1H, H-15), 3.89 (dd, J = 9.0, 7.9 Hz, 1H, H-2), 3.84–3.78 (m, 1H, H-6), 3.78–3.74 (m, 1H, H-6′), 3.73–3.70 (m, 1H, H-4), 3.25–3.14 (m, 4H, H-3, H-13, H-14, H-38), 3.01 (m, 1H, H-13′), 2.93 (ddd, J = 12.8, 5.0, 2.2 Hz, 1H, H-39), 2.79 (d, J = 4.4 Hz, 1H, H-18), 2.71 (dd, J = 12.6, 2.2 Hz, 1H, H-39′), 2.63 (br s, 1H, H-16), 2.37–2.30 (m, 1H, H-8), 2.28 (t, J = 7.3 Hz, 2H, CH_2-34), 2.18 (td, J = 12.5, 4.7 Hz, 1H, H-19), 2.13–2.03 (m, 1H, H-8′), 2.00–1.94 (m, 2H, H-1, H-5), 1.86 (dt, J = 12.6, 3.8 Hz, 1H, H-19′), 1.77 (dd, J = 5.7, 4.5 Hz, 1H, H-7), 1.75–1.58 (m, 4H, CH_2-12, CH_2-35), 1.57 (dd, J = 9.7, 1.7 Hz, 1H, H-20), 1.44 (p, J = 7.7 Hz, 2H, CH_2-36), 1.41–1.27 (m, 4H, CH_2-9, CH_2-37), 1.26–1.06 (m, 4H, CH_2-10, CH_2-11), 1.04 (dt, J = 9.7, 2.2 Hz, 1H, H-20′). ^{13}C NMR (151 MHz, MeOD) δ 176.78 (C-43), 175.98 (C-44), 166.13 (C-42), 160.39 & 160.38 (C-24), 158.63 (C-21, 23), 143.25 & 143.23 (C-25), 141.08 (C-26), 135.72 & 135.71 (C-32), 131.85 (C-27), 129.95 (C-28, 30), 128.41 (C-29, 31), 122.72 (C-17), 121.10 (C-22), 78.44 & 78.43 (C-3), 72.52 & 72.51 (C-2), 72.26 (C-4), 63.37 & 63.35 (C-40), 62.61 & 62.59 (C-6), 61.62 (C-41), 60.64 (C-8), 57.06 & 57.03 (C-38), 48.20 (C-14), 45.88 (C-16), 44.92 & 44.90 (C-7), 44.41 & 44.38 (C-15), 44.19 (C-20), 43.82 (C-33), 42.46 (C-1), 41.78 & 41.76 (C-5), 41.08 (C-39), 40.33 (C-13), 39.87 (C-18), 36.75 & 36.74 (C-34), 33.92 (C-19), 30.41 (C-9), 30.29 (C-37), 29.79 & 29.77 (C-36), 29.51 & 29.48 (C-12), 28.03 & 27.99 (C-10), 27.76 & 27.73 (C-11), 26.91 & 26.89 (C-35). HRMS calculated for $[C_{44}H_{59}N_9O_7S + H]^+$ 858.4331; found 858.4336.

Cell culture

HEK293T cells (ATCC, CRL-3216) were cultured in DMEM high glucose (Sigma-Aldrich) supplemented with 10% FCS, 0.1 (w/v) penicillin/streptomycin, and 1% (v/v) Glutamax at 37 °C at 7% CO_2 and sub-cultured at 1 : 10 ratio twice a week. Normal human dermal fibroblasts (NHDF, Cambrex-Lona, CC-2511) were cultured in HAMF12-DMEM medium (Sigma-Aldrich) supplemented with 10% FCS and 0.1 (w/v) penicillin/streptomycin at 5% CO_2 and subcultured at 1 : 3 ratio once per 1–2 weeks. Culture medium was refreshed every 2–3 days for NHDF.

