



Published in final edited form as:

ACS Sens. 2023 August 25; 8(8): 2996–3003. doi:10.1021/acssensors.3c00409.

Genetically Encoded Boronolactin as a Specific Red Fluorescent UDP-GlcNAc Biosensor

Jing Zhang^{1,2}, Zefan Li^{1,2}, Yu Pang^{2,3}, Yichong Fan^{1,2,†}, Hui-wang Ai^{1,2,3,4,*}

¹Department of Molecular Physiology and Biological Physics, University of Virginia School of Medicine, 1340 Jefferson Park Ave, Charlottesville, Virginia, 22908, USA

²Center for Membrane and Cell Physiology, University of Virginia School of Medicine, Charlottesville, Virginia, 22908, USA

³Department of Chemistry, University of Virginia, Charlottesville, Virginia, 22904, USA

⁴UVA Comprehensive Cancer Center, University of Virginia, Charlottesville, Virginia, 22903, USA

Abstract

There is great interest in developing boronolactins, which are synthetic lectin mimics containing a boronic acid functional group for reversible recognition of diol-containing molecules, such as glycans and ribonucleotides. However, it remains a significant challenge to gain specificity. Here, we present a genetically encoded boronolactin, which is a hybrid protein consisting of a noncanonical amino acid (ncAA) *p*-boronophenylalanine (*p*BoF), natural-lectin-derived peptide sequences, and a circularly permuted red fluorescent protein (cpRFP). The genetic encodability permitted a straightforward protein engineering process to derive a red fluorescent biosensor that can specifically bind uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc), an important nucleotide sugar involved in metabolic sensing and cell signaling. We further characterized the resultant boronic acid- and peptide-assisted UDP-GlcNAc sensor (bapaUGAc) both *in vitro* and in live mammalian cells. Because UDP-GlcNAc in the endoplasmic reticulum (ER) and Golgi apparatus plays essential roles in glycosylating biomolecules in the secretory pathway, we genetically expressed bapaUGAc in the ER and Golgi and validated the sensor for its responses to metabolic disruption and pharmacological inhibition. In addition, we combined bapaUGAc with UGAcS, a recently reported green fluorescent UDP-GlcNAc sensor based on an alternative sensing mechanism, to monitor UDP-GlcNAc level changes in the ER and cytosol simultaneously. We expect our work to facilitate the future development of specific boronolactins

*Corresponding author. huiwang.ai@virginia.edu.

†Present Address: AbbVie Inc., Irvine, CA 92612

Author Contributions

J.Z., Y.F., and Y.P. engineered the sensors. J.Z. performed characterization experiments with the help of Z.L.. J.Z. analyzed data, prepared figures, and wrote the manuscript. H.-w.A. supervised research, revised figures, and wrote the manuscript.

ASSOCIATED CONTENT

Supporting Information

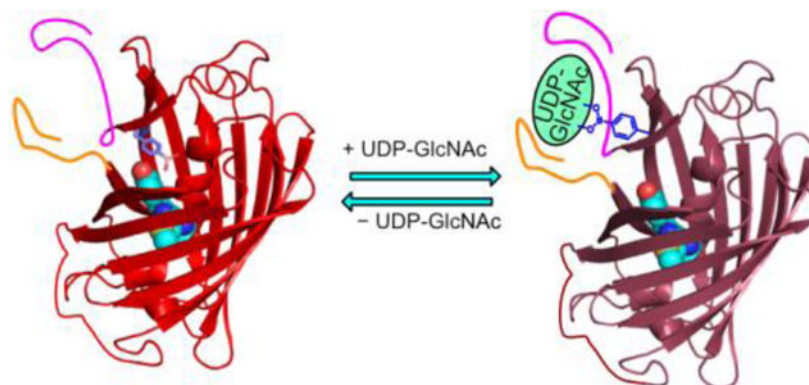
The Supporting Information is available free of charge on the ACS Publications website.

Illustration of the protein engineering process, sequence alignment, additional sensor characterization, structural illustration of key residues, and oligonucleotide sequences (PDF)

The authors declare no competing financial interest. The plasmids pBAD-bapaUGAc (#206173), pMAH-ER-bapaUGAc (#206174), and pMAH-Golgi-bapaUGAc (#206175) have been deposited to Addgene.

for carbohydrates. In addition, this newly developed genetically encoded bapaUGAc sensor will be a valuable tool for studying UDP-GlcNAc and glycobiology.

Graphical Abstract



Carbohydrates constitute one of the four major classes of biomolecules, and they play a crucial role in many physiological and pathological processes. As such, specific carbohydrate binders are highly sought after for their potential applications in research, diagnostics, therapy, and drug delivery.^{1–3} In recent decades, researchers have focused on developing carbohydrate-binding antibodies, lectins, aptamers, and synthetic lectin mimics.^{4–6} Despite these efforts, the development of specific binders for biologically significant carbohydrates remains a significant challenge in the field.

Boronic acids can undergo reversible esterification reactions with molecules that contain cis-diol functional groups (Fig. 1a).⁷ This property allows boronic acids to recognize and bind to carbohydrates, leading to the development of synthetic lectin mimics known as “boronolactins”. However, these interactions are often promiscuous and lack specificity. Researchers have attempted to improve the specificity of boronic acid-carbohydrate interactions by using multiple coordination sites, steric effects, and molecular geometry, but these approaches have only been successful for a few examples.^{6,8–10} A generalizable approach is still lacking.

Genetic code expansion (GCE) is a technology that allows for the convenient expression of proteins containing noncanonical amino acids (ncAAs) in live cells and organisms.^{11,12} GCE uses engineered tRNA and aminoacyl-tRNA synthetase to introduce ncAAs into proteins in a site-specific manner. Previous studies have enabled the genetic encoding of *p*-boronophenylalanine (*p*BoF, Fig. 1b) in *E. coli* and mammalian cells.^{13,14} In addition, *p*BoF has been introduced into a phage display library, leading to the identification of *p*BoF-containing single-chain variable antibody fragments (scFvs) that bind to glucosamine (GlcN).¹⁵ However, the specificity of these *p*BoF-containing scFvs to glucosamine, compared to other structurally similar carbohydrates, was not sufficiently addressed in this previous study.¹⁵

Fluorescent protein (FP)-based biosensors have become popular tools for biological and medical research because they can provide real-time and specific detection of a wide range

of biomolecules and bioactivities.¹⁶ Introducing ncAAs into FPs is an attractive strategy for developing new FP-based biosensors.^{17,18} Previously, we and others incorporated *p*BoF into FPs and developed reaction-based biosensors for hydrogen peroxide and peroxynitrite.^{14,19–21} Despite the progress, *p*BoF-modified FPs have not yet been used to sense glycans.

Based on the knowledge and results described above, we hypothesize that *p*BoF-modified FPs may be engineered into specific biosensors for carbohydrates. Because GCE provides a practical method for the genetic encoding of *p*BoF-modified FPs, specificity may be achieved via genetic protein engineering. To test our hypothesis, we selected uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc, Fig. 1c), an important cell metabolite, as our target.

