Design of a highly specific glutamine sensor by splitting the glutamine-binding protein

Moon-Hyeong Seo[1]*, Yun Lim[1], Youn Hee Jung[1], Jae Hoon Lee[2], Min Seok Baek[4], Je Hyeong Jung[2], Ho-Youn Kim[2], Wookbin Lee[1] and Keunwan Park[3]*

[1]Natural Product Research Center, [2]Smart Farm Research Center and [3]Natural Product Informatics Research Center, Korea Institute of Science and Technology, Gangneung 25451, South Korea,
[4]Department of Neurology, Wonju Severance Christian Hospital, Yonsei University Wonju College of Medicine, Wonju 26426, South Korea

*Corresponding authors

E-mail: mhseo@kist.re.kr (M.-H. Seo) and keunwan@kist.re.kr (K. Park)

Abstract

Biosensors for ligands are versatile tools for biological applications ranging from disease diagnosis to the detection of environmental chemicals. Thus, expanding the repertoire of biosensor toolkits provides novel insights into the unappreciated potentials of molecules. While the natural recognition domains have been employed in current biosensors, their use is limited to the methods such as fluorescence resonance energy transfer sensors or circular permutation. Here, we describe a first-of-its-kind approach that transforms a protein exhibiting ligand-mediated hinge-bending motion into a highly specific fluorescent biosensor. As a proof-of-concept, a glutamine sensor named Q-SHINE (Glutamine sensor via Split HINgE-motion binding protein) was designed by splitting a glutamine-binding protein into two separate stable domains that, in principle, only dimerize in the presence of glutamine. The application of Q-SHINE to determine glutamine concentrations in solution or mouse serum yielded comparable sensitivity and higher specificity to a conventional glutamine assay kit. Moreover, genetically encoded Q-SHINE could effectively monitor intracellular glutamine levels. Aside from highlighting the reliability of a versatile glutamine sensor, our study opens avenues for researchers who wish to explore the application of ligand-binding proteins as biosensors.

Keywords

Biosensor, Genetically encoded sensor, Glutamine, Glutamine-binding protein, Periplasmic binding protein

1. Introduction

Ligands play a vital role in biological systems.[1] The sensing, detection, or *in situ* monitoring of target molecules (clinical biomarkers, signaling molecules, nutrients, metabolites, and environmental chemicals) is a powerful technology that has led to advances in fields, such as biomedical research, food and agricultural industries, and environmental analysis.[2,3] Thus, expanding the range of biosensor toolkits for target molecules whose significance is not sufficiently understood, will provide researchers with deeper insights into the associated biological events. For biosensors, the installation of a selective probe as a target is essential. The function of many proteins in living organisms is initiated by the binding of specific molecules. Accordingly, a large panel of natural ligand-binding proteins, such as periplasmic binding proteins (PBPs)[4,5] and lectins,[6] has been exploited to constitute a molecular recognition unit in a sensor system, followed by combination with a reporter domain that converts a binding event into detectable signals or gene expression.[3,7]

Among ligand-binding proteins, the PBP family is steadily employed as a scaffold for a binding module owing to its intrinsic features, such as its innate narrow specificity and affinity towards corresponding small molecules.[5,8,9] In particular, the large conformational changes in PBPs upon binding or release of a ligand allow for the development of versatile PBP-based fluorescence resonance energy transfer (FRET) sensors[10,11] or protein switches via the insertion of circular permuted output domains.[9] Despite their successful application, current design strategies for PBP-based sensors mostly rely on where and how to conjugate fluorophores (or quenchers) to the PBP domain to optimize sensitivity and dynamic range. Modulation of the affinity or specificity of PBPs against a ligand is further considered in the development of advanced sensor systems. Therefore, PBPs lacking enough bending movement (since the degree of bend the hinge allows varies) or those intolerant to domains inserted for output signals are not available as sensing elements. The barriers that limit the repertoire of PBPs have been challenged by several approaches, such as metagenomics, directed evolution, and computational design strategies.[9] Beyond such efforts, if a novel design strategy is developed, it will enable us to further expand the applications of PBPs.

Here, we introduce a novel method for the utilization of a naturally occurring PBP to generate a biosensor system that recognizes a ligand and fluoresces upon binding. As a proof-of-concept, we designed a novel glutamine (Gln) sensor system using a Gln-binding protein (QBP) from *Escherichia coli*. Although several QBP-FRET sensors have been reported in the past,[12–14]

their low dynamic range and frequent experimental errors have limited their widespread adoption, despite the physiological and clinical importance of Gln.[14–18] The sensing module, termed Q-SHINE (Gln sensor via Split HINgE-motion binding protein), was designed by dividing a Gln-binding protein into two independent domains based on the principle of chemically induced hetero-dimerization. The subsequent recombination of on-demand split reporter systems generated a facile tool for measuring and monitoring Gln. Q-SHINE yielded comparable sensitivity and higher specificity than a conventional Gln assay kit; this could enable reagentless point-of-care diagnosis of diverse diseases related to the levels of Gln in the body, such as cancer,[18–20] diabetes,[21,22] and neurodegeneration.[23–25]. Furthermore, genetically encodable Q-SHINE could allow for live-cell imaging of intracellular Gln concentration.

2. Materials and Methods

2.1. Split design of Q-SHINE domains based on hinge bending motion analysis

The domain movement of QBP was analyzed using DynDom; [26] two heterogeneous QBP conformers were compared (PDB ID: 1GGG and 1WDN). From the analysis, it was found that QBP consists of two motional domains and exhibits open-closed bending motion (rotation angle: 55.7 degree, and closure ratio: 99.8%). The residues situated in the hinge-bending region were 87–89 and 181–184 in the QBP (1WDN). Splitting the QBP at the two connecting strands generated two separate domains, where one domain (QBP_Lg) was fragmented into two segments (**Figure 1A**). The discontinuous fragments were redesigned by modeling a new linker sequence, which connected the cleaved ends to generate a single polypeptide chain. Linker modeling was performed using RosettaRemodel with default options;[27] loop closure trials were attempted 10,000 times for loop lengths ranging from two to five. The designed loops were then filtered using LoopAnalyzerMover in RosettaScript.[28] The Rosetta total score, which summarizes loop quality, was used to select the final loop sequence that stabilized the fused domain. A shorter loop length was also prioritized to minimize flexibility. Structures of the designed domains were predicted using AlphaFold Colab.[29]

2.2. Gene cloning for bacteria, mammalian cell, and plant cell expression

Genes for the sensor components (QBP, mCherry, NanoBiT and other PBPs with nucleotide sequences complementary to pET21a, Novagen, at both ends) were synthesized by IDT gBlocks® Gene Fragments (Integrated DNA Technologies, Iowa, USA) and Gene Fragments (Twist Bioscience, South San Franscisco, USA). For bacterial expression, the pET21a plasmid was linearized by restriction with the NdeI and XhoI enzymes. The synthesized gene fragments were then cloned into the linear pET21a vector using the In-Fusion® HD Cloning Kit (Takara Bio, Kusatsu, Japan). For mammalian expression, the pcDNA3.1(+) vector was linearized using PCR amplification with a pair of primers (pcDNA_Amp_FW and pcDNA_Amp_RV). The split QBP and mCherry constructs were amplified and assembled with their corresponding primers (Table S1) and cloned using the In-Fusion HD Cloning Kit. For plant cell expression, the whole expression cassette from the pcDNA plasmid amplified with primers (QBP_Split_FW_pGO and QBP_Split_RV or QBP_Split_RV_Mut) was cloned into the pTALCOMT plasmid for Agrobacterium-mediated transformation.[30]

2.3. Protein expression and purification

Plasmids containing the sensor constructs were transformed into *E. coli* BL21 (DE3) cells. Protein expression was induced with 0.5 mM isopropyl b-D-1-thiogalactopyranoside at 18 °C. Proteins were purified from the soluble cell lysate using His60 Ni Superflow Resin (Takara Bio) and concentration was measured using NanoDrop. For analytical gel filtration chromatography, the purified proteins from a Ni column were loaded onto a Superdex 75 Increase column equilibrated with 20 mM Tris-HCl (pH 8.0, 150 mM NaCl). Molecular mass standards included conalbumin (75 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa) and ribonuclease (13.7 kDa).

2.4. *In vitro* BiFC

Next, 5 μM of both sensor domains were mixed with different concentrations of Gln or other amino acids in the Nunc Microwell 96-well Microplate (Thermo fisher, Massachusetts, United States), and the mixture was incubated in the dark for 1–5 h at room temperature. Emission spectra of mCherry were read in 1 nm increments from 600 to 620 nm using an Infinite M1000 (Tecan, Männedorf, Switzerland) under 580 nm excitation.

