University of Groningen

Endo- β -Glucosidase Tag Allows Dual Detection of Fusion Proteins by Fluorescent Mechanism-Based Probes and Activity Measurement

Kallemeijn, Wouter W; Scheij, Saskia; Voorn-Brouwer, Tineke M; Witte, Martin D.; Verhoek, Marri; Overkleeft, Hermen S; Boot, Rolf G; Aerts, Johannes M F G

Published in:
ChemBioChem

DOI:
[10.1002/cbic.201600312](https://doi.org/10.1002/cbic.201600312)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Kallemeijn, W. W., Scheij, S., Voorn-Brouwer, T. M., Witte, M. D., Verhoek, M., Overkleeft, H. S., Boot, R. G., & Aerts, J. M. F. G. (2016). Endo- β -Glucosidase Tag Allows Dual Detection of Fusion Proteins by Fluorescent Mechanism-Based Probes and Activity Measurement. *ChemBioChem*, 17(18), 1698-1704. <https://doi.org/10.1002/cbic.201600312>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Endo- β -Glucosidase Tag Allows Dual Detection of Fusion Proteins by Fluorescent Mechanism-Based Probes and Activity Measurement

Wouter W. Kallemeyjn,^[a] Saskia Scheij,^[b] Tineke M. Voorn-Brouwer,^[b] Martin D. Witte,^[c, d] Marri Verhoek,^[a] Hermen S. Overkleeft,^[d] Rolf G. Boot,^[a] and Johannes M. F. G. Aerts*^[a]

β -Glucoside-configured cyclophellitols are activity-based probes (ABPs) that allow sensitive detection of β -glucosidases. Their applicability to detect proteins fused with β -glucosidase was investigated in the cellular context. The tag was *Rhodococcus* sp. M-777 endoglycoceramidase II (EGCaseII), based on its lack of glycans and ability to hydrolyze fluorogenic 4-methylumbelliferyl β -D-lactoside (an activity absent in mammalian cells). Specific dual detection of fusion proteins was possible in vitro and in situ by using fluorescent ABPs and a fluorogenic substrate. Pre-blocking with conduritol β -epoxide (a poor inhibitor of EGCaseII) eliminated ABP labeling of endogenous β -glucosidases. ABPs equipped with biotin allowed convenient purification of the fusion proteins. Diversification of ABPs (distinct fluorophores, fluorogenic high-resolution detection moieties) should assist further research in living cells and organisms.

Visualization of proteins in vivo in their cellular context has enormously contributed to knowledge on physiological function. After the discovery of green fluorescent protein (GFP) from *Aequoera victoria*,^[1] a large palette of fluorescence-emitting proteins (XFPs) has been engineered, thus allowing studies on gene expression, protein localization, and other phenomena.^[2] Typically the nucleotide sequence encoding a protein of interest is (over)expressed in-frame with an XFP to generate a fluorescent chimera. Chromophore photobleaching, photoconversion, and light-induced damage (i.e., by production of reactive oxygen species) to the surrounding biological

medium often limit extended observation of XFPs.^[3] In view of this, other strategies have been developed to label proteins, such as self-labeling protein tags and non-natural amino acids. The various methods have been excellently reviewed.^[4] The use of self-labeling protein tags is an attractive alternative to chimeras with permanent XFPs. Self-labeling tags are relatively small proteins with unique reactivity toward weak electrophiles. Widely used are the commercially available SNAP-tag and HaloTag, which can be detected by incubating cells with fluorescent moieties that are trapped by these tags.^[5] Extended in situ visualization is facilitated by the use of fluorophores such as synthetic boron dipyrromethene (BODIPY), which rarely produces reactive oxygen species.^[6]

We previously developed highly potent activity-based probes (ABPs) that bind specifically and selectively to the catalytic nucleophile of target β -glucosidases, using their enzymatic retaining double-displacement mechanism.^[7] Cyclophellitol β -epoxide ABPs **1** and **2** (Scheme 1 A) bound highly specifically in vivo to the target β -glucosidase lysosomal glucocerebrosidase (GBA).^[7a] Cyclophellitol β -aziridine ABPs **3** and **4** (Scheme 1 B) bound, in addition, to the nucleophile of non-lysosomal β -glucosidase (GBA2), cytosolic β -glucosidase (GBA3), and lactase/phlorizin hydrolase (LPH).^[7b] Both ABP classes irreversibly modify the enzymatic nucleophile by a double-displacement mechanism intrinsic to β -glucosidase (Scheme 1 A and B). The incorporated fluorescent BODIPY enables ultra-sensitive visualization of attomoles of ABP-labeled protein by in-gel imaging and microscopy.^[7]

We envisioned exploiting this new ABP technology for specific in situ detection of the β -glucosidases, in order to visualize proteins of interest fused with an appropriate *endo*- or *exo*- β -glucosidase (Scheme 1 C). We identified *Rhodococcus* sp. M-777 endoglycoceramidase II (EGCaseII) as an attractive tag. Firstly, EGCaseII hydrolyzes the linkage between and ceramide and is consequently likely to be labeled with β -glucoside configured complex oligosaccharide cyclophellitol-type ABPs, similarly to GBA and GBA2.^[7,8] EGCaseII, lacking N-linked glycans, requires no processing to reach its active conformation, thus making it an attractive tag in contrast to GBA or GBA2.^[9] In addition, EGCaseII has a unique catalytic feature: it is able to hydrolyze the fluorogenic substrate 4-methylumbelliferyl β -D-lactoside (4MU- β -D-Lac) to yield 4-methylumbelliferyl (4MU) and lactose (an *endo*- β -glucosidase activity absent in most mammalian cells). Fusion of EGCaseII in-frame with the protein of choice should yield a chimera that can be labeled with fluores-

