# Single-cell chemical imaging of engineered strains reveals heterogeneity in fatty acid production

Nathan Tague[† 1,2], Haonan Lin[† 1,3], Jean-Baptiste Lugagne[1,2], Deeya Burman[1], Wilson W. Wong[1,2], Ji-Xin Cheng* [1,3,4], Mary J. Dunlop* [1,2]

† Co-first authors
* Co-corresponding authors

[1] Department of Biomedical Engineering, Boston University, Boston, MA, USA
[2] Biological Design Center, Boston University, Boston, MA, USA
[3] Photonics Center, Boston University, Boston, MA, USA
[4] Department of Electrical and Computer Engineering, Boston University, Boston, MA, USA

## Abstract
Engineered microbes can produce valuable chemicals, however production strains often require extensive optimization before they can be used at industrial scales. To quantify production, metabolic engineers typically employ mass spectrometry, which offers excellent chemical specificity. However, with this approach samples are derived from bulk cultures, obscuring potential cell-to-cell differences in production, and precluding selection methods that rely on measurements of a single cell's phenotype, such as directed evolution. Here, we use stimulated Raman scattering (SRS) based chemical imaging to directly visualize free fatty acids in metabolically engineered *Escherichia coli*. We uncover substantial heterogeneity in fatty acid production among and within colonies. We then demonstrate longitudinal SRS imaging, allowing fatty acids to be linked with the cells that generated them. Further, using the hyperspectral images, we develop an approach for compositional analysis that determines the chain length and unsaturation of the fatty acids in living cells. Our results demonstrate that cell-to-cell heterogeneity in production of metabolic targets can be substantial and indicate that controlling this to eliminate low producers represents a critical area for future optimization. Further, because SRS imaging is rapid, chemically specific, and can be acquired longitudinally, it is a promising method for metabolic engineers, allowing for direct imaging of biosynthesis at the single-cell level.

## Introduction
Microbial production of chemicals has the potential to provide a sustainable source of products ranging from fuels to specialty materials[1–4]. A major difficulty holding back the replacement of industrial chemicals with bio-based alternatives is that bioproduction often falls short in terms of conversion metrics that dictate economic feasibility, such as titer, rate, and yield. Over the past two decades, researchers have made great strides in identifying metabolic pathways capable of producing a diverse array of useful chemicals[5]. However, the reality is that extensive engineering and optimization are required for any given chemical to compete as an alternative to those sourced from petroleum.

Fatty acid synthesis is an attractive pathway for metabolic engineering because it offers a biological means to synthesize linear hydrocarbons. Fatty acids and their derivatives are high demand chemicals that can be used as fuels, commodities, and specialty chemicals. Numerous

studies have aimed at increasing the efficiency of fatty acid synthesis pathways as well as controlling the species of fatty acid produced[6–9]. In particular, type II fatty acid synthesis in *E. coli* is a flexible pathway that creates hydrocarbons of variable length and unsaturation. Termination enzymes that interface with this pathway can be used to produce a wide variety of high-value fatty acid derivatives such as alkanes, olefins, and alcohols[10].

Producing chemicals in cells offers many advantages, but presents unique industrial challenges. For example, cell-to-cell variation and genetic mutations can result in production heterogeneity during fermentation that limits overall process efficiency. Single-cell variation can stem from a variety of causes, such as stochasticity in the underlying biological processes[11,12], variations in media environments within cultures[13], or selection pressures against high producing cells causing mutational escape variants[14,15]. However, the frequency and impact of production variation are largely unknown because studying heterogeneity in chemical production strains is technically difficult. Methods that enable quantification are a prerequisite to understanding the root cause and implementing designs that mitigate its effect on overall efficiency.

Current methods to measure production strain performance include mass spectrometry, biosensor-based fluorescent systems, and dyes. Gas or liquid chromatography-mass spectrometery (GC-MS/LC-MS) are the current gold standard for quantification because they provide excellent chemical specificity and sensitivity. However, most mass spectrometry-based techniques are limited in their ability to quantify single cells, which means they can overlook valuable information that is key to predicting population stability during scale-up[16–18]. Single-cell production information obtained at the strain development stage could help to determine the potential for successful scale-up of a given strain. For free fatty acids specifically, derivatization to esters is necessary for GC-MS quantification, which adds a laborious step when screening strains[19]. In addition, this process separates the product from the cells through an organic phase extraction, which rules out directed evolution-based approaches that connect a desired phenotype, such as high production, to genotype contained within a cell.

Biosensor-based fluorescent assays, in contrast, are amenable to high throughput screens and are non-destructive[20]. However, well-characterized biosensors are scarce in comparison to the number of chemicals metabolic engineers can produce and they only provide an indirect measurement of production. Additionally, significant optimization is often necessary to fine tune the concentration responsive range of a biosensor[21–23]. In the case of fatty acid production, lipophilic dyes such as Nile red have been used in strain engineering studies[24], however, these stains lack lipid specificity, are indirect measurements of production, and do not account for product exported from cells.

Given the drawbacks of current screening methods, we sought to develop an alternative approach that captures single-cell production information to better understand heterogeneity, while being amenable to high-throughput assays to compliment metabolic engineering workflows. Stimulated Raman scattering (SRS) offers the unique ability to directly capture chemical information from single living cells in a linear concentration-dependent manner. SRS is a high-speed vibrational spectroscopic imaging method that can directly detect chemical compounds based on intrinsic molecular vibrations[25,26]. With SRS, a Raman signal can be gathered in microseconds and provides molecular fingerprints for each biomolecular component, making it possible to perform label-free chemical imaging of several chemicals in near real-time. The speed and small sample sizes for

SRS imaging offer additional potential benefits that could be exploited, as there is an increasing need to scale down experiments to manage throughput with large libraries of variants[27,28].

Here, we use SRS microscopy to study metabolically engineered strains, imaging free fatty acid production in *E. coli* in the carbon-hydrogen (C-H) spectral window. First, we demonstrate that we can clearly distinguish fatty acid production strains from wild type *E. coli* by deconstructing hyperspectral images into maps of their chemical components. With the ability to measure production at the single-cell level, we examine heterogeneity in fatty acid production strains and uncover both colony-level heterogeneity and substantial cell-to-cell differences in production. Next, we optimize imaging parameters to achieve longitudinal SRS imaging to capture fatty acid production over time in growing cells, demonstrating differences in production within the same strain. Lastly, we develop a hyperspectral analysis technique that generates chain length and unsaturation level predictions, allowing for chemical readouts that are analogous to GC-MS. Together with synthetic biology methods, our system has the potential to answer fundamental questions relating to the production of biosynthetic targets at the single-cell level. Further, because the imaging approach is rapid and does not require engineered biosensors, it has the potential to serve as a widely useful platform to boost the pace of strain engineering for a broad range of metabolites.

## Results

### Spectroscopic SRS imaging of fatty acid production strains

Spectral signals from Raman scattering correspond to vibrational energies of covalent bonds. This allows for direct imaging of chemicals without the need for labels such as fluorescent reporters or dyes. Here, we deploy spectroscopic SRS[29–31] to obtain chemical maps of protein and fatty acids. To achieve this, we chirp two broadband femtosecond laser beams (pump and Stokes) using high-dispersion glass rods, producing linear temporal separation of the frequency components (Fig. 1a, Fig. S1). The beating frequency of the two beams is linearly correlated with the temporal delay between the two laser pulses. Using a two-dimensional galvo scanner, the combined laser beam is moved across the x and y dimensions of the sample to generate an image. This process is then repeated for a range of temporal delays, each of which produces a different wavenumber, ultimately producing a hyperspectral SRS image generated in a frame-by-frame manner. The spectral region surrounding the 2900 $cm^{-1}$ wavenumber is typically referred to as the 'C-H region' and has a strong SRS signal. Biomolecules such as proteins and fatty acids, which contain many C-H bonds, show high Raman signal in this region. The strong signal in the C-H region enables high fidelity SRS imaging with low optical powers that are compatible with live-cell imaging. Thus, this configuration can be used to acquire longitudinal images of live cells, resulting in data across four dimensions: space (x and y), wavenumber, and time. We set out to utilize SRS chemical imaging in the C-H region to measure fatty acid production in metabolically engineered strains of *E. coli*.

Previous metabolic engineering efforts have focused on producing free fatty acids in *E. coli* using the native type II fatty acid synthesis pathway[9,21,32]. Introducing a heterologously expressed acyl-acyl coat protein (ACP) thioesterase can catalyze the formation and pooling of free fatty acids from elongating acyl hydrocarbon chains that would otherwise be incorporated into membrane

phospholipids[33,34] (Fig. 1b). We reasoned that SRS imaging could effectively capture fatty acid in production strains due to the C-H-rich carbon chains present in fatty acids. To test this hypothesis, we studied several production strains that were previously engineered to produce high quantities of free fatty acids (Tables 1 and 2). We first focused on the strain *Ab*TE*, which expresses an acyl-ACP thioesterase from *Acinetobacter baylyi*, carrying G17R/A165R mutations that improve enzymatic activity in *E. coli* [35]. SRS images of *Ab*TE* show increased fatty acid production relative to the wild type strain, as evidenced by differences in both the chemical spectra and visible fatty acid droplets around the cells (Fig. 1c). Using spectral standards, SRS images can be decomposed into their major chemical components to produce chemical maps (Fig. 1d). We used standard spectra from pure protein (Bovine serum albumin, BSA), saturated fatty acids (C10:0 and C16:0), and unsaturated fatty acids (C16:1) to decompose the hyperspectral image (Fig. S2). To achieve this, we used a least absolute shrinkage and selection operator (LASSO) linear unmixing analysis to separate the hyperspectral image into its chemical components (Methods). This results in two dimensional chemical maps for protein and fatty acid components. Protein levels were comparable between wild type and *Ab*TE* strains, with slightly elevated levels in the engineered strain. In contrast, the fatty acid signal in the *Ab*TE* strain was significantly stronger than in wild type. Wild type cells contain membrane phospholipids, however these signals are much weaker than those recorded in the *Ab*TE* strain (Fig. S3).

**Quantification of heterogeneity in fatty acid production strains**

One of the enabling features of chemical imaging applied to metabolic engineering is the ability to understand production at the single-cell level. Previous studies have reported sub-populations within production cultures that are less productive and lead to decreased overall performance of the population in a scaled up bioprocess[24,36]. Single-cell chemical imaging with SRS is uniquely suited to quantifying this phenomenon. However, highly productive strains will secrete end-products, making it difficult to track the source of produced chemicals back to the cells that generated them. Therefore, sampling from liquid culture for imaging does not accurately provide production heterogeneity information. To ensure that free fatty acid production is tracked to the cells responsible for production, we first grew cells on agarose pads such that production could be localized to the region containing the cells.

We first quantified fatty acid production from *E. coli* microcolonies of the wild type and 'TesA-FV50 production strain (Fig. 2a). The 'TesA-FV50 strain expresses a leaderless native thioesterase ('TesA) when induced with isopropylthio-β-galactoside (IPTG; Table 1)[37]. Endogenously, TesA contains a leader sequence that localizes the enzyme to the periplasm; deleting the leader peptide sequence allows for interaction with cytosolic acyl-ACPs, enabling the production of free fatty acids[33] (Fig. S4). Additionally, this strain expresses heterologous *fadR* and *vhb50*, which have been shown to increase free fatty acid production[7,37]. FadR is a transcription factor that regulates many genes in the fatty acid synthesis pathway for increased free fatty acid titer when expressed alongside 'TesA. Vhb50 is a *Vitreoscilla* hemoglobin that further increases fatty acid production by increasing oxygen uptake.

