Chapter 6

Growth and Preparation of *Staphylococcus epidermidis* for NMR Metabolomic Analysis

Greg A. Somerville and Robert Powers

Abstract

The "omics" era began with transcriptomics and this progressed into proteomics. While useful, these approaches provide only circumstantial information about carbon flow, metabolic status, redox poise, etc. To more directly address these metabolic concerns, researchers have turned to the emerging field of metabolomics. In our laboratories, we frequently use NMR metabolomics to acquire a snapshot of bacterial metabolomes during stressful or transition events. Irrespective of the "omics" method of choice, the experimental outcome depends on the proper cultivation and preparation of bacterial samples. In addition, the integration of these large datasets requires that these cultivation conditions be clearly defined [1].

Key words Bacterial growth, Growth conditions, Metabolism, Metabolomics, NMR

1 Introduction

Staphylococci live in, and transition between, diverse host environments that can lead to rapid changes in the availability of nutrients necessary to provide energy and biosynthetic intermediates for the synthesis of macromolecules. These changes in nutrient availability alter the bacterial metabolic status, which cause regulatory changes that help bacteria adapt and survive in the new environment [2]. Changes in nutrient availability not only occur when bacteria are growing on, or in, a host environment, but they also occur during non-steady-state cultivation. Since most bacterial cultivation in microbiology laboratories is done using batch cultures, an understanding of the in vitro cultivation of *Staphylococcus epidermidis* is essential for experimental success. Achieving the reproducible cultivation of *S. epidermidis*, or any bacterium, requires consideration of three things: bacterial strain, culture medium, and cultivation conditions [3].

Paul D. Fey (ed.), *Staphylococcus Epidermidis: Methods and Protocols*, Methods in Molecular Biology, vol. 1106,
DOI 10.1007/978-1-62703-736-5_6,

1.1 Bacterial Strain Selection

There are two schools of thought regarding strain selection: the first advocates a homogeneous approach using standard strains, that is to say, have many laboratories studying one common strain in great detail to maximize understanding of that particular strain. The second school of thought uses a heterogeneous approach with multiple strains to gain a broader understanding of the bacterial species population. We will leave it to the readers to decide which approach is more valuable; however, we would note that given the large number of *S. epidermidis* strains in use around the world it seems that most researchers have decided on the heterogeneous approach.

1.2 Culture Medium

There are three types of culture media: chemically defined, complex, and selective. A chemically defined medium (CDM) is one whose exact chemical composition is known [4]. Complex media are composed of digests of chemically undefined substances such as yeast, soy, or meat extracts, for example, casamino acids–yeast extract–glycerophosphate (CYGP) broth [5]. Finally, selective media can be made from either chemically defined or complex media and are used to enhance the isolation of particular bacterial species (e.g., mannitol salt agar). In NMR metabolomic studies, the choice of medium will largely depend on which isotopically labeled metabolite is being followed. For example, when using ^{15}N-glutamate, it is impractical to add labeled glutamate to a complex medium containing an unknown concentration of unlabeled glutamate. In this example, to achieve maximal labeling of the bacteria, it would be best to use a CDM lacking glutamate and glutamine. In our laboratories, we typically label staphylococci using ^{13}C-glucose in the complex medium tryptic soy broth (TSB) that is devoid of unlabeled glucose [6, 7]. This medium allows for maximal biomass generation while assuring that nearly all (~99 %; 1.1 % is due to naturally occurring ^{13}C) of the ^{13}C-labeled metabolites in the metabolome were derived from glucose.

1.3 Cultivation of S. epidermidis

In order to integrate the large quantity of genomic, proteomic, and metabolomic data that is rapidly accumulating in the literature, there must be a common thread between these studies that permits the direct comparison of data. This common thread must start with clearly defined growth conditions [1]. Unfortunately, as microbiology has progressed into the "omics" age, little emphasis has been placed on the necessary rigor needed to cultivate bacteria in a reproducible manner that permits the integration of data from different laboratories. The importance of having clearly defined growth conditions is easily demonstrated using the example of staphylococcal pyruvate catabolism [8]. During microaerobic or anaerobic growth, pyruvate is primarily reduced to lactic acid with the concomitant oxidation of NADH to NAD$^+$. In contrast, during aerobic growth pyruvate undergoes oxidative decarboxylation to produce acetyl-CoA, which is used to synthesize the small

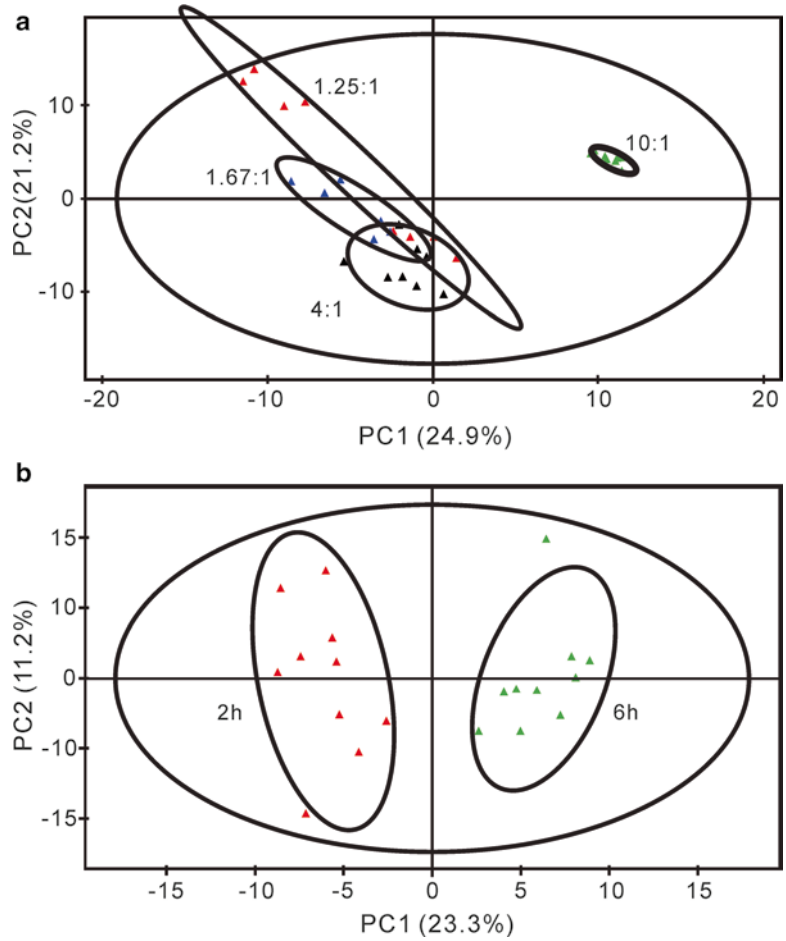

Fig. 1 Staphylococcal metabolism dramatically changes in response to changes in the flask-to-medium ratio and during the transition from the exponential phase of growth to the post-exponential growth phase. (**a**) Principal component analysis (PCA) of *S. aureus* metabolomes grown with differing flask-to-medium ratios (*indicated in the figure*). (**b**) PCA scores plot of the metabolomes of exponential (2 h) and post-exponential (6 h) growth phases for *S. epidermidis* strain 1457 grown at 37 °C with a flask-to-medium ratio of 10:1 and 225 rpm aeration. The ellipses correspond to the 95 % confidence limits from a normal distribution for each cluster

phospho-donor acetyl-phosphate. Acetyl-phosphate serves as the substrate for acetate kinase in substrate-level phosphorylation to generate ATP and acetic acid. It is obvious from this example that the metabolomes of staphylococci grown under microaerobic, anaerobic, or aerobic conditions would be different, hence the need to clearly define the growth conditions (Fig. 1a). At a minimum, microbiologists must define the cultivation medium, temperature, pH buffering (if used), % CO_2 (if used), flask-to-medium ratio, revolutions per minute of agitation (if used), use of baffled or non-baffled flasks, and inoculum dose. Ideally, it would be best to report the growth rates for every set of cultivation conditions [1].