[a] Dr. W. W. Kallemeyjn, M. Verhoek, Dr. R. G. Boot, Prof. Dr. J. M. F. G. Aerts
Department of Medical Biochemistry
Leiden Institute of Chemistry, Leiden University
Einsteinweg 55, 2333 CC Leiden (NL)
E-mail: j.m.f.g.aerts@lic.leidenuniv.nl

[b] S. Scheij, T. M. Voorn-Brouwer
Department of Medical Biochemistry
Academic Medical Center, University of Amsterdam
Meibergdreef 15, 1105 AZ Amsterdam (NL)

[c] Dr. M. D. Witte
Department of Bio-Organic Chemistry
Stratingh Institute for Chemistry, University of Groningen
Nijenborgh 7, 9747 AG Groningen (NL)

[d] Dr. M. D. Witte, Prof. Dr. H. S. Overkleeft
Department of Bio-organic Synthesis
Leiden Institute of Chemistry, Leiden University
Einsteinweg 55, 2333 CC Leiden (NL)

Supporting information for this article can be found under <http://dx.doi.org/10.1002/cbic.201600312>.

Scheme 1. A) Cyclophellitol β -epoxide ABPs **1** and **2** (green and red fluorescence, respectively), and labeling mechanism. B) Cyclophellitol β -aziridine ABPs **3** and **4** (green and red fluorescence, respectively), and labeling mechanism. C) EGCasell-tagging strategy to detect and visualize proteins of interest in vitro and in situ.

cent ABPs or sensitively detected by measurement of the activity towards 4MU- β -D-Lac (Scheme 1 C).

Studies on EGCasell suitability as ABP-labeling tag

We first examined the labeling efficiency of EGCasell with the cyclophellitol β -epoxide and β -aziridine ABPs and compared the results with GBA, the most abundant and most potently ABP-targeted β -glucosidase in mammalian cells. We labeled commercial recombinant EGCasell and GBA with **1** and **3** by incubating 1 μ g of protein with a range of ABP concentrations for 1 h at 37 °C and at the optimal pH for enzymatic activity (pH 5.5 for EGCasell; pH 5.2 for GBA; Figure 1A). The enzymes were comparably detected with just 10 to 100 fM fluorescent ABP **1** or **3**. Two EGCasell protein bands are labeled with both ABPs; the higher-molecular-weight form has been described as a complex of EGCasell with activator II protein and/or aggregated recombinant EGCasell.^[10] Labeling of EGCasell with **1** was more intense at pH 6.0 than at pH 5.0 in the absence of detergent (Figure 1B). In the presence of 0.2% (w/v) sodium taurocholate and 0.1% (v/v) Triton X-100 (required for optimal ABP labeling of GBA), the labeling of EGCasell with **1** was markedly less than of GBA. Optimum labeling of EGCasell and GBA coincided with the optimal conditions for enzymatic activity. EGCasell (but not GBA) was well able to hydrolyze 4MU- β -D-Lac (Figure 1C), an activity that is optimal in the absence of the detergents Triton X-100 and sodium taurocholate (Figure 1D).

Analysis of in vitro results by in silico docking

The observation that EGCasell hydrolyzes 4MU- β -D-Lac markedly better than 4MU- β -D-Glc is consistent with *in silico* dock-

ing analysis of these substrate molecules with the crystal structures of EGCaseII and GBA (PDB IDs: 2OYK and 2V3E, respectively). 4MU- β -D-Lac fits optimally in the active-site groove of EGCaseII, with the C₁ of D-glucose in close proximity to nucleophile Glu351 and the inter-glycosidic bond oxygen pointing towards the acid/base Glu233 (Figure S1A in the Supporting Information).^[8,9] Furthermore, the coordination of 4MU- β -D-Lac in the active site is similar to that found earlier for co-crystallized ganglioside GM3 in complex with EGCaseII co-crystallized GM3 lipid (EGCaseII overlay with (PDB ID: 2OSX; Figure S1B). In contrast, analysis of the lowest free-energy conformer of 4MU- β -D-Lac on GBA revealed that its active site cannot accommodate the entire substrate (i.e., with C₁ of D-glucose within the required 4–5 Å of nucleophile Glu340, and the glucoside bond oxygen within the same distance of acid/base Glu235; Figure S1C and D). Molecular docking of 4MU- β -D-Glc showed that this substrate fits well in the EGCaseII pocket (i.e., close to the reactive residues) with high affinity (-8.6 ± 0.1 kcal mol⁻¹, Figure S1E and F) but with slightly less affinity than for GBA (-9.4 ± 0.1 kcal mol⁻¹, Figure S1G and H).

