

Published in final edited form as:

Future Med Chem. 2012 June ; 4(10): 1273–1306. doi:10.4155/fmc.12.59.

Analysis of bacterial biofilms using NMR-based metabolomics

Bo Zhang and Robert Powers*

Department of Chemistry, University of Nebraska-Lincoln, 722 Hamilton Hall, Lincoln, NE 68588-0304, USA

Abstract

Infectious diseases can be difficult to cure, especially if the pathogen forms a biofilm. After decades of extensive research into the morphology, physiology and genomics of biofilm formation, attention has recently been directed toward the analysis of the cellular metabolome in order to understand the transformation of a planktonic cell to a biofilm. Metabolomics can play an invaluable role in enhancing our understanding of the underlying biological processes related to the structure, formation and antibiotic resistance of biofilms. A systematic view of metabolic pathways or processes responsible for regulating this ‘social structure’ of microorganisms may provide critical insights into biofilm-related drug resistance and lead to novel treatments. This review will discuss the development of NMR-based metabolomics as a technology to study medically relevant biofilms. Recent advancements from case studies reviewed in this manuscript have shown the potential of metabolomics to shed light on numerous biological problems related to biofilms.

Biofilms are a natural part of the ecology of the Earth, and correspond to a ‘social structure’ of microorganisms compared with a planktonic state [1–3]. They are a ubiquitous microbial phenomenon that has been observed for bacteria, fungi, algae, yeasts, protozoa and other organisms. In a biofilm composed of multiple diverse organisms, the interspecies interaction can range from neutral to cooperative, to competitive, and finally, to antagonistic [4]. The diversity of organisms able to self-organize and form biofilms is quite astounding and may provide clues to the evolution of multicellular organisms [5]. Are biofilms a transitional state of evolution and the basis for multicellular organisms? Or are biofilms simply a highly organized state of single-cell organisms? The fact that biofilms provide a significant survival advantage for adapting to the harsh and distinct environmental conditions probably explains its broad adaptation.

How bacterial biofilms are related to human disease

The interest in biofilms is not merely a result of scientific curiosity, but is also derived from practical concerns related to medical science [6], material engineering [7], civil engineering [8] and others [9]. In the area of medicine, research on biofilms has focused on its relationship to bacterial infections and drug resistance. Bacterial infections are a serious disease and major source of deaths worldwide. Especially concerning is the growing resistance to antibiotics that has become a major medical issue in developing countries. Between 1980 and 1992, infectious disease deaths increased by 58%; the major contributors

© 2012 Future Science Ltd

*Author for correspondence: Tel.: +1 402 472 3039, Fax: +1 402 472 2044, rpowers3@unl.edu.

Financial & competing interests disclosure

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

were HIV infections and AIDS, respiratory disease and bloodstream infection [10]. In 2000, a US government report identified infectious diseases as a leading cause of death worldwide and the third leading cause of death in the USA [11]. It has been estimated that 60–80% of human microbial infections are caused by bacteria growing as a biofilm [12]. Certain pathogenic biofilms are of particular concern because of the added issue of drug resistance [13]. Multidrug-resistant pathogens, such as *Enterococcus faecium*, *Klebsiella pneumonia*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* spp., are currently infecting the majority of US hospitals [14]. As methicillin-resistant *Staphylococcus aureus* (MRSA) and other resistant pathogens capable of biofilm formation continue to emerge and propagate, understanding and circumventing biofilm resistance to antibiotics is a paramount necessity [15].

Pathogens can be introduced into the human body through trauma, medical operations, dental procedures or by other means [16–18]. Many surfaces of organs are heavily colonized by microbes that have the potential to cause an infection, especially during any invasive medical procedure. In fact, the ratio of bacteria to mammalian cells living within the human body is ten to one, providing ample opportunity for inducing a bacterial infection from medical procedures or trauma [19]. For example, there are over 500 species of microorganisms identified in typical dental plaque [301]. Correspondingly, dental cavities (caries) are commonly a result of bacterial biofilm infections [20]. Biofilms are also formed on our tongues, cheeks, in our intestines, nasal passages, sinuses and on our skin [301]. These human microbial communities are largely unstudied and their role in infections is largely unknown. But biofilms protect the organisms from both antimicrobials and the host immune response, making infectious biofilms extremely difficult to treat [21]. For instance, staphylococcal biofilm infections have a 10–1000-fold increase in antibiotic resistance [22–24].

A serious source of biofilm infections is heart disease, which is also a major cause of mortality in the USA [25]. Invasive surgical techniques are inevitably required to treat the resulting symptoms of heart disease, which may lead to fatal staphylococcal-infective endocarditis. *S. aureus* and *S. epidermidis* [26] infections stemming from implantable medical devices (e.g., pace makers [27,28], indwelling vascular catheters [29–31], grafts [32] and left ventricular assist devices [33–37]) are common causes of infective endocarditis. Biofilms have also been identified on various other medical devices [12,38] such as contact lenses, endotracheal tubes, central venous catheters, pacemakers and voice prostheses that account for over 80% of microbial infections in the body. Catheter-associated urinary tract infection is also a common source of biofilm infections. It has also been suggested that autoimmune disorders, such as arthritis, chronic fatigue syndrome, fibromyalgia, Crohn's disease and ulcerative colitis, are caused by biofilm infections [39]. In summary, bacterial biofilms pose a serious threat to human health because of the added protection biofilms provide from an immune response and antibiotic treatments, the ease of acquiring an infection from trauma and medical procedures and the rapid emergence of drug resistance among bacteria that form biofilms.

What does a biofilm look like?

The formation and structure of bacterial biofilms have been extensively reviewed and will only be briefly summarized here [1–3,5,8,9,40–48]. A biofilm (Figure 1) is composed of three parts: a living or nonliving substance that provides a moist surface for attachment of the highly organized microbial structure [49–51]; a slim-like matrix made of extracellular DNA, proteins and polysaccharides (β (1–6)-linked *N*-acetylglucosamine polymer) [52,53] that embeds the microorganism [54]; and an aggregate of microorganisms in a community that exchange fluids, nutrients and chemical signals [46]. The life cycle of the biofilm can be

divided into approximately three steps: attachment, growth and propagation (Figure 2). First, a few colonies reversibly adhere to the surface via van der Waals' forces to create an initiation site. Attachment involves lipoteichoic acid anchored to the cell membrane [55]. This is followed by an irreversible attachment of the cells through the production of the exopolysaccharide matrix and cell growth [51]. The cell growth is not uniform and results in the formation of channels [9]. A combination of cell division and recruitment occurs during the maturation stage and only biofilm shape and size are changed [9,52,56]. Finally, detachment of individual cells and dispersion enables the biofilm to spread and colonize new surfaces or to join another biofilm [57,58].

Biofilms are viewed as layers of bacteria encapsulated within different microenvironments due to variations in nutrient availability [59] and differing cell densities [45,55]. Bacteria within the biofilm core exist in a stationary or dormant growth phase [59] and are physiologically distinct from planktonic bacteria [44,60–62]. Thus, antibiotics that target cellular mechanisms associated with growing bacteria have diminished activity against biofilms [63,64]. Biofilms form irregular spatial structures, which are affected by many different ecological, biological, chemical and physical factors. The effect of these factors on biofilm formations have been extensively investigated using a variety of computer simulations [56,57,65–67].

Biofilm formation at the molecular level

There are some general principles regarding biofilm formation that include the need for metabolically active bacteria for surface adhesion, the need for an adequate nutrient supply for cell replication and exopolysaccharide production, and the fact that surfaces coated with organic nutrients stimulate biofilm formation [2,3]. Correspondingly, bacteria biofilms readily form on the surfaces of plastic or metal medical devices in body fluids. Some common bacterial biofilm infections include *P. aeruginosa* in the lung, *Escherichia coli* in the urinary tract, *Vibrio cholera* in the GI tract, *S. epidermidis* in the heart, *S. aureus* in arteries, *Enterococcus* spp. in the urinary tract, and fungi such as *Candida* spp. in the GI tract [58]. The identification of biochemical pathways critical to biofilm formation is an important first step to being able to prevent these bacterial biofilm infections. Even with our general understanding of the basic structure and development of bacterial biofilms, comprehending the underlying processes responsible for inducing the transition from planktonic cells to a biofilm is still unclear. Correspondingly, the planktonic to biofilm transition is a complex and highly regulated process that results in a phenotypic change. Thus, the differential expression and regulation of specific genes are associated with biofilm formation.

Genomics & bacterial biofilms

Genomics analysis of biofilm formation started in the 1990s by first screening for biofilm-defective mutants [68,69]. Such efforts identified a diverse number of genes required for biofilm formation [70–76]. More recently, DNA microarray technology has been used to identify genes up- or down-regulated in bacterial biofilms [77]. Unfortunately, there does not appear to be a clear trend in biofilm-related genes. Instead, multiple pathways to biofilm formation that depend on media, growth conditions and the specific organism are likely [40]. Nevertheless, some broad, common features have been observed, such as the upregulation of genes for polysaccharide production, for various stress-induced pathways, for stationary phase-induced genes, for a prevalence of genes of unknown function and new regulatory pathways [77]. For example, Quoc *et al.* identified 19 genes in *S. aureus* associated with biofilm formation that were not previously observed [74]. Again, this highlights the difficulty encountered with identifying a uniform set of biofilm-related genes. Besides genes involved in polysaccharide intercellular adhesion (PIA) or unknown function, the authors

observe mutations in guanosine-dependent regulation and formation of wall teichoic acid. The complexity of biofilm formation may be attributed to the fact that the transition from planktonic cells to a biofilm is influenced by various and diverse environmental factors such as ethanol [78], oleic acid [79], glucose [80], UDP-*N*-acetylglucosamine [81], sub-inhibitory concentrations of some antibiotics [82], anaerobic conditions [83], iron limitation [84–86], high osmolarity [87] and high temperature [87]. The diversity of these external stimuli suggests a versatile regulation system.

Regulating biofilms: σ factors, two-component systems or quorum sensing

After two decades of research, multiple mechanisms of biofilm regulation have been proposed: σ factors, two-component systems (TCSs) or quorum sensing. σ factors control the expression of various genes, including virulence factors and global regulators, which are related to biofilm formation [88–90]. They are activated when bacteria sense environmental conditions that induce stress (e.g., heat shock, nitrogen-limitation, starvation and high osmolarity) [91]. TCSs are an alternative stimulus–response coupling mechanism that have been shown to regulate diverse metabolic processes, such as the bacterial cell cycle, cell–cell communication, and virulence factors in biofilm formation [92]. In a TCS, a ligand or a signal molecule can stimulate a histidine kinase sensor protein, which undergoes autophosphorylation at a conserved histidine residue. The phosphoryl group is then transferred to the cognate response regulator, which can activate or repress transcription of the target genes [92,93]. Conversely, quorum sensing uses signal molecules for bacterial intercellular communication. Quorum sensing enables bacteria to ‘sense’ cell density and coordinate behavior in response to nutrient availability, toxic compounds, host–immune response, and defense [94,95]. In Gram-negative bacteria, *N*-acyl homoserine lactones (autoinducer-1 [AI-1]) have been identified as the signal molecules [96]. AI-1 is synthesized and sensed by analogous LuxI and LuxR regulatory proteins. The specific AI-1 molecule varies between Gram-negative organisms. For Gram-positive bacteria, autoinducer peptides (with no conserved sequence) have been identified as a signal molecule that involves a two-component signal-transduction system [97]. Furanosyl borate diester (autoinducer-2 [AI-2]) has been identified as a universal interspecies signal molecule that regulates biofilm formation in over 55 Gram-positive and -negative species [98]. Figure 3 illustrates some common regulatory mechanisms of the planktonic to biofilm transition.

Targeting biofilm regulation systems for drug discovery

TCS and autoinducers are promising drug targets for biofilms because of the essential role in cell growth regulation and the unique mechanisms of action compared with conventional antibiotics [99–102]. Importantly, proteins from TCS and quorum sensing are absent in humans, minimizing toxicity concerns. In addition, both biological processes are based on ligand-receptor interactions, which are typical targets for drug discovery and have a reasonable likelihood of success. Histidine kinases and response regulators, such as WalK/WalR, YhcS/YhcR81, HP165/HP166 and MtrB/MtrA, are potential drug targets for bacteria pathogens [102]. Targeting the kinase domain appears to suffer from poor selectivity, but targeting the sensor domain may prove more successful. Furthermore, targeting nonessential TCS proteins that regulate virulence, such as GacS/GacA, PhoQ/PhoP and CorS/CorR, has demonstrated some initial positive results [102].

Alternatively, quorum sensing may be interrupted by targeting the LuxI, LuxR or LuxS transcriptional regulators [103,104], AIP receptors [101] or Lsr transporters [98]. A common approach is to use the three classes of autoinducers as chemical templates to design agonists or antagonists as a starting point for drug design (Figure 4) [105,106]. For instance, TCS proteins QseC/QseB from *E. coli* (EHEC) O157:H7 that responds to AI-3, epinephrine and norepinephrine, are inhibited by LED209 (*N*-phenyl-4-[[[(phenylamino)thioxomethyl]

amino]-benzenesulfonamide) [107]. Additionally, the RNAIII-inhibiting peptide (RIP; YSPWTFN-NH₂) is an inhibitor of the staphylococcal TRAP/*agr* system, which is regulated by autoinducer RNAIII-activating protein [108]. RIP was shown to be active against MRSA graft infections. There has also been significant effort in the design of AI-2 analogs as novel antibiotics because of AI-2's broad activity against multiple species [109,110]. For instance Roy *et al.* explored the activity of C-1 alkyl analogs of AI-2 against multiple bacterial species [110]. Ethyl-4,5-dihydroxy-2,3-pentanedione (DPD) was found to inhibit quorum sensing in both *E. coli* and *Salmonella typhimurium*. In addition, Rui *et al.* explored DPD analogs with a new stereocenter at C-5 (4*S*, 5*R*)-DHD [111]. The compound was also shown to be active against both *E. coli* and *Vibrio harveyi*. Conversely, Tsuchikama *et al.* synthesized carbocyclic analogues of DPD that were inactive against *S. typhimurium* and *V. harveyi*, suggesting the importance of the linear form of DPD and the heterocyclic oxygen atom [112]. Similarly, Lowery *et al.* explored a variety of DPD analogs that resulted in a uniform lower activity, which highlights the general challenge in evolving a small chemical template into a drug [113]. Nevertheless, the unique mechanism of action for TCS and autoinducer targets holds the promise of a valuable new class of drugs that may circumvent biofilm resistance to antibiotics (Figure 4).

An overview of metabolomics

What is metabolomics?

Metabolomics is the study of metabolites, such as amino acids, carbohydrates and lipids that are the end products of cellular regulatory processes, as well as intermediates and other signaling molecules [114,115]. The **metabolome** is the complete collection of all metabolites within a biological cell compartment, cell, tissue, organ or organism examined in the form of a cellular extract or biofluid [116,117]. In general, a molecular weight of 1 kDa is the typical limit that separates metabolites from macromolecules [118].

There are many differences between conventional metabolite measurements and metabolomics. First, metabolomics focuses on a global or broad-based analysis of metabolites through a high-throughput detection methodology compared with a limited and directed analysis of a specific number of individual metabolites [115]. In general, metabolomics does not require the complete separation of individual metabolites. Instead, it captures a 'snapshot' or 'fingerprint' of the state of the metabolome. Thus, metabolomics simplifies metabolite detection by using a single analytical technique to characterize the state of the metabolome. In this manner, metabolomics also provides an unbiased view of changes in metabolism by covering all major pathways. Thus, the systematic analysis of the ultimate response of a biological system has a better chance of describing pleiotropic effects [115]. Second, metabolomics uses a combination of multiple methodologies, such as cellular biology, instrumental analysis, **chemometrics** and bioinformatics to analyze the biological system. This combination of techniques provides a better view of the global role that metabolism plays in cellular functions. Again the analysis of a select set of metabolites does not provide this sort of global picture of cellular activity. However, in theory, it should be possible to correlate metabolic changes in a biochemical pathway with the enzymes involved, and then to the underlying genetic alterations or changes in gene expression or regulation [115]. A computational simulation could also integrate the experimental data to create a systematic view of the effected biochemical pathways and, potentially, the relevant proteins. The identification of specific proteins that are disease-related or, in this case, related to biofilm formation, is a fundamental and critical step of the drug-discovery process.

Benefits of metabolomics to drug discovery

Autoinducers are an excellent example of the significant roles that small molecules can play in bacterial biofilms, where mimicking or inhibiting the mode of action of autoinducers is a potential drug-discovery target. Other inter- or intra-cellular small-molecular-weight molecules may have similar roles in the initiation, progression and survivability of bacterial biofilms. Specifically, biofilms are a spatial distribution of heterogeneous cells, where cells exist in different metabolic states to maximize survival. Thus, understanding biofilms requires a comprehensive characterization of the various metabolic states within a complex cellular community. Correspondingly, metabolomics provides a systematic approach to explain this complex system [119,120].

Compared with genes and proteins, primary metabolites are highly conserved between various cells and organisms [121]. In a similar manner to gene expression, some of the key metabolites, nutrients and signal molecules (autoinducers) have been shown to significantly influence biofilm formation [122]. Thus, metabolite quantification and the pathway modeling of complex biological systems is useful for exploring cell behavior in establishing a biofilm community. Furthermore, metabolomics is an invaluable approach for investigating antibiotic resistance in biofilms. By generating a network of metabolites affected by the drug treatment, it is possible to predict the antibiotic's mechanism of action [123]. Additionally, the phenotype of antibiotic resistance and biofilm strains can be characterized through their relative metabolome differences. Similarly, monitoring metabolic changes can be used to investigate the effects of other environmental stimuli on biofilm formation [124]. Metabolomics can also be used for detecting disease biomarkers [125] and as a supplementary tool for proteomics and transcriptomics. The linkage between metabolomics, mRNA and protein expression makes it possible to visualize the biological state of an organism [126].

Metabolomics is the bridge between genotype and phenotype [115]. Correspondingly, metabolomics provides a better understanding of a disease since it links the pathology to actual changes in the activity of biological processes. Metabolomics provides an approach to diagnose a disease, monitor its progression, evaluate a response to therapy, and identify potential novel drug targets. Thus, metabolomics has a wide range of applications in drug discovery [127], including toxicology [128] and functional genomics [126].

Achievements of NMR-based metabolomics

NMR metabolomics has been applied to identify biomarkers for cardiac disease [129,130], liver disease [131], respiratory disease [132,133], cancer [134–137] and CNS disorders [138–141], among others. NMR metabolomics provides a means to differentiate between a disease and healthy state or between drug treated and untreated. Drug discovery or chemical-lead identification is then based on observing the metabolome change from a disease state to a healthy state or by simply observing that a compound changes the metabolome. For example, Tizianni *et al.* describes using NMR metabolomics in a high-throughput screening platform (96-well plates) to identify kinase inhibitors [142]. They demonstrate that changes in the lactate/pyruvate ratio in human leukemia cells (CCRF-CEM) and human ovarian cancer cells (SKOV-3) was successful in identifying inhibitors of eEF-2, NF- κ B, MK2, PKA, PKC and PKG kinases. Similarly, Halouska *et al.* demonstrate that the *in vivo* mechanism of action of a chemical lead can be inferred by comparing the metabolome changes to a known drug [123]. If two or more drugs have a similar impact on the metabolome then the compounds share a similar target. Additionally, NMR metabolomics is also widely used for drug development and personalized medicine [143–145]. The consortium for metabolomic toxicology, an organization of major pharmaceutical companies, was formed to share metabolomics data from drug studies to characterize

metabolites associated with kidney or liver toxicity. The overall protocol for toxicity analysis is comparable to drug discovery and lead identification. Specifically, biofluid (e.g., urine, serum and saliva) metabolites from animals or patients undergoing drug treatment are analyzed by NMR to identify metabolites known to be associated with drug toxicity or disease biomarkers. If the biofluid metabolites reveal the presence of drug toxicity or the lack of drug efficacy, then an alternative treatment can be prescribed.

Designing a metabolomics experiment: what information is desired?

Metabolomics is routinely combined with alternative strategies to resolve a biological problem and to generate a comprehensive analysis. Although different strategies may require alternative experiments and data analysis, targeted metabolite analysis, metabolic profiling and metabolic fingerprinting share the same general workflow from sample preparation to data collection and analysis. Targeted metabolite analysis is the exclusive study of the direct product of a corresponding enzyme or protein [116]. Metabolic profiling is the identification and quantification of a set of predefined metabolites in a biological sample [146]. The metabolites may belong to a specific class of compounds or a particular metabolic pathway. As an illustration, metabolic fingerprinting can be used to probe different metabolic phenotypes. Metabolic profiling can then be used to provide a detailed analysis of specific metabolite changes between the two phenotypes. Metabolomics can also be referred to as 'metabolic fingerprinting' [147], and is generally designed to rapidly classify biological samples. The combination of different strategies makes metabolomics a flexible and versatile technique for the analysis of various biological systems, such as bacterial biofilms [148,149]. The NMR metabolomics methodologies described for drug discovery are equally applicable to investigating bacterial biofilms, for identifying new drug targets and chemical leads, and evolving lead candidates into new drugs. Again, characterizing and comparing the metabolomic differences between planktonic cells and biofilms provides a means to identify active and relevant biological processes associated with biofilm formation. Correspondingly, proteins involved in these pathways are potential drug targets. Identifying and validating drug leads can then be accomplished by observing chemical-induced changes in the metabolome related to biofilm formation.

Designing a metabolomics experiment: what steps need to be taken?