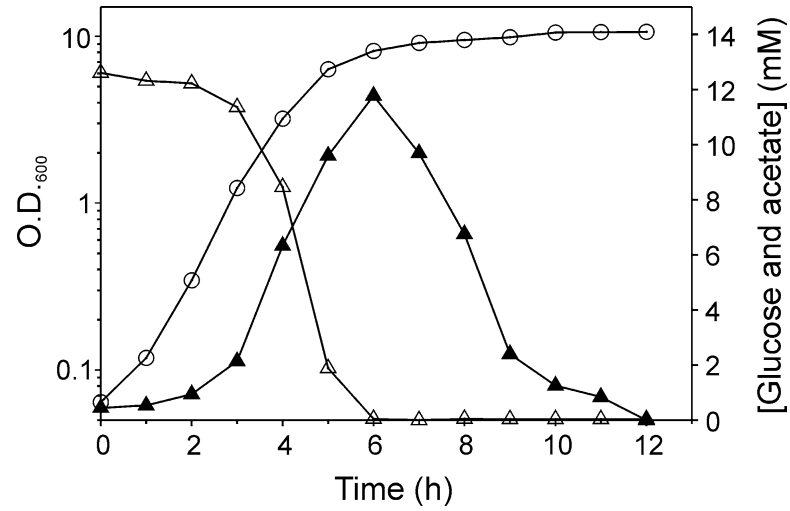

Fig. 2 Growth analysis of *S. epidermidis* showing the temporal depletion of glucose and the accumulation and depletion of acetate. *S. epidermidis* strain 1457 was grown in TSB at 37 °C with a flask-to-medium ratio of 10:1 and 225 rpm aeration

While reporting the cultivation conditions essential for integrating "omics" data, it is equally important to be aware that staphylococcal metabolism also changes during growth-phase transitions and that these changes can have regulatory consequences [2, 8].

1.4 Growth Phase-Dependent Differences in Metabolism

A non-steady-state bacterial growth profile consists of lag, exponential, stationary, and death phases. The lag phase is the time during which bacteria are increasing the cell size and mass in preparation for the rapid growth to follow. After the bacteria have achieved their initiation mass, they enter into a period of exponential, not logarithmic, growth. Exit from the exponential growth phase usually occurs when one or more nutrients are depleted from the medium or when there is an accumulation of a toxic by-product of metabolism. Following the exponential growth phase, the bacteria enter into the stationary phase, which is characterized by slower, or no, growth. Although not technically defined as a growth phase, the period after the end of the exponential growth phase, but prior to the cessation of growth, is often referred to as the post-exponential growth phase. In staphylococci, the exponential and post-exponential growth phases coincide with two very distinct metabolic states (Figs. 1b and 2).

As stated above, for NMR metabolomic studies, we aerobically cultivate *S. epidermidis* in TSB containing uniformly labeled ^{13}C-glucose. In the exponential growth phase, glucose is primarily catabolized through the glycolytic and pentose phosphate pathways. This is due in part to the CcpA-dependent repression of tricarboxylic acid (TCA) cycle activity, which prevents the flow of carbon into the TCA cycle [9]. The aerobic growth conditions together with a

minimally active TCA cycle shunt the pyruvate from glycolysis through the pyruvate dehydrogenase complex to generate acetyl-CoA, which is used in the phosphotransacetylase/acetate kinase pathway to produce acetyl-phosphate [8]. Acetyl-phosphate is used for substrate-level phosphorylation to generate ATP, but it also leads to the accumulation of acetate in the culture medium (Fig. 2). In contrast to aerobic growth conditions that we typically employ, the heterofermentative nature of *S. epidermidis* will lead to the accumulation of lactic acid and small amounts of formate, ethanol, acetate, and other alcohols during microaerobic or anaerobic growth. This aerobic/anaerobic change in catabolic profiles is similar to Louis Pasteur's observation that in the absence of oxygen yeasts consume more glucose than in the presence of oxygen and that this increased consumption coincided with alcoholic and lactic acid fermentation [10]. In the laboratory, the transition from aerobic growth to fermentative growth can occur by simply decreasing the flask-to-medium ratio from 10:1 to 4:1 (Fig. 1a); for this reason, it is critical to state the exact growth conditions.

Whether *S. epidermidis* is aerobically or anaerobically cultivated, the exponential growth phase is characterized by the rapid catabolism of carbohydrates and the accumulation of partially oxidized acidic and alcohol by-products of pyruvate catabolism (Fig. 2). In the absence of oxygen or other electron acceptor(s) (e.g., nitrate), growth will cease shortly after the depletion of carbohydrates due to an inability to complete the oxidation of these by-products. When oxygen is available, the bacteria will catabolize the acidic and alcohol by-products, a process requiring the TCA cycle. This second distinct metabolic state coincides with the post-exponential growth-phase derepression of the TCA cycle, gluconeogenesis, and numerous biosynthetic pathways. Using our example of aerobically cultivated *S. epidermidis*, acetate is taken up from the culture medium (Fig. 2) and enters the TCA cycle in the form of acetyl-CoA via a condensation reaction with oxaloacetate to generate citric acid. The complete oxidation of acetate by way of the TCA cycle creates biosynthetic intermediates and reducing potential to drive biosynthesis and oxidative phosphorylation, which supports post-exponential growth in a medium depleted of readily catabolizable carbohydrates. In summary, the transition between the two distinct staphylococcal metabolic states can be followed by monitoring the bacterial growth phases.

1.5 Spectrophotometric Evaluation of Bacterial Growth

When cultivating staphylococci for metabolomic analysis, the monitoring of bacterial growth is a critical parameter due to the fact that variation in growth conditions, growth phases, or cell numbers will be reflected in the NMR spectra. To assess bacterial growth, there are three commonly used practices: spectrophotometry, microscopic counting, and determination of colony-forming units (cfu)/mL [11]. For a rapidly growing bacterium, the determination of cfu in real time is not possible; hence, this method is better

suited for confirming the number of cfu taken at harvest than it is for determining how many mL of culture must be harvested to achieve equivalent bacterial numbers (*see* **Note 1**). Similarly, microscopic counting of bacterial numbers (e.g., using a Petroff-Hausser counting chamber) is impractical for staphylococci due to the fact that they often form irreducible clusters. This leaves spectrophotometry as the method of choice for monitoring the real-time growth of staphylococci.

In principle, spectrophotometers pass a monochromatic beam of light, with an intensity of I_o, through a gas, solution, or solid having a defined path length (l) and then record the intensity of the light (I) that emerges from the sample. Under these conditions, transmission (T) = I/I_o and the attenuation of the beam are dependent upon two independent properties: the frequency of the light and the nature of the sample. When monitoring bacterial growth using a spectrophotometer, the nature of the sample is such that there is minimal absorbance of light, but there is scattering of the light. The net result of light scattering in a spectrophotometer is that the intensity of the beam is attenuated due to the redirection of light away from the detector, not due to being absorbed by the sample. To accurately reflect the beam attenuation, it is essential to adequately dilute the sample so as to avoid multiple scattering events. The degree to which a sample must be diluted is dependent upon the wavelength and the size and shape of the scattering particles. In addition, instrument limitations must be taken into account so as to maintain the concentration (turbidity) within the linear range of the spectrophotometer. For each spectrophotometer and wavelength, it is appropriate to empirically determine the dilutions necessary to maintain the absorbance (aka, optical density; O.D.) within the linear range of the spectrophotometer. In our laboratories, we typically measure the O.D. using a 600 nm wavelength in a 1 cm path length cuvette. As the culture turbidity approaches the maximum linear range of a spectrophotometer (this is usually between 0.8 and 1.0 $O.D._{600}$ units), the samples are diluted into water or culture medium. Typically, these dilutions are between 1:5 and 1:50 and are done to maintain the $O.D._{600}$ below about 0.6. In rich culture media, the optical density at a wavelength of 600 nm for a culture of *S. epidermidis* will rarely achieve 20 $O.D._{600}$ units; hence, it is unlikely that dilutions greater than 1:50 will be necessary.

