Activity-based protein profiling of glucosidases, fucosidases and glucuronidases

Jiang, J.

7

Exo- and *endo*-retaining β-glucuronidase activities and mechanisms revealed by cyclophellitol aziridine-based inhibitors and probes

Jianbing Jiang, Liang Wu, Chi-Lin Kuo, Wouter W. Kallemeijn, Gijsbert A. van der Marel, Jeroen D. C. Codée, Marco C. van Eijk, Bogdan I. Florea, Johannes M. F. G. Aerts, Herman S. Overkleeft and Gideon J. Davies, **2016**, manuscript in preparation.

7.1 Introduction

GH2 and GH79 β-glucuronidases catalyze the hydrolysis of β-glucuronic acid linkages from a widespread number of substrates and are expressed in numerous different species including bacteria and eukaryotes.[1,2] GH2 and GH79 β-glucuronidases are retaining enzymes that process their substrate to yield β-glucuronic acid residues. β-Glucuronidase mediated hydrolysis proceeds through a Koshland two-step double displacement mechanism (Figure 1A).[3] In the first step the aglycon is protonated through the action of a general acid-base residue positioned at the β-face of the substrate. Concomitantly, in a formal S_N2 nucleophilic displacement the protonated aglycon is expelled via nucleophilic attack by a nucleophilic residue residing at the α-face of the substrate, forming an intermediate enzyme-α-glucuronide adduct. Next, water enters the enzyme active site and in a reversal of steps the glucuronide is released from the enzyme with overall retention of anomeric configuration. This mechanism of action is adopted by both *exo*- and *endo*-β-glucuronidases, and the differences between these enzymes are in their substrate recognition and physiological roles.

Both GH2 and GH79 belong to the GH-A clan.[4] The GH2 family consists of only *exo*-acting enzymes, such as GH2 human β-glucuronidase (GUSB), which is responsible for cleaving β-linked D-glucuronides from the non-reducing end of glycosaminoglycans (GAGs). Deficiency of GUSB causes the autosomal recessive disease: mucopolysaccharidosis type VII (MPSVII), also known as Sly syndrome.[5] The GH79 family consists of both *exo*-acting and *endo*-acting enzymes, such as bacterial β-Glucuronidase from *Acidobacterium capsulatum* (AcaGH79) and human heparanase (HPSE) respectively.[6,7] HPSE is an important factor in the processing and degradation of heparan sulfate, and has been implicated in a variety of processes underlying human pathologies such as inflammation, tumor metastasis and angiogenesis.[8]

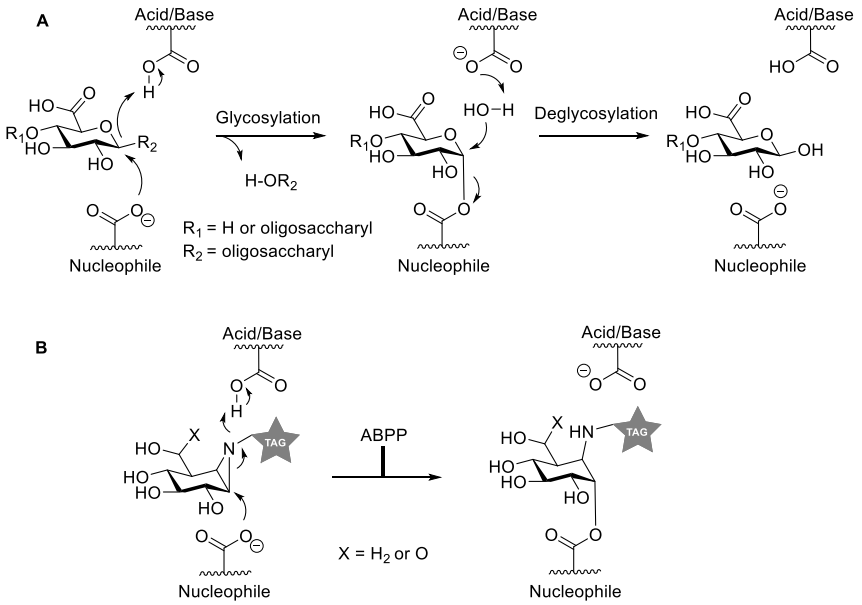


Figure 1. A) Koshland two-step double displacement mechanism employed by GH2 and GH79 retaining β-glucuronidases. B) Cyclophellitol aziridine-derived mechanism-based inhibitors and activity-based probes for retaining β-glucosidases (**1**, X = H_2) and retaining β-glucuronidases (**2**, **3** X = O).

Retaining glycosidases that employ a Koshland double displacement mechanism and that form during substrate processing a covalent enzyme-glycoside adduct are amenable to activity-based protein profiling (ABPP).[9] It has been shown in the past that fluorescent or biotin-conjugated derivatives of the broad-spectrum retaining β-glucosidase inhibitor, *N*-alkyl cyclophellitol aziridine (Figure 1B, **1**, X = H_2) are suitable activity-based probes (ABPs) for profiling retaining β-glucosidases *in vitro*, *in situ* and *in vivo*.[10,11] The efficiency of the ABPs is based on their tight initial binding to the enzyme active site, subsequent protonation of the aziridine nitrogen, and finally S_N2 substitution of the aziridinium ion by the active site nucleophile. The resulting enzyme-cyclitol adduct is comparatively more stable than the acylal

linkage that results from the natural process (as in Figure 1A) and the retaining β-glucosidase is thus inhibited in a mechanism-based manner. The covalent enzyme-inhibitor adduct is stable after protein denaturation, allowing for biochemical (gel electrophoresis) and analytical (mass spectrometry-based proteomics) study of the captured enzyme(s).

Altering the configuration of the cyclophellitol aziridine scaffold yielded probes comparably effective and selective for retaining GH27 α-galactosidases[12] and GH29 α-L-fucosidases,[13] showing the general applicability of the activity-based glycosidase profiling methodology. Modification of the cyclophellitol aziridine core to emulate a glucuronic acid moiety (Figure 1B, **2**,**3**, X = O) yielding ABPs capable of modifying and visualizing β-glucuronidase AcaGH79 have been described in Chapter 6, as well as the design, synthesis and evaluation of a suite of retaining β-glucuronidase inhibitors and ABPs. Here, an in-depth study on the use of the activity-based β-glucuronidase probes in labeling and identification of both *exo*- and *endo*-glucuronidases in various research settings and from various tissues is described. As well, crystal structures of compound **4** in complex with wild type and (E to Q) nucleophile-mutant human HPSE (*endo*) GH79 retaining β-glucuronidases, comparing the complex structure of **4** with bacterial (*exo*) AcaGH79 (Chapter 6) are described in this chapter.

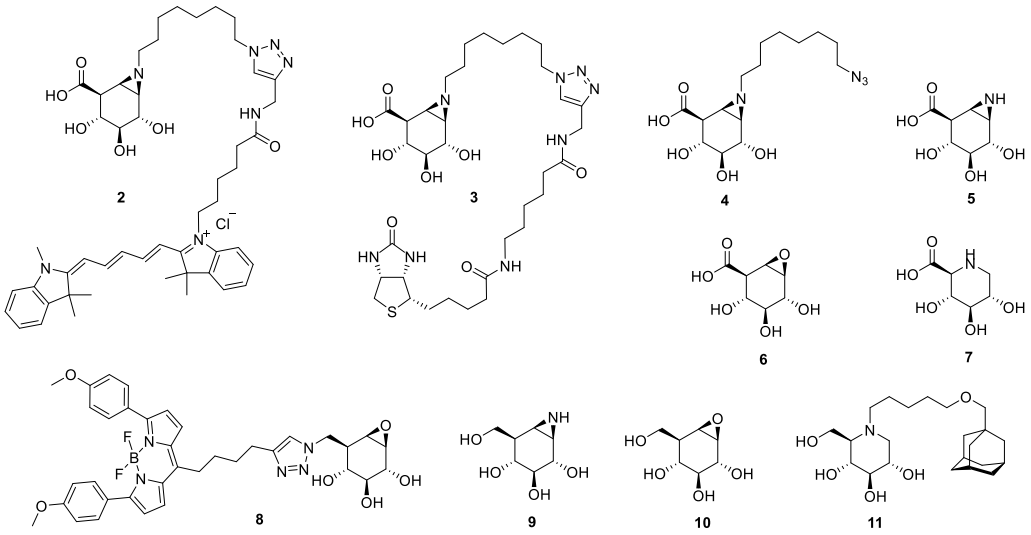


Figure 2. Inhibitors and ABPs used in the here-reported study. **2** and **3** are β-glucuronidase ABPs; **4**, **5**, **6** and **7** are β-glucuronidases inhibitors; **8** is β-glucosidase ABP; **9**, **10** and **11** are β-glucosidase inhibitors.

