Acoustic Biosensors for Noninvasive Imaging of Molecular Processes

Thesis by

Zhiyang Jin

In Partial Fulfillment of the Requirements for
the Degree of
Doctor of Philosophy in Medical Engineering

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2024
(Defended April 18, 2024)

Zhiyang Jin
ORCID: 0000-0002-4411-6991

ACKNOWLEDGEMENTS

For years, I've been imagining this moment of writing the acknowledgements for my thesis and it's probably what I have been looking forward to the most since the beginning of my PhD. Finally, we are here. I really appreciate the help from everyone throughout this journey and let's work on the specifics.

For an unusual start, I would like to thank the TV show, the Big Bang Theory. Thank you for bringing me here for my PhD. I started to watch this show when I started my undergraduate studies, and the science nerds' life it described somehow really appealed to me and later became one of my core motivations to try research. It was also the reason that I was so obsessed with applying to Caltech that I went through every faculty in the department and unexpectedly ran into someone's mind-blowing research.

And that was my advisor Prof. Mikhail G. Shapiro. I emailed him the night I saw his research and he responded within 5 minutes. His responsiveness has been this way since then. I really appreciate it and that drives me to be as responsive as I can. This is just one example. In my mind, he is almost the perfect individual for science: he is smart, diligent, passionate about science, knowledgeable about many fields, great at telling stories, accessible and supportive, also very nice, good-looking, and has that ever-lasting smile on his face which probably is what made him many friends everywhere. Observing him is already a very comprehensive way to learn to be a good researcher. In addition, he is a very good mentor. He really cares about our scientific growth and spends a lot of time meeting with us for both big picture and technical details. Simply put, he is the kind of scientist that I want to become, and I really thank him for helping me make progress on that track. Thank you, Mikhail, for all the things you taught me, both in science and beyond.

Next, I want to thank all of my mentors who guided me through my scientific career. Starting with Prof. Yun Jing, I am really grateful for the opportunities he gave me to do my first research project in his lab and that is the very moment that I became convinced about having science as my life-long pursuit. I also want to thank Prof. Xiasheng Guo, Prof. Juan Tu, and Prof. Dong Zhang who later let me work in their labs for my undergraduate thesis. In

particular, Prof. Guo helped me through many difficult problems and also taught me to think independently instead of always asking for help. Even after I came to the US, we still talked a lot and he has been a mentor for both science and life throughout. I truly appreciate it and I wish to have him as my life-long friend and mentor. At Caltech, I want to thank all my committee members, Prof. Lihong V. Wang, Prof. Henry A. Lester, and Prof. Viviana Gradinaru for all their time helping me shape and build up my thesis project, especially since they are all very busy. At last, I want to thank my graduate mentor Dr. Anupama Lakshmanan. When I joined the Shapiro Lab, I didn't know anything about biology, not even what a plasmid is. Anu basically taught me everything I now know about molecular biology and made my transition from an ultrasound engineer to a bioengineer possible and enjoyable. Beyond that, she taught me many things about being a good scientist, a good mentor, and a good friend. I cannot imagine what I would be doing right now without her. Thank you, Anu.

Furthermore, I had the privilege of working with several super talented mentees who, in a way, are also my mentors. Particularly, two of them really taught me a lot. First, I want to thank Teresa. She, together with Anu, tutored me in biology. I appreciate her patience with me, especially since I was supposed to mentor her. Beyond knowledge sharing, we also shared many moments of tedious experiments together; she taught me about persistence and, more important, how to make these moments fun. I had a great time working with her and I am truly grateful. Second, I would not have enjoyed graduate school this much without working with Ruby. There are many, many things that are awesome about Ruby which I don't have the room to list here, but the most impressive to me was her showing me how productive one could ever be, with zero tolerance of procrastination. I believe it would be one of my life goals to become like her in this respect. I am grateful for the two years with her and I genuinely hope it could be longer. Thank you, Ruby! I also want to thank all of my other amazing mentees: Jee Won Yang, Stefan Wan, Akanksha Yadav, Queen Igomu, and Rosmelle Meglar. Thank you for working with me; it was fun!

Then I want to thank my friends and colleagues in the Shapiro Lab. First, to my lovely officemates, Dr. Ernesto Criado Hidalgo, Dr. Claire Rabut, Jee Won Yang, and Dr. Yuxing Yao, thank you for making it so much fun to go to a workplace. Then, to all the friends who

spent much of their valuable time talking about random things with me, namely Tom Duan, Ishaan Dev, Abdullah Farooq, Justin Lee, Shirin Shivaei, Di Wu, Bill Ling, Nivin N. Nyström, Hao Shen, Josh Chen, Ahyoung Kim, Daniel Tang, Alen Pavlič, George Daghilan, Ann Liu, Erik Schrunk, Mei Yi You, Przemysław Dutka, Maria Paulene Abundo, Rob C. Hurt, Avinoam Bar-Zion, Mohamad Abedi, Danny Sawyer, Arash Farhadi, David Mittelstein, and everyone else in the Shapiro Lab. Thank you so much for making this lab one of my favorite place despite somewhat painful experiences (just kidding). Lastly, to many lab alumni who helped me even when they left the lab, Prof. Amir Salahshoor, Dr. Lealia Xiong, Prof. George Lu, Prof. Jerzy Szablowski, Dr. Hunter Davis, Prof. David Maresca, Dr. Pradeep Ramesh, Dr. Dan Piraner, and Prof. Arnab Mukherjee. Thank you for your support and making the Shapiro Lab feel like a big family.

Now it's time for people outside the work-related context. I have made many life-long friends here at Caltech and I value that as one of the most important things throughout my PhD. I want to thank my roommate Xinhong Chen. We shared many laughs and tears together and that made the time through graduate school much easier. I want to thank Yuxing Yao for always being supportive when I keep talking and talking about my pessimistic feelings, along with all the nights we spent in the office and his car (talking about science, of course). I want to thank Ruby, Teresa, Di, and Mei Yi for all the adventures we had outside the lab. I want to thank Kun Miao, Shumao Zhang, Yuelin Shi, Jiajun Du, Peiwei Chen, Jialiang Lu, William Kuo, Yu-Wei Wu, and others for all the time we spent together. Caltech is a special place in my mind not just because I did my PhD here, but also because I met all of you folks here. That matters much more than just a PhD.

Further away from work, I appreciate all the support from friends across the globe. I thank many of you who constantly come and check my sanity, especially Lisha Huo, Xikun Chen, Danjie Sun, Fan Gao, Yuwei Yang, Chengxin Heng, Tianyang Long, Han Bao, and Zhenyuan Xi. Thank you for always being there when I reached out in weird hours. I had to miss many important life moments of yours because of this degree, but I really wish that we could spend some more time together so I can make it up in the future. Thank you for your support and your understanding.

Now, I want to thank someone who is both a friend and a family member to me, Yan Xu. Thank you for staying with me for almost 15 years. Even though we are upset with each other all the time, I would not dare to imagine a day without you; you have been an indispensable part of my life. The past 7 years of my PhD were not easy, but I always feel secure whenever I think of you. Thank you for everything you did for me. The life ahead doesn't look any easier either, but it doesn't matter, because we will be together.

Lastly but the most importantly, I would like to dedicate this thesis to my family: my parents Ningjiang Zhao and Hui Jin, my aunt Huajiang Zhao, as well as my grandparents Haoming Zhang, Xiuyi Zhao, Fuzhen Lin, and Wanda Jin. Thank you all for your unconditional love and support since I came into this world. I cannot imagine how much effort it took to raise me, and I don't think I can ever pay it back, but I will do everything I can to love you back.

我爱你们呀。

ABSTRACT

Understanding biology in its native context has been a major scientific endeavor. Yet, it is challenging to visualize cellular dynamics at the molecular scale in the context of a living organism at the macroscopic scale. Ultrasound imaging represents a promising candidate to address this challenge, with its unique advantages of large imaging volume, deep penetration, and good spatiotemporal resolution. However, ultrasound was historically limited in retrieving molecular information that biology carries. Until very recently, the discovery of the first ultrasound-interacting biomolecules, gas vesicles (GVs), established a connection between connect cellular function and ultrasound signals, which later enabled ultrasound imaging of gene expression and thus the location of GV-expressing cells. Going beyond location tracking, this thesis describes the engineering of GV-based acoustic biosensors that made it possible to noninvasively image the dynamics of cellular signaling in living organisms.

GVs are genetically encoded intracellular air-filled “balloons” that are encapsulated by protein shells. The acoustic biosensor design leverages the GV surface protein GvpC, which controls GVs' ultrasound scattering by setting the stiffness of their protein shell. We developed the first acoustic biosensors by engineering GvpC to change its confirmation and thereby GVs’ ultrasound contrast in response to the activity or concentration of specific molecules. Specifically, we first built the biosensors for three different types of enzymes and demonstrated noninvasive imaging of enzyme activity inside probiotic cells in the mouse colon in vivo. Next, we engineered the acoustic biosensors for calcium, a ubiquitous signaling molecule that is essential in many cellular processes (e.g., neural activity). With the first generation of this calcium sensor for ultrasound, we demonstrated imaging of receptor-specific calcium signaling deep inside the mouse brain through the intact skull noninvasively, which opened up the possibility of whole-brain neuroimaging that can lead to many breakthroughs in neuroscience. Last, we established a high-throughput engineering platform to develop all these GV-based imaging agents in a much shorter time frame. Collectively, this thesis presents the first demonstration of noninvasively imaging dynamic cellular signaling with acoustic biosensors and the feasibility of efficiently improving them for potential real-world applications with our engineering pipeline, opening up a new route towards understanding biology across scales.

PUBLISHED CONTENT AND CONTRIBUTIONS

Jin, Z., Lakshmanan, A., Zhang, R., Tran, T.A., Rabut, C., Dutka, P., Duan, M., Hurt, R.C., Malounda, D., Yao, Y., et al. (2023). Ultrasonic reporters of calcium for deep tissue imaging of cellular signals. *bioRxiv*. DOI: 10.1101/2023.11.09.566364.

Z.J. conceived the study, conducted experiments, analyzed the data, and prepared the manuscript.

Hurt, R.C.[#], **Jin, Z.**[#], Soufi, M., Wong, K., Sawyer, D., Deshpande, R., Shen, H., Mittelstein, D.R., Shapiro, M.G. Directed evolution of acoustic reporter genes using high-throughput acoustic screening. [#]**Equal contribution**. *bioRxiv*. DOI: 10.1101/2024.03.30.587094.

Z.J. conceived the study with R.C.H., conducted experiments, analyzed the data, and contributed to manuscript preparation with a focus on the development of the acoustic plate reader.

Lakshmanan, A.[#], **Jin, Z.**[#], Nety, S.P., Sawyer, D.P., Lee-Gosselin, A., Malounda, D., Swift, M.B., Maresca, D., and Shapiro, M.G. (2020). Acoustic biosensors for ultrasound imaging of enzyme activity. *Nature Chemical Biology* 16, 988–996. [#]**Equal contribution**. DOI: 10.1038/s41589-020-0591-0.

Z.J. conceived the study with A.L., conducted experiments, analyzed the data, and contributed to manuscript preparation with a focus on the ClpXP biosensors and intracellular sensing.

Hurt, R.C., Buss, M.T., Duan, M., Wong, K., You, M.Y., Sawyer, D.P., Swift, M.B., Dutka, P., Barturen-Larrea, P., Mittelstein, D.R., **Jin, Z.**, et al. (2023). Genomically mined acoustic reporter genes for real-time in vivo monitoring of tumors and tumor-homing bacteria. *Nature Biotechnology*, 1–13. DOI: 10.1038/s41587-022-01581-y.

Z.J. participated in data collection.

Salahshoor, H., Yao, Y., Dutka, P., Nyström, N.N., **Jin, Z.**, Min, E., Malounda, D., Jensen, G.J., Ortiz, M., and Shapiro, M.G. (2022). Geometric effects in gas vesicle buckling under ultrasound. *Biophysical Journal* 121, 4221–4228. DOI: 10.1016/j.bpj.2022.09.004.

Z.J. participated in data collection and analysis.

Yao, Y., **Jin, Z.,** Ling, B., Malounda, D., and Shapiro, M.G. (2021). Self-assembly of protein superstructures by physical interactions under cytoplasm-like conditions. *Biophysical Journal* 120, 2701–2709. DOI: 10.1016/j.bpj.2021.05.007.

Z.J. participated in building computational models, data collection, and analysis.

Rabut, C., Wu, D., Ling, B., **Jin, Z.**, Malounda, D., and Shapiro, M.G. (2021). Ultrafast amplitude modulation for molecular and hemodynamic ultrasound imaging. *Applied Physics Letter* 118, 244102. DOI: 10.1063/5.0050807.

Z.J. participated in experiment design and data collection.

Rabut, C., Yoo, S., Hurt, R.C., **Jin, Z.**, Li, H., Guo, H., Ling, B., and Shapiro, M.G. (2020). Ultrasound technologies for imaging and modulating neural activity. *Neuron* 108, 93–110. DOI: 10.1016/j.neuron.2020.09.003.

Z.J. participated in the writing of the manuscript and the preparation of figures.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Number *Page*

LIST OF TABLES/OTHER SUPPLEMENTARY MATERIAL

Number *Page*

Chapter 1

MOLECULAR IMAGING OF BIOLOGICAL PROCESSES

Sections of this chapter have been adapted from:

Rabut, C., Yoo, S., Hurt, R.C., **Jin, Z.**, Li, H., Guo, H., Ling, B., and Shapiro, M.G. (2020). Ultrasound technologies for imaging and modulating neural activity. *Neuron* 108, 93–110. DOI: 10.1016/j.neuron.2020.09.003.

1.1 Why and What to Image

Observation has always been essential in any scientific research and breakthrough can often be traced to new methods that enable better observation. For example, the invention of microscopes enabled observation of cork's microscopic structures that were described as "cells" which originated the current use of this work in biology[1]. Since then, researchers have been exploring methods to observe biology at different scales, from the molecular scales below nanometers to the macroscopic scales for the entire organism. Many breakthroughs can be traced to the development of these methods, especially ones that focus on the microscopic scale, as one major endeavor in understanding biology is to observe the activity and interaction of molecules and correlate that with biological functions. In the ideal case, we would want to observe both the spatial and temporal dynamics of these molecules in real time, and the technology for this is termed molecular imaging. In addition, it is desirable to image these molecular processes in their native context, usually at the scale of a whole organism, and noninvasively, with minimal disturbance to the normal physiology. Thus, noninvasive molecular imaging at scale is extremely valuable for studying biology. Furthermore, the fruits of molecular imaging research go beyond basic understanding and have a direct effect on medicine. It can enable direct detection of molecular changes caused by diseases, usually before the anatomic and physiologic changes, which then would allow intervention at an earlier stage often leads to better outcomes. Additionally, these changes can be tracked along the course of therapy to inform its effect and guide future treatment.

To implement molecular imaging, we need to (1) identify our targets out of the countless molecules in biological entities and develop probes that can (2) report the molecular information while (3) being precisely delivered to the targets. First, the targets range from small molecules, proteins, to genetic elements, basically anything that correlates to the biological functions of interest, which are being actively screened for. Some of the popular ones include signaling molecules in the extracellular space (e.g., hormones), cell surface markers (e.g., tumor antigens), signaling molecules inside the cells (e.g., calcium ions) and gene expression patterns. Second, we would need probe-imaging pairs that can transform the information (e.g., location, concentration, and activity) of the targets to properties that can read out. Lastly, for the molecules that are accessible outside the cells, probe can be designed to target them directly, while for those intracellular molecules, we need to engineer probes that can specifically recognize what happens inside the cells. As the first (target identification) and the third (delivery) steps are also open questions for other technologies (e.g., drug delivery), this thesis would focus on the second step of developing better molecular imaging technologies.

Now that it is established what tools we need, in this chapter we will first review the existing technologies for molecular imaging and their advantages and disadvantages. Next, we will focus on ultrasound as a promising modality for this purpose and discuss the current state-of-art approaches and their limitations, from using synthetic contrast agents to acoustic biomolecules. Then, we will look into the fundamental features of these acoustic biomolecules that enable us to develop acoustic biosensors for molecular ultrasound imaging. At last, we will go through the organization of this thesis — how we developed these sensors and took a significant leap towards noninvasive molecular imaging at scale.

1.2 Current Technologies for Molecular Imaging

Established molecular imaging techniques include optical imaging, nuclear imaging, magnetic resonance imaging (MRI), photoacoustic imaging, and ultrasound imaging. We will briefly discuss each of them below.

1.2.1 Optical Imaging

Light, dated back to the first microscope, has been the primary read out of molecular imaging, including fluorescence, absorption, and luminescence[2]. A number of chemical dyes[3] and protein-based probes[4] have been developed, and enable major discoveries in biology, in conjunction with advances in imaging methods such as confocal imaging[5], multi-photon imaging[6] and intravital microscope[7]. With all these, optical molecular imaging, such as those based on fluorescence readout, usually provides good specificity, outstanding resolution (sub-micron), and high multiplexity (multi-color imaging). For example, the green fluorescent protein (GFP) and other fluorescent proteins (FPs) are widely applied to image gene expression and a broad spectrum of biosensors derived from FPs enable imaging of important cellular dynamics such as calcium signaling[4]. Another major advantage of these protein-based probes is that they can be targeted to specific cell population, allowing the study of genetically defined cell types. However, while both the probes and the optical methods continue to expand their coverage and sensitivity, it is challenging to implement most of these methods in intact, non-transparent organisms due to the light scattering in biological tissue[8], limiting their usage to very shallow or surgically accessed regions. One exception is luminescence imaging, where light is generated from chemical reaction inside the target region and can be detected outside the body with high specificity and sensitivity due to the low background[9]. However, the scattering issue persists when the light travels out of the body and it dramatically reduces the spatial resolution to the scale of mm. It also requires the chemical substrates for the light emission reaction, and homogenous delivery to certain body region can be difficult.

1.2.2 Nuclear Imaging

Nuclear imaging, such as positron emission tomography (PET) and single photon emission computed tomography (SPECT), primarily relies on the detection of radio-active probes that are either concentrated in the areas with target molecules or turned on by genetically encoded enzymes[2]. It provides both noninvasive access to the whole organisms and molecular specificity. The in vivo molecular-specific information nuclear imaging provides has the potential for early diagnostics, enhancing therapy efficacy and thereby impacting patient

survival. Indeed, it has been applied in the clinics for oncology[10]. However, the resolution for nuclear imaging is not ideal. The spatial resolution is usually limited to millimeter scale[11] and the temporal resolution is typically in the order of seconds or even longer[12]. In addition, the requirement for short-shelf-life radio-active probes and sophisticated imaging instruments make it less accessible to researchers and most of the hospitals, which still needs to be addressed.

1.2.3 MRI

MRI can image biological tissue with unlimited depth penetration and decent spatiotemporal resolution. It has been widely applied in the clinical settings to probe anatomic changes and measure hemodynamics to study functional neural activity[13,14]. It was historically limited in retrieving specific molecular information until the recent development of a field called molecular functional MRI (fMRI), where molecular signals such as neurotransmitter activity is coupled to MRI contrast[15]. However, at this stage, these probes require fairly high concentrations, are difficult to encode genetically and involve acute injection into the target tissue[15].

1.2.4 Photoacoustic Imaging

Photoacoustic imaging is a hybrid imaging modality with light excitation and ultrasound readout[16]. The target is illuminated by very short laser pulse (nanoseconds), where the light is absorbed and converted to heat, leading to a transient expansion of the tissue that generates a local pressure rise and subsequently propagating sound waves. The acoustic signals are used to inform how much light is absorbed by either endogenous chromophores, such as hemoglobin, melanin and lipid, or exogenous contrast agents, including synthetic dye, nanoparticles, photo-sensitive proteins[17,18]. Since sound wave scatters much less in tissue than light, this method can form images solely based on the received ultrasound signals while relying on diffusive light for excitation. It combines molecular specificity of light absorption and the good spatial resolution at depth, making it a promising candidate for noninvasive molecular imaging. Although it has been used in the preclinical and clinical settings[19], it requires sophisticated equipment (e.g., pulsed laser and tomographic ultrasound arrays).

Engineering genetically encoded probes for photoacoustic imaging is also an active field of research[17,18], where strong endogenous absorption presents some obstacles.

1.2.5 Ultrasound Imaging

Ultrasound (US) is an oscillating mechanical wave with wavelengths on the order of 100 µm and penetrates tissues at the scale of centimeters[20] (**Fig. 1-1a, b**). This allows US-based imaging methods to have corresponding resolution and the temporal resolution can go below 1 ms due to the speed of sound at approximately 1500 m/s in soft tissues. The combination of high spatiotemporal resolution, deep penetration, and other features (e.g., safe and accessible) have made US imaging one of the most widely used technologies in the clinics. US imaging is usually a pulse-echo technique involving the transmission of brief pulses of US into tissue and recording the backscattered echoes from objects and interfaces within the tissues. The relative timing of the transmitted pulses and received signals is used to locate objects in space and form an image. Scattering arises from materials with different density and/or compressibility relative to their surrounding medium, including tissue interfaces, blood cells and contrast agents. This regime has been used for anatomic imaging (**Fig. 1-1c**), as well as measuring hemodynamics through the Doppler effect[20] (**Fig. 1-1c**). In addition, ultrasound is also applied to visualize the viscoelastic properties of the tissues in the clinics[21]. Thus, ultrasound represents an excellent imaging modality for noninvasive molecular imaging, and we will cover the existing technologies for molecular ultrasound imaging in the next sections.

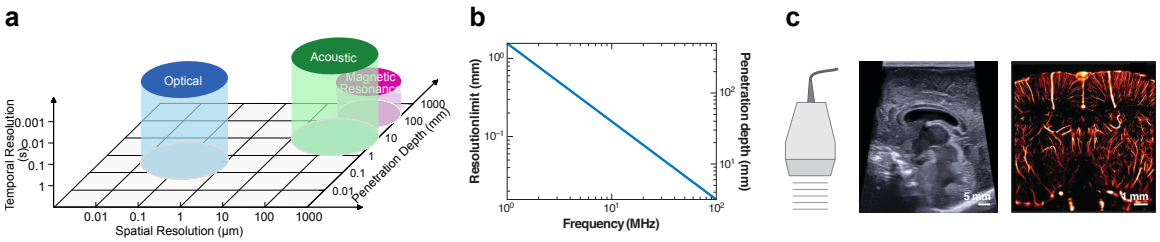


Figure 1-1. Ultrasound imaging. (**a**) Approximate performance characteristics of common imaging modalities. (**b**) Tradeoff between ultrasound resolution and penetration depth as a function of frequency in brain tissue. (**c**) Illustration of ultrasound imaging capabilities; conventional B-mode image of an infant brain with a submillimeter resolution of cerebral structures; 15-MHz super-resolution ultrasound image of the rat brain vasculature. Panel (**a**) adapted from Rabut *et al.*[22] and panel (**b-c**) adapted from Maresca *et al.*[20].

1.3 Synthetic Ultrasound Contrast Agents

As mentioned above, ultrasound measures mechanical properties and many molecular changes would not result in changes in tissue mechanics. Thus, we need probes, or contrast agents to convert molecular signals to properties ultrasound can image. In this section, we will focus on synthetic contrast agents and their applications in molecular imaging. Since the first ultrasound contrast agents were invented in the 1980s, almost all following ones are in a form of encapsulated bubbles. The dominant type of ultrasound contrast agents are called microbubbles: micro-sized bubbles of gas stabilized by a lipid or protein shell[23]. Other types include phase-change nanodroplets that can evaporate into bubbles upon ultrasound activation and sub-micron-sized nanobubbles[24]. Regardless, due to their high compressibility and resonant behaviors, they produce distinct signals from tissues that can be specifically extracted through certain ultrasound pulse sequences, such as amplitude modulation and pulse inversion[25]. In particular, most tissues and endogenous cells are considered as linear scatterers, where the backscattered signals are only at the frequencies of the transmitting sound wave and scale linearly with the transmitting acoustic pressure. While in the case of contrast agents, they demonstrate nonlinear behaviors both in frequency and amplitude domain, where they radiate harmonic frequencies (multiple times of the transmitting frequencies) and demonstrate nonlinear relationship between backscattered signals and transmitting pressure. With these synthetic contrast agents and imaging methods to differentiate them from normal tissues, molecular ultrasound imaging has been implemented with targeting contrast agents, where localized and lasting signals of contrast agents are applied to visualize the presence of the target molecules[24]. Although it has been applied in the clinics for cancer diagnostics[26] and nano-sized contrast agents have been developed to enhance accumulation at the tumor site[24], this has been the only possible method for molecular ultrasound imaging, limiting it only to detecting whether a molecular is there or not in the extracellular space or cell surface. Can ultrasound be used to image the molecular activity in a group of genetically defined cells? Until recently, this idea was considered very unlikely, because it was very difficult to imagine how synthetic contrast agents can be targeted to certain genotypes, internalized and sense dynamic signals. In the next section, we

will talk about the breakthrough in biomolecular engineering that enabled further development of molecular ultrasound imaging — genetically encoded acoustic biomolecules.

1.4 Gas Vesicles as Ultrasound Contrast Agents and Reporter Genes

Looking at the path other modalities have taken, there needs to be a protein that can lead to detectable contrast for genetic targeting and imaging intracellular signals. For example, many fluorescent probes are derived from FPs as genetically encodable fluorophores. Hence, we need proteins that can produce ultrasound contrast. Fortunately, an acoustic biomolecule of this type was identified in 2014 and recently adapted as reporter genes. The "GFP for ultrasound" is based on a unique class genetically encoded air-filled protein nanostructures known as gas vesicles or GVs, which evolved in aquatic photosynthetic microbes as a means to achieve buoyancy[27,28]. GVs comprise a cone-tipped cylindrical compartment of gas with a typical diameter of ~80 nm and length of 500 nm, enclosed by a 2-nm thick protein shell (**Fig. 1-2a, b**)[29–31]. The gas contents of GVs are in fast equilibrium with surrounding media, meaning that GVs are typically filled with air[28]. This unique gas core of GV produces ultrasound contrast because the air has dramatically different mechanical properties from the surrounding water (**Fig. 1-2c**). GVs are encoded in 8-12 genes and their protein shell mostly comprises a crystalline 2D arrangement of a single protein, GvpA, reinforced by an optional external surface protein, GvpC (**Fig. 1-2d, e**)[30,31]. The other genes encode minor structural components, chaperones, or other essential "assembly factors"[27,28]. The purified GVs produce bright backscattered ultrasound contrast[32], and specialized ultrasound imaging paradigms have been developed to detect GVs with maximal sensitivity and specificity[33–35]. As contrast agents, GVs are extremely stable and thus provide much longer circulation time than synthetic microbubbles. This feature makes GVs better in enhancing hemodynamic imaging for neuroimaging than conventional contrast agents[36]. It was also easy to engineer with all the available protein engineering strategies. For example, GVs were modified to target tumor cells[37,38] and theoretically any protein-based targeting motifs can be conjugated on GVs. Furthermore, miniaturized GVs with the size of ~50 nm were developed as the smallest ultrasound contrast agents for theragnostics with their capabilities of extravasating and accumulating in the tumor region[38].

In addition to all the advantages that purified GVs offer as contrast agents, all the components required to make GVs are encoded in those 8 genes, so they can be made inside cells of interest as reporter genes. To date, GV expression has been imaged with ultrasound in tumors[39,40], GI-resident bacteria[41], and tumor-homing bacteria[40]. Both the expression of GVs and their imaging with US is well-tolerated by the cell types tested. In mammalian cells, GV are detectable by US when they occupy < 0.1% of the cytoplasm. This genetic encodability also means that now it is possible to have genotype-specific probes for molecular ultrasound imaging. Altogether, applying GVs as acoustic reporter genes enables real-time tracking of gene expression in specific cell population, which can potentially lead to breakthroughs in neuroscience, immunology, regenerative therapy, and many other fields.

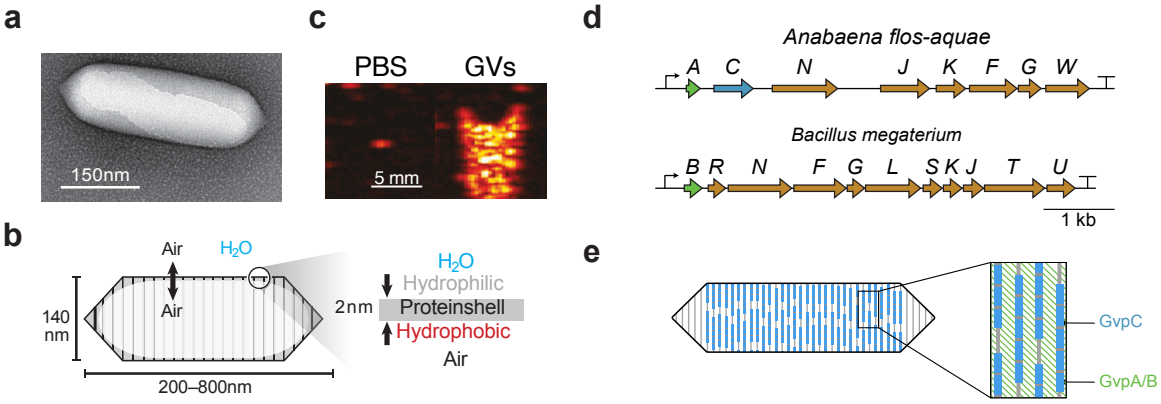


Figure 1-2. Gas vesicles for ultrasound imaging. (**a**) Transmission electron micrograph of individual GV. (**b**) Illustration of GV structure. (**c**) Ultrasound images of purified GVs. (**d**) Examples of gene clusters encoding GVs. (**e**) Illustration of GV structure protein (GvpA/B) and optional surface protein GvpC. Panel (**a-b**) adapted from Maresca *et al.*[20], panel (c) adapted from Shapiro *et al.*[32], and panel (d-e) adapted from Hurt and Jin *et al.*[42].

1.5 Gas Vesicle's Acoustic Properties Mediated by its Surface Protein

With acoustic reporter genes, we now can track the gene expression, so we set out for something even more exciting for molecular. We need to see the real-time dynamics of different molecules — their concentrations and activity — to dissect the intricate network of biology. For example, with the neural circuits for the feelings of rewards functioning at the time scale of seconds, could we capture its underlying molecular processes with ultrasound noninvasively? We believe engineering GVs to change its ultrasound contrast is one of the

most promising ways to achieve this and the molecular structure of GVs holds the key to our journey to make the acoustic biosensors.

As mentioned above, GVs from some organisms have a class of surface proteins GvpC. It is predicted to be a rigid alpha helix (**Fig. 1-3a**) and the function of GvpC is to mechanically strengthen the GV shell, making it stiffer and more resistant to deformation under pressure[37]. Correspondingly, when the GvpC is removed either chemically or genetically, the GV shell becomes softer and would reversibly buckle under acoustic pressure[37,43] (**Fig. 1-3b**). Thus, the state of the GvpC-GV interaction dictate GVs' mechanical properties. This change in mechanics can be read out by ultrasound: the stiff GVs with GvpC behave similar to linear scatterers under ultrasound, while the soft, buckling GVs without GvpC demonstrate nonlinear behavior both in amplitude and frequency, similar to microbubbles[33]. These findings make GvpC the molecular handle that we can use to couple molecular information to ultrasound readout and serve as the cornerstone for the acoustic biosensors — the main topic of this thesis.

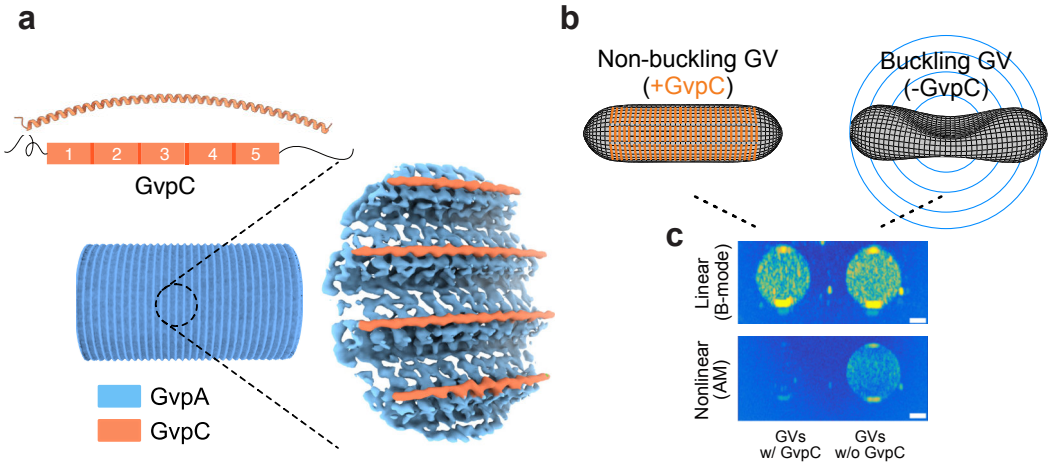


Figure 1-3. GvpC mediates GVs' acoustic properties. (**a**) (Upper) GvpC has homologous repeats and is predicted to be an alpha helix. (Bottom) Cryo-EM density map showing the structure of GV with GvpA and GvpC. (b) Illustration of GvpC mediating GVs' acoustic properties. (c) Linear and nonlinear (amplitude modulation) images of GVs with and without GvpC. Panel (**a**) adapted from Dutka *et al.*[44], and panel (c) adapted from Maresca *et al.*[33].

1.6 Thesis Organization

After all the introduction, we will go through the outline for this thesis. With the GvpC-GV interaction connected to GVs' ultrasound contrast, we set out to design our acoustic biosensors by engineering GvpC to change its binding to GVs in response to certain analytes. In the *Chapter 2*, we will talk about how we made our first proof-of-concept biosensors sensing proteases. Proteases are enzymes that catalyze the hydrolysis of the peptide bonds — cut proteins into pieces. We chose proteases because they are important but also this cleavage is significant modification to GvpC, and we hypothesized that it would be enough to induce changes in GVs' acoustic properties. After confirming the possibility of using GvpC to build acoustic biosensors, we went ahead and built the acoustic biosensors for calcium with a more generalizable design principle — engineering GvpC to have analyte-dependent, reversible conformational changes. In the *Chapter 3*, we will go into the details about how we implemented this idea and used the calcium sensors to image receptor-specific calcium signaling in the mouse brain through the intact skull *in vivo*. After this journey of making acoustic biosensors, we also realized that we need better engineering platform for GVs and GV-based biosensors, so in the *Chapter 4*, I will introduce the high-throughput screening platform we developed and demonstrate how powerful it is through a showcase of applying directed evolution to screen for acoustically "brighter" GVs. At last, I will summarize everything we learned through the process and take my chance to imagine the future about acoustic biosensors, GVs, and biomolecular ultrasound in general. Hope you enjoy this thesis and have fun!

Chapter 2

ACOUSTIC BIOSENSORS FOR IMAGING ENZYME ACTIVITY

Lakshmanan, A.[#], **Jin, Z.[#]**, Nety, S.P., Sawyer, D.P., Lee-Gosselin, A., Malounda, D., Swift, M.B., Maresca, D., and Shapiro, M.G. (2020). Acoustic biosensors for ultrasound imaging of enzyme activity. *Nature Chemical Biology* 16, 988–996. [#]**Equal contribution**. DOI: 10.1038/s41589-020-0591-0.

This chapter is a reformatted version of the above manuscript. My contributions to the work were designing and performing the study, analyzing the data, and preparing the manuscript, with a main focus on the ClpXP biosensors and intracellular sensing *in vitro* and *in vivo*.

2.1 Introduction

Virtually every biological process in living organisms involves dynamic changes in the concentration or activity of specific molecules. Visualizing these changes within the context of intact living tissues is critical to expanding our understanding of biological function and developing next-generation medicines. A large repertoire of genetically encoded fluorescent sensors has been developed to image specific molecular and cellular events[45–48]. However, deploying such biosensors in living organisms is challenging due to the limited penetration of light in tissue[49]. In contrast, non-invasive techniques such as ultrasound are capable of imaging deep tissues with high spatial and temporal resolution (below 100 μm and 1 ms, respectively)[20]. However, ultrasound currently lacks the sensors needed to observe dynamic molecular activity.

Here, we introduce molecular biosensors for ultrasound based on gas vesicles (GVs), a unique class of air-filled protein nanostructures that were recently established as genetically encodable imaging agents for ultrasound[32,41]. GVs evolved in certain aquatic microbes as a means to regulate cellular buoyancy for optimal photosynthetic illumination[9]. GV nanostructures comprise a 2 nm-thick protein shell enclosing an air-filled compartment, with genetically determined widths between 45-250 nm and lengths of several hundred nm[27,28].

The low density and high compressibility of GVs relative to surrounding aqueous media allows these proteins to scatter sound waves and thereby produce ultrasound contrast when injected into the body or expressed heterologously in engineered cells[32,37,41,50].

We hypothesized that we could engineer GV-based biosensors that dynamically change their ultrasound contrast in response to the activity of specific biomolecules. This possibility arises from the recent discovery that GVs' acoustic properties can be modified at the level of their constituent proteins[37]. In particular, the scaffolding protein GvpC, which sits on the GV surface (**Fig. 2-1a**) and provides structural reinforcement[51], can be modified at the level of its amino acid sequence to change GV mechanics. For example, shortening or removing GvpC makes GVs less rigid, allowing them to buckle more easily under acoustic pressure[33,37]. This reversible buckling produces nonlinear ultrasound contrast, which appropriate ultrasound pulse sequences readily distinguish from the linear signals produced by non-buckling GVs and background tissue[33,34].

As an initial target for acoustic biosensor development, we chose proteases — an important class of enzymes involved in many aspects of cellular signaling, homeostasis, disease, therapy, and synthetic biology[52–58]. While these enzymes were the targets of some of the first fluorescent biosensors[59,60], and continue to be a major focus of sensor engineering[61], no acoustic biosensors of protease activity have been developed. We postulated that by engineering variants of GvpC incorporating amino acid sequences that are recognized and acted upon by specific proteases, we could generate GVs whose nonlinear ultrasound contrast becomes activated by protease activity. As representative targets, we selected the constitutively active tobacco etch virus (TEV) endopeptidase, the calcium-dependent mammalian protease calpain, and the processive bacterial protease ClpXP. We set out to test the ability of acoustic biosensors engineered to respond to each of these enzymes to reveal their activity under ultrasound, and to demonstrate biosensor imaging *in vitro*, in living engineered cells, and *in vivo* in the mouse gastrointestinal (GI) tract.

2.2a Engineering an Acoustic Sensor of TEV Endopeptidase

We selected the TEV endopeptidase as our first sensing target because of its well-characterized recognition sequence and widespread use in biochemistry and synthetic biology[62,63]. To sense TEV activity, we engineered a GvpC variant containing the TEV recognition motif ENLYFQ'G (**Fig. 2-1b**), hypothesizing that the cleavage of GvpC into two smaller segments would cause the GV shell to become less stiff, thereby allowing it to undergo buckling and produce enhanced nonlinear ultrasound contrast. We implemented this design *in vitro* using GVs from *Anabaena flos-aque* (Ana), whose native GvpC can be removed after GV isolation, and replaced with new versions expressed heterologously in *Escherichia coli*[37,64]. Ana GvpC comprises five repeats of a predicted alpha-helical polypeptide (**Fig. 2-1a**), and we tested insertions of the TEV recognition sequence, with and without flexible linkers of different lengths, at several locations within this protein. After incubating the engineered GVs with active TEV protease or a heat-inactivated "dead" control (dTEV), we measured their hydrostatic collapse using pressurized absorbance spectroscopy. This technique measures the optical density of GVs (which scatter 500 nm light when intact) under increasing hydrostatic pressure, providing a quick assessment of GV shell mechanics: GVs that collapse at lower pressures also produce more nonlinear contrast[32,37,41,64]. Using this approach, we identified an engineered GV variant that showed ~ 70 kPa reduction in its collapse pressure midpoint upon incubation with the active TEV protease (**Fig. 2-1c and Fig. 2-S1**), and selected it for further characterization. This GV sensor for TEV, hereafter referred to as GVS_{TEV}, has the TEV cleavage site on the second repeat of GvpC, flanked by flexible GSGSGSG linkers on both sides.

TEV cleavage of the GvpC on GVS_{TEV} is expected to produce N- and C-terminal fragments with molecular weights of approximately 9 and 14 kDa, respectively. Indeed, gel electrophoresis of GVS_{TEV} after exposure to active TEV resulted in the appearance of the two cleaved GvpC fragments and a significant reduction in the intact GvpC band intensity (**Fig. 2-1d**). In addition, removal from solution of unbound fragments via buoyancy purification of the GVs resulted in a reduced band intensity for the N-terminal cleavage fragment, indicating its partial dissociation after cleavage (**Fig. 2-1d**). No significant changes

in the GvpC band intensity were observed after incubation with dTEV. Transmission electron microscopy (TEM) images showed intact GVs with similar appearance under both conditions, confirming that protease cleavage did not affect the structure of the underlying GV shell (**Fig. 2-1e**). Dynamic light scattering (DLS) showed no significant difference in the hydrodynamic diameter of the engineered GVs after incubation with dTEV and active TEV protease, confirming that the GVs remain dispersed in solution (**Fig. 2-1f**).

After confirming the desired mechanical and biochemical properties of GVS_{TEV}, we imaged it by ultrasound. Nonlinear imaging was performed in hydrogel samples containing the biosensor, using a recently developed cross-amplitude modulation (x-AM) pulse sequence[34]. x-AM uses pairs of cross-propagating plane waves to elicit highly specific nonlinear scattering from buckling GVs at the wave intersection, while subtracting the linear signal generated by transmitting each wave on its own[34]. Linear images were acquired using a conventional B-mode sequence. As hypothesized, exposing the GVS_{TEV} samples to TEV protease produced a strong nonlinear acoustic response, with a maximal contrast-to-noise ratio (CNR) enhancement of ~ 7 dB at an applied acoustic pressure of 438 kPa (**Fig. 2-1g**). Substantially less nonlinear contrast was observed in controls exposed to dTEV, while, as expected, both samples produced similar linear scattering. Consistent with the pressure-dependent mechanics of the GV shell, the differential nonlinear acoustic response of GVS_{TEV} became evident at pressures above 295 kPa, and kept increasing until 556 kPa, at which point the GVs began to collapse (**Fig. 2-1h and Fig. 2-S1**). As an additional control, we found that GVs with the wild-type GvpC sequence (GV_{WT}) showed no difference in their hydrostatic collapse pressure or nonlinear acoustic contrast in response to TEV protease (**Fig. 2-S1**), and no wild-type GvpC cleavage was seen upon gel electrophoresis (**Fig. 2-S1**). These results established GVS_{TEV} as an acoustic biosensor of the TEV protease enzyme, and additionally provided an experimental template to develop additional sensors.

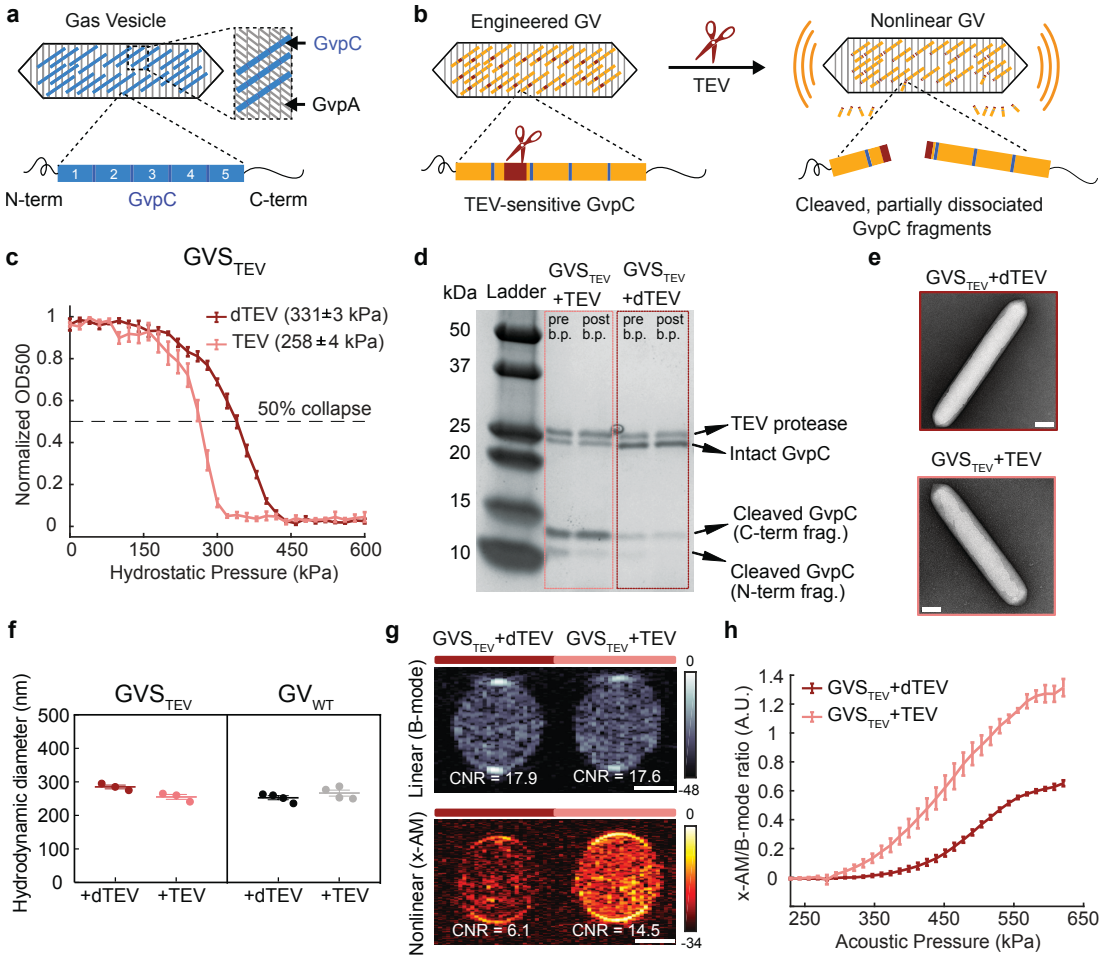


Figure 2-1. Acoustic biosensor of TEV endopeptidase. (**a**) Top: schematic of a gas vesicle (GV), including the primary shell protein GvpA (gray) and the reinforcing protein GvpC (blue). Bottom: schematic of GvpC structure, comprising five 33-amino acid repeats flanked by N-and C-terminal regions. (**b**) Schematic of GVS_{TEV}. (**c**) Normalized OD_{500nm} of GVS_{TEV} as a function of hydrostatic pressure, after incubation with active TEV or heat-inactivated TEV (dTEV). The legend lists the midpoint collapse pressure for each condition (±95% confidence interval), determined from fitting a Boltzmann sigmoid function (N = 3 biological replicates for GVS_{TEV} + TEV and 4 for GVS_{TEV} + dTEV). (**d**) Coomassie-stained SDS-PAGE gel of OD_{500nm}-matched samples of GVS_{TEV} incubated with dTEV or active TEV protease, before and after buoyancy purification (labeled pre b.p. and post b.p., respectively). This experiment was repeated 3 times with similar results. (**e**) Representative TEM images of GVS_{TEV} after incubation with dTEV or active TEV protease (N=3 biological replicates for GVS_{TEV} + TEV and 2 for GVS_{TEV} + dTEV; at least 100 GV particles were imaged for each condition). (**f**) DLS measurements of the average hydrodynamic diameter of GVS_{TEV} and GV_{WT} samples after protease incubation (N = 3 biological replicates for GVS_{TEV} and 4 for GV_{WT}; individual dots represent each N, and thick horizontal line indicates the mean). (**g**) Representative ultrasound images of agarose phantoms containing GVS_{TEV} incubated with TEV or dTEV protease at OD_{500nm} 2.2. The linear (B-mode) image was acquired at 132 kPa and the nonlinear (x-AM) image was acquired at 438 kPa. (**h**) Average ratio of x-AM to B-mode ultrasound signal as a function of applied acoustic pressure for GVS_{TEV}, after incubation with TEV or dTEV protease. N=3 biological replicates, with each N consisting of 2-3 technical replicates for **g** and **h**. For ultrasound images in **g**, CNR stands for contrast-to-noise-ratio, and color bars represent relative ultrasound signal intensity on the dB scale. Solid curves represent the mean in **c** and **h**. Error bars in **c**, **f** and **h** indicate SEM and were calculated from independent biological replicates. Scale bars in **e** represent 100 nm. Scale bars in **g** represent 1 mm. Individual data points for panels **c** and **h** shown as scatter plots in **Figure 2-S1**.