Distinguishing EGCase1 from endogenous β -glucosidase activity

Selective labeling of EGCaseII is essential for its use as tag. Therefore, we investigated whether conduritol β -epoxide (CBE) discriminates between GBA and EGCaseII.^[7] GBA and EGCaseII were equally mixed with each other, and pre-incubated with CBE (0–10 mM) for 30 min at 37 °C. Pre-incubation was performed at pH 5.2 with 0.2% (w/v) sodium taurocholate and 0.1% (v/v) Triton X-100 (Figure 1E; optimal condition for GBA activity) or pH 5.5 without detergents (Figure 1F; optimal for EGCaseII). Subsequently, EGCaseII and GBA were labeled with

Figure 1. In vitro activity and inhibition of recombinant EGCaseII. A) In vitro labeling of recombinant EGCaseII and GBA with **1** and **3**. B) Effect of detergents (0.2% (w/v) sodium taurocholate with 0.1% (v/v) Triton X-100) and pH on in vitro ABP labeling of recombinant EGCaseII and GBA with **3**. C) Hydrolysis of 4MU-conjugated substrates by recombinant GBA and EGCaseII. Data are average $\pm$ SD ($n = 3$). D) Hydrolysis of 4-methylumbelliferyl β -D-lactoside (4MU- β -D-Lac) by EGCaseII or 4-methylumbelliferyl β -D-glucopyranoside (4MU- β -D-Glc) by GBA (circles and squares, respectively) in the presence or absence of detergents (open and closed, respectively) at various pH values. Minimal hydrolysis (< 0.1) of 4MU- β -D-Glc by EGCaseII and 4MU- β -D-Lac by GBA, in the absence and presence of detergents, omitted for clarity. Data are average $\pm$ SD ($n = 3$). E) Blocking of recombinant GBA and recombinant EGCaseII by CBE at pH 5.2 in presence of detergents. Equimolar GBA/EGCaseII mixture was pre-incubated with different CBE concentrations, followed by labeling with **3**. F) Similar blocking of GBA with CBE, but at pH 5.5 and without detergents.

1 μ M **3** (same conditions as for pre-incubation). Pre-incubation at pH 5.2 with CBE and detergents completely prevented ABP labeling of GBA, whereas labeling of EGCaseII was not affected, even by 10 mM CBE (Figure 1E). At pH 5.5 without detergent, labeling of EGCaseII with **1** improved (Figure 1F), and was hardly blocked by CBE (only at > 5 mM is the ABP-labeling of EGCaseII reduced). In sharp contrast, the labeling of GBA was completely inhibited by 10 μ M CBE (Figure 1E and F). Thus, EGCaseII can be selectively labeled with ABPs in the presence of CBE. Based on this, we decided to proceed with generating EGCaseII fusion proteins.

Generation of EGCaseII-appended fusion proteins

We derived a nucleotide sequence based on the published EGCaseII protein sequence,^[11] codon-optimized for mammalian expression and followed this by oligonucleotide synthesis. Ito et al.^[8,9] suggested that a signal peptide is present in EGCaseII; this might affect localization of chimera proteins within mammalian cells. We therefore cloned EGCaseII with and without the predicted signal peptide (EGCaseII, Δ spEGCaseII, respectively) and fused EGFP to its C terminus with a linker sequence (to encode chimeras EGCaseII-EGFP and Δ spEGCaseII-EGFP). We also constructed a fusion of Δ spEGCaseII with the peroxisomal protein catalase (Δ spEGCaseII-catalase). Catalase contains a SKL signal sequence for import into peroxisomes. After over-expression in COS-7 cells, EGCaseII activity in cell lysate was determined by using 4MU- β -D-Lac at pH 5.5 (Figure 2A). All transfected constructs generated active EGCaseII, thus indicating that appending EGCaseII as a fusion tag results in an adequately folded protein. The fusion proteins exhibited similar affinities towards β -epoxide and β -aziridine ABPs, as shown by their IC_{50} values (Figure S2A). As labeling with activity-based probes results in time-dependent irreversible modification, we additionally determined the inhibition constants of **1–4** for EGCaseII (Figure S2B). The k_i was similar for all ABPs, whereas β -epoxides **1** and **2** had a higher K_i than β -aziridines **3** and **4** (483 and 256 nM compared to 19.5 and 19.6 nM, respectively). These results are in line with the observed in vitro IC_{50} values

Figure 2. Enzymatic properties of EGCaseII fusion proteins. A) Hydrolysis of 4MU- β -D-Lac as catalyzed by 1 μ g total protein from COS-7 cells expressing EGCaseII fusion proteins. Data are average $\pm$ SD ($n = 2$). B) Lysates of COS-7 cells expressing EGCaseII constructs were labeled with 1 μ M **1–4** in 150 mM McIlvaine buffer (pH 5.5) for 1 h at 37 $^{\circ}$ C.

(Figure S2A) and suggest that K_i is the determining factor for their inhibitory potency in vitro (in the case of EGCasell).

In vitro labeling of EGCasell fusion proteins with ABPs

Next, the expressed EGCasell fusion proteins were labeled in vitro with 1–4 (Figure 2B). All were labeled with the ABPs. ABP-labeled EGCasell and Δ spEGCasell (predicted: 52.8 and 50.0 kDa, respectively) migrated slightly faster than ABP-labeled endogenous GBA (58–66 kDa). Labeled EGCasell-EGFP and Δ spEGCasell-EGFP (predicted: 83.1 and 80.4 kDa, respectively) migrate more slowly than endogenous ABP-labeled GBA. The EGCasell-EGFP fusion protein was present as two bands, with the lower band close to that for Δ spEGCasell-EGFP. This is possibly caused by partial cleavage of the predicted signal peptide in EGCasell. The signal peptide seems not essential for activity, and thus the use of Δ spEGCasell as tag is attractive. Labeled Δ spEGCasell-catalase migrated as predicted (110.5 kDa).