7.2 Results

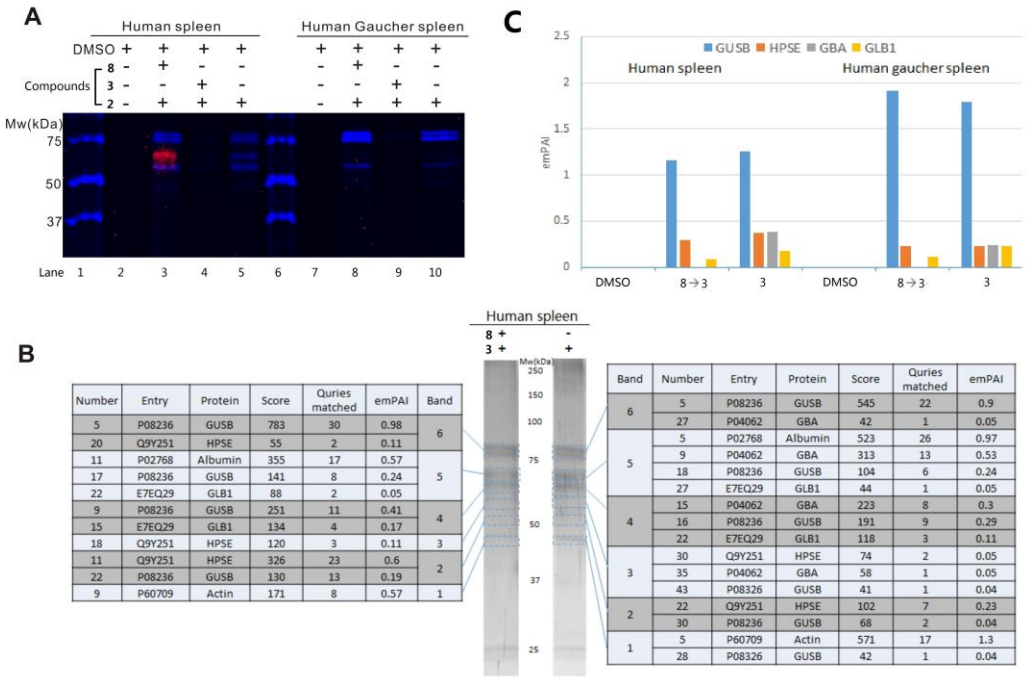
Inhibitor and probe design

The β-glucuronidase inhibitors and probes used in this study are depicted in Figure 2 (see Chapter 6 for their synthesis). ABPs **2** and **3** are composed of β-glucuronic acid-configured

cyclophellitol aziridine, bearing at the aziridine-nitrogen an alkyl spacer that terminates in Cy5 and biotin respectively. Fluorescent ABP **2** is used for in-gel β-glucuronidase detection (fluorescence scanning of gel slabs) and compound **3** for biotin-Streptavidin pull down enrichment-identification through chemical proteomics strategies with mass spectrometry. The set of ABPs are accompanied by azide-modified cyclophellitol aziridine **4** (the precursor in the preparation of ABPs **2**, **3**), unsubstituted cyclophellitol aziridine **5**, cyclophellitol-6-carboxylate **6** as well as trihydroxypipecolic acid **7** as a competitive β-glucuronidases inhibitor. In addition to the β-glucuronidase inhibitors and ABPs, the acid glucosylceramidase (GBA) ABP **8**, mechanism-based GBA inhibitors **9** (cyclophellitol aziridine), **10** (cyclophellitol) and competitive GBA inhibitor **11** are also employed in the studies as described below.

Chemical proteomics reveals GH2 GUSB and GH79 HPSE as targets of ABPs 2 and 3

In a first set of experiments to assess the activity and selectivity of our β-glucuronidase ABPs human spleen tissue extracts were treated with Cy5-modified cyclophellitol aziridine **2**. Human spleen is expected to express high GUSB levels and low HPSE levels, according to the Expression Atlas transcriptome database.[14] Following denaturing of the samples and resolving of their protein content on SDS-PAGE, labeled proteins were visualized by fluorescence scanning of the wet gel slab at 605 and 695 nm (overlay of Cy3 and Cy5 channel, Figure 3A left panel, for separate channel images see the supporting information Figure S1). Several distinct bands at molecular weights roughly between 50 and 80 kD became apparent, all of which were absent in the (DMSO) mock-treated sample (lane 2). The lower and second to highest band was putatively assigned as GUSB and the highest molecular weight band as its pro-form, based on literature information on GUSB expression forms in human spleen.[15-16] The band at 58-62 kD could not be assigned to a known β-glucuronidase, but it was previously reported that labeling human spleen with the close (with respect to the configuration of the cyclitol aziridine) structural analogue, ABP **8**, yielded exclusive labeling of glucocerebroside (GBA) at exactly this position.[17] Indeed, pre-incubation with **8** prior to treatment with **2** yielded a labeling pattern (lane 3) reminiscent of that observed in lane 5, but lacking the band at 58-62 kD. The identity of GUSB and pro-GUSB as labeling-target of ABP **2** was further confirmed by repeating the exact same series of experiments, but in extracts derived from human Gaucher spleen tissue containing less activity of GBA[18] (Figure 3A right panel, note the absence of the bands at 50-80 kD in both lanes 8 and 10). ABP **2** labeling in both tissue extracts could be completely blocked by pre-incubation with biotin-aziridine **3**, revealing that this pull-down chemical proteomics probe targets at least the same set of proteins as fluorescent probe **2**.



Number	Entry	Protein	Score	Quries matched	emPAI	Band
5	P08236	GUSB	783	30	0.98	6
20	Q9Y251	HPSE	55	2	0.11	
11	P02768	Albumin	355	17	0.57	5
17	P08236	GUSB	141	8	0.24	
22	E7EQ29	GLB1	88	2	0.05	
9	P08236	GUSB	251	11	0.41	4
15	E7EQ29	GLB1	134	4	0.17	
18	Q9Y251	HPSE	120	3	0.11	3
11	Q9Y251	HPSE	326	23	0.6	2
22	P08236	GUSB	130	13	0.19	
9	P60709	Actin	171	8	0.57	1

Band	Number	Entry	Protein	Score	Quries matched	emPAI
6	5	P08236	GUSB	545	22	0.9
	27	P04062	GBA	42	1	0.05
5	5	P02768	Albumin	523	26	0.97
	9	P04062	GBA	313	13	0.53
	18	P08236	GUSB	104	6	0.24
	27	E7EQ29	GLB1	44	1	0.05
4	15	P04062	GBA	223	8	0.3
	16	P08236	GUSB	191	9	0.29
	22	E7EQ29	GLB1	118	3	0.11
3	30	Q9Y251	HPSE	74	2	0.05
	35	P04062	GBA	58	1	0.05
	43	P08326	GUSB	41	1	0.04
2	22	Q9Y251	HPSE	102	7	0.23
	30	P08236	GUSB	68	2	0.04
1	5	P60709	Actin	571	17	1.3
	28	P08326	GUSB	42	1	0.04

Figure 3. A) Labeling of human spleen and human Gaucher spleen with ABP **2** and **8** (lanes 1 and 6 are protein marker). B) Silver staining of in-gel digestion and identification of target proteins modified by biotin-ABP **3** with Mascot data parameters table. C) Mascot parameter emPAI value of glycosidases identified after on-bead pull-down and processing of spleen lysate with biotin-ABP **3**.

With the aim to unambiguously establish the targets of the β-glucuronidase ABPs, a number of chemical proteomics experiments were performed. As described above, biotin-aziridine ABP **3** is able to compete with fluorescent ABP **2** (Figure 3A, lane 4 and lane 9) and thus targets at least the set of proteins as seen in the gels in Figure 3A, lane 5 and lane 10. In a first chemical proteomics experiment, human spleen and Gaucher spleen extracts were treated with biotin-aziridine **3**, DMSO (mock treatment), or treatment with **3** after pre-incubating with ABP **8**. Biotinylated proteins were captured on magnetic streptavidin-loaded beads and non-biotinylated proteins were washed off. The samples were separated in two parts: one-third for in-gel digestion and two-thirds for on-bead digestion. The captured proteins for in-gel digestion were liberated from the beads by treatment with Laemmli buffer containing excess biotin. Thus, protein pools were obtained that besides endogenous biotinylated proteins should also contain those proteins able to react covalently and irreversibly with biotin-aziridine **3**. The eluted proteins were resolved on SDS-PAGE, visualized by silver staining (Figure 3B and SI, Figure S1B), stained bands excised from the gel and treated with trypsin. Tryptic peptides were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) and the resolved peptide sequences matched against the Mascot database.[19] As can

be seen (Figure 3B), GUSB was identified in this manner, alongside GBA as the main off-target protein alongside retaining β-galactosidase (GLB1), keratin, and a number of endogenously biotinylated CoA carboxylases (SI, Figure S1C). Interestingly, HPSE in band 2 and 3 of both lanes were also identified in both wild type and Gaucher tissues (SI, Figure S1D), despite low HPSE expression level according to the Expression Atlas database. Essentially, the same results were obtained when performing the on-bead trypsin digestion instead of in-gel (Figure 3C, the Exponentially Modified Protein Abundance Index (emPAI)[20] value offers approximate, label-free, relative quantitation of the proteins in a mixture based on protein coverage by the peptide matches in a database search result). In both experimental set-ups, GBA-specific ABP **8** could compete for GBA labeling/pull down. Identification of the labeled proteins in the chemical proteomics experiments shown in Figure 3B, 3C is based on sequence identification using the Mascot search engine. The software does not recognize the active site fragments of the ABP-modified retaining glycosidases because of their altered molecular structure. Therefore, a manual search was performed for the molecular weight of the putative tryptic fragment containing the active site nucleophile modified by biotin aziridine **3**, both for GUSB and HPSE, in the LC-MS spectra. Neither of the expected masses (*m/z*=4494.19 for GUSB Tyr532-Lys563 and *m/z*=3882.90 for HPSE Val339-Lys368) could be identified and thus an alternative digestion protocol was used for protein digestion. Following this protocol, the trypsinolysis-derived peptide pool was further treated with endoproteinase Glu-C,[21] a protease that cleaves specifically after acidic residues. Although not successful for HPSE, due to the small amount of total protein in the samples, this protocol proved successful in identifying the oligopeptide containing the active site nucleophile of human GUSB in Gaucher spleen extract. As can be seen (SI, Figure S1E, F), a peak corresponding to the predicted mass of this modified peptide from Gaucher spleen extract was observed in the LC-MS trace, and LC-MS/MS fragmentation of this peptide delivered the expected sequence (*m/z*=2261.14) Tyr532-Glu544 with nucleophile active-site Glu540 modified by ABP **3**.