Interestingly, 'TesA-FV50 microcolonies exhibit a high level of colony-to-colony production variation. This intercolony heterogeneity is visible in the fatty acid chemical maps, with strains from the same original source exhibiting high and low producing microcolonies. One possible explanation for these differences in production is variable transcriptional regulation of key

enzymes that are maintained through replication, leading to metabolic bottlenecks[12,38]. Alternatively, the ability to manage toxicity associated with production in the time frame following thioesterase induction may lead to divergent production outcomes[39].

We also examined production heterogeneity in another fatty acid production strain, *Ab*TE*-FV50. Similar to 'TesA-FV50, this strain includes expression of *fadR* and *vhb50* to boost production, however the 'TesA thioesterase is replaced with *Ab*TE*. Strikingly, we observed a very different type of production variation in *Ab*TE*-FV50 (Fig. 2b). Unlike the intercolony heterogeneity in 'TesA-FV50, the *Ab*TE*-FV50 strain has high heterogeneity between cells in a single microcolony. We used the protein channel to segment the image into single cells for analysis (Fig. S5) and quantified single-cell production (Fig. 2c). Our quantification indicates that in this strain a small percentage of cells produce the vast majority of fatty acids. This result is consistent across many fields of view within the microscopy images, suggesting that it is a general feature of this production strain (Fig. S6). It is possible that toxicity related to heterologous thioesterase expression may contribute to the unique production heterogeneity profile. Stochastic gene expression or improper thioesterase protein folding could also play a role in generating the cell-to-cell differences.

**Longitudinal SRS imaging for production dynamics**

Understanding the dynamics of chemical production with single-cell resolution can provide key insights into the emergence of heterogeneity, production bottlenecks, and can guide engineering strategies to maximize metabolic flux. To that end, we sought to adapt the SRS system for longitudinal imaging. While SRS imaging of living cells has been extensively reported[40–42], its application to chemical production over multiple generations of growth has not been demonstrated. We installed an incubator on the microscope stage and grew live cells on agarose pads for at least 16 hours at 31°C. First, we tested whether the routine laser powers we used for endpoint SRS imaging were damaging to live cells (75 mW for 1040 nm Stokes and 15 mW for 800 nm pump at the sample). At the beginning of longitudinal imaging, we captured a bright field transmission image and measured a hyperspectral SRS image in one field of view (Fig. S7 a-b). After 16 hours of incubation, cells that were previously exposed to SRS imaging did not duplicate, nor did they produce significant levels of fatty acids (Fig. S7 c-d). In contrast, cells in a region in the immediate vicinity that had not been exposed to imaging grew into a dense microcolony and produced fatty acid droplets. Although the laser exposure did not induce visible cell damage, the photodamage altered normal cell functions, indicating that these laser powers were too high.

To optimize the imaging conditions to reduce phototoxicity, we performed the same live-cell experiment with lower laser powers. We obtained normal cell growth when we reduced the Stokes power from 75 to 25 mW, while the pump laser at 800 nm was kept as 15 mW (Fig. S8). To illustrate growth and fatty acid production, we measured transmission and SRS images for the same field of view after 3 and 5 hours of incubation, seeing clear evidence of replication even after SRS imaging (Fig. S8). We took a final wide-field image at 6 hours, which showed that cells continued to replicate normally and suggested that these laser power parameters permit growth and are applicable to longitudinal imaging.

With these optimized imaging conditions, we first tracked fatty acid production within the strain 'TesA-FV50. In line with heterogeneity patterns seen in this strain, the production trajectories of

different microcolonies were significantly different (Fig. 3a-b). In one example, fatty acid signals increased significantly in cells by ~12 hours after thioesterase induction (Fig. 3a). After the microcolony reached a high cell density on the agarose pad, we observed significant accumulation of fatty acids. In contrast, a second microcolony of the same strain produced only low levels of fatty acid (Fig. 3b). For comparison, we also tracked the growth and fatty acid production of wild type cells under the same condition (Fig. 3c). Time-lapse wide-field transmission images for the wild type strain show that cells under SRS laser exposure grew well during the entire experiment period (Movie S1). We quantified fatty acid and protein levels of each microcolony. Protein levels in each strain increased at comparable rates (Fig. 3d). Fatty acid levels in the wild type colony increased modestly while the high-producing 'TesA-FV50 microcolony fatty acid levels increased dramatically (Fig. 3e). The low-producing 'TesA-FV50 microcolony produced fatty acids at levels comparable to wild type.

The activity in the high-producing 'TesA-FV50 microcolony is in line with known regulation patterns in *E. coli* fatty acid synthesis. When high cell density is reached in wild type *E. coli*, the pathway is inhibited by a buildup of acyl-ACPs. This mechanism is reported to act through direct inhibitory interactions with key enzymes within the pathway, such as acetyl-CoA carboxylase, FabH, and FabI [43,44]. Additionally, acyl-ACP or acyl-CoA responsive transcription factors, FadR and FabR, respectively, act to regulate transcriptional responses that control fatty acid synthesis[45,46]. In the presence of a cytosolic thioesterase, as in the 'TesA-FV50 strain, this inhibition is released through the conversion of accumulated acyl-ACPs to free fatty acids. However, thioesterase expression is induced starting at t = 0 hr, and significant accumulation of fatty acid does not happen until the microcolony is well established. Even with the 'TesA thioesterase highly expressed, phospholipid metabolism may dominate metabolic flux through the fatty acid synthesis pathway until sufficient density is reached to suppress incorporation of acyl-ACPs into phospholipids. A recent study from Noga *et al.* uncovered a post-translational mechanism that modulates phospholipid biosynthesis through PlsB acyltransferase and ppGpp, which may explain the delay in free fatty acid accumulation[47].

Additionally, we measured the dynamics of the *Ab*TE*-FV50 fatty acid production strain, which produces a variety of medium- and long-chain fatty acids (Fig. S4), with significant heterogeneity in production among cells (Fig. 2b-c). We again observed fatty acids production over time, with similar delays in fatty acid accumulation despite thioesterase induction at t = 0 hr (Fig. S9). In this strain, a few cells within the microcolony produce large amounts of fatty acid. The production dynamics for these few cells are similar to fatty acid production within the 'TesA-FV50 strain, but the remainder of cells remain at low levels of production for the duration of imaging.

**Prediction of enzymatic specificity, chain length distribution, and degree of unsaturation**

SRS imaging uncovers unique information that would otherwise be lost by bulk culture quantification techniques such GC-MS. However, GC-MS offers precise chemical specificity information. For fatty acid quantification, gas chromatography effectively separates fatty acid esters based on chain length and, along with mass/charge spectra, can specifically read out fatty acid ester chain length and unsaturated bonds. From a metabolic engineering perspective, quantification of a fatty acid production strain's chain length distribution and level of unsaturation are critical. For biofuel purposes, chain length and termination chemistry can be tuned to mimic characteristics of fuel sources such as gasoline, diesel, or jet fuel[48]. Alternatively, medium chain

fatty acids (C8-C12) and their derivatives can be sources of many specialty chemicals[49]. With these end point applications in mind, we sought to extend SRS imaging capabilities to capture the specific profiles of free fatty acid production strains.

Although pure fatty acids of different chain lengths have different spectra in the C-H region, they are too similar to accurately decompose using spectral unmixing with LASSO linear regression analysis. However, we expanded our analysis methodology to take advantage of spectral windows that correspond to $CH_2$ or $CH_3$ bonds, which are present in the 2832-2888 $cm^{-1}$ and 2909-2967 $cm^{-1}$ wavenumber regions, respectively[50]. Since a saturated fatty acid has an increasing number of $CH_2$ bonds as the chain length increases, but the terminal $CH_3$ bond number is constant, we reasoned that the ratio of the $CH_2/CH_3$ spectral windows would scale with chain length (Fig. 4a). Using pure saturated fatty acid standards of variable chain length, we observed a nearly linear relationship between chain length and the ratio of $CH_2/CH_3$ area under the curve (Fig. 4b). We next tested whether we could use this relationship to estimate chain length production profiles.

In *E. coli*, fatty acid biosynthesis is carried out through a multistep, enzymatic Claisen reduction[51]. The enzymatic components of type II fatty acid synthesis in *E. coli* are encoded as separate proteins, creating a pathway in which two carbons are added to an elongating acyl-ACP chain with each cycle (Fig. 4c). The number of cycles around this pathway before the elongating acyl chain is cleaved by an acyl-ACP thioesterase determines the resulting fatty acid chain length. The primary factor driving chain length is the enzymatic specificity of the heterologously expressed thioesterase [6,52]. Researchers have carried out numerous efforts to engineer specificity of acyl-ACP thioesterases in order to create desired chain length profiles[9,35,53–55]. Several thioesterases have been shown previously to produce a range of free fatty acid chain length profiles. Three examples are *Cp*FatB1*, *Ab*TE*, and 'TesA. The *Cp*FatB1* and *Ab*TE* thioesterases originate from *Cuphea palustris* and *A. baylyi*, respectively, and the " * " denotes mutants that were engineered to increase activity in *E. coli*[35,56]. 'TesA is *E. coli*'s native thioesterase, where the " ' " denotes deletion of the leader sequence[33].

To test our ability to estimate chain length distributions using imaging, we examined strains *Cp*FatB1*, *Ab*TE*-FV50, and 'TesA-FV50, which each express a different thioesterase (Table 2). We conducted an experiment in which each of the three strains were grown in liquid culture and thioesterase expression was induced for 24 hours to produce free fatty acids. Samples from each production culture were taken in parallel for GC-MS quantification and SRS hyperspectral imaging. Since these experiments focus on endpoint analysis, we returned to the routine laser powers to increase the spectral resolution of our measurements (75 mW for 1040 nm Stokes, 15 mW for 800 nm pump). As expected, GC-MS results show highly variable chain length distributions depending on the thioesterase expressed (Fig. 4d). *Cp*FatB1* primarily produces octanoic acid (C8:0). *Ab*TE*-FV50 produces a mix of medium- and long-chain saturated fatty acids with myristic acid (C14:0) as the largest component. Lastly, 'TesA-FV50 produces long-chain fatty acids with large contributions from both myristic (C14:0) and palmitic acid (C16:0). Since each production strain has a unique chain length profile, they serve as an ideal group of strains to test our ability to predict chain length distributions with SRS imaging.

To implement chain length prediction, we first decomposed the spectra at each pixel into protein and representative fatty acid chemical maps (C10:0, C16:0, C16:1). The protein and unsaturated

fatty acid maps were then subtracted from the raw SRS image to produce a hyperspectral SRS image of saturated fatty acids (Fig. S10), which can be used to estimate the average chain length at each pixel. We introduced a concentration weighting factor using the SRS spectral ensemble intensity at the same pixel. The SRS predicted chain length distributions closely matches the qualitative features of the GC-MS distributions (Fig. 4d). Importantly, the prediction captures whether the strain produces primarily medium- or long-chain fatty acids, or a mixture of both.