1.6 Sample Preparation

NMR metabolomic sample preparation requires three things: speed, quantity, and consistency. *Speed* is required to capture an accurate "snapshot" of the metabolome, particularly, unstable metabolites, such as those having high energy phosphate bonds (e.g., acetyl-phosphate) may experience concentration changes if the sample preparation time is too long. The greater the *quantity* of metabolites harvested, the stronger the NMR signal, and the greater the likelihood that metabolites of lower abundance will be identified. *Consistency* is essential due to the sensitivity of the NMR

to artificial variation. Inconsistency in sample preparation may cause artificial differences that will be reflected in statistical analyses (e.g., principal component analysis (PCA)) or that may obscure small but significant metabolic differences. While there are numerous methods of preparing samples for NMR metabolomic analysis, a number of recent reviews on these methods are available [12, 13]; therefore, we will only describe the technique most commonly used in our laboratories in Subheading 3.

1.7 NMR Data Collection

There are two general goals of an NMR metabolomics study: The first is to capture the global state of the metabolome (metabolic fingerprinting), which is typically accomplished using a one-dimensional (1D) ^{1}H NMR experiment. The second primary goal is to identify the metabolites that differentiate two or more bacteria or bacterial growth conditions, specifically, which metabolites are undergoing a significant concentration change or are completely missing in a particular bacterial type or culture condition. In our laboratories, this is typically accomplished by using two-dimensional (2D) ^{1}H–^{13}C heteronuclear single-quantum coherence (HSQC) NMR experiments. In the following sections, we discuss some important considerations for each type of experiment.

1.7.1 1D ^{1}H NMR

The 1D ^{1}H NMR spectrum contains hundreds to thousands of spectral lines or peaks, where each peak arises from a specific hydrogen nucleus (or proton in NMR parlance) from each metabolite, buffer, or solvent molecule present in the sample (Fig. 3). To minimize obscuring metabolite peaks by buffer or solvent peaks, a PBS buffer made with deuterium oxide (D_2O) is used as the solvent for the NMR metabolomic samples (*see* **Note 2**). Despite using D_2O, there are still some residual H_2O peaks that need to be removed using solvent suppression. To achieve this, we incorporate excitation sculpting into our 1D ^{1}H pulse sequence to remove the water signal and maintain a flat spectral baseline [14]. A common alternative approach employs a 1D ^{1}H NOESY experiment, described in [15]. In addition to water, the presence of proteins or other biomolecules in the sample may also obscure the metabolite peaks. Large-molecular-weight proteins usually cause an underlying broadness to the NMR spectrum and a distorted "smile" shape to the baseline. Any protein contamination can be chemically removed by precipitation with methanol or other equivalent organic solvents; however, this may result in the unintended loss of metabolites. Alternatively, the NMR pulse sequence can be modified by using a Carr–Purcell–Meiboom–Gill (CPMG) spin-echo system [16]. Simply, the CPMG pulse sequence selects for metabolites over protein signals based on the large-molecular-weight difference.

After the removal of water and/or protein peaks, metabolite peaks are assigned chemical shifts that typically range from 0 to 10.0 ppm, which depends on the unique chemical environment of each proton. Maintaining a consistent chemical shift assignment between multiple

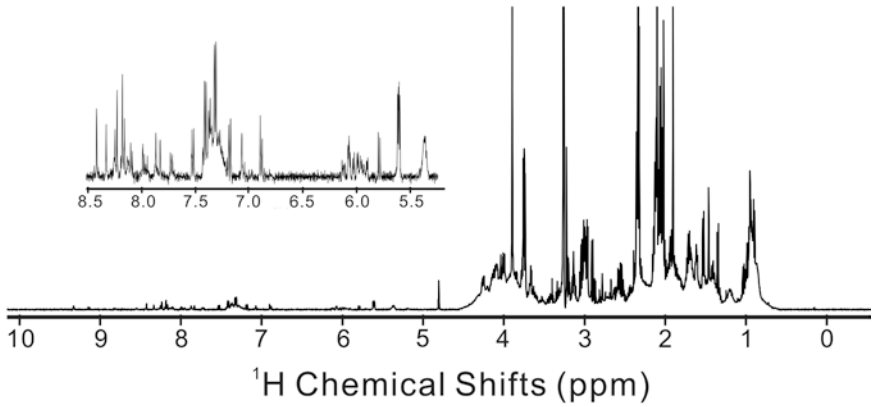

Fig. 3 Illustration of a typical one-dimensional ^{1}H NMR spectrum of an *S. epidermidis* strain 1457 cell-free lysate. The *inset* is an enhanced view of the aromatic region where the peak intensity has been increased by a factor of 12

samples requires using an internal standard. In our laboratories, we add 50 μM 3-(trimethylsilyl)propionic acid-2,2,3,3-d4 (TMSP-d4) to the NMR buffer and assign the single methyl peak to 0 ppm. To further minimize any instrument variation and bias, the NMR spectra are collected in an automated fashion using a robotic sample changer and software that automates sample tuning, matching, shimming and locking, and data collection. Furthermore, replicate samples from each class are randomly interleaved to avoid the introduction of any unintended bias into the spectral data.

1.7.2 Two-Dimensional ^{1}H–^{13}C HSQC

As stated above, we routinely use (2D) ^{1}H–^{13}C HSQC to identify metabolites and to determine the concentrations of those metabolites. Alternatively, 2D ^{1}H-^{1}H total correlation spectroscopy (TOCSY) [17] experiments, and statistical methods such as statistical total correlation spectroscopy (STOCSY) [18] are commonly used for metabolite identification (Fig. 4). The 2D ^{1}H–^{13}C HSQC NMR experiment generally requires including a ^{13}C-labeled metabolite in the culture medium; hence, 2D ^{1}H–^{13}C HSQC NMR monitors the flow of carbon-13 through the metabolome, and only metabolites derived from the ^{13}C-labeled metabolite will be observed in the 2D ^{1}H–^{13}C HSQC spectrum. Since 2D ^{1}H–^{13}C HSQC experiments may require an hour or longer to acquire a spectrum compared to approximately 5 min for 1D ^{1}H NMR experiments, the spectrum is only collected in triplicate. To minimize data acquisition time in 2D NMR experiments, sample sensitivity can be increased by increasing the number of bacteria harvested.

1.8 NMR Data Analysis

An NMR metabolomics study attempts to understand the changes in the cellular metabolome resulting from environmental stress, drug treatment, different bacterial strains or mutations, or any number of experimental conditions or biological systems. The NMR spectrum captures the state of the metabolome in order to make