UDP-GlcNAc is considered an integrator of nutritional and metabolic signals.²² In mammalian cells, UDP-GlcNAc can be synthesized *de novo* from glucose, glutamine, acetyl-coenzyme A (Ac-CoA), adenosine triphosphate (ATP), and uridine triphosphate (UTP) via the hexosamine biosynthetic pathway (HBP).²³ Subsequently, UDP-GlcNAc, along with other nucleotide sugars, is used as glycosyl donors in the process of glycosylation that covalently attach carbohydrate moieties to biomolecules. In the cytosol and nucleus, UDP-GlcNAc is used by O-GlcNAc transferase (OGT) to add O-GlcNAc modifications to proteins, resulting in reversible and regulated cell signaling.²⁴ In the endoplasmic reticulum (ER) and Golgi apparatus, UDP-GlcNAc and other nucleotide sugars are used by glycosyltransferases to form complex glycosylation products, which may be further displayed on the cell surface or secreted to the extracellular space.²⁵

Because of the importance of UDP-GlcNAc in metabolism and signaling, we have devoted efforts to developing fluorescent UDP-GlcNAc biosensors. We recently reported a green fluorescent biosensor, namely UGAcS, which is an insertion of a circularly permuted green FP (cpGFP) into an inactive mutant of an *E. coli* UDP-GlcNAc transferase.²⁶ UGAcS has the right affinity for monitoring the levels of UDP-GlcNAc in the cytosol. In this manuscript, we describe a parallel effort in our lab, which explores a *p*BoF-modified circularly permuted red FP (cpRFP) for the specific detection of UDP-GlcNAc (Fig. 2a). The work has resulted in a novel genetically encoded red fluorescent boronolactin, denoted as bapaUGAc, with a suitable affinity for detecting the levels of UDP-GlcNAc in the ER and Golgi.

MATERIALS AND METHODS

Key reagents and methods.

This study used the racemic form of the *p*-borono-DL-phenylalanine (*p*BoF) amino acid purchased from Syntonix (Syntonix, Waltham, USA). Uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc), uridine diphosphate *N*-acetylgalactosamine (UDP-GalNAc), uridine diphosphate glucose (UDP-Glc), uridine diphosphate galactose (UDP-Gal), cytidine monophosphate *N*-acetylneuraminic acid (CMP-Neu5Ac), GDP-mannose, UTP, UMP, *N*-acetylglucosamine 1-phosphate (GlcNAc-1-P), and D-glucosamine hydrochloride (GlcN) were purchased from MilliporeSigma (St. Louis, MO, USA). UDP, 2-deoxy- D-glucose (2-DG), and L-arabinose were purchased from Cayman Chemical

(Ann Arbor, Michigan, USA). Synthetic DNA oligos were ordered from Integrated DNA Technologies (Coralville, Iowa, USA). DNA sequences were confirmed with Sanger sequencing performed by Eurofins Genomics (Louisville, KY, USA). pEvol-pBoF was a gift from Dr. Peter Schultz (Scripps Research).²⁷ pMAH-POLY and pcDNA-UGAcS were from our previous studies.^{14,26}

Engineering of bapaUGAc.

The cpmApple gene fragment was amplified from R-GECO1 using oligos pBAD-modified-cpmApple-F and pBAD-modified-cpmApple-R (see Table S1 for oligo sequences).^{28,29} The PCR product was digested with Xho I and Hind III and inserted into a pre-digested, modified pBAD plasmid, resulting in pBAD-cpmApple. To further improve the brightness of cpmApple, error-prone PCRs (EP-PCRs) were performed using oligos pBAD-F and pBAD-R, according to a previously described procedure.³⁰ Next, the PCR product was digested with Xho I and Hind III and inserted into the abovementioned, modified pBAD plasmid. The DNA library was used to transform *E. coli* DH10B electrocompetent cells. Cells were plated on 2×YT agar plates supplemented with 100 µg/mL ampicillin and 0.02% (w/v) l-arabinose. Next day, colonies on the agar plates were imaged using a customized imaging system equipped with a Dolan-Jenner Mi-LED Fiber Optic light source, appropriate excitation and emission filters in Thorlabs motorized filter wheels, and a QSI 628 CCD camera. Colonies with higher fluorescence were selected and cultured in 1 mL 2×YT supplemented with 100 µg/mL ampicillin in the wells of a 96-well deep-well plate shaken at 250 r.p.m. and 37 °C overnight. Next, 1 mL of fresh 2xYT media supplemented with 100 µg/mL ampicillin and 0.04% (w/v) l-arabinose was added to the overnight starter culture, and the new mixtures were incubated for additional 48 hr at 250 r.p.m. and 30 °C. Centrifugation at 3800 ×g was used to pellet cells, which were further lysed with 300 µL of a bacterial lysis buffer (5 mg/mL octyl glucoside, 0.1 mg/mL chicken egg lysozyme, and 0.2 U/mL Benzonase in Tris-HCl, pH 8) by shaking at 200 r.p.m. on ice for 1 hr. The fluorescence of the cell lysates was measured on a BioTek Synergy Mx microplate reader with the excitation wavelength at 575 nm and the emission wavelength at 600 nm. Plasmid DNAs were extracted from colonies showing high fluorescence and sequences were determined via Sanger sequencing. Three rounds of directed evolution were carried out to derive ecpApple.

Next, site-directed mutagenesis was used to introduce the amber (TAG) codon to residues 14 and 30 of ecpApple, respectively. Briefly, oligos pBAD-F and Ser14TAG-R, or Ser14TAG-F and pBAD-R, were used to amplify fragments of ecpApple. These fragments were assembled with pBAD-F and pBAD-R in an overlap PCR. The PCR product was digested with Xho I and Hind III and inserted into the pBAD plasmid, resulting in pBAD-ecpApple-S14TAG. Similarly, oligos pBAD-F, Lys30TAG-R, Lys30TAG-F and pBAD-R were used to generate pBAD-ecpApple-K30TAG. The resultant pBAD plasmids, along with pEvol-pBoF,²⁷ were used to co-transform DH10B electrocompetent cells. Protein purification and initial substrate specificity test were carried out, according to the procedures presented in the next section.

pBAD-ecpApple-K30TAG was chosen for further engineering. To derive bapaUGAc0.1, oligos pBAD-UEA-II-F and pBAD-R were used to amplify the ecpApple-K30TAG gene

fragment and add an N-terminal 13-amino-acid peptide sequence (DSYFGKTYNPWDP) derived from the *Ulex europaeus* agglutinin II (UEA-II) lectin. The PCR product was inserted into pBAD as described above. Next, oligos pBAD-F and 5NNK-R were used to amplify bapaUGAc0.1 and add five randomized residues to the C-terminus. The library was screened using a procedure similar to the engineering of ecpApple, except that both high brightness and UDP-GlcNAc responsiveness were used as the selection criteria. This procedure led to bapaUGAc0.2.