To gauge unsaturation levels, we utilized the presence of the Raman peak at ~3000 $cm^{-1}$, which is unique to the $C{=}CH_2$ bonds in unsaturated fatty acids (Fig. 4e). This peak serves as an identifier of unsaturation level and components from this fatty acid source can be unmixed with LASSO regression. To demonstrate our ability to predict unsaturation level from production strains, we tested the same three strains, which have different ratios of unsaturation to saturation (Fig. S4). The ratio of unsaturation from GC-MS data scales linearly with predicted unsaturated ratios from SRS images (Fig. 4f), giving an indication of the ability of this approach to predict the ratio of unsaturation. With the ability to calculate unsaturation level in addition to chain length distributions from SRS images, we cover many aspects of free fatty acid production that are important for metabolic engineers, bringing SRS hyperspectral imaging closer to a form of optical mass spectrometery.

**Discussion**

Chemical imaging can play a key role in the strain engineering process. Current quantification techniques rely either on methods like GC-MS, which are chemically-specific but where information about individual cells and their dynamics are lost, or on fluorescence reporters or dyes, which are indirect readouts and can be difficult to engineer or limited in their specificity. SRS imaging has the potential to dramatically improve this process by providing key insights into chemical production at the single-cell level. Thus, methods that were previously only accessible with single-cell readouts, such as directed evolution or cell-sorting approaches are in principle possible with SRS imaging. The landscape for strain engineering is expanding rapidly, with systems biology approaches to enzyme engineering and novel technologies for quantifying production offer great promise for improving designs. In this study we focus on fatty acid synthesis, which is an important pathway that can be engineered to produce a diversity of valuable chemicals. Development of this pathway towards near theoretical yields will be important to replace many industrial chemicals with sustainable bio-based alternatives[5].

Here, we examined several free fatty acid production strains of *E. coli* using SRS and demonstrated that hyperspectral imaging allows for image decomposition into major chemical components, with the ability to distinguish cells from their chemical product. Visualizing chemical production at the single-cell level reveals important information that would otherwise be obscured by bulk culture quantification methods. We demonstrate this by examining production heterogeneity among different engineered strains, observing both intra- and inter-colony differences in production within microcolonies. These results provoke fundamental questions about the mechanisms leading to cellular heterogeneity, and also suggest that engineering strategies that eliminate low-producers could improve yields. For example, it may be possible to gradually enhance the overall production levels of a strain of engineered *E. coli* through multiple cycles of growth and dilution, with a step that removes low-producers at the end of each cycle.

Furthermore, we established parameters that allow us to extend SRS imaging for longitudinal studies in live cells. Unlike previous phototoxicity studies focusing on acute responses like membrane blebbing[57,58], we directly observe long-term cell functions including cell replication and free fatty acid synthesis. We envision production tracking at the single-cell level will be valuable for metabolic engineering studies. The tradeoff between growth and chemical production could also be examined with unprecedented detail using this approach. Likewise, recent approaches to increase bioproduction involving dynamic regulation, either through transcriptional feedback circuits or optogenetic regulation, show promise to increase strain efficiency[59,60]. Imaging single-cell production dynamics in these strains could increase our understanding of how feedback systems can be used in the context of metabolic engineering.

By incorporating additional analysis, we also introduce an approach that can estimate chain length distribution and unsaturation degree, increasing the amount of information that can be extracted from SRS hyperspectral images. These advances enable a metabolic engineer to examine fatty acid production strains using SRS imaging while maintaining important chemical specificity data. The ability to gauge enzyme specificity through imaging opens up the possibility of screening mutant enzyme libraries in a high throughput fashion to select for optimal free fatty acid profiles. These methodologies, both experimental and analytical, lay the groundwork for future SRS-guided metabolic engineering efforts.

Understanding the connection between production at the single-cell level and bulk culture output will be required to use SRS imaging as a means of screening strains for their overall production. Imaging fields of view sampled from bulk culture can lead to biased overall titer prediction, especially if the product is not soluble in water. Additionally, SRS has sensitivity limits significantly[61] higher than mass spectrometry and thus requires a product to be produced at sufficient quantities before SRS can be used to guide further engineering. Given this limitation, we envision that SRS studies will be most useful for strain optimization rather than enzyme or pathway discovery.

SRS imaging in different spectral regions, such as the fingerprint region (400-1800 $cm^{-1}$), can be adapted to study strains producing non-fatty acid derived chemicals of interest, such as terpenes, to expand the scope of SRS imaging in metabolic engineering[62]. In addition, because the approach is label-free it does not require biosensors with fluorescent reporter readouts, making it amenable to quantification of production in organisms that are recalcitrant to genetic modification. Moreover, instrumentation advancements can enable SRS guided single-cell screening, such as SRS-based cell sorting, which has been demonstrated recently for cell phenotyping[63]. Much like the utility of fluorescence activated cell sorting in synthetic biology applications, we envision that SRS-based cell sorting can increase the throughput of strain screening and enable directed evolution based on chemical production. This work acts as a jumping off point for SRS imaging in metabolic engineering to aid in the development of more efficient strains for renewable chemical production.

## Methods

### Bacterial strains and plasmids

Plasmid and strain information are listed in Tables 1 and 2. The pBbA5c-'tesA-vhb50-8fadR plasmid was a gift from Dr. Fuzhong Zhang. The BW25113 Δ*fadE* strain is from the Keio collection[64], and we used the FLP recombination protocol from Datsenko and Wanner to cure the *kan*$^R$ cassette from the genome[65]. We used golden gate cloning[66] to create the pBbA5c-vhb50-8fadR plasmid by deleting the coding sequence of '*tesA* from pBbA5c-'tesA-vhb50-8fadR. The pBbA5c-CpFatB1.2-M4-287 plasmid was also constructed using golden gate cloning, with the pBbA5c backbone amplified from the BglBrick plasmid library[67] and the coding sequence of *Cp*FatB1.2-M4-287 derived from Hernández Lozada *et al*.[56] and synthetized by Twist Biosciences (South San Francisco, CA). pSS200 was a gift from Dr. Pamela Peralta-Yahya.

**Growth and induction of fatty acid production strains**

For fatty acid production experiments, pre-cultures were grown overnight in LB media and used to inoculate 3 mL M9 minimal media (M9 salts, 2mM $MgSO_4$, 100 μM $CaCl_2$) with 2% glucose and grown at 37°C with 200 rpm shaking. Antibiotics were added to the media where necessary for plasmid maintenance according to resistances in Table 1 (100 μg/ml for carbenicillin and 25 μg/ml for chloramphenicol). The cultures were allowed to grow until approximately $OD_{600}$ = 0.6 before thioesterase expression was induced with IPTG. Induction levels were 500 μM for 'TesA-FV50 and 50 μM for *Ab*TE*, *Ab*TE*-FV50, and *Cp*FatB1*. For imaging from liquid cultures, cells were grown for 24 hours after IPTG induction and then 3 μL of sample was taken for imaging. Samples from liquid culture were placed on 3% agarose pads containing M9 minimal media and sandwiched between glass coverslips to immobilize the cells for imaging. For longitudinal imaging and production heterogeneity experiments, once cells reached $OD_{600}$ = 0.6 in liquid culture, the sample was placed on a 3% low melting point agarose pad (Promega) containing M9 minimal media with 2% glucose, IPTG as specified above, and appropriate antibiotics for plasmid maintenance, as detailed in Table 1.

For the chain length distribution prediction, cultures were induced with IPTG in liquid cultures for 24 hours. At the 24 hour timepoint, we took 3 μL of sample for imaging and another sample of the culture was taken for GC-MS analysis to allow direct comparison of the same culture. Five fields of view were imaged for each culture.

**Fatty acid derivatization and quantification with GC-MS**

Samples for GC-MS quantification were taken at 24 hours post IPTG induction. 400 μL of vortexed culture was taken for fatty acid extraction and derivatization into fatty acid methyl esters as described by Sarria *et al*.[35] with the following minor modifications: Internal standards of nonanoic acid (C9) and pentadecanoic acid (C15) were added to the 400 μL sample at final concentrations of 88.8 mg/L each and vortexed for 5 sec. The following was then added to the sample for fatty acid extraction and vortexed for 30 sec: 50 μL 10% NaCl, 50 μL glacial acetic acid, and 200 μL ethyl acetate. The sample was then centrifuged at 12,000 g for 10 mins. After centrifugation, 100 μL of the ethyl acetate layer was mixed with 900 μL of a 30:1 mixture of methanol:HCl (12N) in a 2 mL microcentrifuge tube. The solution was vortexed for 30 sec followed by an incubation at 50°C for 60 mins for methyl ester derivatization. Once cooled to room temperature, 500 μL hexanes and 500 μL water were added to the 2 mL microcentrifuge tube, vortexed for 10 sec, and allowed to settle. 250 μL of the hexane layer was mixed with 250 μL ethyl acetate in a GC-MS vial for quantification.

The samples were analyzed with an Agilent 6890N/Agilent 5973 MS detector using a DB-5MS column. The inlet temperature was set to 300°C with flow at 4 mL/min. The oven heating program was initially set to 70°C for 1 min, followed by a ramp to 290°C at 30°C/min, and a final hold at 290°C for 1 min. GLC-20 and GLC-30 FAME standard mixes (Sigma) were tested using this protocol to ensure proper capture of all chain lengths and to gauge retention times. Internal standards were used for quantification, with chain lengths C8-C12 quantified with the nonanoic acid internal standard and C14-C18 quantified with the pentadecanoic internal standard.

**Optical setup**

The SRS setup is driven by an 80 MHz femtosecond laser (Insight Deepsee+, Spectra Physics, USA) with two synchronized outputs. One output is fixed at 1040 nm with a pulse duration of ~150 fs, while the other is tunable from 680 - 1300 nm with ~120 fs pulse width. The 1040 nm beam is used as the Stokes and is modulated by an acousto-optical modulator (522c, Isomet, USA) at 2.5 MHz. The tunable output is set to 798 nm to excite the C-H region and is spatially combined with the Stokes by a dichroic mirror. Six 15 cm SF-57 glass rods are used to linearly chirp the femtosecond pulses to ~ 2 ps. Five of the rods are placed on the common path while one is placed on the Stokes path to parallelize the degree of chirping considering its longer wavelength. A motorized delay stage is used to scan the temporal delay between two pulses to tune the excitation frequency. The combined beams are sent to a pair of two-dimensional galvo scanners (GVSM002, Thorlabs, USA) to perform laser scanning imaging. A 40X oil-immersion objective (RMS40X-PFO, Olympus, Japan) is used to focus the laser onto the sample. Powers on the sample are 15 mW for pump and 75 mW (or 25 mW for longitudinal imaging) for Stokes. A home-built resonant amplifier photodiode collects and amplifies the stimulated Raman loss signal at the modulation frequency. A lock-in amplifier (UHFLI, Zurich Instruments, Switzerland) is used to extract the signal and send it to a data collection card (PCIe-6363, National Instruments, USA). Custom LabView (National Instruments, USA) software is used to synchronize the galvo scan with the delay line scan to obtain a hyperspectral SRS image stack in a frame-by-frame manner.