In situ detection of EGCasell fusion proteins in cells

The detection of EGCasell-tagged proteins in intact cells with fluorescent ABPs was studied next. We transfected COS-7 cells with Δ spEGCasell, Δ spEGCasell-EGFP, Δ spEGCasell-catalase, or EGFP-catalase. Transfected cells were then incubated for 2 h with 2 (red fluorescence) to visualize EGCasell fusion proteins by fluorescence microscopy (Figure 3A) and by SDS-PAGE and in-gel fluorescence imaging of the same samples (Figure 3B). To account for fluorescent 2-labeling of endogenous β -glucosidases as GBA, cells were pre-incubated in the presence or absence of 10 μ M CBE. Fluorescence microscopy revealed that mock-transfected COS-7 cells showed lysosomal GBA labeling in the absence of CBE (Figure 3A, top). Pre-incubation with CBE eliminated GBA labeling almost completely. In-gel fluorescence revealed that EGCasell fusion proteins and GBA were labeled in situ (Figure 3B). CBE pre-incubation abrogated labeling of GBA, but not of EGCasell. Cells showed weak background fluorescence, possibly partly due to non-specific membrane association of 2 (Figure 3A, right).^[7] Cytosolic aggregates were seen with Δ spEGCasell and Δ spEGCasell-EGFP (Figure 3A, second and third row). EGFP-catalase showed a punctated localization resembling typical peroxisomal structures (Figure 3A, fifth row). In the case of Δ spEGCasell-catalase, labeling was also cytosolic but less punctated than for EGFP-catalase (Figure 3A, fourth and fifth rows).

EGCasell presence was determined by detection of its enzymatic activity towards 4MU- β -D-Lac: 4MU- β -D-Lac was added to the cell medium (Figure 3C) or to lysate (Figure 3D). All intact EGCasell-transfected cells and lysates produced fluorescent 4MU when not exposed to 2, but the activity was abrogated when cells and lysates were treated with 2 (Figure 3C and D). Pre-incubation of cells with CBE did not affect the activity in lysates or cells, thus indicating that β -glucosidase did not contribute to the release of 4MU.

Convenient purification of EGCasell-tagged proteins

Lastly, we explored the ease of isolation of EGCasell-tagged proteins by using an EGCasell-binding ABP containing a terminal biotin (β -epoxide ABP 5; Figure 4A). Biotin allows specific isolation of 5-labeled EGCasell fusion protein by pull-down with streptavidin-coupled beads or columns. Lysates of cells expressing Δ spEGCasell-catalase were incubated with 10 μ M CBE then with 10 μ M 5. The 5-labeled proteins were isolated by streptavidin-based pull-down and subjected to SDS-PAGE (Figure 4B and C). All Δ spEGCasell-catalase was recovered. Activity towards 4MU- β -D-Lac was completely lost (Figure 4D).

Conclusions

In summary, *Rhodococcus* sp. M-777 endoglycoceramidase II (EGCasell) was selected as fusion-protein tag for a number of reasons. Firstly, it is irreversibly labeled with high affinity by fluorescent cyclophellitol β -epoxide and β -aziridine activity-based probes (ABPs). Secondly, EGCasell does not acquire glycans or disulfide-bridges^[10] (which might affect the targeting of the tagged protein of interest) and as a tag achieves a catalytically active conformation. Thirdly, EGCasell shows *endo*- β -glucosidase activity, thus allowing the production of fluorescent 4MU from 4MU- β -D-Lac, an activity not present in mammalian cells.^[11] Finally, EGCasell can be distinguished from the prominent endogenous β -glucosidase GBA by the inhibitor CBE, which, at 10 μ M, blocks ABP labeling of the latter enzyme only.

Upon ABP labeling, (Δ sp)EGCasell-tagged proteins were very sensitively detected by fluorescence after PAGE. We compared the tagging of protein with EGFP and EGCasell. We also fused EGFP to (Δ sp)EGCasell. We determined the lower limit of detection with fluorescence scanning of EGFP and green fluorescent ABPs 1 and 3, and red fluorescent DsRed and ABPs 2 and 4 (Figure S3). In the case of EGFP, the detection limit was 0.1 pmol, and in the case of ABPs 1 and 3 it was approximately 1.5 pmol. The difference in red fluorescent protein (DsRed) is even greater (0.1 pmol compared to 3 pmol). In other words, the sensitivity of detection of the bound ABP is on a par with, or superior to, EGFP. Another intrinsic advantage of the ABP-based detection of fusion proteins is the possibility of using (sequentially) distinct ABPs (e.g., differing in fluorophore). Such an approach is compatible with pulse-chase experiments as demonstrated earlier.^[7] We also examined EGCasell-tagged fusion proteins with a bi-functional ABP (biotin and fluorophore, ABP 5). The labeling with this ABP was swift and allowed streptavidin pull-down of the fusion protein.