Furthermore, three additional peptide sequences derived from natural GlcNAc-binding lectins were added to the C-terminus of bapaUGAc0.2 via two additional randomized residues. Briefly, to add the second copy of the 13-amino-acid peptide sequence from UEA-II, oligos pBAD-F and cpmApple-UEA-II-R were used to amplify a gene fragment from pBAD-bapaUGAc0.2, and oligos cpmApple-UEA-II-F and UEA-II-R were used in a PCR without additional templates to generate a short fragment; next, the two fragments were assembled using oligos pBAD-F and UEA-II-R via an overlap PCR. The assembled fragment was digested with Xho I and Hind III and inserted into the pBAD plasmid. Similar procedures were used to add peptide sequences derived from the *Griffonia simplicifolia* GS-II lectin, and these sequences were previously reported to interact with GlcNAc.^{31,32} GS-II SLS refers to a 27-amino acid sequence (IVFCEFDLYKNGIDPSYTPHLGINVNQ), while GS-II loop refers to a 14-amino-acid sequence (DLYKNGIDPSYTPH). Screening of these three libraries for high brightness and UDP-GlcNAc responsiveness led to bapaUGAc0.3, which is essentially bapaUGAc0.2 linked to the N-terminus of GS-II loop via a two-amino-acid linker.

From bapaUGAc0.3, three rounds of directed evolution based on EP-PCRs were performed by following the procedure described above. The best-selected clone was bapaUGAc0.4. To further optimize the linker between the N-terminal UEA-II-derived peptide and cpmApple, oligos pBAD-UEA-II-3NNKs-F and pBAD-R were used to amplify bapaUGAc0.4 and randomize the first three residues of this 9-amino-acid linker. The PCR product was digested by Xho I and Hind III and then inserted into the predigested pBAD plasmid. Screening of the library identified an improved mutant, which was used for further engineering. Next, similar procedures were used to randomize the middle and the last three residues of this 9-amino-acid linker. Screening of the libraries led to the final mutant, namely bapaUGAc.

Protein purification and *in vitro* characterization.

To express the bapaUGAc protein, pBAD-bapaUGAc and pEvol-pBoF were used to co-transform DH10B electrocompetent cells. A single colony was used to inoculate 3.0 mL of 2×YT media supplemented with 100 µg/mL ampicillin and 50 µg/mL chloramphenicol at 220 r.p.m. and 37 °C overnight. The saturated starter was added into 300 mL of Terrific Broth medium (TB) supplemented with the same concentrations of antibiotics mentioned above. When the optical density (OD) at 600 nm reached 1.0, 0.02% (w/v) L-arabinose and 2 mM pBOF were added. The culture was incubated at 220 r.p.m. and 30 °C for additional 48 hr. Cells were next pelleted, resuspended in 1× phosphate-buffered saline (PBS, pH 7.4), and lysed on ice by sonication. The lysate was clarified with centrifugation at 16000 ×g and 4 °C for 30 min. The supernatant was subjected to Ni-NTA agarose bead (Pierce, Rockford, IL)

affinity purification. Proteins in ~ 4 mL of 300 mM imidazole-containing buffer were next injected into a HiLoad 16/600 Superdex 200 pg size-exclusion column on an AKTA protein purification system (Cytiva). 1× PBS was used for elution, and the eluted proteins were kept at 4 °C. Protein concentration was determined using the alkali denaturation method by assuming the extinction coefficient of the denatured chromophore to be $44,000 \text{ M}^{-1}\text{cm}^{-1}$.³³

Fluorescence spectra of the purified protein (1 μM) in the PBS buffer supplemented with 1 mM CaCl_2 and 0.5 mM MgCl_2 (PBS++) in the presence and absence of 5 mM UDP-GlcNAc were measured using a monochromator-based BioTek Synergy Mx Microplate Reader. To record the excitation spectra, the emission wavelength was set at 640 nm while the excitation was scanned from 500 nm to 620 nm. To record the emission spectra, the excitation wavelength was set at 530 nm while the emission was scanned from 550 to 700 nm. Absorbance spectra were measured similarly except that 5 μM protein was used. Next, to determine the dose-dependent response of bapaUGAc to UDP-GlcNAc, 1 μM protein was incubated with PBS++ alone or PBS++ containing UDP-GlcNAc in at final concentrations of 0.5, 1, 2, 3, 4, and 5 mM, respectively. In addition, to examine the indicator's specificity, the freshly prepared protein (1 μM) was incubated with chosen chemicals in PBS++ at a final concentration of 5 mM. Endpoint fluorescence intensity measurement was carried out 5 min after mixing all reagents with excitation and emission set at 570 and 600 nm, respectively. The pH titration experiment was performed by mixing 10 μL of 5 μM proteins with 90 μL of pre-made buffers containing 200 mM citric acid and 200 mM phosphate with pH values spanning from 3 to 10. For the experiment in the presence of UDP-GlcNAc, the sensor (5 μM) was pre-complexed with UDP-GlcNAc (100 mM) in PBS++ before being diluted with the pH buffers. The fluorescence of each sample was determined using the plate reader and plotted as a function of pH. The Hill equation was used to fit the data to derive the apparent pK_a .

Plate-reader-based assays were used to evaluate the association and dissociation kinetics of between bapaUGAc and UDP-GlcNAc. The binding assay was performed by adding 50 μL of 10 mM UDP-GlcNAc to 50 μL of 2 μM bapaUGAc protein in PBS++. The ligand addition and mixing took ~20 s, before and after which fluorescence was monitored every 10 s. To compensate for the fluorescence changes caused by sample volume changes, results from the control experiments with the addition of the same volume of PBS++ were used for data normalization. To examine the dissociation kinetics, the bapaUGAc protein (5 μM) pre-complexed with UDP-GlcNAc (5 mM) was diluted 10-fold using PBS++. A fluorescence plate reader was used to monitor the process, and the 10-fold dilution and mixing took ~20 s, before and after which fluorescence was again monitored every 10 s. To compensate for the fluorescence changes caused by sample volume changes, results from the control experiments with the dilution of the bapaUGAc protein (5 μM) alone were used for data normalization.

Construction of mammalian expression plasmids.

To generate a plasmid for bapaUGAc expression in the ER, oligos ER-bapaUGAc-F and bapaUGAc-KDEL-R were used to amplify the bapaUGAc gene from pBAD-bapaUGAc. This step also added a C-terminal ER retention sequence (KDEL). Meanwhile, pMAH-ER-

F and ER-bapaUGAc-R were used in a PCR without additional templates to generate a short fragment encoding the N-terminal ER signal peptide (MLLSVPLLLGLLGLAAAD) derived from calreticulin. The two resultant DNA fragments, along with a pMAH plasmid predigested with Hind III and Apa I, were subjected to a three-fragment Gibson assembly reaction, resulting in the pMAH-ER-bapaUGAc plasmid. A similar procedure was used to construct a plasmid for bapaUGAc expression in the Golgi apparatus. Briefly, oligos Golgi-bapaUGAc-F and MAH-bapaUGAc-R were used to amplify the bapaUGAc gene from pBAD-bapaUGAc. Oligos MAH-Golgi-F and MAH-Golgi-bapaUGAc-R were used to amplify an 81-amino-acid Golgi signal peptide derived from human β -1,4-galactosyltransferase 1, and HeLa cell cDNAs were used as the PCR template. The resultant fragments, along with pMAH predigested with Hind III and Apa I, were again assembled via Gibson assembly, resulting in the pMAH-Golgi-bapaUGAc plasmid. Negative control plasmids, pMAH-ER-ecpApple and pMAH-Golgi-ecpApple were generated using similar procedures from pBAD-ecpApple.