**Chemical map processing with LASSO**

To obtain concentration maps for chemicals, we perform linear unmixing on the raw hyperspectral image stack. Assuming the number of pure components as $K$ and the dimensions of a hyperspectral image as $N_x, N_y, N_\lambda$, the unmixing model can be written as:

$$D = CS + E, \tag{1}$$

where $D \in \mathbb{R}^{N_x N_y \times N_\lambda}$ is the raw data reshaped as a two dimensional matrix in raster order, $C \in \mathbb{R}^{N_x N_y \times K}$ is the collection of concentration maps, $S \in \mathbb{R}^{K \times N_\lambda}$ contains SRS spectra of all the components, while $E$ is the residual term with error and noise. Given the prior knowledge of spectra for all the pure components, the task is reduced to generating chemical maps $C$ via least square fitting. To avoid crosstalk between spectrally overlapped components, we add a $L1$ norm sparsity constraint by observing that at each spatial position, a few components dominate the contribution. The solution for $C$ is found in a pixel-by-pixel manner by solving for the following optimization problem known as the least absolute shrinkage and selection operator (LASSO):

$$\hat{C}_i = \arg\min_{C_i}\{\frac{1}{2}\ \|D(i,:) - C_i S\|^2 + \beta\|C_i\|_1\}\ , \tag{2}$$

where $i$ represents a specific pixel in the hyperspectral image, $\hat{C}_i$ stands for the estimated concentrations for all components at pixel $i$, and $\beta$ is a hyperparameter controlling the level of $L1$ norm regularization at each pixel.

For each imaging experiment, we measured spectra of pure chemical standards for analysis. Specifically, we input the spectra from the following pure components to perform linear unmixing. We use BSA as the protein standard, palmitic acid (C16:0) and capric acid (C10:0) as representative saturated fatty acids, and palmitoleic acid (C16:1) as an unsaturated fatty acid standard. All standards were sourced from Sigma Aldrich, USA.

**Chain length and unsaturation prediction**

To predict chain length distribution, we first processed images with linear unmixing as described above. However, this analysis outputs two-dimensional chemical maps whereas a three-dimensional hyperspectral image is needed for chain length prediction. We created a hyperspectral, saturated fatty acid map by subtracting the protein and unsaturated fatty acid components from the original background-subtracted hyperspectral image (Fig. S10). We then calculated the area under the curve ratio of $CH_2$ to $CH_3$ for each pixel, using 2832 to 2888 nm for $CH_2$ and 2909 to 2967 nm for $CH_3$.

We used the linear relationship of ratio to chain length produced from standards (C6-C20, Sigma Aldrich, USA) to calculate a predicted chain length for each pixel. This prediction was then multiplied by a concentration weighting factor that corresponds to the SRS spectral summation at the same pixel. Thus, if the raw SRS signal is low then the prediction is also low. All pixels' concentration-weighted chain lengths were compiled to create the fatty acid chain length distribution. To calculate the unsaturation ratio, the sum of the C16:1 chemical map generated through linear unmixing was divided by the sum of the hyperspectral saturated chemical map.

**Single cell segmentation**

Segmentation of single cells was implemented in two steps. The protein segmentation map was first sent to CellProfiler to generate an initial segmentation[68]. A customized pipeline was used for the analysis, including illumination correction, background subtraction, and edge enhancements based on the Laplacian of the Gaussian. Then a custom Matlab program was used to manually correct errors in the automated segmentation analysis using the raw SRS and protein chemical maps as a guide.

**Acknowledgements**

This work was supported by the Office of Science (BER) at the U.S. Department of Energy (DE-SC0019387 to M.J.D., J.X.C., W.W.W.), the National Science Foundation (1804096 to MJD), and NIH R35 GM136223 to JXC. We thank Dr. Normal Lee and the Chemical Instrumentation Center for assistance with GC-MS experiments. Dr. Joshua Finkelstein provided valuable input on the manuscript.

**Author Contributions**

N.T. and H.L. performed experiments and conducted data analysis. M.J.D., J.X.C. and W.W.W. provided overall guidance on the project. N.T. was responsible for strain construction, production quantification, and sample preparation. H.L. performed SRS imaging. J.B.L. performed pilot experiments with a production strain. D.B. helped to develop the GC-MS protocol and quantified

strain production. N.T., H.L., and M.J.D. wrote the manuscript with input from J.B.L., W.W.W. and J.X.C.

**Competing Interests**

The authors declare no competing interests.

## Figures

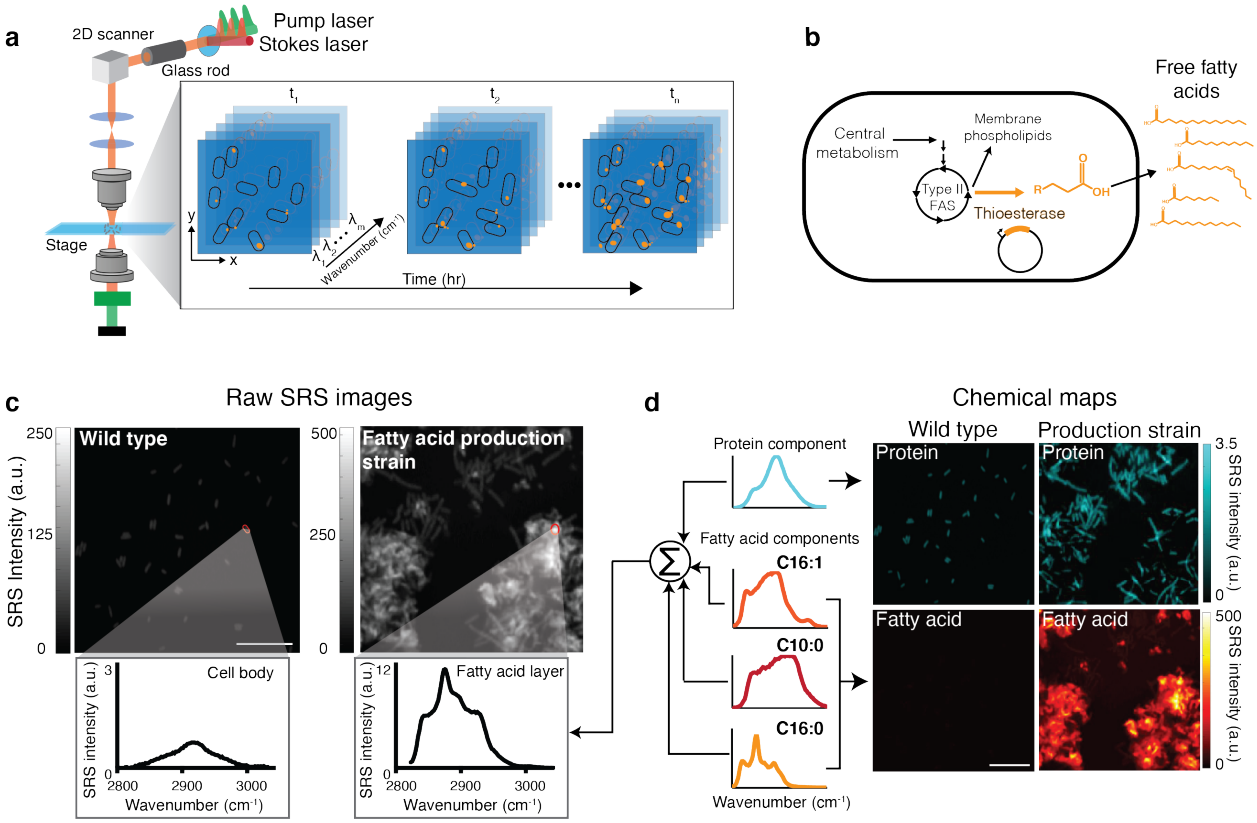


**Figure 1. SRS imaging of *E. coli* production strains shows single-cell free fatty acid levels**. **(a)** Schematic of the optical setup for SRS imaging to produce hyperspectral images using a Stokes and pump laser focused on a live sample. Hyperspectral SRS images contain three-dimensional data: x and y coordinates and wavenumber, which provides spectral information. Longitudinal SRS imaging adds a fourth dimension, time. **(b)** Schematic of free fatty acid production in *E. coli*. Expression of a cytosolic thioesterase results in free fatty acid accumulation through the type II fatty acid synthesis (FAS) pathway. Free fatty acids can vary in chain length and unsaturation, largely dictated by thioesterase specificity. **(c)** Representative raw SRS data from wild type *E. coli* and a strain overexpressing a cytosolic thioesterase (*Ab*TE*). The summation of Raman spectra at each pixel is shown. Representative regions are outlined in red with the corresponding Raman spectra shown below the image. Fatty acids and proteins emit strong Raman signals in the C–H region (~2900 $cm^{-1}$). Note that the y-axis scales are different; Fig. S3 shows them on the same scale. Scale bar, 10 µm. **(d)** Spectra at each pixel of the SRS image can be decomposed to generate chemical maps. Protein and fatty acid components are decomposed using spectral standards to produce chemical maps. Spectral standards shown in schematic are Bovine serum albumin (cyan), palmitoleic acid (C16:1, orange), capric acid (C10:0, red), and palmitic acid (C16:0, yellow). Protein and fatty acid chemical maps for both strains are shown. Scale bar, 10 µm.

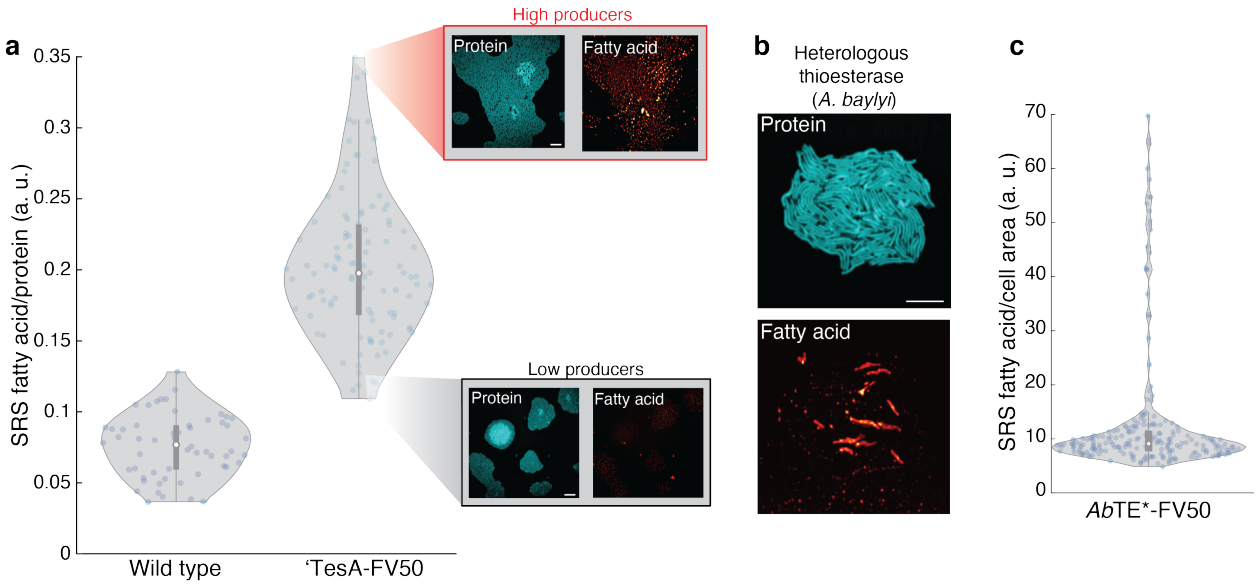


**Figure 2. Inter- and intra-colony heterogeneity profiles of production strains. (a)** Production from 'TesA-FV50 microcolonies (n = 105) are compared to wild type microcolonies (n = 56), revealing inter-colony production heterogeneity. Protein and fatty acid chemical maps are shown for representative high and low producing microcolonies. Scale bar, 10 µm. **(b)** Representative protein and fatty acid chemical maps are shown for a microcolony of the production strain *Ab*TE*-FV50. **(c)** Intra-colony production is quantified for single cells within the microcolony (n = 213) (Fig. S5). Scale bar, 10 µm.