NMR-based metabolomics generally refers to a comprehensive approach to the analysis of metabolomic samples that include specific NMR experiments, sample preparation protocols, and multivariate statistical analysis [150–152]. NMR spectroscopy is used to characterize the metabolic samples by providing both qualitative and quantitative data [117]. NMR-based metabolomics of bacterial biofilms consist of the general procedures outlined in Figure 5: prepare the metabolic samples by culturing the desired bacterial strains under identical conditions (the only variable should be the specific environmental or genetic factor being investigated); prepare the NMR samples by lysing the cells, extracting the metabolites and removing cell debris; detect the metabolites through various NMR techniques; and perform spectral processing, data normalization, statistical analysis and metabolite identification. The success of metabolomics largely depends on accomplishing each step in a highly controlled and uniform manner. Variations in the NMR metabolomics data should result from relevant biological differences between the samples as opposed to artifacts introduced from sample or data handling. For example, extracting the metabolites from the lysed cells should occur quickly and at low temperatures to avoid changes to the metabolome that results from the process of harvesting the cells. In effect, all potential variables, such as the number of cells, growth phase, culture media, experimental conditions, bacterial strain and time, need to remain constant between all bacterial samples [153]. Again, the only difference between the various bacterial cultures should be the specific environmental or genetic factor being investigated. Uniformity is the key to a successful metabolomics experiment. It is

impractical to investigate a heterogeneous sample composed of multiple bacterial organisms since there is no mechanism to associate the majority of the commonly observed metabolites to a specific organism. Alternatively, using flow cytometry or other techniques [154] to separate the cells prior to investigating the metabolome does not solve the problem. The results would be invalid because the time required and the process of separating the cells would perturb the state of the cells and, correspondingly, the metabolome.

Designing a metabolomics experiment: what model is used?

Fundamental to a metabolomics study is the identification of the **classes or groups** of bacterial cells that will be compared. A straightforward application is the comparison between two groups, a bacterial strain in its planktonic state (class one) and its biofilm state (class two). Models of higher complexity include even more classes. For example, Figure 6 illustrates the application of metabolomics to monitor *in vivo* drug activity. The diagram illustrates the general analysis of clustering patterns in a principal component analysis (PCA) **scores plot**. In a scores plot, the metabolome of four different classes are compared: wild-type cells; mutant cells in which the drug-target has been genetically knocked-out; wild-type cells treated with the drug; and mutant cells treated with the drug. The activity and selectivity of the drug is determined by the relative similarity of the four different metabolomes as described in Figure 6. As an illustration, a drug is selective and active if the metabolome of the wildtype cells treated with the drug clusters together with the metabolome obtained from the mutant cells with and without drug treatment, where this cluster forms a separate cluster from the wild-type cells without the drug treatment (Figure 6B). These results indicate the protein target in the wild-type cells was chemically inactivated since the metabolome is identical to the mutant cells where the protein was genetically inactivated. It also differs from the wild-type cells without the drug treatment where the protein is still active. The drug is selective because there is no difference between the metabolomes for the mutant cells with or without drug treatment. This analysis can be easily generalized. The 'drug' in this scenario can also be taken as any environmental condition, while the 'mutant' can be taken as the drug target or any knockout, repressed or overexpressed gene.

Metabolomics sample preparation

Metabolite sample preparation includes cell quenching, cell harvesting, cell disruption and metabolite extraction. An important advantage of NMR-based metabolomics is the minimal and relatively simple sample preparation protocol. Nevertheless, the details of the procedure influence the accuracy, reliability and reproducibility of the metabolomics data [117]. Different approaches to sample preparation have various advantages in terms of speed, capability, consistency, efficiency and metabolite recovery yield [155–157]. Since biofilms can form on a wide range of surfaces or habitats, the experimental conditions for growing and harvesting cells can be highly variable. Therefore, this review will focus on a general discussion of sampling methods for biofilm-related planktonic cells.

A proper metabolite extraction technique is critical

A very critical issue in sample preparation is the need to rapidly and efficiently quench all enzymatic and biological activities in order to capture an accurate 'snap-shot' of the metabolome. This is because metabolites, such as pyruvate, fumarate, oxoglutarate, phosphoenolpyruvate, fructose-6-phosphate and others, have a rapid turnover rate [158]. In addition, it is important to avoid inducing a stress response or cell death that would completely invalidate the study. Thus, a quick quenching step that involves reducing the cell temperature has been shown to be a useful approach to slow down enzyme activity within a cell [155]. Methanol is commonly used because of its low freezing point and minimal

toxicity relative to other organic solvents [156]. However, methanol may only be suitable for Gram-positive bacteria or fungi due to the possibility of cell leakage and the loss of metabolites during the quenching and washing steps [158]. Choosing the proper metabolite extraction protocol is extremely critical since it influences the efficiency and accuracy of the entire metabolomics experiment.

Improperly removing the cell growth medium and washing the cells is an easy way to contaminate the metabolomic samples and generate unreliable data. Filtration and centrifugation are the two main methods of removing the culture medium before collecting the metabolome. Centrifugation takes longer so there are concerns regarding induced stress and metabolome changes, but it has a higher consistency in sample preparation. Conversely, filtration is significantly faster, and it is also easier to quench cells on a filter membrane. However, there are also practical concerns with uniformly and consistently retrieving all the frozen cells from the filter paper. Nevertheless, filtration quenching was demonstrated to have the highest yield for an *S. aureus* metabolomics study [159]. An NMR-based metabolomics study of *P. pastoris* applied a single centrifugation step and demonstrated that there was no benefit to including an additional washing step in the quenching process [160]. Directly growing *E. coli* cells on filter paper may provide an efficient and fast approach to quench cells and extract the metabolome [161]. To date, the choice of technique to separate and wash cells is still very flexible, which implies a necessary optimization step for any metabolomics study. An inability to efficiently arrest all cell processes and purify the cells without inducing leakage or lysis will inadvertently lead to undesirable changes in the state of the system and the metabolome. Thus, choosing system-appropriate washing, quenching and cell separation protocols is the first and most critical step of a metabolomic project and will determine the validity of the entire study.

Cell lysis and metabolite extraction can sometimes be carried out simultaneously. Both mechanical disruptions, such as the Fast-Prep[®] system or organic solvent-based methods are widely used [157,162]. Trichloroacetic acid is a traditional approach for lysing cells from filter paper, but trichloroacetic acid causes a significant background for metabolomics data because it also degrades the filter paper [158]. The optimal extraction buffer should: extract the largest number of metabolites; be nonselective and not exclude molecules with particular physical or chemical properties; and be nondestructive, as well as not modify metabolites through chemical or physical means [163]. Since metabolites are normally dissolved in a D₂O buffer or CDCl₃ for NMR, choosing alternative extraction solvents is not particularly beneficial. A common extraction solvent is a 5:2:2 v/v mixture of methanol, chloroform and water [164]. Other extraction mixtures include a 1:1 v/v mixture of methanol and chloroform, or water and chloroform for metabolite extractions that include lipids [165].

NMR experiments for metabolomics

There are many practical challenges encountered when studying the bacterial metabolome. A cellular metabolome can contain upwards of thousands of metabolites, with a 7–9 order of magnitude range in concentrations (i.e., picomoles to millimoles) [166]. Therefore, it is generally not possible to analyze all cellular metabolites in a single experiment. Also, cellular metabolism is very sensitive to environmental changes, in which the measurement and sampling process can influence the metabolome. Thus, metabolomic measurements are also perturbed by including separation techniques. Correspondingly, each biological system requires experimental optimization to accurately study its metabolome.

NMR- or MS-based metabolomics?

MS and NMR are the primary analytical techniques used for metabolite detection. MS measures the mass-to-charge ratio of charged molecules that can be used to determine the

elemental composition and elucidate the chemical structure of molecules. While an exact mass is determined by MS, the limited molecular-weight dispersion of the metabolome generally requires the use of chromatography [167]. GC, HPLC and CE are common separation techniques used in MS-based metabolomics [168,169]. Unfortunately, the use of chromatography to separate metabolites inevitably leads to changes in the metabolome. The amount that is recovered from the chromatography step will vary for each metabolite, where some metabolites will be lost or chemically modified. MS also requires ionization of the molecule for detection with the corresponding uncertainty that a specific metabolite will ionize. Furthermore, determining a concentration for each metabolite by MS is particularly challenging.

Conversely, NMR determines a molecular structure by measuring nuclear chemical shifts within a magnetic field [170]. Three valuable pieces of information are obtainable from a single peak in an NMR spectrum (Figure 7). The chemical shift is related to the local chemical environment of that specific nucleus (e.g., ^1H , ^{13}C and ^{15}N), and the peak-splitting (J-coupling) identifies chemically bonded nuclei, which are both used to identify the chemical structure of the metabolite [170]. Importantly, the peak intensity is directly proportional to the metabolite's concentration. Typically, multiple distinct NMR resonances are observed per molecule, which eliminates the need for chromatographic separation and increases the accuracy of metabolite identification.

MS is significantly more sensitive than NMR and covers a wider diversity of the metabolome, although the use of cryogenic probes has significantly increased the sensitivity of NMR by a factor of four [171]. In effect, NMR only detects the most abundant metabolites that are present at concentrations greater than 1 to 5 μM . ^1H NMR is typically used for metabolomics since ^1H NMR is 64-times more sensitive than ^{13}C NMR. Nevertheless, NMR cryoprobes can still be used to detect metabolites using naturally occurring ^{13}C , which has an abundance of only 1.1% (^{12}C is not detectable by NMR) [172,173]. MS is a destructive technique, but it requires a significantly lower sample amount (<100 μl) compared with NMR (600 μl). Correspondingly, both approaches are complementary to each other and contribute inherently distinct information to the analysis of a metabolome [174]. The complementary nature of MS and NMR has been demonstrated by a number of metabolomic studies using both techniques [175–179]. In effect, the MS and the NMR data can be combined to create a 3D scores plot. The added dimensionality from complementary data may provide the additional resolution necessary to differentiate between multiple classes or groups.

1D, 2D & solid-state NMR techniques for metabolomics

The application of NMR spectroscopy for metabolomics can be categorized into one of three groups, 1D NMR, 2D NMR and solid-state NMR [170]. 1D and 2D solution-state ^1H NMR experiments are commonly used for global metabolomics analysis of bacterial cell extracts [180]. Conversely, solid-state NMR can be used to analyze intact cells [181–183]. In addition to ^1H , other nuclei are also used in 1D NMR-based metabolomics, such as the metabolic profiling of the carbohydrate cycle using 1D ^{13}C NMR [184,185], or tissue metabolism using 1D ^{31}P NMR [186].

A typical 1D ^1H NMR spectrum of a bacterial cell lysate may contain thousands of sharp lines from low-molecular weight metabolites (Figure 7) [117]. The entire 1D ^1H NMR spectrum is used as a 'fingerprint' to characterize the state of the bacterial cell. A global investigation of the metabolome is based on a comparative analysis of the features present or absent in each 1D ^1H NMR spectrum. A global metabolomic analysis is based on how similar or how different the 1D ^1H NMR spectra are between each class or group. It is not necessary to assign each 1D ^1H NMR spectrum to identify and quantify all the metabolites

present in each sample. Overlapping these relevant NMR resonances and interfering with the analysis may be broad bands from proteins or other biomolecules, which can be readily removed by using a Carr–Purcell–Meiboom–Gill (CPMG) spin-echo sequence [187]. The CPMG pulse sequence takes advantage of the large difference in T_2 relaxation times between small-molecular-weight metabolites and large-molecular weight biomolecules. The NMR resonances from the biomolecules rapidly decay during the CPMG pulse. Alternatively, proteins and other biomolecules can be removed by an appropriate choice of extraction solvents [188,189]. The large interfering signal from water or other buffer components is also eliminated by the use of appropriate NMR solvent-suppression methods and a 100% deuterated buffer [190]. The water NMR resonance is set in the center of the spectrum where selective irradiation and gradient pulses suppress the solvent peak while leaving all other peaks unaffected. In effect, any resonance in the 1D ^1H NMR spectrum that does not originate from the bacterial metabolome will generate a ‘false feature’ that needs to be removed. NMR resonances originating from proteins or solvents are likely to be variable and dominate the spectra relative to metabolite signals. This will lead to an erroneous interpretation of the 1D ^1H NMR spectra and incorrect sample classification. For instance, replicate samples may not cluster together because of a significant variation in the peak height and peak shape of the water resonance despite the overall similarity in all the metabolite NMR peaks.

High-resolution magic angle spinning (HR-MAS) is used to generate *in vivo* 1D ^1H spectra of solids comparable to solution-state NMR [191]. Thus, small, intact and untreated cells or tissues can be directly analyzed by HR-MAS by spinning samples at speeds between 4 and 12 kHz at the ‘magic’ angle of 54.7° relative to the external magnetic field. Spinning the sample significantly reduces NMR line widths by averaging out chemical shift anisotropy, magnetic susceptibility, and dipolar coupling that are prominent in solid samples [170].

Assigning a 1D ^1H NMR spectrum to identify the metabolites present in a sample is challenging because of the large number of peaks, the significant overlap in peaks, the high chemical shift degeneracy (multiple metabolites have some chemical shifts in common), and an incomplete database of NMR reference spectra for metabolites. Again, assigning a 1D ^1H NMR spectrum is not necessary for a global analysis of the metabolome, but identifying the specific metabolites that are changing and are the main contributors to class distinction is extremely valuable for understanding the underlying biological differences. Statistical total correlation spectroscopy (STOCSY) can be used to associate multiple NMR peaks from the same molecule in a complex mixture [192]. This significantly simplifies the assignment problem since most, if not all, of the NMR resonances for a given metabolite can be used together in a database search. A positive identification only occurs when all of the observed chemical shifts match the metabolite’s known chemical shifts in a database. In STOCSY, a series of 1D ^1H NMR spectra is converted into a pseudo-2D spectrum that is based on a correlation of peak intensities. NMR peaks from the same metabolite will change together as the metabolite’s concentration varies across multiple distinct classes. The statistical heterospectroscopy (SHY) is similar in concept to STOCSY [193]. Instead of correlating NMR peak intensities, SHY correlates chemical shifts from NMR with m/z data from MS. Thus, SHY can improve molecular identification by directly cross-correlating NMR chemical shifts with a molecular weight.

More commonly, 2D NMR spectroscopy improves the accuracy of metabolite assignments by significantly increasing spectral resolution by extending chemical shift information into a second frequency dimension. Additionally, 2D NMR experiments can identify the network of resonances associated with a specific metabolite through J-coupling. 2D correlation spectroscopy (COSY) and total correlation spectroscopy (TOCSY) experiments identify spin–spin coupling connectivities that identify chemically bonded pairs of hydrogens,

carbons or nitrogens [194,195]. To a lesser extent, 2D J-resolved NMR experiments are also used for metabolomics studies [196]. In a 2D J-resolved NMR experiment, the data are dispersed into two dimensions based on chemical shifts and the J-coupling pattern [197]. While it is possible to match coupling patterns to identify bonded resonances, this is generally not practical for a complex metabolomics data set. Therefore, a 2D J-resolved NMR experiment has significantly less information than a 2D COSY or TOCSY experiment while requiring the same amount of experimental time. Alternatively, a 1D projection of the 2D J-resolved NMR experiment can be used to simplify the NMR spectra by removing peak multiplicity due to J-coupling [198]. This dramatically reduces the number of peaks and correspondingly reduces peak overlap. Removing peak splitting improves the accuracy of metabolite identification and quantification. Of course, a single 2D NMR experiment may require 1 h or longer to acquire compared with a few minutes for a 1D NMR experiment.

Why 2D ^1H - ^{13}C heteronuclear single-quantum correlation experiments are commonly used for metabolite identification

The 2D ^1H - ^{13}C heteronuclear single-quantum correlation (HSQC) experiment correlates the ^1H and ^{13}C chemical shifts for each C-H pair in a molecule [199]. This provides unique information relative to the COSY, TOCSY or J-resolved experiments. In addition, metabolite assignments are easier with a 2D ^1H - ^{13}C HSQC experiment because of two correlated and distinct chemical shift assignments, the large ^{13}C chemical shift dispersion and the simplified spectrum without splitting from J-coupling. However, due to the low natural abundance (1.1%) of ^{13}C -labeled compounds, such as ^{13}C -methanol, ^{13}C -CO₂ [200], ^{13}C -glycerol [201] and ^{13}C -glucose [202] are required as a bacterial carbon source to enhance the sensitivity of the NMR spectrum [203]. This significantly simplifies and focuses the analysis of the metabolome. Only metabolic intermediates and products of the ^{13}C -labeled materials will be partially or completely enriched with ^{13}C . Correspondingly, only these metabolites will be observed in a **2D ^1H - ^{13}C HSQC spectrum**, which provides a means to follow carbon flow through the metabolome and identify the perturbed metabolic pathways. Standard HSQC experiments are not quantitative because of significant variability in coupling constants and relaxation times (T_1 and T_2) between metabolites. Nevertheless, the newly developed 2D extrapolated time zero ^1H - ^{13}C HSQC (HSQC₀) experiment allows for the calculation of metabolite concentrations [204]. This experiment collects a series of 2D HSQC spectra with an increasing number of the core NMR pulse sequence or HSQC block. Typically, this HSQC block is repeated from one- to three-times. The peak intensity will decrease linearly with the number of HSQC blocks, where a linear fit and extrapolation back to zero HSQC blocks will determine the true peak intensity and metabolite concentration. Examples of 2D NMR spectra used to identify metabolites from biological samples are shown in Figure 8.

Processing NMR data: binning, peak alignment, baseline correction & normalization

For chemometrics (also see the section 'Chemometrics & bioinformatics analysis of metabolomics data'), the 1D ^1H NMR spectra are transformed into a data matrix of integrated peak intensities and corresponding chemical shift values. These data are used to define the classes and to identify NMR spectral features that differentiate the classes. Unfortunately, subtle instrument, temperature and sample condition variability (e.g., pH and ionic strength) can result in chemical shift differences between replicate samples. Correspondingly, misalignments will occur in the NMR data matrix between these replicate samples that will lead to clustering errors independent of any biological differences. One approach to normalize NMR metabolomics samples is the inclusion of a known concentration of chemical shift reference compound, such as the sodium salt of 3-trimethylsilylpropionic acid. The 3-trimethylsilylpropionic acid peak intensity can also be used to calibrate the concentrations of the metabolites in the biological sample. However, an

internal reference only corrects for global differences. It does not correct for individual peak position and shape variability due to subtle experimental differences. Instead, binning or bucketing is commonly used to correct for peak variability between replicate samples [205,206].

In its simplest implementation, bucketing divides the NMR spectrum into equally sized, nonoverlapping subspectral regions with a width or bin size of 0.025 ppm. A typical 1D ^1H NMR spectrum with a width that ranges from 0 to 10 ppm will have 400 buckets. Bin sizes typically range from 0.01 to 0.04 ppm. The peak intensity within each bucket is integrated, and these resulting integrals are then used as input to the chemometrics analysis. In this manner, the buckets or bins smooth out small peak variability with the expectation that the same peaks occur in the same bucket. Unfortunately, bin edges create a second problem: the undesirable splitting of a peak between buckets. Thus, 'intelligent' or 'adaptive' bucketing techniques have been developed that vary the individual bin size to avoid dividing peaks between multiple buckets [207–210]. These methods use a Gaussian function, a recursive algorithm, optimize an objective function using a dynamic programming strategy, or use undecimated wavelet transforms to automatically identify bin edges. In all cases, intelligent bucketing performs significantly better than uniform bucketing, where dynamic adaptive binning was recently shown to perform the best [210]. The use of intelligent bucketing results in a significant improvement in replicate clustering in scores plots since it minimizes spectral differences that are biologically irrelevant. Noise regions of the NMR spectra are typically zeroed or removed from the bucketing [211]. Similarly, buckets resulting from solvent or buffer peaks that are unrelated to the bacterial metabolome are also excluded. Again, this eliminates class distinction, which results from biologically irrelevant data. In essence, the variables used in the chemometrics analysis should be relevant to real variations in the metabolome between the classes [153].

Peak alignment is a more robust and complex alternative to NMR binning [212–215]. The goal is to remove the chemical shift variability between the replicate 1D ^1H NMR spectra by individually aligning each NMR peak to a representative spectrum from each class. The approaches used include fuzzy warping, genetic algorithms, a generalized fuzzy Hough transform approach, a reduced set mapping (PARS) algorithm, or a recursive segment-wise peak alignment. Each approach demonstrated acceptable results on test metabolomic samples and were shown to improve upon the results obtained from bucketing [215]. Nevertheless, intelligent bucketing is still the predominant methodology used in NMR metabolomics.

An NMR spectrum is experimentally collected as a free induction decay (FID) that requires further processing to convert the time-domain data into a frequency-domain spectrum through a Fourier transform. This process can be accomplished using a variety of software packages (Table 1). All of the software packages can import popular NMR data formats, such as Bruker, Varian, Jeol Delta, JCAMP-DX, as well as others. Some NMR software can process both 1D and 2D data, where others only focus on processing 1D (most popular) or 2D data sets. In addition to the Fourier transform step, the processing of NMR data may include zero-filling, phase correction, baseline correction, applying a window function and removal of solvent peaks. The inclusion of any of these steps induces changes in the resulting NMR spectrum that is not biologically relevant. Thus, uniformity in the NMR processing protocol is essential, where minimizing all spectral manipulation is ideal. In general, phase correction is essential in order to obtain purely absorption peak shapes. Uniform zero-filling of the NMR spectra is typically acceptable since it provides a constant improvement in the digital resolution. Similarly, removing residual solvent peaks by simply zeroing the corresponding region of the NMR spectrum does not have any detrimental effect because these solvent regions are not included in the binning process. Conversely, applying

a window function or incorporating baseline correction should be avoided, if possible, since these processes induce significant variable changes in the NMR spectrum. The goal of a window function is to either increase the spectral resolution or the signal-to-noise by multiplying the FID with a mathematical function. As a result, each NMR peak shape is artificially changed. Instead, the signal-to-noise and spectral resolution should be improved experimentally. A baseline correction is required if the baseline for the NMR spectrum is not flat. A flat baseline is necessary for reliable chemometrics analysis. Distorted baselines may result from intense solvent or buffer peaks, from phasing problems, errors in the first data points of the FID, large range of T_1 values (short recycle times), or short acquisition times. Again, baseline problems should be experimentally minimized. Nevertheless, numerous computational approaches to obtain a flat baseline are available and include a linear or polynomial (up to six orders) fit of the baseline, FID reconstruction and spectrum averaging. The choice of a baseline correction method depends on the specifics of the baseline distortion, which is typically determined by trial and error. Importantly, a single baseline correction method must be used for an entire metabolomics dataset in order to avoid inducing class distinctions that are primarily a result of the NMR processing protocol.