Mammalian cell culture, transfection, and live cell imaging.

Human Embryonic Kidney (HEK) 293T cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) containing 4.5 g/L glucose supplemented with 10% fetal bovine serum (FBS) in the humidified incubator containing 5% CO₂ at 37 °C. HEK 293T cells were seeded in 35-mm tissue culture dishes with glass coverslips at the bottom. Transfection was performed on the second day when cell confluency reached ~30%. Briefly, 1.5 μ g pMAH-POLY and 1.5 μ g pMAH-ER-bapaUGAc (or pMAH-Golgi-bapaUGAc) were added to 7.5 μ g of PEI (polyethyleneimine, linear, M.W. 25 kDa) in 600 μ L Opti-MEM (Gibco). For the dual-color imaging experiment, 1.5 μ g pMAH-POLY, 1.5 μ g pMAH-ER-bapaUGAc, 1 μ g pcDNA-UGAcS, and 9 μ g of PEI were used. The plasmid-PEI mixture was incubated at room temperature for 30 min before being added to HEK 293T cells in culture dishes. 16 hr later, fresh DMEM media containing 10% FBS and 2 mM *p*BoF were used to replace the transfection media. After another 24 hr, to deplete intracellular free *p*BoF amino acid, the culture media were again replaced with fresh DMEM containing 10% FBS but no *p*BoF. Imaging was performed 12 hr later. Cells were rinsed with a mammalian cell imaging buffer (140 mM NaCl, 5 mM KCl, 1 mM CaCl₂, 0.4 mM MgSO₄, 0.5 mM MgCl₂, 0.3 mM Na₂HPO₄, and 0.4 mM KH₂PO₄, 6 mM D-glucose, and 20 mM HEPES, pH 7.4) three times and kept in this buffer at room temperature for 30 min for stabilization of metabolism. Cells were next imaged using a Leica DMI8 inverted microscope with both confocal and wide-field imaging capabilities. A Leica TSC SPE module was used to acquire confocal images to confirm subcellular compartmental localization, while time-lapse images were acquired with a Leica EL6000 light source, a TRITC filter cube (545/25 nm bandpass excitation and 605/70-nm bandpass emission), and a Photometrics Prime 95B sCMOS camera. The acquisition intervals were 1 min, the excitation light intensity at the focal plane was ~0.97 W/cm², and the exposure time was 200 ms. To perturb the intracellular UDP-GlcNAc level, cells were treated with either 10 mM GlcN or 10 mM 2-DG pre-dissolved in the imaging buffer. Images were analyzed using the ImageJ software. Backgrounds were subtracted using the software by setting the rolling ball radius to 100 pixels. Intensity means of randomly selected cells were used for analysis. The fluorescence changes were calculated by

normalizing fluorescence intensities (F) to the value at starting time (F_0), and the F/F_0 ratios were plotted against time.

RESULTS AND DISCUSSION

Biosensor engineering

Our biosensor engineering journey started with cpmApple, a cpRFP mutant derived from a previously reported calcium indicator, R-GECO1 (Supporting Information, Fig. S1).^{28,29} Since cpmApple showed relatively low folding efficiency, we randomized the gene using error-prone polymerase chain reactions (EP-PCRs) and screened the library for increased fluorescence in *E. coli* colonies at 37 °C. We repeated this process for three rounds, resulting in an enhanced cpmApple mutant (ecpApple) with three mutations from cpmApple (Fig. S2). Next, we sought to introduce pBoF into ecpApple. Because we aimed at using carbohydrate binding to modulate fluorescence, we selected two residues (S14 and K30, Figs. S2 & S3) in spatial proximity to the chromophore of ecpApple and site-specifically introduced pBoF to each of the two residues. Among the two mutants, ecpApple with pBoF at residue 30 (ecpApple-K30B) showed a reproducible fluorescence turn-off response to 5 mM UDP-GlcNAc, but the response was not specific since several other carbohydrate analogs could also cause fluorescence changes.

To increase the specificity of the response, we appended a 13-amino-acid peptide sequence derived from a natural *N*-acetylglucosamine (GlcNAc)-binding lectin, UEA-II (see bapaUGAc0.1 in Fig. S1).³⁴ Next, we added several randomized residues to the C-terminus, followed by examining the addition of another peptide sequence derived from natural GlcNAc binding lectins. We tested several options, and a 14-amino-acid loop sequence derived from the GS-II lectin^{31,32} led to the best responsiveness and specificity (bapaUGAc0.3 in Fig. S1). Furthermore, we carried out three additional rounds of error-prone PCRs, and in addition, optimized the linker between the UEA-II-derived peptide and the N-terminus of ecpApple. Our effort resulted in the final bapaUGAc mutant (Figs. 2a, S1, and S2).

In vitro characterization

The fluorescence excitation and emission peaks of bapaUGAc as a purified protein were at 575 and 600 nm, respectively (Fig. 2b). The fluorescence decreased by ~50% in response to 5 mM UDP-GlcNAc. We further examined the response using several UDP-GlcNAc concentrations, and the apparent dissociate constant (K_d) was determined to be 2.14 ± 0.29 mM (Fig. 2c). The response displayed some cooperative effect, as the Hill coefficient was 1.7. We further used plate-reader-based assays to roughly assess the speed of ligand association and dissociation. Under our experimental condition, the association and dissociation were completed within 40 s and 1–1.5 min, respectively (Fig. S4), suggesting that bapaUGAc can be used to monitor metabolic changes with a minute-level temporal resolution. In addition, we examined the specificity of bapaUGAc among common nucleotide sugars and nucleotides (Fig. 2d). UDP-GlcNAc caused robust fluorescence turn-off, while other tested compounds induced no or only small fluorescence changes. The small fluorescence decreases caused by UDP-Gal and UDP-GalNAc are unlikely to cause

any interference because both UDP-Gal and UDP-GalNAc are less abundant than UDP-GlcNAc. In live mammalian cells, UDP-GlcNAc and UDP-GalNAc are interconvertible and the concentration of UDP-GlcNAc is maintained to be ~3-fold higher than that of UDP-GalNAc by epimerase.³⁵ Furthermore, the concentration of UDP-Gal is lower than that of UDP-GalNAc by another 100-fold.³⁶ As a negative control, we prepared the ecpApple protein and confirmed its unresponsiveness to UDP-GlcNAc (Fig. S5ab).