2.2b Engineering an Acoustic Sensor of Calpain

After validating our basic acoustic biosensor design using the model TEV protease, we examined its generalizability to other endopeptidases. As our second target, we selected the calcium-dependent cysteine protease calpain, a mammalian enzyme with critical roles in a wide range of cell types[65–67]. The two most abundant isoforms of this protease, known as μ-calpain and m-calpain, are expressed in many tissues and involved in processes ranging from neuronal synaptic plasticity to cellular senescence[65,66]. We designed an acoustic biosensor of μ-calpain by inserting the α-spectrin-derived recognition sequence QQEVY'GMMPRD[68] into Ana GvpC (**Fig. 2-2a**). We screened several versions of GvpC incorporating this cleavage sequence, flanked by GSG or GSGSG linkers, at different positions within the second helical repeat. Pressurized absorbance spectroscopy performed in buffers with and without calpain and Ca^{2+} allowed us to identify a GV sensor for calpain (GVS_{calp}), showing an approximately 50 kPa decrease in hydrostatic collapse pressure in the presence of the enzyme and its ionic activator (**Fig. 2-2b and Fig. 2-S2**). Electrophoretic analysis confirmed cleavage and partial dissociation of the cleaved fragments from the GV surface (**Fig. 2-S2**), while TEM showed no change in GV morphology (**Fig. 2-S2**).

Ultrasound imaging of GVS_{calp} revealed a robust nonlinear acoustic response when both calpain and calcium were present (**Fig. 2-2c, e, g**), but not in negative controls lacking either or both of these analytes. A slight clustering tendency of GVS_{calp} nanostructures, which was attenuated by incubation with activated calpain (**Fig. 2-S2**), resulted in a slightly higher B-mode signal for the negative controls. However, this did not significantly affect the maximal nonlinear sensor contrast of GVS_{calp} of approximately 7dB (**Fig. 2-2c, e, g**). This contrast increased steeply beyond an applied acoustic pressure of 320 kPa (**Fig. 2-2d, f, h and Fig. 2-S2**). Using this biosensor, ultrasound imaging could be used to visualize the dynamic response of calpain to Ca^{2+}, with a half-maximal response concentration of 140 μM (**Fig. 2-2i and Fig. 2-S2**). Additional control experiments performed on GVs with wild-type GvpC showed no proteolytic cleavage, change in GV collapse pressure or ultrasound response, after incubation with calcium-activated calpain (**Fig. 2-S3**). These results show that acoustic

biosensor designs based on GvpC cleavage can be generalized to a mammalian protease and used to sense the dynamics of a conditionally active enzyme.

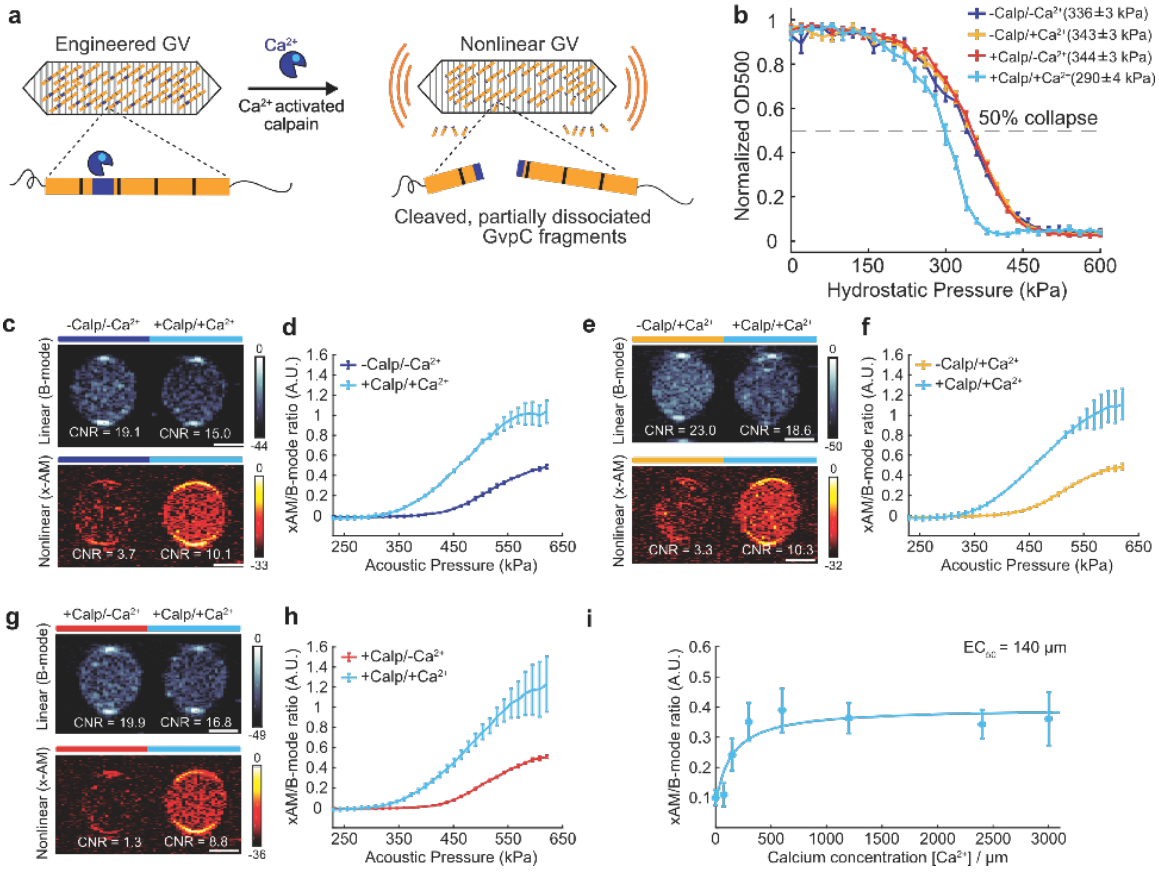


Figure 2-2. Acoustic biosensor of calcium-activated calpain protease. (a) Schematic illustration GVS_{calp}. (**b**) Hydrostatic collapse curves of GVS_{calp} after incubations in the presence or absence of calpain and calcium. The legend lists the midpoint collapse pressure for each condition (±95% confidence interval) determined from fitting a Boltzmann sigmoid function N = 5 biological replicates for +Calp/+Ca^{2+}, 6 for -Calp/+Ca^{2+} and +Calp/-Ca^{2+}, and 7 for -Calp/-Ca^{2+}. (**c, e, g**) Representative ultrasound images of agarose phantoms containing GVS_{calp} incubated with and without calpain and/or calcium at OD_{500nm} 2.2. The B-mode images were taken at 132 kPa for **c**, **e** and **g**, and the x-AM images were taken at 438 kPa for **c**, **e** and at 425 kPa for **g**. CNR stands for contrast-to-noise-ratio, and color bars represent relative ultrasound signal intensity on the dB scale. Scale bars represent 1 mm. (**d, f, h**) Average ratio of x-AM to B-mode ultrasound signal as a function of applied acoustic pressure for GVS_{calp} after incubation in the presence or absence of calpain and/or calcium. N=3 biological replicates, with each N consisting of 2 technical replicates for **c-h**. Solid curves represent the mean and error bars indicate SEM. Statistics were performed on independent biological replicates for **b, d, f** and **h**.(**i**) Calcium-response curve for GVS_{calp} in the presence of µ-calpain, showing the ratio of x-AM to B-mode ultrasound signal at 425 kPa as a function of calcium concentration. The mean values are fitted to a Hill equation with a coefficient of 1, giving a half-maximum response concentration (EC_{50}) of 140 µm (N = 3 biological replicates, individual dots represent the mean values with the solid blue line showing the fitted curve). Error bars indicate SEM. Individual scatter plots for **d**, **f**, **h** and **i** are shown in **Figure 2-S2**.

2.2c Building an Acoustic Sensor of the Protease ClpXP

In addition to endopeptidases, another important class of enzymes involved in cellular protein signaling and homeostasis is processive proteases, which unfold and degrade full proteins starting from their termini[69]. To determine whether GV-based biosensors could be developed for this class of enzymes, we selected ClpXP, a processive proteolytic complex from *E. coli* comprising the unfoldase ClpX and the peptidase ClpP[70]. ClpX recognizes and unfolds protein substrates containing specific terminal peptide sequences called degrons. The unfolded proteins are then fed into ClpP, which degrades them into small peptide fragments[70]. We hypothesized that the addition of a degron to the C-terminus of GvpC would enable ClpXP to recognize and degrade this protein, while leaving the underlying GvpA shell intact, resulting in GVs with greater mechanical flexibility and nonlinear ultrasound contrast (**Fig. 2-3a**).

To test this hypothesis, we appended the ssrA degron, AANDENYALAA, via a short SG linker, to the C-terminus of Ana GvpC, resulting in a sensor that we named GVS_{ClpXP} (**Fig. 2-3a**). We tested the performance of this biosensor *in vitro* using a reconstituted cell-free transcription-translation system comprising *E. coli* extract, purified ClpX, and a ClpP-expressing plasmid. Gel electrophoresis performed after incubating GVS_{ClpXP} with this cell-free extract showed significant degradation of the engineered GvpC, compared to a negative control condition in which the extract was pre-treated with a protease inhibitor (**Fig. 2-3b**). TEM images showed intact GVs under both conditions, confirming that GvpC degradation left the underlying GV shell uncompromised (**Fig. 2-3c**). Pressurized absorbance spectroscopy indicated a substantial weakening of the GV shell upon ClpXP exposure, with the hydrostatic collapse midpoint shifting by nearly 250 kPa (**Fig. 2-3d and Fig. 2-S4**). Ultrasound imaging revealed a 17dB enhancement in the nonlinear contrast produced by GVS_{ClpXP} at an acoustic pressure of 477 kPa, in response to ClpXP activity (**Fig. 2-3e-f and Fig. 2-S4**). Control GVs containing wild type GvpC showed no sensitivity to ClpXP (**Fig. 2-3g-i and Fig. 2-S4**). These results establish the ability of GV-based acoustic biosensors to visualize the activity of a processive protease as turn-on sensors.

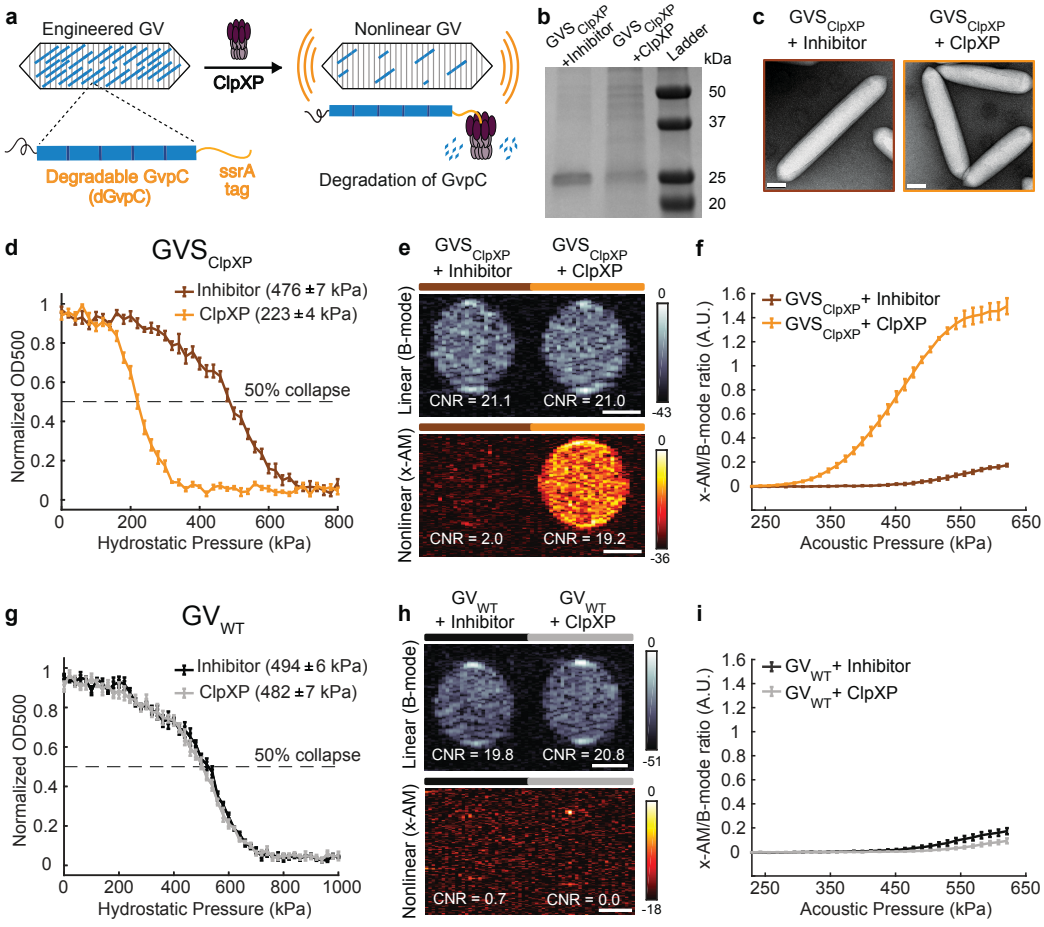


Figure 2-3. Acoustic biosensor of ClpXP protease. (**a**) Schematic of GVS_{ClpXP}. (**b**) Coomassie-stained SDS-PAGE gel of OD_{500nm}-matched GVS_{ClpXP} samples, incubated in a reconstituted cell-free transcription-translation (TX-TL) system containing a protease inhibitor cocktail or ClpXP (N= 3 biological replicates). Additional bands in these gels arise from components of the TX-TL system (**Extended Data Figure 4**) (**c**) Representative TEM images of GVS_{ClpXP} after incubations in the presence of a protease inhibitor or ClpXP. Scale bars represent 100 nm. A minimum of 100 GV particles were imaged for the +ClpXP condition and 50 particles for the +inhibitor control. (**d**) Normalized optical density (OD_{500nm}) measurements of GVS_{ClpXP} as a function of hydrostatic pressure after protease incubation (N=5 biological replicates). (**e**) Representative ultrasound images of agarose phantoms containing GVS_{ClpXP} incubated with the inhibitor cocktail or active ClpXP at OD_{500nm} 2.2. (**f**) Average x-AM/B-mode ratio as a function of applied acoustic pressure for GVS_{ClpXP}, after incubation with the protease inhibitor or active ClpXP. (**g**) Hydrostatic collapse pressure measurements for engineered Ana GVs with WT-GvpC (GV_{WT}) after protease incubation (N=5 biological replicates). For collapse pressure data in **d** and **g**, the legend lists the midpoint collapse pressure for each condition (±95% confidence interval), determined from fitting a Boltzmann sigmoid function. (**h**) Representative ultrasound images of agarose phantoms containing GV_{WT} incubated with the inhibitor cocktail or active ClpXP at OD_{500nm} 2.2. Scale bars in **e** and **h** represent 1mm. CNR stands for contrast-to-noise-ratio, and color bars represent relative ultrasound signal intensity on the dB scale. The B-mode images were acquired at 132 kPa and the x-AM images were acquired at 477 kPa. (**i**) Average ratio of x-AM to B-mode acoustic signal as a function of applied acoustic pressure for GV_{WT} after incubation with the inhibitor cocktail or ClpXP protease. For **e**, **f**, **h** and **i**, N=3 biological replicates, with each N having 3 technical replicates. For **d**, **f**, **g** and **i**, solid curves represent the mean and error bars indicate SEM, which were calculated from independent biological replicates. Individual scatter plots for **d, f**, **g** and **i** are shown in **Figure 2-S4**.

2.2d Constructing Intracellular Acoustic Sensor Genes

After demonstrating the performance of acoustic biosensors *in vitro*, we endeavored to show that they could respond to enzymatic activity inside living cells. As the cellular host, we chose *E. coli* Nissle 1917. This probiotic strain of *E. coli* has the capacity to colonize the mammalian gastrointestinal tract, and is widely used as a chassis for the development of microbial therapeutics[71–73], making it a valuable platform for intracellular biosensors. Recently, an engineered operon comprising GV-encoding genes from *Anabaena flos-aquae* and *Bacillus megaterium* was expressed in Nissle cells as acoustic reporter genes (*ARGs*), allowing gene expression to be imaged with linear B-mode ultrasound[41]. To develop an intracellular acoustic sensor gene targeting ClpXP (ASG_{ClpXP}), we swapped the wild type *gvpC* in the *ARG* gene cluster (ARG_{WT}) with the modified *gvpC* from GVS_{ClpXP} (dGvpC) (**Fig. 2-4a**). For a first test of this intracellular biosensor, we transformed it into wild-type (WT) Nissle cells, which natively express ClpXP protease, hypothesizing that it would show a reduced intracellular collapse pressure and enhanced nonlinear contrast compared to ARG_{WT}. Indeed, pressurized absorbance spectroscopy on intact cells expressing ASG_{ClpXP} revealed a reduction in the hydrostatic collapse pressure midpoint of ~ 160 kPa relative to cells expressing ARG_{WT} (**Fig. 2-S5**). In ultrasound imaging, live cells expressing ASG_{ClpXP} showed an enhancement in nonlinear contrast of approximately 13 dB (**Fig. 2-S5**), while linear B-mode signal was similar. The nonlinear response of ASG_{ClpXP} expressing cells was strongest beyond an acoustic pressure of 784 kPa (**Fig. 2-S5**).

Next, to examine the ability of ASG_{ClpXP} to respond to intracellular enzymatic activity in a dynamic manner, we generated a ClpXP-deficient strain of Nissle cells (Δ*clpXP*) through genomic knock-out of the genes encoding ClpX and ClpP, and created a plasmid containing these two genes under the control of an arabinose-inducible promoter (**Fig. 2-4a**). This allowed us to externally control the activity of the ClpXP enzyme. Δ*clpXP* Nissle cells were co-transformed with an inducible *clpX-clpP* (*clpXP*) plasmid and ASG_{ClpXP}. ClpXP production in these cells after induction with L-arabinose resulted in an approximately 160

kPa reduction in the hydrostatic collapse pressure midpoint (**Fig. 2-4b and Fig. 2-S5**). Under ultrasound imaging, cells with induced ClpXP activity showed substantially stronger nonlinear contrast (+6.7 dB) compared to cells uninduced for this protease (**Fig. 2-4c**), while showing a similar B-mode signal. This enhancement in nonlinear signal was detectable with acoustic pressures above 950 kPa (**Fig. 2-4d and Fig. 2-S5**). These experiments demonstrate the ability of ASG_{ClpXP} to function as an intracellular acoustic sensor to monitor variable enzyme activity.

A major application of dynamic sensors in cells is to monitor the activity of natural or synthetic gene circuits[74–76]. To test if our acoustic sensors could be used to track the output of a synthetic gene circuit in cells, we co-transformed WT Nissle cells with ASG_{ClpXP}, and a separate wild-type *gvpC* gene controlled by anhydrotetracycline (aTc) (**Fig. 2-4e**). Our hypothesis was that induction of this gene circuit only with IPTG would result in the production of GVs with ClpXP-degradable GvpC, resulting in nonlinear contrast, whereas the additional input of aTc would result in the co-production of non-degradable wild-type GvpC, which would take the place of any degraded engineered GvpC on the biosensor shell and lead to reduced nonlinear scattering (**Fig. 2-4e**). Indeed, when we induced cells with just IPTG we observed strong nonlinear contrast. However, when aTc was added to the cultures after IPTG induction, this contrast was reduced by approximately 10 dB (**Fig. 2-4f-g and Fig. 2-S5**). These results, together with our findings in Δ*clpXP* cells with inducible ClpXP, show that acoustic biosensors can be used to visualize the output of synthetic gene circuits.

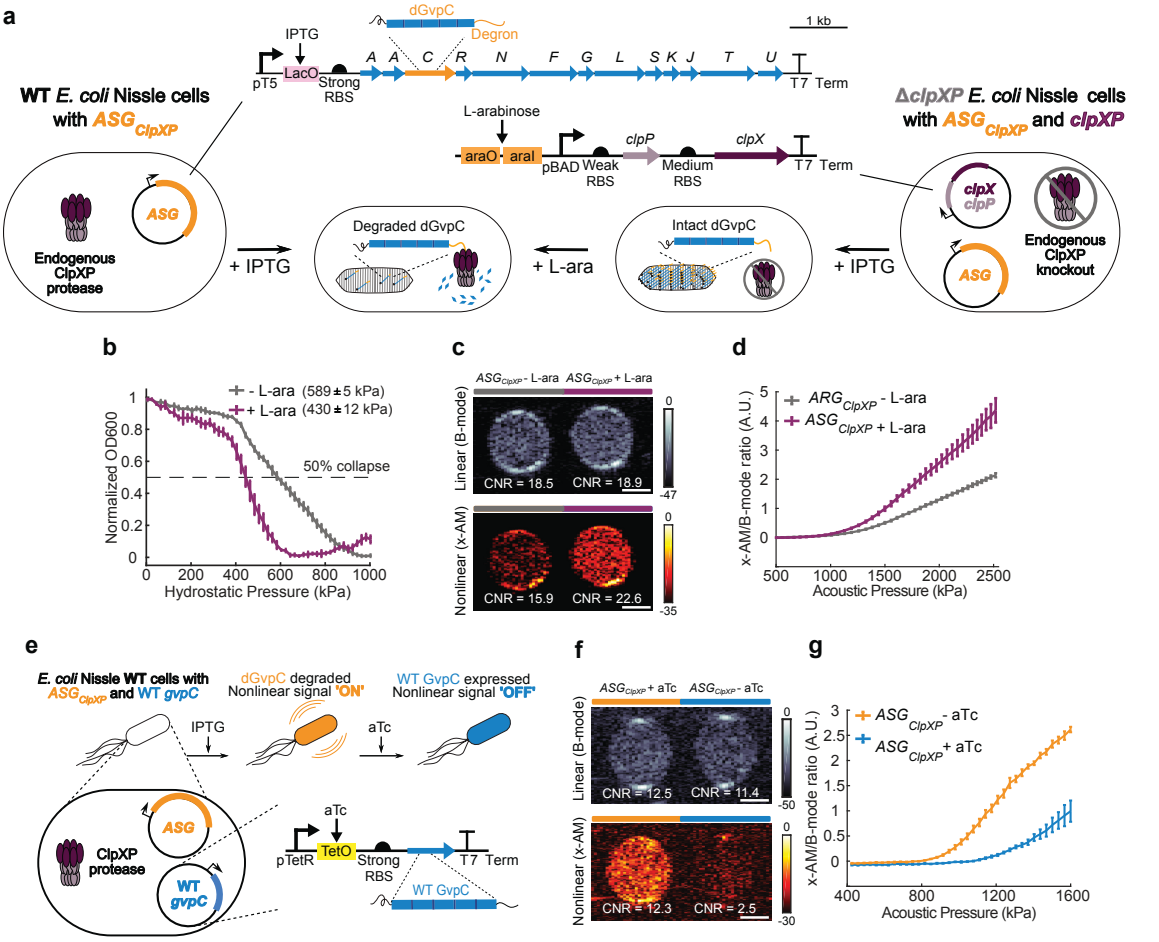


Figure 2-4. Monitoring intracellular protease activity and circuit-driven gene expression in engineered cells. (**a**) Schematic of *E. coli* Nissle cells expressing the acoustic sensor gene construct for ClpXP. In some cases, the Nissle cells are genomically modified to lack the *clpX* and *clpP* genes (Δ*clpXP*), and co-transformed with a plasmid encoding L-arabinose (L-ara) driven ClpXP. (**b**) Normalized pressure-sensitive optical density at 600 nm of Δ*clpXP* Nissle cells expressing ASG_{ClpXP} with or without L-ara induction of ClpXP protease expression. The legend lists the midpoint collapse pressure for each cell type (±95% confidence interval) determined from fitting a Boltzmann sigmoid function (N = 3 biological replicates). (**c**) Representative ultrasound images of Δ*clpXP* Nissle cells expressing ASG_{ClpXP} with or without L-ara induction of ClpXP protease at OD_{600nm} 1.5. (**d**) Average x-AM/B-mode ratio as a function of applied acoustic pressure for Δ*clpXP* Nissle cells expressing ASG_{ClpXP} with or without L-ara induction of ClpXP expression at OD_{600nm} 1.5. N=3 biological replicates, with each N having 3 technical replicates for **c** and **d**. (**e**) Schematic of pT5-LacO driven ASG_{ClpXP} and pTet-TetO driven WT *gvpC* gene circuits co-transformed into Nissle cells for dynamic switching of non-linear acoustic signals from the intracellular GV sensors in response to circuit-driven gene expression. (**f**) Representative ultrasound images of Nissle cells (OD_{600nm} 1) expressing ASG_{ClpXP}, with or without aTc induction to drive expression of WT GvpC. (**g**) Average x-AM/B-mode ratio as a function of applied acoustic pressure for Nissle cells expressing ASG_{ClpXP}, with or without aTc induction. N=5 biological replicates for **f** and **g**. CNR stands for contrast-to-noise-ratio, and color bars represent relative ultrasound signal intensity in the dB scale. The B-mode images were acquired at 309 kPa for (**c**) and 132 kPa for (**f**). The x-AM images were acquired at 1.61 MPa for (**c**), and 1.34 MPa for (**f**). Scale bars in **c** and **f** represent 1 mm.For **b**, **d** and **g**, solid curves represent the mean and error bars indicate SEM. Statistics were performed on data from independent biological replicates. Individual scatter plots for **b,d** and **g** are shown in **Figure 2-S5**.

2.2e Ultrasound Imaging of Intracellular ClpXP Activity *in vivo*

Finally, after establishing the basic principles of acoustic biosensor engineering *in vitro* and demonstrating their performance in living cells, we assessed the ability of our sensor constructs to produce ultrasound contrast within a biologically relevant anatomical location *in vivo*. In particular, approaches to imaging microbes in the mammalian GI tract[41,77–79] are needed to support the study of their increasingly appreciated roles in health and disease[80,81] and the development of engineered probiotic agents[82,83]. The GI tract is also an excellent target for ultrasound imaging due to its relatively deep location inside the animal, and the use of ultrasound in clinical diagnosis and animal models of GI pathology, with appropriate measures taken to minimize potential interference from air bubbles and solid matter[84,85].

To demonstrate the ability of acoustic biosensors to produce nonlinear ultrasound contrast within the *in vivo* context of the mouse GI tract, we first co-injected WT Nissle cells expressing ASG_{ClpXP} and ARG_{WT} into the mouse colon (schematic shown in **Fig. 2-S6**), distributing one cell population along the lumen wall and the other in the lumen center. In these proof-of-concept experiments, the cells are introduced into the colon in a rectally-injected agarose hydrogel to enable precise positioning and control over composition. Using nonlinear ultrasound imaging, we could clearly visualize the unique contrast generated by the protease-sensitive *ASG*s as a bright ring of contrast lining the colon periphery (**Fig. 2-5a**). When the spatial arrangement was reversed, the bright nonlinear contrast was concentrated in the middle of the lumen (**Fig. 2-S7**). A comparison of ultrasound images acquired before and after acoustic collapse of the GVs, using a high-pressure pulse from the transducer, confirmed that the bright ring of nonlinear contrast was emanating from ASG_{ClpXP} -expressing cells (**Fig. 2-5a)**, and this result was consistent across independent experiments in 9 mice (**Fig. 2-5b).**

To demonstrate *in vivo* imaging of enzyme activity, we introduced Δ*clpXP* Nissle cells expressing ASG_{ClpXP} into the mouse colon, with and without transcriptionally activating intracellular ClpXP (schematic shown in **Fig. 2-S6**) . As above, the cells were contained in

an agarose hydrogel. Cells induced to express this enzyme showed enhanced nonlinear contrast compared to cells not expressing ClpXP (**Fig. 2-5c**). Acoustic collapse confirmed the acoustic biosensors as the primary source of nonlinear signal (**Fig. 2-5c**). This performance was consistent across 7 mice and 2 spatial arrangements of the cells (**Fig. 2-5d**). These results demonstrate the ability of acoustic biosensors to visualize enzyme activity within the context of *in vivo* imaging.

Besides molecular sensing, one additional benefit of the nonlinear contrast generated by ASG_{ClpXP} -expressing cells is to make the cells easier to detect relative to background tissue compared to linear B-mode imaging. Indeed, the nonlinear contrast of WT Nissle cells expressing ASG_{ClpXP} had a significantly higher contrast-to-tissue ratio than either the nonlinear contrast of ARG_{WT}-expressing cells, or the B-mode contrast of either of these two species (**Fig. 2-S8**).

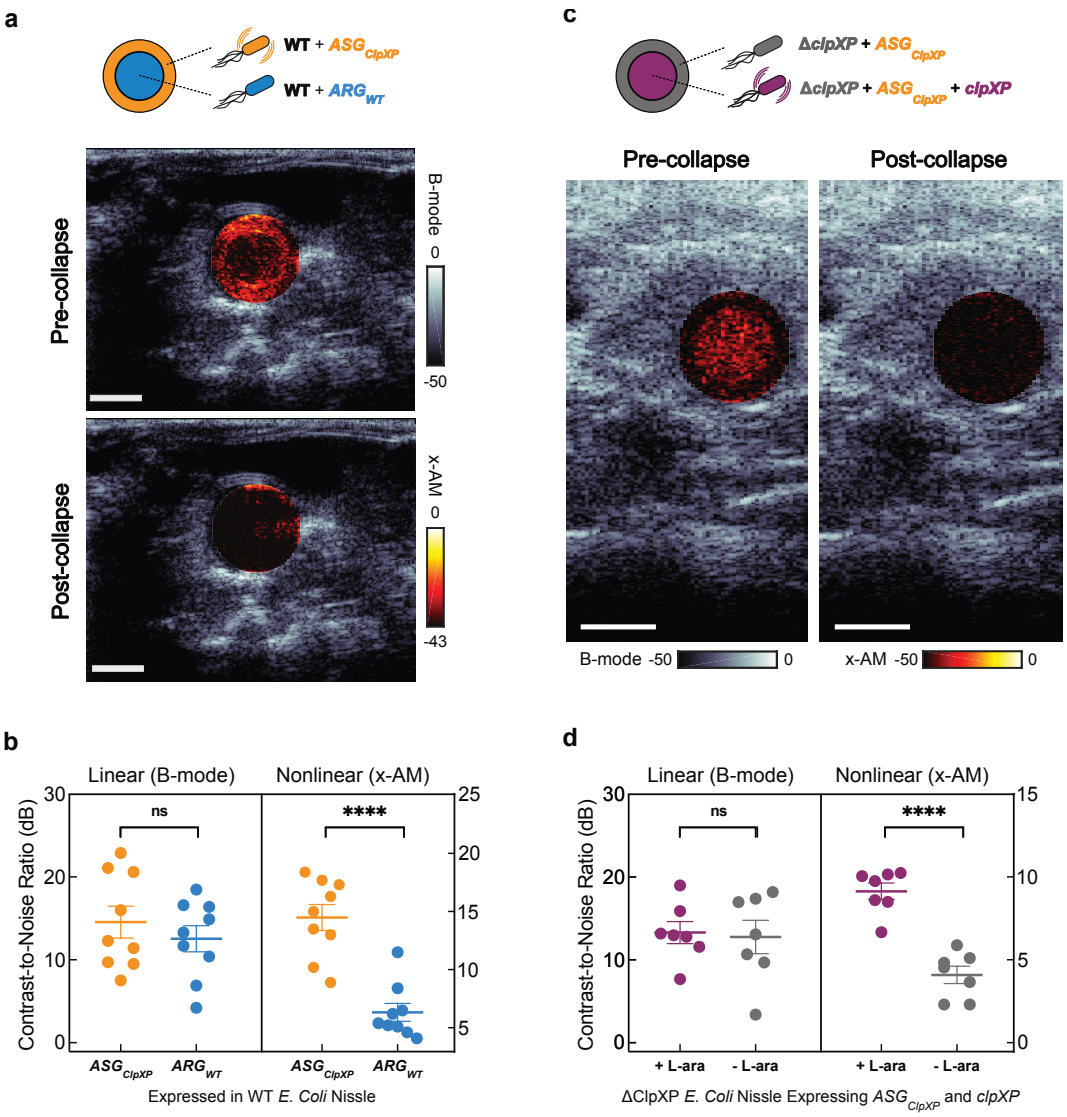


Figure 2-5. Ultrasound imaging of bacteria expressing acoustic sensor genes in the gastrointestinal tract of mice. (**a**) Transverse ultrasound image of a mouse whose colon contains WT Nissle cells expressing *ARG*$_{WT}$ at the center of the lumen and the same strain expressing *ASG*$_{ClpXP}$ at the periphery of the lumen. These imaging experiments were independently repeated 9 times with similar results. (**b**) B-mode and xAM contrast-to-noise ratio (CNR) *in vivo*, for WT Nissle cells expressing *ARG*$_{WT}$ or *ASG*$_{ClpXP}$. N = 9 mice. P = 7.8E-5 for x-AM signal from cells expressing *ASG*$_{ClpXP}$ versus the *ARG*$_{WT}$ control and P = 0.2890 for B-mode signal. (**c**) Transverse ultrasound image of a mouse whose colon contains Δ*clpXP* Nissle cells expressing *ASG*$_{ClpXP}$ with L-ara induction of ClpXP protease expression at the center and without L-ara induction at the periphery of the lumen. These imaging experiments were independently repeated 7 times with similar results. Cells were injected in agarose gel at a final concentration of 1.5E9 cells ml^{-1} for **a** and **c**. Nonlinear (x-AM) images of the colon, acquired at 1.27 MPa for (**a**) and 1.56 MPa for (**c**) before and after acoustic collapse (hot color map), are superimposed on linear (B-mode) anatomical images (bone colormap). Color bars represent relative ultrasound signal intensity on the dB scale. Scale bars represent 2 mm for **a** and **c**. (**d**) B-mode and xAM CNR *in vivo*, for Δ*clpXP* Nissle cells expressing *ASG*$_{ClpXP}$ with or without L-ara induction of ClpXP expression. N = 7 mice. P = 1.8E-5 for x-AM signal from cells expressing *ASG*$_{ClpXP}$ with ClpXP protease expression induced versus non-induced and P = 0.8293 for B-mode signal. Individual dots represent each N, and the thick horizontal line indicates the mean. Error bars indicate SEM. P-values were calculated using a two-tailed paired t-test.

2.3 Discussion and Outlook

Our results establish a paradigm for visualizing protease activity non-invasively with ultrasound imaging. This paradigm is enabled by the dependence of the buckling mechanics of GVs on the reinforcing protein GvpC, and the ability to turn this protein into a protease substrate by incorporating specific internal or terminal peptide sequences. Similar to the earliest work on fluorescent biosensors[59,60], this initial study has focused on proteases due to the importance of this class of enzymes in biology, their relatively compact recognition motifs, and the large impact of their activity on protein structure. Based on our success in sensing the function of three distinct proteases, we anticipate that the basic design strategy presented here should be applicable to many enzymes of this type.

Our study lends itself to numerous future investigations to extend the applications of acoustic protease sensors beyond the proof-of-concept demonstrations shown here. While our experiments in *E. coli* and within the mouse GI tract establish the critical ability of such biosensors to produce ultrasound contrast in relevant biological settings, additional application-centric optimizations would enable the use of these constructs to address specific problems in basic and synthetic biology. For example, purified acoustic biosensors could be designed to sense extracellular proteases, which play homeostatic and disease-causing roles in tissues ranging from extracellular matrix remodeling and blood clot formation to inter-cellular signaling. Meanwhile, the expression of acoustic biosensor genes in cells could be used to monitor natural cellular enzyme activity or serve as the output of synthetic signaling pathways. Intracellular use in bacteria could be particularly relevant in studying microbes in the mammalian GI tract, provided the successful adaptation of acoustic sensor genes to the relevant host species and ensuring successful delivery via oral gavage, colonization and metabolic viability. For potential applications in mammalian cells, acoustic protease sensor designs must be integrated into recently developed genetic programs enabling the expression of GVs in mammalian cells[39]. Successful use of acoustic sensors in this context will require increasing the level of mammalian GV expression to enable non-destructive nonlinear imaging.

In parallel, significant scope exists for further optimizing and generalizing the design of acoustic biosensors. While all three of our sensors produced detectable nonlinear contrast in response to protease activity, the changes exhibited by GVS_{ClpXP} were significantly larger than for the other two constructs. This is not surprising for an enzyme that processively degrades GvpC, and whose recognition motif can be incorporated outside the main GV-binding region of GvpC. Endopeptidase sensors could be optimized to reach similar performance by incorporating more than one cleavage site within the GvpC sequence and tuning the linkers connecting these sites to the rest of the protein. As with other protease biosensors, the irreversibility of proteolysis means that for repeated or continuous sensing, it is necessary for new sensor molecules to be synthesized or delivered. For genetically encoded biosensors, this occurs through gene expression, potentially posing a metabolic burden to the cell. For GVs, this burden could be reduced by re-expressing only the engineered GvpC rather than the full GV, since this protein can be added onto the shell of existing GVs, as demonstrated in this study and previous work[37]. Going beyond proteolytic sensors, we anticipate that our biosensor design strategy could be modified to enable allosteric conformational changes in GvpC, rather than its cleavage, to alter ultrasound contrast, thereby creating acoustic biosensors that respond reversibly to non-cleaving enzymes, ions or other signals of interest.

In addition to optimizing the biosensor constructs, it is also possible to improve the ultrasound techniques used for their visualization. In this study, we monitored the activation of our biosensors using a nonlinear x-AM pulse sequence, quantifying the resulting contrast relative to linear B-mode scattering. This ratiometric signal is advantageous for quantification in scenarios where the sensor concentration may vary. However, the dependence of the x-AM response on applied acoustic pressure introduces a variable that may differ across the ultrasonic field of view, and strategies involving dynamic pressure adjustment may be needed to obtain the optimal signal from each point in the imaged plane. In addition, normalization to B-mode signal in complex *in vivo* contexts may require methods to separate the linear scattering contributions of acoustic sensors from those of background

tissue. With these improvements, acoustic biosensors promise to take dynamic imaging of molecular and cellular function to new depths.

2.4 Material and Methods

Design and cloning of genetic constructs

All gene sequences were codon optimized for *E.Coli* expression and inserted into their plasmid backbones via Gibson Assembly or KLD Mutagenesis using enzymes from New England Biolabs and custom primers from Integrated DNA Technologies. The protease recognition sequences for TEV protease and μ-calpain, flanked by flexible linkers, were introduced by substitution-insertion into the second repeat of the wild-type Ana *gvpC* gene sequence in a pET28a expression vector (Novagen) driven by a T7 promoter and lac operator. The ssrA degradation tag for the ClpXP bacterial proteasome was appended to the C-terminus of Ana *gvpC* using a short flexible linker. The acoustic sensor gene for intracellular protease sensing of ClpXP was constructed by modifying of the acoustic reporter gene cluster *ARG1*[41], by addition of the ssrA degradation tag to the C-terminal of *gvpC* using a linker sequence. For expression in *E.coli* Nissle 1917 cells, the pET28a T7 promoter was replaced by the T5 promoter. For inducible expression of *clpX* and *clpP*, the genes encoding those two proteins were cloned from the *E. coli* Nissle 1917 genome into a modified pTARA backbone under a P_{BAD} promoter and araBAD operon. For dynamic regulation of intracellular sensing, the wild-type GvpC sequence was cloned into a modified pTARA backbone under a pTet promoter and tetracycline operator. The complete list and source of plasmids used in this study is given in **Table 2-S1**. Plasmid constructs were cloned using NEB Turbo *E. Coli* (New England Biolabs) and sequence-validated.

Construction of *clpX*$^{-}$ *clpP*$^{-}$ strain of *E.coli* Nissle 1917 (Δ*clpXP*)

The knockout of *clpX* and *clpP* in *E.coli* Nissle (ECN) was accomplished by Lambda Red recombineering using previously published methods[86]. An FRT-flanked *cat* gene was recombined into ECN genome to replace the *clpX* and *clpP* genes, and the integrated *cat* gene was then removed by the *FLP* recombinase from pE-FLP[87] to yield the Δ*clpXP* strain.

More information on the recombineering plasmids used in this study and their source is provided in **Table 2-S1**.

GV expression, purification, and quantification

For *in vitro* assays, GVs were harvested and purified from confluent *Ana* cultures using previously published protocols[37,64]. Briefly, Ana cells were grown in Gorham's media supplemented with BG-11 solution (Sigma) and 10 mM sodium bicarbonate at 25°C, 1% CO_2 and 100 rpm shaking, under a 14h light and 10h dark cycle. Confluent cultures were transferred to sterile separating funnels and left undisturbed for 2-3 days to allow buoyant *Ana* cells expressing GVs to float to the top and for their subnatant to be drained. Hypertonic lysis with 10% Solulyse (Genlantis) and 500 mM sorbitol was used to release and harvest the Ana GVs. Purified GVs were obtained through 3-4 rounds of centrifugally assisted floatation, with removal of the subnatant and resuspension in phosphate buffered saline (PBS, Corning) after each round.

For expression of acoustic reporter/sensor genes (*ARG/ASG*) in bacteria, wild-type *E. Coli* Nissle 1917 cells (Ardeypharm GmbH) were made electrocompetent and transformed with the genetic constructs. After electroporation, cells were rescued in SOC media supplemented with 2% glucose for 1h at 37°C. Transformed cells were grown for 12-16 hours at 37°C in 5 mL of LB medium supplemented with 50 μg/mL kanamycin and 2% glucose. Large-scale cultures for expression were prepared by a 1:100 dilution of the starter culture in LB medium containing 50 μg/mL kanamycin and 0.2% glucose. Cells were grown at 37°C to an OD_{600nm} of 0.2-0.3, then induced with 3μM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and allowed to grow for 22 hrs at 30°C. Buoyant *E. coli* Nissle cells expressing GVs were isolated from the rest of the culture by centrifugally assisted floatation in 50 mL conical tubes at 300g for 3-4 hrs, with a liquid column height less than 10 cm to prevent GV collapse by hydrostatic pressure.

The concentration of Ana GVs was determined by measurement of their optical density (OD) at 500 nm (OD_{500}) using a Nanodrop spectrophotometer (Thermo Fisher Scientific), using

the resuspension buffer or collapsed GVs as the blank. As established in previous work[64], the concentration of GVs at OD_{500} = 1 is approximately 114 pM and the gas fraction is 0.0417%. The OD of buoyant cells expressing GVs were quantified at 600 nm using the Nanodrop.

Bacterial expression and purification of GvpC variants

For expression of Ana GvpC variants, plasmids were transformed into chemically competent BL21(DE3) cells (Invitrogen) and grown overnight for 14-16 h at 37°C in 5 mL starter cultures in LB medium with 50 μg/mL kanamycin. Starter cultures were diluted 1:250 in Terrific Broth (Sigma) and allowed to grow at 37°C (250 rpm shaking) to reach an OD_{600nm} of 0.4-0.7. Protein expression was induced by addition of 1 mM IPTG, and the cultures were transferred to 30°C. Cells were harvested by centrifugation at 5500g after 6-8 hours. For the GvpC-ssrA variant, expression was carried out at 25°C for 8 hours to reduce the effect of protease degradation and obtain sufficient protein yield.

GvpC was purified from inclusion bodies by lysing the cells at room temperature using Solulyse (Genlantis), supplemented with lysozyme (400 μg/mL) and DNase I (10 μg/mL). Inclusion body pellets were isolated by centrifugation at 27,000g for 15 mins and then resuspended in a solubilization buffer comprising 20 mM Tris-HCl buffer with 500 mM NaCl and 6 M urea (pH: 8.0), before incubation with Ni-NTA resin (Qiagen) for 2 h at 4°C. The wash and elution buffers were of the same composition as the solubilization buffer, but with 20mM and 250 mM imidazole, respectively. The concentration of the purified protein was assayed using the Bradford Reagent (Sigma). Purified GvpC variants were verified to be >95% pure by SDS-PAGE analysis.

Preparation of gas vesicles for *in vitro* protease assays

Engineered GVs having protease-sensitive or wild-type GvpC were prepared using urea stripping and GvpC re-addition[37,64]. Briefly, Ana GVs were stripped of their native outer

layer of GvpC by treatment with 6M urea solution buffered with 100 mM Tris- HCl (pH:8-8.5). Two rounds of centrifugally assisted floatation with removal of the subnatant liquid after each round were performed to ensure complete removal of native GvpC. Recombinant Ana GvpC variants purified from inclusion bodies were then added to the stripped Ana GVs in 6 M urea a 2-3x molar excess concentration determined after accounting for 1:25 binding ratio of GvpC: GvpA. For a twofold stoichiometric excess of GvpC relative to binding sites on an average Ana GV, the quantity of recombinant GvpC (in nmol) to be added to stripped GVs was calculated according to the formula: 2 * OD * 198 nM * volume of GVs (in liters). The mixture of stripped GVs (OD_{500nm} = 1-2) and recombinant GvpC in 6 M urea buffer was loaded into dialysis pouches made of regenerated cellulose membrane with a 6-8 kDa M.W. cutoff (Spectrum Labs). The GvpC was allowed to slowly refold onto the surface of the stripped GVs by dialysis in 4 L PBS for at least 12 h at 4 °C. Dialyzed GV samples were subjected to two or more rounds of centrifugally assisted floatation at 300 g for 3-4 h to remove any excess unbound GvpC. Engineered GVs were resuspended in PBS after subnatant removal and quantified using pressure-sensitive OD measurements at 500 nm using a Nanodrop.

Pressurized absorbance spectroscopy

Purified, engineered Ana GVs were diluted in experimental buffers to an OD_{500nm} ~ 0.2-0.4, and 400 μL of the diluted sample was loaded into a flow-through quartz cuvette with a pathlength of 1 cm (Hellma Analytics). Buoyant *E.Coli* Nissle cells expressing GVs were diluted to an OD_{600nm} of ~ 1 in PBS for measurements. A 1.5 MPa nitrogen gas source was used to apply hydrostatic pressure in the cuvette through a single valve pressure controller (PC series, Alicat Scientific), while a microspectrometer (STS-VIS, Ocean Optics) measured the OD of the sample at 500 nm (for Ana GVs) or 600 nm (for Nissle cells). The hydrostatic pressure was increased from 0 to 1 MPa in 20 kPa increments with a 7 second equilibration period at each pressure before OD measurement. Each set of measurements was normalized by scaling to the Min-Max measurement value, and the data was fitted using the Boltzmann

sigmoid function $f(P) = 1 + e^{(P-P_c)/\Delta P^{-1}}$, with the midpoint of normalized OD change (P_c) and the 95% confidence intervals, rounded to the nearest integer, reported in the figures.