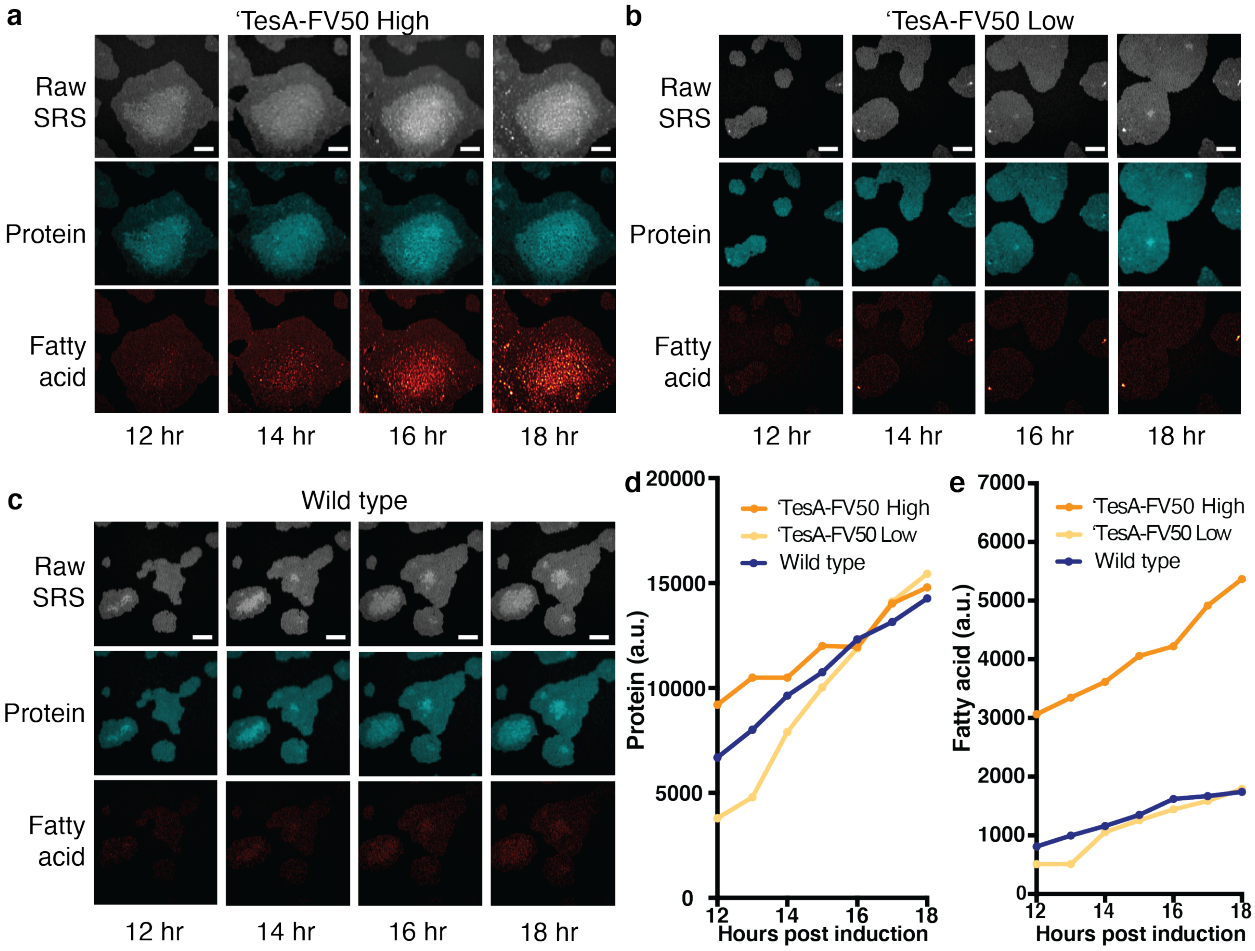


**Figure 3. Longitudinal SRS imaging of production dynamics.** Time-lapse images of **(a)** a 'TesA-FV50 high producing microcolony, **(b)** a 'TesA-FV50 low producing microcolony, and **(c)** a wild type control strain, shown with the raw SRS images (spectral summation of the SRS image stack) and chemical maps corresponding to protein and fatty acid content. Quantification of **(d)** protein and **(e)** fatty acid over time from the microcolonies in (a-c). Scale bars, 10 µm.

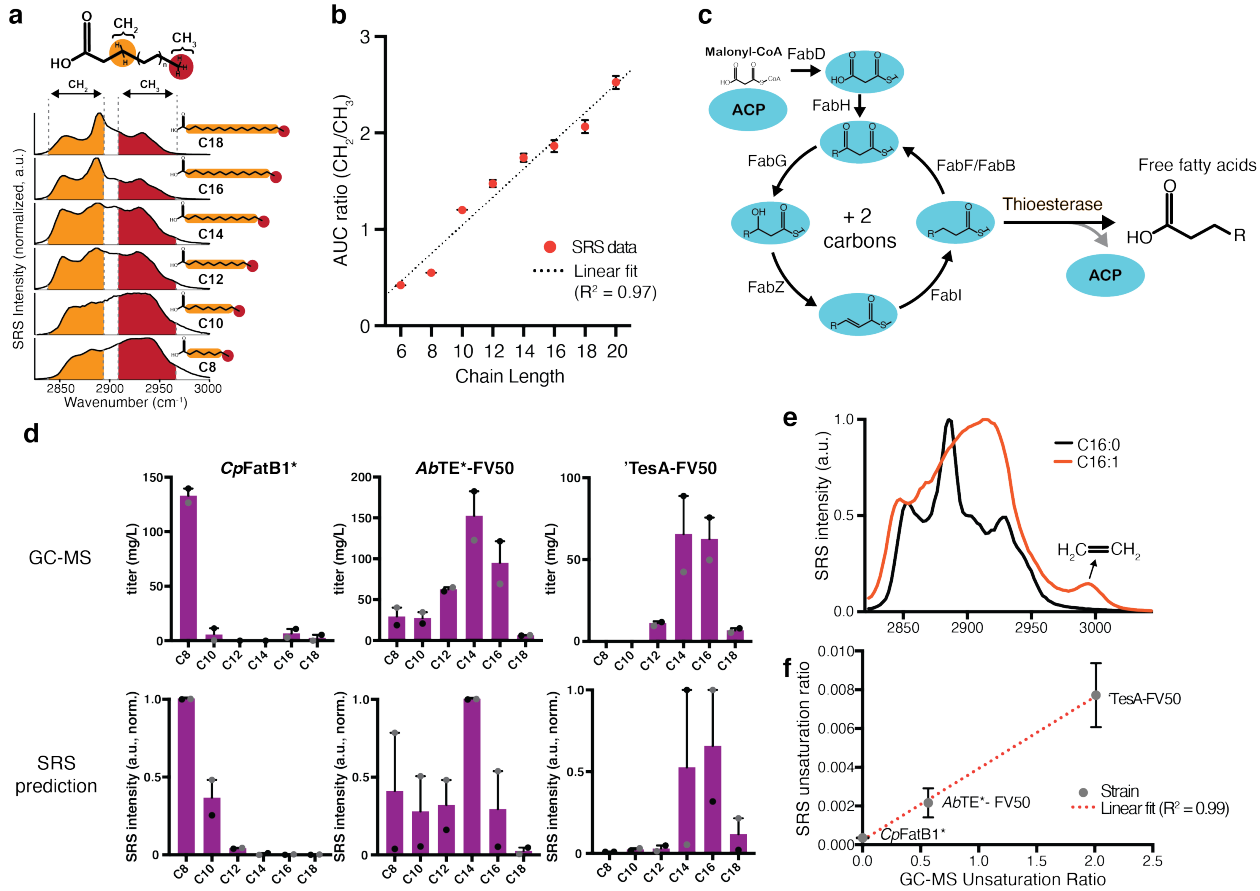


**Figure 4. Chain length distribution prediction from different thioesterase enzymes. (a)** The ratio of internal $CH_2$ and terminal $CH_3$ bonds within a fatty acid is a function of chain length. Raman spectra of pure fatty acid standards are shown for different chain lengths. Specific spectral windows correspond to each bond. **(b)** The ratio of area under the curve (AUC) of $CH_2/CH_3$ bonds scales approximately linearly with chain length. **(c)** Schematic of the type II fatty acid synthesis pathway in *E. coli*. Introduction of an acyl-ACP thioesterase pulls out elongating acyl-ACPs to form free fatty acids. Enzymatic specificity of the thioesterase largely determines the distribution of the fatty acid chain length profile. **(d)** Chain length distribution prediction with GC-MS compared to SRS using $CH_2/CH_3$ ratio analysis (n = 2 biological replicates using 5 fields of view for each replicate). Strains shown are: *Cp*FatB1*, *Ab*TE*-FV50, and 'TesA-FV50 (Table 2). **(e)** SRS spectra of saturated and unsaturated fatty acid standards (C16:0, C16:1). The unique peak at ~3000 $cm^{-1}$ allows for spectral decomposition of unsaturation content. **(f)** Comparing GC-MS unsaturation ratio of produced free fatty acids to SRS production based on spectral analysis Error bars show standard deviation from n = 5 fields of view for each strain.