We also put the approach to the test by generating a fusion protein of the peroxisomal enzyme catalase with EGCasell. Following expression, we labeled intact COS-7 cells with cell-permeable 2 and studied the subcellular distribution of the label by fluorescence microscopy. We observed fluorescently stained vesicles in the cytosol consistent with the location of peroxisomes. Of note, more punctated labeling was observed for EGFP-catalase. This difference might be attributable to the intrinsic differences in the fluorophores (BODIPY versus protein). Further improvement of ABPs for super-resolution microscopy

Figure 3. Detection of EGCasell fusion proteins in situ. A) Fluorescence micrographs of transfected COS-7 cells treated for 18 h in the absence or presence of 10 μM CBE and subsequently incubated with 100 nM **2** (red fluorescence). From top to bottom: mock-transfected COS-7 cells, EGCasell, $\Delta\text{spEGCasell}$, EGCasell-EGFP, $\Delta\text{spEGCasell}$ -EGFP, $\Delta\text{spEGCasell}$ -catalase, and EGFP-catalase. Fluorescence: BODIPY (red), EGFP (green), and DAPI (blue); scale bars: 25 μm . B) SDS-PAGE analysis of ABP-labeled proteins in COS-7 cells transfected with EGCasell constructs. C) Hydrolysis rate of 4MU- β -D-Lac measured in the medium of living COS-7 cells, after treatment with/without 10 μM CBE followed by treatment with/without 100 nM **2**. Data are average $\pm$ SD ($n = 3$). D) Hydrolysis rate present within lysates of aforementioned COS-7 cells. Data are average of $n = 3$, $\pm$ SD.

and for correlative light and electron microscopy is possible. Ideally, the unbound ABP should be non-fluorescent (negligible background), whereas bound ABP can be made to fluoresce, specifically and on demand. For example, cyclophellitol could be fitted with a norbornene at C8 (cyclophellitol numbering) to generate a mechanism-based inhibitor. Cells could then be pulse-labeled with this non-fluorescent probe; after treatment with fluorogenic tetrazine, a fluorescent structure would form (inverse-electron demand Diels–Alder). Fluorogenic dyes could include next-generation moieties such as those developed by Urano and colleagues.^[12]

In conclusion, EGCasell appears an attractive tag for fusion proteins, and allows detection, both in vitro and in intact cells, following labeling with a cell-permeable fluorescent ABP. EGCasell-tagged fusion proteins can be quantified based on their ability to hydrolyze fluorogenic 4MU- β -D-Lac. It might be a useful addition to existing methodologies based on fusion with XFPs and other types of tags.

Experimental Section

General methods: ABPs were synthesized as described earlier.^[7] Recombinant *Rhodococcus* sp. M-777 endoglycoceramidase (EGCasell) was obtained from Takara Bio Inc. (Shiga, Japan). Recombinant GBA (Imiglucerase) was obtained from Genzyme (Cambridge, MA). COS-7 cells were cultured in DMEM/F12 (Thermo Fisher) supplemented with FBS (10%, v/v). EGFP (1 $\mu\text{g } \mu\text{L}^{-1}$ ($\sim 37.175 \mu\text{M}$); BioVision, Milpitas, CA) and DsRed (1 $\mu\text{g } \mu\text{L}^{-1}$ ($\sim 36.232 \mu\text{M}$); BioVision) were sequentially diluted in PBS (from 100 pmol) in a clear-bottom black 96-well plate (Greiner Bio). Fluorescence was scanned in a Typhoon Variable Mode Imager (GE Healthcare Life Sciences): $\lambda_{\text{ex}} = 488 \text{ nm}$, $\lambda_{\text{em}} = 520 \text{ nm}$, 40 nm bandpass filter for EGFP, **1**, and **3**; $\lambda_{\text{ex}} = 532 \text{ nm}$, $\lambda_{\text{em}} = 610 \text{ nm}$, 30 nm bandpass filter for DsRed, **2**, and **4**.

ABP labeling of EGCasell and GBA: Recombinant *H. sapiens* acid β -glucosidase GBA in Mcllvaine buffer, (citric acid/ Na_2HPO_4 (150 mM, pH 5.2), supplemented with sodium taurocholate (0.2%, w/v), Triton X-100 (0.1%, v/v), and *Rhodococcus* sp. M-777 endoglycoceramidase EGCasell (1 μg total protein in 5 μM) in Mcllvaine

Figure 4. Labeling of EGCasell with ABP containing an additional biotin functionality. A) Cyclophellitol β -epoxide **5** containing red fluorescent BODIPY and terminal biotin. B) EGCasell-fusion proteins labeled in COS-7 cell lysate with $1 \mu\text{M}$ **5**. C) COS-7 $\Delta\text{spEGCasell}$ -catalase cell lysate pre-incubated with CBE prior to labeling with excess **5** and subsequent pull-down by streptavidin. Isolated **5**-labeled proteins (pull-down) compared to unlabeled proteins (supernatant). In a separate experiment, CBE was substituted by **3** (supernatant, right). Inter-gel comparisons was facilitated by imiglucerase (asterisks, loading control). D) Hydrolysis rate of 4MU- β -D-Lac measured in the $\Delta\text{spEGCasell}$ -catalase input (4MU- β -D-Lac activity present in the sample prior to pulldown), pulldown fraction, and supernatant labeled with or without **5**. Data are average $\pm$ SD ($n=2$).