TEM sample preparation and imaging

Freshly diluted samples of engineered Ana GVs (OD_{500nm} ~ 0.3) in 10 mM HEPES buffer containing 150 mM NaCl (pH 8) were used for TEM. 2 μL of the sample was added to Formvar/carbon 200 mesh grids (Ted Pella) that were rendered hydrophilic by glow discharging (Emitek K100X). 2% uranyl acetate was added for negative staining. Images were acquired using the FEI Tecnai T12 LaB6 120kV TEM equipped with a Gatan Ultrascan 2k X 2k CCD and 'Leginon' automated data collection software suite.

Dynamic light scattering (DLS) measurements

Engineered Ana GVs were diluted to an OD_{500nm} ~ 0.2 in experimental buffers. 150-200 μL of the sample was loaded into a disposable cuvette (Eppendorf UVette®) and the particle size was measured using the ZetaPALS particle sizing software (Brookhaven instruments) with an angle of 90 ° and refractive index of 1.33.

Denaturing polyacrylamide gel electrophoresis (SDS-PAGE)

GV samples were OD_{500nm} matched and mixed 1:1 with 2x Laemmli buffer (Bio-Rad), containing SDS and 2-mercaptoethanol. The samples were then boiled at 95°C for 5 minutes and loaded into a pre-made polyacrylamide gel (Bio-Rad) immersed in 1x Tris-Glycine-SDS Buffer. 10 uL of Precision Plus Protein™ Dual Color Standards (Bio-Rad) was loaded as the ladder. Electrophoresis was performed at 120V for 55 minutes, after which the gel was washed in DI water for 15 minutes to remove excess SDS and commassie-stained for 1 hour in a rocker-shaker using the SimplyBlue SafeStain (Invitrogen). The gel was allowed to de-stain overnight in DI water before imaging using a Bio-Rad ChemiDoc™ imaging system.

In vitro protease assays

For *in vitro* assays with the TEV endopeptidase, recombinant TEV protease (R&D Systems, Cat. No. 4469-TP-200) was incubated (25% v/v fraction) with engineered Ana GVs resuspended in PBS (final OD_{500nm} in reaction mixture = 5-6) at 30°C for 14-16 h. This corresponds to a TEV concentration of 0.1~0.125 mg/mL (depending on the lot), within the range used in previous studies with this enzyme[88,89]. Engineered GVs with wild-type GvpC and TEV protease heat-inactivated at 80°C for 20-30 mins were used as the controls.

For *in vitro* assays with calpain, calpain-1 from porcine erythrocytes (Millipore Sigma, Cat. No. 208712) was incubated in a 10% v/v fraction with engineered Ana GVs in a reaction mixture containing 50 mM Tris-HCl, 50 mM NaCl, 5 mM 2-mercaptoethanol, 1 mM EDTA and 1 mM EGTA and 5 mM Ca^{2+} (pH: 7.5) This corresponds to a calpain concentration of ≥ 0.168 units per µl, with 1 unit defined by the manufacturer as sufficient to cleave 1 pmol of a control fluorogenic substrate in 1 min at 25˚C. The final concentration of engineered GVs in the reaction mixture was OD_{500nm} ~ 6 and the protease assay was carried out at 25°C for 14-16h. Negative controls included the same reaction mixture without calpain, without calcium, or without calpain and calcium. Engineered GVs with WT-GvpC were used as additional negative controls.

For *in vitro* assays with ClpXP, a reconstituted cell-free transcription-translation (TX-TL) system adapted for ClpXP degradation assays[90] (gift from Zachary Sun and Richard Murray) was used. Briefly, cell-free extract was prepared by lysis of ExpressIQ *E. coli* cells (New England Biolabs), and mixed in a 44% v/v ratio with an energy source buffer, resulting in a master mix of extract and buffer comprising: 9.9 mg/mL protein, 1.5 mM each amino acid except leucine, 1.25 mM leucine, 9.5 mM Mg-glutamate, 95 mM K-glutamate, 0.33 mM DTT, 50 mM HEPES, 1.5 mM ATP and GTP, 0.9 mM CTP and UTP, 0.2 mg/mL tRNA, 0.26 mM CoA, 0.33 mM NAD, 0.75 mM cAMP, 0.068 mM folinic acid, 1 mM spermidine, 30 mM 3-PGA and 2% PEG-8000. For purified ClpX protein, a monomeric N-terminal deletion variant Flag-ClpXdeltaNLinkedHexamer-His6[91] (Addgene ID: 22143) was used. Post Ni-NTA purification, active fractions of ClpX hexamers with sizes above 250 kDa were

isolated using a Supradex 2010/300 column, flash frozen at a concentration of 1.95 μM and stored at -80°C in a storage buffer consisting of: 50 mM Tris-Cl (pH 7.5), 100 mM NaCl, 1mM DTT, 1 mM EDTA and 2% DMSO. The final reaction mixture was prepared as follows: 75% v/v fraction of the master mix, 10% v/v of purified ClpX, 1nm of the purified pBEST-ClpP plasmid and engineered Ana GVs (concentration of OD_{500nm} = 2.5-2.7 in the reaction mixture). The mixture was made up to the final volume using ultrapure H_2O. The reaction was allowed to proceed at 30°C for 14-16 h. As a negative control, a protease inhibitor cocktail mixture (SIGMA*FAST*™, Millipore Sigma) was added to the reaction mixture at 1.65x the manufacturer-recommended concentration and pre-incubated at room temperature for 30 mins.

Dynamic sensing of ClpXP activity in Δ*clpXP E.Coli* Nissle 1917 cells

ClpXP *E. Coli* Nissle 1917 cells were made electrocompetent and co-transformed with the pET expression plasmid (Lac-driven) containing the *ASG* for ClpXP and a modified pTARA plasmid (pBAD-driven) containing the *clpX* and *clpP* genes. Electroporated cells were rescued in SOC media supplemented with 2% glucose for 2h at 37°C. Transformed cells were grown overnight at 37ºC in 5 mL LB medium supplemented with 50 μg/mL kanamycin, 25 μg/mL chloramphenicol, and 2% glucose. Starter cultures were diluted 1:100 in LB medium with 50 μg/mL kanamycin, 25 μg/mL chloramphenicol, and 0.2% glucose and allowed to grow at 37 °C to reach an OD_{600nm} of 0.2-0.3. *ASG* expression was induced with 3μM IPTG and the bacterial culture was transferred to the 30 °C incubator with 250 rpm shaking for 30 minutes. The culture was then split into two halves of equal volume, and one half was induced with 0.5% (weight fraction) L-arabinose for expression of ClpXP protease. Cultures with and without L-arabinose induction were allowed to grow for an additional 22 h at 30°C. Cultures were then spun down at 300 g in a refrigerated centrifuge at 4 °C for 3-4 h in 50 mL conical tubes to isolate buoyant cells expressing GVs from the rest of the culture. The liquid column height was maintained at less than 10 cm to prevent GV collapse by hydrostatic pressure.

Dynamic sensing of circuit-driven gene expression in *E. coli* Nissle 1917 cells

Electrocompetent *E. coli* Nissle cells were co-transformed with the pET expression plasmid (Lac-driven) containing the *ASG* for ClpXP and a modified pTARA plasmid[92] (Tet-driven) containing the WT Ana GvpC gene. Electroporated cells were rescued in SOC media supplemented with 2% glucose for 2h at 37°C. Transformed cells were grown overnight at 37ºC in 5 mL LB medium supplemented with 50 μg/mL kanamycin, 50 μg/mL chloramphenicol and 2% glucose. Starter cultures were diluted 1:100 in LB medium with 50 μg/mL kanamycin, 50 μg/mL chloramphenicol and 0.2% glucose and allowed to grow at 37 °C to reach an OD_{600nm} of 0.2-0.3. *ASG* expression was induced with 3 μM IPTG and the bacterial culture was transferred to 30 °C incubator with 250 rpm shaking for 1.5-2 h. The culture was then split into two halves of equal volume, and one half was induced with 50 ng/mL aTc for expression of WT GvpC. Cultures with and without aTc induction were allowed to grow for an additional 20 h at 30°C. Cultures were then spun down at 300 g in a refrigerated centrifuge at 4 °C for 3-4 h in 50 mL conical tubes to isolate buoyant cells expressing GVs from the rest of the culture. The liquid column height was maintained at less than 10 cm to prevent GV collapse by hydrostatic pressure.

In vitro ultrasound imaging

Imaging phantoms were prepared by melting 1% agarose (w/v) in PBS and casting wells using a custom 3-D-printed template mold containing a 2-by-2 grid of cylindrical wells with 2 mm diameter and 1 mm spacing between the outer radii in the bulk material. Ana GV samples from *in vitro* assays or buoyant Nissle cells expressing GVs were mixed 1:1 with 1% molten agarose solution at 42°C and quickly loaded before solidification into the phantom wells. All samples and their controls were OD-matched using the Nanodrop prior to phantom loading, with the final concentration being OD_{500nm} = 2.2 for Ana GVs and OD_{600nm}= 1.0-1.5 for buoyant Nissle cells. Wells not containing sample were filled with plain 1% agarose. Hydrostatic collapse at 1.4 MPa was used to determine that the contribution to light scattering from GVs inside the cells was similar for those expressing the acoustic sensor gene and its wild-type *ARG* counterpart. The phantom was placed in a custom holder on top

of an acoustic absorber material and immersed in PBS to acoustically couple the phantom to the ultrasound imaging transducer.

Imaging was performed using a Verasonics Vantage programmable ultrasound scanning system and a L22-14v 128-element linear array Verasonics transducer, with a specified pitch of 0.1 mm, an elevation focus of 8 mm, an elevation aperture of 1.5mm and a center frequency of 18.5 MHz with 67% -6 dB bandwidth. Linear imaging was performed using a conventional B-mode sequence with a 128-ray-lines protocol. For each ray line, a single pulse was transmitted with an aperture of 40 elements. For nonlinear image acquisition, a custom cross-amplitude modulation (x-AM) sequence detailed in an earlier study[15], with an x-AM angle (θ) of 19.5° and an aperture of 65 elements, was used. Both B-mode and x-AM sequences were programmed to operate close to the center frequency of the transducer (15.625 MHz) and the center of the sample wells were aligned to the set transmit focus of 5 mm. Transmitted pressure at the focus was calibrated using a Precision Acoustics fiber-optic hydrophone system. Each image was an average of 50 accumulations. B-mode images were acquired at a transmit voltage of 1.6V (132 kPa), and an automated voltage ramp imaging script (programmed in MATLAB) was used to sequentially toggle between B-mode and x-AM acquisitions. The script acquired x-AM signals at each specified voltage step, immediately followed by a B-mode acquisition at 1.6V (132 kPa), before another x-AM acquisition at the next voltage step. For engineered Ana GVs subjected to *in vitro* protease assays, an x-AM voltage ramp sequence from 4V (230 kPa) to 10V (621 kPa) in 0.2V increments was used. For wild-type Nissle cells expressing GVs, an x-AM voltage ramp sequence from 7.5V (458 kPa) to 25V (1.6 MPa) in 0.5V increments was used. Samples were subjected to complete collapse at 25V with the B-mode sequence for 10 seconds, and the subsequent B-mode image acquired at 1.6V and x-AM image acquired at the highest voltage of the voltage ramp sequence was used as the blank for data processing. There was no significant difference between the signals acquired at specific acoustic pressures during a voltage ramp or after directly stepping to the same pressure (**Fig. 2-S9**).

Due to transducer failure, a replacement Verasonics transducer (L22-14vX) with similar specifications was used in experiments with Δ*clpXP* cells. The transmitted pressure at the focus was calibrated in the same way as the L22-14v. B-mode images were acquired at a transmit voltage of 1.6V (309 kPa), and an x-AM voltage ramp sequence from 6V (502 kPa) to 25V (2.52 MPa) was used. The imaging protocol was otherwise unchanged.

In vivo ultrasound imaging

All *in vivo* experiments were performed on C57BL/6J male mice, aged 14–34 weeks, under a protocol approved by the Institutional Animal Care and Use Committee of the California Institute of Technology. No randomization or blinding were necessary in this study. Mice were anesthetized with 1–2% isoflurane, maintained at 37 °C on a heating pad, depilated over the imaged region, and enema was performed by injecting PBS to expel gas and solid contents in mice colon. For imaging of *E. coli* in the gastrointestinal tract, mice were placed in a supine position, with the ultrasound transducer positioned on the lower abdomen, transverse to the colon such that the transmit focus of 5 mm was close to the center of the colon lumen. Prior to imaging, two variants of buoyancy-enriched *E. coli* Nissle 1917 were mixed in a 1:1 ratio with 4% agarose in PBS at 42 °C, for a final bacterial concentration of 1.5E9 cells ml^{-1}. An 8-gauge gavage needle was filled with the mixture of agarose and bacteria of one cell population. Before it solidified, a 14-gauge needle was placed inside the 8-gauge needle to form a hollow lumen within the gel. After the agarose–bacteria mixture solidified at room temperature for 10 min, the 14-gauge needle was removed. The hollow lumen was then filled with the agarose–bacteria of the other cell population. After it solidified, the complete cylindrical agarose gel was injected into the colon of the mouse with a PBS back-filled syringe. For the colon imaging, imaging planes were selected to avoid gas bubbles in the field of view. In all *in vivo* experiments, three transducers were used, including two L22-14v and one L22-14vX, due to transducer failures unrelated to this study. B-mode images were acquired at 1.9V (corresponding to 162 kPa in water) for L22-14v, and 1.6V (309 kPa in water) for L22-14vX. x-AM images were acquired at 20V (1.27 MPa in water) for L22-14v and 15V (1.56 MPa in water) for L22-14vX, with other parameters being the

same as those used for *in vitro* imaging. B-mode anatomical imaging was performed at 7.4V using the 'L22-14v WideBeamSC' script provided by Verasonics.

Image processing and data analysis

All *in vitro* and *in vivo* ultrasound images were processed using MATLAB. Regions of interest (ROIs) were manually defined so as to adequately capture the signals from each sample well or region of the colon. The sample ROI dimensions (1.2 mm × 1.2 mm square) were the same for all *in vitro* phantom experiments. The noise ROI was manually selected from the background for each pair of sample wells. For the *in vivo* experiments, circular ROIs were manually defined to avoid edge effects from the skin or colon wall, and the tissue ROIs were defined as the rest of the region within the same depth range of the signal ROIs. For each ROI, the mean pixel intensity was calculated, and the pressure-sensitive ultrasound intensity ($\Delta I = I_{\text{intact}} - I_{\text{collapsed}}$) was calculated by subtracting the mean pixel intensity of the collapsed image from the mean pixel intensity of the intact image. The contrast-to-noise ratio (CNR) was calculated for each sample well by taking the mean intensity of the sample ROI over the mean intensity of the noise ROI. The x-AM by B-mode ratio at a specific voltage (or applied acoustic pressure) was calculated with the following formula:

$$\frac{\Delta \mathrm{I}_{\mathrm{x-AM}}(\mathrm{V})}{\Delta \mathrm{I}_{\mathrm{B-mode}}(\mathrm{V})}$$

where $\Delta \mathrm{I}_{\mathrm{x-AM}}(\mathrm{V})$ is the pressure-sensitive nonlinear ultrasound intensity acquired by the x-AM sequence at a certain voltage V, and $\Delta \mathrm{I}_{\mathrm{B-mode}}(\mathrm{V})$ is the pressure-sensitive linear ultrasound intensity of the B-mode acquisitions at 1.6V (132 kPa) following the x-AM acquisitions at the voltage V. All images were pseudo-colored (bone colormap for B-mode images, hot colormap for x-AM images), with the maximum and minimum levels indicated in the accompanying color bars.

Statistical analysis

Data is plotted as the mean ± standard error of the mean (SEM). Sample size is N=3 biological replicates in all *in vitro* experiments unless otherwise stated. For each biological replicate, there were technical replicates to accommodate for variability in experimental

procedures such as sample loading and pipetting. SEM was calculated by taking the values for the biological replicates, each of which was the mean of its technical replicates. The numbers of biological and technical replicates were chosen based on preliminary experiments such that they would be sufficient to report significant differences in mean values. Individual data for each replicate is given in **Fig. 2-S1-9** in the form of scatter plots. P values, for determining the statistical significance for the *in vivo* data, were calculated using a two-tailed paired t-test.

2.5 Supplementary Information

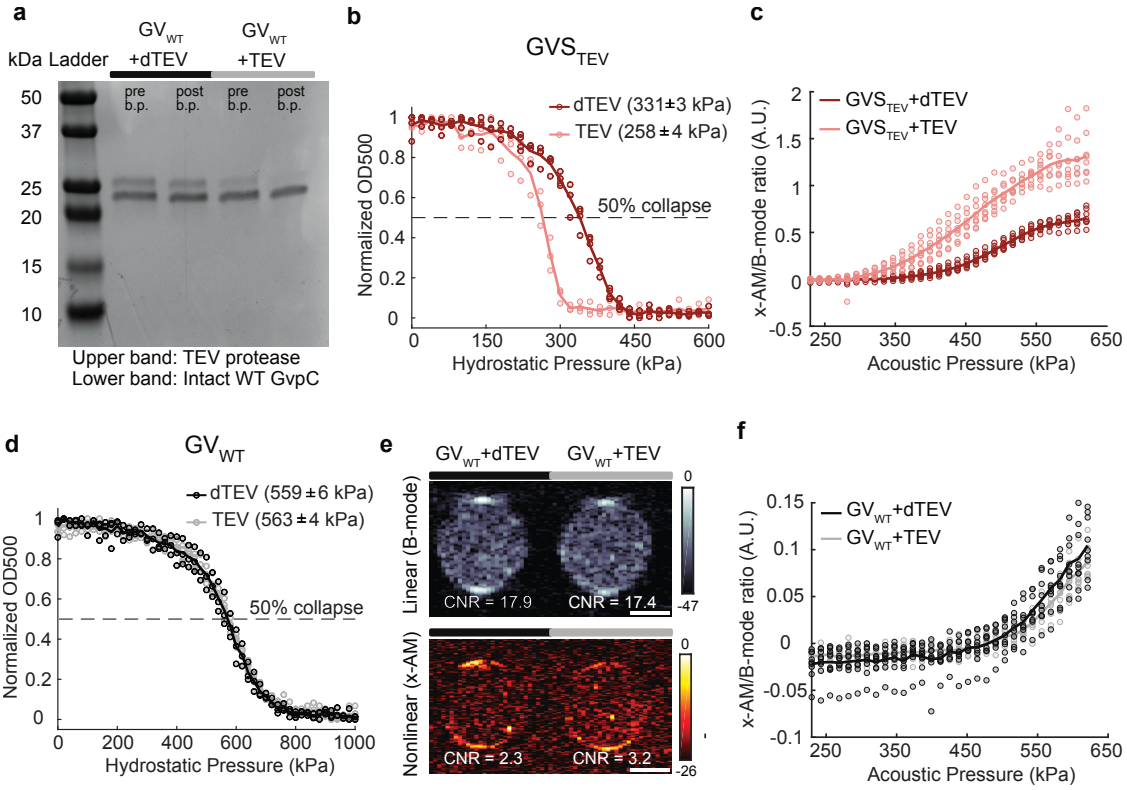


Figure 2-S1. Engineering an acoustic sensor of TEV endopeptidase activity. (**a**) Coomassie-stained SDS-PAGE gel of OD_{500nm}-matched samples of GV_{WT} incubated with dTEV and TEV protease, before and after buoyancy purification (labeled pre b.p. and post b.p., respectively). N = 3 biological replicates. (**b**) Scatter plots showing normalized OD_{500nm} of GVS_{TEV} as a function of hydrostatic pressure. (N = 3 biological replicates for GVS_{TEV} + TEV and N =4 for GVS_{TEV} + dTEV.) (**c**) Scatter plots showing the ratio of nonlinear (x-AM) to linear (B-mode) ultrasound signal as a function of applied acoustic pressure for all the replicate samples used in the x-AM voltage ramp imaging experiments for GVS_{TEV}. N = 3 biological replicates and total number of replicates is 8. (**d**) Scatter plots showing normalized OD_{500nm} of GV_{WT} as a function of hydrostatic pressure. (N = 3 biological replicates for GV_{WT} +dTEV and N =4 for GV_{WT} + TEV.) (**e**) Representative ultrasound images of agarose phantoms containing GV_{WT} incubated with TEV or dTEV protease at OD_{500nm} 2.2. The B-mode image was acquired at 132kPa and the x-AM image at 569 kPa. Similar images acquired for N=3 biological replicates, with each N consisting of 3 technical replicates. CNR stands for contrast-to-noise-ratio, and color bars represent

relative ultrasound signal intensity on the dB scale. Scale bars represent 1 mm (**f**) Scatter plots showing the ratio of nonlinear (x-AM) to linear (B-mode) ultrasound signal as a function of applied acoustic pressure for all the replicate samples used in the x-AM voltage ramp imaging experiments for GV_{WT}. N=3 biological replicates, with each N consisting of 3 technical replicates. Solid line represents the mean of all the replicates.

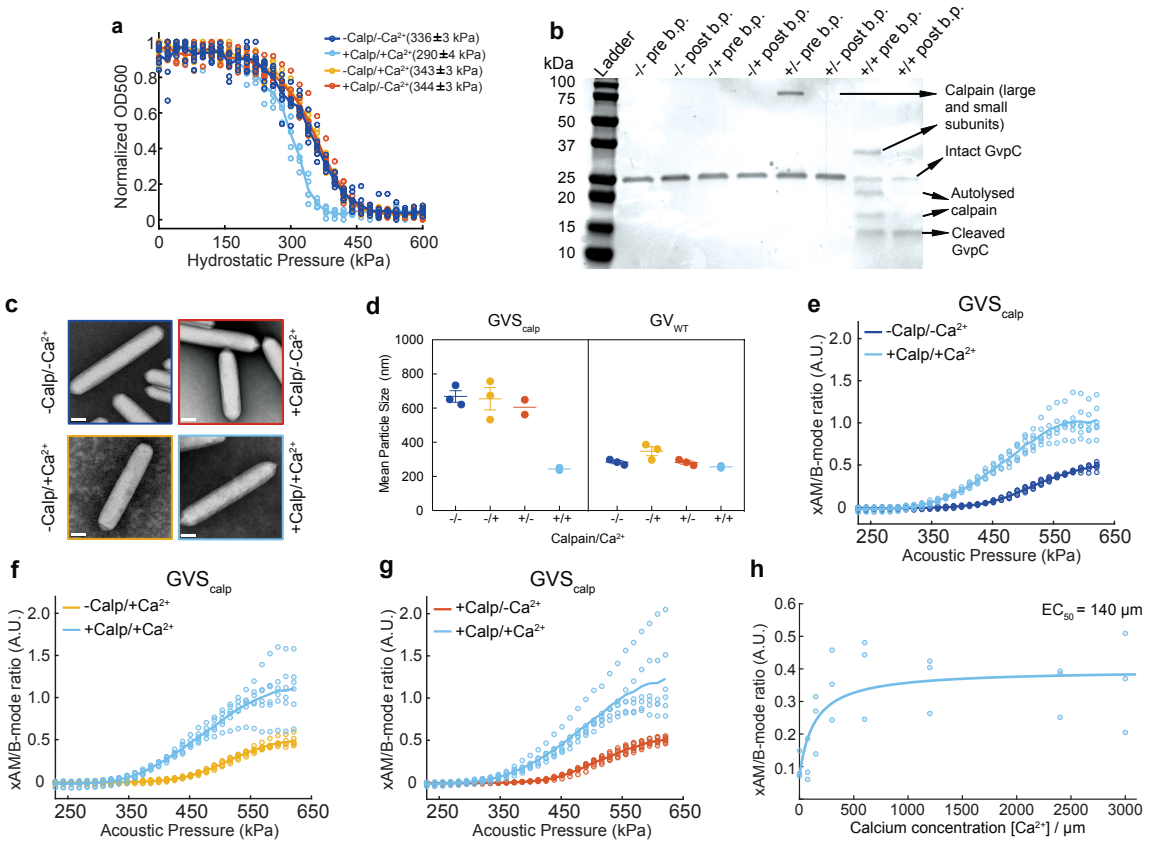


Figure 2-S2. Engineering an acoustic sensor of calpain activity. (a) Individual scatter plots for **Fig. 2(b)**. N = 5 biological replicates for +Calp/+Ca^{2+}, 6 for -Calp/+Ca^{2+} and +Calp/-Ca^{2+}, 7 for -Calp/-Ca^{2+}. **(b)** Coomassie-stained SDS-PAGE gel of OD_{500nm}-matched samples of GVS_{calp} incubated in the presence (+) or absence (-) of calpain (first +/-) and calcium (second +/-), before and after buoyancy purification (labeled pre b.p. and post b.p., respectively). N = 3 biological replicates. (**c**) Representative TEM images of GVS_{calp} after incubations in the presence or absence of calpain and/or calcium. Scale bars represent 100 nm. At least 20 GV particles were imaged for each condition. (**d**) DLS measurements showing the average hydrodynamic diameter of GVS_{calp} and GV_{WT} samples after calpain/calcium incubations (N = 2 biological replicates for GVS_{calp} +/-, +/+, GV_{WT} +/+ and 3 for other conditions, individual dots represent each N and horizontal line indicates the mean). Error bars indicate SEM when N = 3. (**e, f, g**) Individual scatter plots for **Fig. 2(d, f, h)**. N = 3 biological replicates with each N consisting of 2 technical replicates (total number of replicates is 18 for +/+ and 6 for each of the remaining conditions). Solid line represents the mean of all the replicates for (**a, e-g**). (**h**) Scatter plots for **Fig. 2i**; N = 3 biological replicates, individual dots represent each N and solid blue line showing the fitted curve (a Hill equation with a coefficient of 1, with a half maximum response concentration (EC_{50}) of 140 µm).

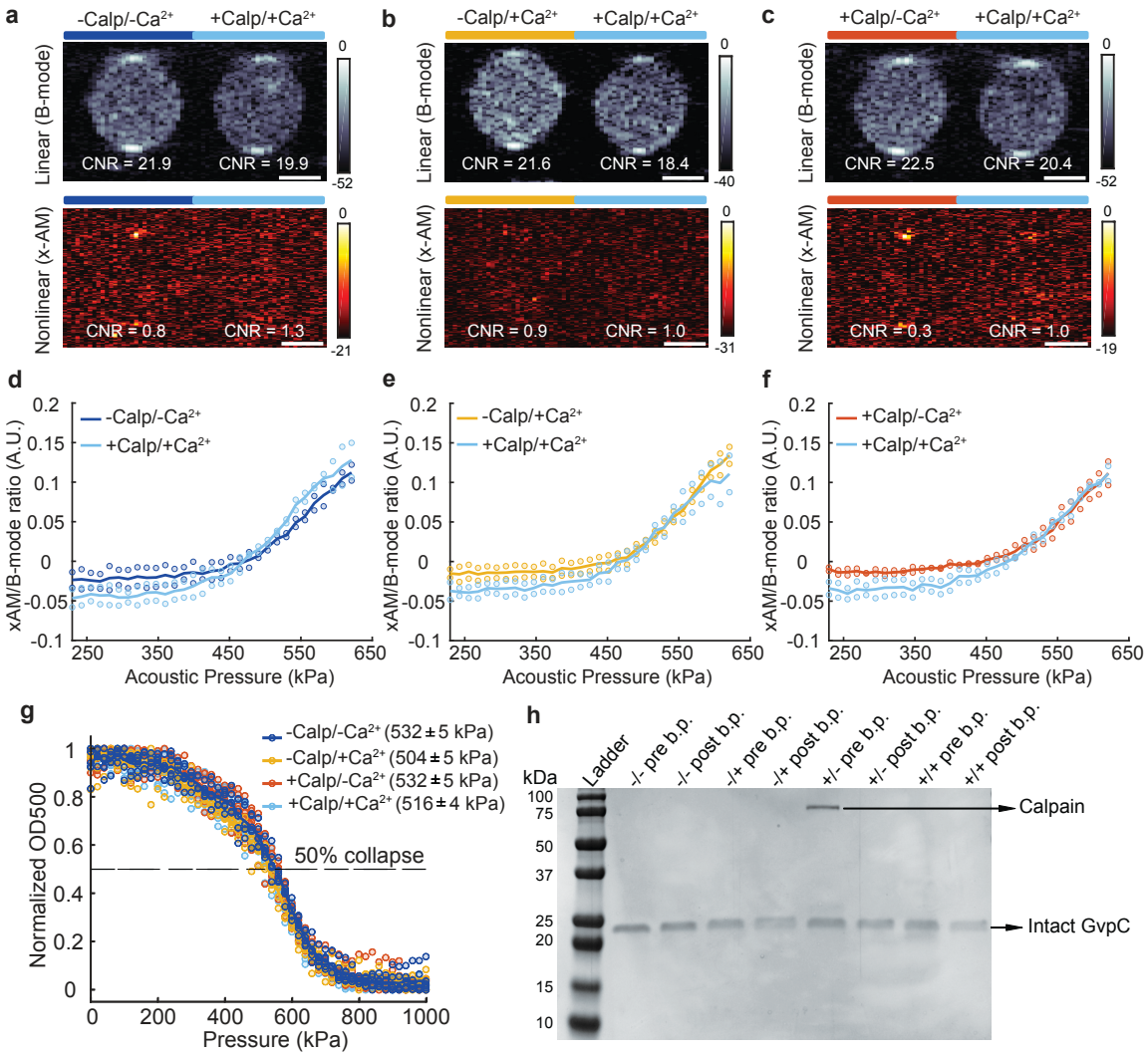


Figure 2-S3. Characterization of GV_{WT} sample with calpain protease. (**a**, **b**, **c**) Representative ultrasound images of agarose phantoms containing GV_{WT} incubated in the presence (+) or absence (-) of calpain (first +/-) and calcium (second +/-), at OD_{500nm} 2.2. The B-mode images were taken at 132 kPa for **a**, **b** and **c** and the x-AM images corresponding to the maximum difference in non-linear contrast between the +/+ sample and the negative controls were taken at 438 kPa for **a** and **b** and at 425 kPa for **c**. CNR stands for contrast-to-noise-ratio and color bars represent ultrasound signal intensity in the dB scale. Scale bars represent 1 mm. N = 2 biological replicates for (**a-c**). (**d**, **e**, **f**) Scatter plots showing the ratio of x-AM to B-mode ultrasound signal as a function of increasing acoustic pressure for GV_{WT} after incubation in the presence or absence of calpain and/or calcium (N = 2 biological replicates). (**g**) Hydrostatic collapse curves of GV_{WT} after incubations in the presence (+) or absence (-) of calpain and/or calcium. The legend lists the midpoint collapse pressure for each condition (±95% confidence interval) determined from fitting a Boltzmann sigmoid function (N = 5 biological replicates for -/+ and N = 6 for other conditions) (**h**) Coomassie-stained SDS-PAGE gel of OD_{500nm}-matched samples of GV_{WT} incubated in the presence (+) or absence (-) of calpain/calcium, before and after buoyancy purification (labeled pre b.p. and post b.p., respectively, N=1). Individual dots in **d**, **e**, **f,** and **g** represent each N and solid line represents the mean of all the replicates.

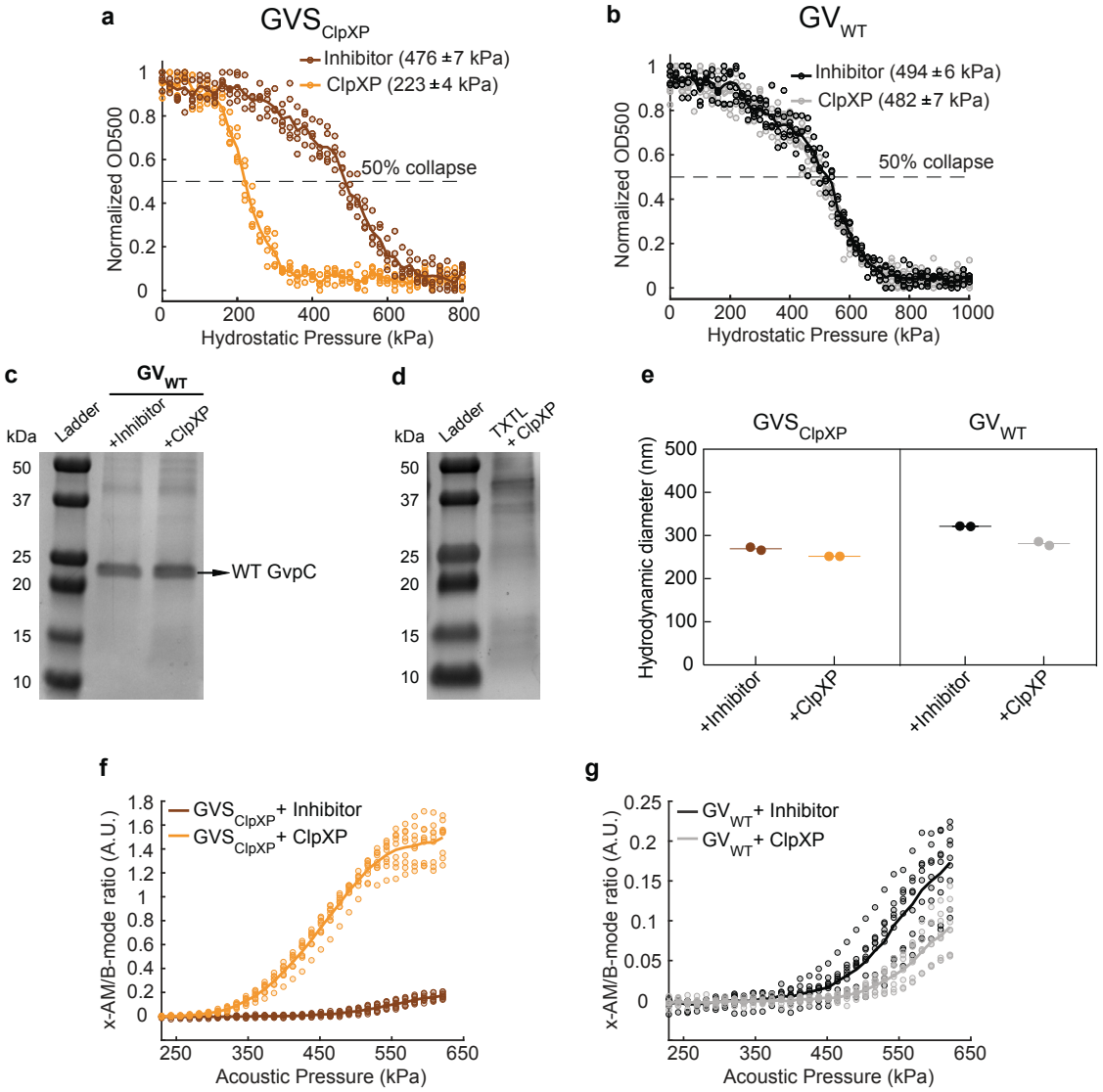


Figure 2-S4. Engineering an acoustic sensor of ClpXP proteolytic activity. (**a, b**) Scatter plots for Figure. 2-3(**d, g**). N = 5 biological replicates. (**c**) Coomassie-stained SDS-PAGE gel of OD_{500nm}-matched GV_{WT} samples incubated in a reconstituted cell-free transcription-translation (TX-TL) system containing a protease inhibitor cocktail or ClpXP. N = 3 biological replicates. (**d**) Coomassie-stained SDS-PAGE gel of 30x diluted content of TX-TL system containing ClpXP. N = 2 biological replicates(**e**) DLS measurements showing the average hydrodynamic diameter of GVS_{ClpXP} and GV_{WT} samples, after incubations with protease inhibitor or ClpXP (N = 2 biological replicates, individual dots represent each N and horizontal line indicates the mean). (**f**, **g**) Scatter plots showing the ratio of x-AM to B-mode acoustic signal as a function of applied acoustic pressure for all the replicate samples used in the x-AM voltage ramp experiments for GVS_{ClpXP} (**f**) and GV_{WT} (**g**). N = 3 biological replicates, with each N consisting of 3 technical replicates. Individual dots represent each N and solid line represents the mean of all the replicates for (**a-b, f-g**).

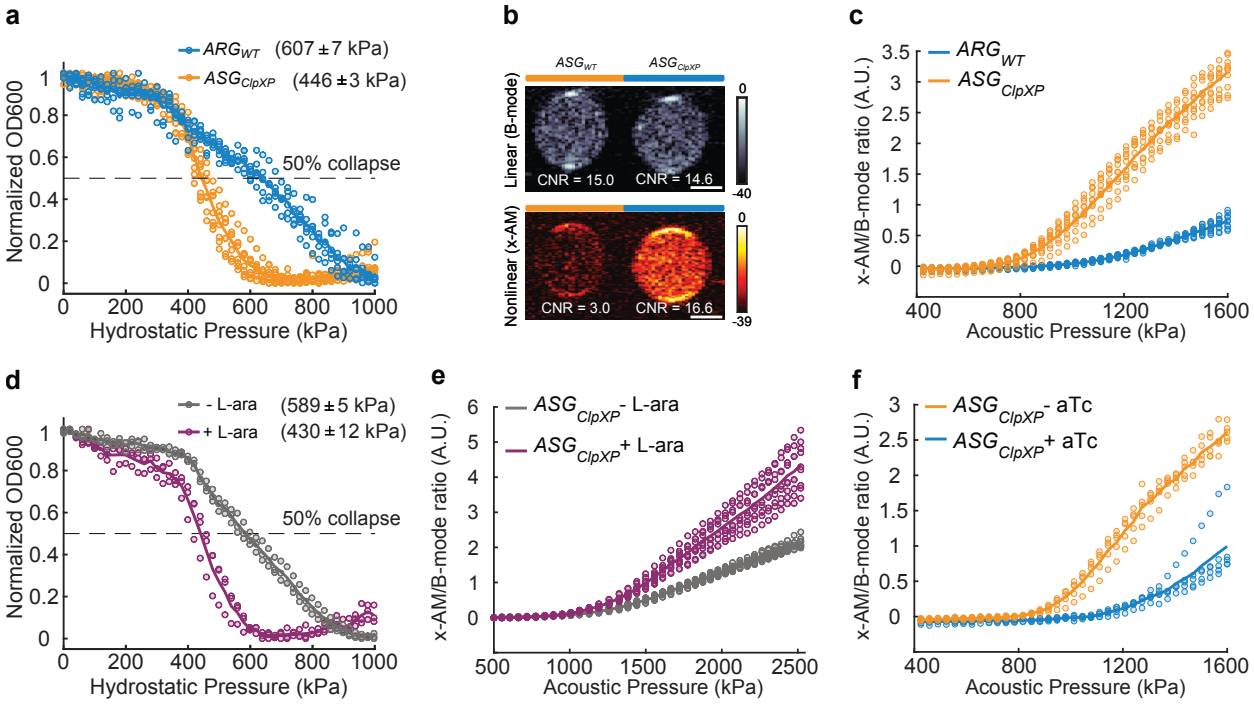


Figure 2-S5. Constructing intracellular acoustic sensor genes for dynamic monitoring of protease activity and circuit-driven gene expression. (**a**) Normalized pressure-sensitive optical density at 600 nm of WT Nissle cells expressing either *ARG*$_{WT}$ or *ASG*$_{ClpXP}$. The legend lists the midpoint collapse pressure for each cell type (±95% confidence interval) determined from fitting a Boltzmann sigmoid function (N = 5 biological replicates and 8 total replicates for *ASG*$_{ClpXP}$; N = 3 biological replicates for *ARG*$_{WT}$ and 6 total replicates). (**b**) Representative ultrasound images of WT Nissle cells expressing either *ARG*$_{WT}$ or *ASG*$_{ClpXP}$ at OD_{600nm} 1.5 (N = 4 biological replicates and the number of total replicates is 10). (**c**) Scatter plots showing x-AM/B-mode ratio as a function of applied acoustic pressure for WT Nissle cells expressing either *ARG*$_{WT}$ or *ASG*$_{ClpXP}$ at OD_{600nm} 1.5 (N = 4 biological replicates and the number of total replicates is 10). (**d**) Scatter plots for Figure 2-4b, N = 3 biological replicates. (**e, f**) Scatter plots showing the ratio of x-AM to B-mode acoustic signal as a function of acoustic pressure for all the replicate samples used in the x-AM voltage ramp experiments for Δ*clpXP* Nissle cells expressing *ASG*$_{ClpXP}$ and araBAD driven *clpXP*, with or without L-arabinose induction (**e**) and WT Nissle cells expressing *ASG*$_{ClpXP}$ and pTet-TetO driven WT *gvpC*, with or without aTc induction (**f**). N = 3 biological replicates, with each N having 3 technical replicates for (**e**) and N = 5 biological replicates for (**f**). Individual dots represent each N and solid line represents the mean of all the replicates for (**a**, **c**, **d**-**f**).

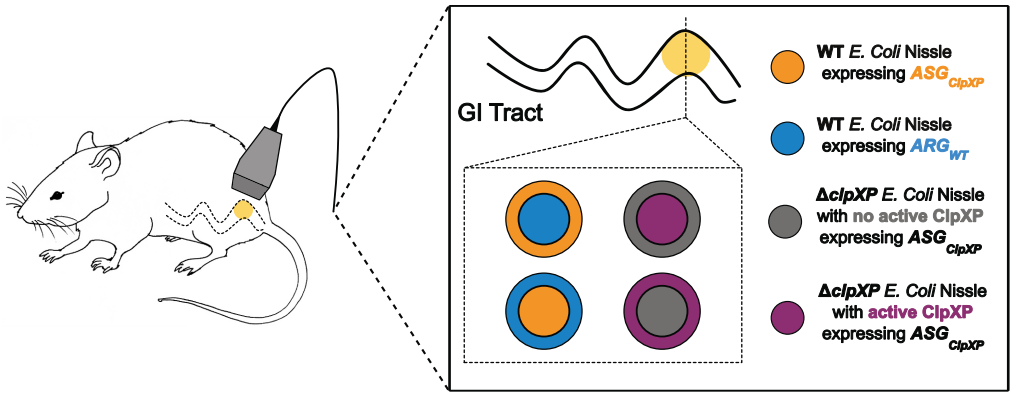


Figure 2-S6. Schematic illustrating the *in vivo* ultrasound imaging experiment. Cells in cylindrical hydrogel with the indicated cross-sectional arrangements were injected into the GI tract of mice and imaged with ultrasound.

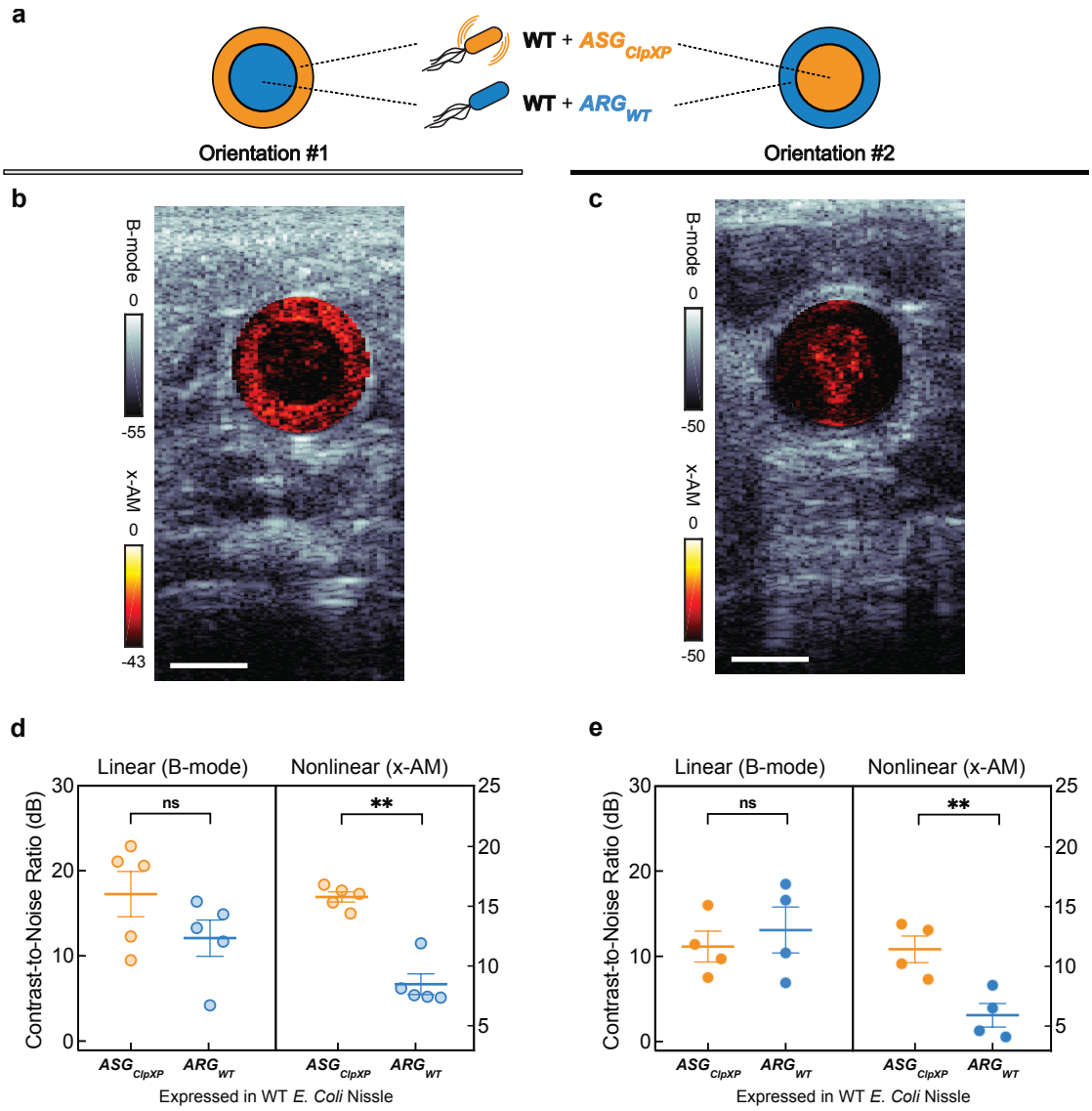


Figure 2-S7. Ultrasound imaging of bacteria expressing acoustic sensor genes in the gastrointestinal tract of mice. (**a**) Schematic illustrating two orientations of the wild type (WT) *E. coli* Nissle cells expressing ARG_{WT} or ASG_{ClpXP} introduced into the mouse colon as a hydrogel. (**b, c**) Representative transverse ultrasound images of the colon for two mice used in the *in vivo* imaging experiments, with orientation #1 (**b**) and with orientation #2. (**c**). Cells are injected at a final concentration of 1.5E9 cells ml^{-1}. B-mode signal is displayed using the bone colormap and x-AM signal is shown using the hot colormap. Color bars represent B-mode and x-AM ultrasound signal intensity in the dB scale. Scale bars represent 2 mm. (**d, e**) B-mode and xAM contrast-to-noise ratio (CNR) *in vivo*, for WT Nissle cells expressing ARG_{WT} or ASG_{ClpXP} in orientation #1 (**d**) and orientation #2. (**e**). N = 5 mice for orientation #1 (**b, d**) and N = 4 mice for orientation #2 (**c, e**). Error bars indicate SEM. P = 0.0014 for x-AM signal from cells expressing ASG_{ClpXP} versus the ARG_{WT} control in orientation #1, and P = 0.0016 for that in orientation #2. P = 0.0570 for B-mode signal in orientation #1 and P = 0.3445 in orientation #2. P-values were calculated using a two-tailed paired t-test. Individual dots represent each N and horizontal line indicates the mean.