In-depth comparative and competitive ABPP on GH2 human GUSB

As evidenced from the proteomics data presented above, the main off-target identified for ABPs **2** and **3** proved to be the retaining β-glucosidase, GBA. GBA was also identified after labeling several mouse tissue extracts with ABP **2** (Figure 4B), and the labeling intensity with ABP **2** in mouse liver extracts appeared optimal at pH 5-6, corresponding with the observation made using human spleen extract (Figure 4C). Previously, it was reported that 6-modified (glucopyranose numbering) cyclophellitol derivative **8**,[17] with a BODIPY-TMR tag positioned at the 6 position is a highly selective and sensitive probe for GBA, which when deficient is at the basis of the lysosomal storage disorder Gaucher disease. The activity of GBA and GUSB as well as the impact of putative inhibitors on these enzymes can be dissected, as shown in Figure 4, by making use of Cy5-aziridine **2** and BODIPY-TMR epoxide **8**. Human spleen extracts were treated with Cy5-aziridine **2**, BODIPY-TMR-epoxide **8**, or a combination thereof, and a series of

competitive and covalent inhibitors **5**-**7** (GUSB inhibitors) and **9**-**11** (GBA inhibitors) was included in a competitive ABPP format. Lanes 1-3 recapitulate the results shown in Figure 4A, with red fluorescence (Cy3) corresponding to GBA labeling and blue fluorescence to GUSB (including its pro-form) labeling. The covalent (**5**, **6**) and competitive (**7**) GUSB inhibitors selectively outcompete ABP **2** (lanes 3-5), whereas the corresponding covalent (**9**, **10**) and competitive (**11**) inhibitors abolish to a large extend labeling with ABP **8**.

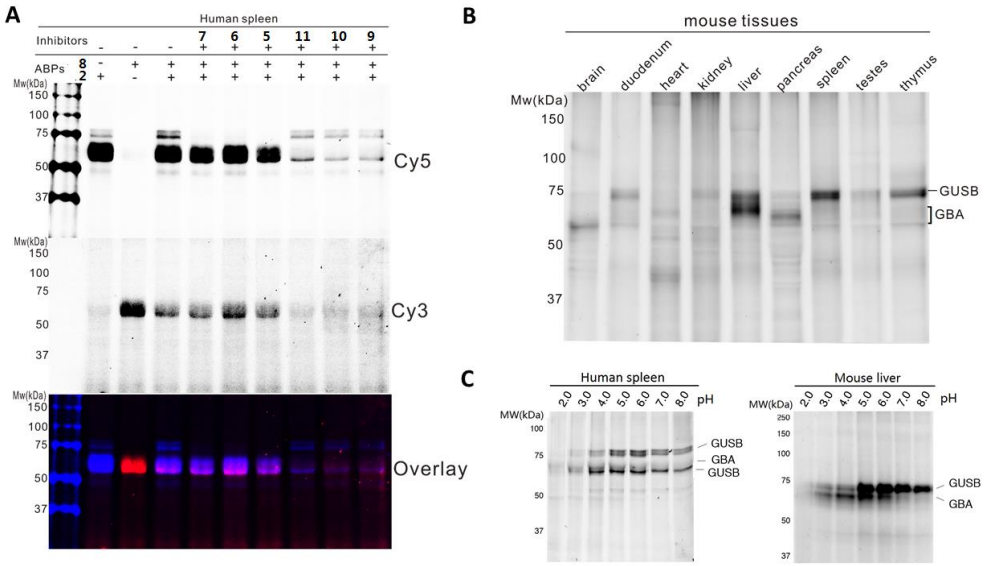


Figure 4. A) *In vitro* competitive ABPP on human spleen lysate after probing for both GBA (with ABP **8**) and GUSB (with ABP **2**) in Cy3, Cy5 and overlay channels. B) *In vitro* labeling of several mouse tissue extracts with ABP **2** for determination of GUSB expression and activity in different tissues. C) Detection of GUSB from human spleen and mouse liver extracts labeling efficiency at various pH with ABP **2**.

In-depth comparative and competitive ABPP on GH79 human HPSE

With the aim to explore Cy5 aziridine ABP **2** to study HSPE – the second β-glucuronidase activity identified in our chemical proteomics experiments (see Figure 3B), we turned our attention to tissues expressing this enzyme in relative high abundance. As HPSE could not be identified from fluorescence scanning on SDS-PAGE gel slabs derived from human spleen extract labeling (see Figure 3A and SI, Figure S2A), we brought HPSE to overexpression in the first instance in HEK293T cells. Labeling extracts of this tissue with Cy5 aziridine **2** showed time-dependent labeling of HPSE both as the mature enzyme and as the pro-enzyme, which is originally synthesized in an inactive 65 kDa form.[22] The identity of these bands was confirmed with immunostaining and Western blotting with an anti-HPSE 1 antibody (Figure 5A). In an independent experiment, primary human fibroblasts were treated with recombinant pro-HSPE, and its internalization was subsequently demonstrated after labeling cell extracts with ABP **2**

(SI, Figure S2B). Labeling of pro-HSPE was somewhat surprising, as literature data indicates that this immature enzyme is not active towards its natural substrate, heparan sulfate.[23] However, comparison of fluorescence images with those obtained from anti-HPSE western blotting shows significantly more pro-HPSE than mature HPSE in the cell lysates, indicating that labeling of the mature protein is more effective than that of the proenzyme.

Human platelets are known to express high HSPE levels and were therefore used to study labeling of this *endo*-glycosidases in more depth.[24] In platelet extracts prepared in the presence of EDTA elevated HPSE activity was observed, as indicated by labeling with ABP **2** and immunostaining (SI, Figure S2C). As can be seen (Figure 5B and SI, Figure S2E), ABP **2** clearly and in a pH-dependent manner labels HSPE. Labeling optimum is at pH 4-5, thus the pH corresponding with lysosomal pH – the natural environment of the enzyme. In contrast, labeling of GUSB in its various forms occurs at comparatively higher pH 5.5-6.5 (SI, Figure S2F). Therefore, GUSB and HPSE labeling can be done selectively in neutral and acidic conditions respectively. In these experiments using endogenous platelet samples no pro-HSPE labeling was observed. HSPE labeling in platelet extracts is optimal at 100 nanomolar probe and 60 minutes incubation (SI, Figure S2D) and can be partially competed by pre-incubation with heparin (Figure 5C), which had no effect on labeling of GUSB. It was observed that compounds **5**, **6**, **7** and known inhibitor Siastatin B[25] in millimolar concentration could completely block both GUSB and HPSE activity in platelet, following by ABP **2** incubation and readout (SI, Figure S2G,H). These results suggest that the GUSB catalytic pocket can only accommodate small monosugar-like inhibitors, whereas the HPSE binding site can accommodate both polysaccharide heparin or monosugar inhibitors.

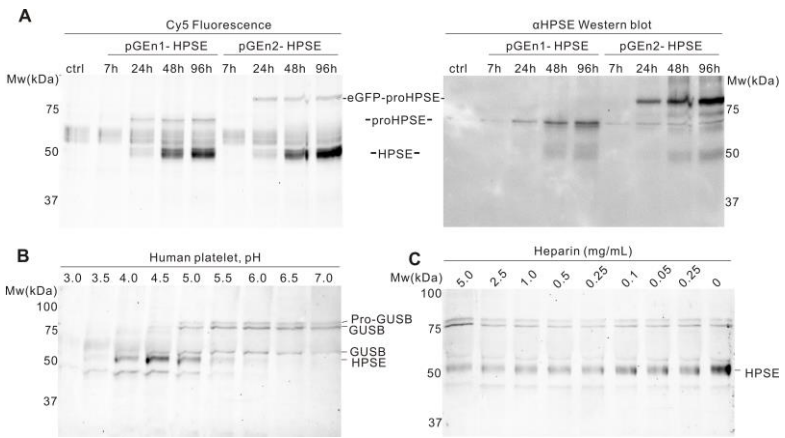


Figure 5. A) Overexpressed HPSE in HEK293 cells visualized with ABP **2** at different time points, by scanning the slab gel in Cy5 fluorescent channel (left) and Western blot analysis with HPSE specific antibody (right). B) Detection of endogenous HSPE in human platelets with ABP **2** at various pH values. C) HSPE labeling in platelets can be partially abolished by incubation with its natural substrate, heparin.