buffer (150 mM, pH 5.5), was labeled with ABP **1** or **3** (0.0001–1000 nM) for 1 h at 37°C . The samples were then denatured with $5\times$ sample denaturation buffer (Tris-HCl (50% v/v, 1 M, pH 6.8), glycerol (50% v/v), DTT (10%, w/v), SDS (10%, w/v), Bromophenol Blue (0.01% w/v)), boiled for 4 min, and separated by electrophoresis on a 7.5% SDS-PAGE gel at 90 V.^[7] Wet slab gels were scanned with the Typhoon Variable Mode Imager ($\lambda_{\text{ex}}=488 \text{ nm}$, $\lambda_{\text{em}}=520 \text{ nm}$ (40 nm bandpass) for **1** and **3**; $\lambda_{\text{ex}}=532 \text{ nm}$, $\lambda_{\text{em}}=610 \text{ nm}$ (30 nm bandpass) for **2**, **4**, and **5**).

In vitro ABP labeling of EGCasell and GBA: Recombinant GBA and EGCasell (1 μg total protein) were incubated in Mcllvaine buffer (150 mM, pH 5.0 or 6.0) with **1** or **3** (100 nM) in the absence or presence of detergents (0.2% (w/v) sodium taurocholate, 0.1% (v/v) Triton X-100), for 1 h at 37°C . Samples were denatured, and ABP-tagged proteins were visualized by fluorescence scanning of the SDS-PAGE slab-gels.

Substrate specificity: The activity of GBA and EGCasell was assayed at 37°C by incubating with saturating concentrations of artificial substrate containing 4-methylumbelliferyl (4MU) as aglycone conjugated to either a β -D-linked xylose, galactose, glucose, mannose, glucuronic acid, *N*-acetyl- β -D-glucosyl- or galactosamine, lactose or α -L-fucose for 0–1 h. After stopping the enzymatic activity by

adding excess NaOH-glycine (pH 10.3), liberated 4MU was quantified by measuring its fluorescence with a LS55 fluorimeter (Perkin-Elmer) using $\lambda_{\text{ex}} 366 \text{ nm}$ and $\lambda_{\text{em}} 445 \text{ nm}$.

Determination of inhibition constants: The kinetic parameters of inhibition were determined by incubating EGCasell with **1** or **2** (0–1000 nM) or **3** or **4** (0–100 nM) in Mcllvaine buffer (150 mM, pH 5.0) for 0–60 min at 37°C . The reaction was stopped at intervals of 5 min with excess NaOH/glycine (pH 10.6). Inhibitory constants k_i and K_i were determined by calculating the K_{obs} per ABP concentration and curve-fitting the data to a one-phase exponential association function (GraphPad Prism 5.0).^[7,13]

In vitro ABP labeling competition: Recombinant EGCasell and GBA were mixed (1:1) and pre-incubated with CBE (0–10 mM) in Mcllvaine buffer (150 mM, pH 5.5 or 5.2) supplemented with sodium taurocholate (0.2%, w/v) and Triton X-100 (0.1%, v/v) for 30 min at 37°C . EGCasell and GBA were labeled by incubation with **1** (1 μM) in corresponding Mcllvaine buffer (150 mM) for 1 h at 37°C , then denatured and ABP-labeled proteins were separated by SDS-PAGE and visualized by fluorescence scanning of the SDS-PAGE slab-gels.

Molecular docking: 4MU- β -D-glucose, 4MU- β -D-lactose, CBE, and **2–5** ligand structures were drawn, then the lowest energy conformers were calculated and saved as.pdb structure files in MarvinDraw (ChemAxon, Budapest, Hungary). Rotatable bonds were defined and structure were saved in a ligand.pdbqt file (AutoDockTools, <http://mgltools.scripps.edu/downloads>). Crystal structures of EGCasell (PDB IDs: 2OSX and 2OYK) and GBA (PDB IDs: 2V3E and 2VT0) were isolated by removal of water molecules and other molecules in the active site (yielding enzymes mimicking a free state, in PyMol); aligned, polar-only residues were protonated and finally saved as a receptor.pdbqt file (AutoDockTools). Flexible docking experiments were run in Autodock Vina (<http://vina.scripps.edu/>) with standard settings, with exhaustiveness adjusted to 12,^[14] and search space boundaries encompassing the complete protein (EGCasell or GBA).

Cloning procedures: The nucleotide sequence for EGCasell was determined from the amino-acid sequence^[11] codon-optimized for mammalian expression, synthesized by GeneArt (Thermo Fisher), and cloned into pcDNA3.1, either without (Δsp) or with the predicted *Rhodococcus* sp. M-777 signal peptide. Fusion proteins were produced by in-frame cloning of coding sequences of EGFP (yielding EGCasell-EGFP, $\Delta\text{spEGCasell}$ -EGFP), or human catalase (GenBank accession NM_001752.3; yielding $\Delta\text{spEGCasell}$ -catalase) after the linker sequence NSAVDGTAGMADSRDPASDQMHWKEQRAAQK (also used to fuse EGFP-catalase).