2.5. Luciferase assay

We mixed 5 μM of two domains of Q-SHINE_NanoBiT with 0–5 mM Gln for 10 min at room temperature in the dark. We also added 25 μL of the Nano-Glo Live Cell reagent included in the Nano-Glo Live Cell Assay System (Promega) to each well and gently mixed for 10 s. Luminescence was measured for a single time point using the GloMax Navigator System (Promega) and an integration time of 1 s.

2.6. Measurement of Gln concentration in serum samples

Female C57BL/6 mice, aged 6–12 weeks, were purchased from Orient-Bio (Seongnam, Korea). The mice were cared for in accordance with institutional guidelines for experimental animals. All experimental protocols were approved by the Animal Care and Use Committee of KIST

(Approval No. KIST-2021-09-104). To collect serum samples from the mice, blood samples were taken from the abdominal caval vein and centrifuged at 300 ×*g* for 10 min at 4 °C after anesthesia was administered and an abdominal incision was created. At least 200 μL serum was collected and stored at -80 °C. For the Q-SHINE measurement, 20 μL of serum sample from each mouse and different concentrations of Gln solution were dissolved in distilled water. The sample was then mixed with 5 μM of two domains of Q-SHINE, followed by incubation in the dark and at room temperature for 5 h. The standard curve is fitted by nonlinear curve fitting method using the Hill1 function in Origin 2020 (OriginLab Corporation, Northampton, USA), and the concentration of Gln in serum samples was calculated using the 'Find X from Y' function based on the fitted graph. For comparison, Gln concentration in mouse serum was simultaneously determined using the EnzyChrom Glutamine Assay Kit (EGLN-100, BioAssay Systems, Hayward, CA, USA) according to the manufacturer's protocol.

2.7. Mammalian cell culture and transfection

Human embryonic kidney 293T (HEK293T) was purchased from the American Type Culture Collection (ATCC). The cell line was tested for mycoplasma contamination and used only within 30 passages. HEK293T was maintained in Dulbecco's Modified Eagle Medium (DMEM; Gibco, Massachusetts, USA) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS; Gibco), 100 U/mL penicillin and 100 μg/mL streptomycin at 37 °C in a humidified 5% CO_2 incubator. The mammalian expression vector was transiently transfected with Lipofectamine 3000 (Invitrogen, Thermo Fisher Scientific, Massachusetts, USA). For the intracellular Bimolecular fluorescence complementation (BiFC) experiments, HEK293T cells were seeded into 6-well cell-culture treated plates (Thermo Fisher Scientific) and grown at 37 °C in 5% CO_2 overnight in DMEM/10% FBS/1% penicillin/streptomycin to 4–50% confluency. After the medium was changed to DMEM (no Gln)/10% FBS/1% PS to deprive the cells of endogenous Gln, cells were transfected with plasmids containing Q-SHINE domains using Lipofectamine 3000. After 16 h of transfection, 0–2.0 mM of Gln were added to the media for induction of Gln-dependent BiFC.

2.8. Fluorescence microscopy and image analysis

Live cells were imaged with the EVOS FL Cell Imaging System (Thermo Fisher Scientific) after the addition of Gln. To obtain separate images of GFP and mCherry, two light cubes (470 nm excitation, 525 nm emission; 530 nm excitation, and 593 nm emission) were used. Images from both channels were overlaid to visualize overall fluorescence. Fluorescence intensities of mCherry and GFP were quantified using Image J software (National Institutes of Health). The images were segmented into each color in grayscale and the intensity of each fluorescence was measured on a median scale after the ROI settings were established through threshold application.

2.9. Flow cytometry

Gln-treated cells were analyzed using the BD FACSVerse (BD Biosciences, New Jersey, USA). Different concentrations of Gln were added to cells 24 h after transfection, and cells were further incubated for 24 h. Before analysis, cells were washed with PBS and diluted in cold PBS to a density of 50,000 cells/mL. A total of 30,000 cells were recorded per sample. GFP and mCherry were excited with a 488 nm laser and their fluorescence was detected using an FITC (530/30 nm) or PE (575/26 nm) emission filter. Dead cells were excluded by staining with the Fixable Viability Dye eFluor 780 (Invitrogen) and only cell populations showing $>10^3$ mCherry fluorescence intensity were considered transfected. The mean fluorescence intensity of the positive population was calculated and data were analyzed using FlowJo v.10.7.0.

2.10. Transient expression and imaging of fluorescent proteins in tobacco leaves

Agrobacterium-mediated transient expression analysis was performed as described by Lee et al.[31] *Agrobacterium tumefaciens* LBA4404 cells carrying different sensor constructs were adjusted to OD600 0.4 in resuspension solution (10 mM $MgCl_2$, 10 mM MES-KOH (pH 5.6), 100 μM acetosyringone) and infiltrated into leaves of 4 to 5-week-old *Nicotiana tabacum* cv. Samsun. Plants were grown and incubated at 24/21 °C (day/night temperatures), for a 16 h photoperiod, with 500 μmol/m^2/s light intensity and at 80% relative humidity. At 5 days post-infiltration, fluorescence intensities from the leaves were determined using an LSM5 Zeiss confocal laser scanning microscope with the ZEN image analysis software (Carl Zeiss, Jena, Germany). The excitation/emission spectra were 488/493–598 nm for GFP and 594/599–650 nm for mCherry.

3. Results and Discussion

3.1. Split design of a Gln-binding protein

PBPs consist of two domains connected by a hinge, where a ligand binds to the region between two lobes. In the apo-form (open conformation) of type II PBPs, there are few interdomain contacts, except for regions adjacent to hinge regions.[32–34] Ligand binding induces the closing and twisting motions of the QBP that bring the lobes closer by hydrogen bonds between the two domains and the ligand. In addition, as shown in our previous study,[35] the conformational energy (the energy required for the transition from an open to closed state) of a PBP is largely affected by the structural configuration of the hinge (i.e. mutations of the hinge sequence induce a significant change in conformational energy). We demonstrated that mutations in two hinge residues could create a 50-fold change in the binding affinity to the ligand while the ligand-binding interfaces remained identical. We assumed that two independent domains of PBP, when cleaved at the hinge, could dimerize in the presence of the ligand through only chemical interactions without entropic cost (conformational energy) of the hinge-bending motion. Therefore, we attempted to convert a PBP scaffold into a biosensor that dimerizes upon ligand binding (**Scheme 1**).

As a proof-of-concept, we began with *E. coli* QBP that assists Gln transport as part of the ABC transporter complex. By comparing the open and closed structures, the domain movement of QBP through the hinge-bending motion was analyzed to estimate the cleavage sites for least structural damage on the resulting split domains, as well as to ensure minimal effect on ligand binding. Residues 87–89 and 181–184 (amino acid numbering in PDB 1WDN) were identified as hinge-bending regions (**Figure 1A**). Two domains were separated by the deletion of Lys87-Ser88 and Ala182. Since cutting two connecting strands of a QBP results in a noncontinuous large domain (Lg; residues 1–86 and 183–226) and a continuous small domain (Sm; residues 89–181), a joining linker was needed for the stable expression of the large domain. Of the three tested loops, a designed linker (TGNG) that closed the gap between Tyr86 and Gln183 in QBP Lg showed superior solubility when expressed in *E.coli* BL21(DE3) (data not shown). The structure of the two divided domains, when predicted with AlphaFold, showed an insignificant difference from the global fold of native QBP (**Figure 1B**). Each split QBP fragment, Lg and Sm, was subsequently recombined with split mCherry fragments (MN159 & MC160),[36] which served as reporter elements, resulting in the final fusion construct, Q-SHINE consisting of Split QBP_Lg – mCherry_C (LgC) and mCherry_N – Split QBP_Sm (SmN) (**Figure 1C**).

Gly-Ser linkers were used to link the split QBP and mCherry fragments.