In the final bapaUGAc construct, the GlcNAc-binding elements including *p*BoF30 and the two lectin-derived peptides are close to the chromophore in the 3-dimensional space (Fig. S3). To further explore the mechanism for the GlcNAc-dependent fluorescence change, we examined the *pH* dependence of bapaUGAc fluorescence but observed no apparent *pK_a* shift in the presence or absence of UDP-GlcNAc (Fig. 5c). In addition, we recorded the absorbance of bapaUGAc. UDP-GlcNAc caused a decrease in the absorbance but did not drastically alter the overall spectral shape (Fig. 5d). Collectively, the results suggest that UDP-GlcNAc binding does not modulate chromophore ionization; instead, it quenches the fluorescence of the chromophore via a yet-to-identified mechanism. We have recently obtained the crystal structure of pnRFP, a similar protein containing *p*BoF at residue 30.³⁷ Assuming that the *p*BoF residue in bapaUGAc, in the absence of UDP-GlcNAc, adopts a similar conformation to that in pnRFP, the binding of UDP-GlcNAc to bapaUGAc via *p*BoF30 and the lectin-derived peptides would require a rotation of the *p*BoF residue from the interior of the FP β -barrel to the exterior (Fig. S6). Such structural change may affect the local electrostatic environment of the chromophore or even rearrange the H-bond network around the chromophore, leading to the modulation of the bapaUGAc fluorescence.

Imaging UDP-GlcNAc changes in live cells

We next explored the use of bapaUGAc to image UDP-GlcNAc level changes in mammalian cells. The *K_d* of bapaUGAc is likely too high compared to the concentrations of UDP-GlcNAc in the cytosol. On the other hand, nucleotide sugars are more concentrated in the ER and Golgi,³⁸ so we tested whether bapaUGAc could sense UDP-GlcNAc in the ER and Golgi. We used signal peptides to genetically express bapaUGAc in the corresponding subcellular compartments of human embryonic kidney (HEK) 293T cells (Fig. 3a,b). We first treated the cells with GlcN, a potent precursor that can boost the biosynthesis of UDP-GlcNAc (Fig. 3c). As expected, the fluorescence of both ER- and Golgi-localized bapaUGAc decreased in response to GlcN (Fig. 3d,e). Next, we treated the cells with 2-DG, a glucose analog inhibiting hexokinase (HK) and phosphoglucose isomerase (PGI) (Fig. 3c). The addition of 2-DG caused a robust rise of bapaUGAc signals in the ER or Golgi (Fig. 3 f,g). In contrast, cells expressing ecpApple in the ER or Golgi were unresponsive to GlcN and 2-DG (Fig. 3 d,e,f,g). Collectively, the results support that bapaUGAc can detect UDP-GlcNAc level changes in the ER and Golgi in living mammalian cells in response to metabolic disruption and pharmacological inhibition.

Although it is widely accepted that the ER and Golgi concentrate nucleotide sugars by 10 to 30-fold,^{22,39} the exact concentrations of nucleotide sugars in these organelles are not yet clearly defined. A previous review estimated the cytosolic UDP-GlcNAc concentration in the range of 2–30 μ M,²² while another study reported the total UDP-GlcNAc concentration

in mammalian cells to be as high as 9 mM.⁴⁰ Furthermore, several other studies determined the total UDP-GlcNAc concentrations to be ~30–800 μ M.^{41–44} Our own work, which employed the green fluorescent UGAcS to image UDP-GlcNAc concentration changes in the cytosol of HEK 293T cells, suggests a cytosolic UDP-GlcNAc concentration of ~70–100 μ M.²⁶ Assuming an average of 20-fold enrichment, this would bring the number to ~1.4–2 mM for the basal UDP-GlcNAc concentration in the ER and Golgi. This number agrees well with the fact that ER- and Golgi-localized bapaUGAc, with the K_d determined to be low millimolar *in vitro* (Fig. 2c), responded to live-cell metabolic disruptions in both directions (Fig. 3d-g).

To visualize UDP-GlcNAc in different subcellular compartments, we paired ER-localized bapaUGAc with our cytosolic UGAcS biosensor. We co-expressed the two biosensors in HEK 293T cells and monitored fluorescence changes in the green and red channels after adding GlcN, a precursor for UDP-GlcNAc synthesis (Fig. 4). Since UGAcS is a fluorescence turn-on sensor, the green fluorescence of UGAcS increased in response to GlcN, indicating an increase in UDP-GlcNAc in the cytosol. Meanwhile, the red fluorescence of bapaUGAc decreased over time, suggesting an accompanied increase in UDP-GlcNAc in the ER. These results support the compatibility of bapaUGAc and UGAcS for dual-color, dual-compartment imaging of UDP-GlcNAc.

CONCLUSIONS

In summary, we have integrated GCE and protein engineering to create a new FP-based boronolactin called bapaUGAc, which can serve as a specific sensor for UDP-GlcNAc. We characterized the functionality of bapaUGAc both *in vitro* and in live cells. We further achieved dual-color monitoring of UDP-GlcNAc levels in two different subcellular compartments. We expect bapaUGAc to be a valuable research tool for studying UDP-GlcNAc and glycobiology. In addition, our results will catalyze the future development of specific boronolactins for other carbohydrate molecules.

bapaUGAc integrates several genetically encoded elements for specific UDP-GlcNAc recognition. First, pBoF was introduced to a residue close to the FP chromophore. Next, we grafted two peptides derived from GlcNAc-binding lectins to the N- and C-termini of the protein to enhance specificity. We chose these small peptides because they were previously reported to retain carbohydrate-binding properties.^{31,34,45,46} In addition, their small sizes and structural flexibility, compared to intact lectins, minimize steric hindrance, allowing UDP-GlcNAc to simultaneously interact with multiple recognition elements. In the absence of UDP-GlcNAc, bapaUGAc is highly fluorescent and likely contains a relatively coplanar chromophore similar to pNRFP (Fig. S6). UDP-GlcNAc binding quenched the fluorescence, resulting in a turn-off response. Would it be possible to use the strategy presented here to re-engineer a fluorescence turn-on sensor? Currently, we do not know any fundamental mechanism that disallows UDP-GlcNAc-induced fluorescence turn-on. In principle, if we can first convert the pBoF-containing protein to be weakly fluorescent in the absence of UDP-GlcNAc, UDP-GlcNAc binding may enhance the fluorescence for a turn-on response. Future studies will be needed to test the feasibility of this proposal.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

Research reported in this publication was supported by the NIH Common Fund Glycoscience Program and the NIH Awards U01CA230817 and R01DK122253.