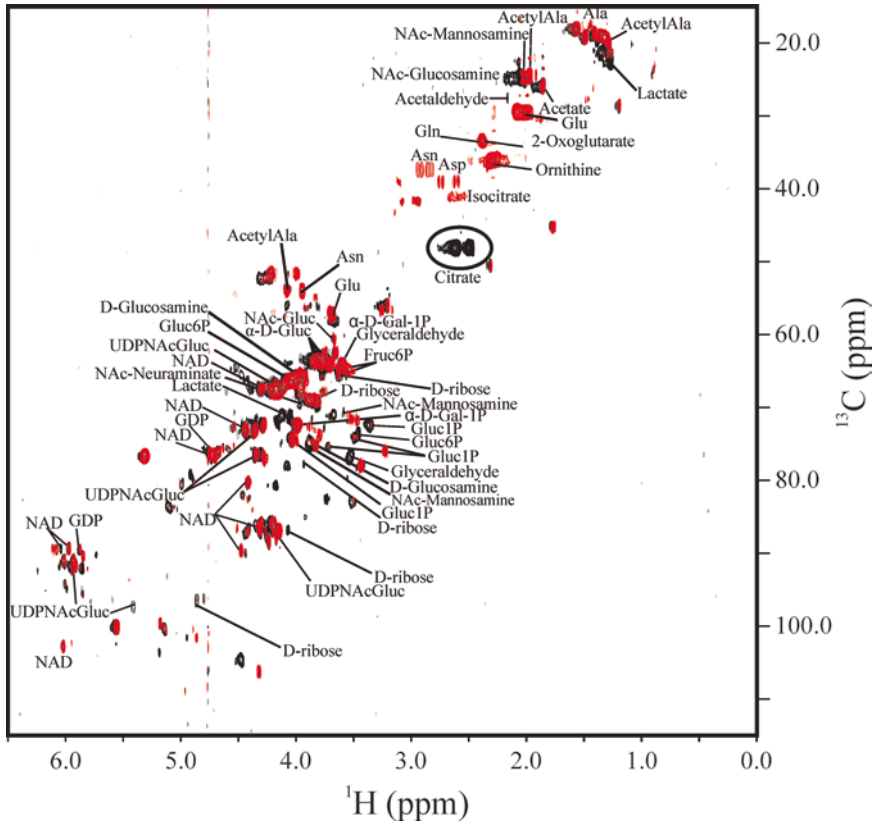

Fig. 4 Overlay of 2D ^{1}H–^{13}C HSQC spectra comparing *S. epidermidis* strain 1457 (*red*) and an isogenic aconitase mutant 1457-*acnA*::*tet*M (*black*) grown for 6 h in TSB medium containing 0.25 % ^{13}C-glucose. NMR resonances corresponding to specific metabolites are labeled, where citrate is *circled*. Reprinted with permission from Zhang B, Halouska S, Schiaffo CE, Sadykov MR, Somerville GA, Powers R: NMR analysis of a stress response metabolic signaling network. J Proteome Res 2011, 10:3743–3754 [49]. Copyright 2011 American Chemical Society

these comparisons. Thus, it is essential that the observed changes in the NMR spectrum are biologically relevant and not a result of sample preparation, data collection, or data processing. Correspondingly, consistency is critical to a metabolomics study. All sample preparation and NMR data collection should follow the identical protocol, use the same equipment and chemicals, and be conducted by the same individual at the same time, within reason. Similarly, uniform pre-processing of the NMR data is essential in order to obtain reliable and interpretable results from multivariate analysis techniques.

Pre-processing of the NMR spectra consists of (1) binning or aligning, (2) normalization, (3) scaling, (4) and noise removal. These issues have been previously discussed in detail (*see* refs. 19–22) and are only briefly summarized here. Chemical shifts or the spectral lines in an NMR spectrum are extremely sensitive to experimental conditions. As a result, it is critical that multiple NMR spectra are properly aligned so we are monitoring changes in the same peaks across the entire set of NMR spectra. The most common

approach, and the method we employ, to solving this issue is the use of "adaptive" or "intelligent" binning [23–26]. The NMR spectrum is divided into a series of bins with widths ranging from 0.025 to 0.040 ppm. The bin edges are selected in a manner to avoid splitting peaks between bins. While binning ensures that we are comparing identical peaks between spectra, each spectrum must also be normalized to prevent the introduction of false concentration differences. To normalize each NMR spectrum, we use bacterial growth (*see* Subheading 1.5) to confirm that each sample has the same number of bacteria. After global normalization, there is still considerable variability in NMR peak intensities, even within a group of replicate samples, due to experimental variation. We address this problem by center averaging (also known as standard normal variate (SVN) or a Z-score) each spectrum that emphasizes correlations [22]:

$$Z = \frac{x_i - \bar{x}}{\sigma}$$

where $\bar{x}$ is the average signal intensity, σ is the standard deviation in the signal intensity, and x_i is the signal intensity within the bin. The next step of the data pre-processing is to remove noise, solvent, or buffer regions [20]. Again, normalization removes biologically irrelevant variations between samples due to differences in the total sample concentration.

The pre-processed NMR data is then interpreted using multivariate statistical techniques. We primarily use PCA and orthogonal partial least-squares discriminant analysis (OPLS-DA) [27–29]. OPLS-DA and PCA are performed using SIMCA-P11.5+ (UMETRICS), which automates scaling of the bins (columns) between each sample (row). SIMCA-P uses unit variance by default, but there are a variety of scaling protocols available depending on the specific goals of the metabolomics study. Scaling diminishes irrelevant contributions of intense peaks to the multivariate statistical analysis. Simply, relatively small percentage changes in intense peaks would dominate large percentage changes in weak peaks in the absence of scaling. This occurs because PCA and OPLS-DA simply emphasize absolute changes in peak intensities. The primary outcome of PCA and OPLS-DA is a scores plot, where each NMR spectrum is represented as a single point in PC space (Fig. 5). Briefly, each NMR spectrum is reduced to a single point in multidimensional space, where each axis is a bin and the value along the axis is the bin integral. Each NMR spectrum is then fitted to a vector $\left(\overrightarrow{PC_1}\right)$ corresponding to the largest variation in the data and then to a second vector orthogonal $\left(\overrightarrow{PC_2}\right)$ to the first. The 2D scores plot represents the projection of each spectra into the hyperplane described by $\left(\overrightarrow{PC_1}\right)$ and $\left(\overrightarrow{PC_2}\right)$. The relative clustering in the resulting scores plot identifies the relative similarities or differences between the 1D ^{1}H NMR spectrum and, correspondingly, the metabolomes. Thus, biological significance is