3.2. Molecular characterization of Q-SHINE

After purification, the Q-SHINE domains were analyzed using gel filtration chromatography. Both domains (LgC and SmN) moved faster than their expected size (**Figure S1**), the reason for which might presumably be the shape of recombinant proteins, which is elongated rather than spherical (**Figure 1C**) since the two independent fragments are connected by flexible linkers. Each fragment can also be destabilized. Interestingly, LgC was isolated as two separate peaks (LgC_1 and LgC_2) while SmN was eluted as a single peak with a minor tail. Even so, when either LgC_1 and LgC_2 were mixed with a fraction of SmN in the presence of Gln, the signal-to-noise ratio (i.e. mCherry fluorescence from the well with the ligand to the without ligand) of each pair displayed insignificant differences. The complementation rate (the increase of fluorescence) was faster when LgC_2 was used (data not shown). Although the underlying molecular mechanism determining the difference of complementation efficiency remains elusive, LgC_2 was used in subsequent experiments. Since BiFC relies on the propensity for self-assembly of nearby split fluorescent proteins, the self-assembly of Q-SHINE was assessed. We compared the association rate of Q-SHINE to Q-SHINE_ΔSm, which had lost the ability of binding Gln due to the removal of QBP Sm. As a result, Q-SHINE exhibited a slower increase in the signal caused by self-assembly than Q-SHINE_ΔSm (**Figure S2**). Fusion of the split QBP domain may hinder the self-complementation of split mCherry fragments, which is beneficial for a higher S/N ratio of Q-SHINE (**Figure S2**).

3.3. Gln measurement using Q-SHINE

To test Q-SHINE for its ability to detect Gln in solution, we mixed equivalent concentrations (5 μM) of Q-SHINE domains, LgC and SmN, with different concentrations of Gln. After incubation for a few hours, Gln-induced QBP dimerization was recognized by red fluorescence, in accordance with the proximity-induced complementation of mCherry. While the signal contrast was maximized during the 3–6 h incubation period, the sensitivity of Q-SHINE increased with a longer incubation time (**Figure 2A**). In terms of analytical sensitivity, Q-SHINE could detect Gln concentrations as low as 1 μM. The emission peak was observed at 610 nm, which is the reported value for the native mCherry protein (**Figure 2B**). The ligand

selectivity of Q-SHINE was further tested using structurally similar amino acids: glutamate, asparagine and D-Gln. As shown in **Figure 2C**, Q-SHINE exhibited a high specific reactivity to Gln, clearly distinguishing it from other amino acids. Importantly, Q-SHINE showed notable sensitivity in the range of hundreds of μM to mM concentrations, demonstrating its suitability for measuring physiological Gln levels.[14,15,19,37]

The broad applicability of our split PBP design for creating a sensing module was further examined by swapping the reporting element. Split QBP domains were fused with a widely used reporter system, NanoBiT. Split luciferase, unlike mCherry, is a reversible complementation system based on enzymatic reactions that enable real-time kinetic analysis of live cells.[38] The split QBP and NanoBiT fragments were recombined in a similar manner (**Table S2**). As expected, an increase in the luminescent signal was observed in proportion to Gln concentration (**Figure 2D**). The sensitivity of the NanoBiT-based system was commensurate with that of the mCherry-based original Q-SHINE because the sensing elements were essentially identical in both systems. The mCherry-based Q-SHINE showed a higher fold change in signals (i.e., high signal-to-noise ratio) owing to cumulative fluorescence by irreversibly complemented mCherry, while NanoBiT gave an immediate response in less than 10 min. This suggests that other split reporter systems,[39–41] such as β-lactamase, kinase, protease, biotin ligase, or Cas9, could be readily combined with Q-SHINE for other applications.

3.4. Measurement of Gln concentration in mouse serum

Since Q-SHINE accurately detected Gln in the sample, we further investigated the performance of Q-SHINE in quantifying Gln in a biological sample. Given that Q-SHINE showed remarkable signal contrast in the 100–1000 μM range and physiological Gln concentrations typically fall in the 200–1400 μM range,[15] we constituted the reaction using 20 μL of mouse serum to make a 100 μL reaction mixture. Thus, 40 μL of each Q-SHINE domain was mixed with 20 μL of four different mouse sera as well as standard Gln solutions from 0 to 2 mM. Fluorescence intensities from the wells were read by excitation at 587 nm and emission at 610 nm 5 h after incubation. The Gln concentration of mouse sera was determined based on the standard curve obtained by plotting the values from standard Gln solutions (**Figure S3**). Q-SHINE provided data comparable to a conventional, commercially available Gln colorimetric assay kit (**Figure 3**). Furthermore, intra-assay analysis verified the high repeatability of Q-

SHINE for serum samples (**Table 1**). Cells, tissues, or clinical samples with high glutamate or protein contents often result in high background signals in conventional colorimetric assays, which require additional pretreatment or background control. In addition, endogenous compounds may interfere with enzymatic reactions. As such, Gln is a less highlighted molecule due to its limited practical sensing system despite the growing number of reports demonstrating its relevance in multiple diseases, including cancer,[18–20] diabetes,[21,22] and neurodegeneration.[23–25] The simple composition of Q-SHINE, containing only two protein domains, enables straightforward measurement without any pretreatment process or expensive equipment. We anticipate that Q-SHINE could produce highly reliable and accurate data in biological and biomedical research that involves the use of conventional Gln assay kits.

3.5. Live-cell imaging to detect intracellular Gln

There has been immense progress in the development of genetically encodable fluorescent sensors by enabling the live imaging of cells. We, therefore, investigated whether Q-SHINE could monitor intracellular Gln concentrations when genetically encoded. For the equimolar production of each Q-SHINE fragment, LgC and SmN were connected by self-cleaving T2A peptides. GFP was attached at the N-terminal of LgC as an internal control, resulting in Q-SHINE-GFP (**Figure S4**). Q-SHINE-GFP was expressed in HEK293T with varying concentration of Gln in the culture medium. Cells were deprived of Gln for 16 h after transfection. Then, the medium was replaced with a fresh one containing different Gln concentrations. After 3, 6, and 9 h, mCherry fluorescence was detected in the HEK293T cells fluorescing GFP as a control (**Figure 4A**). As expected, Q-SHINE-GFP responded in a dose-dependent manner to Gln, resulting in increased red fluorescence in proportion to the incubation time and ligand concentration (**Figure 4B, Figure S4 & S5**). The fluorescence from live cells expressing the sensor was also examined using flow cytometry analysis, where dose-dependent increase of Q-SHINE signal was detected with increasing Gln concentration (**Figure S6**). Similarly, cell populations that fluoresce beyond the gating criteria ($>10^3$) increased with higher Gln level.

We further attempted to express Q-SHINE-GFP in tobacco leaves to determine whether the present PBP-based design strategy could be applied to the detection of plant metabolites. Since the endogenous Gln concentration in plant species, including tobacco, is within the range 2.5–20 mM,[42] which is excessive for measurement purposes, we compared fluorescence from Q-

SHINE-GFP and Q-SHINE-GFP_ΔSm. As shown in **Figure 5**, the signal from Q-SHINE-GFP clearly differed from the mCherry signal of Q-SHINE-GFP_ΔSm, confirming that Q-SHINE could detect the presence of Gln in plant cells as well.

3.6. Split design of other PBPs

To further investigate if the split design widely applies to related Type II PBPs that harbor similar structure, *Campylobacter jejuni* cysteine-binding protein (CBP)[43] and *E. coli* histidine-binding protein (HBP)[44] were tested using the same design strategy. Although QBP, CBP and HBP share homologous fold and chain topology (**Figure 6A**), the detailed modes of ligand binding are clearly different. There are 13 hydrogen bonds between Gln and QBP, nine from the large domain and four from the small domain,[32] and 12 hydrogen bonds/electrostatic interactions between cysteine (Cys) and both domains of CBP,[43,45] whereas in HBP, the small domain (domain 2) alone is unable to bind histidine (His).[34] Two linkers of different lengths were designed for each large domain of split CBP and HBP, followed by recombination with split mCherry domains (**Table S2**). CBP_L3 and CBP_L5 showed Cys-dependent fluorescence despite their limited dynamic range of the signal compared to that of Q-SHINE (**Figure 6B**). Difference of reactivity between CBP_L3 and CBP_L5 indicates that the design of an optimized linker for noncontinuous split domain is crucial in constituting the binding interface with the ligand. However, His-dependent fluorescence was not detected in both HBP_L2 and HBP_L4. Rather, there was reverse reactivity to the ligand in split HBPs. When the split CBPs were incubated with His to test if high concentrations of His affected the self-complementation of split mCherry fragments, it showed trivial effect compared to split HBPs (**Figure S7**). Thus, the binding of His to the large domain of HBP is supposed to hinder the complementation of two domains; however, the molecular mechanism needs to be studied further. Considering the higher binding affinity (30 nM) of HBP than QBP (300 nM) and CBP (100 nM), the imbalance in the binding mode of His against two domains of HBP (**Table 2**) might be accountable for the poor ligand-induced dimerization. Taken together, the results of this experiment suggest that the choice of a suitable PBP scaffold and elaborate design of split domains are critical to the successful application of our design approach for a novel split PBP-based sensor.