Ideally, the overall concentration of the metabolites and the corresponding signal-to-noise of each replicate 1D ^1H NMR spectrum will be essentially identical. Unfortunately, in practice, there may be a significant variability in the signal-to-noise between replicate NMR spectra due to random errors in cell lysing, metabolite extraction, or the number of bacterial cells per sample. Correspondingly, the 1D ^1H NMR spectra need to be normalized [216]. Center averaging is a common normalization technique:

$$Z = \frac{X_i - \bar{X}}{\sigma} \quad \text{Equation 1}$$

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 a bin. Other normalization techniques include normalization to a constant sum, Pareto scaling, mean centering, logarithmic scaling and probabilistic quotient normalization. A recent analysis of NMR processing procedures by De Meyer *et al.* suggests that a combination of intelligent bucketing with probabilistic quotient normalization yields the best results [217]. Of course, it was not practical to explore all the possible combinations of the various processing techniques and, as a result, the analysis did not include center averaging. This highlights a particular challenge with processing NMR metabolomics data: there is no clear consensus for an optimal protocol because of all the possible combinations of parameters that need to be evaluated.

Chemometrics & bioinformatics analysis of metabolomics data

For a detailed understanding of a complex biological system, it is essential to follow the response of an organism to a conditional perturbation at the transcriptome, proteome or metabolome level [126]. Metabolic products are dependent on genotype, environment, time and location [218]. Perturbations in any of these factors may lead to a variety of biological changes that inadvertently affect the metabolome [126]. The primary goal of chemometrics is to reduce the complexity of the NMR-based metabolomics data to understand these global correlations. In essence, chemometrics identifies the major features within an NMR spectrum, the presence or absence of peaks, the change in peak intensity or shape, or the change in chemical shifts that differ between the various classes. A class definition can be based on any variable that affects or changes the bacterial cell culture, such as different bacterial strains (including wild-type and mutant cells) and different experimental conditions (e.g., growth phase, drug dosage, media composition and pH). Bioinformatics is used to uncover and characterize all the associated variables and to reveal the underlying

relationships. Essentially, bioinformatics is used to identify the metabolites that correlate with the major changes in the NMR spectra and to associate these metabolites with specific proteins, enzymes, metabolic pathways and other biological processes.

PCA & orthogonal partial least squares discriminant analysis are the most commonly used chemometrics techniques

For statistical modeling, PCA is the most common multivariate technique for the comparison of metabolomic data [219]. Each multivariable (chemical shift and intensity) NMR spectrum is converted into a single point in multidimensional Cartesian space (Figure 9). Each axis ($V_1, V_2, V_3 \dots V_n$) corresponds to a specific chemical shift where the peak intensity is the value along each axis. PCA determines the vector (*PC1*) corresponding to the largest variation in the data and fits each NMR spectra to this vector. Each NMR spectrum is assigned a value – principal component one (PC1) based on its distance to the vector (*PC1*). The process is repeated for the next largest variation (*PC2*) orthogonal to the first vector. Plotting the resulting principal components (PC1 and PC2) corresponds to a scores plot, where similar spectra (and metabolomes) cluster together in a scores plot. Spectra (and metabolomes) obtained from different classes are expected to form separate clusters in the scores plot. Figure 9 illustrates the hypothetical separation of class 1 from class 2 and the relatively tight clustering of the class replicates. The class separation is based on differences in the 1D ^1H NMR spectra.

Orthogonal partial least squares discriminant analysis (OPLS-DA) is a related, but alternative approach to PCA. Unlike PCA, OPLS-DA is a supervised method. This is an important and significant difference. In PCA, the different classes are not identified and each 1D ^1H NMR spectrum follows the protocol described above. Thus, the clustering of each NMR spectrum in the scores plot is solely determined by the intrinsic differences and similarities in the features of its NMR spectrum relative to all the other NMR spectra. However, in OPLS-DA each 1D ^1H NMR spectrum is assigned to one of two classes, (e.g., biofilm vs planktonic, healthy vs diseased, treated vs untreated, and wild-type vs mutant). More than two class definitions can be used, but it is generally not recommended. As a result, OPLS-DA determines the clustering patterns in the scores plot based on the NMR spectral features and the class definition. OPLS-DA tries to maximize the separation between the classes based on these manual class assignments, while minimizing the within class variation. This results in tighter class clustering in the OPLS-DA scores plot relative to PCA. Fundamentally, if two classes are defined in OPLS-DA then two clusters corresponding to the two assigned classes will be generated in the resulting scores plot. OPLS-DA assesses a relationship between an X matrix (NMR data) and Y matrix (the 0 or 1 class designation). Thus, OPLS-DA will only identify all the spectral features that can be used to distinguish the two classes regardless of any real significance (e.g., noise and small random variability in peak height). Conversely, the class separation in a PCA scores plot depends on a combination of all principal component variables. Since OPLS-DA is a supervised method or biased by the class assignment, it is essential that the model is validated [153]. Is the class separation due to relevant changes in the 1D ^1H NMR spectra resulting from changes in metabolite concentrations?

The quality of the OPLS-DA model can be evaluated by multiple statistical factors and cross validation, the goodness of fit (R^2) and the quality assessment score (Q^2) [220]. A good OPLS-DA model has R^2 values > 0.5 (range of 0–1) that is conceptually similar to simple linear regressions. Similarly, a typical value for Q^2 for a biological model is > 0.4 . The leave-one-out cross-correlation technique is commonly used to calculate Q^2 , where a sub-set of the NMR spectra are left out to calculate a model that is then used to predict the left-out data [221,222]. Q^2 is the consistency between the predicted and original data. Importantly, Q^2 and R^2 values should only be used as a figure of merit and not validation of the model. A

permutation tests is one approach to validate the model [221,222]. The NMR spectra classifications are randomly assigned creating, in principle, a random data set that should have poor class separation and low Q^2 scores. The process is repeated numerous times (>1000 permutations) until a Gaussian distribution of Q^2 scores is obtained for the random data set. The statistical significance of the model with the correct NMR spectra classification can then be obtained by comparing the model Q^2 with the random Q^2 scores and calculating a p-value. In addition to Q^2 , a similar validation can be achieved by using the area-under-a-receiver-operating-characteristic curve or the number of misclassifications.

Further validation of the OPLS-DA model can be assessed by determining if the NMR features that determine the class separation are biologically relevant. S-plots and loading plots are additional outcomes of an OPLS-DA model. An S-plot identifies the relative contribution of each bin (chemical shift and metabolite) to the clustering in the corresponding scores plot, each bin with a covariance of >0.10 or <-0.10 are identified as major contributors to the class separation [223]. Similarly, a loadings plot displays the relative contribution of each bin to the principal components. Thus, a valid OPLS-DA model is supported by being able to assign metabolites to the bins that are associated with the class separation. The model is further validated if multiple bins assigned to the same metabolite are all major contributors to the class separation. Additionally, the identification of multiple correlated metabolites, members of the same metabolic pathway for instance, that are all major contributors to the class separation would further strengthen the reliability of the OPLS-DA model. Again, the overall goal of NMR-based metabolomics is to identify major changes in the NMR spectra that can be associated to a specific set of metabolites with a relationship to a biological process, such as bacterial biofilm formation.

A number of metabolomics software packages have been developed to automate the chemometrics analysis by combining data normalization, data reduction, model prediction and validation, and even metabolite identification, into a single work flow. These programs include: Automics [224], HiRes [225], MetaboAnalyst [226] and the R-package Metabonomic [227]. Among these, Automics is the most versatile and extensive package and includes nine different statistical methods applicable to metabolomics data: feature selection (Fisher's criterion), data reduction (PCA, LDA and ULDA), unsupervised clustering (K-Mean) and supervised regression and classification (PLS/PLS-DA, KNN, SIMCA and SVM). Automics also incorporates processing tools to generate a STOCSY spectrum.

Interpreting a PCA or OPLS-DA scores plot is fundamentally a cluster analysis. If two or more classes cluster together in the scores plot, then the NMR spectra and the corresponding metabolome are considered statistically similar. Conversely, two or more classes that form distinct clusters indicate the samples are significantly different metabolomes. For a simple metabolomics study that involves only two or three different types of samples, the cluster analysis is generally straightforward. An ellipse that corresponds to the 95% confidence limits from a normal distribution for each cluster can be used to define each class in the 2D scores plot to visually determine class separation. Alternatively, Werth *et al.* [228] recently demonstrated the application of metabolomic tree diagrams combined with standard bootstrapping techniques [229] as a more robust statistical analysis of clustering patterns in scores plots. Each PC value (PC1 and PC2) is treated as an axis in a Cartesian coordinate system. An average position is calculated for each class cluster, which is then used to calculate a Euclidean distance between each class to create a distance matrix. The cluster centers and distances between clusters are re-calculated by randomly selecting data points from each class. The process is repeated until 100 different distance matrices are generated, which are then used to generate 100 tree diagrams using Phylip 3.68b [230]. A consensus tree is created, where the bootstrap number is simply the number of times each node appears

in all 100 trees. Bootstrap values below 50% imply a statistically insignificant separation. Further analysis to identify the metabolic processes that led to class separation may require bioinformatics [231]. An example 2D OPLS-DA scores plot and associated metabolomics tree diagram is shown in Figure 10.

Metabolic pathway reconstructions have been widely employed with five major goals: contextualization of high-throughput data; guidance of metabolic engineering; directing hypothesis-driven discovery; interrogation of multispecies relationships; and network property discovery [232]. NMR-based metabolomics are routinely used to generate these metabolic networks for a particular biological system (Figure 11). The first step of the process is to deconvolute and decode the NMR spectra. This is a difficult and time-consuming process because the high degeneracy of NMR chemical shifts makes unambiguously assigning an NMR resonance to a specific metabolite extremely challenging. This is further complicated by the incompleteness of metabolomics data and by several technical problems that include proper peak assignment [180], peak alignment [233,234], absolute concentration determination [204,235] and intensity normalization [236]. A number of metabolomics databases that contain NMR and MS spectra with the associated assignments are available: Metabominer [237], Madison Metabolomics Consortium Database [238], BioMagResBank [239] and Human Metabolome Database [240]. Some efforts to automate the process have also been made [180,224].

The next critical step is to analyze the identified metabolites and find all possible correlations. The presence of metabolites and metabolic pathways in a particular organism can be verified by the KEGG [241] and MetaCyc databases [242]. A metabolic network map can then be generated using Cytoscape [243], and there are many metabolome simulators that are useful for predicting the networks involving hundreds of metabolites [244]. Additionally, Cell Designer is a valuable diagram editor for drawing biochemical networks [245].

Analysis of biofilms with NMR-based metabolomics

An obvious application of NMR-based metabolomics has been used to characterize the difference in the metabolome between planktonic and biofilm cells. Gjersing *et al.* applied ^1H HR-MAS to study the metabolome difference between *P. aeruginosa* chemostat planktonic and biofilm cells [149]. There was no apparent difference in the extracellular metabolite composition when planktonic and biofilm cells were grown under continually feed chemostat mode. Conversely, the 2D PCA scores plot of the intracellular metabolome indicated a clear distinction between the batch and chemostat planktonic and biofilm cells (Figure 12A). The corresponding loadings plot suggests major and complex differences between the two metabolomes (Figure 12B). The metabolite differences were not analyzed in detail, but it was noted that biofilm metabolites were generally lower in concentration, possibly due to the cells closer to the substrate operating at a lower metabolic rate. Similarly, Workentine *et al.* also used ^1H NMR to characterize the phenotype of different biofilm colonies of *Pseudomonas fluorescens* [246]. PCA of the 1D ^1H NMR spectra yielded a 3D PCA scores plot with distinct clustering for the four different *P. fluorescens* colonies (Figure 13A). The class distinction was attributed to changes in amino acids (Asp, Glu, Gly, Met, Phe, Pro, Trp and Val) and central metabolites (acetate, glutathione and pyruvate). This is suggestive of changes in exopolysaccharide production, response to oxidative stress, and an impaired amino acid metabolism. Interestingly, the four different *P. fluorescens* phenotypes exhibited distinctly different metal sensitivity (Figure 13B). Unfortunately, it was not possible to determine whether the metabolome differences were a result of the different metal sensitivity. Booth *et al.* also analyzed *P. fluorescens* cells under metal stress, but compared planktonic to biofilms cells [148]. 1D ^1H NMR was combined with GC-MS to

obtain a detailed analysis of metabolite changes caused by the addition of copper (Figure 14). Planktonic cells responded differently to copper stress compared with biofilms. Planktonic cells experienced an oxidative stress response as indicated by changes in the tricarboxylic acid (TCA) cycle, glycolysis, pyruvate, nicotinate and niacinamide metabolism. Conversely, biofilms exhibited changes in exopolysaccharide-related metabolism suggestive of a protective response instead of the reactive changes that occurred in planktonic cells.

Demonstrating a difference between planktonic cells and biofilms by NMR is an important step towards an application in drug discovery. Does a chemical lead inhibit metabolic pathways associated with biofilm formation? Does drug treatment result in a metabolome more similar to planktonic cells despite conditions that induce biofilm formation? Comparing the metabolome of planktonic cells with and without drug treatment may provide an answer to these key questions. Furthermore, the comparative analysis of planktonic cells and biofilms has already identified changes in the activity of specific cellular process correlated with biofilm formation. Thus, proteins essential to exopolysaccharide production and the response to oxidative stress may be novel targets for disrupting biofilm formation. Additionally, the metabolome differences between planktonic cells and biofilms can be used as a diagnostic tool and in the design of treatments. As an illustration, Hall-Stoodley *et al.* describe the application of microbiological culture, polymerase-chain reaction-based diagnostics, direct microscopic examination, fluorescence *in situ* hybridization and immunostaining to characterize middle-ear mucosa biopsy specimens for biofilm morphology [247]. The authors concluded that chronic otitis media in children that require tympanostomy tube placement is biofilm related. NMR metabolomics could provide a similar analysis of biopsy samples. The treatment of biofilm infections that have colonized on medical implants often requires the removal of the infected device [149]. Investigating changes in bacterial metabolomes in response to the different surfaces (e.g., metals and polymers) may contribute to the development of novel materials resistant to biofilm formation [248,249]. Coating or embedding medical devices with antibiotics is a common approach to prevent biofilm infections, but the overuse of antibiotics incurs the risk of inducing the rapid development of resistance [250,251].

Besides characterizing cellular differences through metabolomics, NMR can also be used to explore the overall structure and function of biofilms. Vogt *et al.* used NMR to describe differences in metabolite diffusion within a biofilm [252]. Pulsed-field gradient NMR was used to measure diffusion coefficients for slowly moving water and other components in a *P. aeruginosa* biofilm. Five groups of components including water, glycerol, and polysaccharides, were observed with diffusion coefficients ranging from 1.8×10^{-9} to 5×10^{-13} (m^2s^{-1}) that indicate locations in the biofilm pores or the extracellular polymeric substance. Correspondingly, the complicated structure of a biofilm is a major obstacle to successful treatment with antibiotics. As the above NMR experiment suggests, a biofilm is a diffusion barrier that hinders the infusion and dispersion of antibiotics within a biofilm [253]. This also suggests that the *in vivo* activity of a chemical lead is determined by both its efficient dispersion throughout the biofilm in addition to its intrinsic inhibitor activity.

A recent study by Rogers *et al.* analyzed the synergistic activity of 2-aminoimidazole-derived compounds, a new class of antibiofilm agents that disperse biofilms [254]. Combining a 2-aminoimidazole-derived compound with known antibiotics resulted in a two- to eight-fold reduction in MICs against biofilms of *P. aeruginosa*, *Acinetobacter baumannii*, *Bordetella bronchiseptica* and *S. aureus*. Importantly, the antibiofilm agent actually resensitized MRSA. Walencka *et al.* observed a synergy between salvisipone and aethiopinone from *Salvia sclarea* hairy roots with β -lactam antibiotics. Improved activity was observed against MRSA and multiresistant *S. epidermidis* (i.e., MRSE) [255].

Salvipisone and aethiopinone are postulated to function by altering cell surface hydrophobicity and cell wall/membrane permeability.

NMR metabolomics has also been used to investigate the mechanism by which *S. epidermidis* and *S. aureus* respond to a diverse set of environmental signals to induce the planktonic to biofilm transition. Bacteria have been shown to form biofilms in response to variations in ethanol, oleic acid, glucose, UDP-*N*-acetylglucosamine, subinhibitory concentrations of some antibiotics, anaerobic conditions, iron limitation, high osmolarity and high temperature. Instead of numerous distinct signaling pathways, a series of detailed NMR and molecular biology experiments has demonstrated the presence of a single flexible metabolic signaling pathway centered on the TCA cycle [124,202,256]. 1D ^1H NMR combined with OPLS-DA was used to compare the metabolome of wild-type *S. epidermidis* 1457 and an aconitase mutant strain 1457-*acnA::tetM* under various environmental stressors known to induce biofilm formation (Figure 10). The change in the metabolome of wild-type *S. epidermidis* 1457 in the presence of 4% ethanol, 2% glucose, Fe-limitation and 0.06 $\mu\text{g}/\text{ml}$ of tetracycline was shown to be essentially identical to the aconitase mutant with an inactive TCA cycle. No change in the metabolome was observed for the aconitase mutant with or without the addition of environmental stress factors. 2D ^1H - ^{13}C HSQC experiments combined with bacteria grown in the presence of ^{13}C -glucose were used to generate a detailed analysis of the changes in the *S. epidermidis* metabolome (Figure 11). Consistent with the global changes in the metabolome, wild-type *S. epidermidis* in the presence of the environmental stressors induced the downregulation of metabolites associated with the TCA cycle and the upregulation of metabolites related to PIA production. These results suggest that biofilm formation is regulated by the activity of the TCA cycle. Inactivating the TCA cycle allows the shuttling of key metabolites into PIA production, which is generally necessary for biofilm formation. Similarly, White *et al.* compared the metabolome of wild-type *Salmonella* and a CsgD deletion mutant that prevents production of an extracellular matrix [257]. GC-MS and NMR were used to analyze the metabolome along with an analysis of gene expression. Metabolites associated with gluconeogenesis and major osmoprotectants were upregulated in wild-type *Salmonella*; whereas metabolites associated with the TCA cycle were upregulated in the mutant. Again, this is consistent with the TCA cycle activity regulating biofilm formation. These results suggest that agonists of the TCA cycle would interfere with biofilm formation. Diets or drugs that modulate the nutrient environment may be an approach to prevent bacterial biofilm. As an example, iron limitation downregulates the TCA cycle and induces biofilm formation [258]. The major source of morbidity and mortality in cystic fibrosis patients is *P. aeruginosa* biofilms formed in the lung. High iron concentrations inhibit *P. aeruginosa* biofilms, where chelated sources of iron combined with antibiotics hold promise as a treatment for cystic fibrosis [259,260].

Metabolomics can also be used to construct metabolic pathways, with contributions from proteomics and genomics information. Liebeck *et al.* provided a time-resolved analysis of *S. aureus* during the transition from exponential growth to glucose starvation [261]. The activity of more than 500 proteins and the concentration of 94 metabolites were followed. 1D ^1H NMR was used for the quantification of compounds in the media before inoculation, and at defined time points during cell growth. Intracellular metabolites were measured by GC-MS/LC-MS. In general, changes in the metabolome correlated with changes in the proteome, where the metabolome displayed a larger dynamic range. The most dramatic changes were observed for amino acids. During initial cell growth, glycolysis and protein synthesis were highly active, but as glucose was exhausted gluconeogenesis and the TCA cycle were activated. Again, this is consistent with the TCA cycle activity regulating biofilm formation due to variations in glucose concentrations [80]. Metabolomics is a valuable approach to characterize the state of a system. Nevertheless, incorporating additional complementary data, such as proteomics, significantly enhances the reliability of the

information. Observing a correlation between a metabolite's concentration and a protein's expression level further substantiates the importance and relevance of the protein to the system, such as biofilm formation. From a drug-discovery perspective, this provides substantial corroboration of a potentially new drug-discovery target.

Conclusion

NMR metabolomics can be used to characterize different cell phenotypes, to investigate the underlying biology of biofilm formation, explore the impact of various environmental stress factors on cell biology, analyze the effect of gene mutations, investigate the spatial and temporal structure of a biofilm, and even to generate a 3D image. Notably, the NMR metabolomics methodologies used to study the biochemistry of bacterial biofilms are directly applicable to a drug-discovery effort. NMR metabolomics has been used to identify disease biomarkers and diagnose a disease, to screen for drugs, to evaluate drug activity and toxicity, to identify new therapeutic targets and design new treatments. While the application of NMR for the analysis of changes in the metabolomics is a relatively new endeavor, the technique has already made some significant contributions to our understanding of bacterial biofilms. More importantly, NMR metabolomics holds great promise to significantly contribute to the diagnosis and treatment of biofilm-related diseases, where it may play an important role in personalized medicine.