Activity of EGCasell fusion proteins: Confluent COS-7 cells were transfected with EGCasell, $\Delta\text{spEGCasell}$, EGCasell-EGFP, $\Delta\text{spEGCasell}$ -EGFP, $\Delta\text{spEGCasell}$ -catalase, or empty pcDNA3.1 plasmid (Mock), by using FuGENE (Promega) and harvested after 72 h by scraping in potassium phosphate buffer (25 mM, pH 6.5) supplemented with Triton X-100 (0.1%, v/v) and protease inhibitor cocktail (Roche). After determination of the total protein concentration with a Pierce BCA kit (Thermo Fisher), total protein (10 μg) was incubated with 4MU- β -D-Lac (20 mM, Sigma-Aldrich) in Mcllvaine buffer (150 mM, pH 5.5) for 0–60 min. At time points, the reaction was stopped by adding excess NaOH/glycine (pH 10.3), liberated 4MU was quantified by measuring its fluorescence with a LS55 fluorimeter (PerkinElmer) by using $\lambda_{\text{ex}} 366 \text{ nm}$ and $\lambda_{\text{em}} 445 \text{ nm}$. The effect of pH was measured by pre-incubating lysate (10 μg total protein) for 30 min at 37°C in Mcllvaine buffer (150 mM, pH 3 to 8), then 4MU- β -D-Lac in corresponding Mcllvaine buffer (150 mM)

was added. After 30 min, the reaction was stopped, and liberated 4MU was quantified. The inhibitory potencies of CBE and **1** to **4** were determined by pre-incubating total lysate (10 µg protein) with a range of inhibitor concentrations at pH 5.5. After 30 min at 37 °C, 4MU-β-D-Lac (20 mM) was added to determine residual EGCasell activity. Alternatively, COS-7 lysate (100 µg) was labeled by incubating with **1** to **4** (1 µM) in Mcllvaine buffer (150 mM, pH 5.5) for 1 h at 37 °C. After sample denaturation, ABP-labeled proteins were visualized by fluorescence scanning of the SDS-PAGE slab-gels.

Pull down of ABP 5-labeled EGCasell-tagged proteins: Lysates of COS-7 cells (100 µg) over-expressing ΔspEGCasell-catalase were pre-incubated with CBE (10 µM) or **3** (100 nM) in Mcllvaine buffer (150 mM, pH 5.5) for 1 h at 37 °C. Labeling was performed with **5** (1 µM in Mcllvaine buffer (150 mM, pH 5.5) for 2 h at 37 °C. Proteins labeled with **5** were pulled down with streptavidin-coupled Dynabeads (Thermo Fisher) for 1 h at room temperature. Proteins were released by incubation in SDS (5% (w/v) for 5 min at 100 °C. Laemmli buffer was added, and proteins were separated by SDS-PAGE and visualized by fluorescence scanning of the slab-gel.

Fluorescence microscopy: Zeiss Axiovert 200 microscope with UV-DAPI filter BP365/12 with beam-splitters FT395 and LP397, TRITC filter BP510–560 with beam-splitters FT580 and LP590, and FITC filter BP450–490 with beam-splitter FT510 and LP515.

Acknowledgements

We thank Dirk Westenberg for technical assistance. This work was supported by The Netherlands Organization for Scientific Research (NWO-CW, Top Grant, to J.M.F.G. A. and H.S.O.) and the European Research Council (ERC AdG, to J.M.F.G. A. and H.S.O.).

Keywords: activity-based probes • EGCasell • fluorescent probes • in-gel imaging • microscopy • protein expression