4. Conclusions

In this study, we described a potentially universal design strategy by transforming a QBP into an atypical Gln sensor system composed of two independent QBP domains. Q-SHINE exhibits a strong signal-to-noise ratio and a detection range ideally suited to physiological Gln concentrations. The ligand-dependent dimerization of Q-SHINE allows for the direct measurement of Gln concentration in bio-fluids, showing comparable sensitivity to and higher specificity than a conventional Gln assay kit, which would facilitate point-of-care diagnosis for several diseases. In addition, we demonstrated the performance of genetically encodable Q-SHINE for the *in situ* monitoring of intracellular Gln in mammalian and plant cells. Indeed, genetically encoded biosensors to monitor intracellular production or depletion of essential nutrients and metabolites have always provided crucial insights by enabling non-invasive diagnosis,[46] *in situ* detection of toxins,[47] and development of biopharmaceuticals[14]. To this end, the split design can be employed with other PBPs for the design of sensor systems for molecules other than Gln. We anticipate that our study will initiate the development of more sophisticated sensing systems by directed evolution of PBPs or the *de nov*o design of a two-domain ligand-binding protein for the detection of other important ligands.

Declaration of competing interest

The authors declare that they have no known competing personal or financial interests that could influence the research reported in this paper.

Acknowledgments

We would like to thank Dr. Sungdo Ha for his valuable support in starting the project. We also thank Jisung Oh and Yebin Won for their help in the experiments.

Appendix A. Supplementary data

Supplementary data to this article can be found online.

Funding sources

This work was supported by an intramural grant from Korea Institute of Science and Technology (2E31882); National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (2021R1C1C1003843).

References

[1] I. Piazza, K. Kochanowski, V. Cappelletti, T. Fuhrer, E. Noor, U. Sauer, P. Picotti, A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication, Cell. 172 (2018) 358-372.e23. https://doi.org/10.1016/j.cell.2017.12.006.

[2] H. Yu, O. Alkhamis, J. Canoura, Y. Liu, Y. Xiao, Advances and Challenges in Small-Molecule DNA Aptamer Isolation, Characterization, and Sensor Development, Angew. Chemie Int. Ed. 60 (2021) 16800–16823. https://doi.org/10.1002/ANIE.202008663.

[3] A.A. Glasgow, Y.M. Huang, D.J. Mandell, M. Thompson, R. Ritterson, A.L. Loshbaugh, J. Pellegrino, C. Krivacic, R.A. Pache, K.A. Barlow, N. Ollikainen, D. Jeon, M.J.S. Kelly, J.S. Fraser, T. Kortemme, Computational design of a modular protein sense-response system, Science. 366 (2019) 1024–1028. https://doi.org/10.1126/science.aax8780.

[4] D. Tavares, V. Maffenbeier, J.R. van der Meer, Engineering of Sensory Proteins with New Ligand-Binding Capacities, in: Handb. Cell Biosens., Springer International Publishing, Cham, 2019: pp. 1–21. https://doi.org/10.1007/978-3-319-47405-2_129-1.

[5] M.A. Dwyer, H.W. Hellinga, Periplasmic binding proteins: a versatile superfamily for protein engineering, Curr. Opin. Struct. Biol. 14 (2004) 495–504. https://doi.org/10.1016/j.sbi.2004.07.004.

[6] M.L.S. Silva, Lectin-based biosensors as analytical tools for clinical oncology, Cancer Lett. 436 (2018) 63–74. https://doi.org/10.1016/J.CANLET.2018.08.005.

[7] S.-G. Woo, S.-J. Moon, S.K. Kim, T.H. Kim, H.S. Lim, G.-H. Yeon, B.H. Sung, C.-H. Lee, S.-G. Lee, J.H. Hwang, D.-H. Lee, A designed whole-cell biosensor for live diagnosis of gut inflammation through nitrate sensing, Biosens. Bioelectron. 168 (2020) 112523. https://doi.org/10.1016/j.bios.2020.112523.

[8] W. Ko, S. Kim, H.S. Lee, Engineering a periplasmic binding protein for amino acid sensors with improved binding properties, Org. Biomol. Chem. 15 (2017) 8761–8769. https://doi.org/10.1039/c7ob02165h.

[9] L.F.C. Ribeiro, V. Amarelle, L.F.C. Ribeiro, M.-E. Guazzaroni, Converting a Periplasmic Binding Protein into a Synthetic Biosensing Switch through Domain Insertion, Biomed Res. Int. 2019 (2019) 1–15. https://doi.org/10.1155/2019/4798793.

[10] L. Lindenburg, M. Merkx, Engineering genetically encoded FRET sensors, Sensors (Switzerland). 14 (2014) 11691–11713. https://doi.org/10.3390/s140711691.

[11] K.A. Edwards, Periplasmic-binding protein-based biosensors and bioanalytical assay platforms: Advances, considerations, and strategies for optimal utility, Talanta Open. 3 (2021) 100038. https://doi.org/10.1016/J.TALO.2021.100038.

[12] J.D. Dattelbaum, J.R. Lakowicz, Optical determination of glutamine using a genetically engineered protein, Anal. Biochem. 291 (2001) 89–95. https://doi.org/10.1006/abio.2001.4998.

[13] H. Yang, M. Bogner, Y.-D. Stierhof, U. Ludewig, H+-Independent Glutamine Transport in Plant Root Tips, PLoS One. 5 (2010) e8917.

https://doi.org/10.1371/journal.pone.0008917.

[14] A. Behjousiar, C. Kontoravdi, K.M. Polizzi, In Situ Monitoring of Intracellular Glucose and Glutamine in CHO Cell Culture, PLoS One. 7 (2012) e34512. https://doi.org/10.1371/journal.pone.0034512.

[15] J.P. Hunt, R.J. Barnett, H. Robinson, M. Soltani, J.A.D. Nelson, B.C. Bundy, Rapid sensing of clinically relevant glutamine concentrations in human serum with metabolically engineered E. coli-based cell-free protein synthesis, J. Biotechnol. 325 (2021) 389–394. https://doi.org/10.1016/j.jbiotec.2020.09.011.

[16] K.N. Rajagopalan, R.J. DeBerardinis, Role of glutamine in cancer: Therapeutic and imaging implications, J. Nucl. Med. 52 (2011) 1005–1008. https://doi.org/10.2967/jnumed.110.084244.

[17] V. Cruzat, M. Macedo Rogero, K. Noel Keane, R. Curi, P. Newsholme, Glutamine: Metabolism and Immune Function, Supplementation and Clinical Translation, Nutrients. 10 (2018) 1564. https://doi.org/10.3390/nu10111564.

[18] C.T. Hensley, A.T. Wasti, R.J. DeBerardinis, Glutamine and cancer: Cell biology, physiology, and clinical opportunities, J. Clin. Invest. 123 (2013) 3678–3684. https://doi.org/10.1172/JCI69600.

[19] J.K. Byun, M. Park, S. Lee, J.W. Yun, J. Lee, J.S. Kim, S.J. Cho, H.J. Jeon, I.K. Lee, Y.K. Choi, K.G. Park, Inhibition of Glutamine Utilization Synergizes with Immune Checkpoint Inhibitor to Promote Antitumor Immunity, Mol. Cell. 80 (2020) 592-606.e8. https://doi.org/10.1016/j.molcel.2020.10.015.

[20] T.Q. Tran, E.A. Hanse, A.N. Habowski, H. Li, M.B. Ishak Gabra, Y. Yang, X.H. Lowman, A.M. Ooi, S.Y. Liao, R.A. Edwards, M.L. Waterman, M. Kong, α-Ketoglutarate attenuates Wnt signaling and drives differentiation in colorectal cancer, Nat. Cancer. 1 (2020) 345–358. https://doi.org/10.1038/s43018-020-0035-5.

[21] N.F. Butte, H.W. Hsu, M. Thotathuchery, W.W. Wong, J. Khoury, P. Reeds, Protein metabolism in insulin-treated gestational diabetes., Diabetes Care. 22 (1999) 806–811. https://doi.org/10.2337/DIACARE.22.5.806.

[22] S.Y. Rhee, E.S. Jung, H.M. Park, S.J. Jeong, K. Kim, S. Chon, S.Y. Yu, J.T. Woo, C.H. Lee, Plasma glutamine and glutamic acid are potential biomarkers for predicting diabetic retinopathy, Metabolomics. 14 (2018) 1–10. https://doi.org/10.1007/s11306-018-1383-3.