Future perspective

NMR-based metabolomics is a relatively new technology for systems biology and, correspondingly, has only had a limited use in the investigation of bacterial biofilms. Nevertheless, the application of NMR-based metabolomics to the investigation of bacterial biofilms is only limited by the creativity of the scientific community. In the near future, the further development of the technology will be a primary focus. While NMR-based metabolomics is straightforward in concept, there are numerous practical considerations that can severely complicate the routine application of the technique. A primary issue is our incomplete knowledge of the metabolome; extensive effort is still required to characterize the metabolome from all organisms and populate databases with reference NMR and MS spectra. Similarly, while some progress has been made, a metabolomics software package is still needed that automates and standardizes the processing of metabolomics data, chemometric analysis and model validation, and metabolite identification. Also, the efficient and accurate extraction of metabolites from cell lysates requires continued optimization, and varies depending on the organism under investigation. Developing these protocols is extremely critical to the future success of metabolomics. Additional methodology advancements that will benefit metabolomics include the routine integration of MS and NMR data, and the efficient quantitation of metabolite concentrations from MS and 2D NMR experiments.

Our understanding of bacteria cell biology is far from complete, where NMR-based metabolomics will be an invaluable addition to the study of biofilms and related processes such as programmed cell death, inter-/intra-species communication and pathogenesis. The systematic analysis of the bacterial genome, transcriptome, proteome and metabolome will enable the construction of a detailed network to describe the regulatory and metabolic pathways associated with biofilm formation, progression, evolution and survivability. In addition to enhancing our basic understanding of bacterial biofilms, NMR-based metabolomics will be an invaluable tool for drug discovery, disease diagnosis and personalized medicine. NMR-based metabolomics can be used as part of a drug-discovery screening protocol. Observing an induced change in the bacterial metabolome due to a drug treatment would further validate a chemical lead identified from standard high-throughput

screens [262,263]. In fact, Halouska *et al.* recently demonstrated how NMR-based metabolomics can be used to identify the *in vivo* mechanism of action of a chemical lead [123]. This addresses a major challenge with drug discovery: identifying new drugs with activity against novel therapeutic targets that also avoid common mechanisms of resistance or toxicity. Similarly, characterizing the metabolome of various pathogenic bacterial strains with a correlation to drug susceptibility provides a means to personalize patient treatments. In essence, the identification of characteristic metabolites from the biofluids of infected patients would identify the bacterial strain and preferred form of treatment.

Acknowledgments

This manuscript was supported in part by funds from the NIH (AI087668), and from the NIH National Center for Research Resources (P20 RR-17675).

Key Term

Metabolome	The complete set of small-molecular-weight compounds (<1 kDa) or metabolites found within bacterial cells other cellular organisms tissues or biological samples.
Chemometrics	Application of multivariate statistics to extract information and describe a chemical or biological system. Chemometrics is used to identify spectral features that differentiate a set of NMR spectra to identify changes between two or more metabolomes.
Classes or groups	Collections of NMR spectra obtained from replicate bacterial cell cultures where one experimental variable differs from the other sets of spectra.
Scores plot	Each NMR spectrum is represented as a single point in a scores plot where its relative location is dependent on the spectrum's fit to a predictive model. The resulting clustering pattern identifies the relative similarity and differences between the NMR spectra.
2D ^1H-^{13}C heteronuclear single-quantum correlation spectrum	An NMR experiment that correlates the chemical shift of a proton (^1H) with the chemical shift of a carbon (^{13}C) where the proton and carbon are chemically bonded.

References

- O'Toole G, Kaplan HB, Kolter R. Biofilm formation as microbial development. *Annu. Rev. Microbiol.* 2000; 54:49–79. [PubMed: 11018124]
- Costerton JW. Overview of microbial biofilms. *J. Ind. Microbiol.* 1995; 15(3):137–140. [PubMed: 8519468]
- Costerton JW, Lewandowski Z, Caldwell DE, Korber DR, Lappin-Scott HM. Microbial biofilms. *Annu. Rev. Microbiol.* 1995; 49:711–745. [PubMed: 8561477]
- Simões LC, Simões M, Vieira MJ. Biofilm interactions between distinct bacterial genera isolated from drinking water. *Appl. Environ. Microbiol.* 2007; 73(19):9.
- Nadell CD, Xavier JB, Foster KR. The sociobiology of biofilms. *FEMS Microbiol. Rev.* 2009; 33(1):206–224. [PubMed: 19067751]
- Hiroiyuki K. Biofilm disease. Its clinical manifestation and therapeutic possibilities of macrolides. *Am. J. Med.* 1995; 99(Suppl. 1)(6):26s–30s. [PubMed: 8585531]

7. Klaus-Joerger T, Joerger R, Olsson E, Granqvist C-G. Bacteria as workers in the living factory. Metal-accumulating bacteria and their potential for materials science. *Trends Biotechnol.* 2001; 19(1):15–20. [PubMed: 11146098]
8. Van Loosdrecht MCM, Eikelboom D, Gjaltema A, Mulder A, Tjihuis L, Heijnen JJ. Biofilm structures. *Water Sci. Technol.* 1995; 32(8):35–43.
9. Lens P, Moran AP, Mahony T, Stoodley P, O'flaherty V. Biofilms in medicine, industry and environmental biotechnology. Characteristics, analysis and control. *Mater. Corros.* 2005; 56(1):55.
10. Hughes JM. Addressing emerging infectious disease threats – accomplishments and future plans. *Emerg. Infect. Dis.* 1998; 4(3):360–361. [PubMed: 9716945]
11. Gannon JC. The global infectious disease threat and its implications for the United States. *Prot. Sci.* 2000; 79(4):1–63.
12. Lewis K. Riddle of biofilm resistance. *Antimicrob. Agents Chemother.* 2001; 45(4):999–1007. [PubMed: 11257008]
13. Davies D. Understanding biofilm resistance to antibacterial agents. *Nat. Rev. Drug Discov.* 2003; 2(2):114–122. [PubMed: 12563302]
14. Fischbach MA, Walsh CT. Antibiotics for emerging pathogens. *Science.* 2009; 325(5944):1089–1093. [PubMed: 19713519]
15. Kennedy AD, Otto M, Braughton KR, et al. Epidemic community-associated methicillin-resistant *Staphylococcus aureus*. Recent clonal expansion and diversification. *Proc. Natl Acad. Sci. USA.* 2008; 105(4):S1327–S1332.
16. Kohn WG, Collins AS, Cleveland JL, Harte JA, Eklund KJ, Malvitz DM. Guidelines for infection control in dental health-care settings – 2003. *MMWR Recomm. Rep.* 2003; 52(RR-7):1–61. [PubMed: 14685139]
17. Hetrick EM, Schoenfisch MH. Reducing implant-related infections. Active release strategies. *Chem. Soc. Rev.* 2006; 35(9):780–789. [PubMed: 16936926]
18. Parsek MR, Singh PK. Bacterial biofilms. An emerging link to disease pathogenesis. *Annu. Rev. Microbiol.* 2003; 57:677–701. [PubMed: 14527295]
19. Savage DC. Microbial ecology of the gastrointestinal tract. *Annu. Rev. Microbiol.* 1977; 31:107–133. [PubMed: 334036]
20. Marsh PD, Moter A, Devine DA. Dental plaque biofilms. communities, conflict and control. *Periodontology* 2000. 2011; 55(1):16–35. [PubMed: 21134226]
21. Nicolle LE. Catheter-related urinary tract infection. *Drugs Aging.* 2005; 22(8):627–639. [PubMed: 16060714]
22. Campoccia D, Montanaro L, Baldassarri L, An YH, Arciola CR. Antibiotic resistance in *Staphylococcus aureus* and *Staphylococcus epidermidis* clinical isolates from implant orthopedic infections. *Int. J. Artif. Organs.* 2005; 28(11):1186–1191. [PubMed: 16353126]
23. Nishimura S, Tsurumoto T, Yonekura A, Adachi K, Shindo H. Antimicrobial susceptibility of *Staphylococcus aureus* and *Staphylococcus epidermidis* biofilms isolated from infected total hip arthroplasty cases. *J. Orthop. Sci.* 2006; 11(1):46–50. [PubMed: 16437348]
24. Laplante KL, Woodmansee S. Activities of daptomycin and vancomycin alone and in combination with rifampin and gentamicin against biofilm-forming methicillin-resistant *Staphylococcus aureus* isolates in an experimental model of endocarditis. *Antimicrob. Agents Chemother.* 2009; 53(9):3880–3886. [PubMed: 19564363]
25. Padera RF. Infection in ventricular assist devices. The role of biofilm. *Cardiovasc. Pathol.* 2006; 15(5):264–270. [PubMed: 16979033]
26. Vadyvaloo V, Otto M. Molecular genetics of *Staphylococcus epidermidis* biofilms on indwelling medical devices. *Int. J. Artif. Organs.* 2005; 28(11):1069–1078. [PubMed: 16353113]
27. Sohail MR, Usfan DZ, Khan AH, et al. Management and outcome of permanent pacemaker and implantable cardioverter-defibrillator infections. *J. Am. Coll. Cardiol.* 2007; 49(18):1851–1859. [PubMed: 17481444]
28. Cunha BA, Eisenstein LE, Hamid NS. Pacemaker-induced *Staphylococcus aureus* mitral valve acute bacterial endocarditis complicated by persistent bacteremia from a coronary stent. Cure with prolonged/high-dose daptomycin without toxicity. *Heart Lung.* 2006; 35(3):207–211. [PubMed: 16701116]

29. Peacock SJ, Curtis N, Berendt AR, Bowler IC, Winearls CG, Maxwell P. Outcome following haemodialysis catheter-related *Staphylococcus aureus* bacteraemia. J. Hosp. Infect. 1999; 41(3): 223–228. [PubMed: 10204125]
30. Knudsen AM, Rosdahl VT, Espersen F, Frimodt-Moller N, Skinhoj P, Bentzon MW. Catheter-related *Staphylococcus aureus* infections. J. Hosp. Infect. 1993; 23(2):123–131. [PubMed: 8097216]
31. Ghanem GA, Boktour M, Warneke C, et al. Catheter-related *Staphylococcus aureus* bacteremia in cancer patients. High rate of complications with therapeutic implications. Medicine. 2007; 86(1): 54–60. [PubMed: 17220756]
32. Hall RE, Ash AS, Ghali WA, Moskowitz MA. Hospital cost of complications associated with coronary artery bypass graft surgery. Am. J. Cardiol. 1997; 79(12):1680–1682. [PubMed: 9202364]
33. Simon D, Fischer S, Grossman A, et al. Left ventricular assist device-related infection. treatment and outcome. Clin. Infect. Dis. 2005; 40(8):1108–1115. [PubMed: 15791509]
34. Holman WL, Rayburn BK, McGiffin DC, et al. Infection in ventricular assist devices. Prevention and treatment. Ann. Thorac. Surg. 2003; 75(6 Suppl.):S48–S57. [PubMed: 12820735]
35. Herrmann M, Weyand M, Greshake B, et al. Left ventricular assist device infection is associated with increased mortality but is not a contraindication to transplantation. Circulation. 1997; 95(4): 814–817. [PubMed: 9054736]
36. Fischer SA, Trenholme GM, Costanzo MR, Piccione W. Infectious complications in left ventricular assist device recipients. Clin. Infect. Dis. 1997; 24(1):18–23. [PubMed: 8994750]
37. Holman WL, Park SJ, Long JW, et al. Infection in permanent circulatory support experience from the rematch trial. J. Heart Lung Transplant. 2004; 23(12):1359–1365. [PubMed: 15607664]
38. Mohamed JA, Huang DB. Biofilm formation by enterococci. J. Med. Microbiol. 2007; 56(Pt 12): 1581–1588. [PubMed: 18033823]
39. Olmstead S, Meiss D, Ralston J. Life on the edge. The clinical implications of gastrointestinal biofilm. Newsletter Spring. 2003; 2003
40. O'Toole GA. To build a biofilm. J. Bacteriol. 2003; 185(9):2687–2689. [PubMed: 12700246]
41. Stewart P, Franklin M. Physiological heterogeneity in biofilms. Nat. Rev. Microbiol. 2008; 6(3): 199–210. [PubMed: 18264116]
42. Davey ME, O'Toole GA. Microbial biofilms. from ecology to molecular genetics. Microbiol. Mol. Biol. Rev. 2000; 64(4):847–867. [PubMed: 11104821]
43. Hall-Stoodley L, Costerton JW, Stoodley P. Bacterial biofilms. From the natural environment to infectious diseases. Nat. Rev. Microbiol. 2004; 2(2):95–108. [PubMed: 15040259]
44. Stoodley P, Sauer K, Davies DG, Costerton JW. Biofilms as complex differentiated communities. Annu. Rev. Microbiol. 2002; 56:187–209. [PubMed: 12142477]
45. Lawrence JR, Korber DR, Hoyle BD, Costerton JW, Caldwell DE. Optical sectioning of microbial biofilms. J. Bacteriol. 1991; 173(20):6558–6567. [PubMed: 1917879]
46. Wickham JR, Rice CV. Solid-state NMR studies of bacterial lipoteichoic acid adsorption on different surfaces. Solid State Nucl. Magn. Reson. 2008; 34(3):154–161. [PubMed: 18617374]
47. Hall-Stoodley L, Stoodley P. Biofilm formation and dispersal and the transmission of human pathogens. Trends Microbiol. 2005; 13(1):7–10. [PubMed: 15639625]
48. Yasuda H. Bacterial biofilms and infectious diseases. Trends Glycosci. Glycotechnol. 1996; 8(44): 409–417.
49. O'Toole GA, Gibbs KA, Hager PW, Phibbs PV Jr, Kolter R. The global carbon metabolism regulator Crc is a component of a signal transduction pathway required for biofilm development by *Pseudomonas aeruginosa*. J. Bacteriol. 2000; 182(2):425–431. [PubMed: 10629189]
50. Stanley PM. Factors affecting the irreversible attachment of *Pseudomonas aeruginosa* to stainless steel. Can. J. Microbiol. 1983; 29(11):1493–1499. [PubMed: 6423248]
51. Watnick PI, Kolter R. Steps in the development of a *Vibrio cholerae* El Tor biofilm. Mol. Microbiol. 1999; 34(3):586–595. [PubMed: 10564499]

52. Maira-Litran T, Kropec A, Abeygunawardana C, et al. Immunochemical properties of the Staphylococcal poly-*N*-acetylglucosamine surface polysaccharide. *Infect. Immun.* 2002; 70(8): 4433–4440. [PubMed: 12117954]
53. Mack D, Fischer W, Krokotsch A, et al. The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear beta-1,6-linked glucosaminoglycan. purification and structural analysis. *J. Bacteriol.* 1996; 178(1):175–183. [PubMed: 8550413]
54. Van Loosdrecht MC, Lyklema J, Norde W, Zehnder AJ. Influence of interfaces on microbial activity. *Microbiol. Rev.* 1990; 54(1):75–87. [PubMed: 2181260]
55. Debeer D, Stoodley P, Roe FL, Lewandowski Z. Effects of biofilm structures on oxygen distribution and mass transport. *Biotech. Bioeng.* 1994; 43(11):1131–1138.
56. Lawrence JR, Neu TR. Confocal laser scanning microscopy for analysis of microbial biofilms. *Methods Enzymol.* 1999; 310:131–144. [PubMed: 10547787]
57. Webb JS, Givskov M, Kjelleberg S. Bacterial biofilms. prokaryotic adventures in multicellularity. *Curr. Opin. Microbiol.* 2003; 6(6):578–585. [PubMed: 14662353]
58. Krasovec R, Jerman I. Bacterial multicellularity as a possible source of antibiotic resistance. *Med. Hypotheses.* 2003; 60(4):484–488. [PubMed: 12615504]
59. Ilyina TS, Romanova YM, Gintsburg AL. Biofilms as a mode of existence of bacteria in external environment and host body. The phenomenon, genetic control, and regulation systems of development. *Russ. J. Genet.* 2004; 40(11):1189–1198.
60. Tanaka G, Shigeta M, Komatsuzawa H, Sugai M, Suganaka H, Usui T. Effect of the growth rate of *Pseudomonas aeruginosa* biofilms on the susceptibility to antimicrobial agents. β -lactams and fluoroquinolones. *Chemotherapy.* 1999; 45(1):28–36. [PubMed: 9876207]
61. Ashby MJ, Neale JE, Knott SJ, Critchley IA. Effect of antibiotics on non-growing planktonic cells and biofilms of *Escherichia coli*. *J. Antimicrob. Chemother.* 1994; 33(3):443–452. [PubMed: 8040110]
62. Picioreanu C, Van Loosdrecht MCM, Heijnen JJ. Discrete-differential modelling of biofilm structure. *Water Sci. Technol.* 1999; 39(7):8.
63. Hao X, Heijnen JJ, Van Loosdrecht MCM. Sensitivity analysis of a biofilm model describing a one-stage completely autotrophic nitrogen removal (CANON) process. *Biotechnol. Bioeng.* 2002; 77(3):266–277. [PubMed: 11753935]
64. Park S, Bae W, Rittmann BE. Multi-species nitrifying biofilm model (MSNB) including free ammonia and free nitrous acid inhibition and oxygen limitation. *Biotechnol. Bioeng.* 2010; 105(6): 1115–1130. [PubMed: 19998282]
65. Chaudhry M, Beg S. A review on the mathematical modeling of biofilm processes. *advances in fundamentals of biofilm modeling.* *Chem. Eng. Technol.* 1998; 21(9):701–710.
66. Nelson YM, Lion LW, Shuler ML, Ghiorse WC. Modeling oligotrophic biofilm formation and lead adsorption to biofilm components. *Environ. Sci. Technol.* 1996; 30(6):2027–2035.
67. Costerton JW, Stewart PS, Greenberg EP. Bacterial Biofilms. A common cause of persistent infections. *Science.* 1999; 284(5418):1318–1322. [PubMed: 10334980]
68. O'Toole GA, Pratt LA, Watnick PI, Newman DK, Weaver VB, Kolter R. Genetic approaches to study of biofilms. *Methods Enzymol.* 1999; 310:91–109. [PubMed: 10547784]
69. Pratt LA, Kolter R. Genetic analyses of bacterial biofilm formation. *Curr. Opin. Microbiol.* 1999; 2(6):598–603. [PubMed: 10607630]
70. Boles BR, Thoendel M, Roth AJ, Horswill AR. Identification of genes involved in polysaccharide-independent *Staphylococcus aureus* biofilm formation. *PLoS ONE.* 2010; 5(4):e10146. [PubMed: 20418950]
71. Chagneau C, Saier MH Jr. Biofilm-defective mutants of *Bacillus subtilis*. *J. Mol. Microbiol. Biotechnol.* 2004; 8(3):177–188. [PubMed: 16088219]
72. Kristich CJ, Nguyen VT, Le T, Barnes AMT, Grindle S, Dunny GM. Development and use of an efficient system for random mariner transposon mutagenesis to identify novel genetic determinants of biofilm formation in the core *Enterococcus faecalis* genome. *Appl. Environ. Microbiol.* 2008; 74(11):3377–3386. [PubMed: 18408066]

73. Li J, Wang N. Genome-wide mutagenesis of *Xanthomonas axonopodis* pv. citri reveals novel genetic determinants and regulation mechanisms of biofilm formation. PLoS ONE. 2011; 6(7):e21804. [PubMed: 21750733]
74. Quoc PHT, Genevaux P, Pajunen M, et al. Isolation and characterization of biofilm formation-defective mutants of *Staphylococcus aureus*. Infect. Immun. 2007; 75(3):1079–1088. [PubMed: 17158901]
75. Richard ML, Nobile CJ, Bruno VM, Mitchell AP. *Candida albicans* biofilm-defective mutants. Eukaryotic Cell. 2005; 4(8):1493–1502. [PubMed: 16087754]
76. Yoshida A, Kuramitsu HK. Multiple *Streptococcus mutans* genes are involved in biofilm formation. Appl. Environ. Microbiol. 2002; 68(12):6283–6291. [PubMed: 12450853]
77. Beloin C, Ghigo J-M. Finding gene expression patterns in bacterial biofilms. Trends Microbiol. 2005; 13(1):16–19. [PubMed: 15639627]
78. Knobloch JK, Bartscht K, Sabottke A, Rohde H, Feucht HH, Mack D. Biofilm formation by *Staphylococcus epidermidis* depends on functional RsbU, an activator of the sigB operon: differential activation mechanisms due to ethanol and salt stress. J. Bacteriol. 2001; 183(8):2624–2633. [PubMed: 11274123]
79. Campbell IM, Crozier DN, Pawagi AB, Buivids IA. *In vitro* response of *Staphylococcus aureus* from cystic fibrosis patients to combinations of linoleic and oleic acids added to nutrient medium. J. Clin. Microbiol. 1983; 18(2):408–415. [PubMed: 6619290]
80. Mack D, Siemssen N, Laufs R. 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. 1992; 60(5):2048–2057. [PubMed: 1314224]
81. Gerke C, Kraft A, Sussmuth R, Schweitzer O, Gotz F. Characterization of the N-acetylglucosaminyltransferase activity involved in the biosynthesis of the *Staphylococcus epidermidis* polysaccharide intercellular adhesin. J. Biol. Chem. 1998; 273(29):18586–18593. [PubMed: 9660830]
82. Rachid S, Ohlsen K, Witte W, Hacker J, Ziebuhr W. Effect of subinhibitory antibiotic concentrations on polysaccharide intercellular adhesin expression in biofilm-forming *Staphylococcus epidermidis*. Antimicrob. Agents Chemother. 2000; 44(12):3357–3363. [PubMed: 11083640]
83. Cramton SE, Ulrich M, Gotz F, Doring G. Anaerobic conditions induce expression of polysaccharide intercellular adhesin in *Staphylococcus aureus* and *Staphylococcus epidermidis*. Infect. Immun. 2001; 69(6):4079–4085. [PubMed: 11349079]
84. Deighton M, Borland R. Regulation of slime production in *Staphylococcus epidermidis* by iron limitation. Infect. Immun. 1993; 61(10):4473–4479. [PubMed: 8406839]
85. Elci S, Atmaca S, Gul K. Effect of iron limitation on the amount of slime produced by strains of *Staphylococcus epidermidis*. Cytobios. 1995; 84(38–339):141–146. [PubMed: 8759151]
86. Evans E, Brown MR, Gilbert P. Iron chelator, exopolysaccharide and protease production in *Staphylococcus epidermidis*: a comparative study of the effects of specific growth rate in biofilm and planktonic culture. Microbiology. 1994; 140(Pt 1):153–157. [PubMed: 8162184]
87. Rachid S, Ohlsen K, Wallner U, Hacker J, Hecker M, Ziebuhr W. Alternative transcription factor sigma(B) is involved in regulation of biofilm expression in a *Staphylococcus aureus* mucosal isolate. J. Bacteriol. 2000; 182(23):6824–6826. [PubMed: 11073930]
88. Rowley G, Spector M, Kormanec J, Roberts M. Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens. Nat. Rev. Microbiol. 2006; 4(5):383–394. [PubMed: 16715050]
89. Chaturongakul S, Raengpradub S, Wiedmann M, Boor KJ. Modulation of stress and virulence in *Listeria monocytogenes*. Trends Microbiol. 2008; 16(8):388–396. [PubMed: 18619843]
90. Sachdeva P, Misra R, Tyagi AK, Singh Y. The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. FEBS J. 2010; 277(3):605–626. [PubMed: 19951358]
91. Senn MM, Bischoff M, Von Eiff C, Berger-Bachi B. sigmaB activity in a *Staphylococcus aureus* hemB mutant. J. Bacteriol. 2005; 187(21):7397–7406. [PubMed: 16237023]
92. Gill SR, Fouts DE, Archer GL, et al. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a

- biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. J. Bacteriol. 2005; 187(7):2426–2438. [PubMed: 15774886]
93. Stock AM, Robinson VL, Goudreau PN. Two-component signal transduction. Annu. Rev. Biochem. 2000; 69:183–215. [PubMed: 10966457]
94. Li M, Villaruz AE, Vadyvaloo V, Sturdevant DE, Otto M. AI-2-dependent gene regulation in *Staphylococcus epidermidis*. BMC Microbiol. 2008; 8:4. [PubMed: 18182108]
95. Miller MB, Bassler BL. Quorum sensing in bacteria. Annu. Rev. Microbiol. 2001; 55:165–199. [PubMed: 11544353]
96. Jayaraman A, Wood TK. Bacterial quorum sensing: signals, circuits, and implications for biofilms and disease. Annu. Rev. Biomed. Eng. 2008; 10:145–167. [PubMed: 18647113]
97. Sturme MH, Kleerebezem M, Nakayama J, Akkermans AD, Vaughn EE, De Vos WM. Cell to cell communication by autoinducing peptides in Gram-positive bacteria. Antonie Van Leeuwenhoek. 2002; 81(1–4):233–243. [PubMed: 12448722]
98. Xavier KB, Bassler BL. LuxS quorum sensing: more than just a numbers game. Curr. Opin. Microbiol. 2003; 6(2):191–197. [PubMed: 12732311]
99. Chauhan N, Calderone R. Two-component signal transduction proteins as potential drug targets in medically important fungi. Infect. Immun. 2008; 76(11):4795–4803. [PubMed: 18765727]
100. Raffa RB, Iannuzzo JR, Levine DR, et al. Bacterial communication (“quorum sensing”) via ligands and receptors. a novel pharmacologic target for the design of antibiotic drugs. J. Pharmacol. Exp. Ther. 2005; 312(2):417–423. [PubMed: 15528454]
101. Kleerebezem M, Quadri LE, Kuipers OP, De Vos WM. Quorum sensing by peptide pheromones and two-component signal-transduction systems in Gram-positive bacteria. Mol. Microbiol. 1997; 24(5):895–904. [PubMed: 9219998]
102. Gotoh Y, Eguchi Y, Watanabe T, Okamoto S, Doi A, Utsumi R. Two-component signal transduction as potential drug targets in pathogenic bacteria. Curr. Opin. Microbiol. 2010; 13(2): 232–239. [PubMed: 20138000]
103. Fuqua C, Parsek MR, Greenberg EP. Regulation of gene expression by cell-to-cell communication: acyl-homoserine lactone quorum sensing. Annu. Rev. Genet. 2001; 35:439–468. [PubMed: 11700290]
104. Taga ME, Bassler BL. Chemical communication among bacteria. Proc. Natl Acad. Sci. USA. 2003; 100(Suppl. 2):14549–14554. [PubMed: 12949263]
105. Passador L, Tucker KD, Guertin KR, Journet MP, Kende AS, Iglewski BH. Functional analysis of the *Pseudomonas aeruginosa* autoinducer PAI. J. Bacteriol. 1996; 178(20):5995–6000. [PubMed: 8830697]
106. Schaefer AL, Hanzelka BL, Eberhard A, Greenberg EP. Quorum sensing in *Vibrio fischeri*: probing autoinducer-LuxR interactions with autoinducer analogs. J. Bacteriol. 1996; 178(10): 2897–2901. [PubMed: 8631679]
107. Rasko DA, Moreira CG, Li DR, et al. Targeting QseC signaling and virulence for antibiotic development. Science. 2008; 321(5892):1078–1080. [PubMed: 18719281]
108. Balaban N, Cirioni O, Giacometti A, et al. Treatment of *Staphylococcus aureus* biofilm infection by the quorum-sensing inhibitor RIP. Antimicrob. Agents Chemother. 2007; 51(6):2226–2229. [PubMed: 17371825]
109. Roy V, Smith JA, Wang J, Stewart JE, Bentley WE, Sintim HO. Synthetic analogs tailor native AI-2 signaling across bacterial species. J. Am. Chem. Soc. 2010; 132(32):11141–11150. [PubMed: 20698680]
110. Roy V, Adams BL, Bentley WE. Developing next generation antimicrobials by intercepting AI-2 mediated quorum sensing. Enzyme Microb. Technol. 2011; 49(2):113–123. [PubMed: 22112397]
111. Rui F, Marques JC, Miller ST, Maycock CD, Xavier KB, Ventura MR. Stereochemical diversity of AI-2 analogs modulates quorum sensing in *Vibrio harveyi* and *Escherichia coli*. Bioorg. Med. Chem. 2012; 20(1):249–256. [PubMed: 22137598]
112. Tsuchikama K, Lowery CA, Janda KD. Probing autoinducer-2 based quorum sensing: the biological consequences of molecules unable to traverse equilibrium states. J. Org. Chem. 2011; 76(17):6981–6989. [PubMed: 21678949]

113. Lowery CA, McKenzie KM, Qi L, Meijler MM, Janda KD. Quorum sensing in *Vibrio harveyi*: probing the specificity of the LuxP binding site. *Bioorg. Med. Chem. Lett.* 2005; 15(9):2395–2398. [PubMed: 15837332]
114. Reo NV. NMR-based metabolomics. *Drug Chem Toxicol.* 2002; 25(4):375–382. [PubMed: 12378948]
115. Fiehn O. Metabolomics – the link between genotypes and phenotypes. *Plant Mol. Biol.* 2002; 48(1–2):155–171. [PubMed: 11860207]
116. Oliver SG, Winson MK, Kell DB, Baganz F. Systematic functional analysis of the yeast genome. *Trends Biotechnol.* 1998; 16(9):373–378. [PubMed: 9744112]
117. Beckonert O, Keun HC, Ebbels TM, et al. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Protoc.* 2007; 2(11):2692–2703. [PubMed: 18007604]
118. Rubakhin SS, Romanova EV, Nemes P, Sweedler JV. Profiling metabolites and peptides in single cells. *Nat. Methods.* 2011; 8(4 Suppl.):S20–S29. [PubMed: 21451513]
119. Roos V, Klemm P. Global gene expression profiling of the asymptomatic bacteriuria *Escherichia coli* strain 83972 in the human urinary tract. *Infect. Immun.* 2006; 74(6):3565–3575. [PubMed: 16714589]
120. Vilain S, Cosette P, Zimmerlin I, Dupont JP, Junter GA, Jouenne T. Biofilm proteome: homogeneity or versatility? *J. Proteome Res.* 2004; 3(1):132–136. [PubMed: 14998174]
121. Peregrin-Alvarez JM, Sanford C, Parkinson J. The conservation and evolutionary modularity of metabolism. *Genome Biol.* 2009; 10(6):R63.
122. Monds, RD.; O'Toole, GA. Metabolites as intercellular signals for regulation of community-level traits. In: Winans, SC.; Bassler, BL., editors. *Chemical Communication Among Bacteria*. Washington DC, USA: ASM Press; 2008. p. 105-130.
123. Halouska S, Fenton RJ, Barletta RG, Powers R. Predicting the *in vivo* mechanism of action for drug leads using NMR metabolomics. *ACS Chem. Biol.* 2012; 7(1):166–171. [PubMed: 22007661]
124. 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(8):3743–3754. [PubMed: 21692534]
125. Shanaiah, N.; Zhang, S.; Desilva, MA.; Raftery, D. NMR-based metabolomics for biomarker discovery. In: Wang, F., editor. *Biomarker Methods in Drug Discovery and Development*. Springer, NY, USA: 2008. p. 341-368.
126. Bino RJ, Hall RD, Fiehn O, et al. Potential of metabolomics as a functional genomics tool. *Trends Plant Sci.* 2004; 9(9):418–425. [PubMed: 15337491]
127. Powers R. NMR metabolomics and drug discovery. *Magn. Reson. Chem.* 2009; 47(Suppl. 1):S2–S11. [PubMed: 19504464]
128. Coen M, Holmes E, Lindon JC, Nicholson JK. NMR-based metabolic profiling and metabonomic approaches to problems in molecular toxicology. *Chem. Res. Toxicol.* 2008; 21(1):9–27. [PubMed: 18171018]
129. Vinaixa M, Angel RM, Rull A, et al. Metabolomic assessment of the effect of dietary cholesterol in the progressive development of fatty liver disease. *J. Proteome Res.* 2010; 9(5):2527–2538. [PubMed: 20402505]
130. Griffin JL, Atherton H, Shockcor J, Atzori L. Metabolomics as a tool for cardiac research. *Nat. Rev. Cardiol.* 2011; 8(11):630–643. [PubMed: 21931361]
131. Li H, Wang L, Yan X, et al. A proton nuclear magnetic resonance metabonomics approach for biomarker discovery in nonalcoholic fatty liver disease. *J. Proteome Res.* 2011; 10(6):2797–2806. [PubMed: 21563774]
132. McClay JL, Adkins DE, Isern NG, et al. 1H Nuclear magnetic resonance metabolomics analysis identifies novel urinary biomarkers for lung function. *J. Proteome Res.* 2010; 9(6):3083–3090. [PubMed: 20408573]
133. Sofia M, Maniscalco M, De LG, Paris D, Melck D, Motta A. Exploring airway diseases by NMR-based metabonomics: a review of application to exhaled breath condensate. *J. Biomed. Biotechnol.* 2011; 2011(2011) 403260.

134. Macintyre DA, Jimenez B, Lewintre EJ, et al. Serum metabolome analysis by ^1H -NMR reveals differences between chronic lymphocytic leukaemia molecular subgroups. *Leukemia*. 2010; 24(4):788–797. [PubMed: 20090781]
135. Carrola J, Rocha CM, Barros AS, et al. Metabolic signatures of lung cancer in biofluids. NMR-based metabonomics of urine. *J. Proteome Res.* 2011; 10(1):221–230. [PubMed: 21058631]
136. Napoli C, Sperandio N, Lawlor RT, Scarpa A, Molinari H, Assfalg M. Urine metabolic signature of pancreatic ductal adenocarcinoma by ^1H nuclear magnetic resonance: identification, mapping, and evolution. *J. Proteome Res.* 2012; 11(2):1274–1283. [PubMed: 22066465]
137. Bertini I, Cacciatore S, Jensen BV, et al. Metabolomic NMR fingerprinting to identify and predict survival of patients with metastatic colorectal cancer. *Cancer Res.* 2012; 72(1):356–364. [PubMed: 22080567]
138. Kork F, Holthues J, Hellweg R, et al. A possible new diagnostic biomarker in early diagnosis of Alzheimer's disease. *Curr. Alzheimer Res.* 2009; 6(6):519–524. [PubMed: 19747162]
139. Sinclair AJ, Viant MR, Ball AK, et al. NMR-based metabolomic analysis of cerebrospinal fluid and serum in neurological diseases – a diagnostic tool? *NMR Biomed.* 2010; 23(2):123–132. [PubMed: 19691132]
140. Blasco H, Corcia P, Moreau C, et al. ^1H -NMR-based metabolomic profiling of CSF in early amyotrophic lateral sclerosis. *PLoS ONE*. 2010; 5(10):e13223. [PubMed: 20949041]
141. Stoop MP, Coulier L, Rosenling T, et al. Quantitative proteomics and metabolomics analysis of normal human cerebrospinal fluid samples. *Mol. Cell. Proteomics.* 2010; 9(9):2063–2075. [PubMed: 20811074]
142. Tiziani S, Kang Y, Choi JS, Roberts W, Paternostro G. Metabolomic high-content nuclear magnetic resonance-based drug screening of a kinase inhibitor library. *Nat. Commun.* 2011; 2:545. [PubMed: 22109519]
143. Lindon JC, Keun HC, Ebbels TMD, Pearce JMT, Holmes E, Nicholson JK. The Consortium for Metabonomic Toxicology (COMET): aims, activities and achievements. *Pharmacogenomics.* 2005; 6(7):691–699. [PubMed: 16207146]
144. Robertson DG, Reilly MD, Baker JD. Metabonomics in pharmaceutical discovery and development. *J. Proteome Res.* 2007; 6(2):526–539. [PubMed: 17269709]
145. Wishart DS. Applications of metabolomics in drug discovery and development. *Drugs Res. Dev.* 2008; 9(5):307–322.
146. Fernie AR, Trethewey RN, Krotzky AJ, Willmitzer L. Metabolite profiling: from diagnostics to systems biology. *Nat. Rev. Mol. Cell Biol.* 2004; 5(9):763–769. [PubMed: 15340383]
147. Ellis DI, Dunn WB, Griffin JL, Allwood JW, Goodacre R. Metabolic fingerprinting as a diagnostic tool. *Pharmacogenomics.* 2007; 8(9):1243–1266. [PubMed: 17924839]
148. Booth SC, Workentine ML, Wen J, et al. Differences in metabolism between the biofilm and planktonic response to metal stress. *J. Proteome Res.* 2011; 10(7):3190–3199. [PubMed: 21561166]
149. Gjersing EL, Herberg JL, Horn J, Schaldach CM, Maxwell RS. NMR metabolomics of planktonic and biofilm modes of growth in *Pseudomonas aeruginosa*. *Anal. Chem.* 2007; 79(21):8037–8045. [PubMed: 17915964]
150. Zulak KG, Weljie AM, Vogel HJ, Facchini PJ. Quantitative ^1H NMR metabolomics reveals extensive metabolic reprogramming of primary and secondary metabolism in elicitor-treated opium poppy cell cultures. *BMC Plant Biol.* 2008; 8:5. [PubMed: 18211706]
151. Pears MR, Cooper JD, Mitchison HM, Mortishire-Smith RJ, Pearce DA, Griffin JL. High resolution ^1H NMR-based metabolomics indicates a neurotransmitter cycling deficit in cerebral tissue from a mouse model of Batten disease. *J. Biol. Chem.* 2005; 280(52):42508–42514. [PubMed: 16239221]
152. Sadykov MR, Zhang B, Halouska S, et al. Using NMR metabolomics to investigate tricarboxylic acid cycle-dependent signal transduction in *Staphylococcus epidermidis*. *J. Biol. Chem.* 2010; 285(47):36616–36624. [PubMed: 20861019]
153. Kjeldahl K, Bro R. Some common misunderstandings in chemometrics. *J. Chemom.* 2010; 24(7–8):558–564.

154. Kumar A, Bhardwaj A. Methods in cell separation for biomedical application: cryogels as a new tool. *Biomed. Mater.* 2008; 3(3) 034008.
155. Bolten CJ, Kiefer P, Letisse F, Portais JC, Wittmann C. Sampling for metabolome analysis of microorganisms. *Anal. Chem.* 2007; 79(10):3843–3849. [PubMed: 17411014]
156. Faijes M, Mars AE, Smid EJ. Comparison of quenching and extraction methodologies for metabolome analysis of *Lactobacillus plantarum*. *Microb. Cell Fact.* 2007; 6:27. [PubMed: 17708760]
157. Maharjan RP, Ferenci T. Global metabolite analysis: the influence of extraction methodology on metabolome profiles of *Escherichia coli*. *Anal. Biochem.* 2003; 313(1):145–154. [PubMed: 12576070]
158. Canelas A, Ras C, Ten Pierick A, Van Dam J, Heijnen J, Van Gulik W. Leakage-free rapid quenching technique for yeast metabolomics. *Metabolomics.* 2008; 4(3):226–239.
159. Wu XH, Yu HL, Ba ZY, Chen JY, Sun HG, Han BZ. Sampling methods for NMR-based metabolomics of *Staphylococcus aureus*. *Biotechnol. J.* 2010; 5(1):75–84. [PubMed: 19824021]
160. Tredwell GD, Edwards-Jones B, Leak DJ, Bundy JG. The development of metabolomic sampling procedures for *Pichia pastoris*, and baseline metabolome data. *PLoS ONE.* 2011; 6(1):e16286. [PubMed: 21283710]
161. Rabinowitz JD. Cellular metabolomics of *Escherichia coli*. *Expert Rev. Proteomics.* 2007; 4(2): 187–198. [PubMed: 17425455]
162. Shaw L, Golonka E, Potempa J, Foster SJ. The role and regulation of the extracellular proteases of *Staphylococcus aureus*. *Microbiology.* 2004; 150(Pt 1):217–228. [PubMed: 14702415]
163. Wittmann C, Kromer JO, Kiefer P, Binz T, Heinze E. Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria. *Anal. Biochem.* 2004; 327(1):135–139. [PubMed: 15033521]
164. Weckwerth W, Wenzel K, Fiehn O. Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their co-regulation in biochemical networks. *Proteomics.* 2004; 4(1):78–83. [PubMed: 14730673]
165. Wu H, Southam AD, Hines A, Viant MR. High-throughput tissue extraction protocol for NMR- and MS-based metabolomics. *Anal. Biochem.* 2008; 372(2):204–212. [PubMed: 17963684]
166. Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* 2009; 5(8):593–599. [PubMed: 19561621]
167. Kell DB. Metabolomics and systems biology: making sense of the soup. *Curr. Opin. Microbiol.* 2004; 7(3):296–307. [PubMed: 15196499]
168. Blow N. Metabolomics: biochemistry's new look. *Nature.* 2008; 455(7213):697–700. [PubMed: 18833281]
169. Wilson ID, Plumb R, Granger J, Major H, Williams R, Lenz EM. HPLC-MS-based methods for the study of metabolomics. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 2005; 817(1): 67–76.
170. Freeman, R. *Magnetic Resonance in Chemistry and Medicine*. London, UK: Oxford University Press; 2003.
171. Moskau D, Ritcher C, Kovaks H, et al. Highest sensitivity for cutting-edge NMR applications. 600 MHz CryoProbes. *Spectra Analyse.* 2003; 32(231):39–41.
172. Keun HC, Beckonert O, Griffin JL, et al. Cryogenic probe ¹³C NMR spectroscopy of urine for metabolomic studies. *Anal. Chem.* 2002; 74(17):4588–4593. [PubMed: 12236374]
173. Carrieri D, Mcneely K, De Roo AC, Bennette N, Pelczer I, Dismukes GC. Identification and quantification of water-soluble metabolites by cryoprobe-assisted nuclear magnetic resonance spectroscopy applied to microbial fermentation. *Magn. Reson. Chem.* 2009; 47(S1):S138–S146. [PubMed: 19415773]
174. Pan Z, Raftery D. Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Analyt. Bioanal. Chem.* 2007; 387(2):525–527. [PubMed: 16955259]
175. Lanza IR, Zhang S, Ward LE, Karakelides H, Raftery D, Nair KS. Quantitative metabolomics by ¹H-NMR and LC-MS/MS confirms altered metabolic pathways in diabetes. *PLoS ONE.* 2010; 5(5):e10538. [PubMed: 20479934]