Structural analysis reveals mechanistic similarities between two GH79 β-glucuronidases: AcaGH79 (*exo*) and HSPE (*endo*)

To further prove the covalent labeling of both *exo*- and *endo*- GH79 β-glucuronidases by alkyl aziridine ABPs and analyze the catalytic mechanism employed by the enzymes, the crystal structures of AcaGH79 (see Chapter 6) and HPSE with mechanism-based inhibitor **4** were obtained. The resulting crystal structures clearly showed the formation of enzyme-**4** adducts, with a covalent O-C bond between the enzyme active-site nucleophile (E287 for AcaGH79, E343 for HPSE) and the carbon equivalent to the anomeric carbon of a bona fide substrate in **4** (Figure 6A). The alkylated linker on aziridine nitrogen shows considerable flexibility and accommodates the space normally occupied by enzyme substrates. During catalytic hydrolysis of both enzymes, aziridine rings on **4** open in trans-diaxial fashion and the substituted cyclohexane adopts $^{4}C_{1}$ conformation in both *exo*-AcaGH79 and *endo* HPSE catalytic pocket. The corresponding mutant enzymes (E287Q for AcaGH79, E343Q for HPSE) crystal structures with **4** were also obtained, revealing the umambiguous electron density for unhydrolyzed **4** in $^{4}H_{3}$ conformation (Figure 6B and Chapter 6 Figure 4). These results are consistent with the expected $^{1}S_{3}$-$^{4}H_{3}$-$^{4}C_{1}$ catalytic itinerary.[26,27] It is also observed that overlay of AcaGH79-**4** and HPSE-**4** complex exhibited the similar position and configuration of **4** or residues in catalytic pockets (Figure 6C). Moreover, although the *exo*-GH2 GUSB-**4** complex crystal structure (which would provide structural insights into covalent labeling of GH2 family β-glucuronidase) is not obtained yet, the overlay of active site residues of a GH2 *E. coli* β-glucuronidase[28] with two GH79 β-glucuronidases revealed the structural similarity between their active sites. Therefore, **4** could also fit in the catalytic cleft of GH2 GUSB well with similar hydrogen bond between amino acid residues and hydroxyl groups of cyclohexane for complex stability as GH79 β-glucuronidases-**4** complexes.

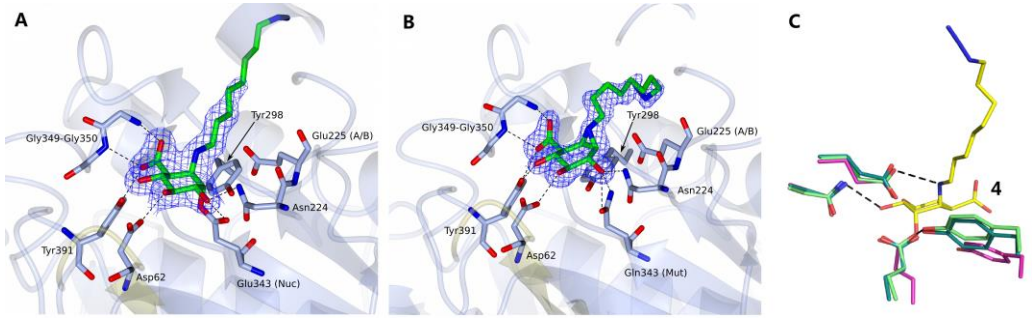


Figure 6. Crystal structures of A) HPSE-**4** complex and B) mutant (E343Q) HPSE-**4** complex. C) The overlay of HPSE (dark green)-**4** (yellow), AcaGH79 (light green)-**4** complexes and *E.coli* GH2 β-glucuronidase active site residues (purple).

7.3 Discussion

Activity-based probes (ABPs) have been proven to be a powerful tool to investigate retaining GHs activity in biological complex systems. GH2 and GH79 β-glucuronidases can be labeled

with high efficiency using β-glucuronic acid-configured cyclophellitol aziridine ABPs **2** and **3**, which is consistent with the efficacy in labeling of GH30 β-glucosidases, GH27 α-galactosidases and GH29 α-L-fucosidases with *N*-alkyl aziridine ABPs corresponding to the configuration and substitution pattern of their substrate glycosides.[11] Chemical proteomics applied by biotin ABP **3** in affinity enrichment of human spleen extracts for mass spectrometry analysis identified not only the main target enzyme GUSB, but also GBA as the main off target. Fluorescent ABP **2** labeled GBA in the same position as the specific GBA ABP **8** when analyzed by SDS-PAGE. GBA recognition of substrate with sugar C5 position modification has much more tolerance, comparing and former GBA inhibitors or probes.[10,18] Similarly, α-glucoside-configurated aziridine ABPs of GH31 α-glucosidases also have ability to label GBA, which has been described in Chapter 5. Based on the broad spectrum labeling of ABP **2**, more advanced application of this probe such as selective inhibitor discovery can be performed, Figure 4 shows an example to selectively distinguish between GBA and GUSB labeling with suitable concentration of the inhibitors **5**-**7** (used for GUSB inhibition) and **9**-**11** (prefer to GBA inhibition) in competitive ABPP strategy.

Gel based labeling of human platelet by ABP **2** (Figure 5 and SI, Figure S2) and chemical proteomics identification of human spleen by ABP **3** gave the first evidence of both *exo*- and *endo*-β-glucuronidases (GUSB and HPSE) labeling by ABPs presented in Figure 3. Both enzyme labeling are pH dependent. Interestingly, their maximum labeling pH is different, pH 4.0-5.0 for HPSE and pH 5.5-6.5 for GUSB (Figure 4). Thus, a pH-selective way could be used to label either GUSB or HPSE in a mixture of protein lysates. Based on these results, it is proposed that the monosaccharide-based ABPs can also label *endo*-glycosidase which cleaves interally within oligosaccharide chains. This point was also proved by competitive ABPP labeling on human platelet by ABP **4**, the monosugar inhibitors **5**, **6**, **7** and Sistatin B can inhibit both *exo*-GUSB and *endo*-HPSE activity, but oligosaccharide heparin can only block HPSE activity (Figure 5C and SI, Figure S2G,H). Therefore, although *exo*- and *endo*-β-glucuronidases share the same two-step double displacement mechanism, the catalytic pocket of *exo*-β-glucuronidase could only be occupied by monosugar from the terminal end of glycoconjugates, whereas for *endo*-β-glucuronidase, the catalytic cleft is suitable for both monosaccharide and oligosaccharides.

Although a structure of a glucuronic cyclophellitol aziridine bound to a GH2 family enzyme was not obtained during this project, the active site peptide identification data after ABP **3** pull down clearly shows the covalent modification of GH2 GUSB nucleophile Glu540 by ABP **3**. Active site comparisons between GH79 and GH2 enzymes allow us to infer the likely configuration of aziridine probes within a GH2 active site. In the CAZy classification[1,4], both GH79 and GH2 enzymes belong to the wider GH-A clan, enzymes which are characterized by the presence of a $(\beta/\alpha)_8$ barrel , which is a conserved protein fold consisting of eight α-helices and eight parallel β-strands that alternate along the peptide backbone. The active site of both

GH2 and GH79 enzymes are located in this $(\beta/\alpha)_8$ barrel. Furthermore, clan GH-A enzymes show functional conservation of several key amino acids within their active sites, which are involved in the catalytic process.[29]

Crystallographic study of GH79 *exo*-β-glucuronidase (AcaGH79)[6] complex with aziridine **4** has been viewed in Chapter 6. Active site overlays of **4** labelled AcaGH79 and HPSE (GH79) clearly show the conservation of four residues within their active sites, which appear invariant between GH2 and GH79 family enzymes: two glutamates which comprise the catalytic nucleophile and acid/base, an asparagine which lies –1 from the acid/base, and a tyrosine which lies adjacent to the nucleophile (Figure 6C). These are the key amino acids which form the actual catalytic machinery of these enzymes, and are involved in attacking the anomeric centre of glucuronide substrates, or the aziridine ABPs. In contrast, the main structural difference between the active sites of GH2 and GH79 enzymes appears to be in their method of recognising the C8 carboxylate of their glucuronide substrates. GH79 enzymes employ a conserved Tyrosine backbone amide motif to hold the C8 carboxylate, via primarily H-bonding interactions (Figure 6A). GH2 enzymes have been suggested to interact with C6 carboxylates in a more electrostatic fashion, using a charged Asn-X-Lys motif (where X is typically His, Lys or Arg)[30] (Figure 6C). Whilst these differences towards the C8 position may affect the initial non-covalent interaction between ABPs and GH79 vs GH2 enzymes, they are unlikely to play a significant role in the subsequent irreversible covalent bond formation step. Thus given the similarities between the active site machinery of GH2 and GH79 enzymes, it is postulated that ABPs bound to GH2 GUSB should show a similar configuration in the active as observed for AcaGH79 and HPSE.

Four crystal structures of wild type and mutant human HPSE in complex with compound **4** provide strong evidence for the covalent binding of cyclophellitol aziridine to active *endo*-β-glucuronidases and therefore the validity of the cyclophellitol aziridine design for activity-based profiling of retaining glycosidases that employ the Koshland double displacement mechanism. HPSE-**4** complex formation also provides evidence that monosugar type inhibitors or probes are able to effectively bind with *endo*-acting glycosidases. In all cases, aziridines on cyclohexane rings preferably open in a *trans*-diaxial fashion through a $^4H_3$ transition state, and reaction with the enzymes nucleophile takes place at the cyclohexane anomeric equivalent center. This corresponds to ring opening to yield a $^4C_1$ chair conformation, as is observed in the trapped enzyme catalytic cleft in the co-crystals. ABPs bound to AcaGH79 and HPSE in a similar configuration in the active site involving the typical $^1S_3$-$^4H_3$-$^4C_1$ catalytic itinerary of β-glycosidases, and it is infered that GH2 GUSB complex with aziridine ABPs will show the same $^4C_1$ configuration. Last but not least, comparative and competitive ABPP experiments, for monitoring retaining *exo*- and *endo*-β-glucuronidases activities, and for the discovery of

selective inhibitors provide promising tools to study enzyme function *in vivo*, which may also contribute to opening avenues for treating glycosaminoglycan metobolism diseases.