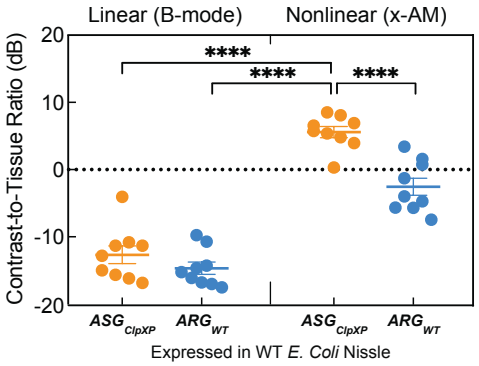


Figure 2-S8. *ASG$_{ClpXP}$* -expressing cells showed higher contrast to tissue with nonlinear imaging. B-mode and xAM contrast-to-tissue ratio (CTR) *in vivo*, for WT Nissle cells expressing *ARG$_{WT}$* or *ASG$_{ClpXP}$* in both orientations. P < 0.0001 for the CTR from xAM imaging of cells expressing *ASG$_{ClpXP}$* versus CTR from xAM imaging of cells expressing *ARG$_{WT}$* (P = 7.8E-5), B-mode imaging of cells expressing *ASG$_{ClpXP}$* (P = 1.4E-6) and *ARG$_{WT}$* (P = 4.9E-7). Individual dots represent each N, and the thick horizontal line indicates the mean. Error bars indicate SEM. N = 9 mice. P-values were calculated using a two-tailed paired t-test for each comparison independently. Individual dots represent each N and horizontal line indicates the mean.

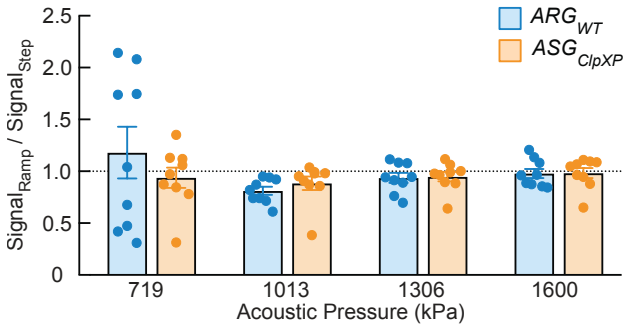


Figure 2-S9. Absence of memory effect from imaging at sequentially increasing acoustic pressure. Ratio of sensor-specific signal (xAM/B-mode) acquired at the indicated acoustic pressures in the process of voltage ramping (comprising 36 points from 458 kPa to 1.6 MPa) or stepping the transducer output directly to corresponding pressure in a single step, for WT Nissle cells expressing either *ARG$_{WT}$* or *ASG$_{ClpXP}$*. N =3 biological replicates, with each N having 3 technical replicates. Individual dots represent each replicate, and the thick horizontal line indicates the mean. Error bars indicate SEM derived from biological replicates (see Online Methods).

Plasmid Name	Description and Purpose	Transcriptional regulators	Output gene product(s)	Insertions/Tags (including linkers)	Reference Information
WT C-His Ana GvpC in pET28a	Ana *gvpC* used as wild-type control for TEV and calpain sensor	pT7, LacO	WT C-His Ana GvpC	SLE-His6 at C-terminus	Addgene ID# 85732
WT N-His Ana GvpC in pET28a	Ana *gvpC* used as wild-type control for ClpXP sensor	pT7, LacO	WT N-His-Ana-GvpC	G-His6-SG at N-terminus	Addgene ID# 153294
C-His-GvpC-TEV	Ana *gvpC* with TEV cleavage site	pT7, LacO	C-His Ana GvpC with TEV cleavage site	SLE-His6 at C-terminus, GSGSGSG-ENLYFQG-SGSGSG in GvpC repeat 2	Addgene ID# 153296
C-His-GvpC-Calpain	Ana *gvpC* with calpain cleavage site	pT7, LacO	C-His Ana GvpC with calpain cleavage site	SLE-His6 at C-terminus, GSGSG-QQEVYGMMPRD-GSGSG in GvpC repeat 2	Addgene ID# 153295
N-His-GvpC-ssrA	Ana *gvpC* with ssrA degradation tag	pT7, LacO	N-His Ana GvpC with ssrA degradation tag	G-His6-SG at N-terminus, SG-AANDENYALAA at C-terminus	Addgene ID# 153297
pBEST_OR2_OR1-Pr-UTR1_ClpP-T500	*clpP* plasmid for use in the cell-free TX-TL system	OR2-OR1-Pr	ClpP		Addgene ID# 153302
pACYC-FLAG-dN6-His	*clpX* plasmid for use in the cell-free TX-TL system	pT7, LacO	Flag-ClpXdeltaNLinkedHexamer-His6	Flag tag at N-terminus His6 at C-terminus L20 linkers	Addgene ID# 22143
pET28a_T5-ARG1	Original acoustic reporter gene construct (ARG_{WT})	pT5, LacO	Ana GvpA, WT Ana GvpC, Mega GvpR-U		Addgene ID # 106476
pET28a-T5-ASG_ClpXP	Acoustic sensor gene for ClpXP (ASG_{ClpXP})	pT5, LacO	Ana GvpA, dGvpC, Mega GvpR-U	SG-AANDENYALAA at C-terminus	Addgene ID# 153299
pKD3	Frt-flanked cat cassette for recombineering		CAT		Addgene ID # 45604
pKD46	Plasmid that carries the Lambda Red recombineering system	pBAD, araBAD operon	Gam, Beta, Exo		Coli Genetic Stock Center
pE-FLP	*FLP* recombinase to remove the integration	pE	FLP		Addgene ID # 45978

	module in reombineering				
araBAD-BCD20-ClpP-BCD17-ClpX	Expression of *clpX* and *clpP* under araBAD promoter	pBAD, araBAD operon	ClpX, ClpP		Addgene ID # 153301
pTetR-BCD2-Ana GvpC	Wild-type Ana *gvpC* under Tet promoter	pTet, TetO	WT Ana GvpC		Addgene ID # 153298

Table 2-S1. List and features of genetic constructs used in Chapter 2.
pKD3[86] was a gift from Barry L. Wanner (Addgene plasmid # 45604 ; http://n2t.net/addgene:45604 ; RRID:Addgene_45604). pKD46[86] was obtained from the Coli Genetic Stock Center (CGSC, https://cgsc.biology.yale.edu/Site.php?ID=64672).
pE-FLP[87] was a gift from Drew Endy & Keith Shearwin (Addgene plasmid # 45978;http://n2t.net/addgene:45978 ; RRID:Addgene_45978).
The pBEST_OR2-OR1-Pr-UTR1_ClpP-T500 was a gift from Zachary Sun and Richard Murray[90].
pACYC-FLAG-dN6-His[91] was a gift from Robert Sauer (Addgene plasmid # 22143 ; http://n2t.net/addgene:22143 ; RRID:Addgene_22143)
The pTARA backbone was modified to make the araBAD-BCD20-ClpP-BCD17-ClpX and pTEtR-BCD2-Ana GvpC constructs. pTARA[92] was a gift from Kathleen Matthews (Addgene plasmid # 31491; http://n2t.net/addgene:31491 ; RRID:Addgene_31491).

.

Chapter 3

ULTRASONIC REPORTERS OF CALCIUM (UROCS) FOR IMAGING DEEP TISSUE CELLULAR DYNAMICS

Jin, Z., Lakshmanan, A., Zhang, R., Tran, T.A., Rabut, C., Dutka, P., Duan, M., Hurt, R.C., Malounda, D., Yao, Y., et al. (2023). Ultrasonic reporters of calcium for deep tissue imaging of cellular signals. *bioRxiv*. DOI: 10.1101/2023.11.09.566364 10.1101/2023.11.09.566364.

This chapter is a reformatted version of the above manuscript. My contribution to the work were designing and performing the study, analyzing the data, and preparing the manuscript.

3.1 Introduction

Calcium plays a central role in many essential cellular functions, ranging from egg fertilization[93] and developmental politization to synaptic transmission[94,95], hormone secretion[96,97], immune activation[98–100], cytotoxicity[101,102] and G-protein coupled receptor (GPCR) signal transduction across a wide variety of cell types[103]. Due to the importance of this signaling molecule, calcium imaging with the state-of-the-art synthetic[104] and genetically encoded calcium indicators (GECIs)[100,105–108] has proven to be of immense value in biological research[109–113]. However, light scattering in tissue makes it challenging to use fluorescent calcium sensors at scale in intact organisms[8]. In most cases, it is limited to single-cell imaging in optically clear or surgically accessed tissues, or lower-sensitivity and lower-resolution imaging in larger and deeper areas in non-transparent organisms[114]. The ability to image calcium at sufficient resolution and depth in the context of intact living organisms would promote our understanding of fundamental physiology and facilitate the development of novel diagnostic and therapeutic agents.

Compared to other prevalent non-invasive imaging modalities such as magnetic resonance imaging[115,116] and photoacoustics[106,117–119], ultrasound provides an unparalleled combination of penetration depth (centimeters), imaging volume (multiple cm^3), spatiotemporal resolution (~100 μm and ~1 ms), accessibility, and compatibility with freely moving

experimental subjects[20,22]. Recently, the first genetically encoded ultrasound contrast agents were introduced based on gas vesicles (GVs), a unique class of air-filled protein nanostructures derived from buoyant microbes[120]. GVs produce ultrasound contrast due to the low density and high compressibility of their gaseous core relative to aqueous tissues[120], and their heterologous expression in bacteria[40,41] or mammalian cells[40,121] allows GVs to serve as acoustic reporter genes.

GVs' ultrasound contrast depends on the composition and mechanics of their protein shell[33,37]. In particular, an alpha-helical protein called GvpC sits on the outside of the shell and stiffens the GV against deformation under acoustic pressure; removal of this protein results in reversible GV buckling and enhanced nonlinear contrast[37,33,34,44,31]. The first GV-based acoustic biosensors — of protease activity — were developed by engineering GvpC to contain specific protease recognition sites, such that cleavage by the cognate protease makes the GV shell more flexible, leading to stronger nonlinear contrast[122]. While these protease sensors heralded an important advance in biomolecular ultrasound[20,22,123], they produced only a one-time, irreversible change in acoustic contrast due to the permanent covalent modification of the sensor and were only shown to function in bacteria.

Here, we set out to develop the first dynamic, reversible, allosteric acoustic biosensor – an ultrasonic reporter of calcium (URoC) – to enable noninvasive, deep-tissue calcium imaging in mammalian cells (**Fig. 3-1a**). Inspired by fluorescent GECIs[124], we hypothesized that we could engineer GvpC to incorporate calmodulin (CaM) and a calmodulin-binding peptide (CBP), such that calcium binding would result in a reversible conformational change that weakens GvpC's binding to the GV shell. This would make the GV more flexible and increase its nonlinear acoustic response (**Fig. 3-1a-b**). After engineering and systematically characterizing a range of URoC designs, we obtained an acoustic biosensor that produces a 4.7-fold increase in nonlinear contrast in response to calcium *in vitro*, with a sensitivity midpoint of 113 nM. We showed that this construct can be fully genetically encoded and functional in mammalian cells, producing a contrast enhancement of more than 170% in response to elevated intracellular calcium. To validate the performance of the URoC *in vivo*, we used it to image intracellular calcium signaling resulting from drug-induced activation of

a G protein-coupled receptor (GPCR) in cells implanted deep in the mouse brain. We showed that the pharmacodynamic response of these cells could be monitored noninvasively in real time, through intact skull, with a spatial resolution of ~100 μm. This first-generation URoC demonstrates that acoustic biosensors can be engineered to respond reversibly to intracellular signals, opening the window for ultrasound to follow a variety of dynamic cellular processes noninvasively in intact living animals.

3.2a Design and Characterization of URoCs

As the first step, we designed the GvpC to undergo a calcium-dependent conformational change that alters its binding to the GV shell and, consequently, its ability to strengthen the GV. Inspired by the fluorescent GECIs[124], our URoC design comprises a CBP incorporated in the middle of GvpC's alpha-helical structure and CaM at its C-terminus (**Fig. 3-1b**). We hypothesized that the calcium-free CaM would have negligible effect on GvpC-GV binding, while Ca^{2+}-bound CaM would interact with the CBP, triggering the engineered GvpC to undergo an allosteric conformational change resulting in reduced GV shell stiffness, increased buckling under acoustic pressure, and elevated non-linear ultrasound contrast **(Fig. 3-1b**). This mechanism is expected to be fully reversible and compatible with continuous imaging of calcium dynamics.

Based on previous experience with GvpC engineering[122], we replaced part of the second repeat of a 3-repeat Ana GvpC (3R GvpC) with a CBP and fused CaM to the C-terminus (**Fig. 3-1b, c**). We built and tested a library of 15 initial URoC designs, incorporating the CBP from CaMKI[125] at different positions within the second GvpC repeat and fusing a mutant CaM from GCaMP6f[126,127] (CaM6f-EF1-KO) to the C-terminus of GvpC using a long flexible linker (8 repeats of G4S, or GGGGS) (**Fig. 3-1c**). We first tested these GvpC designs using purified GVs isolated from *Anabaena flos-aquae* (Ana) by biochemically replacing their GvpC[64]. We measured the Ca^{2+}-dependent nonlinear ultrasound contrast of these engineered GV sensors in agarose phantoms at 37˚C using a cross-propagating amplitude modulation pulse sequence (xAM) at 15.625 MHz[34]. Prior to imaging, samples were incubated for 10 minutes at 37˚C with 200 μM Ca^{2+}, 5 mM of the calcium chelator EGTA, or first with 200 μM Ca^{2+} and then reversed with 5 mM EGTA. As controls, GVs with

unmodified 3R GvpC or without GvpC were tested under the same conditions and showed virtually no calcium-dependent nonlinear contrast (**Fig. 3-S1a**).

Most of our 15 variants stiffened the GV shell in the calcium-free condition, showing low nonlinear signal similar to control GVs with 3R GvpC (**Fig. 3-1a, Fig. 3-S1a**). Nine variants (#1-3, #8-10, #12-14) showed significant enhancement in nonlinear contrast in response to calcium (**Fig. 3-1d, Fig. 3-S1b**). However, among these nine, only variants #1 and #9 showed a complete return to the calcium-free baseline upon sequestration of free Ca^{2+} by EGTA (**Fig. 3-1d, Fig. 3-S1c**), demonstrating fully reversibility (within 10 minutes, see Methods). Of these two, the construct with the CBP at position #9 showed the largest dynamic range – its nonlinear signal in calcium-free or EGTA reversal condition was as low as the GV controls with 3R GvpC, while its signal in the calcium-saturated condition was similar to GVs lacking GvpC (**Fig. 3-1d, Fig. 3-S1a**). We selected this construct for further characterization, naming it URoC1a.

Consistent with results from the initial screening, URoC1a produced reversible ultrasound contrast in response to calcium at the physiological temperature of 37˚ C, with a maximum enhancement of approximately 13.5 decibels (dB), or a fold change of 4.7x, in nonlinear ultrasound signal-to-background ratio (SBR) (**Fig. 3-1e, f**). A control "dead" variant of URoC1a – dURoC1a – with all its calcium-binding EF-hands disabled through point mutations, produced low nonlinear signals without any Ca^{2+}-dependence (**Fig. 3-1e, Fig. 3-S1d**). URoC1a showed minimal sensitivity to magnesium (**Fig. 3-S1e),** while maintaining full calcium-sensing functionality within the physiological range of intracellular pH between 7.0 and 7.5[128] (**Fig. 3-S1f**), with diminished performance outside this range, similar to fluorescent GECIs[124,129]. We characterized URoC1a's sensitivity to calcium and measured a half-maximal response (EC50) at around 469 nM (**Fig. 3-1g**), which fits into the physiological range of calcium concentration inside mammalian cells[130].

To measure the kinetics of URoC1a, we built a custom stopped-flow ultrasound imaging apparatus where the URoC1a solution and calcium or EGTA solution were equilibrated to 37˚ C, then delivered into a mixing tee connector, rapidly mixed, and injected into

acoustically transparent tubing for imaging with ultrasound (**Fig. 3-1h-i**). With this apparatus, we were able to measure the kinetics of URoC with a dead time down to 2 seconds (see Methods, **Fig. 3-S1g**), and found that the calcium-saturated half-rise time of URoC1a is ~ 16 seconds, while the half-decay time for reversal with EGTA is around 20 seconds (**Fig. 3-1j**). These results established URoC1a as an acoustic biosensor for calcium with a large dynamic range, physiologically relevant calcium sensitivity and kinetics on the order of 10 seconds[103] – and potential for further optimization.

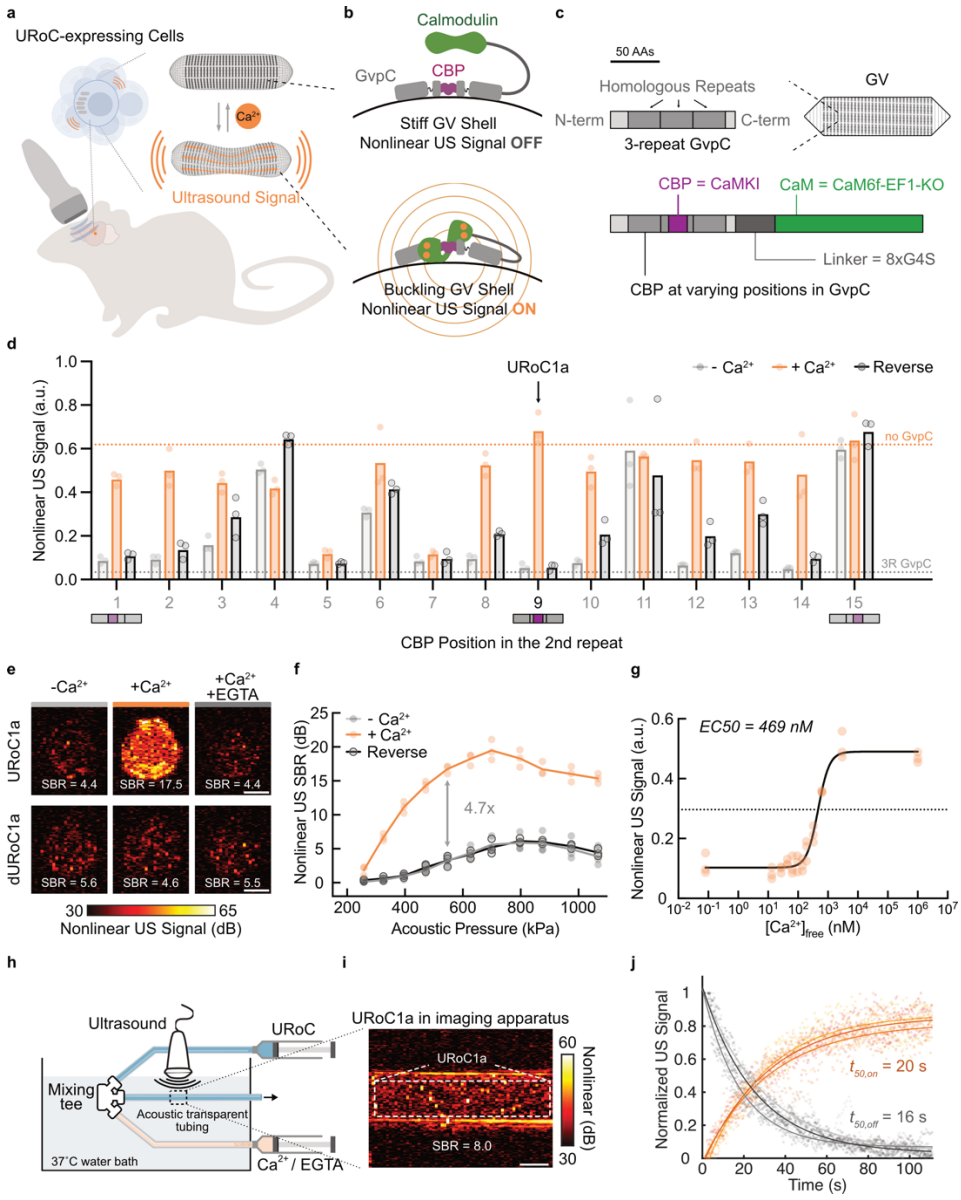


Figure 3-1. Ultrasonic reporter of calcium dynamics. (**a**) Schematics of URoC transducing intracellular calcium dynamics to ultrasound signal, enabling deep-tissue imaging of cellular function with ultrasound. (**b**) Schematic of URoC and its proposed mechanism of action. (**c**) Top: schematic of a 3-repeat Ana GvpC (3R GvpC) structure, comprising three 33-amino acid repeats flanked by N- and C-terminal regions. Bottom: schematic of the molecular design of URoC and its engineering components. (**d**) Nonlinear signal after background subtraction and GV concentration normalization (see Methods) of URoC variants with different CBP insertion positions in the second repeat of GvpC. Samples at OD_{500} 1.8 were incubated with 200 μM $CaCl_2$, 5 mM EGTA, or first with 200 μM $CaCl_2$ and then with 5 mM EGTA before imaging with xAM at 547 kPa. The dashed lines indicate the mean signal of GVs with 3R GvpC (gray) or without GvpC (orange). (**e**) Representative ultrasound images of agarose phantom containing URoC1a or control dURoC1a GVs at OD_{500} 1.8 with 200 μM $CaCl_2$, 5 mM EGTA, or first with 200 μM $CaCl_2$ and then with 5 mM EGTA. The images were acquired with xAM at 547 kPa. (**f**) Nonlinear signal-to-background ratio (SBR) as a function of applied acoustic pressure for URoC1a after incubation with 200 μM $CaCl_2$, 5 mM EGTA, or first with 200 μM $CaCl_2$, and then with 5 mM EGTA. Solid curves represent the mean of all biological replicates. (**g**) Nonlinear signal of URoC1a as a function of free calcium concentration. EC50 of URoC1a determined from fitting a 4-variable Hill equation (solid line). (**h**) Schematic of stopped-flow imaging apparatus. (**i**) Representative ultrasound image of calcium-saturated URoC1a in the stopped-flow apparatus, acquired with xAM at 472 kPa. (**j**) Min-max normalized nonlinear ultrasound signal of URoC1a after being mixed into 1 mM $CaCl_2$ or 5 mM EGTA. Data from each biological replicate was normalized and fitted to an exponential function to estimate the listed mean half-rise and half-decay. Solid lines represent fitted curves from each biological replicate. Dots represent individual measurements of each replicate for **d, g,** and **j**. Dots represent the mean of two technical replicates for **f**. Scale bars = 1 mm. N = 3 biological replicates for (**d-g**, **j**) and each N consists of 2 technical replicates for (**e-f**) and 3 for (**j**).

3.2b Mechanism and Optimization of URoCs

To improve the performance of URoC, we first examined its molecular mechanism of action. In our initial hypothesis, the engineered GvpC would partially dissociate from the GV shell upon calcium binding while part of GvpC would act as an anchor to keep it in proximity to the shell. However, we discovered that the calcium-bound URoC GvpC fully dissociated from the shell, as shown by both cryo-electron microscopy (cryo-EM) and gel electrophoresis (**Fig. 3-2a, Fig. 3-S2a, b**). In cryo-EM images of the longitudinal cross-section of the URoC1a GVs, we observed a dot-shape GvpC density above the α2 helices of GvpA in samples incubated without calcium before freezing (**Fig. 3-2a**). In contrast, when the URoC1a GVs were pre-incubated with calcium, the density of the GvpC disappeared (**Fig. 3-2a**), representing a dissociation of GvpC from the GV shell. This finding was confirmed by denaturing gel electrophoresis, which also showed the disappearance of GvpC from buoyancy-purified (see Methods) GVs after calcium incubation (**Fig. 3-2a**). The control dURoC1a GVs, on the other hand, showed no signs of GvpC dissociation in the presence

and absence of calcium, as seen in both cryo-EM and gel electrophoresis experiments (**Fig. 3-S2a, b**).

After determining that calcium-dependent nonlinear ultrasound contrast originates from the complete dissociation of GvpC, we postulated that the sensitivity and kinetics of URoC1a should depend on the concentration of GvpC relative to its available docking sites on GVs. Indeed, we found that the calcium EC50 of URoC1a decreased to 231 nM with a GvpC concentration of 0.5x relative to the estimated docking sites[37] (**Fig. 3-2b**). In comparison, once the GvpC concentration exceeded the docking site concentration, the EC50 stayed almost constant around 469 nM (**Fig. 3-2b**). In kinetics measurements, increasing the relative GvpC concentration resulted in lower rate constants for the forward reaction (GvpC dissociation), while accelerating the reverse reaction (**Fig. 3-2c**). In both cases, the rate constant showed reasonably linear dependence on GvpC concentration (R squared = 0.77 and 0.95 for the forward and reverse reaction, respectively). These results pointed to calcium-dependent GvpC dissociation and reassociation being rate-limiting steps in sensor function.

Building on these insights, we tested several modifications in the molecular design of URoC1. We started with the length of the linker connecting CaM to GvpC, which presumably dictates the local concentration of CaM near the CBP. Characterizing URoC variants with different lengths of flexible G4S linkers, we found that shorter linkers resulted in higher calcium sensitivity (**Fig. 3-2d**). A clear step-change was observed between 4xG4S and 8xG4S, where the EC50 for the variants with 2xG4S or 4xG4S were 236 nM and 264 nM, respectively, while those for variants with longer linkers were all slightly below 500 nM. Using an estimated length of 3.8 Å per amino acid for G4S[131], this step-change happened between 76 Å and 152 Å. Fittingly, we estimated the distance between the C-terminus of the URoC GvpC and the N-terminus of the CaM in the calcium-bound state to be around 92 Å, assuming GvpC holds a rigid alpha-helical structure (see Methods) (**Fig. 3-S2c**). We hypothesize that when the linker is shorter than this length, the calcium-bound URoC GvpC may adopt a conformation where part of the helical structure of GvpC is disrupted, leading to lower affinity to the GV shell and thus higher sensitivity to calcium. We also found that shorter linkers provided faster forward kinetics, where the shortest 2xG4S linker improved

the sensor kinetics by approximately a factor of 2 to a half-rise time of 10 seconds (**Fig. 3-2e**), while linker length had a negligible effect on reverse kinetics.

Additionally, we tested how EF-hand mutations on the CaM affect the URoC performance, as such mutations were shown to improve the kinetics of certain fluorescent GECIs[127]. We constructed URoC variants with each of the four EF-hands disabled through point mutations (EF*x*-KO, *x* = 1-4) and with the intact CaM from GCaMP6f[126] (No-KO). The EF2-KO and No-KO variants showed similar characteristics, both with better sensitivity and faster kinetics than the URoC1a with EF1-KO (**Fig. 3-2f, g**). The EC50 for EF2-KO and No-KO were 350 nM and 315 nM, respectively, and the half-rise time of both variants decreased to ~10 seconds, with the reverse kinetics insignificantly faster than that of URoC1a (**Fig. 3-2f, g)**. The EF3-KO and EF4-KO variants had lower sensitivity and slower forward kinetics. These results suggest that the second half of CaM (C lobe) may be more dominant than the first half (N lobe) in the actuation of the URoC GvpC.

Hypothesizing that the effects of the linker length and CaM mutations can be synergetic, we combined the shortest linker (2xG4S) and the CaM with the second EF-hand knocked out (EF2-KO) to create URoC1b. As predicted, both the sensitivity and forward kinetics of URoC1b were substantially improved, with an EC50 of 113 nM and a half-rise time of just 3.5 seconds (**Fig. 3-2h, i**), approaching the limits of our stopped-flow kinetics measurement. In the meantime, there was no significant change in the reverse rate constant (half-decay ~18 seconds) (**Fig. 3-2i**). These results provide an improved biosensor — URoC1b — with sensitivity and kinetics suitable for many cellular calcium imaging applications[103]. In addition, they showcase how the modular design of URoCs may enable the future construction of acoustic biosensors with different calcium sensitivity and kinetics for various applications, following a path similar to optical GECIs[132].

Our optimization experiments suggest a conceptual model of URoC operation to inform future engineering. For the forward “on” reaction in response to calcium, the association of calcium-activated CaM and CBP could be rate-limiting, since we observed that the effectively higher local concentration of CaM in constructs with shorter linkers accelerated

the sensor's response (**Fig. 3-2e, j**). In addition, the production of nonlinear contrast may be rate-limited by GvpC dissociation from the GV shell, evidenced by the inverse relationship between excess free GvpC concentration and forward response rates (**Fig. 3-2c, j**). Meanwhile, for the reverse reaction after the removal of free calcium, we identified the re-association of GvpC to the GV shell as a major rate-limiting factor, with baseline return accelerating with higher GvpC concentration (**Fig. 3-2c, k**). In addition, since some EF-hand mutations in the CaM (without altering the GvpC segment) led to dramatically slower reverse kinetics, we speculate that CBP dissociation from the CaM could also be rate-limiting (**Fig. 3-2g, k**). We consider calcium-binding to have a minor contribution to the URoC's response time, since CaM alone binds to calcium ions and changes its conformation within tens of milliseconds[133,134] and even faster with the CBP of CaMKI[135]. Our optimization results, combined with this hypothetical reaction scheme, provide a basis for future URoC engineering.

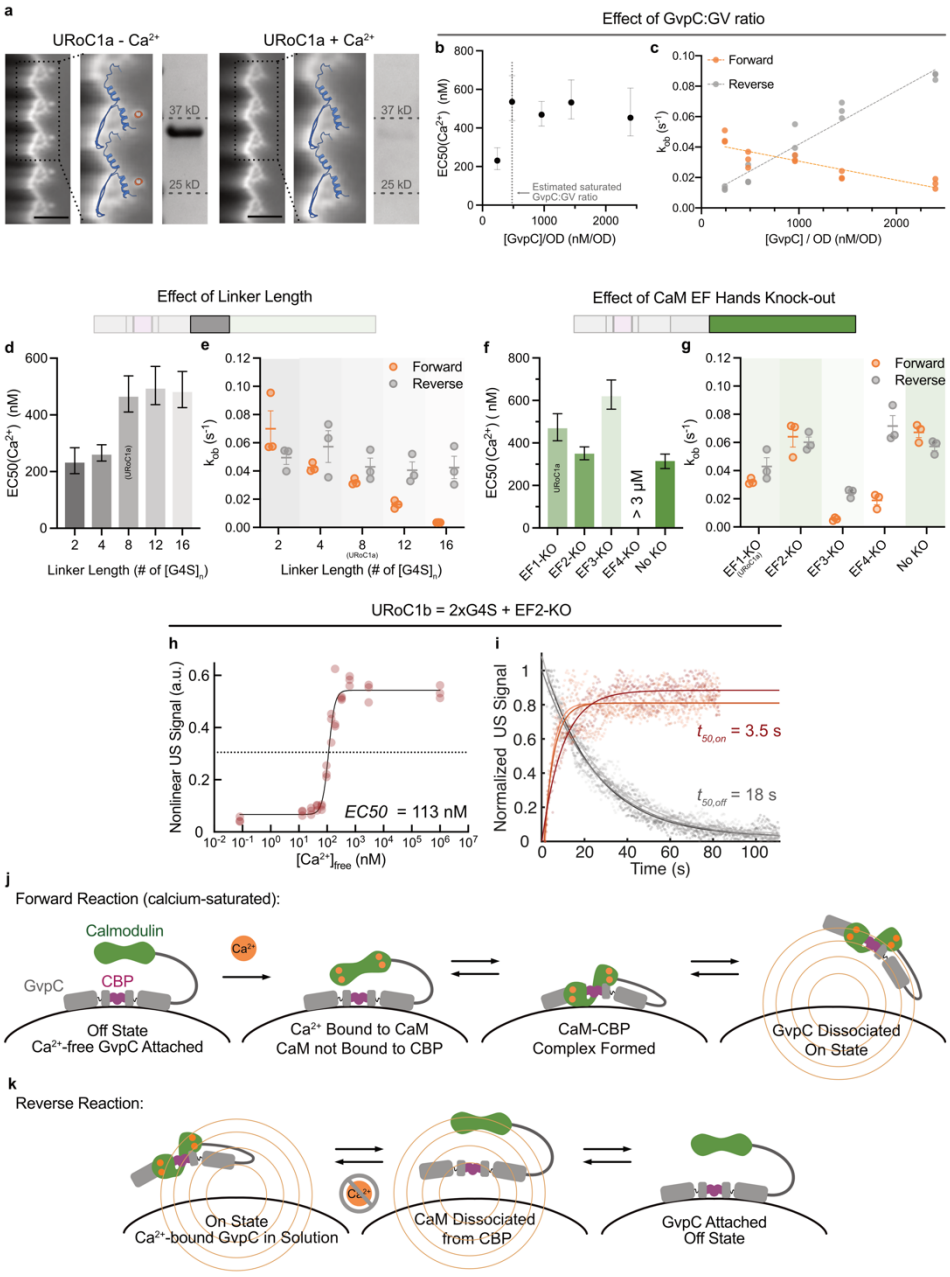


Figure 3-2. Molecular mechanisms and optimization of URoC. (**a**) Left of each column: cryo-EM 2D density map of the side view of the URoC1a GV shell incubated with 200 µM $CaCl_2$ or 5 mM EGTA prior to freezing and an overlaid integrative model of the Ana GvpA:GvpC complex (PDB: 8GBS) . Right of each column: Coomassie-stained SDS-PAGE gel of concentration-matched URoC1a GVs with unbound GvpC removed through buoyancy purification after incubation with calcium or EGTA. The band corresponds to the size of the URoC1a GvpC. (**b**) The EC50 of URoC1a as function of the ratio of GvpC to GV concentration. The dashed line indicates the estimated GvpC concentration needed to saturate all its docking sites on an average GV. (**c**) Observed rate constants of URoC1a as a function of GvpC concentration ratio. The data were fitted to a linear function (dashed lines). (**d-e**) The EC50 (**d**) and observed rate constants (**e**) of URoC variants as a function of the linker length between GvpC and CaM. (**f-g**) The EC50 (**f**) and observed rate constants (**g**) of URoC variants

as a function of individual EF-hand knock-outs of the CaM. The EF4-KO did not reach its half-maximal contrast at the second highest calcium concentration (3 μM) used in our titration. (**h**) Nonlinear signal of URoC1b as a function of free calcium concentration. N = 3 biological replicates. EC50 estimated from fitting a 4-variable Hill equation (black line). Dots represent individual measurements of each replicate. (**i**) Min-max normalized nonlinear ultrasound signal of URoC1b after being mixed with 1 mM $CaCl_2$ or 5 mM EGTA. N = 3 biological replicates, each with 3 technical replicates. Data from each biological replicate were normalized and fitted to an exponential function (solid lines), resulting in the indicated mean half-rise and half-decay times. Dots represent individual measurements of each replicate. (**j**, **k**) Schematics of the proposed molecular mechanisms of the forward (**j**) and reverse (**k**) sensor transitions. Ultrasound data was acquired at 547 kPa for (**b**, **d**, **f**, **h**) and at 472 kPa for (**c**, **e**, **g**, **i**). Scale bars = 5 nm for **a**. Dots represent fitted rate constants from each of the biological replicates for **c**, **e**, and **g**. Error bars represent 95% confidence interval (CI) estimated from the fitting for **b**, **d**, **f**, and standard error of the mean (SEM) for **e** and **g**.

3.2c Ultrasound Imaging of Calcium Dynamics in Mammalian Cells

After demonstrating the performance of URoCs *in vitro* using purified proteins, we endeavored to demonstrate their functionality inside mammalian cells. Recently, genetic constructs were developed for robust mammalian expression of Ana GVs, enabling ultrasound imaging of gene expression[40]. To express URoCs in mammalian cells as intracellular sensors, we generated Ana GV constructs co-expressing the URoC GvpC along with the other genes encoding Ana GVs. These constructs include one polycistronic plasmid encoding the assembly factor genes *gvpNJKFGWV* linked through P2A self-cleaving peptides and a second plasmid encoding the main structural protein GvpA, followed by an internal ribosome entry site (IRES) and the URoC GvpC (**Fig. 3-3a**). This architecture allowed us to maintain a constant GvpA to GvpC ratio while tuning their ratio relative to assembly factors for robust expression[40].

We expressed URoC1b in HEK293T cells, an established cell line for biosensor development[136]. We transiently transfected the cells with the aforementioned two plasmids, induced changes in their intracellular calcium concentration $[Ca^{2+}]_i$ with the ionophore ionomycin[137,138] and imaged them with nonlinear ultrasound. We expected $[Ca^{2+}]_i$ rising to above 1 μM upon incubation with ionomycin [138] and calcium to induce maximal nonlinear signal from the calcium-saturated URoC, and $[Ca^{2+}]_i$ depletion with ionomycin and EGTA to induce minimal signal (**Fig. 3-3b**). We also tested cells incubated in media with physiological calcium concentration without ionomycin to test sensor response to resting state $[Ca^{2+}]_i$ (**Fig. 3-3b**).

As hypothesized, cells incubated with ionomycin and calcium showed higher nonlinear ultrasound contrast across a range of acoustic pressures compared to cells incubated with ionomycin and EGTA, with the peak dynamic range at 326 kPa showing a SBR enhancement of 9 dB, or a fold change of 2.7x in nonlinear signal (**Fig. 3-3c, d, e**). Relative to cells with physiological resting state $[Ca^{2+}]_i$, the ionomycin-calcium incubation provided an 8 dB increase or a 2.3x fold change in SBR (**Fig. 3-3c, d, e**). As additional controls, we tested cells transfected with the mutated calcium-insensitive sensor dURoC1b, which did not produce any significant $[Ca^{2+}]_i$-dependent contrast (**Fig. 3-3c, d, e, Fig. 3-S3a**).

To evaluate the tolerability of URoC1b expression, we measured the viability of cells expressing the biosensors and found no substantial difference compared to cells expressing the fluorescent GECI jRCaMP1b[105] (**Fig. 3-S3b**). Taken together, these results demonstrate the first ultrasound imaging of calcium dynamics in mammalian cells and present URoC1b as a robust tool for cellular calcium imaging.

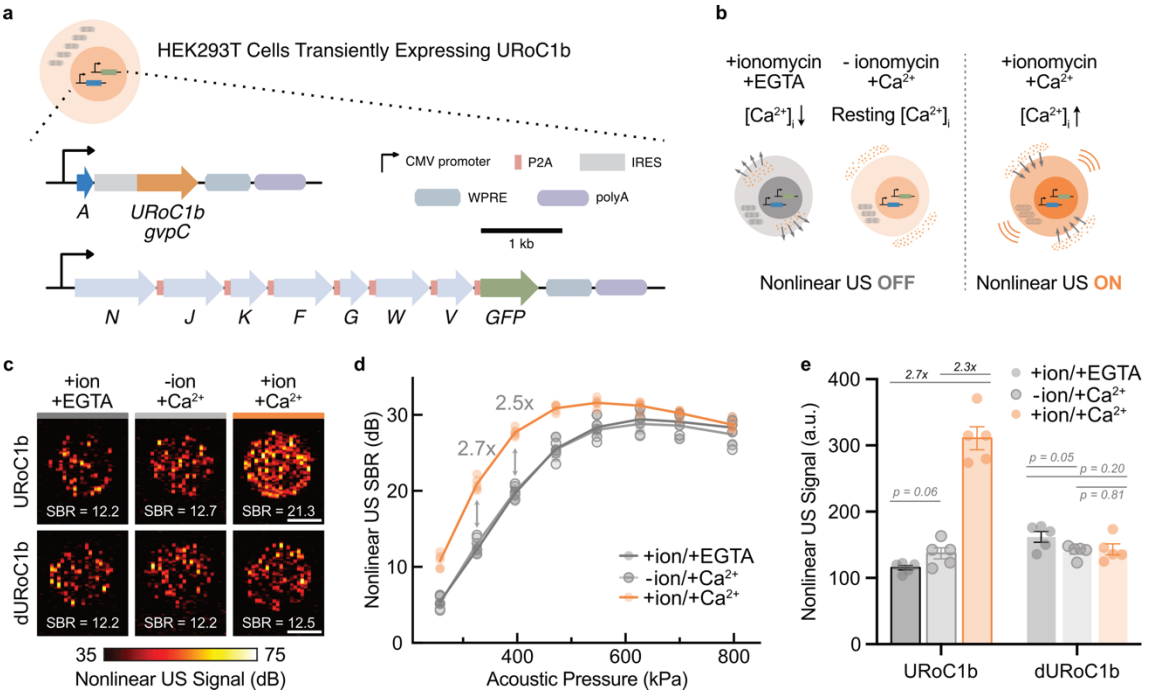


Figure 3-3. Ultrasound imaging of calcium in mammalian cells. (**a**) Schematics of the genetic constructs for transient expression of URoC1b in HEK293T cells. (**b**) Schematics of applying URoC1b to image ionomycin-induced intracellular calcium. (**c**) Representative ultrasound images of agarose phantom containing cells expressing URoC1b or control dURoC1b GVs at 5 million cells per mL with ionomycin and EGTA, ionomycin

and calcium, or calcium only. (**d**) Nonlinear SBR in dB scale as a function of applied acoustic pressure for cells expressing URoC1b after incubation with ionomycin and EGTA, ionomycin and calcium, or calcium only. (**e**) Nonlinear signal of cells expressing URoC1b in phantoms with ionomycin and EGTA, ionomycin and calcium, or calcium only. Ultrasound data were acquired with xAM at 326 kPa for **c**, **e**. Dots represent the mean of two technical replicates for **d**, **e**. N = 5 biological replicates with each N consisting of 2 technical replicates for **c**-**e**. Solid curves represent the mean of all replicates. The bars indicate the mean of all replicates. Scale bars = 1 mm.

3.2d Ultrasound Imaging of Drug-induced Calcium Signaling *in vivo*

After successfully establishing the basic design principles of URoC and characterizing their performance *in vitro* and inside the cells, we assessed its capacity for imaging physiologically relevant calcium dynamics *in vivo*. As a model system, we chose GPCR-driven calcium signaling, which plays critical functional roles in a large variety of cell types[139–142]. In particular, G_q-coupled GPCRs elevate $[Ca^{2+}]_i$ through inositol triphosphate (IP_3) signaling to the endoplasmic reticulum (ER), leading to the release of ER calcium into the cytoplasm[142,143]. To demonstrate *in vivo* URoC performance in a tightly controlled model system, we generated a stable HEK293T cell line co-expressing URoC1b with the hM3D(G_q) DREADD – a muscarinic receptor engineered to respond to bio-orthogonal ligands such as descloroclozapine (DCZ)[144,145] (**Fig. 3-4a**). We chose HEK293T cells for this purpose due to their relatively low endogenous GPCR signaling[146] and previous use as *in vivo* cell-based biosensors[146–148]. We hypothesized that by implanting these cells in tissue such as in the brain, we would be able to follow the calcium response of these cells to a systemically administered ligand such as DCZ, thereby monitoring its pharmacodynamics (**Fig. 3-4a**). In addition to providing a well-controlled proof of concept for *in vivo* URoC functionality, such monitoring has intrinsic utility for drug and receptor engineering.

We constructed a stable HEK293T cell line expressing the hM3D(G_q) receptor through lentiviral infection, then used the PiggyBac transposase system[149] to add doxycycline-inducible[150] genes encoding URoC1b (**Fig. 3-4a**), resulting in DRUM (Drug-Receptor Ultrasound Monitoring) cells. In parallel, we used the non-calcium sensing mutant dURoC1b to produce negative control $DRUM_{mut}$ cells. After generating these cell lines, we implanted DRUM cells unilaterally into the mouse thalamus, with $DRUM_{mut}$ cells implanted contralaterally. (**Fig. 3-4b, c**). Seven days after implantation and 3 days after starting daily doxycycline induction, we performed nonlinear ultrasound imaging through an acoustically

transparent polymer cranial window[151,152] (**Fig. 3-4b**). We located the DRUM and $DRUM_{mut}$ cells through their baseline nonlinear contrast, imaged their baseline signal for 5 minutes, then administered intraperitoneal DCZ (100 μg/kg) and continued to acquire ultrasound images over 20 minutes. At the end of the imaging session, we acquired a Doppler vasculature map as an anatomic reference.

The nonlinear signal of the DRUM cells started to increase shortly after systemic DCZ administration and reached a plateau after several minutes (**Fig. 3-4d, Video 3-S1**). Our images easily spatially resolved the DRUM and contralateral $DRUM_{mut}$ cells. Within the DRUM implant, we observed region-specific kinetics, where certain locations showed a decreasing signal trend after reaching the peak while others remained steady after reaching a plateau (**Fig. 3-4d, Fig. 3-S4a-c**), potentially due to differences in vascularization, receptor expression, downstream signaling, and desensitization[153]. In addition, there were parts of the DRUM cell implant that did not show any response to the stimulation, potentially due to inhomogeneity in GPCR expression or vascular accessibility.

Immunofluorescence imaging of brain tissue collected after the imaging session showed good alignment of ultrasound signal and fluorescence from the implanted cells (**Fig. 3-S5a-b**), confirming expression of both the URoCs and the hM3D(G_q), while also revealing variability in hM3D(G_q) expression and some necrotic areas around in which the ligand might not be easily accessible (**Fig. 3-S5b**), consistent with our ultrasound imaging results.

Across 7 mice, DRUM cells demonstrated a significant increase of approximately 25% in nonlinear ultrasound signal (**Fig. 3-4e, f, Fig. 3-S6a**) upon the administration of DCZ, compared to an average baseline fluctuation of 1% (**Fig. 3-S6b**). Our sensors detected a characteristic onset time of around 128 seconds (**Fig. 3-4e, Fig. 3-S6a**), similar to the 3 minutes previously estimated with invasive measurements[145]. The $DRUM_{mut}$ cell implants did not show a substantial response to DCZ, and both the DRUM and $DRUM_{mut}$ cell implants showed no significant response in control saline injections (**Fig. 3-4g, Fig. 3-S6c**). With a similar DCZ stimulation experiment, we also tested the performance of DRUM implants transcranially and showed in 5 mice that we could achieve a 22% peak

enhancement with similar kinetics through the intact skull without any additional surgeries (**Fig. 3-S6d-e**). These results demonstrate the ability of URoC1b to noninvasively visualize the spatiotemporal dynamics of GPCR-driven calcium signaling across a large area in the mouse brain, thus validating the basic *in vivo* ultrasound imaging capabilities of this technology.