176. Pan Z, Raftery D. Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Anal. Bioanal. Chem.* 2007; 387(2):525–527. [PubMed: 16955259]
177. Gu H, Pan Z, Xi B, Asiago V, Musselman B, Raftery D. Principal component directed partial least squares analysis for combining nuclear magnetic resonance and mass spectrometry data in metabolomics: application to the detection of breast cancer. *Anal. Chim. Acta.* 2011; 686(1–2): 57–63. [PubMed: 21237308]
178. Atherton HJ, Bailey NJ, Zhang W, et al. A combined ^1H -NMR spectroscopy- and mass spectrometry-based metabolomic study of the PPAR- α null mutant mouse defines profound systemic changes in metabolism linked to the metabolic syndrome. *Physiol. Genomics.* 2006; 27(2):178–186. [PubMed: 16868074]
179. Wang H, Manicke NE, Yang Q, et al. Direct analysis of biological tissue by paper spray mass spectrometry. *Anal. Chem.* 2011; 83(4):1197–1201. [PubMed: 21247069]
180. Weljie AM, Newton J, Mercier P, Carlson E, Slupsky CM. Targeted profiling: quantitative analysis of ^1H NMR metabolomics data. *Anal. Chem.* 2006; 78(13):4430–4442. [PubMed: 16808451]
181. Moestue S, Sitter B, Bathen TF, Tessem M-B, Gribbestad IS. HR MAS MR spectroscopy in metabolic characterization of cancer. *Curr. Top. Med. Chem.* 2011; 11(1):2–26. [PubMed: 20809888]
182. Sitter B, Bathen TF, Singstad TE, et al. Quantification of metabolites in breast cancer patients with different clinical prognosis using HR MAS MR spectroscopy. *NMR Biomed.* 2010; 23(4): 424–431. [PubMed: 20101607]
183. Li M, Song Y, Cho N, et al. An HR-MAS MR metabolomics study on breast tissues obtained with core needle biopsy. *PLoS ONE.* 2011; 6(10):e25563. [PubMed: 22028780]
184. Portais JC, Delort AM. Carbohydrate cycling in micro-organisms: what can $(13)\text{C}$ -NMR tell us? *FEMS Microbiol. Rev.* 2002; 26(4):375–402. [PubMed: 12413666]
185. Zamboni N, Sauer U. Novel biological insights through metabolomics and ^{13}C -flux analysis. *Curr. Opin. Microbiol.* 2009; 12(5):553–558. [PubMed: 19744879]
186. Desilva MA, Shanaiah N, Nagana Gowda GA, Rosa-Perez K, Hanson BA, Raftery D. Application of ^{31}P NMR spectroscopy and chemical derivatization for metabolite profiling of lipophilic compounds in human serum. *Magn. Reson. Chem.* 2009; 47(Suppl. 1):S74–S80. [PubMed: 19610016]
187. Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum.* 1958; 29(8):4.
188. Kaiser KA, Barding GA Jr, Larive CK. A comparison of metabolite extraction strategies for ^1H -NMR-based metabolic profiling using mature leaf tissue from the model plant *Arabidopsis thaliana*. *Magn. Reson. Chem.* 2009; 47(Suppl. 1):S147–S156. [PubMed: 19551810]
189. Martineau E, Tea I, Loaec G, Giraudeau P, Akoka S. Strategy for choosing extraction procedures for NMR-based metabolomic analysis of mammalian cells. *Anal. Bioanal. Chem.* 2011; 401(7): 2133–2142. [PubMed: 21837464]
190. William, S. Water signal suppression in NMR spectroscopy. In: Webb, GA., editor. *Annual Reports on NMR Spectroscopy*. Boston, MA, USA: Academic Press; 1999. p. 289–354.
191. Beckonert O, Coen M, Keun HC, et al. High-resolution magic-angle-spinning NMR spectroscopy for metabolic profiling of intact tissues. *Nat. Protoc.* 2010; 5(6):1019–1032. [PubMed: 20539278]
192. Cloarec O, Dumas M-E, Craig A, et al. Statistical total correlation spectroscopy. an exploratory approach for latent biomarker identification from metabolic ^1H NMR data sets. *Anal. Chem.* 2005; 77(5):1282–1289. [PubMed: 15732908]
193. Crockford DJ, Holmes E, Lindon JC, et al. Statistical heterospectroscopy, an approach to the integrated analysis of NMR and UPLC-MS data sets. application in metabonomic toxicology studies. *Anal. Chem.* 2005; 78(2):363–371. [PubMed: 16408915]
194. Xi Y, De Ropp J, Viant M, Woodruff D, Yu P. Automated screening for metabolites in complex mixtures using 2D COSY NMR spectroscopy. *Metabolomics.* 2006; 2(4):221–233.
195. Ludwig C, Ward DG, Martin A, et al. Fast targeted multidimensional NMR metabolomics of colorectal cancer. *Magn. Reson. Chem.* 2009; 47(Suppl. 1):S68–S73. [PubMed: 19790200]

196. Ludwig C, Viant MR. Two-dimensional J-resolved NMR spectroscopy: review of a key methodology in the metabolomics toolbox. *Phytochem. Anal.* 2010; 21(1):22–32. [PubMed: 19904730]
197. Aue WP, Karhan J, Ernst RR. Homonuclear broad band decoupling and two-dimensional J-resolved NMR spectroscopy. *J. Chem. Phys.* 1976; 64(10):4226–4227.
198. Fonville JM, Maher AD, Coen M, Holmes E, Lindon JC, Nicholson JK. Evaluation of full-resolution J-resolved ^1H NMR projections of biofluids for metabolomics information retrieval and biomarker identification. *Anal. Chem.* 2010; 82(5):1811–1821. [PubMed: 20131799]
199. Palmer AI, Cavanagh J, Wright P, Rance M. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson.* 1991; 93(1):151–170.
200. Peyrauda R, Kiefera P, Christena P, Massoub S, Portaisb J, Vorholta J. Demonstration of the ethylmalonyl-CoA pathway by using ^{13}C metabolomics. *Proc. Natl Acad. Sci. USA.* 2009; 106(12):4846–4851. [PubMed: 19261854]
201. Jans AW, Willem R. ^{13}C -NMR study of glycerol metabolism in rabbit renal cells of proximal convoluted tubules. *Eur. J. Biochem.* 1988; 174(1):67–73. [PubMed: 3371365]
202. Sadykov MR, Zhang B, Halouska S, et al. Using NMR metabolomics to investigate tricarboxylic acid cycle dependent signal transduction in *Staphylococcus epidermidis*. *J. Biol. Chem.* 2010; 285(47):36616–36624. [PubMed: 20861019]
203. Shanaiah N, Desilva MA, Nagana Gowda GA, Raftery MA, Hainline BE, Raftery D. Class selection of amino acid metabolites in body fluids using chemical derivatization and their enhanced ^{13}C NMR. *Proc. Natl Acad. Sci. USA.* 2007; 104(28):11540–11544. [PubMed: 17606902]
204. Hu K, Westler WM, Markley JL. Simultaneous quantification and identification of individual chemicals in metabolite mixtures by two-dimensional extrapolated time-zero ^1H - ^{13}C HSQC (HSQC0). *J. Am. Chem. Soc.* 2011; 133(6):1662–1665. [PubMed: 21247157]
205. Holmes E, Foxall PJD, Nicholson JK, et al. Automatic data reduction and pattern recognition methods for analysis of ^1H nuclear magnetic resonance spectra of human urine from normal and pathological states. *Anal. Biochem.* 1994; 220(2):284–296. [PubMed: 7978270]
206. Ross A, Schlotterbeck G, Klaus W, Senn H. Automation of NMR measurements and data evaluation for systematically screening interactions of small molecules with target proteins. *J. Biomol. NMR.* 2000; 16(2):139–146. [PubMed: 10723993]
207. Anderson PE, Reo NV, Delraso NJ, Doom TE, Raymer ML. Gaussian binning: a new kernel-based method for processing NMR spectroscopic data for metabolomics. *Metabolomics.* 2008; 4(3):261–272.
208. Davis RA, Charlton AJ, Godward J, Jones SA, Harrison M, Wilson JC. Adaptive binning: An improved binning method for metabolomics data using the undecimated wavelet transform. *Chemom. Intell. Lab. Syst.* 2007; 85(1):144–154.
209. De Meyer T, Sinnaeve D, Van Gasse B, et al. NMR-based characterization of metabolic alterations in hypertension using an adaptive, intelligent binning algorithm. *Anal. Chem.* 2008; 80(10):3783–3790. [PubMed: 18419139]
210. Anderson PE, Mahle DA, Doom TE, Reo NV, Del RNJ, Raymer ML. Dynamic adaptive binning: an improved quantification technique for NMR spectroscopic data. *Metabolomics.* 2011; 7(2): 179–190.
211. Halouska S, Powers R. Negative impact of noise on the principal component analysis of NMR data. *J. Magn. Reson.* 2006; 178(1):88–95. [PubMed: 16198132]
212. Wu W, Daszykowski M, Walczak B, et al. Peak alignment of urine NMR spectra using fuzzy warping. *J. Chem. Inf. Model.* 2006; 46(2):863–875. [PubMed: 16563018]
213. Forshed J, Schuppe-Koistinen I, Jacobsson SP. Peak alignment of NMR signals by means of a genetic algorithm. *Anal. Chim. Acta.* 2003; 487(2):189–199.
214. Veselkov KA, Lindon JC, Ebbels TMD, et al. Recursive Segment-wise peak alignment of biological ^1H NMR spectra for improved metabolic biomarker recovery. *Anal. Chem.* 2009; 81(1):56–66. [PubMed: 19049366]

215. Csenki L, Alm E, Torgrip RJO, et al. Proof of principle of a generalized fuzzy Hough transform approach to peak alignment of one-dimensional ^1H NMR data. *Anal. Bioanal. Chem.* 2007; 389(3):875–885. [PubMed: 17701402]
216. Craig A, Cloarec O, Holmes E, Nicholson JK, Lindon JC. Scaling and normalization effects in NMR spectroscopic metabonomic data sets. *Anal. Chem.* 2006; 78(7):2262–2267. [PubMed: 16579606]
217. De Meyer T, Sinnavee D, Van Gasse B, et al. Evaluation of standard and advanced preprocessing methods for the univariate analysis of blood serum ^1H -NMR spectra. *Anal. Bioanal. Chem.* 2010; 398(4):1781–1790. [PubMed: 20714889]
218. Oliver, F. Cellular metabolimics: the quest for pathway structure. In: Lindon, JC.; Nicholson, JK.; Holmes, E., editors. *The Handbook of Metabonomics and Metabolomics*. The Netherlands: Elsevier Science, Amsterdam; 2007. p. 20
219. Lindon, JC.; Nicholson, JK.; Holmes, E. *The Handbook of Metabonomics and Metabolomics* (1st Edition). The Netherlands: Elsevier, Amsterdam; 2007.
220. Basant A, Rege M, Sharma S, Sonawat HM. Alterations in urine, serum and brain metabolomic profiles exhibit sexual dimorphism during malaria disease progression. *Malaria J.* 2010; 9(1):110.
221. Westerhuis JA, Hoefsloot HCJ, Smit S, et al. Assessment of PLSDA cross validation. *Metabolomics.* 2008; 4(1):81–89.
222. Rubingh CM, Bijlsma S, Derks EPPA, et al. Assessing the performance of statistical validation tools for megavariate metabolomics data. *Metabolomics.* 2006; 2(2):53–61.
223. Keun HC, Ebbels TM, Antti H, et al. Analytical reproducibility in (1)H NMR-based metabonomic urinalysis. *Chem. Res. Toxicol.* 2002; 15(11):1380–1386. [PubMed: 12437328]
224. Wang T, Shao K, Chu Q, et al. Automics. An integrated platform for NMR-based metabonomics spectral processing and data analysis. *BMC Bioinform.* 2009; 10:83.
225. Zhao Q, Stoyanova R, Du S, Sajda P, Brown TR. HiRes – a tool for comprehensive assessment and interpretation of metabolomic data. *Bioinformatics.* 2006; 22(20):2562–2564. [PubMed: 16895927]
226. Xia J, Psychogios N, Young N, Wishart DS. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* 2009; 37:W652–W660. [PubMed: 19429898]
227. Izquierdo-Garcia J, Rodriguez I, Kyriazis A, et al. A novel R-package graphic user interface for the analysis of metabonomic profiles. *BMC Bioinf.* 2009; 10(1):363.
228. Werth MT, Halouska S, Shortridge MD, Zhang B, Powers R. Analysis of metabolomic PCA data using tree diagrams. *Anal. Biochem.* 2010; 399(1):58–63. [PubMed: 20026297]
229. Efron B, Halloran E, Holmes S. Bootstrap confidence levels for phylogenetic trees. *Proc. Natl Acad. Sci. USA.* 1996; 93(14):13429–13434. [PubMed: 8917608]
230. Retief JD. Phylogenetic analysis using PHYLIP. *Methods Mol. Biol.* 1999; 132(2):16.
231. Shulaev V. Metabolomics technology and bioinformatics. *Briefings Bioinf.* 2006; 7(2):128–139.
232. Oberhardt MA, Palsson BO, Papin JA. Applications of genome-scale metabolic reconstructions. *Mol. Syst. Biol.* 2009; 5:320. [PubMed: 19888215]
233. Robinette SL, Ajredini R, Rasheed H, et al. Hierarchical alignment and full resolution pattern recognition of 2D NMR spectra. Application to nematode chemical ecology. *Anal. Chem.* 2011; 83(5):1649–1657. [PubMed: 21314130]
234. Vu TN, Valkenburg D, Smets K, et al. An integrated workflow for robust alignment and simplified quantitative analysis of NMR spectrometry data. *BMC Bioinf.* 2011; 12:405.
235. Mo H, Raftery D. Solvent signal as an NMR concentration reference. *Anal. Chem.* 2008; 80(24):9835–9839. [PubMed: 19007190]
236. Craig A, Cloarec O, Holmes E, Nicholson JK, Lindon JC. Scaling and normalization effects in NMR spectroscopic metabonomic data sets. *Anal. Chem.* 2006; 78(7):2262–2267. [PubMed: 16579606]
237. Xia J, Bjorn Dahl TC, Tang P, Wishart DS. MetaboMiner – semi-automated identification of metabolites from 2D NMR spectra of complex biofluids. *BMC Bioinf.* 2008; 9:507.
238. Cui Q, Lewis IA, Hegeman AD, et al. Metabolite identification via the Madison Metabolomics Consortium Database. *Nat. Biotechnol.* 2008; 26(2):162–164. [PubMed: 18259166]

239. Ulrich EL, Akutsu H, Doreleijers JF, et al. BioMagResBank. Nucleic Acids Res. 2008; 36:D402–D408. [PubMed: 17984079]
240. Wishart DS, Knox C, Guo AC, et al. HMDB. a knowledge-base for the human metabolome. Nucleic Acids Res. 2009; 37:D603–D610. [PubMed: 18953024]
241. Kanehisa M, Araki M, Goto S, et al. KEGG for linking genomes to life and the environment. Nucleic Acids Res. 2008; 36:D480–D484. [PubMed: 18077471]
242. Caspi R, Foerster H, Fulcher CA, et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res. 2008; 36:D623–D631. [PubMed: 17965431]
243. Cline MS, Smoot M, Cerami E, et al. Integration of biological networks and gene expression data using Cytoscape. Nat. Protoc. 2007; 2(10):2366–2382. [PubMed: 17947979]
244. Viswanathan GA, Seto J, Patil S, Nudelman G, Sealfon SC. Getting started in biological pathway construction and analysis. PLoS Comput. Biol. 2008; 4(2):e16. [PubMed: 18463709]
245. Funahashi A, Jouraku A, Matsuoka Y, Kitano H. Integration of Cell Designer and SABIO-RK. In Silico Biol. 2007; 7(Suppl. 2):S81–S90. [PubMed: 17822394]
246. Workentine ML, Harrison JJ, Weljie AM, et al. Phenotypic and metabolic profiling of colony morphology variants evolved from *Pseudomonas fluorescens* biofilms. Environ. Microbiol. 2010; 12(6):1565–1577. [PubMed: 20236162]
247. Hall-Stoodley L, Hu FZ, Gieseke A, et al. Direct detection of bacterial biofilms on the middle-ear mucosa of children with chronic otitis media. J. Am. Med. Assoc. 2006; 296(2):202–211.
248. Kim YH, An ES, Song BK, Kim DS, Chelikani R. Polymerization of cardanol using soybean peroxidase and its potential application as anti-biofilm coating material. Biotechnol. Lett. 2003; 25(18):1521–1524. [PubMed: 14571976]
249. Carlson RP, Taffs R, Davison WM, Stewart PS. Anti-biofilm properties of chitosan-coated surfaces. J. Biomater. Sci. PolymEd. 2008; 19(8):1035–1046.
250. Khoo X, Grinstaff MW. Novel infection-resistant surface coatings: a bioengineering approach. MRS Bull. 2011; 36(5):357–366.
251. Shunmugaperumal T. Microbial colonization of medical devices and novel preventive strategies. Recent Pat. Drug Deliv. Formul. 2010; 4(2):153–173. [PubMed: 20236065]
252. Vogt M, Flemming HC, Veeman WS. Diffusion in *Pseudomonas aeruginosa* biofilms: a pulsed field gradient NMR study. J. Biotechnol. 2000; 77(1):137–146. [PubMed: 10674220]
253. Hoyle BD, Alcantara J, Costerton JW. *Pseudomonas aeruginosa* biofilm as a diffusion barrier to piperacillin. Antimicrob. Agents Chemother. 1992; 36(9):2054–2056. [PubMed: 1416900]
254. Rogers SA, Huigens RW 3rd, Cavanagh J, Melander C. Synergistic effects between conventional antibiotics and 2-aminoimidazole-derived antibiofilm agents. Antimicrob. Agents Chemother. 2010; 54(5):2112–2118. [PubMed: 20211901]
255. Walencka E, Rozalska S, Wysokinska H, Rozalski M, Kuzma L, Rozalska B. Salvipisone and aethiopinone from *Salvia sclarea* hairy roots modulate staphylococcal antibiotic resistance and express anti-biofilm activity. Planta Med. 2007; 73(6):545–551. [PubMed: 17650545]
256. Sadykov MR, Olson ME, Halouska S, et al. Tricarboxylic acid cycle dependent regulation of *Staphylococcus epidermidis* polysaccharide intercellular adhesin synthesis. J. Bacteriol. 2008; 130(23):7621–7632. [PubMed: 18820013]
257. White AP, Weljie AM, Apel D, et al. A global metabolic shift is linked to *Salmonella* multicellular development. PLoS ONE. 2010; 5(7):e11814. [PubMed: 20676398]
258. Theodore TS, Schade AL. Carbohydrate metabolism of iron-rich and iron-poor *Staphylococcus aureus*. J. Gen. Microbiol. 1965; 40(3):385–395. [PubMed: 5864890]
259. Musk DJ Jr, Hergenrother PJ. Chelated iron sources are inhibitors of *Pseudomonas aeruginosa* biofilms and distribute efficiently in an *in vitro* model of drug delivery to the human lung. J. Appl. Microbiol. 2008; 105(2):380–388. [PubMed: 18284482]
260. Moreau-Marquis S, O'Toole GA, Stanton BA. Tobramycin and FDA-approved iron chelators eliminate *Pseudomonas aeruginosa* biofilms on cystic fibrosis cells. Am. J. Respir. Cell Mol. Biol. 2009; 41(3):305–313. [PubMed: 19168700]

261. Liebeke M, Dorries K, Zuhlke D, et al. A metabolomics and proteomics study of the adaptation of *Staphylococcus aureus* to glucose starvation. *Mol. Biosyst.* 2011; 7(4):1241–1253. [PubMed: 21327190]
262. Persidis A. High-throughput screening. *Nat. Biotechnol.* 1998; 16(5):488–489. [PubMed: 9592401]
263. Sittampalam GS, Kahl SD, Janzen WP. High-throughput screening: advances in assay technologies. *Curr. Opin. Cell Biol.* 1997; 1(3):384–391.
264. Bordi C, De Bentzmann S. Hacking into bacterial biofilms: a new therapeutic challenge. *Ann. Intensive Care.* 2011; 1(1):1–8. [PubMed: 21906322]
265. Roy V, Smith JA, Wang J, Stewart JE, Bentley WE, Sintim HO. Synthetic analogs tailor native AI-2 signaling across bacterial species. *J. Am. Chem. Soc.* 2010; 132(32):11141–11150. [PubMed: 20698680]
266. Rui F, Marques JC, Miller ST, Maycock CD, Xavier KB, Ventura MR. Stereochemical diversity of AI-2 analogs modulates quorum sensing in *Vibrio harveyi* and *Escherichia coli*. *Bioorg. Med. Chem.* 2012; 20(1):249–256. [PubMed: 22137598]
267. Fargue P, Halouska S, Werth M, Xu K, Harris S, Powers R. NMR Metabolic profiling of *Aspergillus nidulans* to Monitor drug and protein activity. *J. Proteome Res.* 2006; 5(8):1916–1923. [PubMed: 16889413]
268. Teng Q, Huang W, Collette TW, Ekman DR, Tan C. A direct cell quenching method for cell-culture based metabolomics. *Metabolomics.* 2009; 5(2):199–208.
269. Fogh RH, Vranken WF, Boucher W, Stevens TJ, Laue ED. A nomenclature and data model to describe NMR experiments. *J. Biomol. NMR.* 2006; 36(3):147–155. [PubMed: 17031528]
270. Pons JL, Malliavin TE, Delsuc MA, Gifa V. A complete package for NMR data set processing. *J. Biomol. NMR.* 1996; 8(4):445–452. [PubMed: 20859778]
271. Van Beek JD. matNMR: a flexible toolbox for processing, analyzing and visualizing magnetic resonance data in Matlab. *J. Magn. Reson.* 2007; 187(1):19–26. [PubMed: 17448713]
272. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRpipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* 1995; 6(3):277–293. [PubMed: 8520220]
273. Lewis IA, Schommer SC, Markley JL. rNMR: open source software for identifying and quantifying metabolites in NMR spectra. *Magn. Reson. Chem.* 2009; 47(Suppl. 1):S123–S126. [PubMed: 19821464]
274. Wilson M, Reynolds G, Kauppinen RA, Arvanitis TN, Peet AC. A constrained least-squares approach to the automated quantitation of *in vivo* (1)H magnetic resonance spectroscopy data. *Magn. Reson. Med.* 2011; 65(1):1–12. [PubMed: 20878762]

Websites

301. Biofilms. The hypertextbook. v4.3. 2001. www.biofilmbook.com
302. Advanced Chemistry Development, Inc. www.acdlabs.com
303. The Collaborative Computing Project for NMR. www-old.ccpn.ac.uk/documentation/analysis
304. Felix NMR, Inc. www.felixnmr.com
305. PERCH Solutions Ltd. www.perchsolutions.com
306. GIFA. NPK program. www.cbs.cnrs.fr/GIFA/
307. Nucleomatica, Inc. www.inmr.net
308. matNMR. <http://matnmr.sourceforge.net/>
309. Mestrelab research. <http://mestrelab.com>
310. NMR Science. www.nmrscience.com
311. One Moon Scientific, Inc. www.onemoonscientific.com
312. Acorn NMR Inc. www.acornnmr.com
313. Grandinetti Laboratory. www.grandinetti.org/Software/RMN
314. rNMR Downloads and Installation. <http://rnmr.nmrfam.wisc.edu/pages/download/download.html>

315. Sparky. NMR assignment and integration software. www.cgl.ucsf.edu/home/sparky
316. Spinworks Software. www.columbia.edu/cu/chemistry/groups/nmr/SpinWorks.html
317. Tarquin Software. <http://tarquin.sourceforge.net/>
318. Bruker Instruments, Inc. www.uni-ulm.de/orgchem/AK_Siehl/www_AK_Siehl_old/Bruker_WWW/index.htm

Executive summary

- Bacterial biofilms are a health concern because they increase antibiotic resistance and protect the bacteria from the host immune response system, which may lead to a serious and untreatable infection.
- Bacterial biofilms can attach to a variety of surfaces that include numerous medical devices, which is a primary source of biofilm infections.
- A biofilm is a highly organized microbial structure that contains a slim-like matrix that embeds the microorganism, but allows for the exchange of fluids, nutrients and chemical signals for cellular communication.
- The three main stages of a biofilm life cycle are attachment, growth and propagation. Biofilm formation depends on the bacteria's response to changing environmental conditions, which is regulated by quorum sensing, two-component systems or σ factors.
- Quorum sensing enables bacteria to 'sense' cell density and coordinate behavior in response to nutrient availability and environmental conditions through signal molecules. These signal molecules are prime targets for developing drugs that inhibit biofilm formation.
- Metabolomics is the global analysis of small-molecule concentration changes within a cell using NMR spectroscopy or MS. It provides a direct means of monitoring changes in the state of the cell resulting from activities such as drug treatments or biofilm formation.
- Metabolomics provides a better understanding of a disease since it links the pathology to actual changes in the activity of biological processes. Any observed change in the metabolome is a direct consequence of a change in protein activity.
- The process of monitoring the metabolome includes rapidly quenching all cell activities, efficiently lysing the cell, and quickly extracting the metabolites. Speed and consistency are essential to a successful metabolic study. The goal is to avoid perturbing the state of the metabolome during the collection process.
- 1D ^1H NMR, 2D ^1H - ^{13}C HSQC and 2D ^1H - ^1H TOCSY are the NMR experiments commonly used for metabolomics. The goal of the 1D ^1H NMR experiments is to generate a 'fingerprint' of the state of the cellular metabolome. The primary goal of the 2D NMR experiments is metabolite identification.
- MS is also used for metabolomics, where NMR and MS are complementary techniques.
- There are many considerations for the proper processing of NMR data for a metabolomics study. These include binning, peak alignments, baseline corrections and normalization.
- Chemometrics is commonly used to analyze NMR metabolomics data with the principal goal of class discrimination. The metabolomics data are interpreted by a simple cluster analysis, NMR spectra that yield distinct clusters in a scores plot have different metabolomes.