[23] H. Al-Otaish, L. Al-Ayadhi, G. Bjørklund, S. Chirumbolo, M.A. Urbina, A. El-Ansary, Relationship between absolute and relative ratios of glutamate, glutamine and GABA and severity of autism spectrum disorder, Metab. Brain Dis. 33 (2018) 843–854. https://doi.org/10.1007/s11011-018-0186-6.

[24] J. Chen, Y. Chen, G. Vail, H. Chow, Y. Zhang, L. Louie, J. Li, R.P. Hart, M.R. Plummer, K. Herrup, The impact of glutamine supplementation on the symptoms of ataxia-telangiectasia: A preclinical assessment, Mol. Neurodegener. 11 (2016) 1–14. https://doi.org/10.1186/s13024-016-0127-y.

[25] C. Madeira, C. Vargas-Lopes, C.O. Brandão, T. Reis, J. Laks, R. Panizzutti, S.T. Ferreira, Elevated Glutamate and Glutamine Levels in the Cerebrospinal Fluid of

Patients With Probable Alzheimer's Disease and Depression, Front. Psychiatry. 9 (2018) 561. https://doi.org/10.3389/fpsyt.2018.00561.

[26] D. Taylor, G. Cawley, S. Hayward, Quantitative method for the assignment of hinge and shear mechanism in protein domain movements, Bioinformatics. 30 (2014) 3189–3196. https://doi.org/10.1093/BIOINFORMATICS/BTU506.

[27] P.S. Huang, Y.E.A. Ban, F. Richter, I. Andre, R. Vernon, W.R. Schief, D. Baker, RosettaRemodel: A Generalized Framework for Flexible Backbone Protein Design, PLoS One. 6 (2011) e24109. https://doi.org/10.1371/JOURNAL.PONE.0024109.

[28] S.J. Fleishman, A. Leaver-Fay, J.E. Corn, E.M. Strauch, S.D. Khare, N. Koga, J. Ashworth, P. Murphy, F. Richter, G. Lemmon, J. Meiler, D. Baker, RosettaScripts: A Scripting Language Interface to the Rosetta Macromolecular Modeling Suite, PLoS One. 6 (2011) e20161. https://doi.org/10.1371/JOURNAL.PONE.0020161.

[29] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature. 596 (2021) 583–589. https://doi.org/10.1038/s41586-021-03819-2.

[30] J.H. Jung, F. Altpeter, TALEN mediated targeted mutagenesis of the caffeic acid O-methyltransferase in highly polyploid sugarcane improves cell wall composition for production of bioethanol, Plant Mol. Biol. 92 (2016) 131–142. https://doi.org/10.1007/s11103-016-0499-y.

[31] J.H. Lee, H.J. Won, E.S. Oh, M.H. Oh, J.H. Jung, Golden Gate Cloning-Compatible DNA Replicon/2A-Mediated Polycistronic Vectors for Plants, Front. Plant Sci. 11 (2020) 1617. https://doi.org/10.3389/FPLS.2020.559365/BIBTEX.

[32] Y.J. Sun, J. Rose, B.C. Wang, C.D. Hsiao, The structure of glutamine-binding protein complexed with glutamine at 1.94 Å resolution: comparisons with other amino acid binding proteins, J. Mol. Biol. 278 (1998) 219–229. https://doi.org/10.1006/JMBI.1998.1675.

[33] C.D. Hsiao, Y.J. Sun, J. Rose, B.C. Wang, The crystal structure of glutamine-binding protein from Escherichia coli, J. Mol. Biol. 262 (1996) 225–242. https://doi.org/10.1006/JMBI.1996.0509.

[34] B.C.H. Chu, T. De Wolf, H.J. Vogel, Role of the Two Structural Domains from the Periplasmic Escherichia coli Histidine-binding Protein HisJ *, J. Biol. Chem. 288 (2013) 31409–31422. https://doi.org/10.1074/JBC.M113.490441.

[35] M.H. Seo, J. Park, E. Kim, S. Hohng, H.S. Kim, Protein conformational dynamics dictate the binding affinity for a ligand, Nat. Commun. 5 (2014) 3724. https://doi.org/10.1038/ncomms4724.

[36] G.S. Filonov, K.D. Piatkevich, L.M. Ting, J. Zhang, K. Kim, V. V. Verkhusha, Bright and stable near-infrared fluorescent protein for in vivo imaging, Nat. Biotechnol. 29 (2011) 757–761. https://doi.org/10.1038/nbt.1918.

[37] M.B. Ishak Gabra, Y. Yang, H. Li, P. Senapati, E.A. Hanse, X.H. Lowman, T.Q. Tran, L. Zhang, L.T. Doan, X. Xu, D.E. Schones, D.A. Fruman, M. Kong, Dietary glutamine supplementation suppresses epigenetically-activated oncogenic pathways to inhibit melanoma tumour growth, Nat. Commun. 11 (2020) 1–15. https://doi.org/10.1038/s41467-020-17181-w.

[38] A.S. Dixon, M.K. Schwinn, M.P. Hall, K. Zimmerman, P. Otto, T.H. Lubben, B.L. Butler, B.F. Binkowski, T. Machleidt, T.A. Kirkland, M.G. Wood, C.T. Eggers, L.P. Encell, K. V. Wood, NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells, ACS Chem. Biol. 11 (2015) 400–408. https://doi.org/10.1021/ACSCHEMBIO.5B00753.

[39] S.S. Shekhawat, I. Ghosh, Split-protein systems: Beyond binary protein-protein interactions, Curr. Opin. Chem. Biol. 15 (2011) 789–797. https://doi.org/10.1016/j.cbpa.2011.10.014.

[40] S.A. Lim, J.A. Wells, Split enzymes: Design principles and strategy, in: Methods Enzymol., Academic Press Inc., 2020: pp. 275–296. https://doi.org/10.1016/bs.mie.2020.05.001.

[41] M. Chen, C. Yan, Y. Ma, X.E. Zhang, A tandem near-infrared fluorescence complementation system with enhanced fluorescence for imaging protein–protein interactions in vivo, Biomaterials. 268 (2021) 120544. https://doi.org/10.1016/j.biomaterials.2020.120544.

[42] V.R. Chellamuthu, E. Ermilova, T. Lapina, J. Lüddecke, E. Minaeva, C. Herrmann, M.D. Hartmann, K. Forchhammer, A Widespread Glutamine-Sensing Mechanism in the Plant Kingdom, Cell. 159 (2014) 1188–1199. https://doi.org/10.1016/J.CELL.2014.10.015.

[43] A. Müller, G.H. Thomas, R. Horler, J.A. Brannigan, E. Blagova, V.M. Levdikov, M.J. Fogg, K.S. Wilson, A.J. Wilkinson, An ATP-binding cassette-type cysteine transporter in Campylobacter jejuni inferred from the structure of an extracytoplasmic solute receptor protein, Mol. Microbiol. 57 (2005) 143–155. https://doi.org/10.1111/J.1365-2958.2005.04691.X.

[44] N. Yao, S. Trakhanov, F.A. Quiocho, Refined 1.89-Å Structure of the Histidine-Binding Protein Complexed with Histidine and Its Relationship with Many Other Active Transport/Chemosensory Proteins, Biochemistry. 33 (1994) 4769–4779. https://doi.org/10.1021/bi00182a004.

[45] H. Bulut, S. Moniot, A. Licht, F. Scheffel, S. Gathmann, W. Saenger, E. Schneider, Crystal Structures of Two Solute Receptors for l-Cystine and l-Cysteine, Respectively, of the Human Pathogen Neisseria gonorrhoeae, J. Mol. Biol. 415 (2012) 560–572. https://doi.org/10.1016/J.JMB.2011.11.030.

[46] E.C. Greenwald, S. Mehta, J. Zhang, Genetically encoded fluorescent biosensors illuminate the spatiotemporal regulation of signaling networks, Chem. Rev. 118 (2018) 11707–11794. https://doi.org/10.1021/ACS.CHEMREV.8B00333/SUPPL_FILE/CR8B00333_SI_001.XLSX.

[47] M.J. Bick, P.J. Greisen, K.J. Morey, M.S. Antunes, D. La, B. Sankaran, L. Reymond,

K. Johnsson, J.I. Medford, D. Baker, Computational design of environmental sensors for the potent opioid fentanyl, Elife. 6 (2017) e28909. https://doi.org/10.7554/eLife.28909.