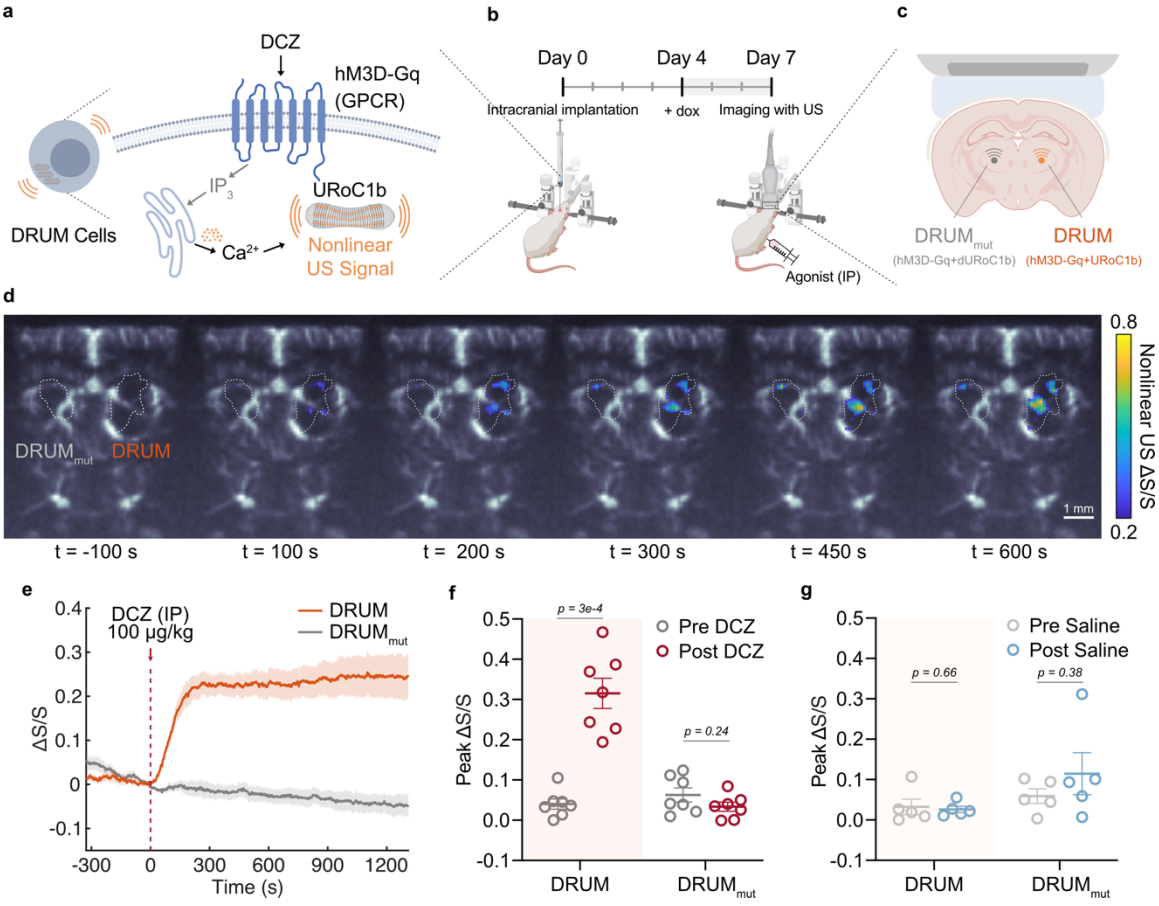


Figure 3-4. Ultrasound imaging of ligand-induced GPCR-driven calcium dynamics in vivo. (**a**) Schematics of engineered cells stably expressing hM3D(G_q) and URoC1b as DRUM cells. (**b**) Schematics of timeline for intracranial implantation, induction for URoC1b expression, and ultrasound imaging. (**c**) Schematics of the implant layout of DRUM and $DRUM_{mut}$ cells in the mouse brain. (**d**) Representative nonlinear US images of DRUM and $DRUM_{mut}$ implants in the brain of mouse 4N. The relative signal change at t = -100/100/200/300/450/600 s compared to t = 0 s was overlaid on a Doppler image. The white dashed lines represent the contour of the regions that showed baseline nonlinear signal above the noise threshold (see Methods) at t = 0. Scale bar = 1 mm. Color bars represent relative change of ultrasound nonlinear signal compared to t = 0. The intensity of the Doppler image represents the blood volume that circulated in each voxel within an integration time of 400 ms (see Methods). (**e**) Time traces of nonlinear ultrasound contrast of the implanted cells. DCZ was injected intraperitoneally (I.P.) at t = 0 s at the dashed line. The nonlinear ultrasound signal was acquired with xAM at 426 kPa (calibrated in water). (**f, g**) Peak relative signal change of DRUM and

$DRUM_{mut}$ implants before and after DCZ (**f**) or saline (**g**) stimulation. For **e**, dark curve = mean and shaded region = SEM. For (**f**, **g**), lines = mean and error bars = SEM. N = 7 mice for the DCZ stimulation (**e, f**) and N = 5 mice for the saline control (**g**). Imaging was conducted through polymer cranial windows for (**d-g**).

3.3 Discussion and Outlook

Our results establish a paradigm for visualizing intracellular calcium dynamics with ultrasound. This paradigm is enabled by engineering the GV-stiffening protein GvpC to undergo a calcium-dependent, allosteric conformational change, leading to increased nonlinear ultrasound signal. To materialize this concept, we engineered and optimized a series of URoCs, resulting in a biosensor with large dynamic range, high sensitivity, full reversibility, and genetically encoded functionality inside mammalian cells.

The *in vivo* application demonstrated in this work provides a fundamental proof of concept for real-time ultrasound calcium imaging inside a living animal. As an immediate application, the DRUM approach could be used to map the region-specific pharmacodynamics of various ligands acting on GPCRs and correlate it with behavioral responses or neuroimaging (e.g., with functional ultrasound[152,154,155]). Similar cellular constructs could be developed to image the action of a large variety of pharmaceutical agents or drugs of abuse.

Additional work is needed to adapt URoCs to a wider range of calcium imaging applications. With parallel, ongoing efforts in GV gene cluster engineering and gene delivery methods, it should in the foreseeable future become possible to express URoCs in primary cells such as neurons *in vivo*. This would enable brain-wide calcium imaging in genetically defined neuronal populations in intact animals, providing an unprecedented view of brain circuit function. Likewise, expression of URoCs in pancreatic beta cells or immune cells would enable the *in vivo* imaging of critical aspects of endocrine and immune activity in basic biology and facilitate the development and monitoring of cell-based therapies. With these and other potential applications, URoCs could have a transformative impact on diverse areas of biology and medicine.

Following the initial invention of fluorescent calcium biosensors, GECIs were optimized over multiple generations to achieve their current outstanding performance, and we envision

a similar development path for URoCs. Although the first-generation biosensors presented in this study have sufficient performance for many envisioned applications[103], there is ample room for improvement. Inspired by the approaches taken to improve GECIs, future work could optimize URoCs through rational engineering and directed evolution of components such as the CBP-CaM pair, fusion positions and linkers, and the underlying GV proteins GvpC and GvpA. For example, based on our mechanistic understanding, engineering the calcium-bound GvpC to remain partly anchored to the GV is expected to improve kinetics. Further improvements could arise from engineering the stoichiometry and identity of GvpC, GvpA and other GV genes[156], leading to URoC GVs with varying size, shape or mechanical properties. In parallel with molecular engineering, improvements in ultrasound instruments and pulse sequences, such as 3D volumetric imaging[155], wearable long-term probes[157,158], advanced nonlinear imaging and super-resolution imaging[159] will provide new and improved ways to visualize URoCs *in vivo*. This will mirror the co-evolution of GECIs with advances in optical microscopy[160–162].

Going beyond calcium imaging, we anticipate that URoCs will serve as a template for the development of other acoustic biosensors based on allosteric conformational changes in GvpC. This would make it possible to sense a larger variety of ions, enzymes, and other biological signals – helping ultrasound rock and roll toward new breakthroughs in biology and medicine.

3.4 Material and Methods

Design and cloning of genetic constructs

All the plasmids in this study were constructed through a combination of polymerase chain reaction (PCR) using Q5 polymerase and Gibson assembly[163] or KLD mutagenesis. All the reagents were from New England Biolabs (NEB) and custom primers were from Integrated DNA Technologies (IDT). For screening and characterization of purified proteins, all *gvpC* gene sequences were codon-optimized for *E. coli* expression and driven by a T7 promoter and lac operator in a pET28a expression vector (Novagen). Specifically, the construct encoding the 3-repeat WT Ana *gvpC* gene was first generated through deletion of the fourth and fifth repeats from an existing plasmid encoding the 5-repeat WT Ana *gvpC* gene with a

C-terminus 6xHis tag through KLD mutagenesis. Second, the codon-optimized CaM sequence from GCaMP6f (synthesized by IDT) and an 8xG4S linker sequence were introduced through Gibson assembly. Last, the sequence of CBP from CaMKI was introduced by substitution/insertion into the second repeat of the *gvpC* gene sequence through KLD mutagenesis to generate the URoC *gvpC* genes. Additional modifications such as changing the linker length or CaM mutations were also achieved through KLD mutagenesis.

For the transient expression of URoC in mammalian cells, all sequences were codon-optimized for expression in human cells and driven by a CMV promoter in a pCMVSport vector with WPRE-hGH polyA. Similar steps were taken to modify the WT *gvpC* sequence to the URoC1b/control *gvpC*, which were then introduced along with an IRES sequence into an existing plasmid encoding the *gvpA* gene (Addgene #197588) through Gibson assembly. An existing plasmid was used for the expression of all 7 chaperones (Addgene #197589). For generating the DRUM or $DRUM_{mut}$ stable cell lines, the lentiviral transfer plasmid constitutively expressing hM3D(G_q)-mCherry was constructed as follows: the sequence encoding hM3D(Gq)-mCherry was amplified by PCR from a plasmid from Addgene (#50474) and assembled into a lentiviral backbone with a human EF1α promoter, an IRES followed by a puromycin resistance gene (*PuroR*) for selection and a WPRE element. The PiggyBac transposon plasmids for URoC expression were constructed by PCR-amplifying the region between the start codon of *gvpNJKFGWV* or *gvpA*-IRES-*gvpC* and the end of the hGH polyA from the transient plasmids. The amplified regions were assembled into the PiggyBac transposon backbone (System Biosciences) between a TRE3G promoter (Takara Bio) for doxycycline-inducible expression and a constitutive EF1α core promoter driving a blasticidin resistance gene *BSD* or hygromycin resistance gene *HygR* gene for selection. The complete list and source of plasmids used in this study are given in **Table 3-S1** and amino acid sequences of URoC variants are given in **Table 3-S2**. All the plasmids were cloned using NEB Turbo *E. coli* (New England Biolabs) and sequence verified.

Preparation of purified GVs for *in vitro* assays

For the *in vitro* purified protein assays, GVs were collected and purified from confluent Ana cultures using previously published protocols[37,64]. The concentration of Ana GVs was determined by measurement of their optical density (OD) at 500 nm (OD500) using a Nanodrop spectrophotometer (Thermo Fisher Scientific), using the resuspension buffer as the blank. As established in previous work[164], the concentration of GVs at OD500 = 1 is approximately 184 pM and the gas fraction is 0.0417%. The engineered GvpC was expressed and purified following a previously published protocol with minor modifications. Briefly, for the expression and purification of URoC GvpC variants, plasmids were transformed into chemically competent BL21(DE3) cells (Invitrogen) and grown overnight for 12–16 h at 37°C in 5 ml starter cultures in 2xYT medium with 50 μg/mL of kanamycin. Starter cultures were inoculated 1:100 into auto-induction Terrific Broth (Novagen 71491) with 50 μg/mL of kanamycin and allowed to grow at 30°C (250 r.p.m. shaking) for 20-24 hours for protein expression. Cells were then collected by centrifugation at 5,500 RCF and lysed at room temperature using SoluLyse (Amsbio L200125), Supplemented with protease inhibitor cocktail (10 μL per 1 mL of culture, Sigma P8849), lysozyme (400 μg/mL) and DNase I (10 μg/mL). GvpC inclusion bodies were isolated by centrifugation at 15,000 RCF for 10 min and then resuspended in a solubilization buffer comprising 20 mM of Tris-HCl buffer with 500 mM of NaCl and 6 M of urea (pH: 8.0), before his-tag purification with Ni-NTA (QIAGEN) and wash and elution buffers of the same composition as the solubilization buffer, but with 20 mM and 250 mM imidazole, respectively. The concentration of the purified protein was assayed using the Bradford Reagent (Bio-rad).

Engineered GVs with the URoC GvpC or wild-type GvpC were prepared using published protocols with urea stripping and GvpC re-addition[37,64]. Briefly, Ana GVs were stripped of their native GvpC through treatment with a 6 M urea solution buffered with 100 mM Tris-HCl (pH:8-8.5), followed by two rounds of centrifugally assisted floatation with removal of the subnatant liquid after each round. The recombinant engineered GvpC variants purified from inclusion bodies were then mixed with the stripped Ana GVs in 6 M urea with varying molar excess concentration determined after accounting for 1:15 binding ratio of 3-repeat

GvpC:GvpA. For an *n*-fold (2-fold unless specified) stoichiometric excess of GvpC relative to binding sites on an average Ana GV, the concentration of recombinant GvpC (in nmol) to be added to stripped GVs was calculated according to the formula: n * OD_{500nm} * 480 nM. The mixture of stripped GVs (OD_{500nm} = 4) and recombinant GvpC in 6 M urea buffer was loaded into dialysis pouches made of regenerated cellulose membrane with a 6-8 kDa M.W. cutoff (Spectrum Labs). The GvpC was allowed to slowly refold onto the surface of the stripped GVs by dialysis in 4 L of calcium-free MOPS buffer (30 mM MOPS, 100 mM KCl**,** pH 7.2) with 1 mM EDTA for the first round, followed by a second round of dialysis in 4 L of MOPS buffer with 30 μM EGTA. Both rounds were done at 4°C for 8-12 hours each. The dialyzed GV samples were directly used for further characterization.

In vitro assay for calcium response

All data were acquired at 37˚C unless specified. For the steady-state calcium response experiments including pH and magnesium sensitivity, engineered GVs at OD_{500} = 3.6 were mixed 1:1 with 1% w/v agarose (Lonza, #50070) in MOPS buffer with 400 μM $CaCl_2$ or 10 mM EGTA (final OD_{500} = 1.8) and then loaded into hydrogel phantoms made of 1% agarose in MOPS buffer with 200 μM $CaCl_2$ or 5 mM EGTA, followed by incubation and ultrasound imaging at 37˚C. For the calcium-free and calcium-saturated conditions, GV samples were loaded directly into the phantoms for incubation. For the reversal condition, 1 M $CaCl_2$ solution was first added to the GV samples after dialysis to make a final concentration of 200 μM $CaCl_2$. The samples then were incubated at 37˚C for 10 minutes before being 1:1 mixed with 1% agarose in MOPS buffer with 10 mM EGTA and loaded into a phantom containing 5 mM EGTA. After loading, the phantoms with GV samples were incubated at 37˚C for 10 minutes and then imaged by ultrasound.

For the calcium titration experiments, the zero free calcium buffer (30 mM MOPS, 100 mM KCl, 10 mM EGTA, pH 7.2) and 39 μM free calcium buffer (30 mM MOPS, 100 mM KCl, 10 mM CaEGTA, pH 7.2) were made in lab and serial dilution was performed to generate buffers with different free calcium concentrations. These buffers were calibrated, with and

without 1:1 dilution with MOPS with 30uM EGTA, using fluo-4 (Invitrogen F14200) and a commercial calcium calibration buffer (Invitrogen C3008MP). GV samples at final OD_{500} = 1.8 were loaded into the phantoms made of 1% agarose (w/v, Lonza, #50070) in those buffers and incubated at 37˚C for >10 minutes before ultrasound imaging.

For measurements of kinetics, GV samples at OD_{500} = 3.6 and the MOPS buffer with 2 mM $CaCl_2$ (forward reaction) or 10 mM EGTA (reverse reaction) were loaded into 1 mL syringes that were both controlled by a syringe pump (Kent Scientific). For the reverse reaction, samples were first incubated with 200 μM [Ca^{2+}] for > 30 minutes at 37˚C before the measurement. The samples and buffer were first delivered into a silicone tubing (OD 1/8", ID 1/16") for prewarming at 37˚C for 2 minutes. Next, both solutions were injected into a 1:1 mixing tip (MIXPAC T-mixer, Medmix) at a flow rate of 100 μL/s (post-mixing at 200 μL/s) and then into a 1/12" polyolefin tubing for ultrasound imaging after mixing. All the tubing was immersed in a water bath at 37˚C. The URoC1a samples pre-incubated with 1 mM $CaCl_2$ were mixed with MOPS buffer with 1 mM $CaCl_2$ to estimate the dead time. The practical dead time was estimated to be around 2 seconds, considering the following two effects. Dead volume of the setup (150 uL) generates dead time of 0.75 s; Additionally, it took approximately 1.2 seconds for the flow to fully stop (**Figure. 3-S1g**).

Cryo-EM characterization and image analysis

A sample of Ana GVs with addition of GvpC variants was incubated for 30 min at 37°C in presence of 200 μM $CaCl_2$ or 5 mM EGTA. A 3 μL volume of sample was applied to C-Flat 2/2 - 3C grids (Protochips) that were glow-discharged (Pelco EasiGlow, 10 mA, 1 min). GV samples were frozen using a Mark IV Vitrobot (FEI, now Thermo Fisher Scientific) (37°C, 60% humidity, blot force 3, blot time 4 s). Before applying the GV sample, the grid was incubated in the Vitrobot chamber for 5 minutes to equilibrate to 37°C. Movie stacks were acquired on a 300 kV Titan Krios microscope (Thermo Fisher Scientific) equipped with a K3 6k × 4k direct electron detector (Gatan). Multi-frame images were collected using SerialEM 3.39 software[165]. Approximately 20 movies were acquired for each condition at a pixel size of 1.4 Å (64,000× magnification) with varying defocus from -1.0 to -3.0 μm. To

generate side-view averages of the GV shell, acquired movies were imported into cryoSPARC[166]. Motion and CTF corrected micrographs were used for subsequent data processing. Using template matching, a few thousand particles of the GV shell's edges were picked, extracted, and subjected to iterative 2D classification. For visualization purposes, an integrative model of the Ana GvpA:GvpC (PDB: 8GBS)[44] complex was overlaid on the GV shell density in the 2D class averages.

Denaturing polyacrylamide gel electrophoresis (SDS-PAGE)

GV samples were incubated with 200 μM $CaCl_2$ or 5 mM EGTA for 30 minutes at 37˚C and then centrifuged at 300 RCF for 2-4 hours at 37˚C. The subnatant liquid was aspirated to remove any unbound GvpC molecules, and the floating layer of GVs were resuspended in MOPS buffer with 200 μM $CaCl_2$ or 5 mM EGTA. This step was repeated for 3 times and the samples were concentration-matched at OD_{500nm} = 10 and mixed 1:1 with 2x Laemmli buffer (Bio-Rad), containing SDS and 2-mercaptoethanol. The samples were then boiled at 95°C for 5 minutes and 20 μL of the samples were loaded into a pre-made polyacrylamide gel (Bio-Rad) immersed in 1x Tris-Glycine-SDS Buffer. 10 uL of Precision Plus Protein™ Dual Color Standards (Bio-Rad) was loaded as the ladder. Electrophoresis was performed at 120V for 55 minutes, after which the gel was washed in DI water for 15 minutes to remove excess SDS and coomassie-stained for 1 hour on a rocker using the SimplyBlue SafeStain (Invitrogen). The gel was allowed to de-stain overnight in DI water before imaging using a Bio-Rad ChemiDoc™ imaging system.

Structure prediction of URoC GvpC

The structure of the GvpC segment of URoC1a (1-127) was predicted using AlphaFold2 (ColabFold[167]) and the structure with the highest confidence was used to align with the existing structure of the CaMKI-CaM complex (PDB: 1MXE)[125]. After the alignment, the distance was measured between the second carbon atom of the G127 in the GvpC and the first carbon atom of the L1 of the CaM using ChimeraX[168].

HEK293T cell culture and transient transfection

HEK293T cells (American Type Culture Collection (ATCC), CLR-2316) were seeded in 6-well plates at 2.5×10^5 cells per well at 37 °C and 5% CO_2 in a humidified incubator in 2 ml of DMEM (Corning, 10-013-CV) with 10% FBS (Gibco), 10 mM HEPES (Cytia) and 1× penicillin–streptomycin 24 hours before transfection. Transient transfection mixtures were created by mixing 2 μg of plasmid mixture with polyethyleneimine (PEI-MAX, Polysciences) at 4 μg of polyethyleneimine per microgram of DNA. The mixture was incubated for 12 minutes at room temperature and added drop-wise to HEK293T cells. Media was changed after 12–16 hours and daily thereafter. For the transient expression of URoC1b and dURoC1b, control GVs, 280 fmol of the *gvpA*-IRES-*gvpC* plasmid and 70 fmol of the plasmid encoding *gvpNJKFGWV* was added into the DNA mixture, and pUC19 plasmid DNA was supplemented to make the total amount of DNA 2 μg. For the transient expression of jRCaMP1b, DNA of the same mass as the GV genes (*i.e.*, *gvpA*-IRES-*gvpC* and *gvpNJKFGWV*) was mixed with the pUC19 plasmid DNA to make the total mass 2 μg. Transfected cells were assayed 72 hours after the transfection.

Preparation of cells for ionomycin stimulation

After 3 days of expression, cells were dissociated using Trypsin/EDTA (Corning 25-053-CI), centrifuged at 300 RCF for 6 minutes at room temperature and resuspended in HEPES-buffered Hanks' balanced salt solution without calcium (HHBSS, 20 mM HEPES, 2 g/L D-glucose in 1x HBSS without calcium). The cells were counted using an automated cell counter (Countess™ 3, Thermo Fisher) and all the samples were concentration-matched to 10 million cells per milliliter in HHBSS without calcium. HHBSS phantoms were prepared with 1% agarose (w/v, Lonza, #50070) supplemented with 10 μM ionomycin (Sigma, I9657) and 5 mM EGTA, 10 μM ionomycin and 2 mM $CaCl_2$, or 2 mM $CaCl_2$. Cells were diluted 1:1 with the 1% agarose containing 2x concentration of the corresponding reagents to result in a final concentration of 5 million cells per milliliter before loading into their respective phantoms. The phantoms were incubated at 37˚C for 10 minutes before ultrasound imaging in a water bath with the same buffer content at 37˚C.

In vitro ultrasound imaging

As described above, GV or cell samples were loaded into imaging phantoms made of 1% agarose (w/v) in specified buffers or injected into an acoustically transparent tubing for imaging (**Fig. 3-1j**). All the samples were placed in the same buffer as the phantom during the imaging session and maintained at 37˚C by a custom water bath unless specified.

Imaging was performed using a Verasonics Vantage programmable ultrasound scanning system and a L22-14vX 128-element linear array Verasonics transducer, with a specified pitch of 0.1 mm, an elevation focus of 8 mm, an elevation aperture of 1.5mm and a center frequency of 18.5 MHz with 67% -6 dB bandwidth. For nonlinear image acquisition, a custom cross-amplitude modulation (xAM) sequence detailed in an earlier study [34] was used with an xAM angle (θ) of 19.5°, an aperture of 65 elements, and a transmitting frequency at 15.625 MHz. The center of the sample wells was placed at a depth of 5 mm with a conventional ray-line scanning B-mode pulse sequence with parabolic focusing at 10 mm and an aperture of 40 elements. The focus was set to be far from the sample position to reduce the acoustic pressure to avoid collapsing the samples. The transmitted pressure at the sample position at 5 mm was calibrated using a Precision Acoustics fiber-optic hydrophone system. Each image was an average of 50 accumulations. B-mode images were acquired at a transmit voltage of 1.6V (86 kPa), and an automated voltage ramp imaging script (programmed in MATLAB) was used to conduct xAM acquisitions at different acoustic pressures. For purified GV samples, an xAM voltage ramp sequence from 3V (258 kPa) to 8V (1068 kPa) in 0.5V step increments was used. For HEK293T cells expressing GVs, an xAM voltage ramp sequence from 3 V (258 kPa) to 6.5V (797 kPa) in 0.5V step increments was used. Samples were subjected to complete collapse at 25V with the B-mode sequence for 10 seconds, and the subsequent post-collapse xAM images acquired at the same voltage steps were used as the blank for data processing. For purified GV samples, B-mode images at the beginning of the pressure ramp and at the end of the post-collapse ramp were acquired for concentration normalization.

Viability assay

The viability of the cells transfected with URoC1b and jRCaMP1b was assayed with alamarBlue™ (Invitrogen DAL1025) for resazurin reduction, the luminescent ATP detection assay kit (Abcam, ab113849) for the ATP content and Trypan Blue for the membrane permeability, all following the manufacturers' protocols. In particular, for the reducing power assay, the cells were incubated with the alamarBlue™ reagent for 1 hour at 37˚C and 5% CO_2 before the read-out with a plate reader (Tecan). For the Trypan Blue assay, cells were first dissociated using Trypsin/EDTA (Corning 25-053-CI), and the culture media with 10% FBS was added to the wells at a 1:1 ratio to quench the reaction. Cells were resuspended in the trypsin-media mixture and then proceeded for viability quantification using an automated cell counter (Countess™ 3, Thermo Fisher) after 1:1 dilution with the Trypan Blue.

Construction of DRUM and $DRUM_{mut}$ cell lines

HEK293T cells were cultured as described above and seeded in 6-well plates at 2.5×10^5 cells per well. After 24 hours, cells were lentivirally transduced with jmL-95_pEF1α-hM3D(G_q)-mCherry-IRES-PuroR-WPRE at a multiplicity of infection (MOI) of 4 with 10 μg/mL of polybrene. The media containing the viruses and polybrene was removed 12 hours after the infection and the cells were passaged 36 hours after the infection. Next, the cells were expanded to a surface-treated T-75 flask and treated with 2 μg/mL of puromycin (Invivogen) for 2 weeks for selection of the transduced cells to generate the cell line expressing hM3D(G_q), HEK-jmL95.

After the puromycin selection, HEK-jmL95 was seeded in 2 separate wells in a 6-well plate at 2.5×10^5 cells per well without puromycin. 22 hours after the seeding, the culture media was replaced with DMEM (Corning, 10-013-CV) containing 2% FBS (Gibco), 10 mM HEPES (Cytia) and 1× penicillin–streptomycin. After 2 hours of incubation with reduced-serum media, cells were transfected with 3 μg of PiggyBac transposon:transposase plasmid mixture (2143 ng of URoC plasmids with a molar ratio of 4:1 PB-gvpA-IRES-URoC1b/dURoC1b:PB-gvpNV transposons and 857 ng of PiggyBac transposase) using PEI-MAX with the same protocol for the transient transfection. Media was changed to 10%

FBS culture media after 12-16 hours of incubation and daily thereafter for 3 days. The transfected cells were expanded to T-75 flasks and cultured with 2 μg/mL of puromycin (Invivogen), 10 μg/mL of blasticidin (Invivogen) and 200 μg/mL of hygromycin (Invivogen) for selection of cells transduced with the URoC1b genes. After 2 weeks, the selected cells were expanded in media without puromycin, blasticidin or hygromycin and frozen in Recovery Cell Culture Freezing Medium (Gibco) using Mr. Frosty cell freezing container (Nalgene) filled with isopropanol at −80 ˚C and then stored in liquid nitrogen (LN2) vapor phase until use.

Intracranial implantation of DRUM and $DRUM_{mut}$ cells

Transcranial injection of DRUM and $DRUM_{mut}$ cells was conducted in NSG mice (Jackson Laboratory) aged 6 weeks for those imaged transcranially or 6-9 weeks for others imaged through cranial windows, all under a protocol approved by the Institutional Animal Care and Use Committee of the California Institute of Technology. No randomization or blinding was necessary in this study. DRUM and $DRUM_{mut}$ cells were recovered from LN_2 storage and cultured without drug selection for at least 2 passages before implantation. On the day of implantation, cells were dissociated using Trypsin/EDTA (Corning 25-053-CI), centrifuged at 500 g for 5 minutes, and resuspended in culture media without antibiotics at a concentration of 60 million cells per milliliter. The cells were then stored on ice before injection.

During the surgery, mice were anesthetized with 1–2% isoflurane, weighed before the surgery, maintained at 37 °C on a heating pad, weighed before the surgery, and placed on a stereotaxic instrument. The cells were injected intracranially at the coordinate of Anterior-Posterior (AP) –2 mm, Media-Lateral (ML) ±1.5 mm (–1.5 mm for $DRUM_{mut}$ and +1.5 mm for DRUM cells), and Dorsal-Ventral (DV) –3.5 mm, relative to the bregma. The injection was conducted through a microliter syringe (Hamilton) with a needle of 33G (World Precision Instrument) controlled by a micro syringe pump (World Precision Instrument) at a flow rate of 7 nL per second. A volume of 3300 nL containing 200 thousand cells was injected on each side.

For the mice to be imaged transcranially, the skin was closed after the injection with a tissue adhesive (GLUture). Animals recovered quickly and remained bright, alert, and responsive before the ultrasound imaging.

For the mice to be imaged through cranial windows, the craniotomy was performed after the injection using a micro drill steel burr (Burr number 19007-07, Fine Science Tools) from approximately AP –1 mm to AP –3 mm and from ML –2.5 mm to ML +2.5 mm. A 0.125 mm thick polymethylpentene TPX® film (Sigma) was cut to cover the cranial opening and then attached to the skull through a light-cured composite (Tetric EvoFlow). Dental cement (C&B METABOND) was used to close between the skin and the cranial window.

Ultrasound imaging of GPCR signaling *in vivo*

4 days after the implantation and every 12 hours thereafter, the mice were induced for GV expression through intraperitoneally (I.P.) injection of 150 μL of saline containing 150 μg of doxycycline. The ultrasound imaging was conducted 7 days after the surgery (induced for 3 days). The mice were first anesthetized with 4% isoflurane, and then placed on a stereotaxic instrument with 1.5% isoflurane for maintenance anesthesia, and their core temperature was monitored with a rectal probe and maintained at 37˚C with a heating pad.

The same instruments and pulse sequences with the *in vitro* ultrasound imaging were used, and the ultrasound transducer L22-14vX was held by a custom holder mounted on the right arm of the stereotaxic instrument. For the mice with the cranial windows, the ultrasound coupling gel (Aquasonic) was directly applied to the cranial windows, and the transducer surface was placed ~1 mm above the middle of the cranial window in the coupling gel. The B-mode pulse sequence described in the in vitro imaging methods was applied to fine-tune the initial position of the transducer, and the xAM pulse sequence was applied at 426 kPa (calibrated in water, higher than the optimal pressure in vitro to account for attenuation through the cranial window and tissue) to scan across the cranial window to locate the DRUM and $DRUM_{mut}$ implants via their off-state nonlinear signal. The coronal planes with the highest signal of both implants were chosen for imaging. After locating the imaging plane, the imaging was paused (no ultrasound transmitted or received), with the mice staying

in the imaging apparatus for 10 minutes to ensure full equilibrium of mouse body temperature, breathing, and other potential effects of isoflurane.

After 10 minutes, the imaging session started using xAM with the same parameters as described above — 65 elements aperture, 426 kPa (calibrated in water), and a frame rate of approximately 1.7 frames per second. The baseline signal was acquired for 500 frames, after which 100 μg/kg of DCZ (in the format of 30 μg/mL of deschloroclozapine dihydrochloride in saline) or 3.3 μL/g of saline was injected I.P. The nonlinear signal was recorded for 2000 frames after the injection. 7 mice were stimulated with DCZ, and 5 mice were injected with saline as controls. For 4 mice stimulated with DCZ, a vasculature map for anatomic reference was acquired through a plane wave power Doppler pulse sequence at 6V at the end of the imaging session.

For mice imaged transcranially, similar procedures were followed except for three steps. First, the skin was cut open and the ultrasound gel (Aquasonic) was applied to the skull to couple the transducer to the skull. Second, an amplitude modulation pulse sequence with a parabolic focus at 6 mm and an aperture of 40 elements at a higher peak positive pressure of 1.28 MPa (calibrated in water) was used for imaging to account for the attenuation of the intact skull. Third, the plane wave Doppler imaging was acquired at 25V to account for the intact skull. 5 mice were imaged transcranially for the DCZ stimulation.

Doppler image acquisition

The power Doppler images mapping local changes in cerebral blood volume (CBV) were acquired at 15.625 MHz as previously described[169,170]. Briefly, the pulse sequence contains 15 tilted plane waves varying from -14˚ to 14˚ at a 500 Hz pulse repetition frequency. A block of 200 coherently compounded frames was processed using an SVD clutter filter to separate tissue signal from blood signal (cutoff of 40) to obtain a final power Doppler image exhibiting CBV in the whole imaging plane[171,172].

Collection of brain tissue

After the imaging session, all the animals were perfused with 30 mL of phosphate-buffered saline (PBS), followed by 30 mL of 10% formalin solution. The brains were resected and placed in 10% formalin solution for 36 hours at 4˚C after which they were transferred to PBS for long-term storage at 4˚C. The brain tissue for one of the mice stimulated with DCZ (mouse 3N) was accidently lost during sample transportation.

Histology of implanted cells in mouse brain

A representative brain from a mouse stimulated with DCZ (mouse 4N with the cranial window) was sectioned into 100 μm slices using a vibrating microtome (Leica). The slices were mounted and stained with DAAPI nuclear stain using ProLong Diamond antifade mountant with DAPI (Thermo Fisher P36962) and sealed with acrylic resins. The mounted slices were imaged using a Zeiss LSM 980 confocal microscope with ZEN Blue. Images were processed and exported using the ZEN Blue software.

Image processing and data analysis

All *in vitro* and *in vivo* ultrasound images were processed using MATLAB. Regions of interest (ROIs) were manually defined so as to adequately capture the signals from each sample well or region of the implants in the brain. For the steady state calcium response, the sample ROI dimensions (1.2 mm × 1.2 mm square) were the same for all *in vitro* phantom experiments except the URoC initial screening and calcium titration experiments, which were centered at each well. The background ROI was manually selected from the background for each pair of sample wells. For the URoC initial screening and calcium titration experiments, 96-well plate layout phantoms with larger well size were used, so ROIs with a size of 2 mm (lateral) × 1 mm (axial) were chosen for analysis. For each ROI, the mean pixel intensity I_{ROI} was calculated, and the pressure-sensitive ultrasound intensity ($\Delta I = I_{intact} - I_{collapsed}$) was calculated by subtracting the mean pixel intensity of the collapsed image from the mean pixel intensity of the intact image. The signal-to-background ratio (SBR) in was calculated for each sample well by $SBR = 20 \times log_{10}(\frac{I_{Sample\ ROI}}{I_{Background\ ROI}})$. For purified

protein characterization, due to the correlation between the GV concentration and B-mode signal [120], the pressure-sensitive B-mode intensity ($\Delta I_{B-mode} = I_{Bmode,pre-ramp} - I_{Bmode,collapsed}$) was used to normalize the small variability in concentrations. The background-subtracted concentration-normalized nonlinear signal at a specific voltage (or acoustic pressure) was calculated with the following formula: $\frac{\Delta I_{xAM}(V)}{\Delta I_{B-mode}}$. For the kinetics measurement, the ROIs were chosen to cover the interior of the tubing and raw mean intensity was used for min-max normalization within each biological replicates. The data after normalization were fitted to an exponential equation $Ae^{Bt} + C$ with data from the first 3 frame post-mixing not used for curve fitting to account for the dead time (**Fig. S1G**). The fitted parameter B was used as the observed rate constant.

For the *in vivo* experiment, rectangle ROIs were selected to contour the brain region with the above-background nonlinear signal, while tissue background ROI was selected at the same depth with the implants. For quantification, the mean intensity within the ROIs were used to calculate the relative signal change as $\frac{\Delta S}{S} = \frac{I_{Sample\ ROI}(t) - I_{Sample\ ROI}(t=0)}{I_{Sample\ ROI}(t=0)}$ and a moving average filter of size 5 was applied to smooth the curve. For image display, this calculation was done on individual pixels to generate relative change images. Then, a median filter with a [1,4] (lateral, axial) neighborhood and a Gaussian filter with a standard deviation of 1 were applied to reduce the noise of the images. The images were overlaid on top of the Doppler images for display. All images were pseudo-colored (bone colormap for Doppler images, hot colormap for xAM images and parula for the relative signal changes), with the maximum and minimum levels indicated in the accompanying color bars. For the contour showing the location of implants, the threshold was calculated as the mean of the background ROI plus twice its standard deviation based on the image at t = 0 s.

Statistical analysis

Error bars indicate the mean standard error of the mean (SEM) unless otherwise specified. The sample size is N=3 biological replicates in all *in vitro* experiments unless otherwise stated. For each biological replicate, there were technical replicates to accommodate for

variability in experimental procedures such as sample loading and pipetting. SEM was calculated by taking the values for the biological replicates, each of which was the mean of its technical replicates. The numbers of biological and technical replicates were chosen based on preliminary experiments such that they would be sufficient to report significant differences in mean values. P values were calculated using a two-tailed paired t-test.

3.5 Supplementary Information

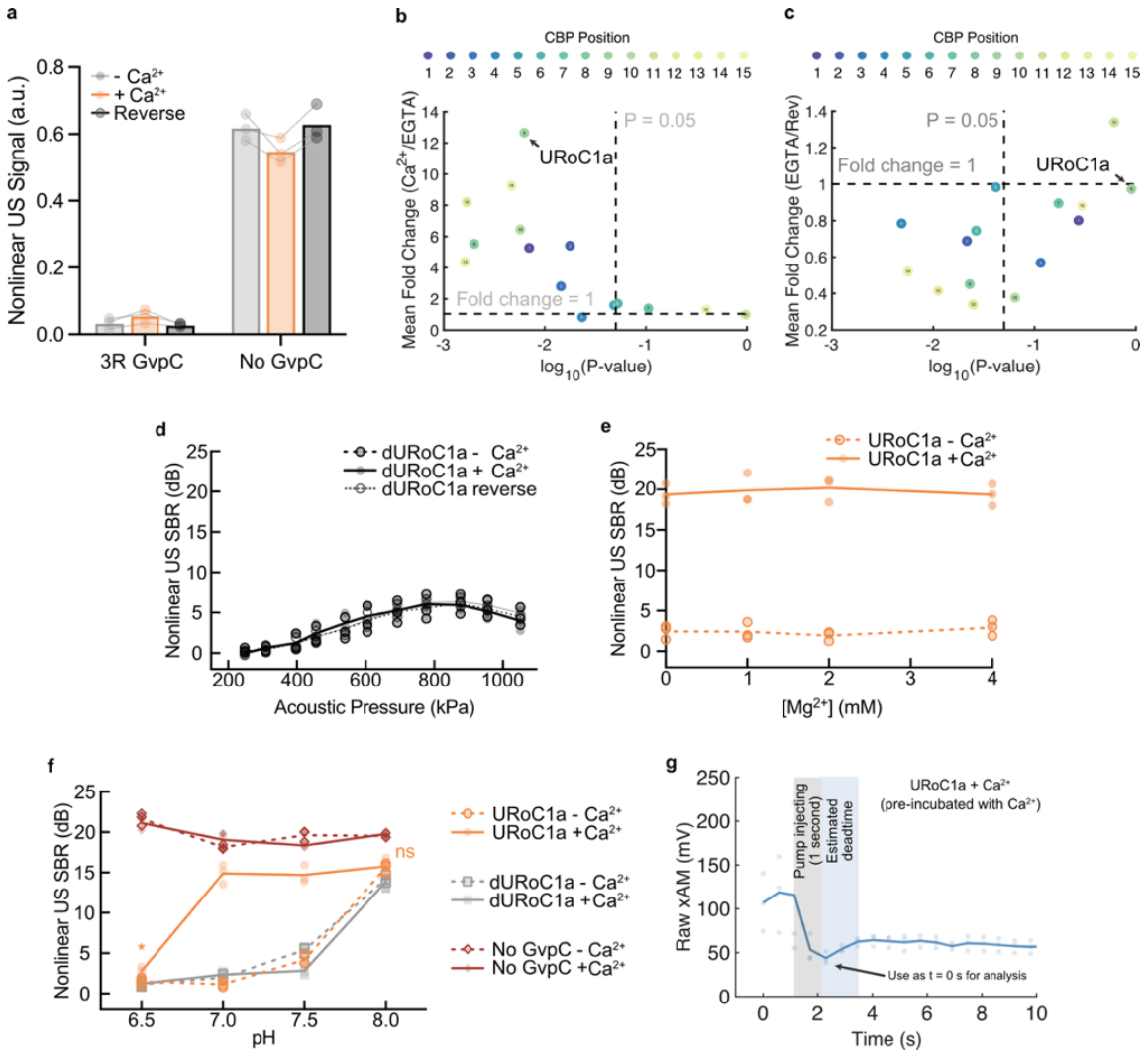


Figure 3-S1. URoC screening and URoC1a characterization. (**a**) Nonlinear signal after background subtraction and GV concentration normalization of GVs with 3-repeat Ana GvpC (3R GvpC) or no GvpC after incubation with EGTA, calcium or first with calcium and then with EGTA. (**b**, **c**) The nonlinear signal mean fold change of URoC variants with different CBP insertion sites as a function the p-value, comparing conditions with calcium and with EGTA (**b**) or first with calcium and then reversed with EGTA and only with EGTA (**c**). The vertical dash lines represent the significance threshold of p = 0.05 and the horizontal lines represent no calcium-dependent change (**b**) or fully reversible (**c**). (**d**) Nonlinear SBR in dB scale as a function of applied

acoustic pressure for control dURoC1a GVs (with all EF hands knocked out) after incubation with 200 μM $CaCl_2$, 5 mM EGTA or first with 200 μM $CaCl_2$ and then with 5 mM EGTA. Solid curves represent the mean of all biological replicates. (**e**) Nonlinear SBR in dB scale as a function of magnesium concentration for URoC1a after incubation with 1 mM $CaCl_2$ (solid lines) or 5 mM EGTA (dash lines). Curves represent the mean of all biological replicates. (**f**) Nonlinear SBR in dB scale as a function of pH for URoC1a, control GVs or GVs without GvpC after incubation with 200 μM $CaCl_2$ (solid lines) or 5 mM EGTA (dash lines). (**g**) Nonlinear ultrasound signal of URoC1a preincubated with 200 μM $CaCl_2$ as a function of time after being mixed into 200 μM $CaCl_2$. The time needed for the signal to become stable was used to estimate the dead time for the kinetics measurement. Ultrasound data were acquired with xAM at 547 kPa for (**a-c**, **e-f**) and at 472 kPa for (**g**). Curves represent the mean of all biological replicates. N = 3 biological replicates for all panels and each biological replicate has 2 technical replicates for **d, e, f**. Dots represent individual measurement of each replicate for **a** and **g**, or the mean of two technical replicates for **d, e, f**.

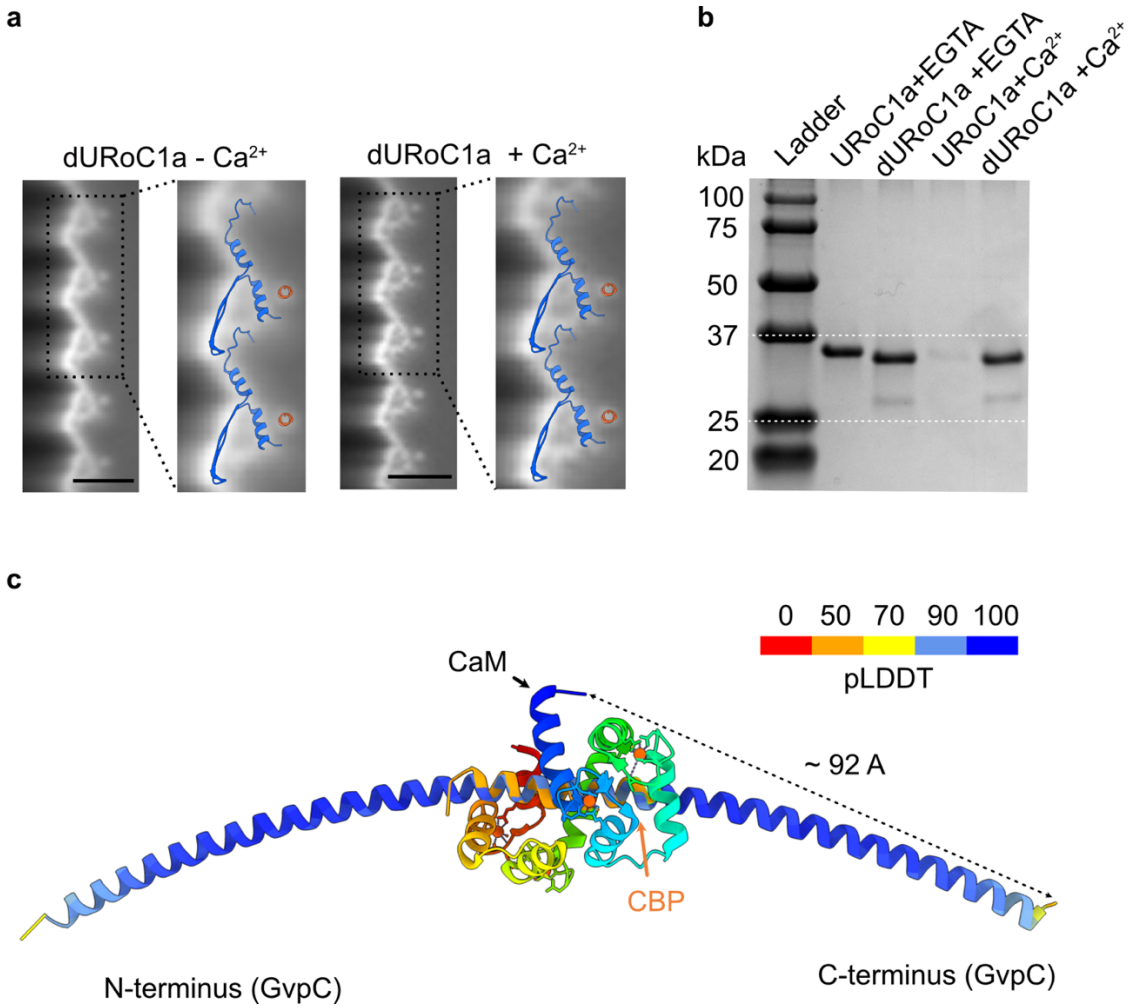


Figure 3-S2. Molecular mechanism of URoC. (**a**) Cryo-EM 2D density map of the side view of the dURoC1a GV shell incubated with 200 μM $CaCl_2$ or 5 mM EGTA prior to freezing and an integrative model of the Ana GvpA:GvpC (PDB: 8GBS)[44] complex was overlaid on the GV shell density in the 2D class averages. (**b**) Coomassie-stained SDS-PAGE gel of OD_{500nm}-matched URoC1a GVs and dURoC1a with unbound GvpC molecules removed through buoyancy purification after incubation with calcium or EGTA at 37˚C. (**c**) Structure prediction of calcium-saturated CaM bound to the URoC GvpC without the linker. The GvpC was colored by the pLDDT confidence from AlphaFold prediction. The CaM was colored in rainbow from blue to red (N-terminus to C-terminus) and the CBP in the CaM-CBP complex was colored orange. The dash line indicates

the distance between the center carbon of the C-terminal glycine of GvpC and the that of the N-terminal aspartic acid of CaM.

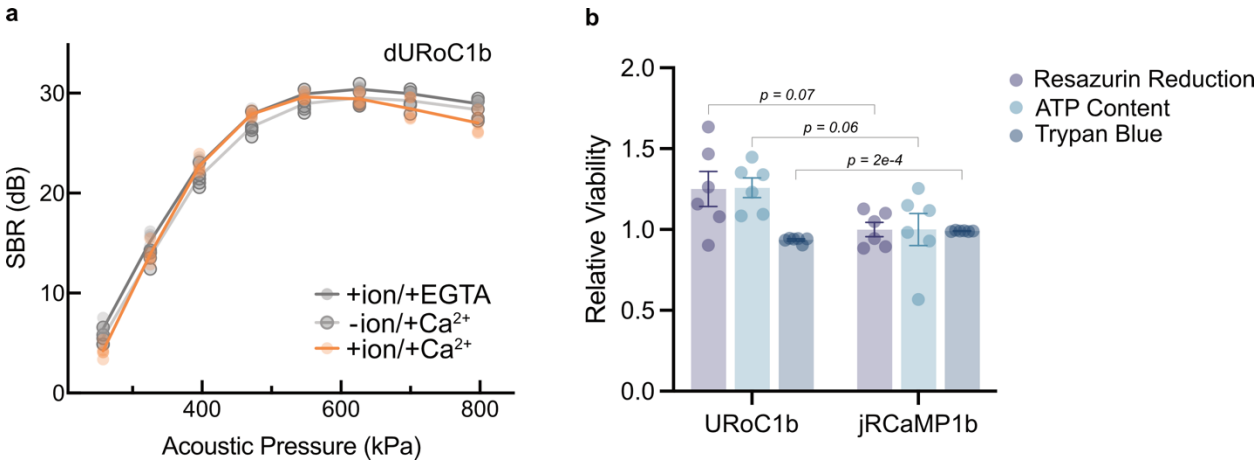


Figure 3-S3. Ultrasound imaging of control non-calcium-sensing GVs in mammalian cells and viability assay. (**a**) Nonlinear SBR in dB scale as a function of applied acoustic pressure for cells expressing control dURoC1b after incubation with ionomycin and EGTA, ionomycin and calcium or calcium only. (**b**) Viability assays of cells transiently expressing URoC1b or jRCaMP1b. The data were normalized to the mean of each measurement done with jRCaMp1b. The difference in viability measured by Trypan Blue is statistically different but both are above 93% (93.5% for URoC1b and 99% for jRCaMP1b).