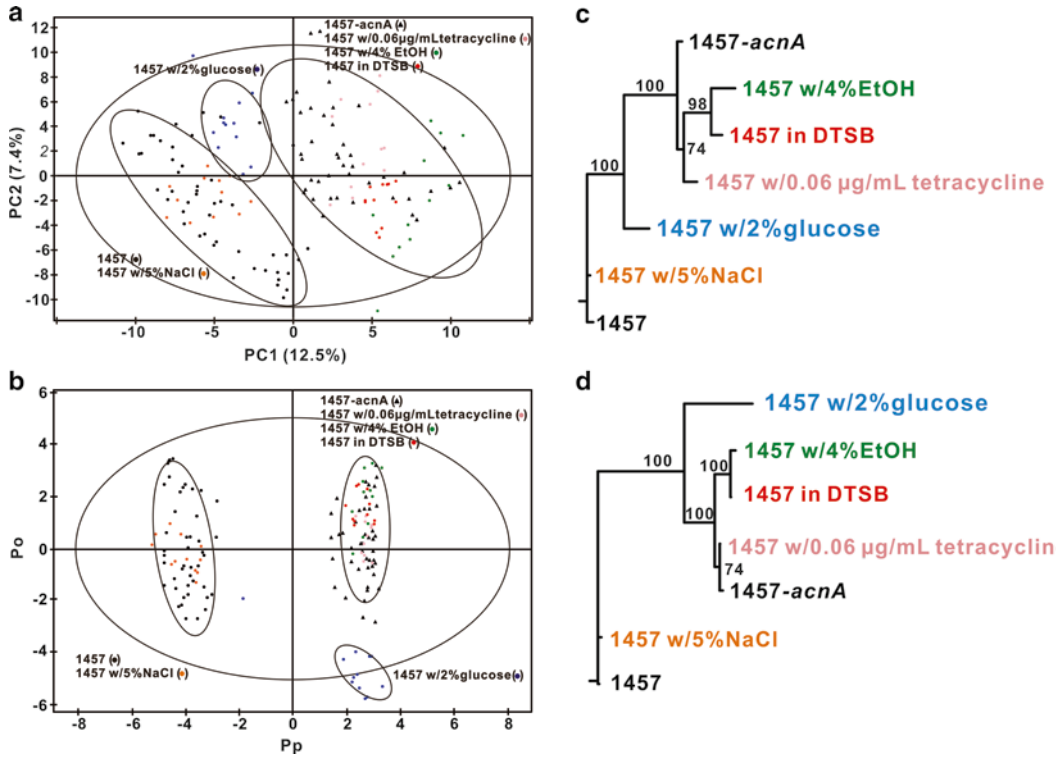

Fig. 5 Statistical analysis of *S. epidermidis* metabolomes prepared from stressed cultures aids in assessing the relationship between the stressors. (**a**) 2D PCA scores plot and (**b**) 2D OPLS-DA scores plot comparing *S. epidermidis* strain 1457 grown for 6 h in TSB medium (*black-filled circle*), with *S. epidermidis* 1457 cells grown 6 h in iron-depleted media (DTSB) (*red-filled circle*), with the addition of 4 % ethanol (*green-filled circle*),with the addition of 2 % glucose (*blue-filled circle*), with the addition of 0.06 µg/mL tetracycline (*pink-filled circle*), and with the addition of 5 % NaCl (*orange-filled circle*), and 6-h growth of aconitase mutant strain 1457-*acnA*::*tet*M in standard TSB medium (*black-filled triangle*). The ellipses correspond to the 95 % confidence limits from a normal distribution for each cluster. For the OPLS-DA scores plot, the 6-h growth of wild-type *S. epidermidis* 1457 (*black-filled circle*) was designated the control class and the remainder of the cultures were designated as treated. The OPLS-DA used 1 predictive component and 4 orthogonal components to yield a R^2X of 0.637, R^2Y of 0.966, and Q^2 of 0.941. CV-ANOVA validation yielded an effective *p*-value of 0.0. Metabolomic tree diagram generated from the (**c**) 2D PCA scores plot depicted in (**a**) and (**d**) 2D OPLS-DA scores plot depicted in (**b**). The *label colors* match the *symbol colors* from the 2D scores plot. Each node is labeled with the boot-strap number, where a value above 50 indicates a statistically significant separation. Reprinted with permission from [49], copyright 2011 American Chemical Society

inferred by the relative clustering or class discrimination in the PCA or the OPLS-DA scores plot. To infer statistical relevance to the clustering pattern, we use both an ellipse that defines the 95 % confidence limit from a normal distribution for each class and hierarchical clustering similar in concept to phylogenetic tree diagrams (Fig. 5) [30, 31]. The hierarchal clustering is based on the distance in the scores plot between each defined cluster or class. An updated version of our PCAtoTree software package is distributed under version 3.0 of the GNU General Public License and is freely available at http://bionmr.unl.edu/pca-utils.php.

In contrast to PCA, OPLS-DA is a supervised technique in which each NMR spectrum is typically assigned to one of the two binary classes, control (0) and treated (1). The interpretation is thus "biased" by the class definition and requires the OPLS-DA model to be validated. Essentially, OPLS-DA will identify class separation even for random data [32]. The identification of a similar class separation between the PCA and OPLS-DA models supports the reliability of the OPLS-DA model, but it is not sufficient for model validation. We use the leave-*n*-out method [33, 34] that reports quality assessment (Q^2) and quality of fit (R^2) values. These statistics simply reflect the consistency between the data and the model. A Q^2 value of ≥ 0.4 infers an acceptable model for a biological system [35], but large Q^2 value can also result from invalid or irrelevant models; therefore, we also use CV-ANOVA [36] to further validate the model with a standard *p*-value.

After identifying a class separation based on the PCA or the OPLS-DA scores plot, the next goal is to identify which metabolites primarily contribute to the class distinction. One common approach we use is to identify significant metabolite concentration changes by using the time-zero HSQC experiment [37]. The 2D ^{1}H–^{13}C HSQC NMR correlates carbon and hydrogen chemical shifts for every bonded C–H pair. The higher spectral resolution (2D vs. 1D) and the multiple chemical shift information (^{1}H and ^{13}C) significantly improve the reliability of identifying metabolites. Simply, the experimental chemical shifts obtained from the 2D ^{1}H–^{13}C HSQC spectrum are matched against reference spectral data in databases such as the Madison Metabolomics Consortium Database [38], the BioMagResBank [39], and the Human Metabolome Database (HMDB) [40]. There are two practical challenges: (1) Metabolomics is a relatively new field, so the databases are incomplete, and (2) ^{1}H and ^{13}C chemical shift tolerances of 0.05 and 0.50 ppm, respectively, are used to identify a match between the experimental and database chemical shifts because of experimental variation, potentially leading to multiple possible matches. For these reasons, it is important to incorporate biological information to refine the metabolite assignments.

Other important outcomes of OPLS-DA are an S-plot and loadings plot (Fig. 6). An S-plot identifies the chemical shift

Fig. 6 (continued) Each point was identified to a specific metabolite using the Human Metabolomics Database and Madison Metabolomics Database. All the identified metabolites are associated with TCA cycle inactivation. (**b**) OPLS-DA S-plot comparing the mutant strain 1457-*acnA*::*tetM* grown for 2 and 6 h. The metabolites identified are associated with variations in the utilization of glucose. (**c**) OPLS-DA loading plot comparing *S. epidermidis* 1457 and aconitase mutant strain 1457-*acnA*::*tetM* where both cell cultures were grown for 6 h. Negative values indicate a decrease in peak intensity when comparing the wild-type to the aconitase mutant, while positive values indicate an increase in peak intensity. (**d**) OPLS-DA loading plot comparing the mutant strain 1457-*acnA*::*tetM* grown for 2 and 6 h. Reprinted with permission from [49], copyright 2011 American Chemical Society

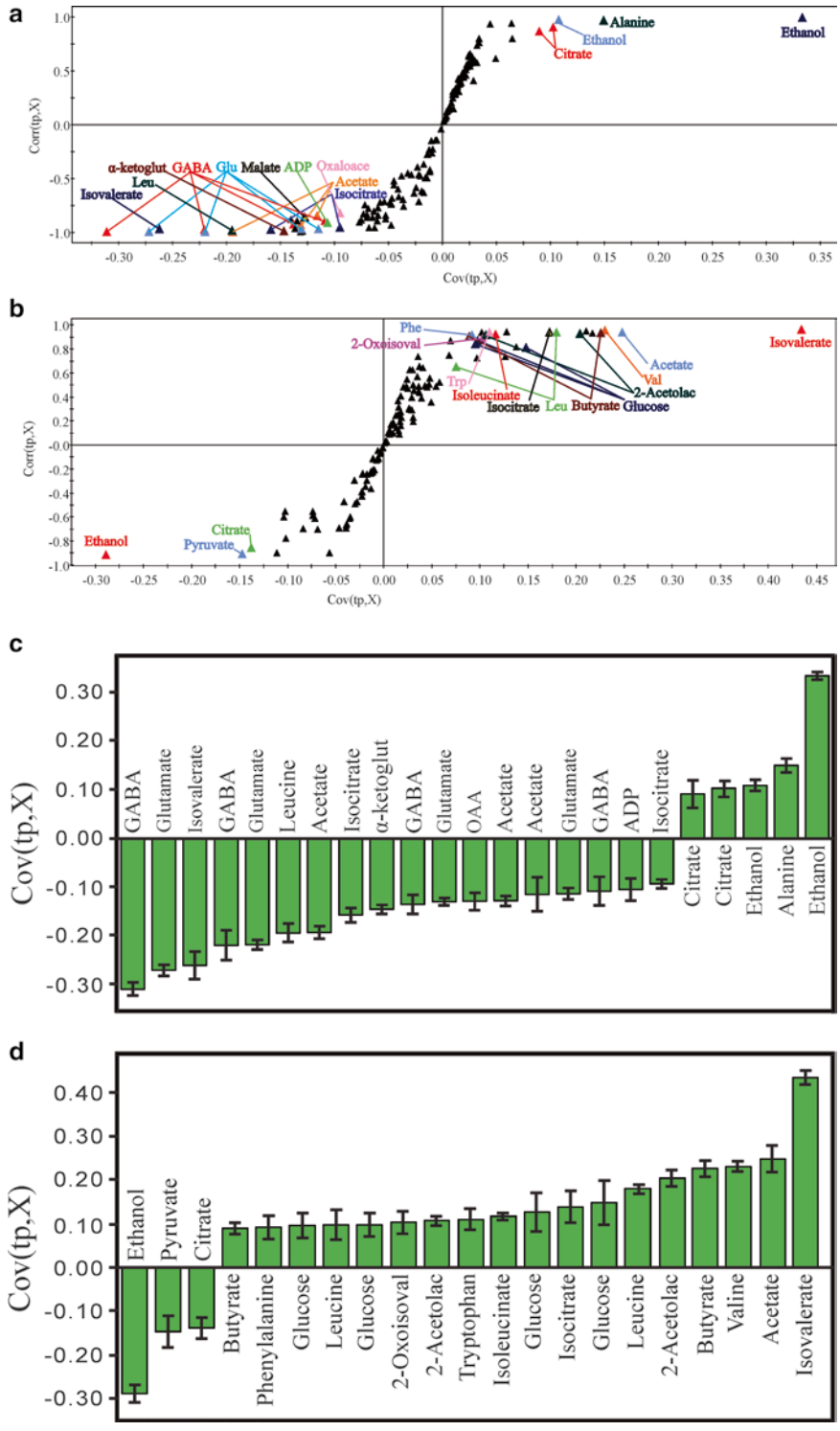

Fig. 6 (**a**) OPLS-DA S-plot comparing the post-exponential (6 h) metabolomes of *S. epidermidis* strain 1457 and its isogenic aconitase mutant (1457-*acnA*::*tetM*). The metabolomes were shown to be separated along the PC1 axis in Fig. 5a. Each point in the S-plot represents a specific bin containing a chemical shift range of 0.025 ppm, where the points at the extreme ends of the S-plot are the major contributors to the class distinction.