- A primary challenge with metabolomics is validating that the clustering pattern in a scores plot is biologically relevant as opposed to experimentally induced changes to the metabolome or the NMR data.
- A number of examples illustrating the application of NMR-based metabolomics to the investigation of bacterial biofilms are presented.

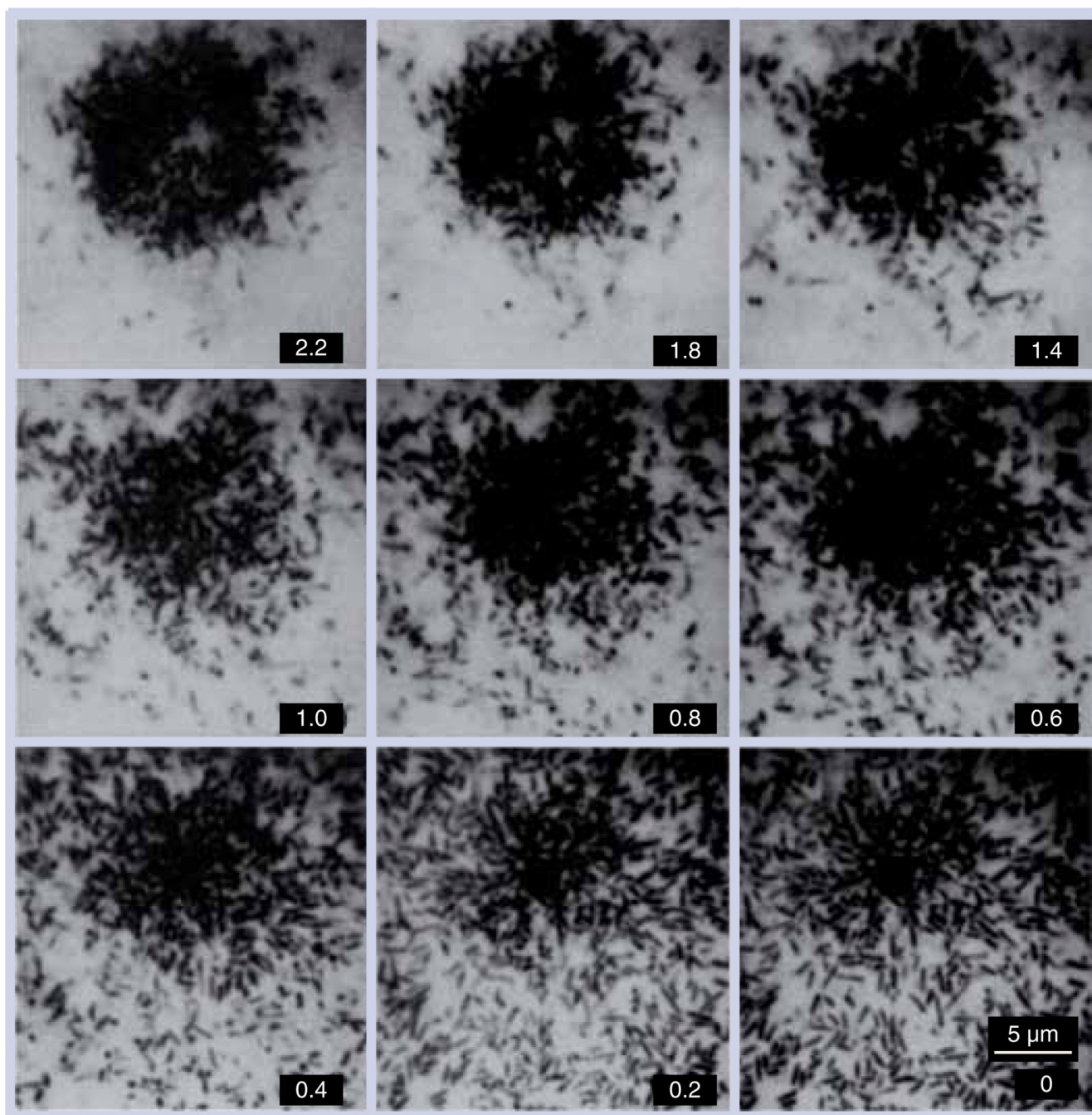


Figure 1. Horizontal optical thin sections (0–2.6 μm) of the *Pseudomonas aeruginosa* biofilm obtained by scanning confocal laser microscopy

The biofilm was negatively stained with 0.1% fluorescein. The horizontal sections show the removal of out-of-focus information and reveal aspects of the internal structure of the biofilm.

Reprinted with permission from [45] © American Society for Microbiology (1991).

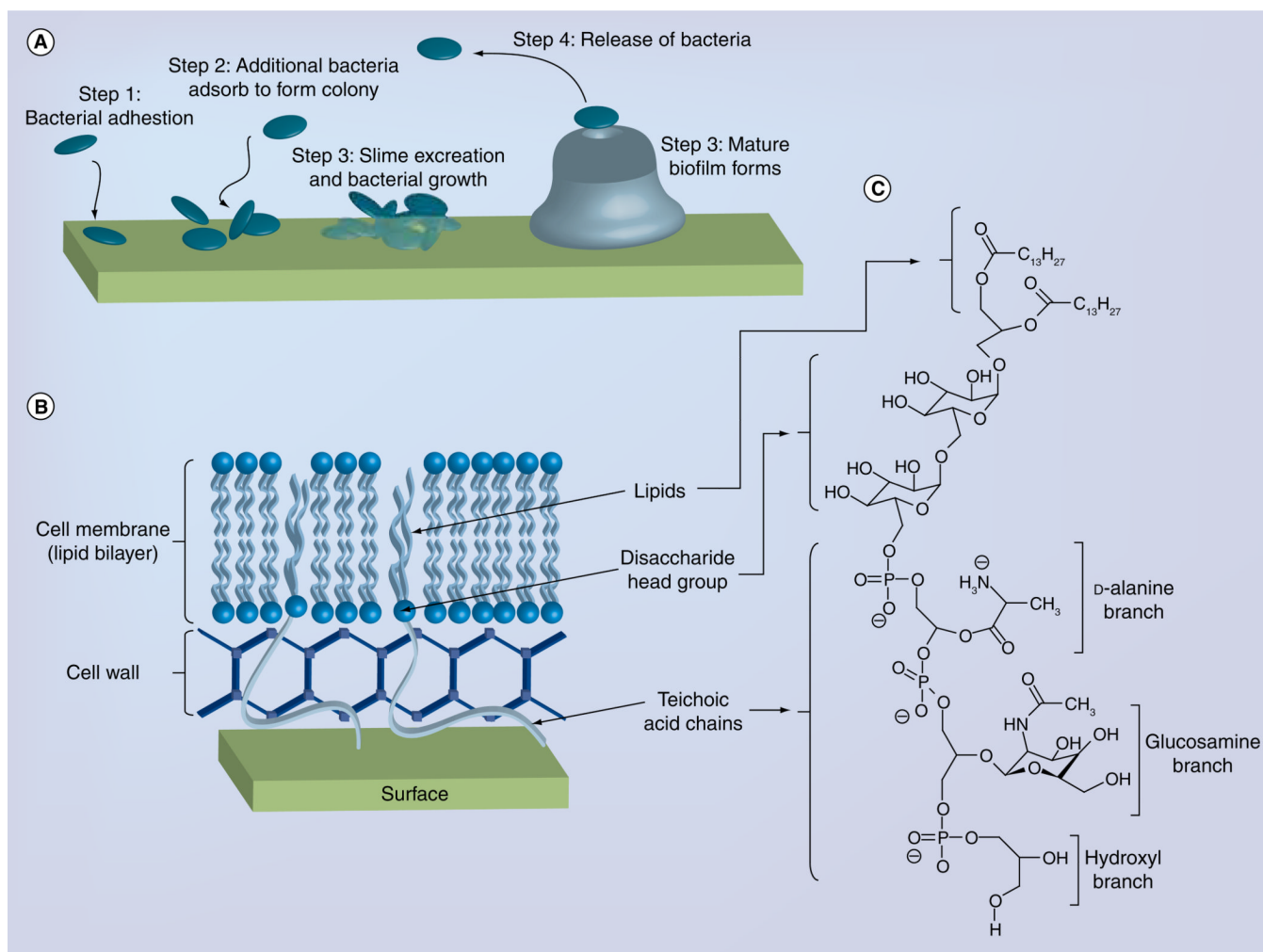


Figure 2. Bacterial biofilm formation and growth

(A) The four-step process involved in the formation of biofilms. **(B)** The lipoteichoic acid is anchored to the cell membrane and extends out past the edge of the cell wall; it is the initial molecular contact between the bacterium and other materials. **(C)** The chemical structure of lipoteichoic acid.

Reprinted with permission from [46] © Elsevier (2008).

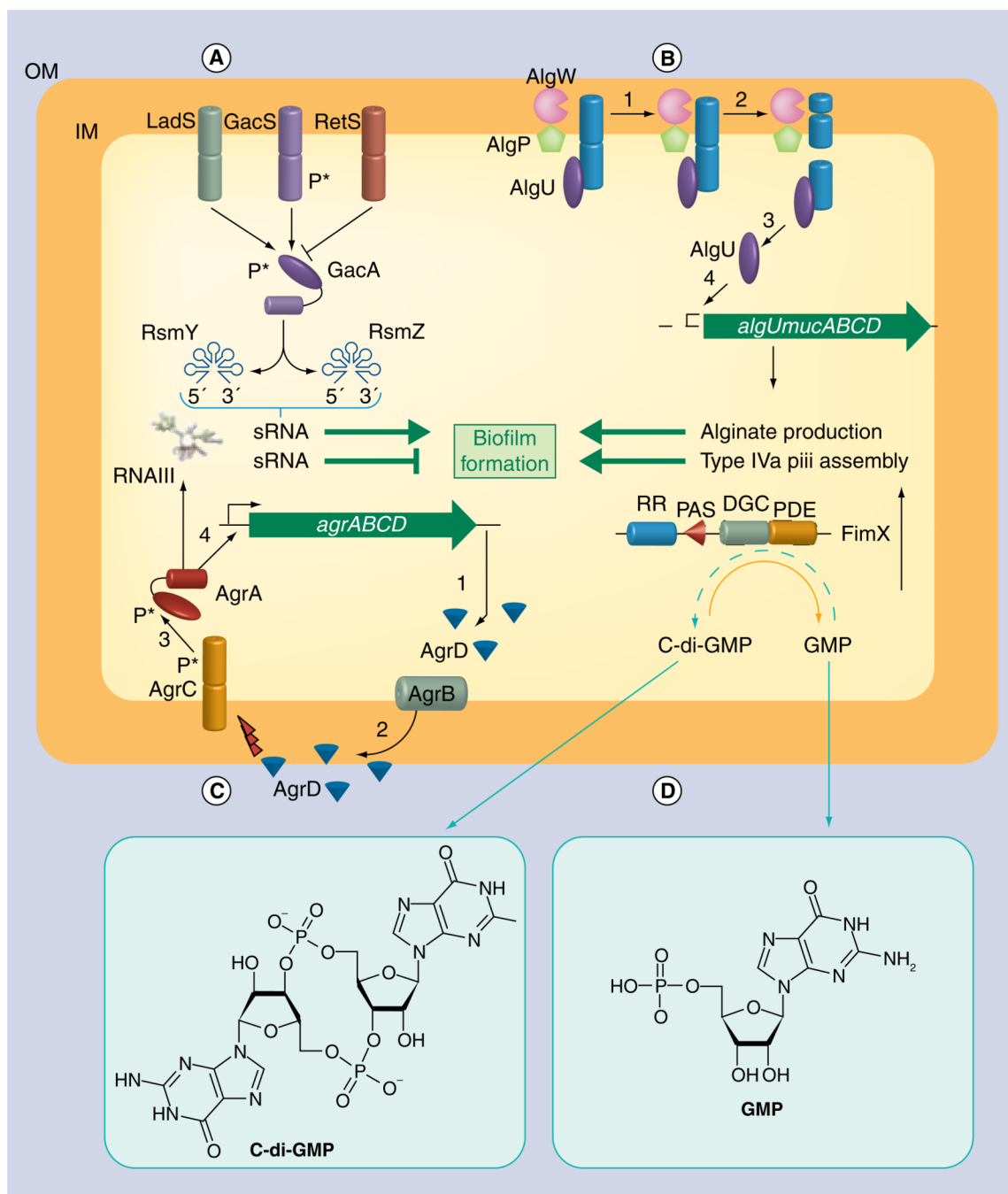


Figure 3. Regulatory networks controlling the transition between planktonic and biofilm lifestyle
The external frames illustrate the bacterial envelope with one or two membranes according to (C) Gram-positive and (A, B & D) Gram-negative bacteria, respectively. (A) Control of biofilm formation in *Pseudomonas aeruginosa* through the two-component system GacS (HK)/GacA (RR) mediated by sRNA *rsmY* and *rsmZ* gene transcription and modulated by RetS and LadS, two additional HK in *P. aeruginosa*. (B) Control of exopolysaccharide alginate in *P. aeruginosa*, which further impacts biofilm architecture by the system ECF σ factor AlgU-anti- σ MucA-AlgP (IM)-AlgW (periplasmic) complex: (1) activation of AlgW/AlgP; (2) cleavage of MucA; (3) release of AlgU; (4) activation of the *alg*

UmuCABCD operon. **(C)** Control of *Staphylococcus aureus* biofilm formation through the Agr QS system: (1) AgrD autoinducer production; (2) AgrD autoinducer accumulation in the extracellular medium where it reaches a threshold; (3) activation of the two-component system AgrCA by AgrD at the threshold concentration; (4) AgrA-dependent activation of the sRNA RNA III expression repressing expression of genes involved in biofilm formation together with amplification loop of *agrABCD*. **(D)** Control of *P. aeruginosa* biofilm formation through the intracellular second messenger c-di-GMP level controlled by the FimX protein having DGC and PDE domains, a RR domain, and a PAS domain. Note that in FimX protein, only PDE activity is detectable (continuous arrow), whereas DGC activity is undetectable (dotted arrow).

DGC: Diguanilatecyclase; IM: Inner membrane; OM: Outer membrane; RR: Response regulator.

Reprinted with permission from [264] © Springer (2011).

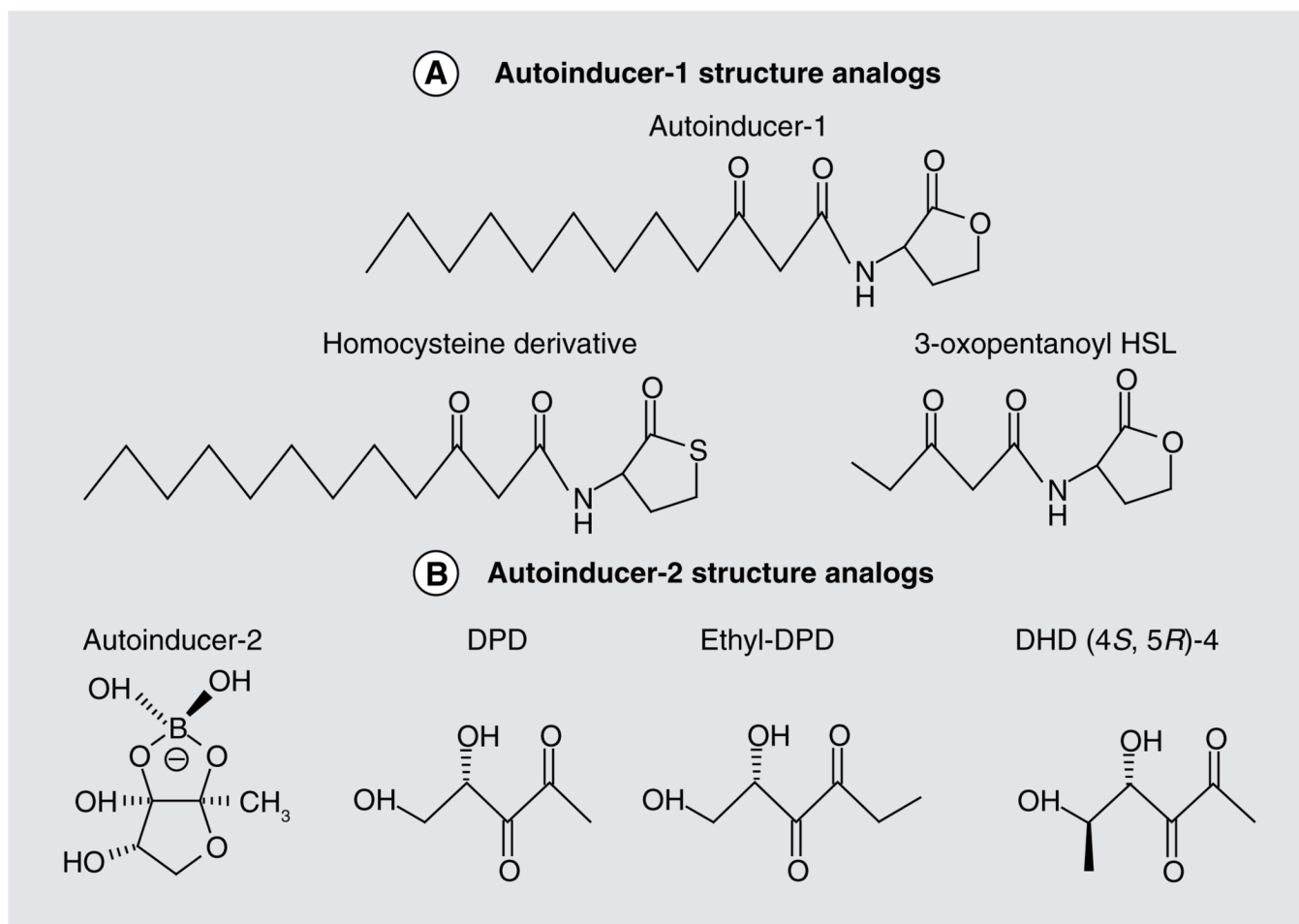


Figure 4. Autoinducers and analogs

(A) Autoinducer-1 chemical structure and its analogs homocysteine derivative [105] and 3-oxopentanoyl HSL [106].

(B) Autoinducer-2 chemical structure, its precursor DPD and its analogs ethyl-DPD [265] and DHD (4*S*, 5*R*)-4 [266].

DPD: 4,5-dihydroxy-2,3-pentanedione.