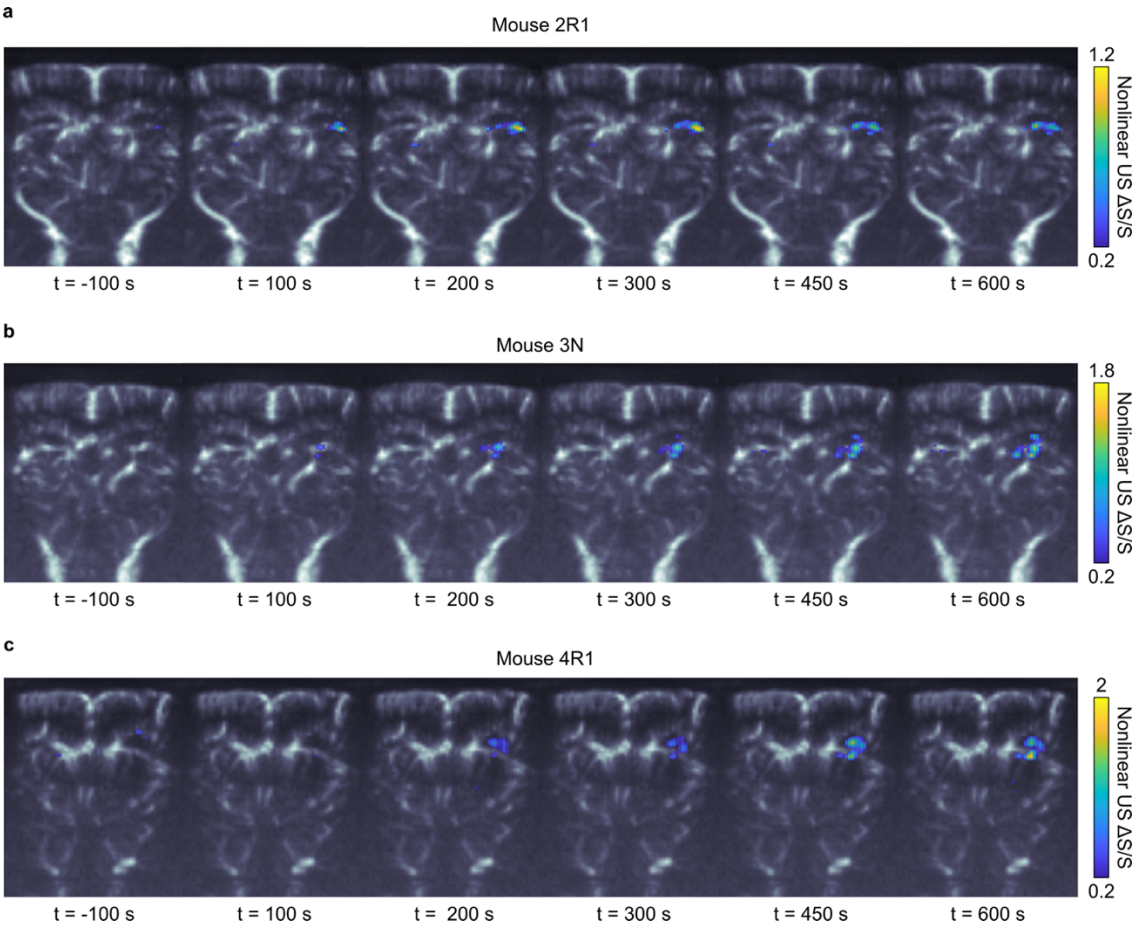


Figure 3-S4. Ultrasound imaging of ligand-induced GPCR-driven calcium dynamics *in vivo*. (**a-c**) Representative nonlinear US images of DRUM and DRUM$_{mut}$ implants in the brain of mouse 2R1 (**a**), 3N (**b**) and 4R1 (**c**). The relative signal change at t = -100/100/200/300/450/600 s compared to t = 0 s was overlaid on a Doppler image. The Scale bar = 1 mm. Color bars represent relative change of ultrasound nonlinear signal compared to t = 0. The intensity of the Doppler image represents the blood volume that circulated in each voxel within an integration time of 400 ms (see Methods). Imaging was conducted through polymer cranial windows.

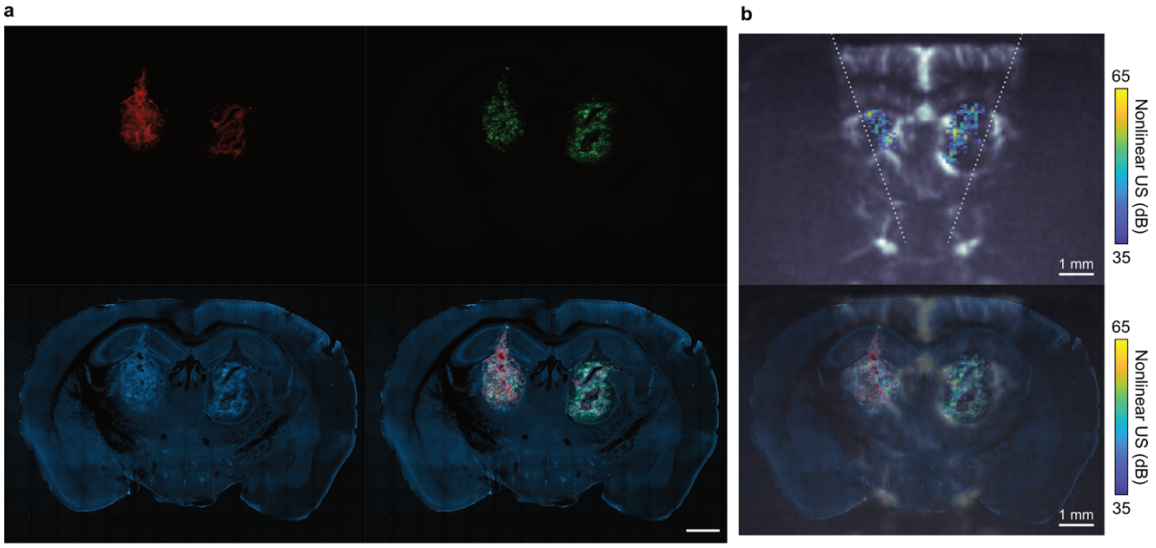


Figure 3-S5. Immunofluorescence characterization of DRUM and DRUM$_{mut}$ implants *in vivo*. (**a**) Representative immunofluorescence micrograph of a 100-μm-thin brain section. Red color shows mCherry fluorescence from the hM3D(Gq) receptor (direct fusion); green color shows GFP fluorescence from the *gvpNV* genes (P2A chained); blue shows DAPI nuclear stain. (**b**) Top: the nonlinear ultrasound image of the baseline signals from the implants overlaid on a Doppler image for the same mouse before DCZ injection (t = 0 s). The dashed line indicates the effective field of view of xAM due to the side aperture being partially blocked by the edge of the skull/cranial window. Bottom: The ultrasound image overlaid with the fluorescence micrograph. All the scale bars = 1 mm. Color bars represent ultrasound signal intensity in the dB scale. The nonlinear ultrasound signal was acquired with xAM at 426 kPa (calibrated in water) and through polymer cranial windows.

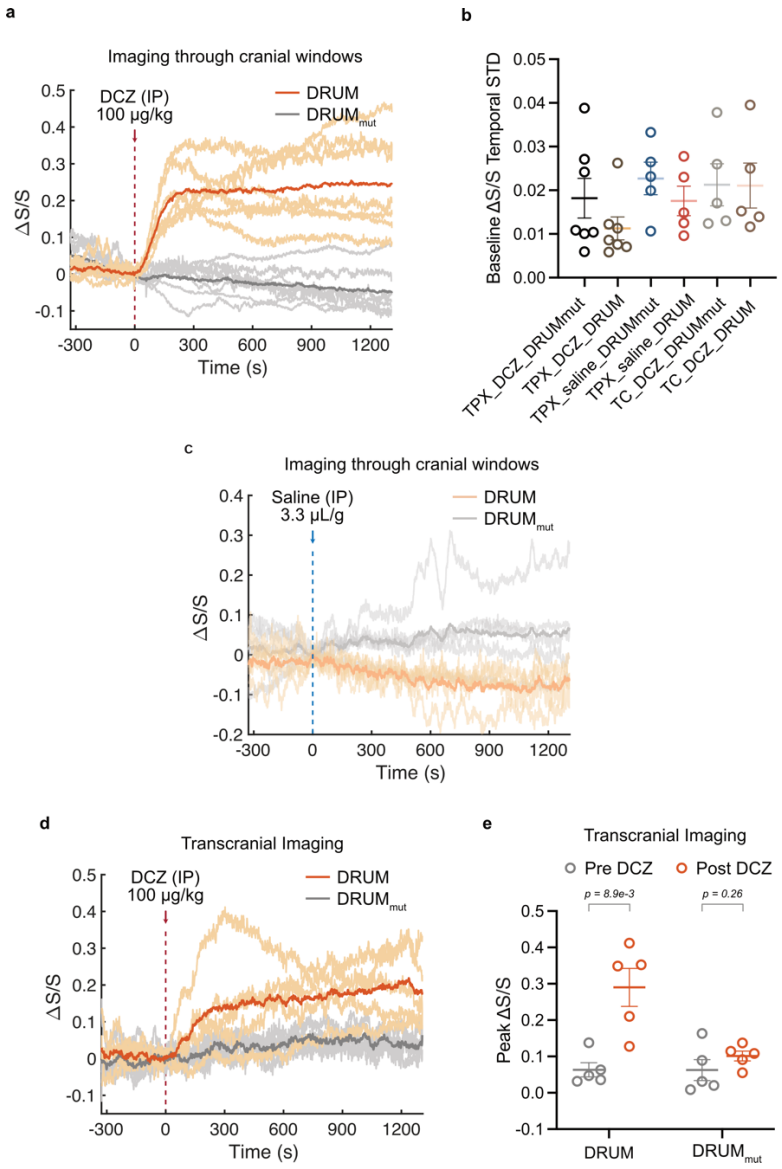


Figure 3-S6. Quantification of the performance of DRUM and DRUM$_{mut}$ implants *in vivo*. (**a**) Time traces of nonlinear ultrasound contrast of the implanted cells through cranial windows in individual animals. DCZ was injected I.P. at t = 0 s at the dash line. Dark curve represents the mean of all traces and light curves represent data from individual animals. (**b**) Temporal standard deviation of the baseline relative change of nonlinear signal of DRUM or DRUM$_{mut}$ implants from t = -327 s to t = 0 s. N = 7 for imaging through TPX cranial window and stimulated with DCZ, N = 5 for the experiment with TPX cranial window and saline stimulation, and transcranially (TC) with DCZ stimulation. Dots = individual trial, lines represent mean and error bars represent SEM. (**c**) Time traces of nonlinear ultrasound contrast of the implanted cells through the cranial windows in response to saline. Saline was injected I.P. at t = 0 s at the dash line. Dark curve represents the mean of all traces and light curves represent data from individual animals. (**d**) Time traces of nonlinear ultrasound contrast of implanted cells through intact skull. DCZ was injected I.P. at t = 0 s at the dash line. Dark curve represents the mean of all traces and light curves represent data from individual animals. (**e**) Peak relative signal change of DRUM and DRUM$_{mut}$ implants before and after DCZ stimulation when imaging through the intact skulls. Lines = mean and error bars = SEM. N = 7 mice for (**a**) and N = 5 mice for (**c-e**).

Plasmid Name	Purpose	Reference Information
jbC-76_pET28a_GvpC-r45d	Bacterial expression of wild type 3-repeat Ana GvpC with 6xHis tag on the C-terminus for purification in pET28a backbone	Fig. 1d; Extended Data Fig. 1a (available on Addgene upon publication).
jbC-78_pET-28a_URoC0-p1	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #1 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-79_pET-28a_URoC0-p2	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #2 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c
jbC-80_pET-28a_URoC0-p3	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #3 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-81_pET-28a_URoC0-p4	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #4 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-8_pET-28a_URoC0-p5	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #5 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-82_pET-28a_URoC0-p6	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #6 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-7_pET-28a_URoC0-p7	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #7 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-83_pET-28a_URoC0-p8	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #8 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-6_pET-28a_URoC1a	Bacterial expression of URoC1a with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #9 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d-g, j; Fig. 2a-g; Extended Data Fig. 1b-c, e-g; Extended Data Fig. 2b (available on Addgene upon publication).
jbC-84_pET-28a_URoC0-p10	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting	Fig. 1d; Extended Data Fig. 1b-c.

	from the #10 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	
jbC-85_pET-28a_URoC0-p11	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #11 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-86_pET-28a_URoC0-p12	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #12 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-87_pET-28a_URoC0-p13	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #13 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-88_pET-28a_URoC0-p14	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #14 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-89_pET-28a_URoC0-p15	Bacterial expression of URoC variant with 8xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #15 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 1d; Extended Data Fig. 1b-c.
jbC-3_pET-28a_dURoC1a	Bacterial expression of control URoC1 with 8xG4S, CaM with all EF hands mutated out and CaMKI CBP substitute-insertion at the #9 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Extended Data Fig. 1d, f; Extended Data Fig. 2a-b (available on Addgene upon publication).
jbC-41_pET-28a_URoC1-FL2	Bacterial expression of URoC1 variant with 2xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #9 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 2d-e.
jbC-38_pET-28a_URoC1-FL4	Bacterial expression of URoC1 variant with 4xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #9 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 2d-e.
jbC-47_pET-28a_URoC1-FL12	Bacterial expression of URoC1 variant with 12xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #9 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 2d-e.
jbC-49_pET-28a_URoC1-FL16	Bacterial expression of URoC1 variant with 16xG4S, CaM-EF1KO and CaMKI CBP substitute-insertion starting from the #9 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 2d-e.
jbC-50_pET-28a_URoC1-EF2KO	Bacterial expression of URoC1 variant with 8xG4S, CaM-EF2KO and CaMKI CBP substitute-insertion starting from the #9 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 2f-g.

jbC-51_pET-28a_URoC1-EF3KO	Bacterial expression of URoC1 variant with 8xG4S, CaM-EF3KO and CaMKI CBP substitute-insertion starting from the #9 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 2f-g.
jbC-52_pET-28a_URoC1-EF4KO	Bacterial expression of URoC1 variant with 8xG4S, CaM-EF4KO and CaMKI CBP substitute-insertion starting from the #9 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 2f-g.
jbC-43_pET-28a_URoC1-nKO	Bacterial expression of URoC1 variant with 8xG4S, CaM6f and CaMKI CBP substitute-insertion starting from the #9 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 2F-G.
jbC-53_pET-28a_URoC1b	Bacterial expression of URoC1b with 2xG4S, CaM-EF2KO and CaMKI CBP substitute-insertion starting from the #9 residue of the second repeat of a 3-repeat GvpC in pET28a backbone	Fig. 2h-i (available on Addgene upon publication).
jmT-89_pCMV-A-IRES-URoC1b-WPRE-hGH	Transient mammalian expression of gvpA gene and URoC1b gvpC	Fig. 3c-e; Extended Data Fig. 3b (available on Addgene upon publication).
jmT-80_pCMV-A-IRES-dURoC1b-WPRE-hGH	Transient mammalian expression of gvpA gene and dURoC1b gvpC	Fig. 3e, Extended Data Fig.3a (available on Addgene upon publication).
jmL-95_pLV-EF1a-hM3D(Gq)-mCherry-IRES-PuroR-WPRE	Lenti viral construct encoding DREADD receptor hM3D(Gq) tethered with mCherry and puromycin resistance marker.	Fig. 4d-g; Extended Data Fig. 4-6 (available on Addgene upon publication).
jmT-88_pPB-TRE-A-IRES-URoC1b-WPRE-mEF1a-rtTA-T2A-HygR	PiggyBac transposon plasmid for inducible expression of gvpA and URoC1b gvpC genes with a hygromycin resistance gene	Fig. 4d-g; Extended Data Fig. 4-6 (available on Addgene upon publication).
jmT-87_pPB-TRE-A-IRES-dURoC1b-WPRE-mEF1a-rtTA-T2A-HygR	PiggyBac transposon plasmid for inducible expression of gvpA and dURoC1b gvpC genes with a hygromycin resistance gene	Fig. 4d-g; Extended Data Fig. 4-6 (available on Addgene upon publication).
jmT-45_pPB-TRE-NV(allP2A)-emGFP-WPRE-mEF1a-BSD	PiggyBac transposon plasmid for inducible expression of Ana GV chaperone genes with a blasticidin resistance gene	Fig. 4d-g; Extended Data Fig. 4-6 (available on Addgene upon publication).
mARGAna (cassette 2, transient)	Second-generation mammalian acoustic reporter gene (cassette 2), used for transient transfection in this work.	Addgene #197589
mARGAna (cassette 1, transient)	Second-generation mammalian acoustic reporter gene (cassette 1), used as cloning backbone.	Addgene #197588
mARGAna (cassette 1)	Second-generation mammalian acoustic reporter gene (cassette 1) on the PiggyBac transposon plasmid, used as cloning backbones.	Addgene #191341

mARGAna (cassette 2)	Second-generation mammalian acoustic reporter gene (cassette 2) on the PiggyBac transposon plasmid, used as cloning backbones.	Addgene #191342
pAAV-hSyn-hM3D(Gq)-mCherry	Gq-coupled hM3D DREADD fused with mCherry under the control of human synapsin promoter, used as cloning templates to generate jmL-95.	Addgene #50474

Table 3-S1. List and features of genetic constructs used in Chapter 3.
pAAV-hSyn-hM3D(Gq)-mCherry[144] was a gift from Bryan Roth (Addgene plasmid # 50474 ; http://n2t.net/addgene:50474 ; RRID:Addgene_50474)

Chapter 4

IMPROVING ACOUSTIC BIOMOLECULES WITH HIGH-THROUGHPUT ENGINEERING PLATFORM

Hurt R.C. [#], **Jin Z.** [#], Soufi M., Wong K., Sawyer D., Deshpande R., Shen H., Mittelstein D.R., Shapiro M.G. Directed evolution of acoustic reporter genes using high-throughput acoustic screening. [#]**Equal contribution**. *bioRxiv*. DOI: 10.1101/2024.03.30.587094

This chapter is a reformatted version of the above manuscript. My contribution to the work were designing and performing the study, analyzing the data, and preparing the manuscript in collaboration with Hurt R.C., with a focus on the development of the acoustic plate reader and the related high-throughput engineering pipeline.

4.1 Introduction

Acoustic reporter genes (ARGs) — genetically encoded reporters that enable the imaging of gene expression using ultrasound (US) — were first introduced to bacteria in 2018[173] and subsequently to mammalian cells in 2019.[174] ARGs are based on genetically encoded, gas-filled protein nanostructures called gas vesicles (GVs) that originally evolved in buoyant microbes.[175,176] GVs scatter US due to the difference in the density and compressibility of their gaseous interior relative to a surrounding aqueous medium.[177] GVs have been the subject of intense study[176–183], development,[184] application[185–194] in recent years.[195–197] ARGs have received considerable attention due to their ability to enable noninvasive, long-term, real-time imaging of gene expression in both bacterial and mammalian cells deep inside living organisms: in particular, ARGs have been used to image tumor growth[174,184] and colonization by therapeutic bacteria,[184] protease activity,[185] phagolysosomal function,[178] and intracellular Ca^{2+} dynamics.[179] However, despite several successful efforts to engineer the acoustic and expression properties of ARGs, further improvements to the performance of ARGs are needed to enable their most impactful applications.

Unfortunately, the methods currently available for ARG engineering and acoustic characterization are low-throughput, complex to implement, and require a great deal of hands-on time per sample. In particular, manual loading and imaging of individual samples limits throughput to a handful of samples per day. In contrast, the state-of-the-art high-throughput methods used to engineer fluorescent proteins can process far larger libraries in shorter times, with less intervention from users: plate readers can assay thousands of samples per run, and flow cytometers have been used to screen libraries of 10^8 mutants in a single experiment.[180] In the past few decades, a growing suite of protein engineering techniques have been developed[181] and applied with remarkable success to improving fluorescent proteins, opsins, Cas proteins, and other biotechnology tools, but these methods often require the screening of libraries containing thousands of members or more.[182] Thus, the low throughput of current acoustic screening methods prevents the effective use of most of the tools needed to unlock the full potential of ARGs.

In this study, we developed a high-throughput, semi-automated pipeline for acoustic screening of ARGs, and used it to evolve two ARG clusters to improve their nonlinear acoustic signals. Our acoustic plate reader (APR) system is capable of collecting acoustic data on up to 1152 ARG samples in a single automated scan and includes graphical user interfaces (GUIs) for data collection and processing. The APR workflow facilitates faster, more reliable, and more standardized acoustic screening of ARG samples, requiring significantly less hands-on time than current methods. Using this pipeline, we improved the nonlinear acoustic signal produced by two ARG clusters—derived from *Anabaena flos-aquae* and *Bacillus megaterium*—by 5- and 14-fold, respectively, when expressed at physiological temperature. Microscopy revealed that these evolved ARG clusters produce more GVs per cell than their parents.

4.2a A High-throughput Workflow for Acoustic Reporter Genes

GVs are known to respond to US in three regimes, depending on the input pressure applied: linear scattering, nonlinear scattering, and collapse[183,194] (**Fig. 4-1a**). Of particular interest for *in vivo* imaging is the nonlinear scattering regime in which GVs produce significantly more contrast than tissue, putatively by "buckling" of their shells.[183,191–194] This effect has

been exploited previously to non-destructively image GV-expressing bacterial and mammalian cells *in vivo* with high specificity,[184] and enhancing this nonlinear US scattering phenotype is a top priority of current ARG engineering efforts.

The primary GV structural protein — GvpA or its homolog GvpB — creates the cone-tipped cylindrical body of the GV, and optionally GvpC may attach to the outside of this structure and reinforce it mechanically (**Fig. 4-1b**). It has already been shown that engineering GvpC to reduce its binding to GvpA can result in GVs with increased nonlinear signal or decreased collapse pressure,[190] but GvpC serves as a limited target for engineering these phenotypes because not all GV types include GvpC. We chose to explore whether altering the primary structure of the main GV structural protein — GvpA in the *A. flos-aquae* cluster and GvpB in the *B. megaterium* cluster — could increase the amount of nonlinear US contrast produced by *E. coli* expressing either ARG type. We selected the GV gene clusters obtained from these species as our starting points based on the previous use of the *B. megaterium* cluster as a bacterial ARG[173] and the use the *A. flos-aquae* cluster in reconstituted contrast agents and mammalian ARGs,[174,179,190] making it desirable to obtain their efficient bacterial expression. Starting without the benefit of the recently published structures and structural models of these proteins,[188,189] we chose an approach based on random mutagenesis and high-throughput acoustic screening of ARG mutants.

As starting points for evolution, we chose the minimal versions of the WT *B. megaterium* ATCC 19213 cluster[187] (lacking *gvpA*, *gvpP*, and *gvpQ*) and the WT *A. flos-aquae* cluster (with only one copy of *gvpA*, and lacking *gvpV*) (**Fig. 4-1c**). To engineer the desired nonlinear signal and collapse pressure phenotypes, we developed a method for high-throughput, semi-automated characterization of US contrast and GV collapse pressure in *E. coli* (**Fig. 4-1d**).

First, we constructed scanning site saturation libraries of *gvpA* or *gvpB* in these clusters, and performed a selection for high levels of GV expression by inducing transformants on Petri dishes and picking only colonies that appeared white (GV-expressing bacteria appear white because GVs scatter light, in addition to US).[186,196] These mutants were then expressed in

liquid cultures in 96-well format and loaded into agarose phantoms. We imaged these phantoms using an automated scanning setup in which a software-controlled 3D translating stage raster-scans an US transducer above the submerged phantoms (**Fig. 4-1e**), producing a set of US images in which samples with high GV expression appear bright. This pipeline allowed us to generate and acoustically screen several mutant libraries, from which we identified mutants with significantly enhanced acoustic phenotypes. We also created graphical user interfaces (GUIs) to simplify and standardize data acquisition (**Fig. 4-1f**) and analysis. We termed this setup the "Acoustic Plate Reader" (**Fig. 4-S1**, **Video 4-S1**).

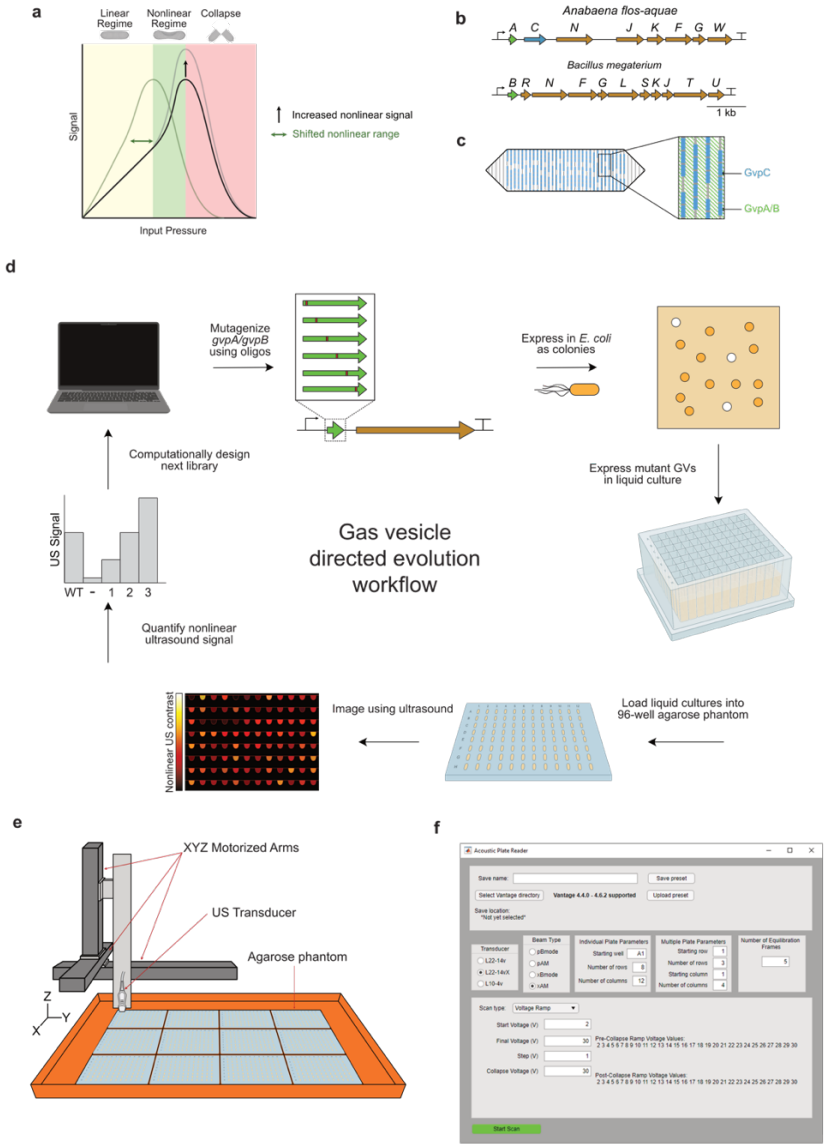


Figure 4-1. High-throughput directed evolution workflow for ARGs. (**a**) Three regimes of GV response to US. (**b**) Roles of the structural proteins GvpA/B and GvpC in GV structure. (**c**) Diagrams of the gene clusters used as starting points for evolution. (**d**) Schematic of directed evolution workflow for ARGs. The starting point GV structural protein is mutagenized, then expressed in *E. coli* as colonies on Petri dishes. Colonies that turn white are assumed to produce GVs, and are picked and expressed in liquid culture. Cultures of GV-expressing E. coli are then loaded into agarose phantoms and imaged using US at 15.625 MHz. The resulting nonlinear US intensity data are used to rank the performance of mutants and select the most promising ones for further mutagenesis. (**e**) Schematic of the Acoustic Plate Reader (APR), which is used for automated US image collection of up to 1152 samples of GV-expressing *E. coli* arrayed in 96-well agarose phantoms. (**f**) Image of the graphical user interface for the APR.

4.2b Optimizing Acoustic Reporter Gene Expression

Before engineering the structural proteins, we optimized the expression of the WT *A. flos-aquae* and *B. megaterium* gene clusters in *E. coli* at 37°C. For each cluster, we cloned three origins of replication (ORIs) of different strengths (~40, ~20, and ~5 copies/cell)[197] (**Fig. 4-2a, b**), and assessed their performance in liquid culture as a function of inducer concentration. For both clusters, the strongest ORI tested gave the highest nonlinear US signal (**Fig. 4-2c, d**), and was chosen for future experiments. With the optimal ORIs selected for expression (**Fig. 4-2e, f**), we then sought to optimize the autoinduction conditions to maximize nonlinear signal (in autoinduction media, increasing the concentration of glucose increases the cell density at which induction occurs, while increasing the concentration of the inducing sugar increases the level to which the transcription unit is induced). We performed titrations of glucose and arabinose and assessed the resulting nonlinear signal from the expressed constructs (**Fig. 4-2g, h**); we decided on concentrations of 0.25% glucose and 0.05% arabinose for induction of these constructs in future experiments, as these conditions yielded high GV expression from both constructs while leaving enough induction dynamic range to tune expression of mutants later without the need to alter any regulatory elements. We observed that US signal from the *A. flos-aquae* cluster peaked at a moderate arabinose concentration (**Fig. 4-2c, g**), while expression from the *B. megaterium* cluster was highest at the maximum concentration (**Fig. 4-2d, h**). We suspect that the signal decline from the *A. flos-aquae* cluster at high arabinose concentrations is due to the high metabolic burden associated with expressing so many non-native proteins in *E. coli*.

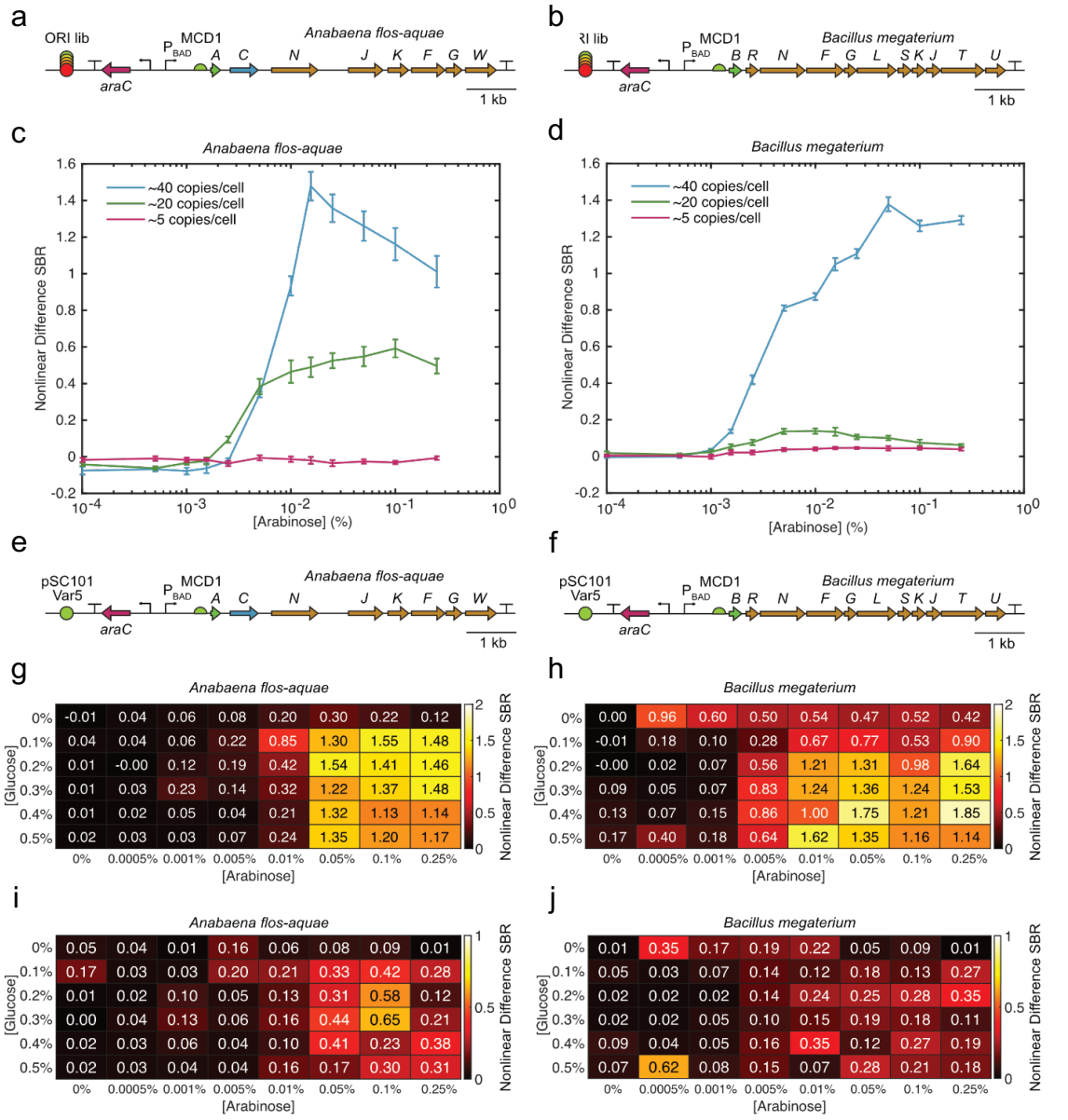


Figure 4-2. Optimization of GV expression from the WT *A. flos-aquae* and *B. megaterium* gene clusters. (**a-b**) Diagrams of the WT *A. flos-aquae* and *B. megaterium* gene clusters with libraries of origins of replication (ORIs) of different strengths. (**c-d**) Nonlinear US signal produced from expression of both clusters at three different copy numbers as a function of inducer concentration. The nonlinear difference SBR is the difference in signal-to-background ratio between pre- and post-collapse images of each sample (see Methods for details). Error bars represent standard error. N=8 biological samples (each an average of 3 technical replicates). (**e-f**) Diagrams of the optimized WT *A. flos-aquae* and *B. megaterium* gene clusters used for directed evolution, both of which used the pSC101-var5 ORI (~40 copies/cell). (**g-h**) Mean and STD (**i-j**) nonlinear US signal produced by both WT clusters as a function of the concentrations of glucose and arabinose used for autoinduction. The concentrations selected for GV expression during library screening were 0.25% glucose and 0.05% arabinose. N=3 biological samples (each an average of 3 technical replicates).

4.2c Directed Evolution of Acoustic Reporter Genes

To improve the nonlinear signal from the WT *A. flos-aquae* and *B. megaterium* clusters in *E. coli*, we designed scanning site-saturation libraries of the genes encoding the primary GV structural protein for each (*i.e.*, *gvpA* for *A. flos-aquae*; *gvpB* for *B. megaterium*) (**Fig. 4-3a, b**). This resulted in libraries containing 1400 and 1740 members for *gvpA* and *gvpB*, respectively (**Table 4-S1**, Library Round 1). We constructed these libraries using a Golden Gate-based version of cassette mutagenesis,[198] in which mutagenic oligonucleotides that tile the gene of interest are synthesized and cloned into an acceptor vector (**Fig. 4-S2a, b**; see methods for details). We chose this approach over error-prone PCR because of its ability to generate defined libraries which have a controllable number of mutations per member and which lack unwanted mutants (*i.e.*, premature stop codons and multiple codons that code for the same mutant).

When induced in solid culture, these libraries produce three distinct types of colonies: 1) blue colonies, in which the dropout chromoprotein was not excised during assembly, returning the original acceptor vector; 2) low-opacity colonies that lack GV expression or express small amounts, either because they contain a mutant that reduces GV expression or because the mutant gene did not insert correctly during assembly; 3) high-opacity colonies with high GV expression. Colony opacity corresponds to GV expression because the low index of refraction of air inside GVs relative to surrounding aqueous media results in light scattering.[186,199] We used this readout to select only the mutants with high GV expression for further study. We then expressed these mutants (384 from each of the two libraries) in 96-well liquid cultures, and imaged them in the APR in 96-well agarose phantoms (**Fig. 4-1d** and **Fig. 4-S1**). Among the GvpA mutants, only a small number showed significantly higher nonlinear US signal than the WT (**Fig. 4-3c**), while many GvpB mutants showed an increase (**Fig. 4-3d**). This was likely because the GvpA construct fails to produce strongly opaque colonies when grown in solid culture, making it impossible to enrich for functional mutants prior to US screening; thus, the mutants screened via US from the GvpA library represent a random subset of the library, while those from the GvpB library are enriched for GV-producing sequences.

We chose up to 10 unique mutants with the highest US signal from each library and re-cloned them (see methods) for validation and further characterization of their nonlinear acoustic signal (**Fig. 4-3e, f** and **Fig. 4-S3a, b**) and OD600 (**Fig. 4-S3c, d**). The top two mutants from each library — GvpA-T6A and -K22R, and GvpB-K22D and -S87I — generated nonlinear US signals 3.07-, 3.44-, 8.54-, and 7.41-fold higher than their parents, respectively, while growing to similar densities in liquid culture. The mutations found in the top 5 and top 10 variants from the GvpA and GvpB libraries, respectively, are shown in **Fig. 4-3g, h**. These mutations cluster in the N-terminal linker and bridge domains, as well as the hinge and wall domains, and the C-terminal tail.[188,189] Notably, no mutations occur in the C-terminal stabilization domain.

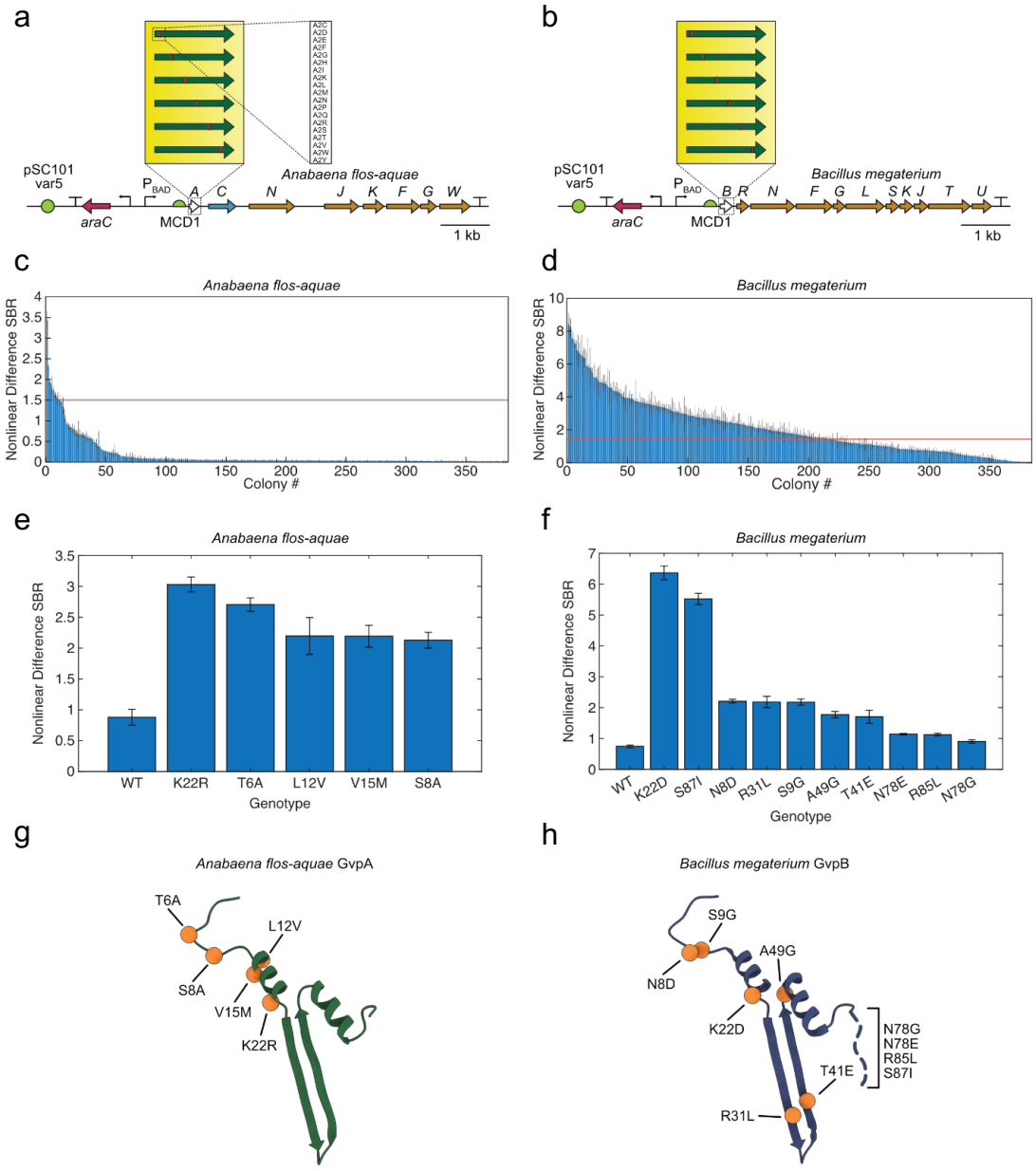


Figure 4-3. First round of directed evolution of *A. flos-aquae* and *B. megaterium* structural proteins. (**a-b**) Diagrams of the mutagenized *A. flos-aquae* and *B. megaterium* gene clusters, depicting the scanning site saturation libraries screened in the first round of evolution. (**c-d**) Nonlinear US difference signal-to-background ratio (SBR) from all screened mutants of both clusters. Red lines indicate the difference SBR of the WT for that cluster. Error bars represent standard error. N=3 technical replicates of one biological sample. (**e-f**) Nonlinear US difference SBR for the WT and top mutants for each cluster. Error bars represent standard error. N=4 biological samples (each an average of 3 technical replicates). (**g-h**) Locations of top mutations from **e-f** in the GvpA/GvpB structure (PDB: 8GBS and 7R1C).

We next performed a second round of directed evolution on these clusters by generating three distinct libraries: two scanning site saturation libraries of the top two mutants of *A. flos-aquae gvpA* (T6A and K22R) and a paired recombination library of the top 10 unique mutants of *B. megaterium gvpB* (**Fig. 4-4a, b**) (though some members of this library contained three mutations due to a well-documented issue with amplifying oligonucleotide pools; see Methods for explanation). We cloned and screened these libraries using the same methods described for the first round of evolution (**Fig. 4-1d**), and identified several mutants with greatly improved signal over their parents in both libraries (**Fig. 4-4c, d**).

We characterized the top 10 unique mutants from each library in terms of their nonlinear acoustic signal (**Fig. 4-4e, f** and **Fig. 4-S4a, b**) and OD600 (**Fig. 4-S4c, d**), and identified GvpA-T6A-L40A, GvpA-T6A-I48V, and GvpB-S9G-R31L-R85L as the top-performing variants. These mutants generated nonlinear signals 5.32-, 5.37-, and 13.93-fold higher than their parents, respectively, while allowing the bacteria expressing them to grow to similar densities in liquid culture. The mutations found in the top 2 and top 1 variants from the second-round GvpA and GvpB libraries, respectively, are shown in **Fig. 4-4g, h**. Similar to the mutations identified from the first-round libraries, these mutations cluster in the N-terminal linker domain, as well as the hinge and wall domains, and the C-terminal tail, but not the C-terminal stabilization domain.[188,189] Representative nonlinear US images of GvpA-T6A-L40A and GvpB-S9G-R31L-R85L, as well as the WT parents, are shown in **Fig. 4-4i, j**. In addition to showing increased nonlinear contrast (**Fig. 4-S5a, b**), the top variants have slightly higher collapse pressure than their WT parents when normalized for nonlinear contrast (**Fig. 4-S5c, d**).