bins and, correspondingly, the metabolites that are the major contributors to the class discrimination. Similarly, the loadings plot identifies the magnitude of each bin's contribution to class discrimination, which can be interpreted as a relative measure of metabolite concentration changes. Typically, the metabolites identified from the S-plot are consistent with the major changes observed in the 2D ^{1}H–^{13}C HSQC spectrum. In fact, the 2D ^{1}H–^{13}C HSQC spectrum is used to aid the assignments of metabolites from the S-plots because of the large overlap in 1D ^{1}H NMR spectra. Nevertheless, the results from the S-plots and 2D ^{1}H–^{13}C HSQC spectrum need not be identical since the 2D ^{1}H–^{13}C HSQC experiments only monitor a subset of the metabolome defined by the ^{13}C-labeled metabolite used in the culture medium. Conversely, the 1D ^{1}H NMR spectra capture all the major metabolites (>10 μM) present in the metabolome.

2 Materials

2.1 Growth Medium and Cultivation Devices

1. Bacterial strain: *S. epidermidis* strain 1457 [41].
2. $^{13}C_6$-glucose (Cambridge Isotope Laboratories, Andover, MA, USA) or non-labeled glucose (0.25 % w/v) is added to bacto tryptic soy broth without dextrose (BD, Franklin Lakes, NJ, USA). The dry culture medium plus glucose is dissolved in ultrapure water (18.2 MΩ resistance) and filter sterilized using a 0.22 μm rapid-flow sterile disposable filter unit (*see* **Note 3**).
3. *S. epidermidis* are cultivated in Nalgene polypropylene Erlenmeyer flasks or in 14 mL disposable culture tubes.
4. Shaking incubator.
5. Spectrophotometer to monitor growth.

2.2 Reagents and Devices for Metabolomic Sample Preparation

1. 20 mM phosphate buffer, pH 7.2–7.4.

For 500 mL:

(a) 1.045 g Dipotassium hydrogen phosphate.

(b) 0.527 g Potassium dihydrogen phosphate.

Dissolve salts in ultrapure water (18.2 MΩ resistance), and adjust volume to 500 mL.

2. Millipore Microfil V filtration devices (EMD Millipore, Billerica, MA, USA), filtration manifold (Thermo Fisher Scientific), vacuum pump (KNF Neuberger, Inc., Trenton, NJ, USA), and filter forceps.
3. Corning 50 mL graduated plastic centrifuge tubes (Corning, Corning, NY, USA).
4. Liquid nitrogen and ice.

5. FastPrep-24 instrument and FastPrep lysing matrix B tubes (MP Biomedicals, Solon, OH, USA).
6. Refrigerated microcentrifuge and 1.5 and 2.0 mL microcentrifuge tubes.
7. 20 mM phosphate buffer in 100 % D_2O (Sigma-Aldrich) at pH 7.2 (uncorrected).
8. Lyophilizer.
9. TMSP-d4.

2.3 NMR Instrumentation

1. Bruker 500-MHz Avance spectrometer equipped with a triple-resonance, z-axis gradient cryoprobe and an automatic tune and match (ATM) system.
2. BACS-120 sample changer with Bruker Icon software.

2.4 Data Analysis Software

1. SIMCA 11.5+ (UMETRICS) for PCA and OPLS-DA.
2. NMRViewJ [42] and Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco) for NMR processing and analysis.
3. Metabominer, Madison Metabolomics Consortium Database (MMCD), the BioMagResBank (BMRB), and HMDB were used to identify metabolites [38–40, 43].
4. KEGG and Metacyc databases were used to verify metabolites and metabolic pathways [44, 45].
5. Cytoscape and Metscape were used to generate metabolic network maps [46, 47].

3 Methods

3.1 Bacterial Growth

1. Two days prior to the planned experiment, streak a TSB agar plate with *S. epidermidis* from a glycerol stock stored at –80 °C. Incubate overnight at 37 °C.
2. Inoculate a 14 mL disposable culture tube containing 2 mL of TSB with a single colony of *S. epidermidis* (*see* **Note 4**) and incubate at 37 °C on a 45° angle with 225 rpm of agitation. Cultivate overnight. This will be the culture used for starting the pre-culture.
3. Inoculate 25 mL of pre-warmed TSB, containing 0.25 % (w/v) glucose, in a 250 mL flask with 250 µl of the overnight culture from **step 2** in Subheading 3.1 and incubate at 37 °C with 225 rpm of agitation. Cultivate for 1.5–2 h. This will be the culture used for starting the primary cultures (*see* **Note 5**).
4. While the pre-culture is incubating, prepare the primary cultures for inoculation. Typically, we harvest 20 $O.D._{600}$ units for each growth condition, strain, or growth phase; therefore, the

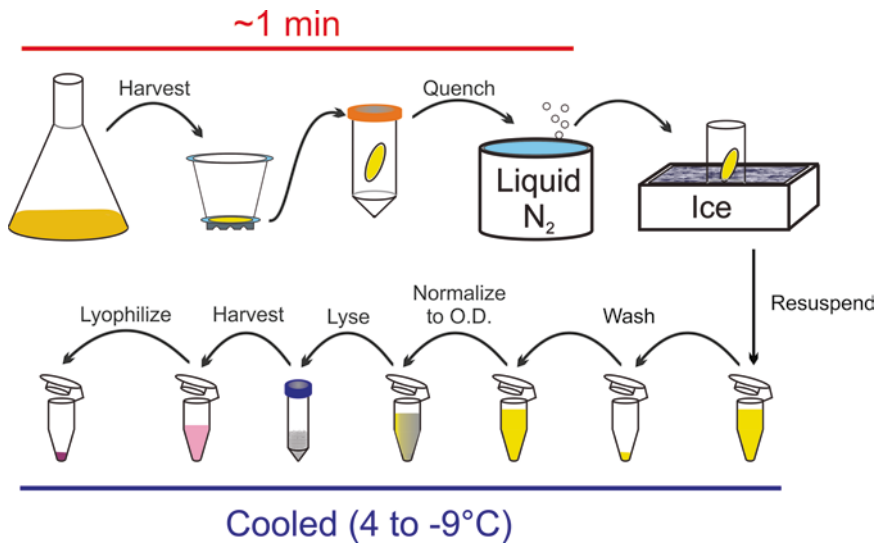

Fig. 7 A schematic representation of the sampling and preparation of *S. epidermidis* cell-free lysates for use in NMR metabolomics

volume of the primary culture will depend on the achievable optical density for the specific circumstances (*see* **Note 6**). Choose an appropriately sized flask to accommodate the volume of culture medium so as to maintain the correct flask-to-medium ratio. Add culture medium to the flasks, and pre-warm the medium to 37 °C (*see* **Note 7**).