Enzyme activity assays

Recombinant murine GALC was cloned and expressed in HEK293 cells. The produced protein which was secreted to the culture medium (DMEM high glucose, Gibco) was directly used (5 µL volume) in enzyme activity studies. Human fibroblast lysates were prepared in KPi buffer (25 mM K_2HPO_4/KH_2PO_4, pH 6.5, 0.1% (v/v) Triton X-100, protease inhibitor cocktail (Roche, version 12, EDTA-free)), and 5.5 µg of total protein was used. β-Galactosidase from *Cellvibrio japonicus* (*Cj*GH35A) was prepared in KPi buffer, and 8.9 ng of protein was used. Enzyme activity assays were performed in 96-well plates (Greiner, black, flat-bottom, medium-binding). All samples were diluted in McIlvaine buffer (150 mM citrate/Na_2HPO_4, pH 4.5) in a total volume of 25 µL, before being incubated with 100 µL of substrate mixture (1 mM 4-methylumbelliferyl β-D-galactopyranoside (4-MU-β-D-Gal) in 150 mM McIlvaine buffer (pH 4.5) with 0.2 M NaCl) for 30 min at 37 °C. Reactions were stopped and analysed as described earlier.[45] To discriminate GLB1 and GALC activity, samples were incubated with McIlvaine buffer with or without 11 µM $AgNO_3$, before being incubated with substrates. To determine the apparent IC_{50} values, enzymes were equilibrated in 12.5 µL of McIlvaine buffer (pH 4.5) and incubated with a range of inhibitor dilutions (12.5 µL) for 30 min at 37 °C, before subsequent incubation with substrates. All assays were performed in triplicate wells, and IC_{50} assays were additionally performed in experimental duplicate per compound.

Measurement of enzyme kinetic parameters

The kinetic parameters of compounds toward *Cj*GH35A were measured with the method described below and the data are shown in the ESI.† An enzyme stock was prepared by diluting *Cj*GH35A to 0.71 ng μL^{-1} in McIlvaine buffer (pH 4.5, 0.1% (w/v) BSA) and pre-warmed at 37 °C for 10 min in a thermoshaker. A series of 2 mL Eppendorf tubes containing compounds **1**, **3**, and **6** (162.5 μL, various compound concentrations) were added with 4-MU-β-D-Gal (1300 μL, reaction [4-MU-β-gal] = 2345 μM) and identically pre-warmed. The $t = 0$ samples were prepared by transferring 112.5 μL from the [compound + 4-MU-β-D-Gal] samples to the first two columns of a 96-well plate (in duplicate for each compound concentration), after which stop buffer (200 μL, glycine-NaOH 1M, pH 10.3) and lastly the enzyme (12.5 μL) were added. Then, the enzyme (137.5 μL) was added ($t = 0$) to each of the 2 mL Eppendorf tubes (containing mixtures of compounds and substrates). At t = 2, 4, 6, 8 and 10 min, an aliquot of 125 μL from each of the reaction mixture was taken in duplicate and added to the plate containing stop buffer. After 10 minutes, the plate was measured with an LS-55 Fluorescence Spectrometer (PerkinElmer) using BL Studio with excitation at 366 nm and emission at 445 nm. Kinetic experiments to determine IC_{50} values were performed in duplicate sets, each set with technical duplicate. The reaction was stopped by adding stop buffer and samples were measured for 4-MU fluorescence. Results were processed and analysed using GraphPad Prism 7.0.

Gel-based ABPP

25 μg of protein from extracts of mouse kidneys was diluted in 10 μL of McIlvaine buffer (150 mM, pH 4.5) and pre-incubated with the beta-glucosidase suicide inhibitor JJB70 (see the ESI† for its structure). Samples were subsequently incubated with ABP **6** (1 μM) for 30 min at 37 °C. Samples were denatured, resolved by SDS-PAGE, and scanned for fluorescence.