7.4 Experimental section

Biological assays:

Materials

Cyclophellitol ABP **8** and inhibitors **9**, **10** and **11** were prepared as described earlier.[31] Siastain B was bought from Sigma. Human spleen, human Gaucher spleen and fibroblast lysates were prepared same as previous report.[10,19] Cell lines were cultured in HAMF12-DMEM medium (Invitrogen) supplied with 10% (v/v). Rabbit anti-heparanase 1 antibody (ab59787) was obtained from Abcam. Human platelet was from healthy blood donors, mixed with the anti-coagulent (EDTA or citrate) and centrifuged for 15 min at 230 g at 37 °C, transferred plasma containing the platelet to a 15 ml tube and centrifuged for 10 min at 2200 g at 37 °C, collected the enriched platelet pellet fraction and frozen them in small aliquots at -80 °C. All the tissue/cell lysates were prepared in lysis buffer (25 mM potassium phosphate buffer in pH 6.5, supplemented with protease inhibitor 1x cocktail (Roche)) via homogenization with silent crusher S equipped with Typ 7 F/S head (30,000 rpm, 3 × 7 sec) on ice and lysate concentration was determined with Qubit 2.0 Fluorometer assay (Invitrogen). The protein fractions were stored in small aliquots at -80 °C until use.

ABP pull-down and LC-MS/MS analysis

2.0 mg total protein from human fibroblast lysate or 3 mg total protein from human spleen or Gaucher spleen lysate was incubated with either 0.1% (v/v) DMSO, 10 µM **3**, or firstly with 10 µM **8** followed by 10 µM **3**, each step taking 30 min at 37 °C, in a total volume of 0.5 mL McIlvaine buffer, pH 5.0, subsequently denatured through the addition of 10% (w/v) SDS 125 µL and boiling for 5 min at 100 °C. From here on, samples were prepared for pull-down with streptavidin beads as published earlier.[20] After pull-down the samples were divided, 2/3 for on-bead digestion and 1/3 for in-gel digestion. On-bead digestion samples were treated by the trypsin digestion buffer (a mixture containing 100 mM Tris-HCl pH 7.8, 100 mM NaCl, 1.0 mM $CaCl_2$, 2% acetonitrile and 10 ng/µL trypsin) and the bead suspension was incubated in a shaker at 37 °C overnight. The supernatant containing the trypsin-digested peptides was desalted using stage tips, followed by evaporation of MeCN and dilution in 70 µL sample solution (H_2O/MeCN/TFA, 95/3/0.1, v/v/v) for LC-MS analysis. The beads containing active-site peptides were treated with endoproteinase Glu-C digestion buffer (100 ng/µL in PBS solution); incubated in a shaker at 37 °C overnight after which the supernatant was desalted using stage tips and for LC-MS. In-gel digestion samples were eluted by boiling the beads at 100 °C in 30 µL of 1x Laemmli buffer containing 10 µM biotin. The eluted proteins were separated on 10% protein gels at 200 V for 1 h, and the protein gels were silver stained using the Invitrogen kit,[4] and visualized by Bio-rad Chemi-Doc MP Imager using the silver stain channel. Bands were excised with a surgery knife by hand and treated with gel digestion buffer (10 mM NH_4HCO_3, 5% MeCN, 1.0 mM $CaCl_2$, 10 ng/µL trypsin). The supernatant containing the trypsin-digested peptides was desalted using stage tips and prepared for LC-MS. All the peptide samples were analyzed with a 2 h gradient of 5%→25% ACN on nano-LC, hyphenated to an LTQ-Orbitrap and identified via the Mascot protein search engine.[20]

Human HPSE gene cloning, expression and protein purification

Mature recombinant HPSE expression, purification was carried out as previously described strategy.[7]

Bacterial AcaGH79 gene cloning, expression and protein purification

The coding sequence of the AcGH79 gene was cloned into the pET28a (Novagen) expression vector with an N-terminal His6 tag. The E287N mutant was produced by polymerase chain reaction (PCR) using the primer 5'-CCTGACCCAAACGAATTC-3' (forward primer) and 5'-GAATTCGTTTGGGTCAGG-3' (reverse primer). Both the wild-type and mutant proteins were overexpressed in *Escherichia coli* strain BL21 (DE3)GOLD using LB medium. The transformed cells were grown at 37 °C in LB media containing 50 μg mL^{-1} kanamycin unitl the A600 nm reached 0.8. Expression of the recombinant proteins was induced by the addition of 1.0 mM isopropyl β-D-1-thiogalactopyranoside for 12 hr at 25 °C. The cells were harvested by centrifugation at 8000 ×g for 30 min and resuspended in 50 mL lysis buffer (20 mM HEPES, NaCl 200 mM, imidazole 5.0 mM, pH 7.0). After 20 min of sonication and 30 min of centrifugation at 12000 g, the filtered supernatant containing His_6-AcGH79 was loaded onto a His Trap column (GE Healthcare), equilibrated with the lysis buffer. The column was washed with lysis buffer and the His_6-AcGH79 protein was eluted with the same buffer with supplement of 400 mM imidazole over a gradient of 100 mL. The fractions containing the His_6-AcGH79 were then loaded onto a Hiload 16/60 Superdex 75 column (GE Healthcare). The fractions containing the His_6-AcGH79 were pooled and concentrated to the final concentration of 14.5 mg/mL.

***In vitro* labeling and SDS-PAGE analysis and fluorescence scanning**

For labeling procedures see below. All the labeling samples were samples resolved on 10% SDS-PAGE. Electrophoresis in sodium dodecylsulfate containing 10% polyacrylamide gels was performed as earlier described.[32] Wet slab-gels were then scanned for ABP-emitted fluorescence using a Bio-Rad ChemiDoc MP imager using green Cy2 (λ_{EX} 470 nm, bandpass 30 nm; λ_{EM} 530 nm, bandpass 28) for **2** and blue Cy5 (λ_{EX} 625 nm, bandpass 30 nm; λ_{EM} 695 nm, bandpass 55) for **8**. All samples were denatured with 5× Laemmli buffer (50% (v/v) 1.0 M Tris–HCl, pH 6.8, 50% (v/v) 100% glycerol, 10% (w/v) DTT, 10% (w/v) SDS, 0.01% (w/v) bromophenol blue), boiled for 4 min at 100 °C, and separated by gel electrophoresis on 10% (w/v) SDS-PAGE gels running continuously at 90 V for 30 min and 200 V for 50 min.

A) Human/Gaucher spleen extracts labeling: Human spleen lysates from control or Gaucher patient (20 μg total protein per sample) were accommodated in McIlvaine buffer pH 5.0 (10 μL total volume) for 5 min on ice. To selectively inhibit GBA or beta-glucuronide, samples were pre-incubated with 1.0 μM **8**, 1.0 μM **3**, or DMSO for 30 min at 37 °C, followed by 30 min incubation with 100 nM probe **2** or DMSO for another 30 min. For competitive ABPP using ctrl spleen to screen inhibitor selectivity, the samples were pre-incubated with 1.0 mM **7**, 100 μM **6**, **5**, 1.0 mM **11**, 100 μM **10** and 100 μM **9** for 30 min at 37°C, followed by 30 min incubation of 100 nM **2** or 1.0 μM **8** or probe mixture of 100 nM **2** and 1.0 μM **8** for another 30 min at 37 °C.

B) Mouse tissues extracts labeling: For direct labeling in mouse tissue extracts, 20 μg total protein per sample was accommodated in McIlvaine buffer pH 5.0 (10 μL total volume) for 5 min on ice and treated with 1.0 μM **2** (10 μL 2x solution in McIlvaine buffer). For ABPP on lysate of C57Bl6/J mouse liver, 50 μg in total was incubated with 1.0 μM **2** for 30 min in 150 mM McIlvaine buffer, pH 5.0, whilst at 37 °C.

C) Human recombinant HPSE labeling: All the recombinant HPSE labeling was performed by incubation of 200 ng HPSE with 100 nM **2** for 30 min at 37 °C. The influence of pH on ABP labeling was established by pre-incubation of 200 ng HPSE at pH 3.0–7.0 citric acid – Na_2HPO_4 buffer solutions for 5 min on ice, and incubation with 100 nM **2** for 30 min at 37 °C.

D) Human cell lines labeling: Direct labeling in overexpressed HEK293 cell lysate from different timepoints or internalized human fibroblast cell lysate (20 µg total protein per sample) were accommodated in 150 mM McIlvaine buffer pH 5.0 (10 µL total volume) for 5 min on ice and treated with 100 nM **2**, incubated for 30 min at 37 °C, followed by Western blot.

E) Human platelet extracts labeling: Direct labeling with ABP **2** of human platelet-EDTA or platelet-citric lysate (20 µg total protein per sample) was accommodated in 150 mM McIlvaine buffer pH 5.0 (10 µL total volume) for 5 min on ice and treated with 100 nM **2** for 30 min at 37 °C. ABP concentration and time limit labeling was analyzed by incubation of 20 µg total protein with 0–100 nM **22** in 150 mM McIlvaine buffer, pH 5.0, for 10 min, 30 min and 60min at 37 °C respectively. Influence of pH on ABP labeling was established by pre-incubation of 20 µg total protein at pH 3.0–7.0 for 5 min on ice and incubation with 100 nM **2** for 30 min at 37 °C. For competitive ABPP on platelet, 100 ng total protein was pre-incubated with decreasing amount of heparin (5.0 – 0.25 mg/mL) and then treated with 100 nM **2** for 30 min at 37 °C.