**Tables**

**Table 1.** List of plasmids used in this study.

| Plasmid | Origin | Overexpressed operon | Resistance | Reference |
|---|---|---|---|---|
| **pSS200** | pMB1 | $P_{trc}$-*Ab*te:G17R/A165R | $Amp^R$ | Sarria *et al.* [35] |
| **pBbA5c-'tesA-vhb50-8fadR** | p15a | $P_{lacUV5}$-'tesA-vhb50, $P_{BAD}$-fadR | $Cm^R$ | Liu *et al.* [37] |
| **pBbA5c-vhb50-8fadR** | p15a | $P_{lacUV5}$-vhb50, $P_{BAD}$-fadR | $Cm^R$ | This study |
| **pBbA5c-CpFatB1.2-M4-287** | p15a | $P_{lacUV5}$-*Cp*fatB1.2-M4-287 | $Cm^R$ | This study, mutant enzyme from Hernandez Lozada, *et al*. [56] |

**Table 2.** List of *E. coli* strains used in this study.

| Strain | Relevant genotype | Reference |
|---|---|---|
| **BW25113 (wild type)** | $F^-$ Δ(araD-araB)567 ΔlacZ4787(::rrnB-3) $\lambda^-$ rph-1 Δ(rhaD-rhaB)568 hsdR514 | Baba *et al.* [64] |
| **BW25113 Δ*fadE*** | *E. coli* BW25113 Δ*fadE*, cured from Keio collection | Baba *et al.* [64] |
| **MG1655** | $F^-$, $\lambda^-$, rph-1 | Blattner, *et al*. [69] |
| ***Ab*TE*** | *E. coli* MG1655; pSS200 | Sarria *et al.* [35] |
| **'TesA-FV50** | *E. coli* BW25113 Δ*fadE*; pBbA5c-'tesA-vhb50-8fadR | Liu *et al.* [37] |
| ***Ab*TE*-FV50** | *E. coli* MG1655; pBbA5c-vhb50-8fadR, pSS200 | This study |
| ***Cp*FatB1*** | *E. coli* MG1655; pBbA5c-CpfatB1.2-M4-287 | This study |

## Supplementary Figures

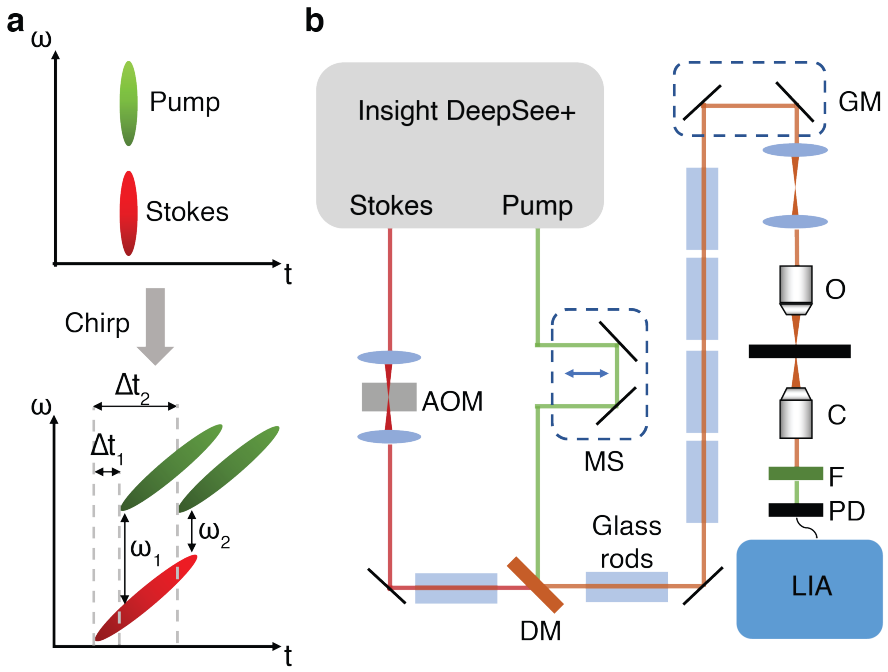


**Figure S1.** Hyperspectral SRS setup. **(a)** Concept of hyperspectral SRS using spectral focusing. The pump and Stokes lasers are linearly chirped by high dispersion glass rods to temporally separate the spectral components. Each temporal delay between the two pulses corresponds to a Raman vibrational mode. **(b)** Optical setup. AOM, acousto-optic modulator; MS, motorized stage; DM, dichroic mirror; GM, galvo mirrors; O, objective; C, condenser; F, filter; PD, photodiode; LIA, lock-in amplifier.

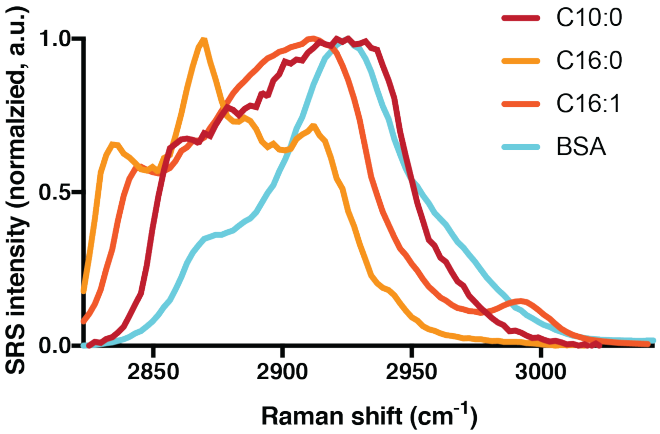


**Figure S2.** SRS spectra of pure standards used to analyze hyperspectral images to produce chemical maps. (BSA: bovine serum albumin, C10:0: decanoic acid, C16:0: palmitic acid, C16:1: palmitoleic acid).

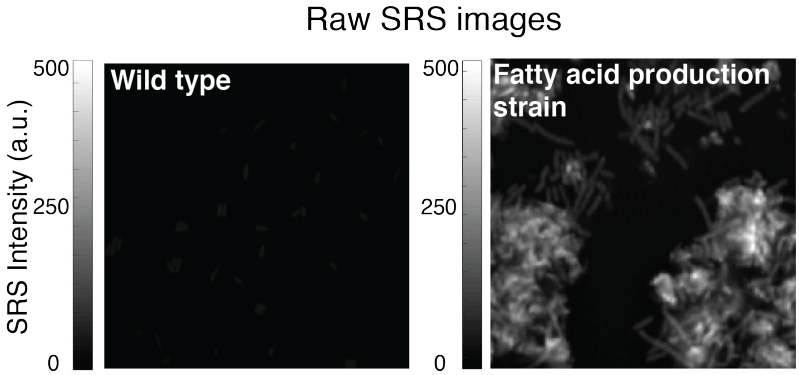


**Figure S3.** Raw SRS images shown in Fig. 1c of wild type and a strain overexpressing a cytosolic thioesterase (*Ab*TE*), but with both images scaled with the same color axis for direct comparison.

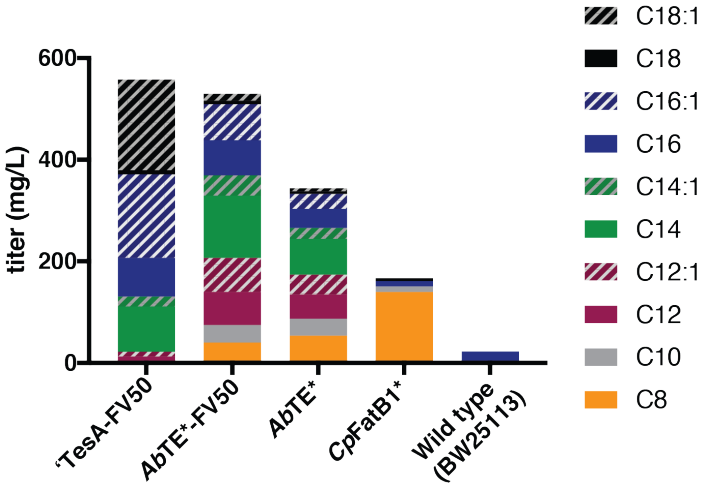


**Figure S4.** Fatty acid production quantification for strains in this study. GC-MS quantified fatty acid production data for each strain. Cells were grown 24 hours post thioesterase induction in liquid culture. For chain length prediction, these exact cultures were taken for SRS imaging at the same timepoint.

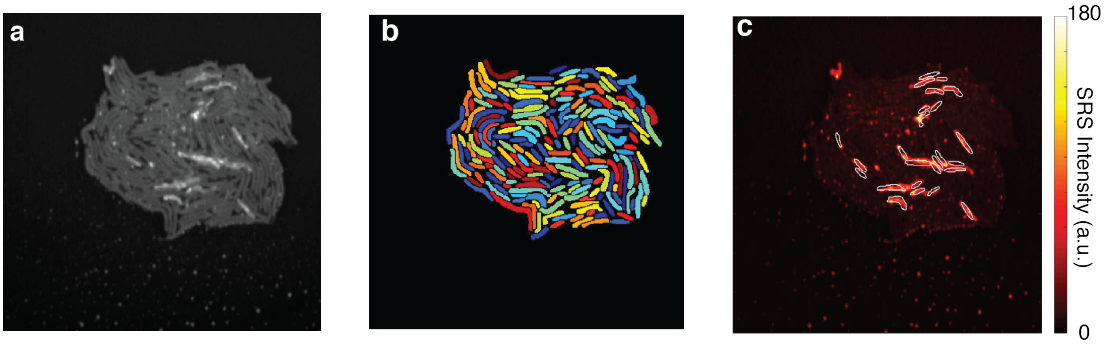


**Figure S5.** Single cell segmentation of a microcolony. **(a)** Raw SRS images are used to segment microcolonies to perform single cell analysis shown in Fig. 2c. **(b)** Segmentation of microcolony in (a). **(c)** Segmentation of the top 25 highest producing cells overlaid on the fatty acid map of the microcolony.

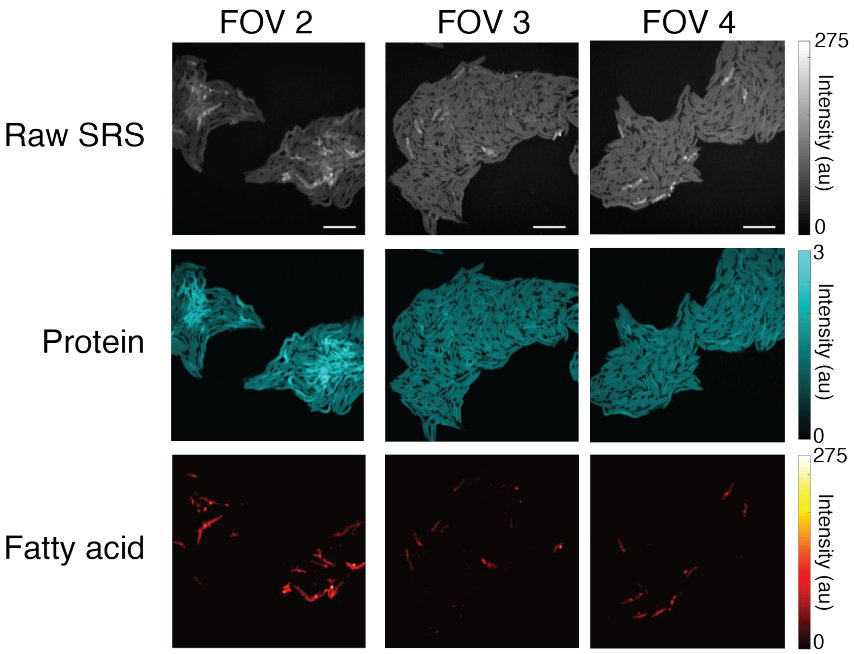


**Figure S6.** Intra-colony heterogeneity of the *Ab*TE* strain. **(a)** Three additional fields of view (FOV) of the *Ab*TE*-FV50 strain shown in Fig. 2b. Raw SRS, protein, and fatty acid chemical maps are shown for all. Scale bars, 10 µm.