Chemical proteomics

For pull-down with ABP **5**, 1 mg of total protein from mouse kidney extracts were diluted with McIlvaine buffer (750 mM, pH 4.5) to a total volume of 240 μL for each sample and incubated with 20 μM ABP **5** at 37 °C for 1.5 h. For control, DMSO was used in place of ABP. For competition with compound **3**, the mouse kidney extracts were incubated with compound **3** for 1 h at 37 °C, followed by incubation with 20 μM ABP **5** at 37 °C for 1.5 h. After ABP incubation, samples were denatured with 10% (w/v) SDS and subjected to chloroform/methanol precipitation (C/M), reduction/alkylation, C/M again, and pull-down with 75 μL of agarose avidin beads in a volume of 3400 μL of pull-down buffer (50 mM Tris–HCl, pH 7.4, with 150 mM NaCl) at 4 °C overnight, following previously described procedures.[46] Afterwards, the samples were subjected to on-bead digestion and desalted using stage-tips.[47] Desalted peptide samples were reconstituted in 30 μL of LC-MS solution (97 : 3 : 0.1 H_2O, ACN, FA) containing 10 fmol μL^{-1} yeast enolase digest (cat. 186002325, Waters) as injection control.

The desalted peptides were separated on an UltiMate 3000 RSLCnano system set in a trap–elute configuration with a nanoEase M/Z Symmetry C18 100 Å, 5 μm, 180 μm × 20 mm (Waters) trap column for peptide loading/retention and a nanoEase M/Z HSS C18 T3 100 Å, 1.8 μm, 75 μm × 250 mm (Waters) analytical column for peptide separation. Samples were injected on the trap column at a flow rate of 15 μL min^{-1} for 2 min with 99% mobile phase A (0.1% FA in ULC-MS grade water (Biosolve)) and 1% mobile phase B (0.1% FA in ULC-MS grade acetonitrile (Biosolve)) as an eluent. The 85 min LC method, using mobile phase A and mobile phase B controlled with a flow sensor at 0.3 μL min^{-1} with an average pressure of 400–500 bar (5500–7000 psi), was programmed as gradient with linear increment to 1% B from 0 to 2 min, 5% B at 5 min, 22% B at 55 min, 40% B at 64 min, 90% B at 65 to 74 min and 1% B at 75 to 85 min. The eluent was introduced by electrospray ionization (ESI) *via* a nanoESI source (Thermo) using stainless steel Nano-bore emitters (40 mm, OD 1/32″, ES542, Thermo Scientific).

The QExactive HF was operated in positive mode with data dependent acquisition without the use of lock mass, with default charge of 2+ and with external calibration with LTQ Velos ESI positive ion calibration solution (88323, Pierce, Thermo Scientific) every 5 days to less than 2 ppm. The tune file for the survey scan was set to the scan range of 350–1400 m/z, 120 000 resolution (m/z 200), 1 microscan, automatic gain control (AGC) of 3e6, max injection time of 100 ms, no sheath, auxiliary or sweep gas, spray voltage ranging from 1.7 to 3.0 kV, a capillary temperature of 250 °C and an S-lens value of 80. For the 10 data dependent MS/MS events the loop count was set to 10 and the general settings were resolution to 15 000, AGC target 1e5, max IT time 50 ms, isolation window of 1.6 m/z, fixed first mass of 120 m/z and normalized collision energy (NCE) of 28 eV. For individual peaks the data dependent settings were 1.00e3 for the minimum AGC target yielding an intensity threshold of 2.0e4 that needs to be reached prior to triggering an MS/MS event. No apex trigger was used, unassigned, +1 and >+8 charges were excluded with peptide match mode preferred, isotope exclusion on, and dynamic exclusion of 10 s. In between experiments, routine wash and control runs were done by injecting 5 μl of LC-MS solution containing 5 μL of 10 fmol μL^{-1} BSA or enolase digest and 1 μL of 10 fmol μL^{-1} angiotensin III (Fluka, Thermo)/oxytocin (Merck) to check the performance of the platform on each component (nano-LC, the mass spectrometer (mass calibration/quality of ion selection and fragmentation) and the search engine).