[48] G.W. Foight, Z. Wang, C.T. Wei, P. Greisen, K.M. Warner, D. Cunningham-Bryant, K. Park, T.J. Brunette, W. Sheffler, D. Baker, D.J. Maly, Multi-input chemical control of protein dimerization for programming graded cellular responses, Nat. Biotechnol. 37 (2019) 1209–1216. https://doi.org/10.1038/s41587-019-0242-8.

Table 1. Intra-assay variability of Q-SHINE for measuring Gln concentrations in mouse serum.

Serum sample	Mean (μM)	SD [a)] (μM)	CV [b)] (%)
1 (N = 6) [c)]	200	5.7	2.9
2 (N = 6)	373	12.3	3.3
3 (N = 6)	537	18.7	3.5
4 (N = 6)	281	11.5	4.1

[a)] SD: Standard Deviation
[b)] CV: Coefficient of Variability
[c)] N = 6 technical repeats)

Table 2. Comparison of binding energy between the ligands and each domain of PBPs.

PBP	PDB ID	AA	ΔΔG [a)] (Lg+Sm)	ΔΔG (Lg)	ΔΔG (Sm)	Binding ratio to Lg [b)]	Binding ratio to Sm	Max/Min [c)]
QBP	1WDN	Gln	-28.457	-13.481	-14.975	0.474	0.526	1.110
CBP	1XT8	Cys	-12.410	-4.911	-7.499	0.396	0.604	1.525
HBP	1HSL	His	-11.788	-7.711	-4.077	0.654	0.346	1.890

[a)] The ligand binding energy (ΔΔG, Rosetta Energy Unit) is calculated by RosettaScript ddG mover after energy minimization by REF2015 score function.

[b)] Binding ratio is defined as binding energy ratio of ligand for a particular domain. For example, binding ratio to Lg is calculated as ΔΔG (Lg) / ΔΔG (Lg+Sm).

[c)] Max/Min is the ratio of maximum to minimum binding ratio value, which represents the extent of balance of ligand-binding by each split domain.

Scheme 1. Split design of a PBP scaffold for a novel biosensor system. (A) Based on the open and closed structures of a PBP, hinge residues were selected by analyzing the domain movement. (B) For stable expression and binding to the ligand, suitable linker sequence was designed for noncontinuous fragments. (C) The separate domains were fused with split reporter fragments, such as split fluorescent protein and luciferase. (D) Two components of the final sensor proteins were mixed with samples or genetically expressed in cells. Resulting signals (fluorescence or luminescence) represent the level of the ligand.

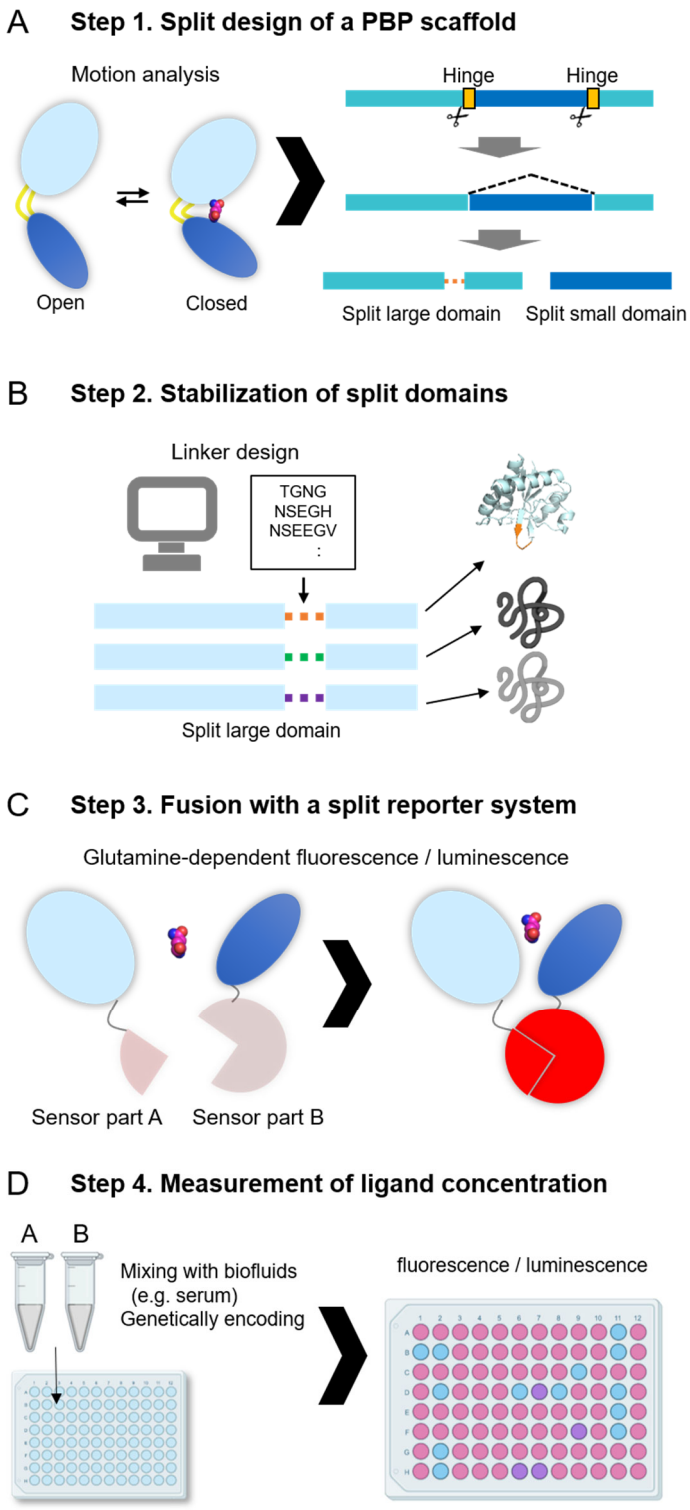


Figure 1. Design of Q-SHINE. (A) Based on the hinge-bending motion of *E. coli* QBP (PDB 1WDN), hinge residues (yellow) were selected for cleavage. One noncontinuous fragment was connected by the design of a new linker (red), resulting in two separate domains of QBP: QBP_Lg (light blue) and QBP_Sm (blue). (B) Structures of split domains were predicted by AlphaFold. When aligned with the holoprotein, two domains of QBP showed the same global fold as in native QBP. (C) Each split QBP domain was recombined with a split mCherry fragment for the construction of the complete Q-SHINE system: Split QBP_Lg – mCherry_C (LgC) and mCherry_N – Split QBP_Sm (SmN). Gln-induced heterodimerization of QBP gave rise to red fluorescence by complementation of split mCherry fragments.