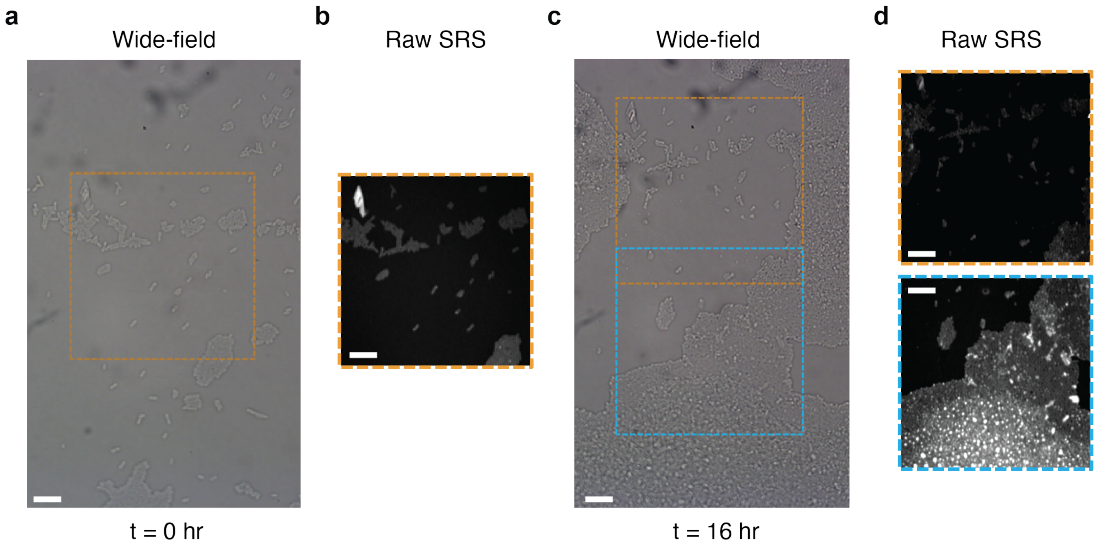


**Figure S7.** Testing photodamage on live *E. coli* cells. **(a)** Wide-field transmission image of *E. coli* cells at the start of the cell incubation (t = 0 hr). **(b)** Hyperspectral SRS image of the region highlighted with a yellow rectangle in (a). **(c)** Wide-field transmission image of the same field of view after incubation (t = 16 hr). **(d)** Hyperspectral SRS images of the previously scanned region (yellow rectangle in (c)) and an adjacent region without previous SRS laser exposure (blue rectangle in (c)). Scale bars, 10 µm.

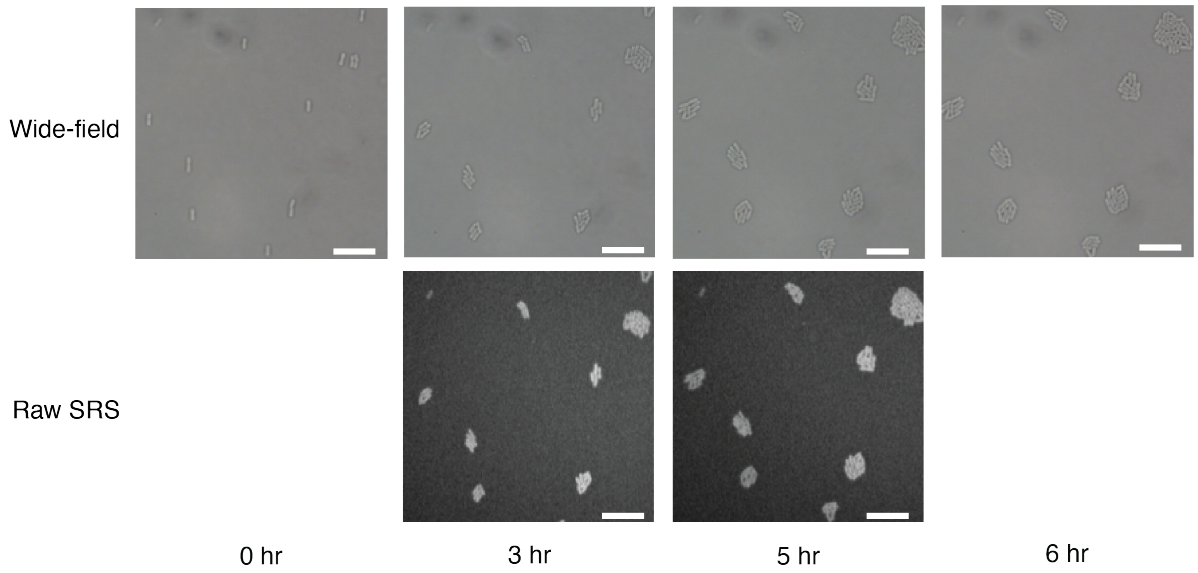


**Figure S8.** Optimized SRS laser powers enable live cell imaging of *E. coli*. Wide-field transmission image of *E. coli*, with raw hyperspectral SRS images of the same region for the t = 3 and 5 hr timepoints. Spectral summation is shown. Scale bars, 10 µm.

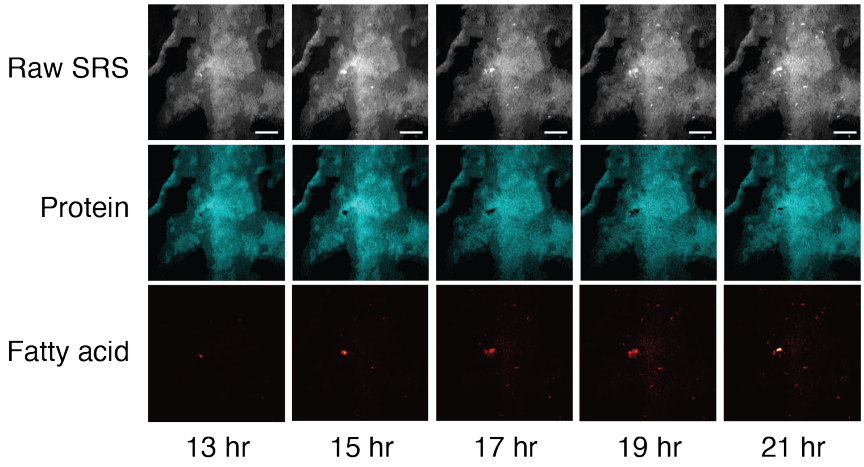


**Figure S9.** Time-lapse images of fatty acid production in the *Ab*TE*-FV50 strain. **(a)** Raw SRS images, protein, and fatty acid chemical maps are shown. Time values represent time grown on the agarose pad after IPTG induction. Scale bars, 10 µm.

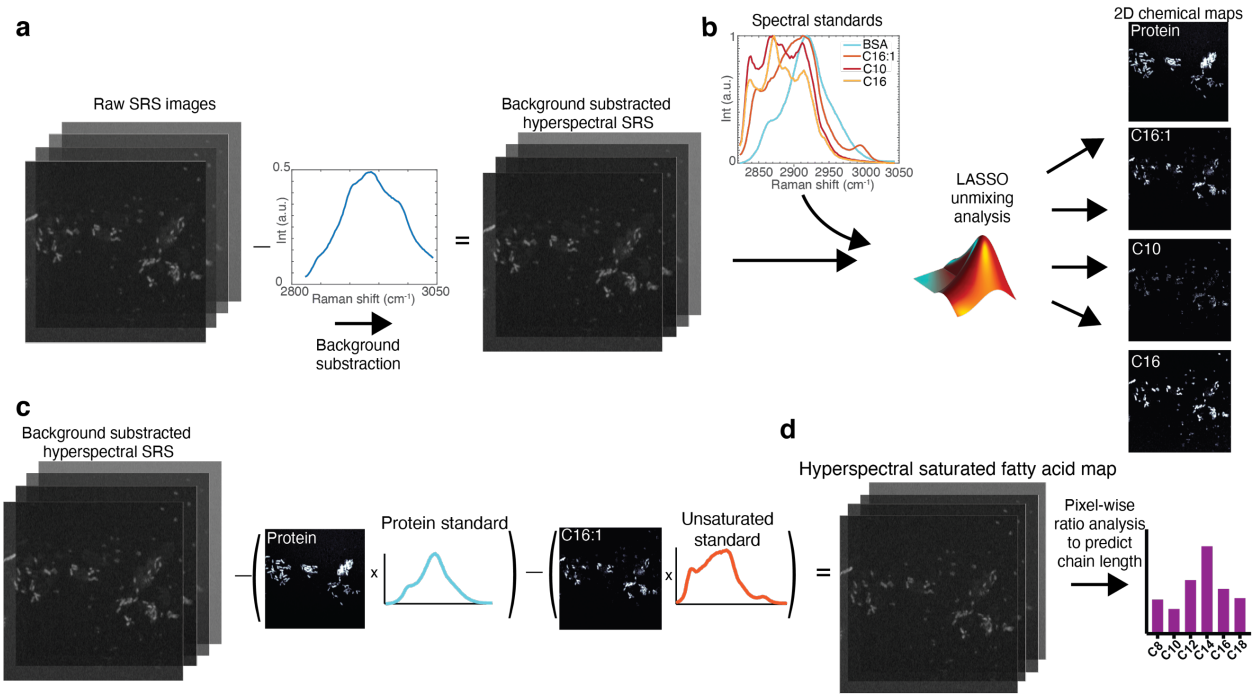


**Figure S10.** Analysis workflow for chain length prediction from hyperspectral SRS images. **(a)** Raw SRS images are first background subtracted. **(b)** Background subtracted images are unmixed using chemical standards. Protein, BSA; unsaturated fatty acid, C16:1; medium chain fatty acid, C10; and long chain fatty acid, C16. C10 and C16 maps are used to represent a mixture of saturated fatty acids. **(c)** Protein and unsaturated fatty acid maps are multiplied by their respective standard spectra and subtracted from the background-subtracted hyperspectral image to produce a three-dimensional saturated fatty acid map. **(d)** Ratio analysis is performed on each pixel to calculate chain length and weighted by raw intensity to predict chain length distribution of the field of view.

## Supplementary Movies

**Movie S1.** Time-lapse wide-field transmission images of the wild type strain during the live cell SRS imaging shown in Fig. 3c. The white box indicates the SRS imaging region.