Raw files were analyzed with MaxQuant (Version 2.0.1.0). The following changes were made to the standard settings of MaxQuant: label-free quantification was enabled with an LFQ minimal ratio count of 1. Matches between runs and iBAQ quantification were enabled. Searches were performed against the Uniprot database of the *Mus musculus* (mouse) proteome (UPID: UP000000589) including yeast enolase (P00924).

Author contributions

C.-L.K., Q.S. and A.R.A.M. performed the biochemistry and analytical chemistry experiments under the guidance of B.I.F. and R.G.B. A.M.C.H.N., T.J.M.B., A.J.S. and L.I.W. synthesised the inhibitors and probes under the guidance of G.A.v.d.M. and J.D.C.C. W.A.O. performed the structural studies under the guidance of G.J.D. H.S.O. and J.M.F.G.A. conceived and supervised the research and C.-L.K. wrote the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank The Netherlands Organization for Scientific Research (NWO-CW, ChemThem grant to J. M. F. G. A. and H. S. O.) and the European Research Council (ERC-2011-AdG290836 "Chembiosphing", to H. S. O., and ERC-2020-SyG-951231 Carbocentre, to H. S. O. and G. J. D.). G. J. D. is funded by the Royal Society Ken Murray Research Professorship.

References

1 T. Hennet, *Cell. Mol. Life Sci.*, 2000, **59**, 1081.
2 Y. Suzuki and E. Nanba, in *The metabolic and molecular basis of inherited disease*, McGraw-Hill, New York, USA, 2000.
3 K. Suzuki and Y. Suzuki, *Proc. Natl. Acad. Sci. U. S. A.*, 1970, **66**, 302.
4 V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho and B. Henrissat, *Nucleic Acid Res.*, 2014, **42**, D490.
5 A. d'Azzo, A. Hoogeveen, A. J. Reuser, D. Robinson and H. Galjaard, *Proc. Natl. Acad. Sci. U. S. A.*, 1982, **79**, 4535.
6 A. van der Spoel, E. Bonten and A. d'Azzo, *J. Biol. Chem.*, 2000, **275**, 10035.
7 R. Kreutzer, M. Kreutzer, M. J. Pröpsting, A. C. Sewell, T. Leeb, H. Y. Naim and W. Baumgärtner, *J. Cell. Mol. Med.*, 2008, **12**, 1661.
8 U. Ohto, K. Usui, T. Ochi, K. Yuki, Y. Satow and T. Shimizu, *J. Biol. Chem.*, 2012, **287**, 1801.
9 N. J. Galjart, N. Gillemans, A. Harris, G. T. van der Horst, F. W. Verheijen, H. Galjaard and A. d'Azzo, *Cell*, 1988, **54**, 755.
10 D. A. Wenger, M. A. Rafi and P. Luzi, *Hum. Mutat.*, 1997, **10**, 268.
11 S. Nagano, T. Yamada, N. Shinnoh, H. Furuya, T. Taniwaki and J. Kira, *Clin. Chim. Acta*, 1998, **276**, 53.
12 J. E. Deane, S. C. Graham, N. N. Kim, P. E. Stein, R. McNair, M. B. Cachón-González, T. M. Cox and R. J. Read, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 15169.
13 C. H. Hill, S. C. Graham, R. J. Read and J. E. Deane, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 20479.