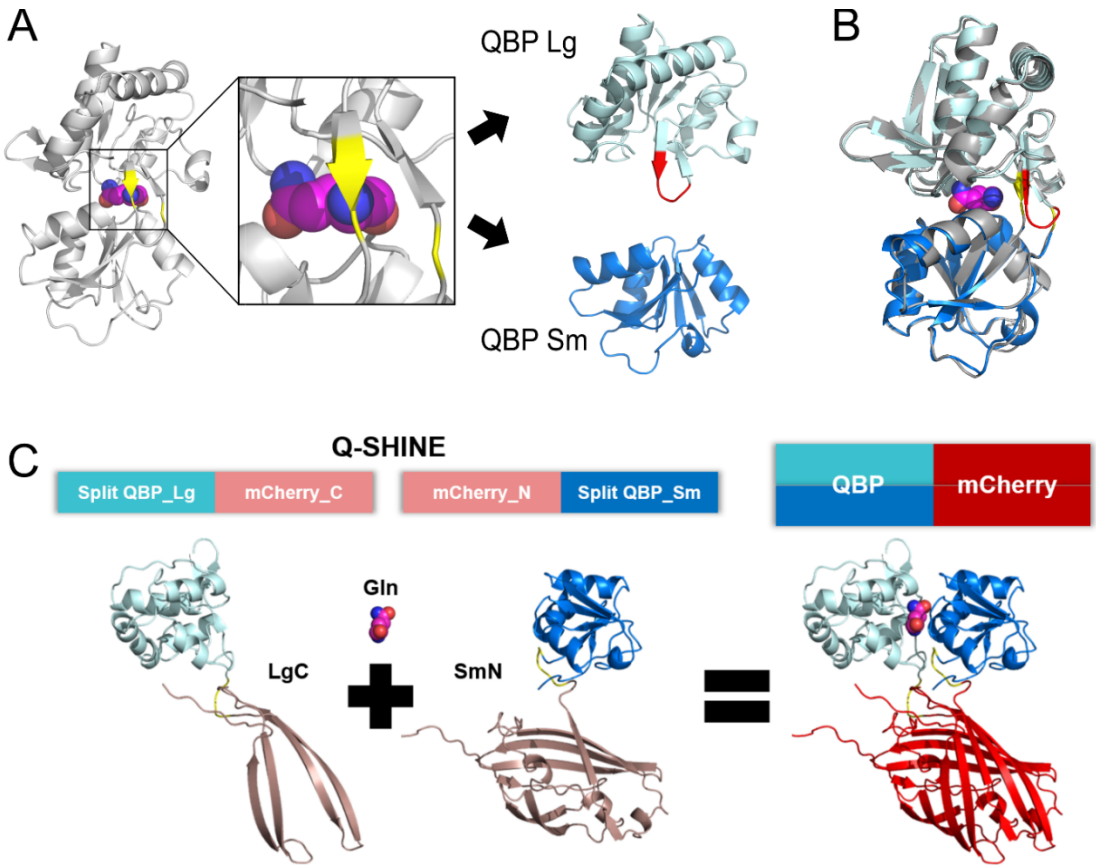


Figure 2. Functional evaluation of Q-SHINE. (A) Gln-dependent BiFC signal was measured with the marked concentration of Gln 2–94 h of incubation. Each signal intensity was normalized to the well without Gln. (B) The representative graph of Q-SHINE fluorescence scan (580–640 nm, 1 nm increment, after 20 h incubation of the mixture). Ligand selectivity of Q-SHINE was examined with other amino acids with similar structure. Reversible enzymatic construction of Q-SHINE showed similar sensitivity with BiFC-based system. The x-axis represents the corresponding Gln concentration in the reaction mixture (C, D). Data were normalized to the signal from the wells no amino acid added. Error bars represent the SD in triplicate experiments.

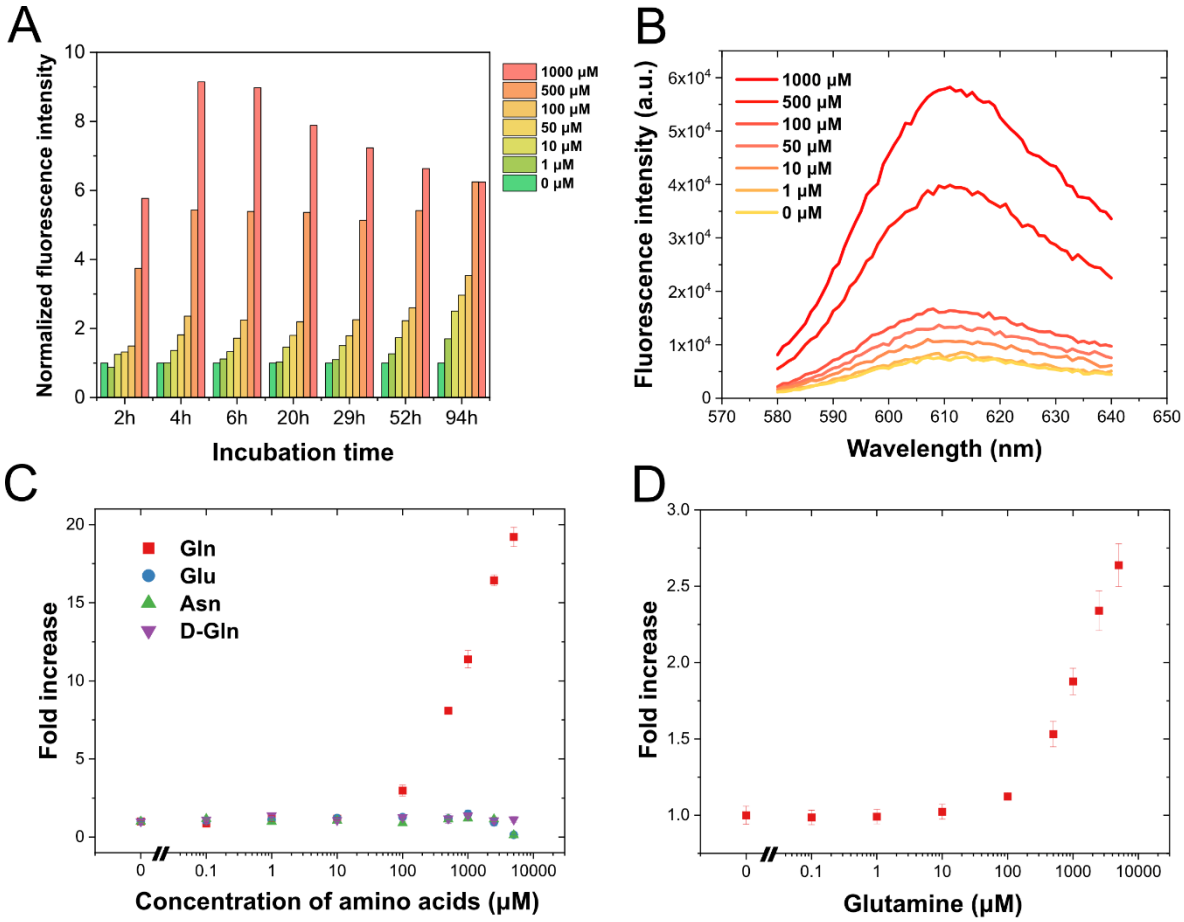
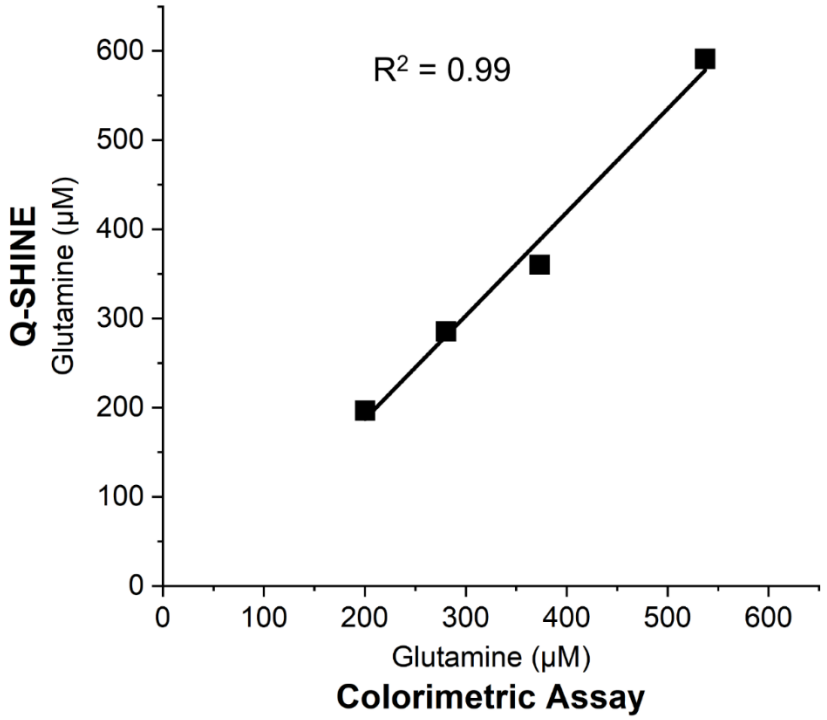


Figure 3. Comparison of Q-SHINE with a commercially available colorimetric Gln assay kit measuring Gln concentration in mouse serum.



$$[\text{L-Glutamine}]_{\text{Q-SHINE}} = 1.16 \text{ X } [\text{L-Glutamine}]_{\text{Colorimetric Assay}} - 45.1$$

Figure 4. Figure shows the changes in cytoplasmic glutamine concentrations. (A) Fluorescence microscopy of HEK293T cells transfected with Q-SHINE-GFP with increasing Gln concentrations. Overlay images from green and red channels are shown. Scale bars, 200 µm. (B) Quantification of mCherry/GFP fluorescence intensity ratio by Q-SHINE-GFP at different Gln concentrations. Gln and time-dependent increase in the fluorescence ratio by Q-SHINE-GFP was detected. The ratio of Q-SHINE-GFP_ΔSm remained unchanged. The number of ROIs ranged from 33 to 173. The levels of significance are indicated as follows: *P< 0.05; **P< 0.01; ***P< 0.001; ****P< 0.0001; n.s., not significant.