- [1] a) O. Shimomura, F. H. Johnson, Y. Saiga, *J. Cell. Comp. Physiol.* **1962**, *59*, 223–239; b) R. Y. Tsien, *Annu. Rev. Biochem.* **1998**, *67*, 509–544.
- [2] N. C. Shaner, M. Z. Lin, M. R. McKeown, P. A. Steinbach, K. L. Hazelwood, M. W. Davidson, R. Y. Tsien, *Nat. Methods* **2005**, *2*, 905–909.
- [3] a) R. Dixit, R. Cyr, *Plant J.* **2003**, *36*, 280–290; b) L. Greenbaum, C. Rothmann, R. Lavie, Z. Malik, *Biol. Chem.* **2000**, *381*, 1251–1258; c) A. F. Bell, D. Stoner-Ma, R. M. Wachter, P. J. Tonge, *J. Am. Chem. Soc.* **2003**, *125*, 6919–6926; d) A. Jiménez-Banzo, S. Nonell, J. Hofkens, C. Flors, *Biophys. J.* **2008**, *94*, 168–172.
- [4] a) Q. Yan, M. P. Bruchez, *Cell Tissue Res.* **2015**, *360*, 179–194; b) C. Jing, V. W. Cornish, *Acc. Chem. Res.* **2011**, *44*, 784–792; c) N. Johnsson, N. George, K. Johnsson, *ChemBioChem* **2005**, *6*, 47–52; d) L. Wang, P. G. Schultz, *Angew. Chem. Int. Ed.* **2004**, *43*, 34–56; *Angew. Chem.* **2004**, *117*, 34–68; e) L. Davis, J. W. Chin, *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 168–182; f) L. Xue, I. A. Karpenko, J. Hiblot, K. Johnsson, *Nat. Chem. Biol.* **2015**, *11*, 917–923.
- [5] a) Y. Zhang, M.-k. So, A. M. Loening, H. Yao, S. S. Gambhir, J. Rao, *Angew. Chem. Int. Ed.* **2006**, *45*, 4936–4940; *Angew. Chem.* **2006**, *118*, 5058–5062; b) L. P. Encell, R. Friedman Ohana, K. Zimmerman, P. Otto, G. Vidugiris, M. G. Wood, G. V. Los, M. G. McDougall, C. Zimprich, N. Karassina, R. D. Learish, R. Hurst, J. Hartnett, S. Wheeler, P. Stecha, J. English, K. Zhao, J. Mendez, H. A. Benink, et al., *Curr. Chem. Genomics* **2012**, *6*, 55–71.
- [6] a) A. C. Benniston, G. Copley, *Phys. Chem. Chem. Phys.* **2009**, *11*, 4124–4131; b) N. Boens, V. Leen, W. Dehaen, *Chem. Soc. Rev.* **2012**, *41*, 1130–1172.
- [7] a) 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; b) W. W. Kallemeijn, K.-Y. Li, M. D. Witte, A. R. A. Marques, J. Aten, S. Scheij, J. 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. Codee, G. A. van der Marel, J. M. F. G. Aerts, H. S. Overkleeft, *Angew. Chem. Int. Ed.* **2012**, *51*, 12529–12533; *Angew. Chem.* **2012**, *124*, 12697–12701; c) W. W. Kallemeijn, J. M. F. G. Aerts, *J. Biol. Chem.* **2014**, *289*, 35351–35362; d) 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. Codee, J. M. F. G. Aerts, G. A. van der Marel, H. S. Overkleeft, *Eur. J. Org. Chem.* **2014**, 6030–6043; e) M. D. Witte, M. T. C. Walvoort, K.-Y. Li, W. W. Kallemeijn, W. E. Donker-Koopman, R. G. Boot, J. M. F. G. Aerts, J. D. C. Codee, G. A. van der Marel, H. S. Overkleeft, *ChemBioChem* **2011**, *12*, 1263–1269.
- [8] a) M. E. C. Caines, M. D. Vaughan, C. A. Tarling, S. M. Hancock, R. A. J. Warren, S. G. Withers, N. C. J. Strynadka, *J. Biol. Chem.* **2007**, *282*, 14300–14308; b) K. Sakaguchi, N. Okino, H. Izu, M. Ito, *Biochem. Biophys. Res. Commun.* **1999**, *260*, 89–93; c) Y.-T. Li, C.-W. Chou, S.-C. Li, U. Kobayashi, Y.-h. Ishibashi, M. Ito, *Glycoconjugate J.* **2009**, *26*, 929–933.
- [9] J. R. Rich, A. M. Cunningham, M. Gilbert, S. G. Withers, *Chem. Commun.* **2011**, *47*, 10806–10808.
- [10] a) M. Ito, Y. Ikegami, T. Yamagata, *J. Biol. Chem.* **1991**, *266*, 7919–7926; b) M. Ito, Y. Ikegami, A. Omori, T. Yamagata, *J. Biochem.* **1991**, *110*, 328–332; c) M. Ito, Y. Ikegami, T. Tadashi, T. Yamagata, *Eur. J. Biochem* **1993**, *218*, 637–643; d) M. Ito, Y. Ikegami, T. Yamagata, *Eur. J. Biochem* **1993**, *218*, 645–649; e) M. Ito, T. Yamagata, *J. Biol. Chem.* **1989**, *264*, 9510–9519.
- [11] H. Izu, Y. Isumi, Y. Kurome, M. Sano, A. Kondo, I. Kato, M. Ito, *J. Biol. Chem.* **1997**, *272*, 19846–19850.
- [12] a) Y. Urano, *Curr. Opin. Chem. Biol.* **2012**, *16*, 602–608; b) G. T. Dempsey, J. C. Vaughan, K. H. Chen, M. Bates, X. Zhuang, *Nat. Methods* **2011**, *8*, 1027–1036; c) R. S. Erdmann, H. Takakura, A. D. Thompson, F. Rivera-Molina, E. S. Allgeyer, J. Brewersdorf, D. K. Toomre, A. Schepartz, *Angew. Chem. Int. Ed.* **2014**, *53*, 10242–10246; *Angew. Chem.* **2014**, *126*, 10407–10412; d) G. Lukinavičius, K. Umezawa, N. Olivier, A. Hongigmann, G. Yang, T. Plass, V. Mueller, L. Reymond, I. R. Corrêa, Jr., Z. G. Luo, C. Schultz, E. A. Lemke, P. Heppenstall, C. Eggeling, S. Manley, K. Johnsson, *Nat. Chem.* **2013**, *5*, 132–139; e) J. C. T. Carlson, L. G. Meimetis, S. A. Hilderbrand, R. Weissleder, *Angew. Chem. Int. Ed.* **2013**, *52*, 6917–6920; *Angew. Chem.* **2013**, *125*, 7055–7058.
- [13] A. G. Marangoni, *Enzyme Kinetics: A Modern Approach*, Wiley, Hoboken, **2003**, pp. 70–78.
- [14] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455–461.

Manuscript received: June 1, 2016

Accepted article published: July 7, 2016

Final article published: August 4, 2016