Western blot analysis

After SDS-PAGE analysis and fluorescence scanning, proteins were transferred from gel to a PVDF membrane using a Trans-Blot® Turbo system (Bio-Rad). Human HPSE was stained using rabbit anti-heparanase 1 antibody as primary antibody, and goat-anti-rabbit HRP as secondary antibody. The blot was developed in the dark using a 10 mL luminal solution, 100 µL ECL enhancer and 3.0 µL 30% H_2O_2 solution. Chemiluminescence was visualized using a ChemiDoc XRS (Bio-Rad).

Overexpression of HPSE in HEK293T cells

pGEn1-HPSE and pGEn2-HPSE plasmids were obtained from the DNASU repository.[33] HEK293T cells were grown in DMEM media supplemented with 10% newborn calf serum (NBCS). 20 µg DNA was transfected into low passage HEK293T cells at ~80% confluency using PEI at a ratio of 3:1 (PEI : DNA). After 7 h, 24 h, 48 h and 72 h cells were washed with ice cold PBS, harvested using a cell scraper, and pelleted. Cell pellets were stored at -80 °C prior to use in labeling experiments.

proHPSE fibroblast internalized experiment

Primary human fibroblasts were grown in F12 media supplemented with 10% NBCS and penicillin/streptomycin. Cells at 70-80% confluency were washed with PBS, before addition of F12 media supplemented with proHPSE at 10 µg/mL final concentration. At relevant time points, cells were washed with ice cold PBS twice, harvested using a cell scraper, and pelleted. Cell pellets were stored at -80 °C prior to use in labeling experiments.

proHPSE bacmid preparation and protein production

Baculovirus encoding for N-6xHis-TEV-proHPSE was sub-cloned from the pGEn1-HPSE plasmid. cDNA encoding for the mellitin signal peptide and TEV-proHPSE were extracted by PCR, along with synthetic DNA encoding for 6xHis-TEV. DNA fragments were ligated using T4 DNA ligase (NEB), before cloning into a pOMNI vector (Geneva biotech) using SLIC.[34] All PCR steps were carried out using Phusion polymerase (Thermo). Primers and synthetic DNA sequences are listed in (Table 1). Subsequent bacmid and expression steps were carried out as described for HPSE.

Table 1 Primers used for cloning proHPSE construct.

pOMNI-Mellitin F primer	ccatcgggcgcggatccatgaaatttttggtg
Mellitin-XmaI R primer	gatcgcccgggtccgcataaatgtagctaatg
XmaI-6xHis-KpnI F sequence	ctagcccgggccatcatcaccaccatcatggtaccgatc
XmaI-6xHis-KpnI R sequence	gatcggtaccatgatggtggtgatgatggcccgggctag
KpnI-TEV-proHPSE F primer	gatcggtaccgcagaaaacttgtactttcaaggccagg
proHPSE-pOMNI R primer	gtacttctcgacaagctttcagatgcaagcagcaactttg

Purification of pro-HPSE

3.0 L of conditioned media was cleared of cells by centrifugation at 400 g for 15 min at 4 °C, followed by further clearing of debris by centrifugation at 4000 g for 60 min at 4 °C. DTT (1.0 mM) and AEBSF (0.10 mM) were added to cleared media, which was then loaded onto a pre-equilibrated HiTrap Sepharose SP FF 5.0 mL column (GE healthcare). The SP column was washed with 10 CV of IEX buffer A (20 mM HEPES (pH = 7.4), 100 mM NaCl, 1.0 mM DTT), and eluted with a linear gradient over 30 CV using IEX buffer B (20 mM HEPES (pH = 7.4), 1.5 mM NaCl, 1.0 mM DTT). proHPSE containing fractions were pooled and diluted 10 fold into HisTrap buffer A (20 mM HEPES (pH = 7.4), 500 mM NaCl, 20 mM imidazole, 1.0 mM DTT), before loading onto a pre-equilibrated HisTrap FF 1.0 mL column (GE healthcare). The HisTrap column was washed with 10 CV HisTrap buffer A, and eluted with a linear gradient over 30 CV using HisTrap buffer B (20 mM HEPES (pH = 7.4), 500 mM NaCl, 1.0 M imidazole, 1.0 mM DTT). proHPSE containing fractions were pooled and concentrated to ~2.0 mL using a 30 kDa cutoff Vivaspin concentrator (GE Healthcare), and treated with 5.0 µL EndoH (NEB) and 5.0 uL AcTEV protease (Invitrogen) for 4 h. Digested protein was purified by size exclusion chromatography (SEC) using a Superdex S75 16/600 column (GE Healthcare) in SEC buffer (20 mM HEPES (pH = 7.4), 200 mM NaCl, 1.0 mM DTT). proHPSE containing fractions were concentrated to 1.0 mg/mL using a 30 kDa Vivaspin concentrator, and flash frozen for use in further experiments.

HPSE-4 complex crystallography and structure solution

HPSE complex with **4** was generated by soaking the *apo* crystal in 250 µM **4** in mother liquor (100 mM MES (pH = 5.5), 100 mM $MgCl_2$, 17% PEG3350). Crystals were soaked overnight at room temperature, transferred to cryoprotectant solution (100 mM MES (pH = 5.5), 100 mM $MgCl_2$, 17% PEG3350, 25% DMSO), and subsequently flash frozen in liquid N_2 for data collection. Complexes were solved by refining directly against the *apo* structure using REFMAC5,[35] before further rounds of manual model building and refinement using Coot[36] and REFMAC5. Ligand coordinates were built using jLigand.[37] Crystal structure figures were generated using ccp4mg.[38]

7.5 References

[1] V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho, B. Henrissat, *Nucleic Acids Res.* **2014**, *42*, D490-D495.

[2] L. Arul, G. Benita, D. Sudhakar, B. Thayumanavan, P. Balasubramanian, *Bioinformation* **2008**, *3*, 194-197.

[3] D. E. Koshland, *Biol. Rev.* **1953**, *28*, 416-436.

[4] B. Henrissat, G. J. Davies, *Curr. Opin. Struct. Biol.* **1997**, *7*, 637-644.

[5] W. S. Sly, Vogler, C. *Nat. Med.* **1997**, *3*, 719-720.

[6] M. Michikawa, H. Ichinose, M. Momma, P. Biely, S. Jongkees, M. Yoshida, T. Kotake, Y. Tsumuraya, S. G. Withers, Z. Fujimoto, S. Kaneko, *J. Bio. Chem.* **2012**, *287*, 14069-14077.

[7] L. Wu, C. M. Viola, A. M. Brzozowski, G. J. Davies, *Nat. Struct. Mol. Biol.* **2015**, 22, 1016–1022

[8] I. Vlodavsky, P. Beckhove, I. Lerner, C. Pisano, A. Meirovitz, N. Ilan, M. Elkin, *Cancer Microenvironment* **2012**, *5*, 115-132.

[9] B. F. Cravatt, A. T. Wright, J. W. Kozarich, *Annu. Rev. Biochem.* **2008**, *77*, 383-414.

[10] W. W. Kallemeijn, K. Y. Li, M. D. Witte, A. R. Marques, J. Aten, S. Scheij, J. Jiang, L. I. Willems, T. M. Voorn-Brouwer, C. P. van Roomen, R. Ottenhoff, R. G. Boot, H. van den Elst, M. T. Walvoort, B. I. Florea, J. D. Codee, G. A. van der Marel, J. M. Aerts, H. S. Overkleeft, *Angew. Chem. Int. Ed.* **2012**, *51*, 12529-12533.

[11] J. Jiang, T. J. M. Beenakker, W. W. Kallemeijn, G. A. van der Marel, H. van den Elst, J. D. C. Codee, J. M. F. G. Aerts, H. S. Overkleeft, *Chem. Eur. J.* **2015**, *21*, 10861-10869.

[12] L. I. Willems, T. J. M. Beenakker, B. Murray, S. Scheij, W. W. Kallemeijn, R. G. Boot, M. Verhoek, W. E. Donker-Koopman, M. J. Ferraz, E. R. van Rijssel, B. I. Florea, J. D. C. Codee, G. A. van der Marel, J. M. F. G. Aerts, H. S. Overkleeft, *J. Am. Chem. Soc.* **2014**, *136*, 11622-11625.

[13] J. Jiang, W. W. Kallemeijn, D. W. Wright, A. M. C. H. van den Nieuwendijk, V. C. Rohde, E. C. Folch, H. van den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schurmann, D. Mink, R. G. Boot, J. D. C. Codee, G. A. van der Marel, G. J. Davies, J. M. F. G. Aerts, H. S. Overkleeft, *Chem. Sci.* **2015**, *6*, 2782-2789.

[14] R. Petryszak, T. Burdett, B. Fiorelli, N. A. Fonseca, M. Gonzalez-Porta, E. Hastings, W. Huber, S. Jupp, M. Keays, N. Kryvych, J. McMurry, J. C. Marioni, J. Malone, K. Megy, G. Rustici, A. Y. Tang, J. Taubert, E. Williams, O. Mannion, H. E. Parkinson, A. Brazma, *Nucleic Acids Res.* **2014**, *42*, D926-D932.

[15] M. Ono, N. Taniguchi, A. Makita, M. Fujita, C. Sekiya, M. Namiki, *J. Biol. Chem.* **1988**, *263*, 5884-5889.

[16] M. R. Islam, J. H. Grubb, W. S. Sly, *J. Biol. Chem.* **1993**, *268*, 22627-22633.