14 N. Brunetti-Pierri and F. Scaglia, *Mol. Genet. Metab.*, 2008, **94**, 391.
15 A. Tessitore, M. P. del Martin, R. Sano, Y. Ma, L. Mann, A. Ingrassia, E. D. Laywell, D. A. Steindler, L. M. Hendershot and A. d'Azzo, *Mol. Cell*, 2004, **15**, 753.
16 A. Takamura, K. Higaki, K. Kajimaki, S. Otsuka, H. Ninomiya, J. Matsuda, K. Ohno, Y. Suzuki and E. Nanba, *Biochem. Biophys. Res. Commun.*, 2008, **367**, 616.
17 R. Sano, I. Annunziata, A. Patterson, S. Moschiach, E. Gomero, J. Opferman, M. Forte and A. d'Azzo, *Mol. Cell*, 2009, **36**, 500.
18 R. Spiegel, G. Bach, V. Sury, G. Mengistu, B. Meidan, S. Shalev, Y. Shneor, H. Mandel and M. Zeigler, *Mol. Genet. Metab.*, 2005, **84**, 160.
19 J. S. Won, A. K. Singh and I. Singh, *J. Neurosci. Res.*, 2016, **94**, 990.
20 J. S. Won, J. Kim, M. K. Paintlia, I. Singh and A. K. Singh, *Brain Res.*, 2013, **1508**, 44.
21 A. C. Graziano, R. Parenti, R. Avola and V. Cardile, *Apoptosis*, 2016, **21**, 25.
22 F. Deodato, E. Procopio, A. Rampazzo, R. Taurisano, M. A. Donati, C. Dionisi-Vici, A. Caciotti, A. Morrone and M. Scarpa, *Metab. Brain Dis.*, 2017, **32**, 1529.
23 E. Zaccariotto, M. B. Cachón-González, B. Wang, S. Lim, B. Hirth, H. Park, M. Fezoui, S. P. Sardi, P. Mason, R. H. Barker Jr. and T. M. Cox, *Biomed. Pharmacother.*, 2022, **149**, 112808.
24 J. Condori, W. Acosta, J. Ayala, V. Katta, A. Flory, R. Martin, J. Radin, C. L. Cramer and D. N. Radin, *Mol. Genet. Metab.*, 2016, **117**, 199.
25 M. Gupta, H. Pandey and S. Sivakumar, *ACS Omega*, 2017, **2**, 9002.
26 M. Schalli, P. Weber, C. Tysoe, B. M. Pabst, M. Thonhofer, E. Paschke, A. E. Stütz, M. Teschernutter, W. Windischhofer and S. G. Withers, *Bioorg. Med. Chem. Lett.*, 2017, **27**, 3431.
27 M. A. Hossain, K. Higaki, S. Saito, K. Ohno, H. Sakuraba, E. Nanba, Y. Suzuki, K. Ozono and N. Sakai, *J. Hum. Genet.*, 2015, **60**, 539.
28 R. C. Baek, M. L. Broekman, S. G. Leroy, L. A. Tierney, M. A. Sandberg, A. d'Azzo, T. N. Seyfried and M. Sena-Esteves, *PLoS One*, 2010, **5**, e13468.
29 M. J. Przybilla, L. Ou, A. F. Tăbăran, X. Jiang, R. Sidhu, P. J. Kell, D. S. Ory, M. G. O'Sullivan and C. B. Whitley, *Mol. Genet. Metab.*, 2019, **126**, 139.
30 M. A. Rafi, H. Z. Rao, P. Luzi, A. Luddi, M. T. Curtis and D. A. Wenger, *Mol. Genet. Metab.*, 2015, **114**, 459.
31 C. J. Folts, N. Scott-Hewitt, C. Pröschel, M. Mayer-Pröschel and M. Noble, *PLoS Biol.*, 2016, **14**, e1002583.
32 Y. Li, Y. Xu, B. A. Benitez, M. S. Nagree, J. T. Dearborn, X. Jiang, M. A. Guzman, J. C. Woloszynek, A. Giaramita, B. K. Yip, J. Elsbernd, M. C. Babcock, M. Lo, S. C. Fowler, D. F. Wozniak, C. A. Vogler, J. A. Medin, B. E. Crawford and M. S. Sands, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 20097.