**References**


1. Liu, Y. *et al.* Biofuels for a sustainable future. *Cell* **184**, 1636–1647 (2021).
2. Choi, S. Y. *et al.* Metabolic engineering for the synthesis of polyesters: A 100-year journey from polyhydroxyalkanoates to non-natural microbial polyesters. *Metab. Eng.* **58**, 47–81 (2020).
3. Cao, M., Gao, M., Suástegui, M., Mei, Y. & Shao, Z. Building microbial factories for the production of aromatic amino acid pathway derivatives: From commodity chemicals to plant-sourced natural products. *Metab. Eng.* **58**, 94–132 (2020).
4. Chae, T. U. *et al.* Metabolic engineering for the production of dicarboxylic acids and diamines. *Metab. Eng.* **58**, 2–16 (2020).
5. Lee, S. Y. *et al.* A comprehensive metabolic map for production of bio-based chemicals. *Nat. Catal.* **2**, 18–33 (2019).
6. Torella, J. P. *et al.* Tailored fatty acid synthesis via dynamic control of fatty acid elongation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11290–11295 (2013).
7. Zhang, F. *et al.* Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metab. Eng.* **14**, 653–660 (2012).
8. Cao, Y. *et al.* Production of free monounsaturated fatty acids by metabolically engineered Escherichia coli. *Biotechnol. Biofuels* **7**, 1–11 (2014).
9. Choi, Y. J. & Lee, S. Y. Microbial production of short-chain alkanes. *Nature* **502**, 571–574 (2013).
10. Yan, Q. & Pfleger, B. F. Revisiting metabolic engineering strategies for microbial synthesis of oleochemicals. *Metab. Eng.* **58**, 35–46 (2020).
11. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science (80-. ).* **297**, 1183–1186 (2002).
12. Kiviet, D. J. *et al.* Stochasticity of metabolism and growth at the single-cell level. *Nature* **514**, 376–379 (2014).
13. Heins, A. L. & Weuster-Botz, D. Population heterogeneity in microbial bioprocesses: origin, analysis, mechanisms, and future perspectives. *Bioprocess Biosyst. Eng.* **41**, 889–916 (2018).
14. Rugbjerg, P., Myling-Petersen, N., Porse, A., Sarup-Lytzen, K. & Sommer, M. O. A. Diverse genetic error modes constrain large-scale bio-based production. *Nat. Commun.* **9**, (2018).
15. Rugbjerg, P., Dyerberg, A. S. B., Quainoo, S., Munck, C. & Sommer, M. O. A. Short and long-read ultra-deep sequencing profiles emerging heterogeneity across five platform Escherichia coli strains. *Metab. Eng.* **65**, 197–206 (2021).
16. Delvigne, F., Zune, Q., Lara, A. R., Al-Soud, W. & Sørensen, S. J. Metabolic variability in bioprocessing: Implications of microbial phenotypic heterogeneity. *Trends Biotechnol.* **32**, 608–616 (2014).
17. Delvigne, F. & Goffin, P. Microbial heterogeneity affects bioprocess robustness: Dynamic single-cell analysis contributes to understanding of microbial populations. *Biotechnol. J.* **9**, 61–72 (2014).
18. Takors, R. Scale-up of microbial processes: Impacts, tools and open questions. *J. Biotechnol.* **160**, 3–9 (2012).
19. Chiu, H. H. & Kuo, C. H. Gas chromatography-mass spectrometry-based analytical strategies for fatty acid analysis in biological samples. *J. Food Drug Anal.* **28**, 60–73 (2020).

20. Zhang, J., Jensen, M. K. & Keasling, J. D. Development of biosensors and their application in metabolic engineering. *Curr. Opin. Chem. Biol.* **28**, 1–8 (2015).
21. Zhang, F., Carothers, J. M. & Keasling, J. D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* **30**, 354–359 (2012).
22. Xu, P., Li, L., Zhang, F., Stephanopoulos, G. & Koffas, M. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 11299–11304 (2014).
23. Tang, S. Y. & Cirino, P. C. Design and application of a mevalonate-responsive regulatory protein. *Angew. Chemie - Int. Ed.* **50**, 1084–1086 (2011).
24. Xiao, Y., Bowen, C. H., Liu, D. & Zhang, F. Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis. *Nat. Chem. Biol.* **12**, 339–344 (2016).
25. Zhang, C. & Cheng, J. X. Perspective: Coherent Raman scattering microscopy, the future is bright. *APL Photonics* **3**, (2018).
26. Hu, F., Shi, L. & Min, W. Biological imaging of chemical bonds by stimulated Raman scattering microscopy. *Nat. Methods* **16**, 830–842 (2019).
27. Tyo, K. E., Alper, H. S. & Stephanopoulos, G. N. Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends Biotechnol.* **25**, 132–137 (2007).
28. Jeschek, M., Gerngross, D. & Panke, S. Combinatorial pathway optimization for streamlined metabolic engineering. *Curr. Opin. Biotechnol.* **47**, 142–151 (2017).
29. Fu, D., Holtom, G., Freudiger, C., Zhang, X. & Xie, X. S. Hyperspectral imaging with stimulated Raman scattering by chirped femtosecond lasers. *J. Phys. Chem. B* **117**, 4634–40 (2013).
30. Hellerer, T., Enejder, A. M. K. & Zumbusch, A. Spectral focusing: High spectral resolution spectroscopy with broad-bandwidth laser pulses. *Appl. Phys. Lett.* **85**, 25–27 (2004).
31. Liao, C.-S. *et al.* Stimulated Raman spectroscopic imaging by microsecond delay-line tuning. *Optica* **3**, 1377–1380 (2016).
32. Lennen, R. M. & Pfleger, B. F. Engineering Escherichia coli to synthesize free fatty acids. *Trends Biotechnol.* **30**, 659–667 (2012).
33. Cho, H. & Cronan, J. E. Defective export of a periplasmic enzyme disrupts regulation of fatty acid synthesis. *J. Biol. Chem.* **270**, 4216–4219 (1995).
34. Jiang, P. & Cronan, J. E. Inhibition of fatty acid synthesis in Escherichia coli in the absence of phospholipid synthesis and release of inhibition by thioesterase action. *J. Bacteriol.* **176**, 2814–2821 (1994).
35. Sarria, S., Bartholow, T. G., Verga, A., Burkart, M. D. & Peralta-Yahya, P. Matching Protein Interfaces for Improved Medium-Chain Fatty Acid Production. *ACS Synth. Biol.* **7**, 1179–1187 (2018).
36. Rugbjerg, P. & Sommer, M. O. A. Overcoming genetic heterogeneity in industrial fermentations. *Nat. Biotechnol.* **37**, 869–876 (2019).
37. Liu, D., Wan, N., Zhang, F., Tang, Y. J. & Wu, S. G. Enhancing fatty acid production in Escherichia coli by Vitreoscilla hemoglobin overexpression. *Biotechnol. Bioeng.* **114**, 463–467 (2017).
38. Huang, J. F. *et al.* Systematic Analysis of Bottlenecks in a Multibranched and Multilevel Regulated Pathway: The Molecular Fundamentals of l -Methionine Biosynthesis in Escherichia coli. *ACS Synth. Biol.* **7**, 2577–2589 (2018).

39. Lennen, R. M. *et al.* Membrane stresses induced by overproduction of free fatty acids in Escherichia coli. *Appl. Environ. Microbiol.* **77**, 8114–8128 (2011).
40. Lee, H. J. *et al.* Label-Free Vibrational Spectroscopic Imaging of Neuronal Membrane Potential. *J. Phys. Chem. Lett.* **8**, 1932–1936 (2017).
41. Wakisaka, Y. *et al.* Probing the metabolic heterogeneity of live Euglena gracilis with stimulated Raman scattering microscopy. *Nat. Microbiol.* **1**, 8–11 (2016).
42. Wei, L. *et al.* Live-Cell Bioorthogonal Chemical Imaging: Stimulated Raman Scattering Microscopy of Vibrational Probes. *Acc. Chem. Res.* **49**, 1494–1502 (2016).
43. Heath, R. J. & Rock, C. O. Regulation of fatty acid elongation and initiation by acyl-acyl carrier protein in Escherichia coli. *J. Biol. Chem.* **271**, 1833–1836 (1996).
44. Davis, M. S. & Cronan, J. Inhibition of Escherichia coli acetyl coenzyme A carboxylase by acyl-acyl carrier protein. *J. Bacteriol.* **183**, 1499–1503 (2001).
45. My, L., Ghandour Achkar, N., Viala, J. P. & Bouveret, E. Reassessment of the Genetic Regulation of Fatty Acid Synthesis in Escherichia coli: Global Positive Control by the Functional Dual Regulator FadR. *J. Bacteriol.* **197**, 1862–1872 (2015).
46. Feng, Y. & Cronan, J. E. Complex binding of the FabR repressor of bacterial unsaturated fatty acid biosynthesis to its cognate promoters. *Mol. Microbiol.* **80**, 195–218 (2011).
47. Noga, M. J. *et al.* Posttranslational control of plsb is sufficient to coordinate membrane synthesis with growth in escherichia coli. *MBio* **11**, 1–15 (2020).
48. Peralta-Yahya, P. P., Zhang, F., Del Cardayre, S. B. & Keasling, J. D. Microbial engineering for the production of advanced biofuels. *Nature* **488**, 320–328 (2012).
49. Sarria, S., Kruyer, N. S. & Peralta-Yahya, P. Microbial synthesis of medium-chain chemicals from renewables. *Nat. Biotechnol.* **35**, 1158–1166 (2017).
50. Czamara, K. *et al.* Raman spectroscopy of lipids: A review. *J. Raman Spectrosc.* **46**, 4–20 (2015).
51. Heath, R. J. & Rock, C. O. The Claisen condensation in biology. *Nat. Prod. Rep.* **19**, 581–596 (2002).
52. Ruppe, A. & Fox, J. M. Analysis of Interdependent Kinetic Controls of Fatty Acid Synthases. *ACS Catal.* **8**, 11722–11734 (2018).
53. Lo, Y. C., Lin, S. C., Shaw, J. F. & Liaw, Y. C. Crystal structure of Escherichia coli thioesterase I/protease I/lysophospholipase L1: Consensus sequence blocks constitute the catalytic center of SGNH-hydrolases through a conserved hydrogen bond network. *J. Mol. Biol.* **330**, 539–551 (2003).
54. Grisewood, M. J. *et al.* Computational Redesign of Acyl-ACP Thioesterase with Improved Selectivity toward Medium-Chain-Length Fatty Acids. *ACS Catal.* **7**, 3837–3849 (2017).
55. Deng, X. *et al.* Structure-guided reshaping of the acyl binding pocket of 'TesA thioesterase enhances octanoic acid production in E. coli. *Metab. Eng.* **61**, 24–32 (2020).
56. Hernández Lozada, N. J. *et al.* Highly Active C 8 -Acyl-ACP Thioesterase Variant Isolated by a Synthetic Selection Strategy. *ACS Synth. Biol.* **7**, 2205–2215 (2018).
57. Zhang, D., Slipchenko, M. N. & Cheng, J. X. Highly sensitive vibrational imaging by femtosecond pulse stimulated raman loss. *J. Phys. Chem. Lett.* **2**, 1248–1253 (2011).
58. Fu, Y., Wang, H., Shi, R. & Cheng, J.-X. Characterization of photodamage in coherent anti-Stokes Raman scattering microscopy. *Opt. Express* **14**, 3942 (2006).
59. Hartline, C. J., Schmitz, A. C., Han, Y. & Zhang, F. Dynamic control in metabolic engineering: Theories, tools, and applications. *Metab. Eng.* **63**, 126–140 (2021).

60. Zhao, E. M. *et al.* Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature* **555**, 683–687 (2018).
61. Ozeki, Y., Dake, F., Kajiyama, S., Fukui, K. & Itoh, K. Analysis and experimental assessment of the sensitivity of stimulated Raman scattering microscopy. *Opt. Express* **17**, 3651–3658 (2009).
62. Lin, H. *et al.* Microsecond fingerprint stimulated Raman spectroscopic imaging by ultrafast tuning and spatial-spectral learning. *Nat. Commun.* **12**, 1–12 (2021).
63. Nitta, N. *et al.* Raman image-activated cell sorting. *Nat. Commun.* **11**, 1–16 (2020).
64. Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol. Syst. Biol.* **2**, (2006).
65. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6640–6645 (2000).
66. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647 (2008).
67. Anderson, J. C. *et al.* BglBricks: A flexible standard for biological part assembly. *J. Biol. Eng.* **4**, 1–12 (2010).
68. Carpenter, A. E. *et al.* CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, (2006).
69. Blattner, F. R. *et al.* The complete genome sequence of Escherichia coli K-12. *Science (80-. ).* **277**, 1453–1462 (1997).