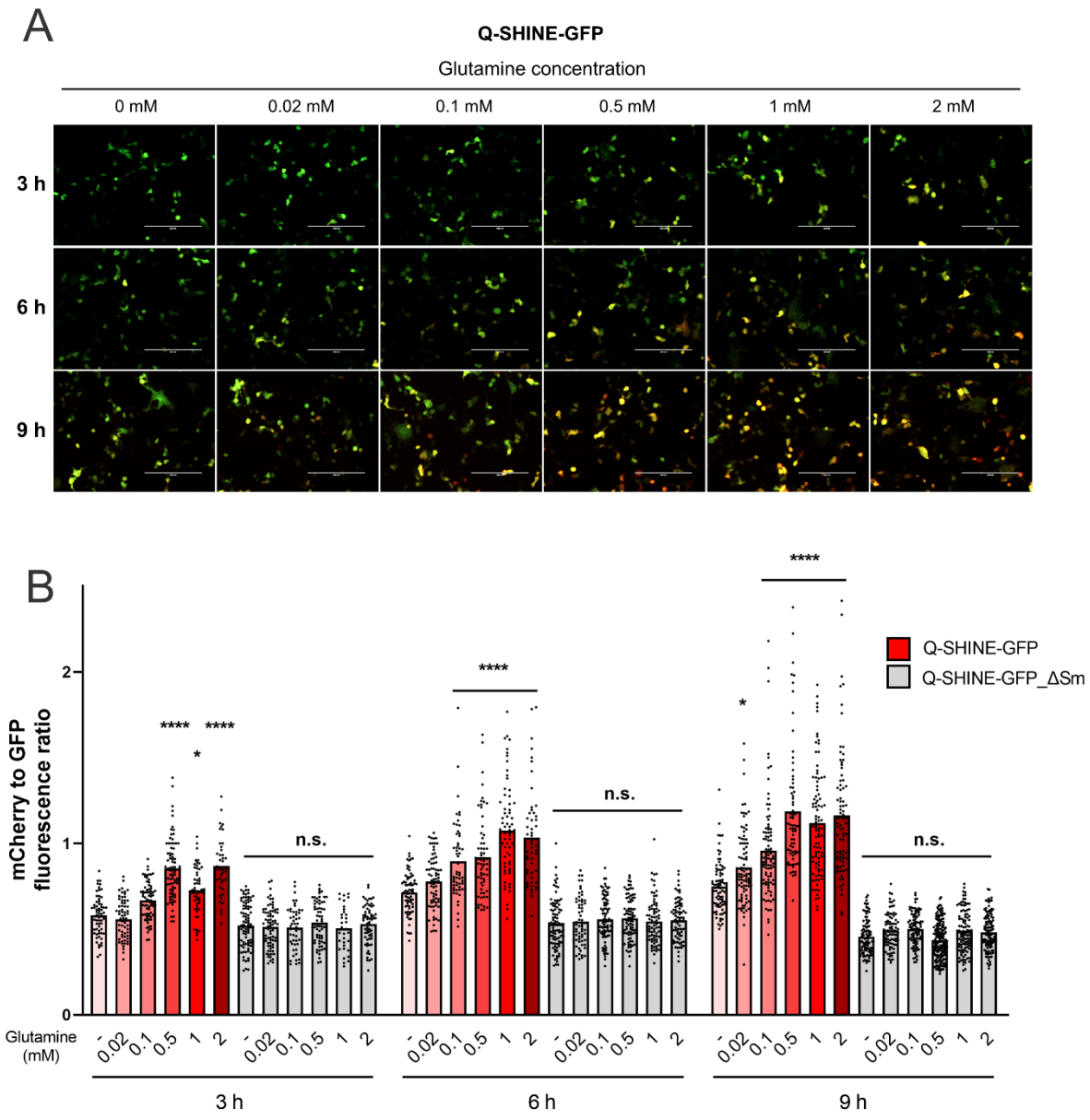


Figure 5. Detection of Gln in tobacco leaves with genetically encoded Q-SHINE-GFP. mCherry-to-GFP intensity ratio is marked. Mean of six images; three biological replicates with two technical replicates.

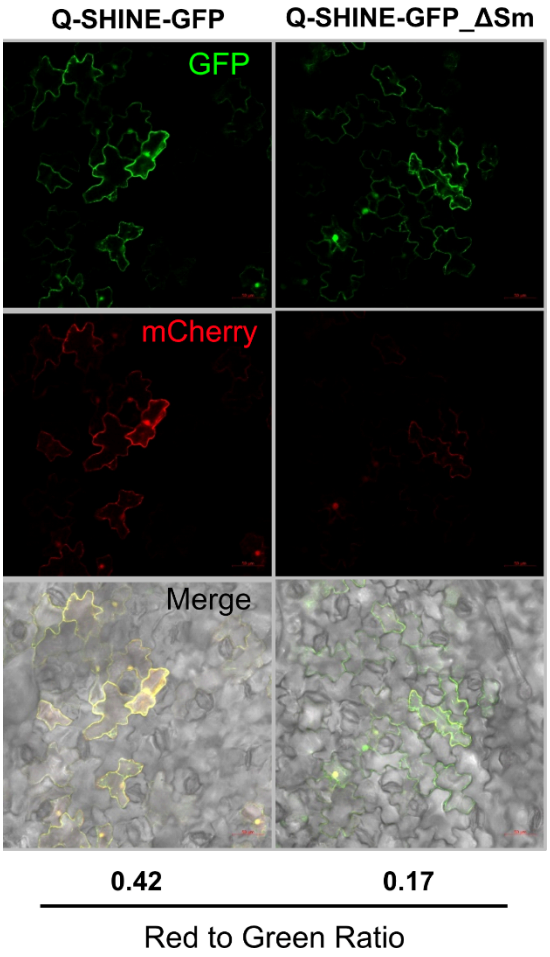
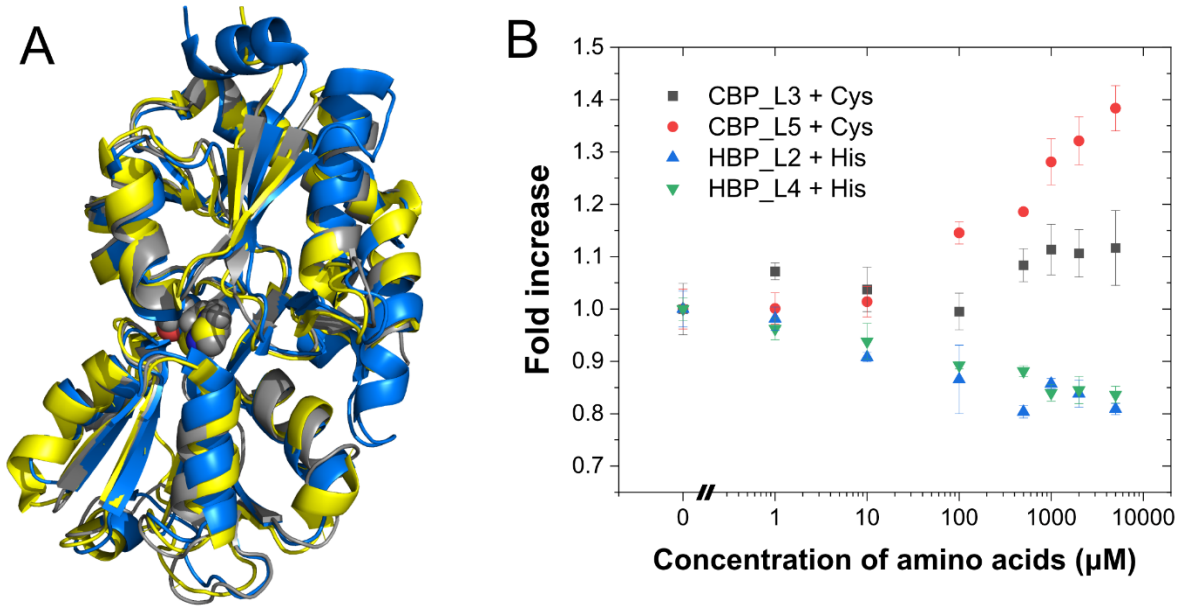



Figure 6. Split design of other PBPs for Cys and His detection. (A) Superposition of QBP (PDB ID: 1WDN; grey), CBP (PDB ID:1XT8, blue) and HBP (PDB ID: 1HSL; yellow). (B) Ligand-dependent BiFC signal from 2 pairs of CBP and HBP were measured with the corresponding concentration of marked amino acids. Results are averages ± s.d. (N = 3) of fluorescence intensity normalized to 0 μM conditions.