REFERENCES

1. Jin S; Cheng Y; Reid S; Li M; Wang B, Carbohydrate recognition by boronolactins, small molecules, and lectins. *Med. Res. Rev* 2010, 30 (2), 171–257. [PubMed: 19291708]
2. Bies C; Lehr CM; Woodley JF, Lectin-mediated drug targeting: history and applications. *Adv. Drug Deliv. Rev* 2004, 56 (4), 425–435. [PubMed: 14969751]
3. Ghazarian H; Itoni B; Oppenheimer SB, A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. *Acta Histochem.* 2011, 113 (3), 236–247. [PubMed: 20199800]
4. Heimburg-Molinari J; Rittenhouse-Olson K, Development and characterization of antibodies to carbohydrate antigens. *Methods Mol. Biol* 2009, 534, 341–357. [PubMed: 19277540]
5. Li M; Lin N; Huang Z; Du L; Altier C; Fang H; Wang B, Selecting aptamers for a glycoprotein through the incorporation of the boronic acid moiety. *J. Am. Chem. Soc* 2008, 130 (38), 12636–12638.
6. Edwards NY; Sager TW; McDevitt JT; Anslyn EV, Boronic acid based peptidic receptors for pattern-based saccharide sensing in neutral aqueous media, an application in real-life samples. *J. Am. Chem. Soc* 2007, 129 (44), 13575–13583.
7. Springsteen G; Wang B, A detailed examination of boronic acid-diol complexation. *Tetrahedron* 2002, 58 (58), 5291–5300.
8. James TD; Shinkai S, Artificial receptors as chemosensors for carbohydrates. *Host-Guest Chem.* 2002, 159–200.
9. Williams GT; Kedge JL; Fossey JS, Molecular Boronic Acid-Based Saccharide Sensors. *ACS Sens.* 2021, 6 (4), 1508–1528. [PubMed: 33844515]
10. Zhao J; Davidson MG; Mahon MF; Kociok-Köhn G; James TD, An enantioselective fluorescent sensor for sugar acids. *J. Am. Chem. Soc* 2004, 126 (49), 16179–16186.
11. Liu CC; Schultz PG, Adding new chemistries to the genetic code. *Annu. Rev. Biochem* 2010, 79, 413–444. [PubMed: 20307192]
12. Ai HW, Biochemical analysis with the expanded genetic lexicon. *Anal. Bioanal. Chem* 2012, 403 (8), 2089–2102. [PubMed: 22322380]
13. Brustad E; Bushey ML; Lee JW; Groff D; Liu W; Schultz PG, A genetically encoded boronate-containing amino acid. *Angew. Chem. Int. Ed. Engl* 2008, 47 (43), 8220–8223. [PubMed: 18816552]
14. Chen ZJ; Ren W; Wright QE; Ai HW, Genetically encoded fluorescent probe for the selective detection of peroxynitrite. *J. Am. Chem. Soc* 2013, 135 (40), 14940–14943.
15. Liu CC; Mack AV; Brustad EM; Mills JH; Groff D; Smider VV; Schultz PG, Evolution of proteins with genetically encoded “chemical warheads”. *J. Am. Chem. Soc* 2009, 131 (28), 9616–9617. [PubMed: 19555063]
16. Greenwald EC; Mehta S; Zhang J, Genetically Encoded Fluorescent Biosensors Illuminate the Spatiotemporal Regulation of Signaling Networks. *Chem. Rev* 2018, 118 (24), 11707–11794.
17. Niu W; Guo J, Expanding the chemistry of fluorescent protein biosensors through genetic incorporation of unnatural amino acids. *Mol. Biosyst* 2013, 9 (12), 2961–2970. [PubMed: 24080788]
18. Chen S; Chen ZJ; Ren W; Ai HW, Reaction-based genetically encoded fluorescent hydrogen sulfide sensors. *J. Am. Chem. Soc* 2012, 134 (23), 9589–9592. [PubMed: 22642566]

19. Chen ZJ; Tian Z; Kallio K; Oleson AL; Ji A; Borchardt D; Jiang DE; Remington SJ; Ai HW, The N-B Interaction through a Water Bridge: Understanding the Chemoselectivity of a Fluorescent Protein Based Probe for Peroxynitrite. *J. Am. Chem. Soc* 2016, 138 (14), 4900–4907. [PubMed: 27019313]
20. Chen Z; Zhang S; Li X; Ai HW, A high-performance genetically encoded fluorescent biosensor for imaging physiological peroxynitrite. *Cell Chem. Biol* 2021, 28 (11), 1542–1553. [PubMed: 33581056]
21. Wang F; Niu W; Guo JT; Schultz PG, Unnatural Amino Acid Mutagenesis of Fluorescent Proteins. *Angew. Chem. Int. Edit* 2012, 51 (40), 10132–10135.
22. Bond MR; Hanover JA, A little sugar goes a long way: the cell biology of O-GlcNAc. *J. Cell Biol* 2015, 208 (7), 869–880. [PubMed: 25825515]
23. Akella NM; Ciraku L; Reginato MJ, Fueling the fire: emerging role of the hexosamine biosynthetic pathway in cancer. *BMC Biol.* 2019, 17 (1), 52. [PubMed: 31272438]
24. Schwein PA; Woo CM, The O-GlcNAc Modification on Kinases. *ACS Chem. Biol* 2020, 15 (3), 602–617. [PubMed: 32155042]
25. Khoder-Agha F; Sosicka P; Escrive Conde M; Hassinen A; Glumoff T; Olczak M; Kellokumpu S, N-acetylglucosaminyltransferases and nucleotide sugar transporters form multi-enzyme-multi-transporter assemblies in golgi membranes in vivo. *Cell Mol. Life Sci* 2019, 76 (9), 1821–1832. [PubMed: 30737517]
26. Li Z; Zhang J; Ai HW, Genetically Encoded Green Fluorescent Biosensors for Monitoring UDP-GlcNAc in Live Cells. *ACS Cent. Sci* 2021, 7 (10), 1763–1770. [PubMed: 34729420]
27. Young TS; Ahmad I; Yin JA; Schultz PG, An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol* 2010, 395 (2), 361–374. [PubMed: 19852970]
28. Zhao Y; Araki S; Wu J; Teramoto T; Chang YF; Nakano M; Abdelfattah AS; Fujiwara M; Ishihara T; Nagai T; Campbell RE, An expanded palette of genetically encoded Ca^{2+} indicators. *Science* 2011, 333 (6051), 1888–1891. [PubMed: 21903779]
29. Fan Y; Chen Z; Ai HW, Monitoring redox dynamics in living cells with a redox-sensitive red fluorescent protein. *Anal. Chem* 2015, 87, 2802–2810. [PubMed: 25666702]
30. Wu T; Pang Y; Ai HW, Circularly Permuted Far-Red Fluorescent Proteins. *Biosensors* 2021, 11 (11), 438. [PubMed: 34821654]
31. Zhu K; Bressan RA; Hasegawa PM; Murdock LL, Identification of N-acetylglucosamine binding residues in Griffonia simplicifolia lectin II. *FEBS Lett.* 1996, 390 (3), 271–274. [PubMed: 8706875]
32. Zhu-Salzman K; Shade RE; Koiwa H; Salzman RA; Narasimhan M; Bressan RA; Hasegawa PM; Murdock LL, Carbohydrate binding and resistance to proteolysis control insecticidal activity of Griffonia simplicifolia lectin II. *Proc. Natl. Acad. Sci. U.S.A* 1998, 95 (25), 15123–15128.
33. Campbell RE; Tour O; Palmer AE; Steinbach PA; Baird GS; Zacharias DA; Tsien RY, A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A* 2002, 99 (12), 7877–7882. [PubMed: 12060735]
34. Loris R; De Greve H; Dao-Thi MH; Messens J; Imberty A; Wyns L, Structural basis of carbohydrate recognition by lectin II from *Ulex europaeus*, a protein with a promiscuous carbohydrate-binding site. *J. Mol. Biol* 2000, 301 (4), 987–1002. [PubMed: 10966800]
35. Robinson KA; Weinstein ML; Lindenmayer GE; Buse MG, Effects of diabetes and hyperglycemia on the hexosamine synthesis pathway in rat muscle and liver. *Diabetes* 1995, 44 (12), 1438–1446. [PubMed: 7589852]
36. Broussard A; Florwick A; Desbiens C; Nischan N; Robertson C; Guan Z; Kohler JJ; Wells L; Boyce M, Human UDP-galactose 4'-epimerase (GALE) is required for cell-surface glycome structure and function. *J. Biol. Chem* 2020, 295 (5), 1225–1239. [PubMed: 31819007]
37. Yu P; Mian H; Yichong F; Hsien-Wei Y; Ying X; Ho Leung N; Hui-wang A, Development, Characterization, and Structural Analysis of a Genetically Encoded Red Fluorescent Peroxynitrite Biosensor. *bioRxiv* 2023, DIO: 10.1101/2023.1104.1113.536636.
38. Parker JL; Newstead S, Gateway to the Golgi: molecular mechanisms of nucleotide sugar transporters. *Curr. Opin. Struct. Biol* 2019, 57, 127–134. [PubMed: 30999236]