5. After ~2 h of incubating the pre-culture (*see* **step 3** in Subheading 3.1), remove an aliquot and determine the $O.D._{600}$. Normalize the inocula so as to have a uniform starting $O.D._{600}$ of 0.06 in the primary cultures (*see* **Note 8**). Inoculate the pre-warmed primary cultures with the appropriate volume of pre-culture, swirl the flasks, remove a 1 mL sample and place these into individual microcentrifuge tubes, put primary cultures into the incubator, start agitation, and then verify that the starting optical densities were 0.06 ± 0.02 $O.D._{600}$ units.

3.2 Harvest (Summarized in Fig. 7)

1. Assemble the filtration apparatus, and pre-wash the filter with the non-deuterated phosphate buffer (*see* **item 1** in Subheading 2.2).
2. Harvest the volume of culture that is predicted to yield slightly (10 %) more than 20 $O.D._{600}$ units. At this stage, speed is more important than accuracy. Harvesting more than is required will compensate for bacterial loss during the wash steps. In addition, the optical density will be normalized in subsequent steps.
3. Disassemble the Millipore Microfil V filtration device, and, using filter forceps, transfer the filter to a Corning 50 mL graduated tube (precooled to −70 °C). Close the 50 mL tube and

immediately place it into a Dewar containing liquid nitrogen (*see* **Note 9**). Samples can remain in liquid nitrogen as other samples are being processed.

4. Transfer the Corning 50 mL graduated tubes to an ice bucket, and wash/scrape the bacteria off the filter using 1 mL of ice-cold phosphate buffer (*see* **item 1** in Subheading 2.2) by repeatedly pipetting the buffer over the filter.
5. Transfer the bacterial suspension to a 1.5 mL microcentrifuge tube (precooled to −20 °C), place the microcentrifuge tubes into a microcentrifuge precooled to −9 °C, and centrifuge for 1 min at 16,100 RCF.
6. After 1 min, place the microcentrifuge tubes into an ice bucket, and then remove and discard the supernatants. Suspend the cell pellet in 1 mL of ice-cold phosphate buffer and hold on ice (*see* **Note 10**).
7. Remove a 5 μl aliquot and add this to 995 μl of phosphate buffer (room temperature) in a disposable spectrophotometer cuvette, mix, and measure the $O.D._{600}$. Normalize the optical densities of the samples using ice-cold phosphate buffer such that all samples contain 20 $O.D._{600}$ units and have a 1 mL final volume.
8. Transfer the samples to FastPrep lysing matrix B tubes that have been cooled to −20 °C, and place the tubes in a FastPrep instrument. Lyse the bacteria at a speed of 6 for 40 s. Remove the tubes and place in an ice bucket to rest for 5 min, and then repeat the lysing step.
9. Centrifuge the lysing matrix B tubes for 2 min at 16,100 RCF in a microcentrifuge precooled to −9 °C.
10. Transfer 0.7 mL of the supernatant (cell-free lysate) to a pre-cooled (−20 °C) 2 mL microcentrifuge tube.
11. Add 1 mL of ice-cold phosphate buffer to each lysing matrix B tube, and suspend the pellet containing the glass beads and cell debris by repeated pipetting.
12. Centrifuge the lysing matrix B tubes for 2 min at 16,100 RCF in a microcentrifuge precooled to −9 °C.
13. Transfer 0.9 mL of the supernatant to the appropriate tubes from **step 10** in Subheading 3.2. The pooled volume of cell-free lysate should be 1.6 mL.
14. Centrifuge the pooled cell-free lysates for 1 min at 16,100 RCF in a microcentrifuge precooled to −9 °C.
15. Transfer 1.5 mL of the cell-free lysates to a new 2 mL micro-centrifuge tubes that have been pre-chilled at −20 °C, and then freeze the samples in dry ice. This step will remove any glass beads and/or cell debris that were carried over from the lysing matrix B tubes.
16. Lyophilize the cell-free lysates at −50 °C.

3.3 Preparation and Data Acquisition of Metabolomic Samples by NMR

1. For 1D ^{1}H NMR, the lyophilized cell-free lysates (*see* **step 16** in Subheading 3.2) are suspended in 600 μl of 50 mM phosphate buffer (PBS) in 99.8 % D_2O at pH 7.2 (uncorrected), containing 50 μM TMSP-d4, which is transferred to an NMR tube. Typically, ten replicate samples are prepared for each class or cell condition to maximize statistical significance.
2. Our 1D ^{1}H NMR spectra are routinely collected on a Bruker 500 MHz Avance DRX spectrometer equipped with a cryoprobe to maximize sensitivity and using BACS-120 sample changer and auto-tune and match (ATM) for automation. 1D ^{1}H NMR spectra are collected at 298 K with a spectrum width of 5,482.5 Hz and 32 K data points. A total of 16 dummy scans and 64 scans are used to obtain each spectrum.
3. As with the 1D ^{1}H experiments, the 2D ^{1}H–^{13}C HSQC spectra are collected in an automated fashion on a Bruker 500 MHz Avance DRX spectrometer with a cryoprobe. The 2D ^{1}H–^{13}C HSQC spectra are collected with solvent presaturation and a relaxation delay of 0.5 s. A total of 2,048 data points with a spectrum width of 4,734.9 Hz and 128 data points with a spectrum width of 13,834.3 Hz are collected in the ^{1}H and ^{13}C dimensions, respectively. A total of 8 dummy scans and 128 scans are used to obtain each of the 2D ^{1}H–^{13}C HSQC NMR spectra.

3.4 NMR Metabolomic Data Analysis

Due to the wide array of methods, protocols, and software available for NMR data analysis, we have not included a detailed protocol describing our approach to data analysis. Instead, we have included in the introductory section some general considerations to keep in mind when doing NMR data analysis (*see* Subheading 1.8).

4 Notes

1. When assessing bacterial numbers, it is important to remember that a certain percentage of bacteria will be nonviable or dead, leading to an overestimation in the number of live bacteria. Usually, this does not create a significant obstacle unless the bacteria are being harvested in the stationary or the death phases or when using mutants that affect cell growth, division, and/or autolysis. Under circumstances where viability may be compromised, the percentage of live bacteria must be determined using methods such as live–dead stains.
2. PBS made with D_2O is used as the solvent because there are no observable hydrogen nuclei in the buffer due to the fact that the majority of hydrogen atoms have been replaced by deuterium.
3. Filter sterilizing culture media prevents pyrolysis of heat labile amino acids and carbohydrates, such as that which occurs

during autoclaving. Pyrolysis is particularly problematic when differences are introduced into culture media autoclaving procedures, specifically, differences in heat exposure time and load density. For these reasons, our experience is that filter sterilization increases the consistency of growth profiles and produces a more uniform metabolome sample.

4. As staphylococci are prone to acquiring mutations during passage and storage [48], use colonies from freshly streaked plates. Using colonies from plates that have been stored at room temperature or 4 °C for more than a few days increases the risk that adaptive mutagenesis will have induced undesirable mutations.
5. By using a pre-culture that is in the exponential growth phase (i.e., has been growing for only 1.5–2 h), the lag phase will be minimal because the bacteria reached their initiation mass during the pre-culture.
6. Prior to initiating metabolomic studies, growth rates and optical densities must be determined empirically for each growth condition, strain, or growth phase. With this information it is possible to estimate the volume of culture that will need to be harvested; however, the actual number of $O.D._{600}$ units used in preparing the metabolome samples will be adjusted based on the optical density. As an example, if we want to examine the metabolome of *S. epidermidis* strain 1457 during the exponential growth phase and we have determined that strain 1457 can achieve an $O.D._{600}$ of 0.4 (Fig. 2), then it is necessary to harvest 50 mL to reach 20 $O.D._{600}$ units.
7. Pre-warming the culture medium will prevent a temperature shock during inoculation.
8. To prevent the introduction of "too much" non-labeled glucose into the primary culture medium, it may be preferable to concentrate the bacteria by centrifugation and suspend in a smaller volume of pre-warmed medium.
9. As discussed, speed is critical for achieving an accurate representation of the metabolome; hence, **steps 2** and **3** in Subheading 3.2 combined should take about 1 min.
10. Washing the bacteria is essential for the removal of metabolites found in the culture medium, which can introduce artifacts into the NMR spectra.