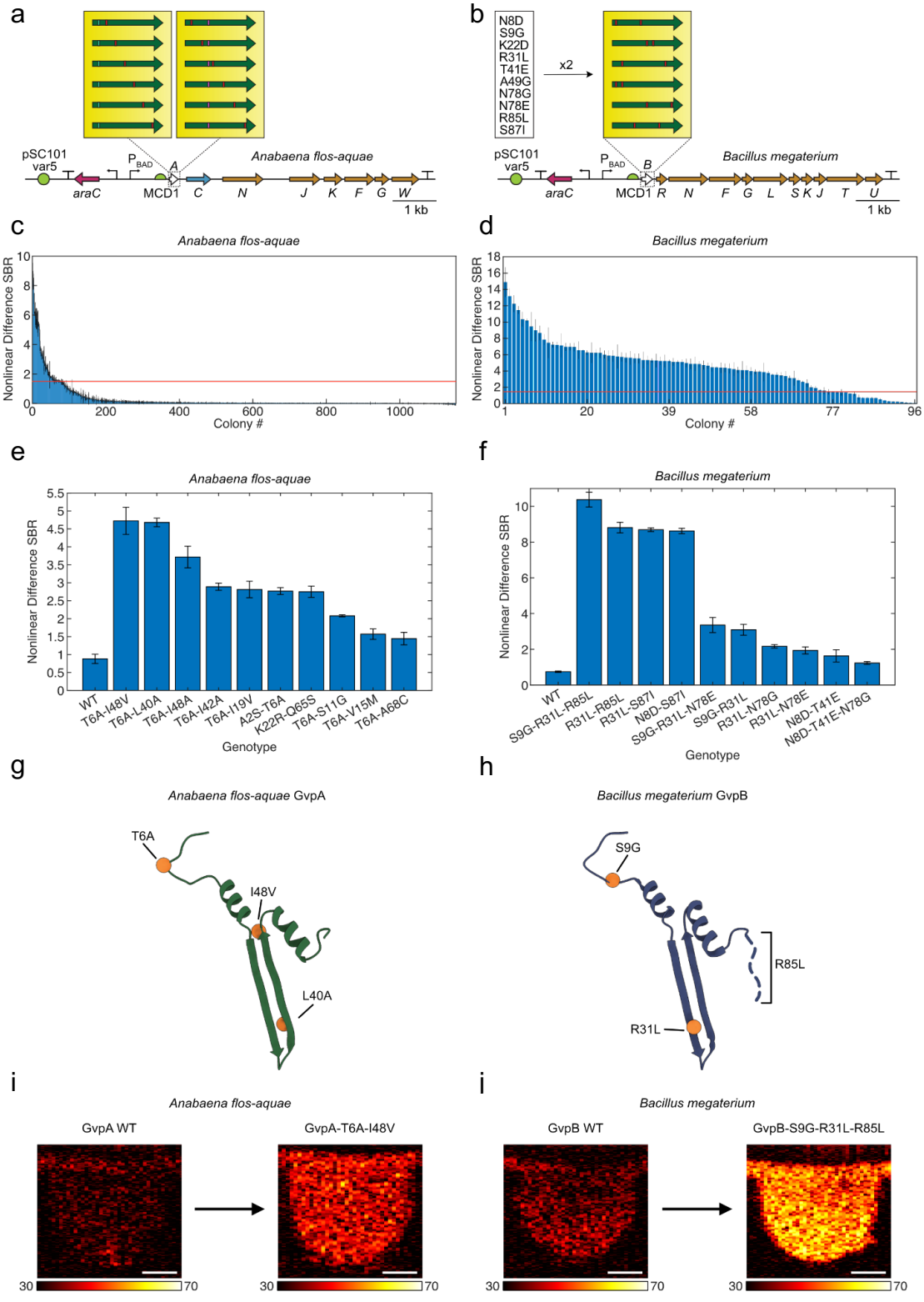

a
pSC101 var5
araC
MCD1
A
C
N
J K F G W
1 kb
Anabaena flos-aquae
b
N8D
S9G
K22D
R31L
T41E
A49G
N78G
N78E
R85L
S87I
x2
pSC101 var5
araC
MCD1
B
R N F G L S K J T U
1 kb
Bacillus megaterium
c
Anabaena flos-aquae
Nonlinear Difference SBR
Colony #
d
Bacillus megaterium
Nonlinear Difference SBR
Colony #
e
Anabaena flos-aquae
Nonlinear Difference SBR
WT
T6A-I48V
T6A-L40A
T6A-I48A
T6A-I42A
T6A-I19V
A2S-T6A
K22R-Q65S
T6A-S11G
T6A-V15M
T6A-A68C
Genotype
f
Bacillus megaterium
Nonlinear Difference SBR
WT
S9G-R31L-R85L
R31L-R85L
R31L-S87I
N8D-S87I
S9G-R31L-N78E
S9G-R31L
R31L-N78G
R31L-N78E
N8D-T41E
N8D-T41E-N78G
Genotype
g
Anabaena flos-aquae GvpA
T6A
I48V
L40A
h
Bacillus megaterium GvpB
S9G
R85L
R31L
i
Anabaena flos-aquae
GvpA WT
GvpA-T6A-I48V
30
70
j
Bacillus megaterium
GvpB WT
GvpB-S9G-R31L-R85L
30
70

Figure 4-4. Second round of directed evolution of *A. flos-aquae* and *B. megaterium* structural proteins. (**a-b**) Diagrams of the mutagenized *A. flos-aquae* and *B. megaterium* gene clusters used in the second round of evolution. The best two mutants of *A. flos-aquae* gvpA were used as parents for another scanning site saturation library, and the best ten mutants of *B. megaterium* gvpB (listed in figure) were used to create a paired recombination library. (**c-d**) Nonlinear US difference signal-to-background ratio (SBR) from all screened mutants of both clusters. Red lines indicate the difference SBR of the WT for that cluster. Error bars represent standard error. N=3 technical replicates of one biological sample. (**e-f**) Nonlinear US difference SBR for the WT and top ten mutants for each cluster. Error bars represent standard error. N=4 biological samples (each an average of 3 technical replicates). (**g-h**) Locations of mutations from the top mutants from **e-f** in the GvpA/GvpB structure. (PDB: 8GBS and 7R1C) (**i-j**) Representative nonlinear US images of the brightest mutants identified in this study, as well as their respective WT parents. Scale bars 1 mm.

4.2d Expression Characteristics of Top Mutants

We performed whole-cell transmission electron microscopy (TEM) on *E. coli* expressing either WT or mutant ARGs to evaluate changes in expression levels. TEM revealed that these mutations increased the expression levels of both ARG types, either by increasing both the typical and maximum number of GVs per cell (in the case of GvpA-T6A-L40A and GvpA-T6A-I48V) or by making the number of GVs expressed per cell more consistent across all cells in the culture (in the case of GvpB-S9G-R31L-R85L) (**Fig. 4-5a-e**, **Fig. 4-S6, S7**).

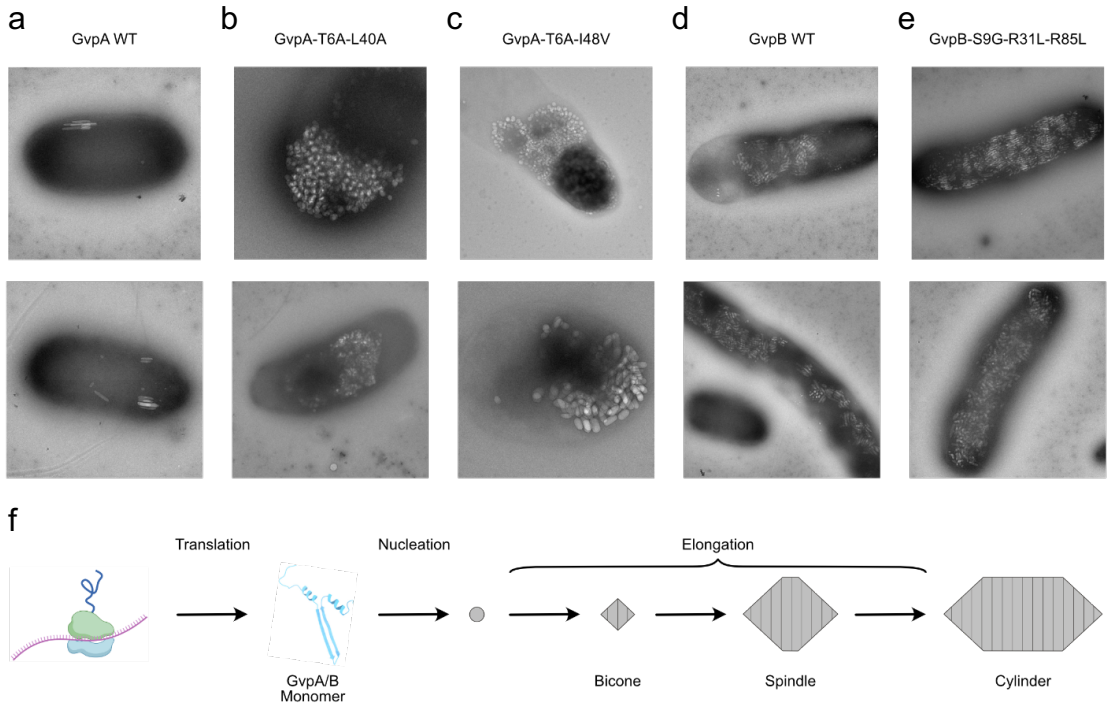


Figure 4-5. TEM of *E. coli* after expressing top-performing *A. flos-aquae* GvpA and *B. megaterium* GvpB mutants. (**a-e**) TEM images of WT and mutant GVs expressed in *E. coli*. (**f**) Diagram of the GV formation process.

4.3 Discussion and Outlook

Our results establish the first method for high-throughput, semi-automated acoustic screening of biomolecules expressed in cells. When used to evolve two ARG clusters—those from *A. flos-aquae* and *B. megaterium*—this method yielded ARG constructs which show 5- to 14-fold improvements in their nonlinear acoustic scattering.

The mutations identified in this study appear to increase nonlinear US signal by increasing the maximum number of GVs produced per cell and/or making GV production more consistent across a cell population. These changes could be due to improved expression of GvpA/GvpB monomers or their incorporation into growing GVs (**Fig. 4-5f**). In addition, it is possible that some mutations contribute to increased nonlinear scattering of individual GVs by altering their mechanical properties.[183,193,194,200]

These results represent a major advance in the way that acoustic biomolecules can be engineered. In the same way that high-throughput screening tools such as plate readers and flow cytometry enabled the engineering of fluorescent proteins and the many sensors derived from them by dramatically increasing the sizes of libraries that can be screened in these experiments, so too will the increased throughput, reliability, and standardization introduced by the Acoustic Plate Reader enable the engineering of next-generation ARGs and their derivatives.

While these evolved constructs represent substantial improvements over their parents, further improvements are required. First, both ARGs could benefit from further improvements in nonlinear contrast; this will likely be achieved through a combination of protein engineering (including not only the structural proteins engineered in this study, but also the assembly factors that assist in GV formation) and expression tuning (ORI, RBSs, and promoter) aimed at increasing both the amount of nonlinear contrast produced per GV and the number of GVs produced per cell. Relatedly, it would be desirable to engineer GVs with higher collapse pressures or ones whose collapse pressure is unchanged while having a significantly lower buckling threshold.

Additional engineering is needed to ensure the mutational stability of these constructs for *in vivo* applications, for example through chromosomal integration or inclusion of plasmid stability elements.[184] APR screening could facilitate any tuning required at the transcriptional (promoter) and translational (RBS) levels. Such tuning would potentially make the more compact *A. flos-aquae* and *B. megaterium*-derived ARGs competitive with the larger *Serratia*-derived ARGs, which currently provide the best *in vivo* imaging performance.[184] To further accelerate ARG development, we need a deeper understanding of how mutations to GvpA/GvpB affect both their structures[188,189] and the protein-protein interactions in which they participate during GV assembly[201–204], as well as biochemical methods to characterize intermediate steps that cannot be assayed by ultrasound, such as GV nucleation (**Fig. 4-5f**).

By enabling the large-scale generation and high-throughput acoustic screening of ARG variants, the APR and its associated protocols allow the suite of modern protein engineering techniques to be applied to ARGs.

4.4 Material and Methods

Plasmid construction via MoClo

The EcoFlex MoClo system[205] was used to create all vectors cloned in this study, including existing parts (Addgene Kit # 1000000080) and custom-made parts (**Table 4-S2**). Custom-made parts were introduced into the existing EcoFlex system as follows: 1) ORIs were selected from the pSC101-varX series[197]; promoters were selected from the Marionette series[206]; RBSs were selected from the MCDX series[207]; terminators were selected from the ECK and LXSXPX series[208]; 2) parts were either synthesized as fragments (Twist Bioscience) and subsequently PCRed using Q5 (NEB), or synthesized as duplex oligos (IDT); 3) parts were cloned into the corresponding part acceptor vector (**Table 4-S2**) via Golden Gate to ensure that they received the appropriate assembly overhangs. EcoFlex assemblies were conducted as described in **Note 4-S1** and electroporated into NEB Stable E. coli (except for the MetClo-based library acceptor vectors, which were transformed into DH10B-M.Osp807II[209]). Transformations were recovered for 2 hr in 1 mL of SOC at 37°C

and 250 RPM, and plated on Petri dishes containing Lennox LB with 1% agar, 100 ug/mL kanamycin, and 1% glucose (for catabolite repression of the PBAD promoter). Colonies were picked into 1.5 mL liquid cultures of Lennox LB with 100 ug/mL kanamycin and 1% glucose in 96-well format and grown overnight to saturation. These cultures were then miniprepped using reagents and a protocol from Qiagen, a lysate clearing plate from Bio Basic (SD5006), and a DNA-binding plate from Epoch Life Sciences (2020-001). All constructs were verified by whole-plasmid sequencing (Primordium Labs).

Liquid culture GV expression in E. coli

GVs were expressed in E. coli liquid cultures in 96-well format according to the following general protocol, with modifications for specific experiments described below.

Miniprepped DNA was electroporated into NEB Stable *E. coli*, and transformations were recovered for 2 hr at 37°C in 1 mL of SOC. Transformations were then inoculated at a dilution of 1:100 into autoinduction Lennox LB containing 100 μg/mL kanamycin, 0.6% glycerol, and the appropriate concentrations of glucose and inducer for the experiment (see below). These expression cultures were set up in 500 uL volumes in deep-well 96-well plates (square wells used for maximum culture aeration; USA Scientific 1896-2800) sealed with porous tape (Qiagen 19571) and incubated at 37°C and 350 RPM for 20 hr. Cultures were stored at 4C until being loaded into phantoms for Acoustic Plate Reader scans. For the concentrations of glucose and arabinose described below, in experiments where titrations were used, 100X stocks of these sugars were prepared in 1X PBS and diluted 1:100 into the cultures when setting up the experiments.

The following concentrations were used for the ORI titration experiments shown in Figure 2A-D: glucose: 0.25%; arabinose: 0, 0.0005, 0.001, 0.00155, 0.0025, 0.005, 0.01, 0.0155, 0.025, 0.05, 0.1, 0.25%.

The following concentrations were used for the parent expression optimization experiments shown in Figure 2E-H: glucose: 0, 0.1, 0.2, 0.3, 0.4, 0.5%; arabinose: 0, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.25%.

The following modifications were made for the library screening experiments shown in **Fig. 4-3a-c** and **4-4a-c**; 1) assembled libraries were transformed multiple times into NEB Stable *E. coli*, and it was ensured that the number of transformants produced was at least 100X the number of unique sequences expected in the library; 2) prior to expression in liquid culture, libraries were expressed in solid culture as colonies on Lennox LB with 100 ug/mL kanamycin, 0.6% glycerol, 0.25% glucose, and 0.05% arabinose at a density of ~100 colonies/dish. Colonies were grown for 48 hr at 37°C, and 380 opaque colonies were picked for each library, as well as 4 colonies for the library's parent, into the wells of 96-well PCR plates containing 100 uL of Lennox LB with 100 µg/mL kanamycin and 1% glucose, and grown to saturation overnight at 30°C. These saturated liquid cultures, rather than transformations, were used to set up expression cultures as described above; 3) 0.25% glucose, and 0.05% arabinose were used to induce expression in these experiments.

The following concentrations were used for the mutant expression experiments shown in **Fig. 4-3e-h** and **4-4e-h**: glucose: 0.25%; arabinose: 0.05%.

The following concentrations were used for the multiplexing experiments shown in **Fig. 4-5a-b**: glucose: 0.25%; arabinose: 0.05%.

Scanning site saturation and recombination library generation

Scanning site saturation libraries were generated via a Golden Gate-based version of the cassette mutagenesis strategy previously described.[210] Briefly, the *A. flos-aquae* GvpA and *B. megaterium* GvpB coding sequences were divided into sections that tiled the gene, and oligos were designed to have a variable middle region with flanking constant regions against which PCR primers were designed (these primers also contain the evSeq[211] inner adapters for optional deep sequencing of the library). Depending on the library being created (*i.e.*, scanning site saturation or recombination), the variable region was designed to either sequentially saturate each residue or recombine pairs of the mutations listed in Figure 4B (mutations identified during screening of the first round of scanning site saturation GvpB). The MATLAB scripts used to generate the oligo sequences for both the scanning site saturation and recombination libraries are available in the Supplemental Electronic Material,

and the oligo sequences themselves are listed in **Table 4-S1**. Oligos were synthesized as a pool by Twist Biosciences or Integrated DNA Technologies, and were amplified by PCR (both to make them double-stranded and to generate enough DNA for Golden Gate assembly) using KAPA HiFi HotStart ReadyMix according to the manufacturer's instructions, but with 10 cycles, 100 ng of oligo pool template, and 1 uM of each primer. PCR products were run on a 2% agarose gel and purified using Qiagen reagents according to the manufacturer's instructions, but with a 5 uL final elution volume of water. Fragments were then assembled with the corresponding library acceptor vector (**Table 4-S1**) in a Golden Gate reaction using reagents from New England Biolabs according to the manufacturer's instructions. Assemblies were then expressed (first in solid culture and then in liquid culture) according to the protocol above.

It is important to note that oligo pools whose members have very high sequence similarity (as was the case in the pools used in this study, in which members differed by only a few bp) have a high likelihood of mutation swapping during PCR which increases with the number of cycles used. The manufacturer proposes that this is due to template swapping from one cycle to the next between incompletely-copied strands. We notice this often in our libraries (*i.e.*, libraries synthesized to have two mutations per member would contain a small number of sequences with zero or three mutations per member after PCR), and we minimized the number of PCR cycles used to amplify these libraries. However, some of the best round 2 GvpB mutants contained three mutations for this reason.

Acoustic Plate Reader scans

The general protocol for preparing and scanning liquid cultures samples of GV-expressing E. coli in 96-well format is described in **Fig. 4-S1** and the corresponding figure caption. Detailed instructions on how to build and use this system, as well as troubleshooting and bug-reporting information, are provided at https://github.com/shapiro-lab/acoustic-plate-reader.

The specific US pulse sequence parameters used for collecting the data shown in each figure are presented in **Table 4-S3**.

For pre-/post-collapse and voltage ramp scans, the nonlinear difference SBR was calculated as: [(pre-collapse sample mean) - (post-collapse sample mean)] / (post-collapse background mean), where means are calculated from the nonlinear signal in a region of interest containing either the sample or an empty region of the phantom. For voltage ramp scans, this quantity was calculated for each pre-collapse image; for simple pre-/post-collapse scans, this quantity was calculated only once for the single pre-collapse image. Importantly, in all cases the two images being compared in each calculation were acquired at the same voltage (*i.e.*, the pre- and post-collapse images were collected under the same imaging conditions).

For collapse ramp scans, the nonlinear SBR was calculated as: (sample mean) / (background mean), where means are calculated from the nonlinear signal in a region of interest containing either the sample or an empty region of the phantom. This quantity was calculated for each image at each voltage.

Validation of best mutants

Selected mutants from each library were miniprepped and sequenced as described above. Unique mutants were then re-cloned using MoClo (see above) before undergoing validation testing to avoid the possibility that these plasmids accrued expression-reducing mutations during the GV expression steps performed during library screening. To prepare fragments for these MoClo assemblies, *gvpA*/*gvpB* mutant CDSs were PCRed using the primers described in **Table 4-S4** (which were selected based on the sequence of the mutant being amplified) and prepared for Golden Gate assembly as described above.

OD600 measurements

OD 600 culture measurements were performed on a Tecan Spark plate reader using the "Absorbance" protocol with the following settings: 600 nm measurement wavelength, 10 flashes, 50 ms settle time. Measurements were collected for 200 uL of culture and normalized to a 1 cm path length using the built-in "Pathlength Correction" feature.

Negative stain TEM imaging

Three microliters of *E. coli* culture expressing GVs were applied to a freshly glow-discharged (Pelco EasiGlow, 15 mA, 1 min) Formvar/carbon-coated, 200-mesh copper grid (Ted Pella), and then incubated for 1 minute. Excess solution was blotted with filter paper, and the grids were washed three times with buffer (20 mM HEPES buffer; pH 7.5, 100 mM NaCl). Subsequently, the sample was stained with a 2% uranyl acetate solution for 1 min, blotted, and air-dried. Images were acquired using a Tecnai T12 electron microscope (FEI, now Thermo Fisher Scientific) operating at 120 kV and equipped with a Gatan Ultrascan 2k × 2k CCD.

Abbreviations

US: Ultrasound

GV: Gas Vesicle

ARG: Acoustic Reporter Gene

APR: Acoustic Plate Reader

SBR: Signal-to-Background Ratio

Data Availability

Selected plasmids are available through Addgene (202023, 202024, 202025). Detailed instructions on how to build and use the Acoustic Plate Reader, as well as troubleshooting and bug-reporting information, are provided at https://github.com/shapiro-lab/acoustic-plate-reader. All other data and code are available from the corresponding author upon reasonable request.

4.5 Supplementary Information

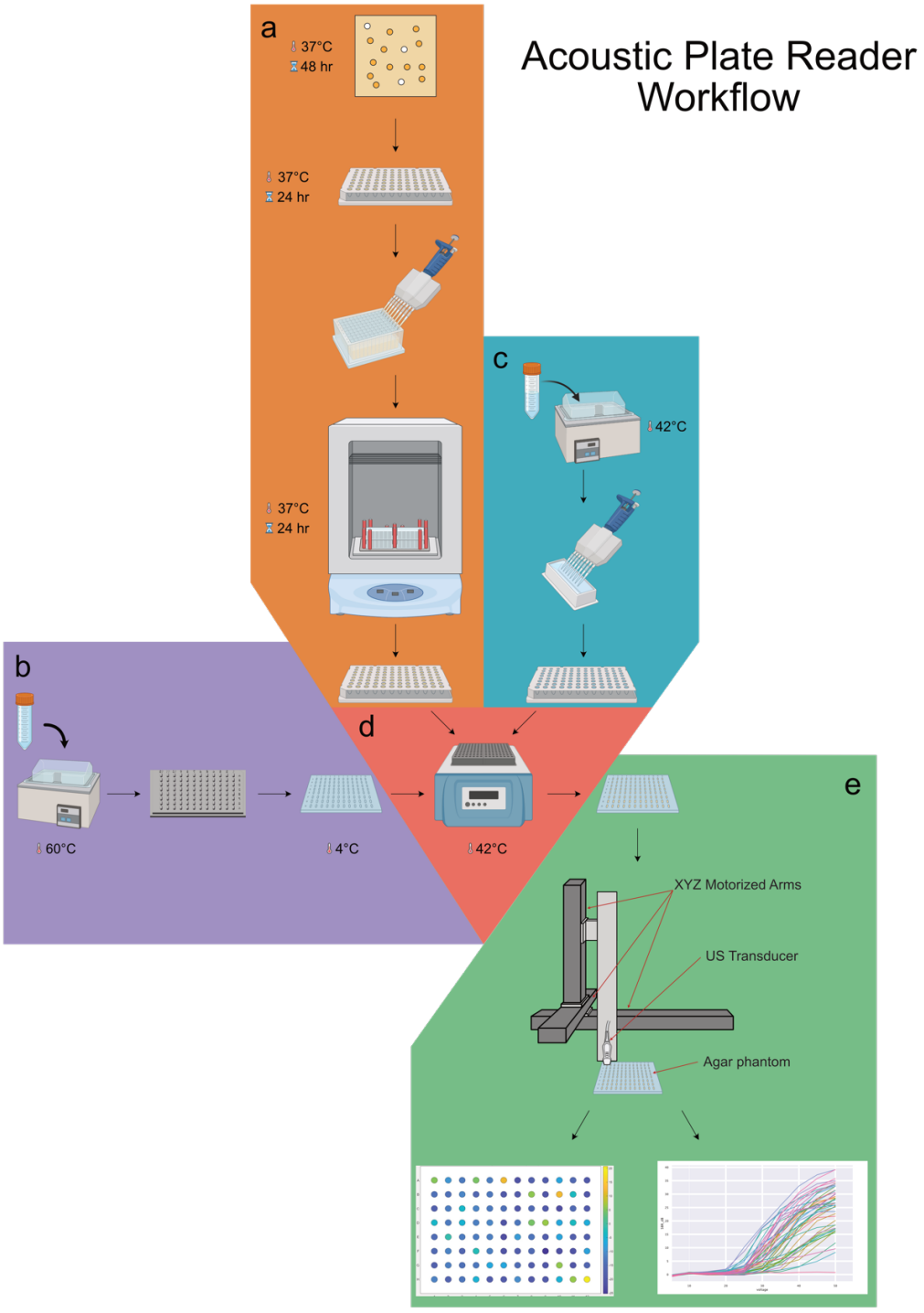
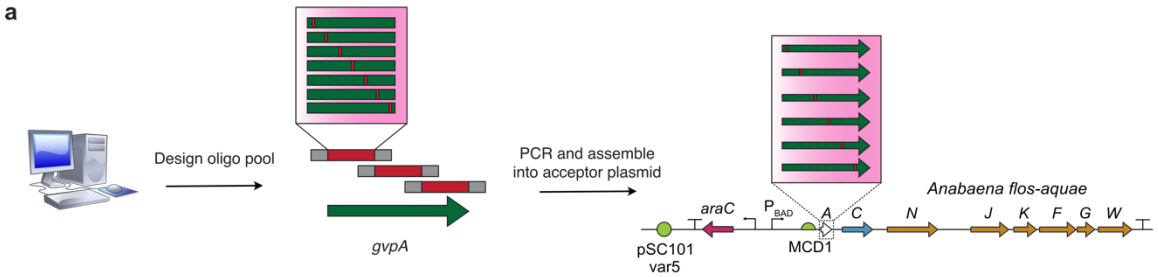


Figure 4-S1. Detailed diagram of the acoustic plate reader workflow. (**a**) GVs are expressed in *E. coli* as colonies on Petri dishes for 48 hr at 37°C, then colonies are picked into LB and grown to saturation in liquid culture for 24 hr at 37°C. These saturated liquid cultures are then diluted 1:100 into autoinduction LB and expressed for 24 hr at 37°C in 500 uL cultures in deep-well 96-well plates (square wells used for maximum culture aeration; USA Scientific 1896-2800). Aliquots of these cultures are aliquoted into an un-skirted 96-well PCR plate for subsequent loading into phantoms. (**b**) A solution of 2% Ultrapure Agarose (Invitrogen, 16500500) is prepared in 1X PBS and incubated at 60C for at least 12 hr to degas. Agarose phantoms are then made by pouring 75 mL of this solution into a 96-well phantom mold and incubating at 4C for 10 min. (**c**) A solution of 1% low-melting-temperature agarose (Goldbio, A-204-100) is prepared in 1X PBS and incubated at 60˚C for at least 12 hr to degas. This solution is then aliquoted into an un-skirted 96-well PCR plate to be used for phantom loading. (**d**) Phantoms from **b** are loaded by placing the 96-well PCR plates from **a** and **c** into 96-well heat block at 42˚C, and combining equal volumes of culture and agarose before pipetting into the empty phantom. (**e**) Phantoms from **d** are scanned using the acoustic plate reader, which generates US data for each sample and can image up to 12 96-well phantoms in a single scan.



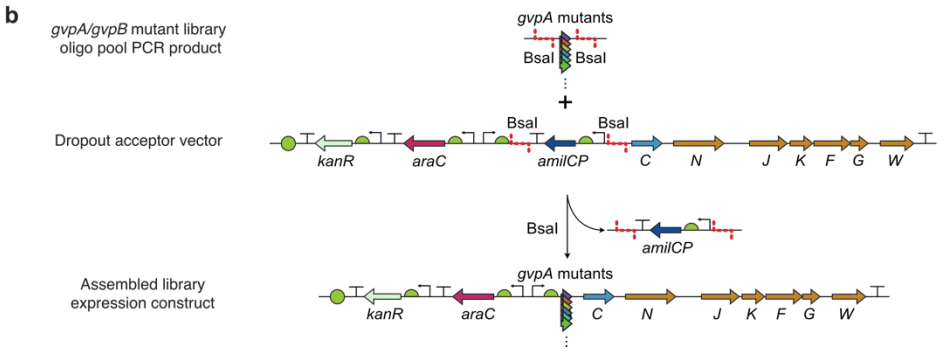


Figure 4-S2. Details of *gvpA*/*gvpB* mutant library construction and screening. (**a**) Overview of workflow for creating either scanning site saturation or recombination libraries. (**b**) Details of library assembly via a Golden Gate-based version of cassette mutagenesis (see Methods).

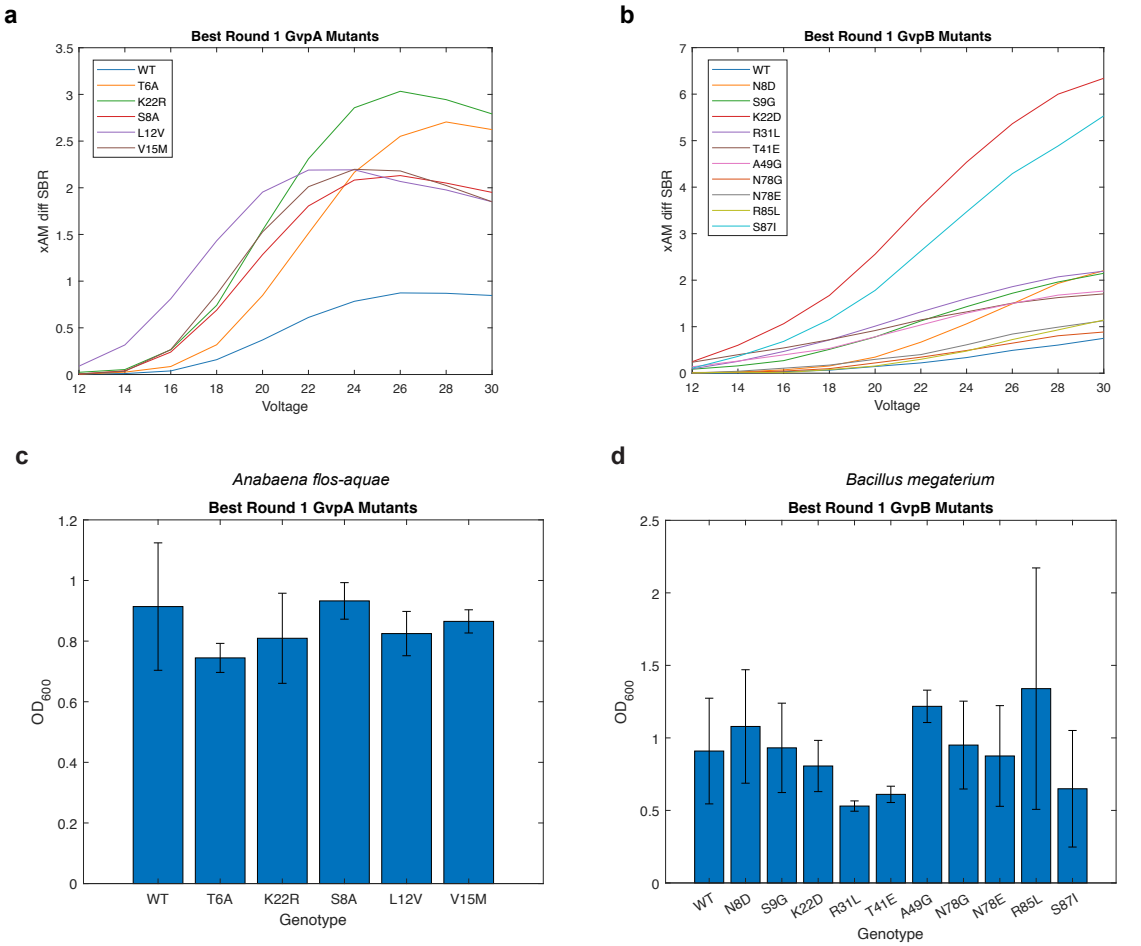


Figure 4-S3. Characterization of the top mutants from Round 1 of evolution. (**a-b**) xAM difference SBR as a function of pressure for each of the top mutants. N=4 biological samples (each an average of 3 technical replicates). (**c-d**) OD600 measurements for the mutants shown in **a-b**. N=4 biological samples.

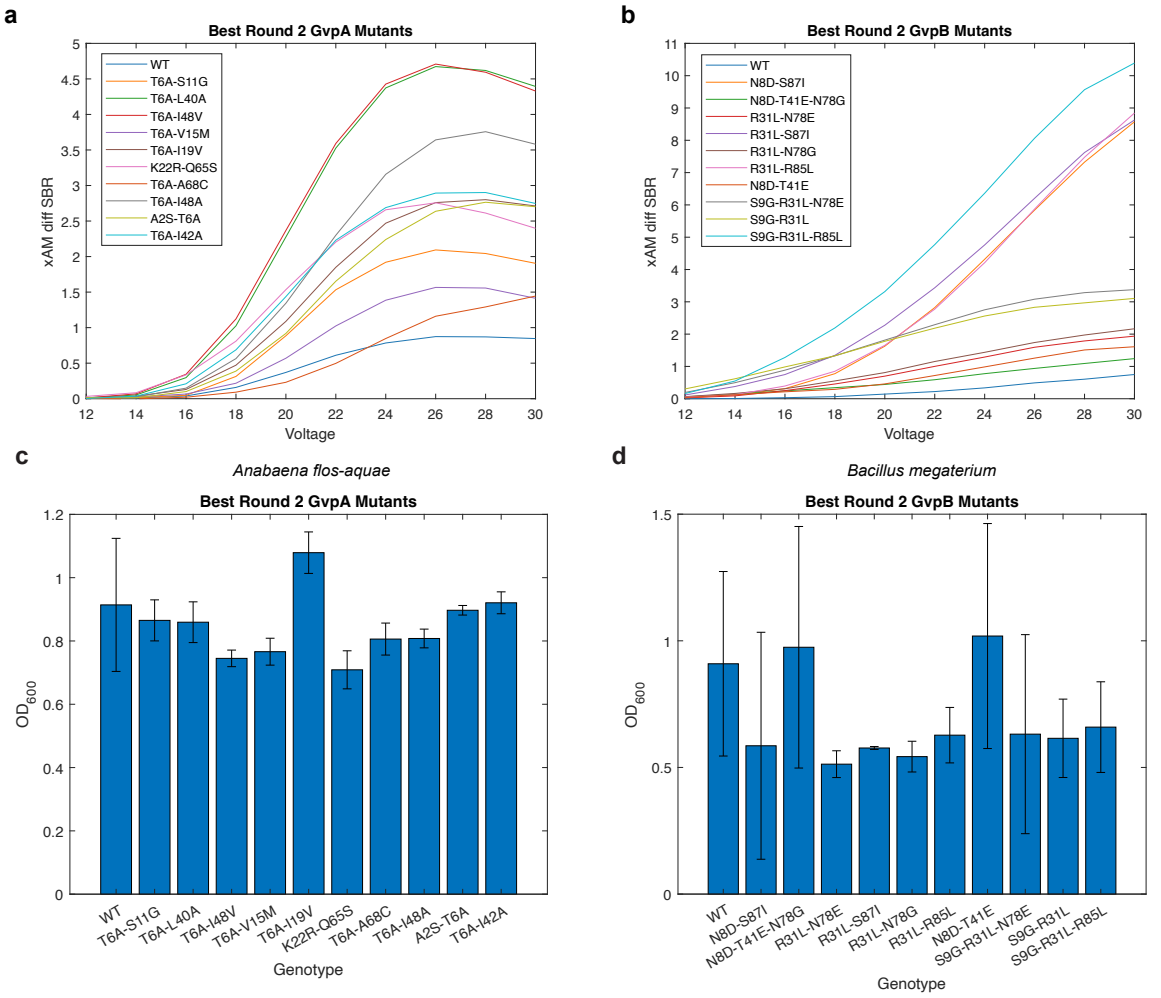


Figure 4-S4. Characterization of the top mutants from Round 2 of evolution. (**a-b**) xAM difference SBR as a function of pressure for each of the top mutants. N=4 biological samples (each an average of 3 technical replicates). (**c-d**) OD600 measurements for the mutants shown in **a-b**. N=4 biological samples.

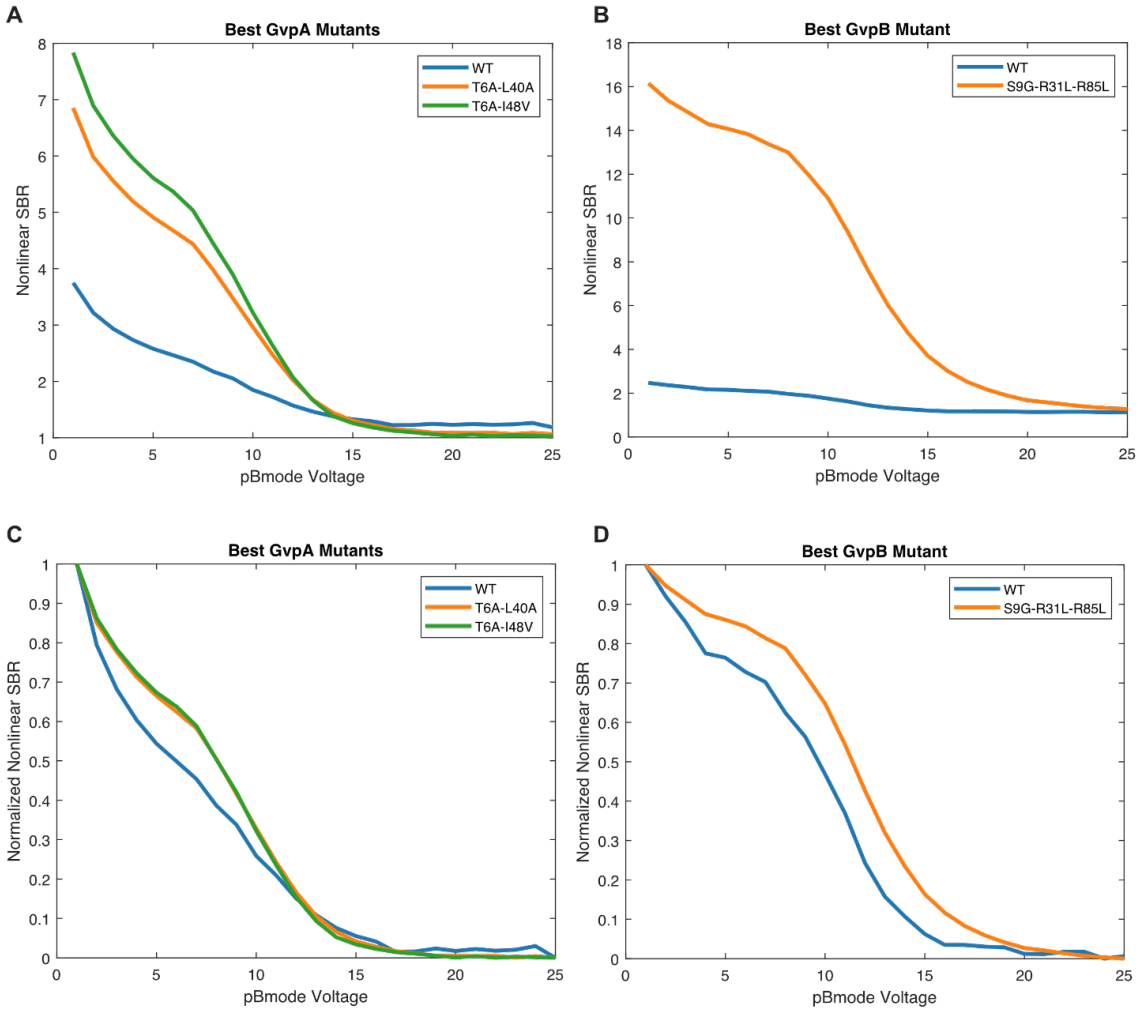


Figure 4-S5. Acoustic collapse pressure curves for the best mutants identified in this study. (**a-b**) xAM acoustic collapse pressure curves for the top-performing mutants identified in this study. (**c-d**) Data from **a-b** normalized to the same min and max. N=4 biological samples.

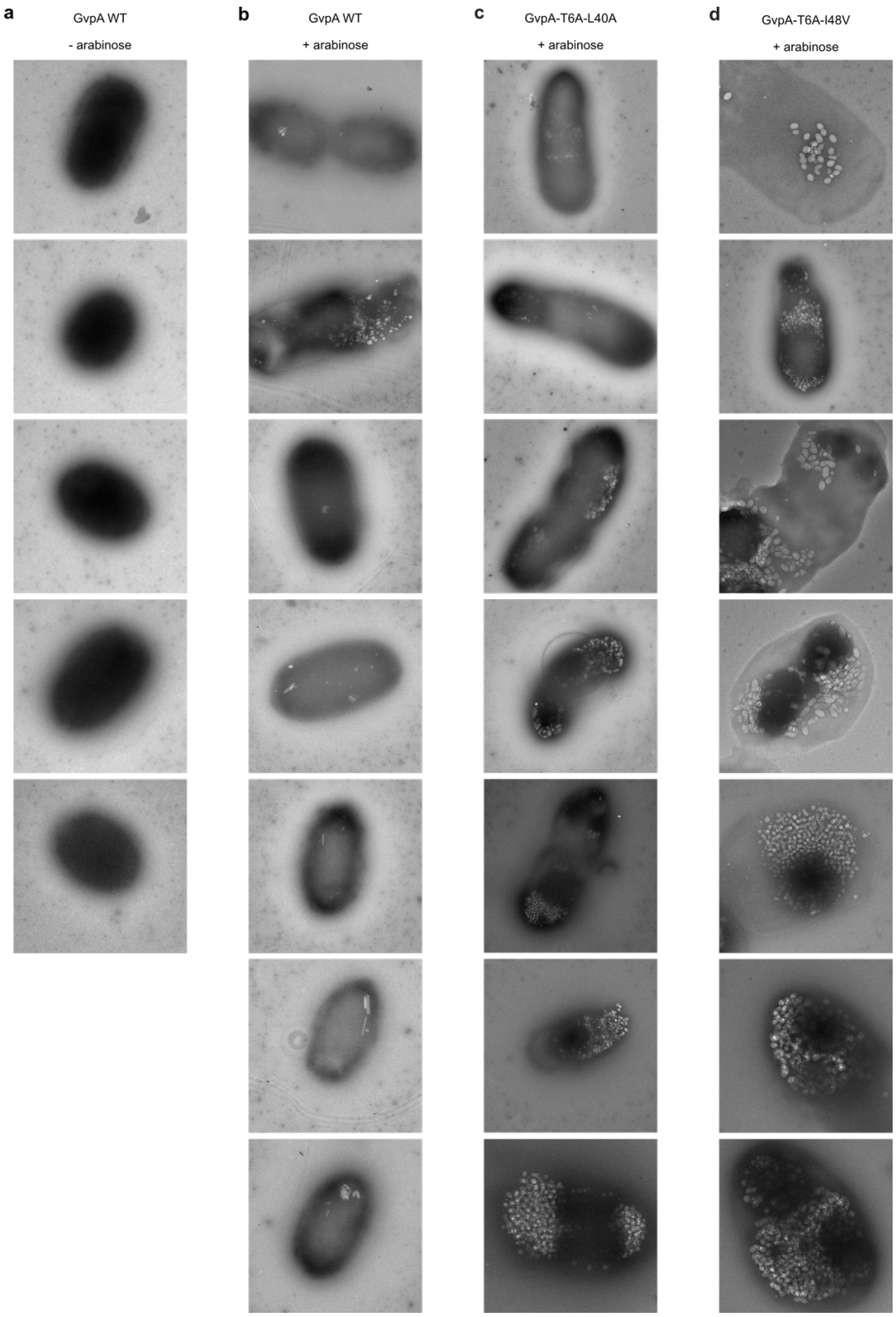


Figure 4-S6. TEM images of *E. coli* cells expressing WT or mutant *A. flos-aquae* GVs. For each sample, 5-25 images were collected; a representative set is shown, ordered from least to most GVs produced per cell.

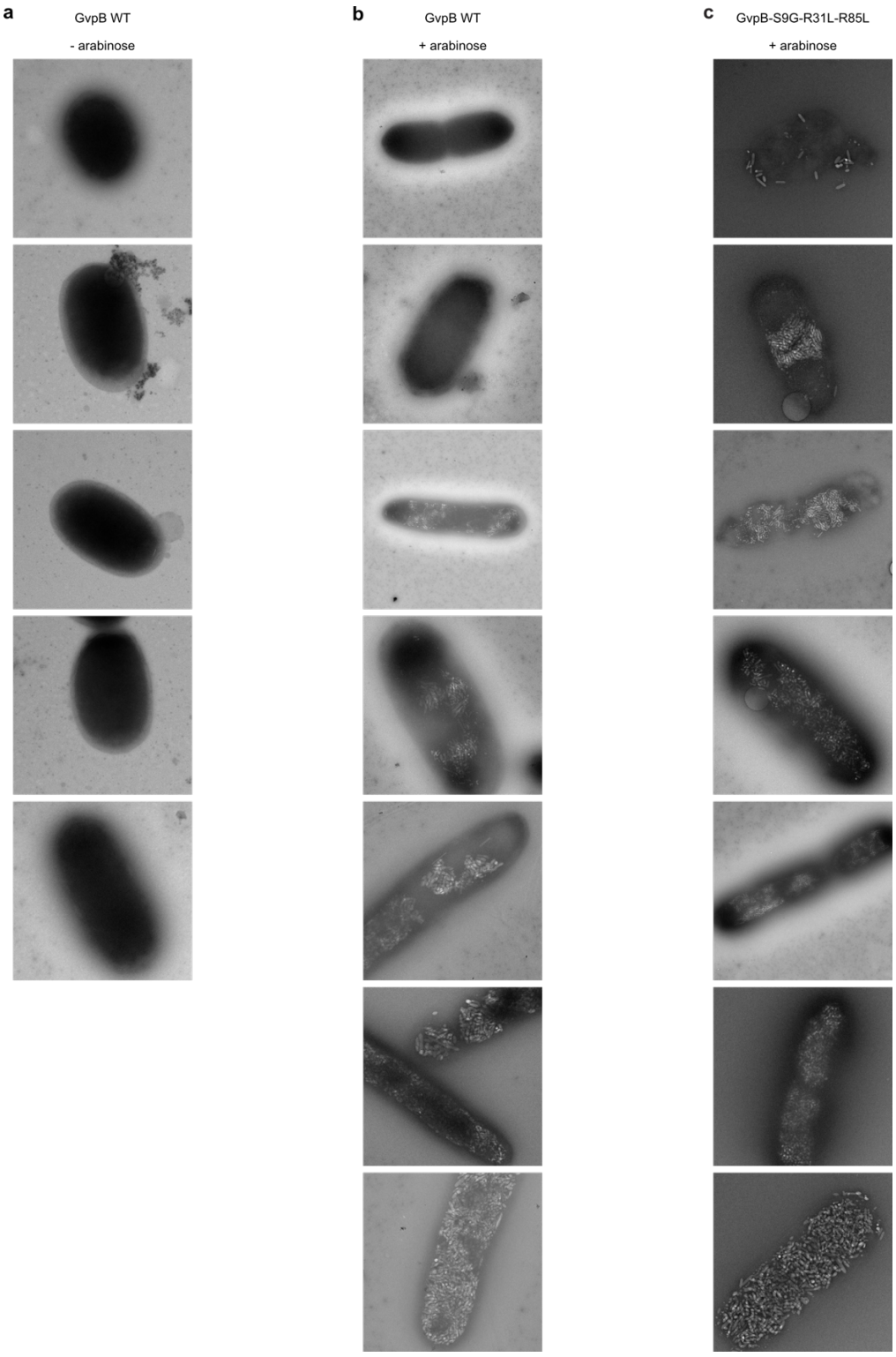


Figure 4-S7. TEM images of *E. coli* cells expressing WT or mutant *B. megaterium* GVs. For each sample, 5-25 images were collected; a representative set is shown, ordered from least to most GVs produced per cell.

Note 4-S1. Golden Gate reactions.
Master mix recipes

Component	Amount per reaction	Stock concentration (NEB)	For 66 reactions
T4 Ligase Buffer	1 uL	10X	66 uL
Hi-T4 DNA Ligase	500 U	400 U/uL	83 uL
BsmBI-v2	15 U	10 U/uL	100 uL
Water	to 5 uL		81 uL

Component	Amount per reaction	Stock concentration (NEB)	For 66 reactions
T4 Ligase Buffer	1 uL	10X	66 uL
Hi-T4 DNA Ligase	500 U	400 U/uL	83 uL
BsaI-HF-v2 or BbsI-HF	15 U	20 U/uL	50 uL
Water	to 5 uL		131 uL

To set up reactions, combine 75 ng of the backbone part with 150 ng of each insert part in a PCR tube with 5 uL of the appropriate master mix and fill to 10 uL with water. Miniprepped parts give higher assembly efficiencies than linear PCR products.

Golden Gate thermocycler protocol
20 min 37/42°C

3 minutes 37/42°C
4 minutes 16°C
Cycle 2-3 x30
10 minutes 50°C
10 minutes 80°C
Hold 4°C

Video 4-S1. Example Acoustic Plate Reader scan. (Right) The Acoustic Plate Reader is scanning six 96-well phantoms. (Left) The computer screen displays the real-time images of linear (left) and nonlinear (middle) contrast, as well as the Verasonics control interface (right).

Table 4-S1. Oligos used for mutagenesis. Sequences of the oligos that composed the four oligo pools used to create the GvpA/GvpB libraries. “Library Round” indicates the round of screening (first or second) in which the oligo was used, and “Sub-Library” indicates the pool in which it was synthesized.