[17] M. D. Witte, W. W. Kallemeijn, J. Aten, K.-Y. Li, A. Strijland, W. E. Donker-Koopman, A. M. C. H. van den Nieuwendijk, B. Bleijlevens, G. Kramer, B. I. Florea, B. Hooibrink, C. E. M. Hollak, R. Ottenhoff, R. G. Boot, G. A. van der Marel, H. S. Overkleeft, J. M. F. G. Aerts, *Nat. Chem. Biol.* **2010**, *6*, 907-913.

[18] R. G. Boot, C. E. M. Hollak, M. Verhoek, P. Sloof, B. J. H. M. Poorthuis, W. J. Kleijer, R. A. Wevers, M. H. J. van Oers, M. M. A. M. Mannens, J. M. F. G. Aerts, S. van Weely, *Human Mutant.* **1997**, *10*, 348-358.

[19] N. Li, C.-L. Kuo, G. Paniagua, H. van den Elst, M. Verdoesm, L. I. Willems, W. A. van der Linden, M. Ruben, E. van Genderen, J. Gubbens, G. P. van Wezel, H. S. Overkleeft, B. I. Florea, *Nat. Protocols* **2013**, *8*, 1155-1168.

[20] Y. Ishihama, Y. Oda, T. Tabata, T. Sato, T. Nagasu, J. Rappsilber, M. Mann, *Mol. Cell Proteomics* **2005**, *4*, 1265-1272.

[21] J. J. Birktoft, K. Breddam, *Methods Enzymol.* Academic Press: **1994**, *244*, 114-126.

[22] M. B. Fairbanks, A. M. Mildner, J. W. Leone, G. S. Cavey, W. R. Mathews, R. F. Drong, J. L. Slightom, M. J. Bienkowski, C. W. Smith, C. A. Bannow and R. L. Heinrikson, *J.Biol. Chem.* **1999**, *274*, 29587-29590.

[23] L. Nadav, A. Eldor, O. Yacoby-Zeevi, E. Zamir, I. Pecker, N. Ilan, B. Geiger, I. Vlodavsky, B. Z. Katz, *J. Cell Sci.***2002**, *115*, 2179-2187.

[24] I. Shafat, A. B. Barak, S. Postovsky, R. Elhasid, N. Ilan, I. Vlodavsky and M. W. Arush, Neoplasia **2007,** *9*, 909-916.

[25] Y. Nishimura, *J. Antibiot.* **2009**, *62*, 407-423.

[26] G. J. Davies, A. Planas, C. Rovira, *Accounts Chem. Res.* **2012**, 45, 308-316.

[27] A. Ardevol, C. Rovira, *J. Am. Chem. Soc.* **2015**, *137*, 7528-7547.

[28] B. D. Wallace, H. Wang, K. T. Lane, J. E. Scott, J. Orans, J. S. Koo, M. Venkatesh, C. Jobin, L. A. Yeh, S. Mani, M. R. Redinbo, *Science* **2010**, *330*, 831-835.

[29] P. Durand, P. Lehn, I. Callebaut, S. Fabrega, B. Henrissat, J. P. Mornon, *Glycobiology* **1997**, 7, 277-284.

[30] B. D. Wallace, A. B. Roberts, R. M. Pollet, J. D. Ingle, K. A. Biernat, S. J. Pellock, M. K. Venkatesh, L. Guthrie, S. K. O'Neal, S. J. Robinson, M. Dollinger, E. Figueroa, S. R. McShane, R. D. Cohen, J. Jin, S. V. Frye, W. C. Zamboni, C. Pepe-Ranney, S. Mani, L. Kelly, M. R. Redinbo, *Chem. Biol.* **2015**, 22, 1238-1249.

[31] K.-Y. Li, J. Jiang, M. D. Witte, W. W. Kallemeijn, H. van den Elst, C.-S. Wong, S. D. Chander, S. Hoogendoorn, T. J. M. Beenakker, J. D. C. Codée, J. M. F. G. Aerts, G. A. van der Marel, H. S. Overkleeft, *Eur. J. Org. Chem.* **2014**, *2014*, 6030-6043.

[32] J. M. Aerts, A. W. Schram, A. Strijland, S. van Weely, L. M. Jonsson, J. M. Tager, S. H. Sorrell, E. I. Ginns, J. A. Barranger, G. J. Murray, *Biochim. Biophys. Acta* **1988**, *964*, 303-308.

[33] C. Y. Seiler, J. G. Park, A. Sharma, P. Hunter, P. Surapaneni, C. Sedillo, J. Field, R. Algar, A. Price, J. Steel, A. Throop, M. Fiacco, J. LaBaer, *Nucleic Acids Res.* **2014**, *42*, D1253-D1260.

[34] M. Z. Li, S. J. Elledge, *Methods Mol. Biol.* **2012**, *852*, 51-59.

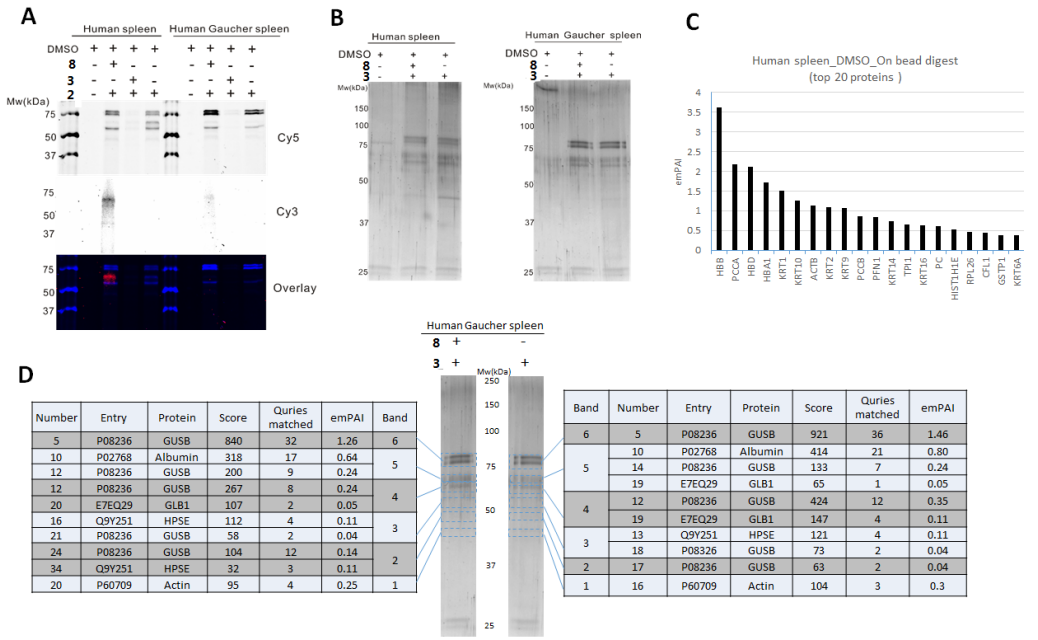
[35] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. D* **2011** *67*, 355-367.

[36] P. Emsley, K. Cowtan, *Acta Crystallogr. D* **2004**, *60*, 2126-2132.

[37] A. A. Lebedev, P. Young, M. N. Isupov, O.V. Moroz, A. A. Vagin, G. N. Murshudov, *Acta Crystallogr. D* **2012**, *68*, 431-440.

[38] S. McNicholas, E. Potterton, K. S. Wilson, M. E. Noble, *Acta Crystallogr. D* **2011**, *67*, 386-394.

7.7 Supporting Information

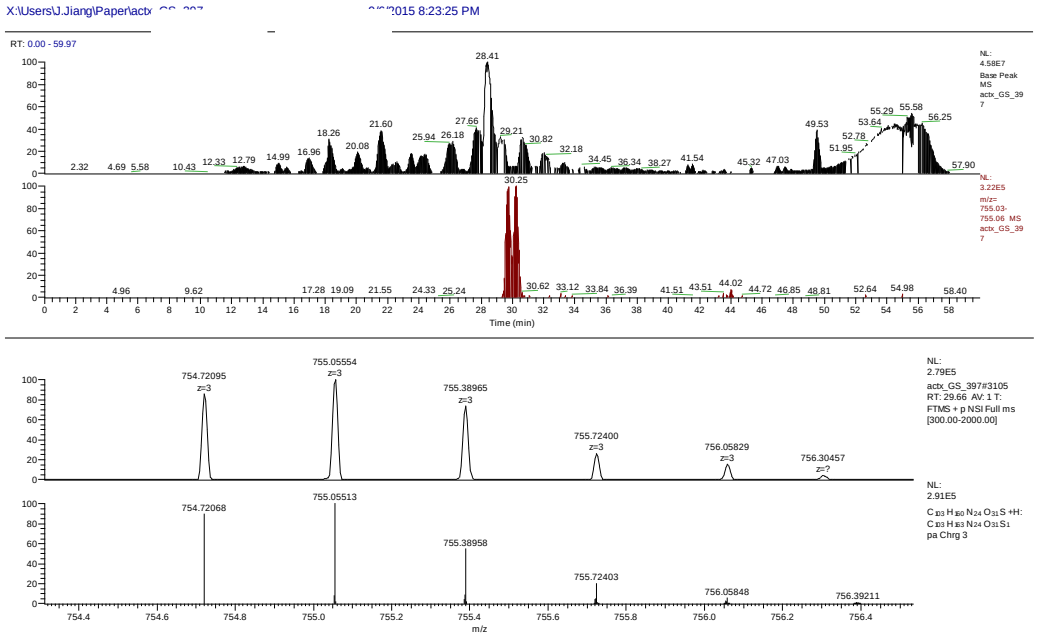


D

Number	Entry	Protein	Score	Quries matched	emPAI	Band
5	P08236	GUSB	840	32	1.26	6
10	P02768	Albumin	318	17	0.64	5
12	P08236	GUSB	200	9	0.24	
12	P08236	GUSB	267	8	0.24	4
20	E7EQ29	GLB1	107	2	0.05	
16	Q9Y251	HPSE	112	4	0.11	3
21	P08236	GUSB	58	2	0.04	
24	P08236	GUSB	104	12	0.14	2
34	Q9Y251	HPSE	32	3	0.11	
20	P60709	Actin	95	4	0.25	1