Acknowledgements

G.A.S. and R.P. were supported by funds provided through the National Institutes of Health (AI087668). R.P. was also supported by funds through the National Institutes of Health National Center for Research Resources (P20 RR-17675) and the American Heart

Association (0860033Z). The research was performed in facilities renovated with support from the National Institutes of Health (NIH, RR015468-01). We are grateful to our past and present students and postdoctoral fellows for their hard work and assistance—thanks, for all that you do!

References

1. Neidhardt FC (2006) Apples, oranges and unknown fruit. Nat Rev Microbiol 4:876
2. Somerville GA, Proctor RA (2009) At the crossroads of bacterial metabolism and virulence factor synthesis in Staphylococci. Microbiol Mol Biol Rev 73:233–248
3. Neidhardt FC, Bloch PL, Smith DF (1974) Culture medium for enterobacteria. J Bacteriol 119:736–747
4. Hussain M, Hastings JG, White PJ (1991) A chemically defined medium for slime production by coagulase-negative staphylococci. J Med Microbiol 34:143–147
5. Novick RP (1991) Genetic systems in staphylococci. Methods Enzymol 204:587–636
6. Sadykov MR, Olson ME, Halouska S et al (2008) Tricarboxylic acid cycle-dependent regulation of *Staphylococcus epidermidis* polysaccharide intercellular adhesin synthesis. J Bacteriol 190:7621–7632
7. Sadykov MR, Zhang B, Halouska S et al (2010) Using NMR metabolomics to investigate tricarboxylic acid cycle dependent signal transduction in *Staphylococcus epidermidis*. J Biol Chem 285:36616–36624
8. Somerville GA, Said-Salim B, Wickman JM et al (2003) Correlation of acetate catabolism and growth yield in *Staphylococcus aureus*: implications for host-pathogen interactions. Infect Immun 71:4724–4732
9. Sadykov MR, Hartmann T, Mattes TA et al (2011) CcpA coordinates central metabolism and biofilm formation in *Staphylococcus epidermidis*. Microbiology 157:3458–3468
10. Krebs HA (1972) The Pasteur effect and the relations between respiration and fermentation. Essays Biochem 8:1–34
11. Miles AA, Misra SS, Irwin JO (1938) The estimation of the bactericidal power of the blood. J Hyg (Lond) 38:732–749
12. van Gulik WM (2010) Fast sampling for quantitative microbial metabolomics. Curr Opin Biotechnol 21:27–34
13. Zhang B, Powers R (2012) Analysis of bacterial biofilms using NMR-based metabolomics. Future Med Chem 4:1273–1306
14. Hwang T-L, Shaka AJ (1995) Water suppression that works. Excitation sculpting using arbitrary waveforms and pulsed field gradients. J Magn Reson A 112:275–279
15. Beckonert O, Keun HC, Ebbels TM et al (2007) Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. Nat Protoc 2:2692–2703
16. Meiboom S, Gill D (1958) Modified spin-echo method for measuring nuclear relaxation times. Rev Sci Intrum 29:4
17. Bingol K, Bruschweiler R (2011) Deconvolution of chemical mixtures with high complexity by NMR consensus trace clustering. Anal Chem 83:7412–7417
18. Sands CJ, Coen M, Ebbels TM et al (2011) Data-driven approach for metabolite relationship recovery in biological 1H NMR data sets using iterative statistical total correlation spectroscopy. Anal Chem 83:2075–2082
19. Craig A, Cloarec O, Holmes E et al (2006) Scaling and normalization effects in NMR spectroscopic metabonomic data sets. Anal Chem 78:2262–2267
20. Halouska S, Powers R (2006) Negative impact of noise on the principal component analysis of NMR data. J Magn Reson 178: 88–95
21. Sysi-Aho M, Katajamaa M, Yetukuri L et al (2007) Normalization method for metabolomics data using optimal selection of multiple internal standards. BMC Bioinformatics 8:93
22. van den Berg RA, Hoefsloot HC, Westerhuis JA et al (2006) Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics 7:142
23. Anderson PE et al (2011) Dynamic adaptive binning: an improved quantification technique for NMR spectroscopic data. Metabolomics 7:179–190
24. Anderson PE et al (2008) Gaussian binning: a new kernel-based method for processing NMR spectroscopic data for metabolomics. Metabolomics 4:261–272

25. Davis RA et al (2007) Adaptive binning: an improved binning method for metabolomics data using the undecimated wavelet transform. Chemometr Intell Lab Syst 85:144–154
26. De Meyer T, Sinnaeve D, Van Gasse B et al (2008) NMR-based characterization of metabolic alterations in hypertension using an adaptive, intelligent binning algorithm. Anal Chem 80:3783–3790
27. Bylesjo M et al (2006) OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. J Chemometr 20: 341–351
28. Barker M, Rayens W (2003) Partial least squares for discrimination. J Chemometr 17: 166–173
29. Rannar S et al (1994) A Pls Kernel algorithm for data sets with many variables and fewer objects. 1. Theory and algorithm. J Chemometr 8:111–125
30. Werth MT, Halouska S, Shortridge MD et al (2010) Analysis of metabolomic PCA data using tree diagrams. Anal Biochem 399:58–63
31. Worley B, Halouska S, Powers R (2013) Utilities for quantifying separation in PCA/PLS-DA scores plots. Anal Biochem 15:102–104
32. Kjeldahl K, Bro R (2010) Some common misunderstandings in chemometrics. J Chemometr 24:558–564
33. Golbraikh A, Tropsha A (2002) Beware of q(2)! J Mol Graph Model 20:269–276
34. Shao J (1993) Linear-model selection by cross-validation. J Am Stat Assoc 88:486–494
35. Westerhuis JA et al (2008) Assessment of PLSDA cross validation. Metabolomics 4:81–89
36. Eriksson L, Trygg J, Wold S (2008) CV-ANOVA for significance testing of PLS and OPLS (R) models. J Chemometr 22:594–600
37. Hu K, Westler WM, Markley JL (2011) Simultaneous quantification and identification of individual chemicals in metabolite mixtures by two-dimensional extrapolated time-zero (1) H-(13)C HSQC (HSQC(0)). J Am Chem Soc 133:1662–1665
38. Cui Q, Lewis IA, Hegeman AD et al (2008) Metabolite identification via the Madison Metabolomics Consortium Database. Nat Biotechnol 26:162–164
39. Ulrich EL, Akutsu H, Doreleijers JF et al (2008) BioMagResBank. Nucleic Acids Res 36:D402–D408
40. Wishart DS, Jewison T, Guo AC et al (2007) HMDB: the Human Metabolome Database. Nucleic Acids Res 35:D521–D526
41. Mack D, Siemssen N, Laufs R (1992) Parallel induction by glucose of adherence and a polysaccharide antigen specific for plastic-adherent *Staphylococcus epidermidis*: evidence for functional relation to intercellular adhesion. Infect Immun 60:2048–2057
42. Johnson BA (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol Biol 278:313–352
43. Xia J, Bjorndahl TC, Tang P et al (2008) MetaboMiner—semi-automated identification of metabolites from 2D NMR spectra of complex biofluids. BMC Bioinformatics 9:507
44. Caspi R, Altman T, Dale JM et al (2010) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 38: D473–D479
45. Kanehisa M, Araki M, Goto S et al (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res 36:D480–D484
46. Karnovsky A, Weymouth T, Hull T et al (2012) Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data. Bioinformatics 28:373–380
47. Killcoyne S, Carter GW, Smith J et al (2009) Cytoscape: a community-based framework for network modeling. Methods Mol Biol 563: 219–239
48. Somerville GA, Beres SB, Fitzgerald JR et al (2002) In vitro serial passage of *Staphylococcus aureus*: changes in physiology, virulence factor production, and *agr* nucleotide sequence. J Bacteriol 184:1430–1437
49. Zhang B, Halouska S, Schiaffo CE et al (2011) NMR analysis of a stress response metabolic signaling network. J Proteome Res 10: 3743–3754