Table 4-S2. Custom-made MoClo parts. Inventory of the MoClo parts added to the base EcoFlex system and used for cloning the constructs in this study.

Table 4-S3. Ultrasound pulse sequences. List of the parameters entered into the APR GUI to perform each scan in this study.

Table 4-S4. PCR primers. Sequences of the primers used to either amplify the oligo pools used to create the libraries, or to re-clone the best *gvpA/gvpB* mutants into Level 0 MoClo part vectors for assembly into expression constructs and subsequent validation.

Chapter 5

CONCLUSION AND FUTURE DIRECTIONS

5.1 Molecular Imaging with Acoustic Biosensors

As we are getting to the end of my thesis, here is a brief recap and the conclusion. We started with the basic concepts and the importance of molecular imaging, getting to know why we need it and what we want to image. Then, we reviewed some widely applied modalities for molecular imaging, with a focus on ultrasound that represents a promising candidate for the ultimate goal. Later, we mentioned its current capabilities and the remaining challenges for molecular ultrasound imaging, which led us to the introduction of gas vesicles and their unique features that enable us to build acoustic biosensors, opening up the possibilities of imaging cellular dynamics noninvasively with ultrasound. In the next three chapters, we went on a journey to make our acoustic biosensors, the main topic of this thesis. We started with our first-of-any-kind, proof-of-concept biosensors sensing enzyme activity, and for the first time, we showed ultrasound imaging of intracellular signals in the complex environment of gastrointestinal tract in living animals[122]. After verifying that it is indeed possible to make acoustic biosensors based on GVs, we went ahead and developed a more generalizable design principle — using analyte-dependent conformational changes of GvpC to couple molecular processes to ultrasound contrast. Based on this design, we engineered the ultrasonic reporters of calcium, or URoCs. Now, we can image the dynamics of calcium signaling inside mammalian cells. We demonstrated noninvasive imaging of receptor-specific calcium dynamics deep in the mouse brain through the intact skull[179]. With the high-throughput engineering platform introduced in the last chapter, this ultrasound calcium imaging probably is just the beginning of all the technologies we have been dreaming for: whole-brain neuroimaging, tracking the activity of immune cells, monitoring the function of islet transplants and a long list that I cannot fit here. Before we start to imagine more about the future of acoustic biosensors and biomolecular ultrasound, here is the conclusion of this thesis:

1. Ultrasound is awesome.

2. Biomolecular engineering is making ultrasound more and more powerful.

3. We can build acoustic biosensors by engineering GvpC to have analyte-dependent conformational changes that lead to changes in GVs' acoustic properties.

4. We can take advantage of ultrasound instrumentation and synthetic biology to accelerate the development of GV-based imaging agents.

5. It is becoming possible to visualize dynamic molecular processes noninvasively with acoustic biosensors.

This thesis marks the starting point of acoustic biosensors. Personally, I believe this also marks the starting point of a revolution in molecular imaging, leading to a future where we can noninvasively "see" the activity of a diverse selection of signaling molecules at the scale of the whole organisms. It would result in many new discoveries in biology, and I am looking forward to that future. In the next two sections, I will share with you how I think we may get there.

5.2 Future Directions for Acoustic Biosensors

Let us start with the more specific topic about acoustic biosensors. In this thesis, we established the initial concepts, the proof-of-concept results and some exciting yet preliminary *in vivo* applications. Moving forward, we should first improve our engineering platform for higher efficiency and better context. More specifically, all of the current sensors were manually constructed and characterized *in vitro* in purified protein format. The disadvantages are: (1) the throughput is very limited, capped around 12 variants per week and (2) the purified format might not reflect what is happening inside the cells where the sensors are supposed to work. Both are the well-learned lessons from our fluorescent counterparts: throughput and screening context matter. Thus, the first step towards better acoustic biosensors would be to develop a better screening platform that would enable intracellular characterization of the sensors at a higher throughput. For the throughput, we

can further improve our acoustic plate reader, from sequential characterization to parallel characterization like some fluorescence plate reader — having one ultrasound probe for each well of the 96-well plates and plates with even more wells. We can also integrate stimulation system into this acoustic plate readers to enable sensor screening and characterization. For example, we can add an automated injector for chemical stimulation or LEDs for light-based stimulation. Meanwhile, we can establish cell lines that can enable external control of the specific analytes inside the cells. For calcium sensors, as an example, we can add optogenetic components[212] into the cells so that we can induce calcium influx with light in a temporally controlled manner. Combining both, we would be able to characterize the performance of our calcium sensors in mammalian cells and do so in a high-throughput manner. This is just one example; we can implement this idea of high-throughput cell-based screening in every sensor we designed and the ones to be designed. Certainly, with this system, we can easily add in the directed evolution or other strategies that require a larger library size.

With the right tools, the current sensors, both the protease and calcium sensors, can be dramatically improved in different aspects. For the protease sensors, it can be expanded to other more physiologically relevant proteases for real-world applications, such as matrix metalloproteinases (MMPs) for tumor diagnostics[213] and caspases for apoptosis[214]. For the calcium sensors, they are good enough for many applications, but the limiting factor is to express them robustly in the host cells, so the top priority to improve the sensors is, in a way, irrelevant to the sensors themselves, which is the topic we will cover in the later section. However, it would be more ideal to have a faster calcium sensor, with the kinetics down to hundreds or even tens of milliseconds, for neuroscience applications — with the ultimate goal to image the neural activity across the entire brain in rodents or even non-human primates. Besides existing sensors, we should develop sensors with more diverse targets: kinases, ATP, neuromodulators, and even membrane potential. The good news is that we can always learn from the fluorescent biosensors about what analytes are important and what sensing domains to use. The not-so-great news is that engineering GvpC to change conformation based on these analytes might not be as easy as engineering a circular-permuted GFP. Thus, I believe we should gain better understanding about GvpC-GV

interaction. What are the docking sites? Could we predict the affinity and kinetics of certain GvpC molecules attaching to GVs? What conformations of GvpC would lead to a stiff shell and what would lead to a soft, buckling shell? All these questions should be answerable with the GV structure solved and many computational tools available, but we should answer them before diving too much into making new sensors.

Going beyond the design principles mentioned in thesis, we may want to figure out how to reduce the number of sensing units per GVs. Currently, every single GvpC plays a part in response to the analytes but that might not be ideal. There can be a significant buffering effect when sensing analytes with extremely low concentration and in some more extreme cases, we might not be able to have all the GvpC accessible to the analytes (e.g., for trans-membrane signals like voltage). Potentially, we need a "chain reaction" or an amplifier that can change GV's acoustic properties based on the changes in fewer proteins, or ultimately just one protein for a single GV. However, it is not clear to me how one can achieve this at this moment, but I believe it is a direction worth exploring.

At last, is GvpC the only solution? Not necessarily. My colleagues and I have been entertaining about ideas to sense through GV clustering, but we all agreed that it is very challenging to control GVs' clustering inside the cells because they might not diffuse through the cytoskeleton very well with their micron-scale sizes. Can GvpA, or the shell itself change its mechanical properties upon stimulation? There is no physical limitation that prohibit this strategy from working, but again, it is not clear to me how it would be implemented, especially with our very limited understanding about the GV assembly and GvpA mechanical properties. Thus, similar to GvpC, I would propose that the future of GV-based acoustic biosensors relies on the fundamental understanding about GV proteins and even though there might be concerns about imminent applications and potential challenges in getting funding, we should put this as one of priorities.

5.3 Future Directions for Biomolecular Ultrasound Imaging

In the last section, I want to imagine about the future of biomolecular ultrasound imaging. Hopefully with the development of more and better GV-based sensors, we will start to gain

more understanding about biology through better observation. However, there are some non-sensor-related challenges that we need to address, including (1) better GV expression across cell types, (2) smaller gene cluster for better delivery, (3) multiplexing capabilities.

First, expressing GVs in the host cells can be difficult and can be disturbing to the host cells. We need to understand the effect of GV expression on cells and tune the expression level of individual proteins correspondingly instead of expression them at equal amount, which can lead to cellular stress. We also need to better regulate the physical footprint of GVs inside the cells to avoid GVs blocking essential diffusion processes in the cytosol. Second, the size of the smallest gene cluster that encode GVs is around 5000 base pairs, which cannot be packaged into a single lentivirus or adeno-associated virus (AAV). It is thereby challenging to deliver GV genes *in vivo*. As we have tried most of the engineering approaches (e.g., deleting individual ones) to shrink the gene cluster, to further reduce the genetic cargo, we may have to make changes to these proteins or engineer novel chaperones to replace them. Third, frequency-dependence is one of the most promising ways to enable multiplexing in ultrasound, but the resonant frequencies of GVs are in the GHz range[192] and it require engineering of the shell mechanics or GVs' ensemble behavior to reduce the resonant frequencies to the medically-relevant range. To me, all these challenges and potential solutions all point to one single direction: we should understand GVs better for future engineering. Once we understand what individual protein is doing and how they work together to assemble the GVs, we would be better prepared to address the aforementioned challenges and set out for a future where we can use ultrasound to noninvasively image everything that the fluorescence microscope can image today, or even more than that.

Is GV the only option for biomolecular ultrasound imaging? I don't think so. It could be a more promising direction to engineer other ultrasound-interacting proteins from the bottom up to fulfill our needs as well: well-understood assembly process, controlled cellular footprint, small genetic cargo, and multiplexing capability. For example, we can explore to generate a single protein bubble with a defined size inside each cell, which can potentially give us a narrow frequency response for multiplexing. However, as we learned from many cases (e.g.,

GFP, channelrhodopsin, and gene editing proteins), nature is better at making proteins than us, at least at this moment. It is not trivial to design never-existing proteins that have similar or better functions than naturally-evolved ones. The silver lining is that the recent advances of computational methods and other protein engineering tools may have started to change this. Although we have not managed to design a totally different but much better fluorescent proteins or optogenetic elements, I believe we will get there eventually. As that day approaches, it is worth thinking about how to build ultrasound-interacting proteins based on our own understanding about sound physics.

Regardless of which route one may take, I genuinely believe that there would be a future when we can understand biology at the molecular scale across the entire organism with the help of ultrasound. I am very excited to have my thesis as the part of this journey and thank you so much for reading it to the end. Let us catch up in that future we are working towards!

BIBLIOGRAPHY

1. Wollman, A. J. M., Nudd, R., Hedlund, E. G. & Leake, M. C. From animaculum to single molecules: 300 years of the light microscope. *Open Biology* **5**, 150019 (2015).
2. Weissleder, R. & Mahmood, U. Molecular imaging. *Radiology* **219**, 316–333 (2001).
3. Fei, X. & Gu, Y. Progress in modifications and applications of fluorescent dye probe. *Progress in Natural Science* **19**, 1–7 (2009).
4. Zhou, X., Mehta, S. & Zhang, J. Genetically encodable fluorescent and bioluminescent biosensors light up signaling networks. *Trends in Biochemical Sciences* **45**, 889–905 (2020).
5. Elliott, A. D. Confocal microscopy: principles and modern practices. *Current Protocols in Cytometry* **92**, e68 (2020).
6. Hoover, E. E. & Squier, J. A. Advances in multiphoton microscopy technology. *Nature Photonics* **7**, 93–101 (2013).
7. Entenberg, D., Oktay, M. H. & Condeelis, J. S. Intravital imaging to study cancer progression and metastasis. *Nature Reviews Cancer* **23**, 25–42 (2023).
8. Taroni, P., Pifferi, A., Torricelli, A., Comelli, D. & Cubeddu, R. In vivo absorption and scattering spectroscopy of biological tissues. *Photochemical & Photobiological Sciences* **2**, 124–129 (2003).
9. Mezzanotte, L., Root, M. van 't, Karatas, H., Goun, E. A. & Löwik, C. W. G. M. In vivo molecular bioluminescence imaging: new tools and applications. *Trends in Biotechnology* **35**, 640–652 (2017).
10. Unterrainer, M. *et al.* Recent advances of pet imaging in clinical radiation oncology. *Radiation Oncology* **15**, 88 (2020).
11. Moses, W. W. Fundamental limits of spatial resolution in pet. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **648**, S236–S240 (2011).
12. Wang, G. High temporal-resolution dynamic pet image reconstruction using a new spatiotemporal kernel method. *IEEE Transactions on Medical Imaging* **38**, 664–674 (2019).
13. Ogawa, S., Lee, T. M., Kay, A. R. & Tank, D. W. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences* **87**, 9868–9872 (1990).
14. Kim, S.-G. & Ogawa, S. Biophysical and physiological origins of blood oxygenation level-dependent fmri signals. *Journal of Cerebral Blood Flow & Metabolism* **32**, 1188–1206 (2012).
15. Wei, H., Frey, A. M. & Jasanoff, A. Molecular fmri of neurochemical signaling. *Journal of Neuroscience Methods* **364**, 109372 (2021).
16. Wang, L. V. & Yao, J. A practical guide to photoacoustic tomography in the life sciences. *Nature Methods* **13**, 627–638 (2016).
17. Yao, J. & Wang, L. V. Recent progress in photoacoustic molecular imaging. *Current Opinion in Chemical Biology* **45**, 104–112 (2018).

18. Li, M. *et al.* Sound out the deep colors: photoacoustic molecular imaging at new depths. *Molecular Imaging* **19**, 1536012120981518 (2020).
19. Lin, L. & Wang, L. V. The emerging role of photoacoustic imaging in clinical oncology. *Nature Reviews Clinical Oncology* **19**, 365–384 (2022).
20. Maresca, D. *et al.* Biomolecular ultrasound and sonogenetics. *Annual Review of Chemical and Biomolecular Engineering* **9**, 229–252 (2018).
21. Shiina, T. *et al.* WFUMB guidelines and recommendations for clinical use of ultrasound elastography: part 1: basic principles and terminology. *Ultrasound in Medicine & Biology* **41**, 1126–1147 (2015).
22. Rabut, C. *et al.* Ultrasound technologies for imaging and modulating neural activity. *Neuron* **108**, 93–110 (2020).
23. Frinking, P., Segers, T., Luan, Y. & Tranquart, F. Three decades of ultrasound contrast agents: a review of the past, present and future improvements. *Ultrasound in Medicine & Biology* **46**, 892–908 (2020).
24. Zlitni, A. & Gambhir, S. S. Molecular imaging agents for ultrasound. *Current Opinion in Chemical Biology* **45**, 113–120 (2018).
25. Averkiou, M., Powers, J., Skyba, D., Bruce, M. & Jensen, S. Ultrasound contrast imaging research. *Ultrasound Quarterly* **19**, 27 (2003).
26. Smeenge, M. *et al.* First-in-human ultrasound molecular imaging with a vegfr2-specific ultrasound molecular contrast agent (br55) in prostate cancer: a safety and feasibility pilot study. *Investigative Radiology* **52**, 419 (2017).
27. Pfeifer, F. Distribution, formation and regulation of gas vesicles. *Nature Reviews Microbiology* **10**, 705–715 (2012).
28. Walsby, A. E. Gas vesicles. *Microbiological Reviews* **58**, 94–144 (1994).
29. Dutka, P. *et al.* Measuring gas vesicle dimensions by electron microscopy. *Protein Science* **30**, 1081–1086 (2021).
30. Dutka, P. *et al. Structure of Anabaena Flos-Aquae Gas Vesicles Revealed by Cryo-ET*. http://biorxiv.org/lookup/doi/10.1101/2022.06.21.496981 (2022) doi:10.1101/2022.06.21.496981.
31. Huber, S. T., Terwiel, D., Evers, W. H., Maresca, D. & Jakobi, A. J. Cryo-em structure of gas vesicles for buoyancy-controlled motility. *Cell* **186**, 975-986.e13 (2023).
32. Shapiro, M. G. *et al.* Biogenic gas nanostructures as ultrasonic molecular reporters. *Nature Nanotechnology* **9**, 311–316 (2014).
33. Maresca, D. *et al.* Nonlinear ultrasound imaging of nanoscale acoustic biomolecules. *Applied Physics Letters* **110**, 073704 (2017).
34. Maresca, D., Sawyer, D. P., Renaud, G., Lee-Gosselin, A. & Shapiro, M. G. Nonlinear x-wave ultrasound imaging of acoustic biomolecules. *Physical Review X* **8**, 041002 (2018).
35. Sawyer, D. P. *et al.* Ultrasensitive ultrasound imaging of gene expression with signal unmixing. *Nature Methods* **18**, 945–952 (2021).
36. Maresca, D. *et al.* Acoustic biomolecules enhance hemodynamic functional ultrasound imaging of neural activity. *NeuroImage* **209**, 116467 (2020).
37. Lakshmanan, A. *et al.* Molecular engineering of acoustic protein nanostructures. *ACS Nano* **10**, 7314–7322 (2016).

38. Ling, B. *et al.* Truly tiny acoustic biomolecules for ultrasound imaging and therapy. *Advanced Materials* **n/a**, 2307106.
39. Farhadi, A., Ho, G. H., Sawyer, D. P., Bourdeau, R. W. & Shapiro, M. G. Ultrasound imaging of gene expression in mammalian cells. *Science* **365**, 1469 (2019).
40. Hurt, R. C. *et al.* Genomically mined acoustic reporter genes for real-time in vivo monitoring of tumors and tumor-homing bacteria. *Nature Biotechnology* 1–13 (2023) doi:10.1038/s41587-022-01581-y.
41. Bourdeau, R. W. *et al.* Acoustic reporter genes for noninvasive imaging of microorganisms in mammalian hosts. *Nature* **553**, 86–90 (2018).
42. Hurt, R. C. *et al.* Directed evolution of acoustic reporter genes using high-throughput acoustic screening. 2024.03.30.587094 Preprint at https://doi.org/10.1101/2024.03.30.587094 (2024).
43. Cherin, E. *et al.* Acoustic behavior of halobacterium salinarum gas vesicles in the high-frequency range: experiments and modeling. *Ultrasound in Medicine & Biology* **43**, 1016–1030 (2017).
44. Dutka, P. *et al.* Structure of anabaena flos-aquae gas vesicles revealed by cryo-et. *Structure* **31**, 518-528.e6 (2023).
45. Lin, M. Z. & Schnitzer, M. J. Genetically encoded indicators of neuronal activity. *Nature Neuroscience* **19**, 1142–1153 (2016).
46. Palmer, A. E., Qin, Y., Park, J. G. & McCombs, J. E. Design and application of genetically encoded biosensors. *Trends in Biotechnology* **29**, 144–152 (2011).
47. Rodriguez, E. A. *et al.* The growing and glowing toolbox of fluorescent and photoactive proteins. *Trends in Biochemical Sciences* **42**, 111–129 (2017).
48. Miyawaki, A. & Niino, Y. Molecular spies for bioimaging—fluorescent protein-based probes. *Molecular Cell* **58**, 632–643 (2015).
49. Piraner, D. I. *et al.* Going deeper: biomolecular tools for acoustic and magnetic imaging and control of cellular function. *Biochemistry* **56**, 5202–5209 (2017).
50. Farhadi, A. *et al.* Recombinantly expressed gas vesicles as nanoscale contrast agents for ultrasound and hyperpolarized mri. *AIChE Journal* **64**, 2927–2933 (2018).
51. Hayes, P. K., Buchholz, B. & Walsby, A. E. Gas vesicles are strengthened by the outer-surface protein, gvpc. *Archives of Microbiology* **157**, 229–234 (1992).
52. López-Otín, C. & Bond, J. S. Proteases: multifunctional enzymes in life and disease. *Journal of Biological Chemistry* **283**, 30433–30437 (2008).
53. Drag, M. & Salvesen, G. S. Emerging principles in protease-based drug discovery. *Nature Reviews Drug Discovery* **9**, 690–701 (2010).
54. Sauer, R. T. & Baker, T. A. AAA+ proteases: atp-fueled machines of protein destruction. *Annual Review of Biochemistry* **80**, 587–612 (2011).
55. Turk, B., Turk, D. & Turk, V. Protease signalling: the cutting edge. *The EMBO Journal* **31**, 1630–1643 (2012).
56. Stein, V. & Alexandrov, K. Protease-based synthetic sensing and signal amplification. *Proceedings of the National Academy of Sciences* **111**, 15934–15939 (2014).
57. Fernandez-Rodriguez, J. & Voigt, C. A. Post-translational control of genetic circuits using potyvirus proteases. *Nucleic Acids Research* **44**, 6493–6502 (2016).
58. Gao, X. J., Chong, L. S., Kim, M. S. & Elowitz, M. B. Programmable protein circuits in living cells. *Science* **361**, 1252–1258 (2018).

59. Mitra, R. D., Silva, C. M. & Youvan, D. C. Fluorescence resonance energy transfer between blue-emitting and red-shifted excitation derivatives of the green fluorescent protein. *Gene* **173**, 13–17 (1996).
60. Heim, R. & Tsien, R. Y. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology* **6**, 178–182 (1996).
61. Ong, I. L. H. & Yang, K.-L. Recent developments in protease activity assays and sensors. *Analyst* **142**, 1867–1881 (2017).
62. Phan, J. *et al.* Structural basis for the substrate specificity of tobacco etch virus protease. *Journal of Biological Chemistry* **277**, 50564–50572 (2002).
63. Parks, T. D., Leuther, K. K., Howard, E. D., Johnston, S. A. & Dougherty, W. G. Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Analytical Biochemistry* **216**, 413–417 (1994).
64. Lakshmanan, A. *et al.* Preparation of biogenic gas vesicle nanostructures for use as contrast agents for ultrasound and mri. *Nature Protocols* **12**, 2050–2080 (2017).
65. Goll, D. E., Thompson, V. F., Li, H., Wei, W. & Cong, J. The calpain system. *Physiological Reviews* **83**, 731–801 (2003).
66. Ono, Y. & Sorimachi, H. Calpains — an elaborate proteolytic system. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1824**, 224–236 (2012).
67. Ono, Y., Saido, T. C. & Sorimachi, H. Calpain research for drug discovery: challenges and potential. *Nature Reviews Drug Discovery* **15**, 854–876 (2016).
68. Suzuki, S. *et al.* Development of an artificial calcium-dependent transcription factor to detect sustained intracellular calcium elevation. *ACS Synthetic Biology* **3**, 717–722 (2014).
69. Sauer, R. T. *et al.* Sculpting the proteome with aaa+ proteases and disassembly machines. *Cell* **119**, 9–18 (2004).
70. Baker, T. A. & Sauer, R. T. ClpXP, an atp-powered unfolding and protein-degradation machine. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1823**, 15–28 (2012).
71. Sonnenborn, U. & Schulze, J. The non-pathogenic escherichia coli strain nissle 1917 – features of a versatile probiotic. *Microbial Ecology in Health and Disease* **21**, 122–158 (2009).
72. Danino, T. *et al.* Programmable probiotics for detection of cancer in urine. *Science Translational Medicine* **7**, 289ra84-289ra84 (2015).
73. Blum-Oehler, G. *et al.* Development of strain-specific pcr reactions for the detection of the probiotic *escherichia coli* strain nissle 1917 in fecal samples. *Research in Microbiology* **154**, 59–66 (2003).
74. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
75. Khalil, A. S. & Collins, J. J. Synthetic biology: applications come of age. *Nature Reviews Genetics* **11**, 367–379 (2010).
76. Tigges, M., Marquez-Lago, T. T., Stelling, J. & Fussenegger, M. A tunable synthetic mammalian oscillator. *Nature* **457**, 309–312 (2009).

77. Mark Welch, J. L., Hasegawa, Y., McNulty, N. P., Gordon, J. I. & Borisy, G. G. Spatial organization of a model 15-member human gut microbiota established in gnotobiotic mice. *Proceedings of the National Academy of Sciences* **114**, E9105–E9114 (2017).
78. Geva-Zatorsky, N. *et al.* In vivo imaging and tracking of host–microbiota interactions via metabolic labeling of gut anaerobic bacteria. *Nature Medicine* **21**, 1091–1100 (2015).
79. Foucault, M.-L., Thomas, L., Goussard, S., Branchini, B. R. & Grillot-Courvalin, C. In vivo bioluminescence imaging for the study of intestinal colonization by escherichia coli in mice. *Applied and Environmental Microbiology* **76**, 264–274 (2010).
80. Round, J. L. & Mazmanian, S. K. The gut microbiota shapes intestinal immune responses during health and disease. *Nature Reviews Immunology* **9**, 313–323 (2009).
81. Derrien, M. & van Hylckama Vlieg, J. E. T. Fate, activity, and impact of ingested bacteria within the human gut microbiota. *Trends in Microbiology* **23**, 354–366 (2015).
82. Steidler, L. *et al.* Treatment of murine colitis by lactococcus lactis secreting interleukin-10. *Science* **289**, 1352–1355 (2000).
83. Daniel, C., Roussel, Y., Kleerebezem, M. & Pot, B. Recombinant lactic acid bacteria as mucosal biotherapeutic agents. *Trends in Biotechnology* **29**, 499–508 (2011).
84. Muradali, D. & Goldberg, D. R. US of gastrointestinal tract disease. *RadioGraphics* **35**, 50–68 (2015).
85. Machtaler, S., Knieling, F., Luong, R., Tian, L. & Willmann, J. K. Assessment of inflammation in an acute on chronic model of inflammatory bowel disease with ultrasound molecular imaging. *Theranostics* **5**, 1175–1186 (2015).
86. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in escherichia coli k-12 using pcr products. *Proceedings of the National Academy of Sciences* **97**, 6640–6645 (2000).
87. St-Pierre, F. *et al.* One-step cloning and chromosomal integration of dna. *ACS Synthetic Biology* **2**, 537–541 (2013).
88. Raran-Kurussi, S., Cherry, S., Zhang, D. & Waugh, D. S. Removal of affinity tags with tev protease. in *Heterologous Gene Expression in E.coli: Methods and Protocols* (ed. Burgess-Brown, N. A.) 221–230 (Springer, New York, NY, 2017). doi:10.1007/978-1-4939-6887-9_14.
89. Sequeira, A. F. *et al.* Gene design, fusion technology and tev cleavage conditions influence the purification of oxidized disulphide-rich venom peptides in escherichia coli. *Microbial Cell Factories* **16**, 4 (2017).
90. Sun, Z., Kim, J., Singhal, V. & Murray, R. M. *Protein Degradation in a TX-TL Cell-Free Expression System Using ClpXP Protease*. http://biorxiv.org/lookup/doi/10.1101/019695 (2015) doi:10.1101/019695.
91. Martin, A., Baker, T. A. & Sauer, R. T. Rebuilt aaa + motors reveal operating principles for atp-fuelled machines. *Nature* **437**, 1115–1120 (2005).
92. Wycuff, D. R. & Matthews, K. S. Generation of an arac-*ara*bad promoter-regulated t7 expression system. *Analytical Biochemistry* **277**, 67–73 (2000).
93. Santella, L., Lim, D. & Moccia, F. Calcium and fertilization: the beginning of life. *Trends in Biochemical Sciences* **29**, 400–408 (2004).

94. Grienberger, C. & Konnerth, A. Imaging calcium in neurons. *Neuron* **73**, 862–885 (2012).
95. Spitzer, N. C., Gu, X. & Olson, E. Action potentials, calcium transients and the control of differentiation of excitable cells. *Current Opinion in Neurobiology* **4**, 70–77 (1994).
96. Rorsman, P., Braun, M. & Zhang, Q. Regulation of calcium in pancreatic α- and β-cells in health and disease. *Cell Calcium* **51**, 300–308 (2012).
97. Campbell, J. E. & Newgard, C. B. Mechanisms controlling pancreatic islet cell function in insulin secretion. *Nature Reviews Molecular Cell Biology* **22**, 142–158 (2021).
98. Trebak, M. & Kinet, J.-P. Calcium signalling in t cells. *Nature Reviews Immunology* **19**, 154–169 (2019).
99. Vig, M. & Kinet, J.-P. Calcium signaling in immune cells. *Nature Immunology* **10**, 21–27 (2009).
100. Thestrup, T. *et al.* Optimized ratiometric calcium sensors for functional in vivo imaging of neurons and t lymphocytes. *Nature Methods* **11**, 175–182 (2014).
101. Halle, S. *et al.* In vivo killing capacity of cytotoxic t cells is limited and involves dynamic interactions and t cell cooperativity. *Immunity* **44**, 233–245 (2016).
102. Weigelin, B. *et al.* Cytotoxic t cells are able to efficiently eliminate cancer cells by additive cytotoxicity. *Nature Communications* **12**, 5217 (2021).
103. Berridge, M. J., Lipp, P. & Bootman, M. D. The versatility and universality of calcium signalling. *Nature Reviews Molecular Cell Biology* **1**, 11–21 (2000).
104. Lock, J. T., Parker, I. & Smith, I. F. A comparison of fluorescent ca2+ indicators for imaging local ca2+ signals in cultured cells. *Cell Calcium* **58**, 638–648 (2015).
105. Dana, H. *et al.* Sensitive red protein calcium indicators for imaging neural activity. *eLife* **5**, e12727 (2016).
106. Qian, Y. *et al.* A genetically encoded near-infrared fluorescent calcium ion indicator. *Nature Methods* **16**, 171–174 (2019).
107. Zhang, Y. *et al.* Fast and sensitive gcamp calcium indicators for imaging neural populations. *Nature* **615**, 884–891 (2023).
108. Su, Y. *et al.* An optimized bioluminescent substrate for non-invasive imaging in the brain. *Nature Chemical Biology* **19**, 731–739 (2023).
109. Hou, X. H. *et al.* Central control circuit for context-dependent micturition. *Cell* **167**, 73-86.e12 (2016).
110. Wang, Y. *et al.* GPCR-induced calcium transients trigger nuclear actin assembly for chromatin dynamics. *Nature Communications* **10**, 5271 (2019).
111. Hamilton, A. *et al.* Adrenaline stimulates glucagon secretion by tpc2-dependent ca2+ mobilization from acidic stores in pancreatic α-cells. *Diabetes* **67**, 1128–1139 (2018).
112. Ortiz-Muñoz, G. *et al.* In situ tumour arrays reveal early environmental control of cancer immunity. *Nature* **618**, 827–833 (2023).
113. Hausmann, D. *et al.* Autonomous rhythmic activity in glioma networks drives brain tumour growth. *Nature* **613**, 179–186 (2023).
114. Girven, K. S. & Sparta, D. R. Probing deep brain circuitry: new advances in in vivo calcium measurement strategies. *ACS Chemical Neuroscience* **8**, 243–251 (2017).
115. Barandov, A. *et al.* Sensing intracellular calcium ions using a manganese-based mri contrast agent. *Nature Communications* **10**, 897 (2019).

116. Thiabaud, G. D. *et al.* Texaphyrin-based calcium sensor for multimodal imaging. *ACS Sensors* (2023) doi:10.1021/acssensors.3c01387.
117. Zhang, R. *et al.* Multiscale photoacoustic tomography of neural activities with gcamp calcium indicators. *Journal of Biomedical Optics* **27**, 096004 (2022).
118. Gottschalk, S. *et al.* Rapid volumetric optoacoustic imaging of neural dynamics across the mouse brain. *Nature Biomedical Engineering* **3**, 392–401 (2019).
119. Shemetov, A. A. *et al.* A near-infrared genetically encoded calcium indicator for in vivo imaging. *Nature Biotechnology* (2020) doi:10.1038/s41587-020-0710-1.
120. Shapiro, M. G. *et al.* Biogenic gas nanostructures as ultrasonic molecular reporters. *Nature Nanotechnology* **9**, 311–316 (2014).
121. Farhadi, A., Ho, G. H., Sawyer, D. P., Bourdeau, R. W. & Shapiro, M. G. Ultrasound imaging of gene expression in mammalian cells. 7 (2019).
122. Lakshmanan, A. *et al.* Acoustic biosensors for ultrasound imaging of enzyme activity. *Nature Chemical Biology* **16**, 988–996 (2020).
123. Heiles, B., Terwiel, D. & Maresca, D. The advent of biomolecular ultrasound imaging. *Neuroscience* **474**, 122–133 (2021).
124. Nakai, J., Ohkura, M. & Imoto, K. A high signal-to-noise ca2+ probe composed of a single green fluorescent protein. *Nature Biotechnology* **19**, 137–141 (2001).
125. Clapperton, J. A., Martin, S. R., Smerdon, S. J., Gamblin, S. J. & Bayley, P. M. Structure of the complex of calmodulin with the target sequence of calmodulin-dependent protein kinase i: studies of the kinase activation mechanism. *Biochemistry* **41**, 14669–14679 (2002).
126. Chen, T.-W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
127. Helassa, N. *et al.* Fast-response calmodulin-based fluorescent indicators reveal rapid intracellular calcium dynamics. *Scientific Reports* **5**, 15978 (2015).
128. Casey, J. R., Grinstein, S. & Orlowski, J. Sensors and regulators of intracellular ph. *Nature Reviews Molecular Cell Biology* **11**, 50–61 (2010).
129. van der Linden, F. H. *et al. A Turquoise Fluorescence Lifetime-Based Biosensor for Quantitative Imaging of Intracellular Calcium*. http://biorxiv.org/lookup/doi/10.1101/2021.06.21.449214 (2021) doi:10.1101/2021.06.21.449214.
130. Gleichmann, M. & Mattson, M. P. Neuronal calcium homeostasis and dysregulation. *Antioxidants & Redox Signaling* **14**, 1261–1273 (2011).
131. Huston, J. S. *et al.* Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain fv analogue produced in escherichia coli. *Proceedings of the National Academy of Sciences* **85**, 5879–5883 (1988).
132. Inoue, M. Genetically encoded calcium indicators to probe complex brain circuit dynamics in vivo. *Neuroscience Research* **169**, 2–8 (2021).
133. Park, H. Y. *et al.* Conformational changes of calmodulin upon ca2+ binding studied with a microfluidic mixer. *Proceedings of the National Academy of Sciences* **105**, 542–547 (2008).
134. Faas, G. C., Raghavachari, S., Lisman, J. E. & Mody, I. Calmodulin as a direct detector of ca2+ signals. *Nature Neuroscience* **14**, 301–304 (2011).

135. Wang, Q. *et al.* Protein recognition and selection through conformational and mutually induced fit. *Proceedings of the National Academy of Sciences* **110**, 20545–20550 (2013).
136. Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved gcamp calcium indicators. *Nature Methods* **6**, 875–881 (2009).
137. Liu, C. & Hermann, T. E. Characterization of ionomycin as a calcium ionophore. *Journal of Biological Chemistry* **253**, 5892–5894 (1978).
138. Ribeiro, C. M. P., McKay, R. R., Hosoki, E., Bird, G. St. J. & Putney, J. W. Effects of elevated cytoplasmic calcium and protein kinase c on endoplasmic reticulum structure and function in hek293 cells. *Cell Calcium* **27**, 175–185 (2000).
139. Kofuji, P. & Araque, A. G-protein-coupled receptors in astrocyte–neuron communication. *Neuroscience* **456**, 71–84 (2021).
140. Lämmermann, T. & Kastenmüller, W. Concepts of gpcr-controlled navigation in the immune system. *Immunological Reviews* **289**, 205–231 (2019).
141. Capote, L. A., Mendez Perez, R. & Lymperopoulos, A. GPCR signaling and cardiac function. *European Journal of Pharmacology* **763**, 143–148 (2015).
142. Mizuno, N. & Itoh, H. Functions and regulatory mechanisms of gq-signaling pathways. *Neurosignals* **17**, 42–54 (2009).
143. Sánchez-Fernández, G. *et al.* Gαq signalling: the new and the old. *Cellular Signalling* **26**, 833–848 (2014).
144. Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S. & Roth, B. L. Evolving the lock to fit the key to create a family of g protein-coupled receptors potently activated by an inert ligand. *Proceedings of the National Academy of Sciences* **104**, 5163–5168 (2007).
145. Nagai, Y. *et al.* Deschloroclozapine, a potent and selective chemogenetic actuator enables rapid neuronal and behavioral modulations in mice and monkeys. *Nature Neuroscience* **23**, 1157–1167 (2020).
146. Nguyen, Q.-T. *et al.* An in vivo biosensor for neurotransmitter release and in situ receptor activity. *Nature Neuroscience* **13**, 127–132 (2010).
147. Muller, A., Joseph, V., Slesinger, P. A. & Kleinfeld, D. Cell-based reporters reveal in vivo dynamics of dopamine and norepinephrine release in murine cortex. *Nature Methods* **11**, 1245–1252 (2014).
148. Xiong, H. *et al.* Probing neuropeptide volume transmission in vivo by simultaneous near-infrared light-triggered release and optical sensing**. *Angewandte Chemie* **134**, e202206122 (2022).
149. Doherty, J. E. *et al.* Hyperactive piggybac gene transfer in human cells and in vivo. *Human Gene Therapy* **23**, 311–320 (2012).
150. Loew, R., Heinz, N., Hampf, M., Bujard, H. & Gossen, M. Improved tet-responsive promoters with minimized background expression. *BMC Biotechnology* **10**, 81 (2010).
151. Rahal, L. *et al.* Ultrafast ultrasound imaging pattern analysis reveals distinctive dynamic brain states and potent sub-network alterations in arthritic animals. *Scientific Reports* **10**, 10485 (2020).
152. Brunner, C. *et al.* Whole-brain functional ultrasound imaging in awake head-fixed mice. *Nature Protocols* **16**, 3547–3571 (2021).
153. Roth, B. L. DREADDs for neuroscientists. *Neuron* **89**, 683–694 (2016).

154. Mace, E. *et al.* Functional ultrasound imaging of the brain. *Nat Methods* **8**, 662–4 (2011).
155. Rabut, C. *et al.* 4D functional ultrasound imaging of whole-brain activity in rodents. *Nature Methods* **16**, 994–997 (2019).
156. Duan, M., Dev, I., Lu, A., You, M. Y. & Shapiro, M. G. *Stoichiometric Expression of Messenger Polycistrons by Eukaryotic Ribosomes (SEMPER) for Compact, Ratio-Tunable Multi-Gene Expression from Single mRNAs*. http://biorxiv.org/lookup/doi/10.1101/2023.05.26.541240 (2023) doi:10.1101/2023.05.26.541240.
157. Wang, C. *et al.* Bioadhesive ultrasound for long-term continuous imaging of diverse organs. *Science* **377**, 517–523 (2022).
158. Hu, H. *et al.* A wearable cardiac ultrasound imager. *Nature* **613**, 667–675 (2023).
159. Errico, C. *et al.* Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging. *Nature* **527**, 499–502 (2015).
160. Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging. *eLife* **5**, e14472 (2016).
161. Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M. & Keller, P. J. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nature Methods* **10**, 413–420 (2013).
162. Skocek, O. *et al.* High-speed volumetric imaging of neuronal activity in freely moving rodents. *Nature Methods* **15**, 429–432 (2018).
163. Gibson, D. G. *et al.* Enzymatic assembly of dna molecules up to several hundred kilobases. *Nature Methods* **6**, 343–345 (2009).
164. Salahshoor, H. *et al.* Geometric effects in gas vesicle buckling under ultrasound. *Biophysical Journal* **121**, 4221–4228 (2022).
165. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology* **152**, 36–51 (2005).
166. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: algorithms for rapid unsupervised cryo-em structure determination. *Nature Methods* **14**, 290–296 (2017).
167. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nature Methods* **19**, 679–682 (2022).
168. Goddard, T. D. *et al.* UCSF chimerax: meeting modern challenges in visualization and analysis. *Protein Science* **27**, 14–25 (2018).
169. Maresca, D. *et al.* Acoustic biomolecules enhance hemodynamic functional ultrasound imaging of neural activity. *NeuroImage* **209**, 116467 (2020).
170. Ling, B. *et al.* Biomolecular ultrasound imaging of phagolysosomal function. *ACS Nano* **14**, 12210–12221 (2020).
171. Demené, C. *et al.* Spatiotemporal clutter filtering of ultrafast ultrasound data highly increases doppler and fultrasound sensitivity. *IEEE Transactions on Medical Imaging* **34**, 2271–2285 (2015).
172. Mace, E. *et al.* Functional ultrasound imaging of the brain: theory and basic principles. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **60**, 492–506 (2013).

173. Bourdeau, R. W. *et al.* Acoustic reporter genes for noninvasive imaging of microorganisms in mammalian hosts. *Nature* **553**, 86–90 (2018).
174. Farhadi, A., Ho, G. H., Sawyer, D. P., Bourdeau, R. W. & Shapiro, M. G. Ultrasound imaging of gene expression in mammalian cells. *Science* **365**, 1469–1475 (2019).
175. Pfeifer, F. Distribution, formation and regulation of gas vesicles. *Nature Reviews Microbiology* **10**, 705–715 (2012).
176. Walsby, A. E. Gas vesicles. *Microbiological Reviews* **58**, 94–144 (1994).
177. Shapiro, M. G. *et al.* Biogenic gas nanostructures as ultrasonic molecular reporters. *Nature Nanotechnology* **9**, 311–316 (2014).
178. Ling, B. *et al.* Biomolecular ultrasound imaging of phagolysosomal function. *ACS Nano* **14**, 12210–12221 (2020).
179. Jin, Z. *et al. Ultrasonic Reporters of Calcium for Deep Tissue Imaging of Cellular Signals*. http://biorxiv.org/lookup/doi/10.1101/2023.11.09.566364 (2023) doi:10.1101/2023.11.09.566364.
180. Yang, G. & Withers, S. G. Ultrahigh-throughput facs-based screening for directed enzyme evolution. *ChemBioChem* **10**, 2704–2715 (2009).
181. Kazlauskas, R. J. & Bornscheuer, U. T. Finding better protein engineering strategies. *Nature Chemical Biology* **5**, 526–529 (2009).
182. Packer, M. S. & Liu, D. R. Methods for the directed evolution of proteins. *Nature Reviews Genetics* **16**, 379–394 (2015).
183. Maresca, D. *et al.* Nonlinear ultrasound imaging of nanoscale acoustic biomolecules. *Applied Physics Letters* **110**, 073704 (2017).
184. Hurt, R. C. *et al.* Genomically mined acoustic reporter genes for real-time in vivo monitoring of tumors and tumor-homing bacteria. *Nature Biotechnology* (2023) doi:10.1038/s41587-022-01581-y.
185. Lakshmanan, A. *et al.* Acoustic biosensors for ultrasound imaging of enzyme activity. *Nature Chemical Biology* **16**, 988–996 (2020).
186. Lu, G. J. *et al.* Genetically encodable contrast agents for optical coherence tomography. *ACS Nano* **14**, 7823–7831 (2020).
187. Li, N. & Cannon, M. C. Gas vesicle genes identified in bacillus megaterium and functional expression in escherichia coli. *Journal of Bacteriology* **180**, 2450–2458 (1998).
188. Huber, S. T., Terwiel, D., Evers, W. H., Maresca, D. & Jakobi, A. J. Cryo-em structure of gas vesicles for buoyancy-controlled motility. *Cell* **186**, 975-986.e13 (2023).
189. Dutka, P. *et al.* Structure of anabaena flos-aquae gas vesicles revealed by cryo-et. *Structure* **31**, 518-528.e6 (2023).
190. Lakshmanan, A. *et al.* Molecular engineering of acoustic protein nanostructures. *ACS Nano* **10**, 7314–7322 (2016).
191. Rabut, C. *et al.* Ultrafast amplitude modulation for molecular and hemodynamic ultrasound imaging. *Applied Physics Letters* **118**, 244102 (2021).
192. Zhang, S. *et al.* The vibration behavior of sub-micrometer gas vesicles in response to acoustic excitation determined via laser doppler vibrometry. *Advanced Functional Materials* **30**, 2000239 (2020).

193. Maresca, D., Sawyer, D. P., Renaud, G., Lee-Gosselin, A. & Shapiro, M. G. Nonlinear x-wave ultrasound imaging of acoustic biomolecules. *Physical Review X* **8**, 041002 (2018).
194. Cherin, E. *et al.* Acoustic behavior of halobacterium salinarum gas vesicles in the high-frequency range: experiments and modeling. *Ultrasound in Medicine and Biology* **43**, 1016–1030 (2017).
195. Maresca, D. *et al.* Biomolecular ultrasound and sonogenetics. *Annual Review of Chemical and Biomolecular Engineering* **9**, 229–252 (2018).
196. Lakshmanan, A. *et al.* Preparation of biogenic gas vesicle nanostructures for use as contrast agents for ultrasound and mri. *Nature Protocols* **12**, 2050–2080 (2017).
197. Segall-Shapiro, T. H., Sontag, E. D. & Voigt, C. A. Engineered promoters enable constant gene expression at any copy number in bacteria. *Nature Biotechnology* **36**, 352–358 (2018).
198. Wells, J. A., Vasser, M. & Powers, D. B. Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites. *Gene* **34**, 315–323 (1985).
199. Farhadi, A. *et al.* Genetically encoded phase contrast agents for digital holographic microscopy. *Nano Letters* **20**, 8127–8134 (2020).
200. Salahshoor, H. *et al.* Geometric effects in gas vesicle buckling under ultrasound. *Biophysical Journal* **121**, 4221–4228 (2022).
201. Iburg, M. *et al. Elucidating the Assembly of Gas Vesicles by Systematic Protein-Protein Interaction Analysis*. http://biorxiv.org/lookup/doi/10.1101/2023.07.23.550228 (2023) doi:10.1101/2023.07.23.550228.
202. Jost, A. & Pfeifer, F. Interaction of the gas vesicle proteins gvpa, gvpc, gvpn, and gvpo of halobacterium salinarum. *Frontiers in Microbiology* **13**, 971917 (2022).
203. Völkner, K., Jost, A. & Pfeifer, F. Accessory gvp proteins form a complex during gas vesicle formation of haloarchaea. *Frontiers in Microbiology* **11**, 610179 (2020).
204. Winter, K., Born, J. & Pfeifer, F. Interaction of haloarchaeal gas vesicle proteins determined by split-gfp. *Frontiers in Microbiology* **9**, 1897 (2018).
205. Moore, S. J. *et al.* EcoFlex: a multifunctional moclo kit for *e. coli* synthetic biology. *ACS Synthetic Biology* **5**, 1059–1069 (2016).
206. Meyer, A. J., Segall-Shapiro, T. H., Glassey, E., Zhang, J. & Voigt, C. A. Escherichia coli “marionette” strains with 12 highly optimized small-molecule sensors. *Nature Chemical Biology* **15**, 196–204 (2019).
207. Mutalik, V. K. *et al.* Precise and reliable gene expression via standard transcription and translation initiation elements. *Nature Methods* **10**, 354–360 (2013).
208. Chen, Y.-J. *et al.* Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nature Methods* **10**, 659–664 (2013).
209. Lin, D. & O’Callaghan, C. A. MetClo: methylase-assisted hierarchical dna assembly using a single type iis restriction enzyme. *Nucleic Acids Research* (2018) doi:10.1093/nar/gky596.
210. Ravikumar, A., Arzumanyan, G. A., Obadi, M. K. A., Javanpour, A. A. & Liu, C. C. Scalable, continuous evolution of genes at mutation rates above genomic error thresholds. *Cell* **175**, 1946-1957.e13 (2018).

211. Wittmann, B. J., Johnston, K. E., Almhjell, P. J. & Arnold, F. H. EvSeq: cost-effective amplicon sequencing of every variant in a protein library. *ACS Synthetic Biology* **11**, 1313–1324 (2022).
212. Adamantidis, A. *et al.* Optogenetics: 10 years after chr2 in neurons—views from the community. *Nature Neuroscience* **18**, 1202–1212 (2015).
213. Kessenbrock, K., Plaks, V. & Werb, Z. Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell* **141**, 52–67 (2010).
214. Lei, Q. *et al.* Biosensors for caspase-3: from chemical methodologies to biomedical applications. *Talanta* **240**, 123198 (2022).