Band	Number	Entry	Protein	Score	Quries matched	emPAI
6	5	P08236	GUSB	921	36	1.46
5	10	P02768	Albumin	414	21	0.80
	14	P08236	GUSB	133	7	0.24
	19	E7EQ29	GLB1	65	1	0.05
4	12	P08236	GUSB	424	12	0.35
	19	E7EQ29	GLB1	147	4	0.11
3	13	Q9Y251	HPSE	121	4	0.11
	18	P08326	GUSB	73	2	0.04
2	17	P08236	GUSB	63	2	0.04
1	16	P60709	Actin	104	3	0.3

E



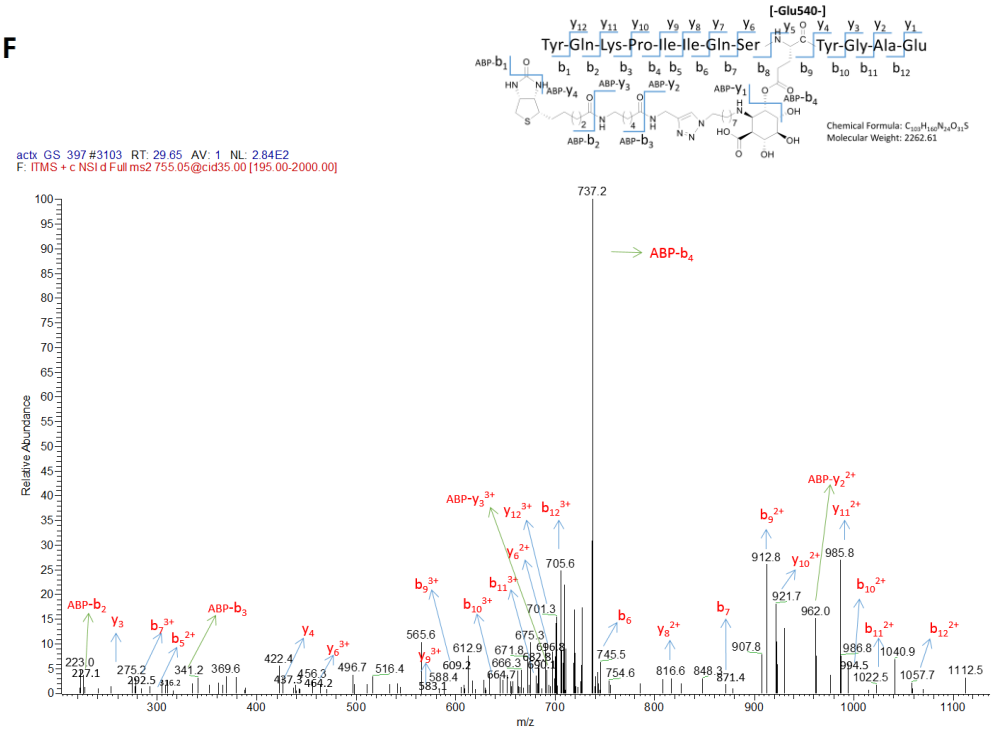


Figure S1. A) Human spleen and human Gaucher spleen extracts labeled with ABP **3** and with or without pre-incubation with **8**, independently shown in Cy5 and Cy3 fluorescent channels. B) Silver stain gel of human spleen and human Gaucher spleen extracts after pull-down with ABP **3** without pre-incubation with **8**; these gels were used for in-gel digestion. C) In-gel trypsin digestion proteins table from human Gaucher spleen extracts involving target glycosidases and off-target proteins. D) Human spleen incubation with DMSO sample after pull-down on bead digest results (emPAI values) of captured proteins. E) Electrospray ionization spectra of ABP **3** modified GUSB nucleophile active-site peptide in 2 charge statement. F) Identification of the GUSB active site fragment modified with biotin-ABP **3** after on-bead capture, trypsinolysis and LC-MS/MS analysis of the tryptic fragment containing the modified active site nucleophile.

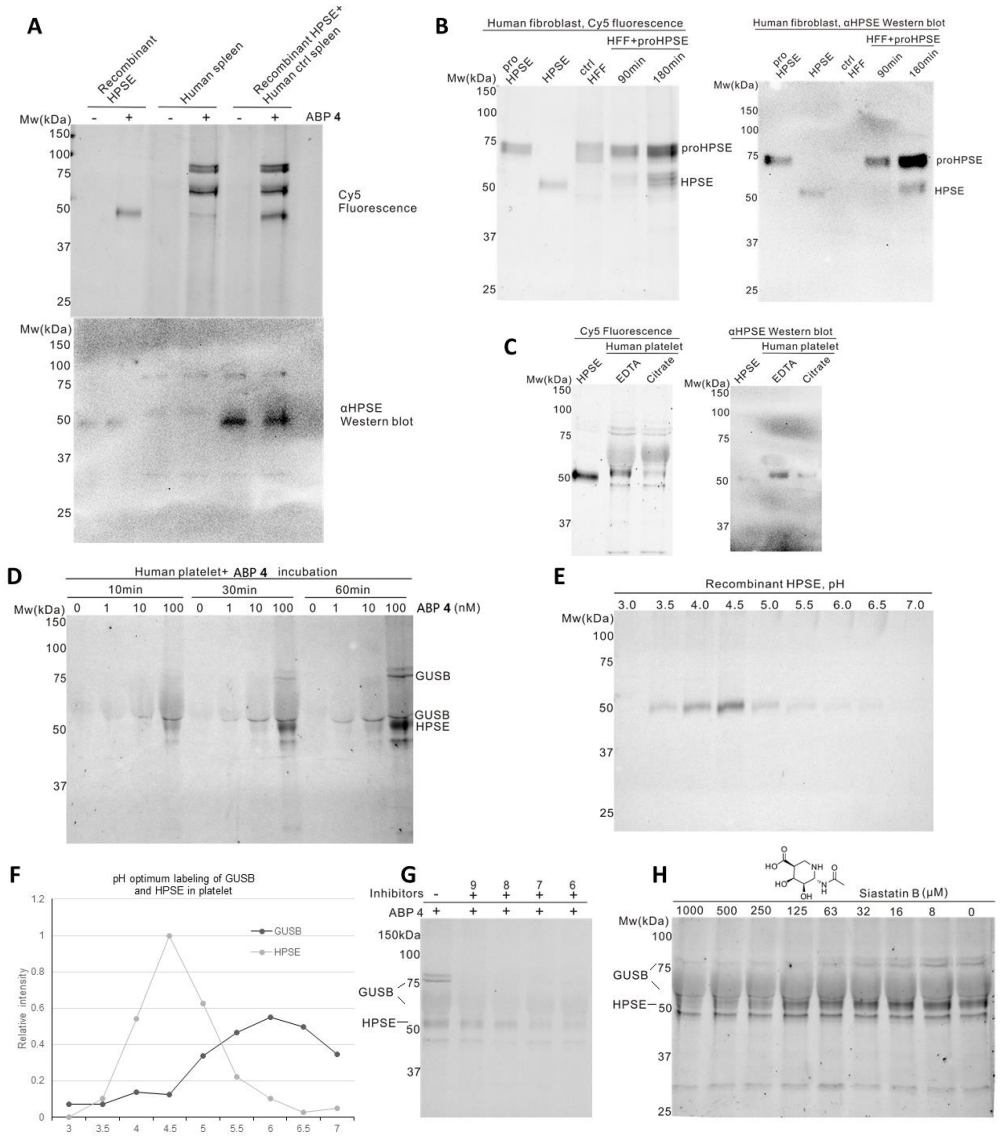


Figure S2. A) Recombinant HPSE, human spleen extracts and their mixture labeled with ABP **2**. SDS-PAGE of Cy5 fluorescent channel is shown on the left and Western blot of PVDF membrane scanning of chemiluminescence channel is shown on the right. B) ABP **2** labels internalized HPSE in human fibroblast (HF). Fibroblast cells are able to internalize and process proHPSE in culture media. ABP labeling intensity reflects intracellular accumulation of HPSE over time. C) Endogenous HPSE labeling from human platelet EDTA and citrate buffer by ABP **2**. SDS-PAGE of Cy5 fluorescent channel is shown on the left and Western blot of PVDF membrane scanning of chemiluminescence channel is shown on the right. D) Time- and concentration-dependent labeling of human platelet extract by ABP **2** as visualized by SDS-PAGE and Western blot. E) pH optimum labeling of human platelet GUSB and HPSE, with maximum labeling for GUSB at pH 5.5-6.5, and that of HPSE at pH 4.0-5.0. F) Detection of recombinant HPSE labeling efficiency at various pH. G) HSPE labeling in platelets is abolished by pre-incubation with known inhibitor Siastatin B.