39. Hirschberg CB, My journey in the discovery of nucleotide sugar transporters of the Golgi apparatus. *J. Biol. Chem* 2018, 293 (33), 12653–12662.
40. Park JO; Rubin SA; Xu YF; Amador-Noguez D; Fan J; Shlomi T; Rabinowitz JD, Metabolite concentrations, fluxes and free energies imply efficient enzyme usage. *Nat. Chem. Biol* 2016, 12 (7), 482–489. [PubMed: 27159581]
41. Namboori SC; Graham DE, Enzymatic analysis of uridine diphosphate N-acetyl-D-glucosamine. *Anal. Biochem* 2008, 381 (1), 94–100. [PubMed: 18634748]
42. Kochanowski N; Blanchard F; Cacan R; Chirat F; Guedon E; Marc A; Goergen JL, Intracellular nucleotide and nucleotide sugar contents of cultured CHO cells determined by a fast, sensitive, and high-resolution ion-pair RP-HPLC. *Anal. Biochem* 2006, 348 (2), 243–251. [PubMed: 16325757]
43. Haltiwanger RS; Grove K; Philipsberg GA, Modulation of O-linked N-acetylglucosamine levels on nuclear and cytoplasmic proteins in vivo using the peptide O-GlcNAc-beta-N-acetylglucosaminidase inhibitor O-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-N-phenylcarbamate. *J. Biol. Chem* 1998, 273 (6), 3611–3617. [PubMed: 9452489]
44. Ritter JB; Genzel Y; Reichl U, Simultaneous extraction of several metabolites of energy metabolism and related substances in mammalian cells: optimization using experimental design. *Anal. Biochem* 2008, 373 (2), 349–369. [PubMed: 18036549]
45. Konami Y; Yamamoto K; Osawa T, Purification and characterization of carbohydrate-binding peptides from *Lotus tetragonolobus* and *Ulex europeus* seed lectins using affinity chromatography. *J Chromatogr* 1992, 597 (1–2), 213–219. [PubMed: 1517320]
46. Yamamoto K; Konami Y; Osawa T; Irimura T, Carbohydrate-binding peptides from several anti-H(O) lectins. *J. Biochem* 1992, 111 (4), 436–439. [PubMed: 1618731]

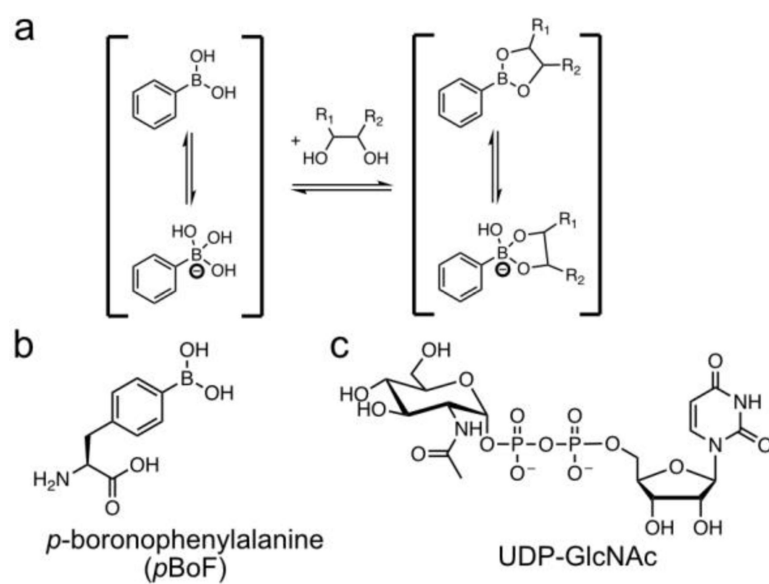
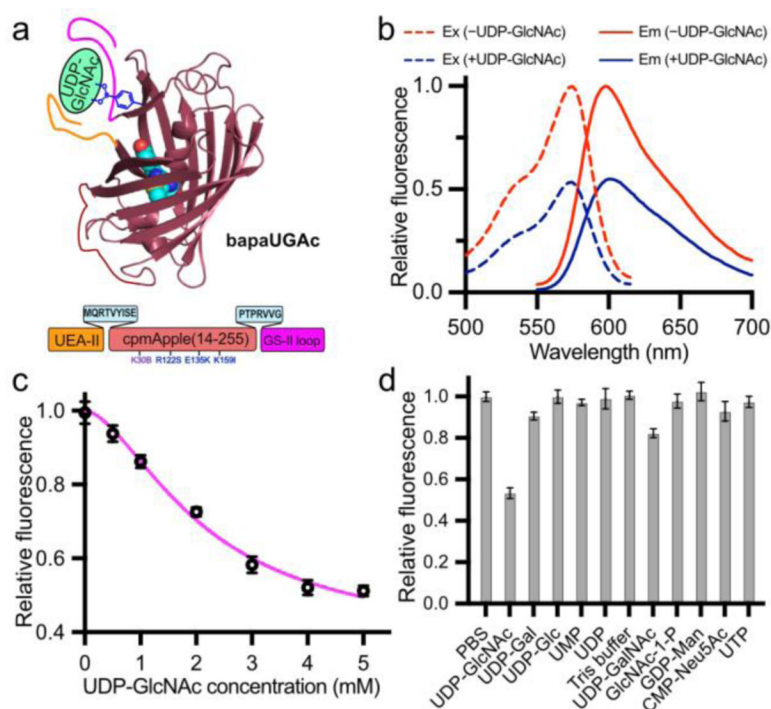


Figure 1.
(a) Chemical reaction between phenylboronic acid and diol. (b, c) Chemical structures of the indicated compounds.