33 D. Koshland, *Biol. Rev.*, 1953, **28**, 416.
34 L. Wu, Z. Armstrong, S. P. Schröder, C. de Boer, M. Artola, J. M. F. G. Aerts, H. S. Overkleeft and G. J. Davies, *Curr. Opin. Chem. Biol.*, 2019, **53**, 25.
35 A. R. Marquez, L. I. Willems, D. Herrera Moro, B. I. Florea, S. Scheij, R. Ottenhoff, C. P. van Roomen, M. Verhoek, J. K. Nelson, W. W. Kallemeijn, A. Bielsa-Banas, O. R. Martin, M. B. Cachón-González, N. N. Kim, T. M. Cox, H. S. Overkleeft and J. M. Aerts, *ChemBioChem*, 2017, **18**, 402.
36 S. P. Schröder, J. W. van de Sande, W. W. Kallemeijn, C. L. Kuo, M. Artola, E. J. van Rooden, J. Jiang, T. J. M. Beenakker, B. I. Florea, W. A. Offen, G. J. Davies, A. J. Minnaard, J. M. F. G. Aerts, J. D. C. Codée, G. A. van der Marel and H. S. Overkleeft, *Chem. Commun.*, 2017, **53**, 12528.
37 L. I. Willems, T. J. M. Beenakker, B. Murray, B. Gagestein, H. van den Elst, E. R. van Rijssel, J. D. C. Codée, W. W. Kallemeijn, J. M. F. G. Aerts, G. A. van der Marel and H. S. Overkleeft, *Eur. J. Org. Chem.*, 2014, 6044.
38 Y. Harrak, C. M. Barra, A. Delgado, A. R. Castaño and A. Llebaria, *J. Am. Chem. Soc.*, 2014, **133**, 12079.
39 N. K. Devaraj, R. Weissleder and S. A. Hilderbrand, *Bioconjugate Chem.*, 2008, **12**, 2297.
40 R. J. Rowland, Y. Chen, I. Breen, L. Wu, W. A. Offen, T. J. Beenakker, Q. Su, A. M. C. H. van den Nieuwendijk, J. M. F. G. Aerts, M. Artola, H. S. Overkleeft and G. J. Davies, *Chem. – Eur. J.*, 2021, **27**, 16377.
41 S. Martino, R. Tiribuzi, A. Tortori, D. Conti, I. Visigalli, A. Lattanzi, A. Biffi, A. Gritti and A. Orlacchio, *Clin. Chem.*, 2009, **55**, 541.
42 W. W. Kallemeijn, K.-Y. Li, M. D. Witte, A. R. A. Marques, J. Aten, S. Scheij, J.-B. Jiang, L. I. Willems, T. M. Voorn-Brouwer, C. P. A. A. van Roomen, R. Ottenhoff, R. G. Boot, H. van den Elst, M. T. C. Walvoort, B. I. Florea, J. D. C. Codée, G. A. van der Marel, J. M. F. G. Aerts and H. S. Overkleeft, *Angew. Chem., Int. Ed.*, 2012, **51**, 12529.
43 F. Percheron, M. J. Folglietti, M. Bernard and B. Ricard, *Biochemie*, 1992, **74**, 5.
44 N. G. S. McGregor, C.-L. Kuo, T. J. M. Beenakker, C.-S. Wong, Z. Armstrong, B. I. Florea, J. D. C. Codée, H. S. Overkleeft, J. M. F. G. Aerts and G. J. Davies, *Org. Biomol. Chem.*, 2022, **20**, 877.
45 L. I. Willems, N. Li, B. I. Florea, M. Ruben, G. A. van der Marel and H. S. Overkleeft, *Angew. Chem., Int. Ed.*, 2012, **51**, 4431.
46 C. L. Kuo, W. W. Kallemeijn, L. T. Lelieveld, M. Mirzaian, I. Zoutendijk, A. Vardi, A. H. Futerman, A. H. Meijer, H. P. Spaink, H. S. Overkleeft, J. M. F. G. Aerts and M. Artola, *FEBS J.*, 2019, **286**, 584.
47 N. Li, C.-L. Kuo, G. Paniagua, H. van den Elst, M. Verdoes, L. I. Willems, W. A. van der Linden, M. Ruben, E. van Genderen, J. Gubbens, G. P. van Wezel, H. S. Overkleeft and B. I. Florea, *Nat. Protoc.*, 2013, **8**, 1155.