**Figure 2.**

(a) Schematic illustration of bapaUGAc and its interaction with UDP-GlcNAc. The cpRFP chromophore is displayed as cyan spheres. The bottom shows the domain arrangement with the linkers and mutations highlighted, and B represents the *pBoF* residue. (b) Fluorescence excitation (dashed line) and emission (solid line) spectra of bapaUGAc in the presence (blue) and absence (red) of UDP-GlcNAc. (c) Dose-dependent fluorescence response. (d) Specificity characterization with 1 μ M bapaUGAc protein and 5 mM of each compound. Data are presented as mean and s.d. of three technical replicates.

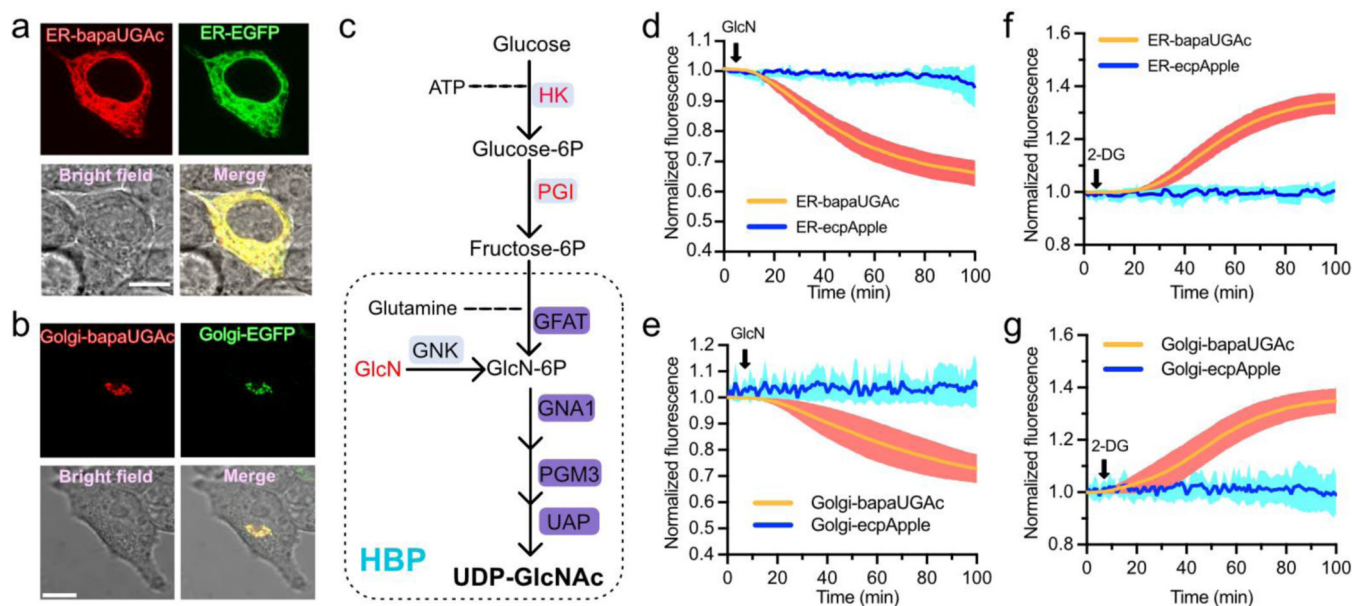


Figure 3. Monitoring UDP-GlcNAc level changes in the ER and Golgi in live mammalian cells using bapaUGAc. (a, b) Confocal images of the co-expression of ER-bapaUGAc and ER-EGFP (a), or Golgi-bapaUGAc and Golgi-EGFP (b) in HEK 293T cells. Scale bar, 10 μ m. (c) Schematic illustration of the hexosamine biosynthetic pathway (HBP). Glucosamine (GlcN), as well as hexokinase (HK) and phosphoglucose isomerase (PGI) that are enzymes inhibited by 2-deoxy-D-glucose (2-DG), are highlighted in red. (d, e) Time-lapse responses of ER- (d) or Golgi- (e) localized bapaUGAc in HEK 293T to extracellular addition of GlcN (10 mM) (f, g) Time-lapse responses of ER- (f) or Golgi- (g) localized bapaUGAc in HEK 293T to extracellular addition of 2-DG (10 mM). The interval for image acquisition was 1 min. Data are presented as mean and s.d. of eight individual cells from three technical replicates.

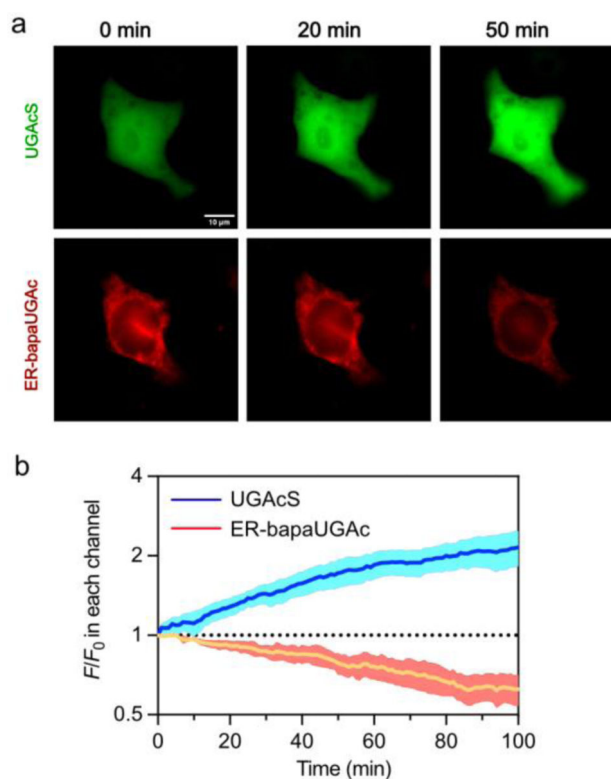


Figure 4.

Dual-color wide-field imaging of UDP-GlcNAc level changes in the cytosol and ER using green fluorescent UGAcS and red fluorescent bapaUGAc. (a) Representative fluorescence images. Scale bar, 10 μm . (b) Quantitation of the responses of UGAcS and bapaUGAc to extracellular addition of GlcN (10 mM). The intervals for image acquisition were 1 min. F/F_0 represents the fluorescence intensities of individual cells normalized to the value at starting time. Data are presented as mean and s.d. of 12 individual cells from three technical replicates.