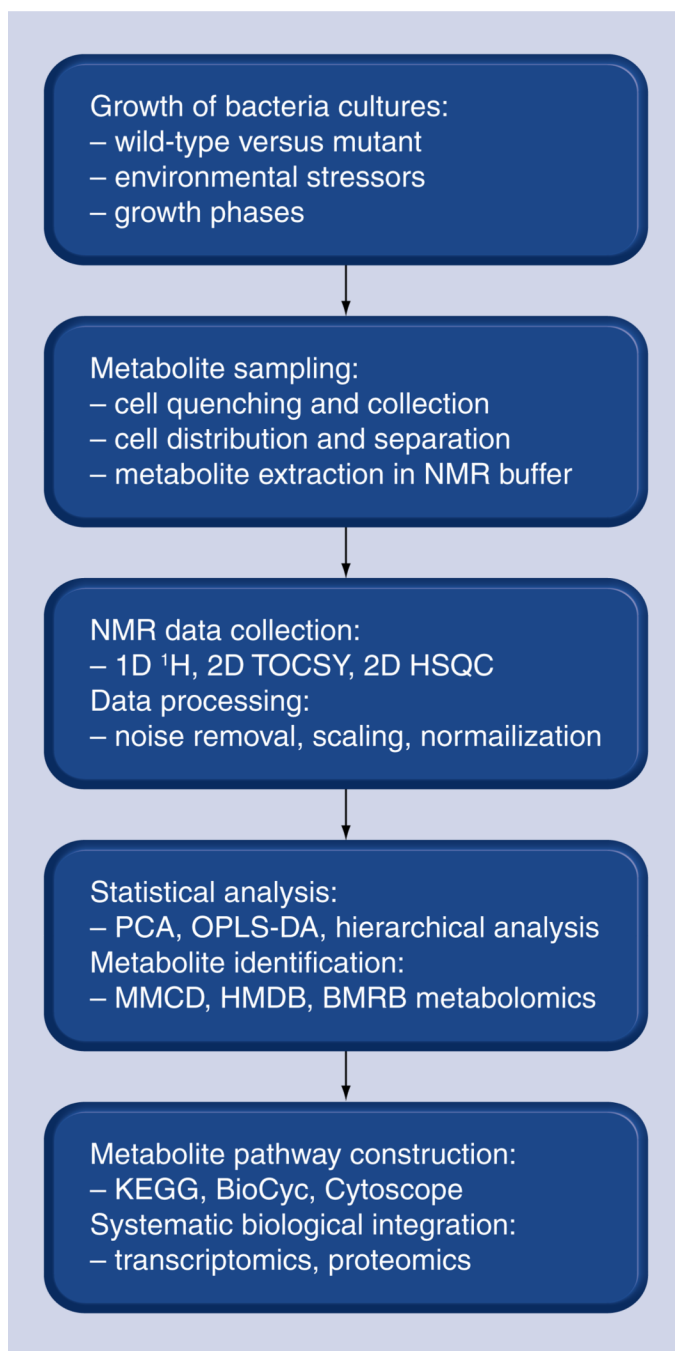


Figure 5. General protocols for an NMR metabolomics study

BMRB: Biological Magnetic Resonance Data Bank; HMDB: Human Metabolome Database; HSQC: Heteronuclear single-quantum correlation; KEGG: Kyoto Encyclopedia of Genes and Genomes; MMCD: Madison Metabolomics Consortium Database; OPLS-DA: Orthogonal partial least squares discriminant analysis; PCA: Principal component analysis; TOCSY: Total correlation spectroscopy.

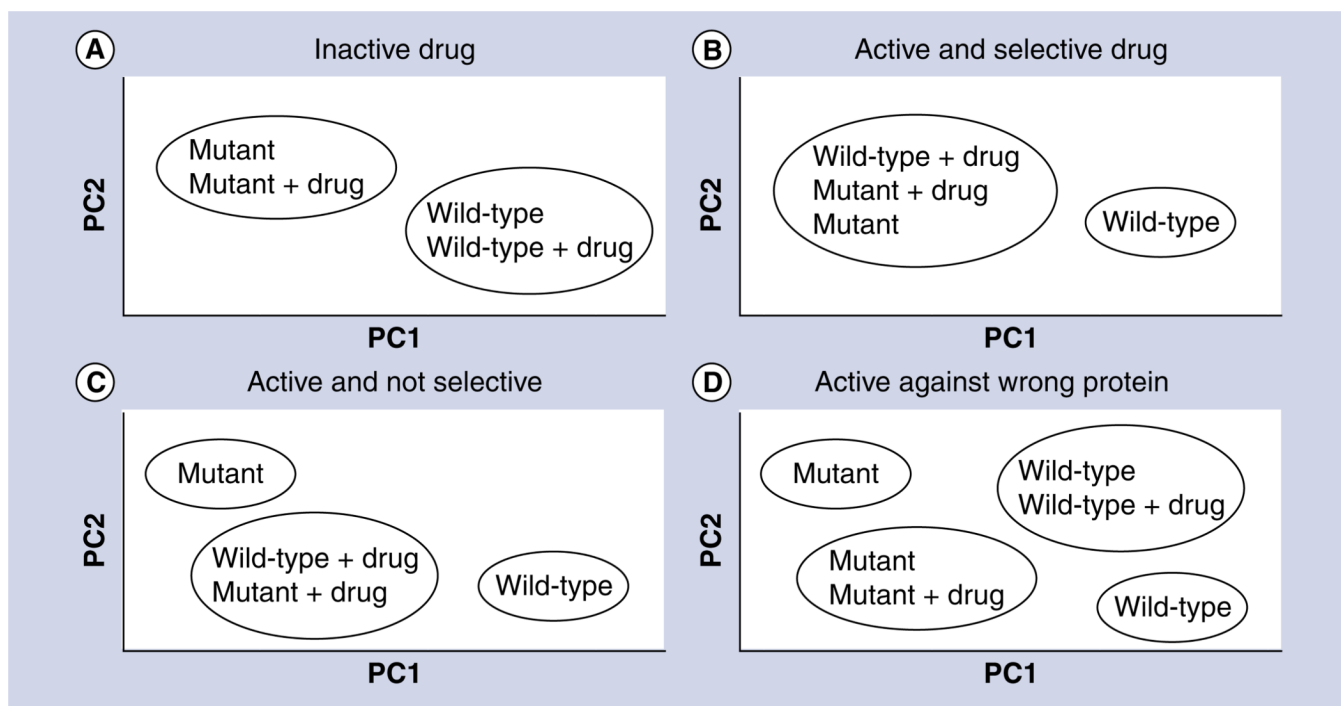


Figure 6. Hypothetical principal component analysis scores plot for several scenarios
(A) Inactive compound; (B) active and selective inhibitor; (C) active, nonselective inhibition of target and secondary protein; and (D) active, nonselective preferential inhibition of secondary protein.

PC: Principal component.

Reprinted with permission from [267] © American Chemical Society (2006).

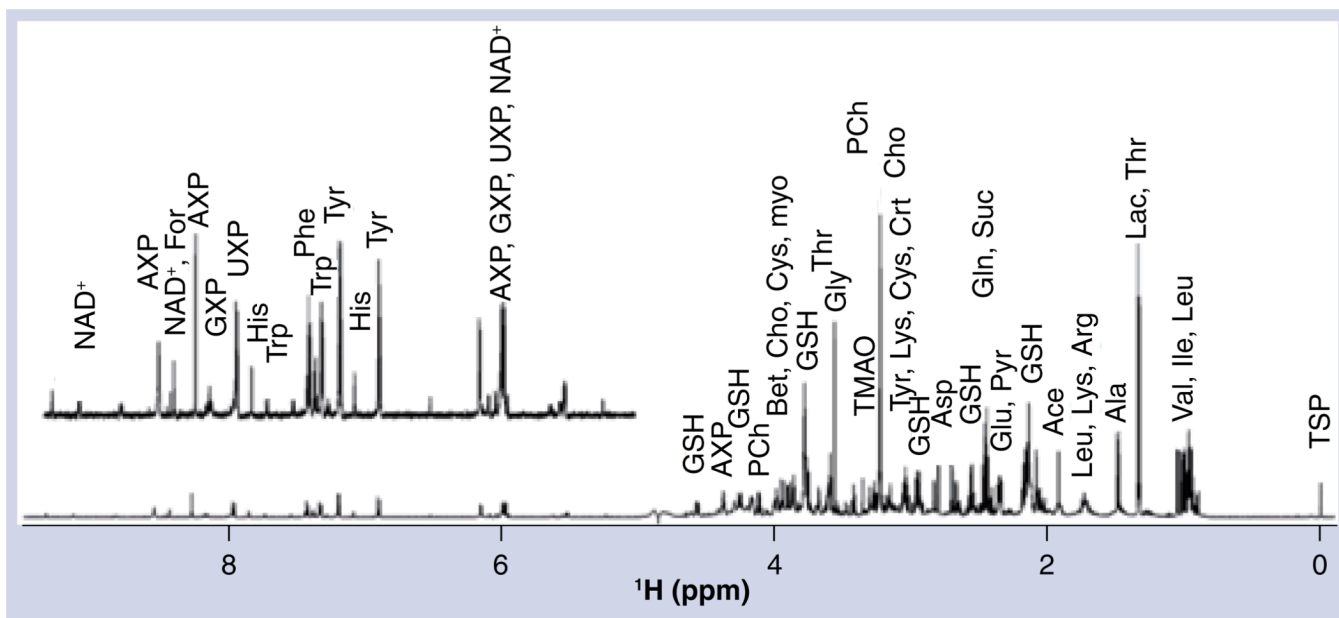


Figure 7. ^1H NMR spectra of the intracellular metabolic profiles of aqueous extract of MCF-7 cells

The sample used for spectrum contained approximately 6×10^5 cells. The spectra shown in the inset were also plotted using the same expanded x-axis scale.

Ace: Acetate; Bet: Betaine; Cho: Choline; For: Formate; GSH: Glutathione; Lac: Lactate; Myo: Myo-inositol; PCh: Phosphocholine; Pyr: Pyruvate; Suc: Succinate; TMAO: Trimethylamine *N*-oxide; TSP: 3-trimethylsilylpropionic acid.

Reprinted with permission from [268] © Springer (2009).



NIH-PA Author Manuscript

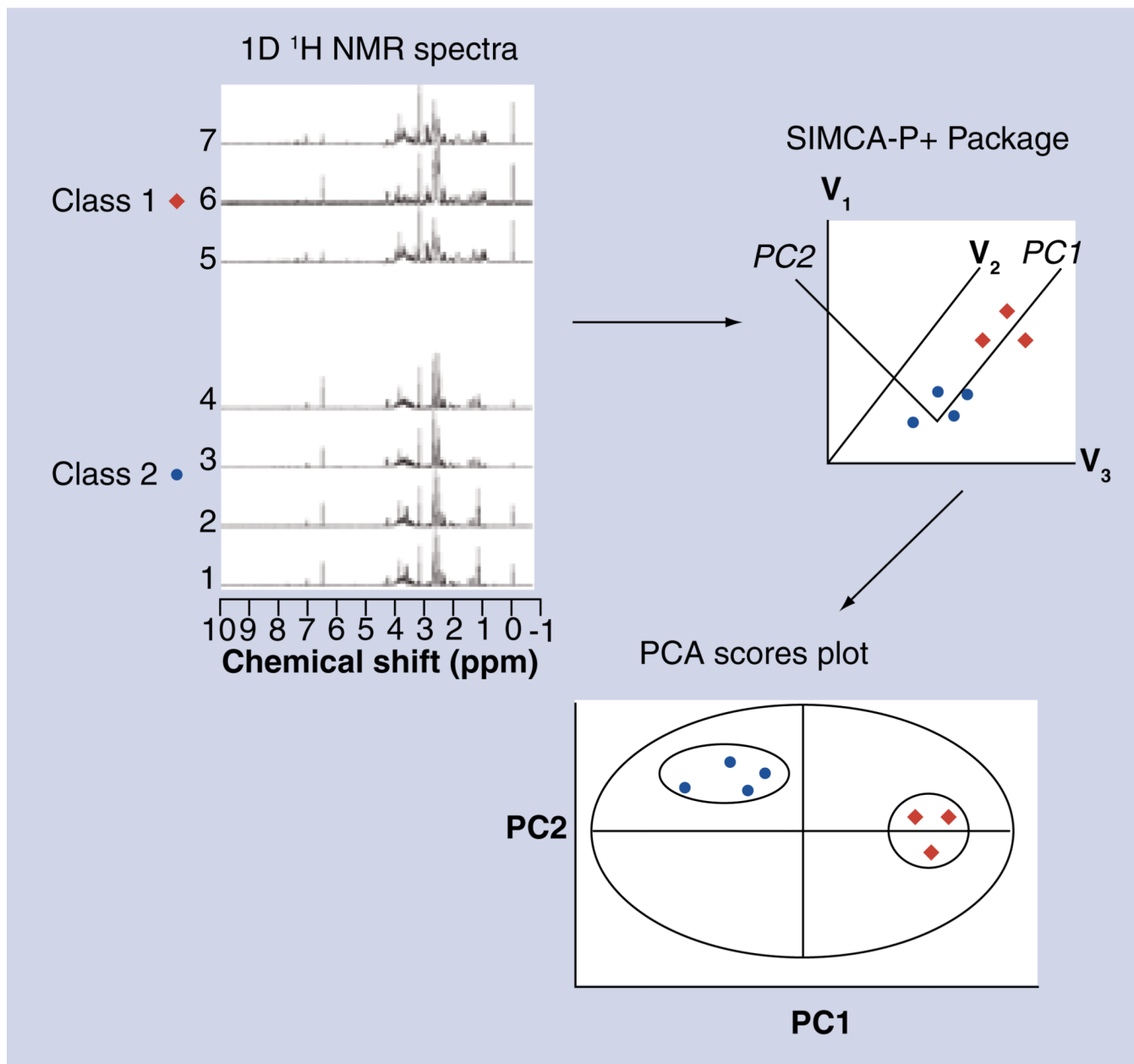


Figure 9. Conceptual illustration of the principal component analysis of NMR-spectral data
 Each 1D ^1H NMR spectrum is converted into a single point in multidimensional space. Each axis (e.g., V_1 , V_2 and V_3) corresponds to a chemical shift (ppm) and the value along the axis is the peak intensity or bin integral. A vector ($PC1$) corresponds to the largest variation in the data. The second vector ($PC2$) is orthogonal to the first vector and corresponds to the second largest variation in the data. The scores correspond to the fit of each point (spectrum) to each PC vector, where the resulting 2D scores plot identifies the relative similarities and differences between the NMR spectra based on the clustering pattern.
 PC: Principal component; PCA: Principal component analysis.

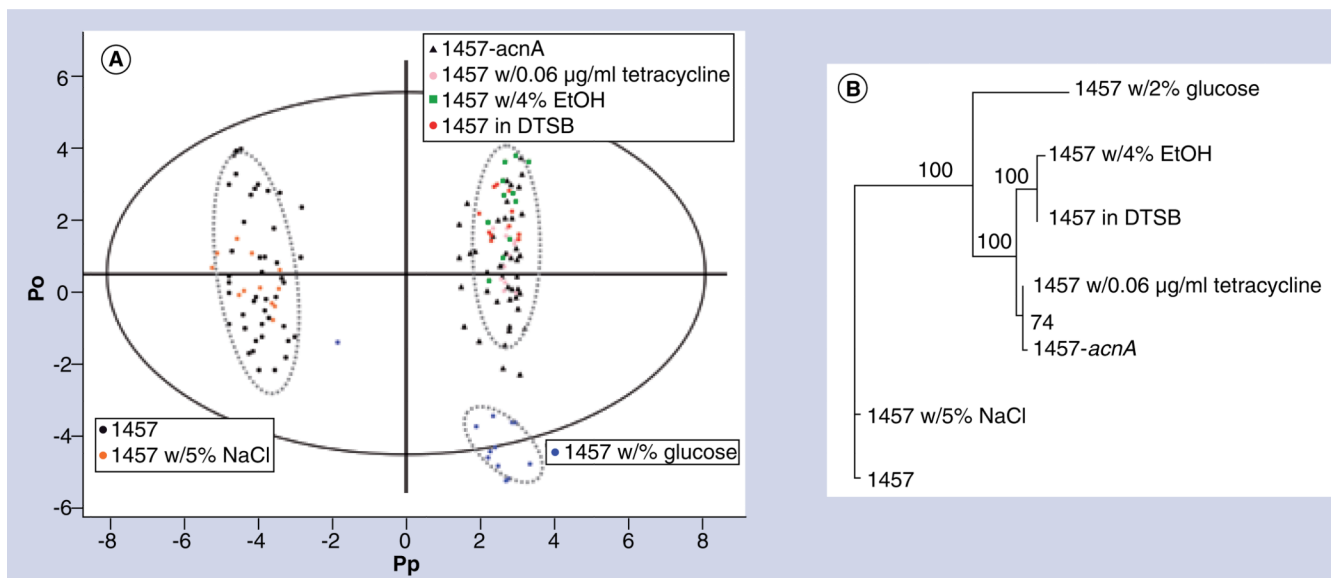


Figure 10. *Staphylococcus epidermidis* biofilm formation, regulated by the tricarboxylic acid cycle

(A) 2D orthogonal partial least squares discriminant analysis (OPLS-DA) comparing wild-type *Staphylococcus epidermidis* 1457 cells grown 6 h in standard tryptic soy broth media, with *S. epidermidis* 1457 cells grown 6 h in iron-depleted media (DTSB), with the addition of 4% ethanol, with the addition of 2% glucose, with the addition of 0.06 µg/ml tetracycline, with the addition of 5% NaCl, and 6 h growth of aconitase mutant strain 1457-*acnA::tetM* in standard tryptic soy broth media. 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 was designated the control class and the remainder of the cells were designated as treated. The OPLS-DA used one predictive component and four orthogonal components to yield a R^2X of 0.637, R^2Y of 0.966 and Q^2 of 0.941. (B) Metabolomic tree diagram generated from the the 2D OPLS-DA scores plot depicted in (A). Each node is labeled with the bootstrap number, where a value above 50 indicates a statistically significant separation.

Reprinted with permission from [124] © American Chemical Society (2011).

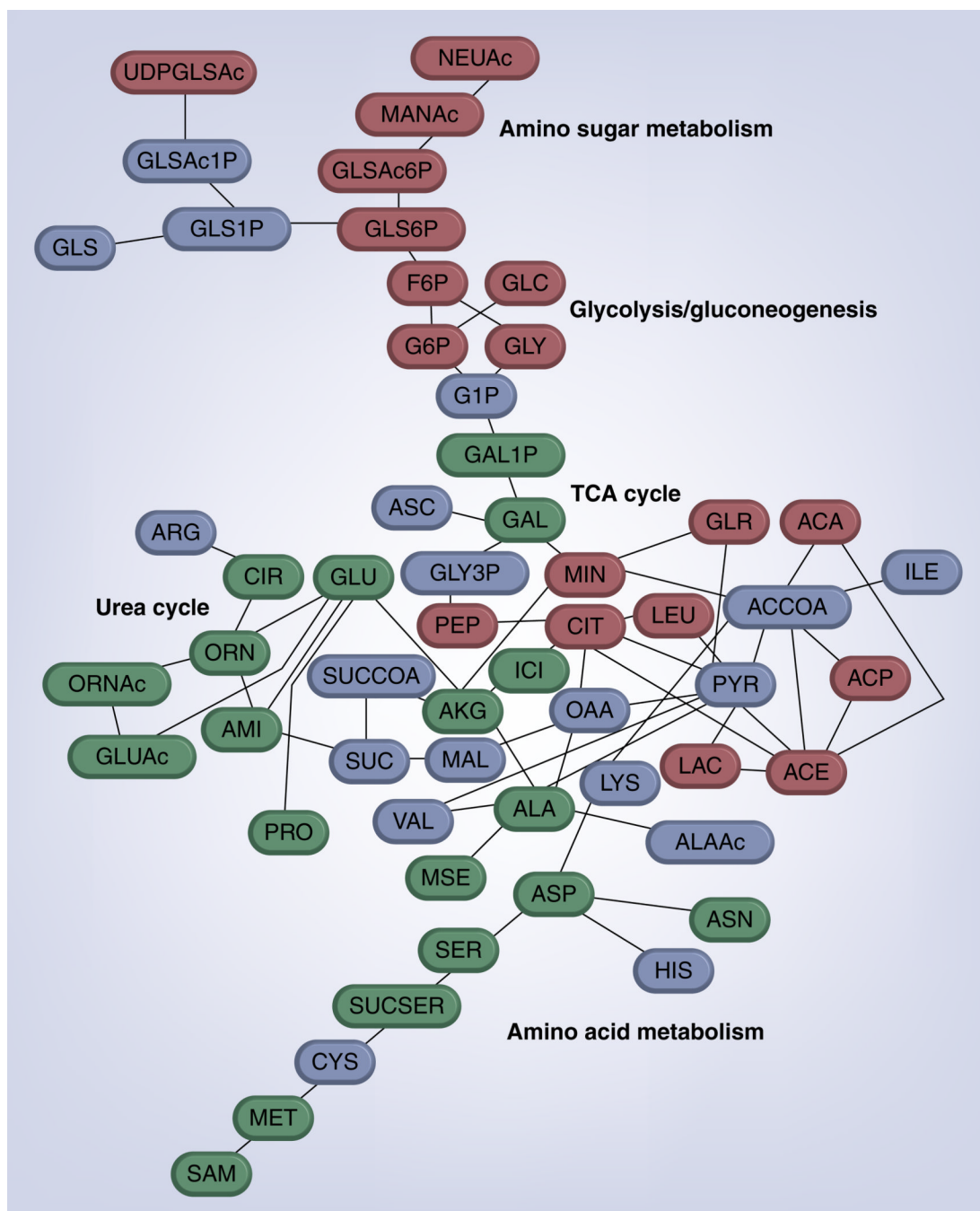


Figure 11. Cytoscape network depicting the metabolite concentration changes caused by the inactivation of the tricarboxylic acid cycle

Nodes colored red correspond to metabolites with an increase in concentration due to inactivated TCA cycle. Nodes colored green correspond to metabolites with a decrease in concentration due to TCA cycle inactivation. Nodes colored blue correspond to metabolites that are not observed in the NMR spectra, do not have a reference NMR spectrum (or assignment) or did not exhibit a significant concentration change. Metabolic pathways are labeled on the network.

ACA: Acetaldehyde; ACE: Acetate; ACP: Acetyl-P; AKG: α -ketoglutarate; ALAAc: Acetyl-alanine; AMI: 4-aminobutanoate; CIR: Citrulline; CIT: Citrate; ETH: Ethanol; F6P:

Fruc-6P; G1P: Gluc-1P; G6P: Gluc-6P; GAL: Galacturonic acid; GAL1P: R-D-gala-1P; GLR: Glucuronate; GLS: D-glucosamine; GLS6P: Glucosamine-6P; GLSAC: N-Ac-D-glucosamine; GLSAC6P: Acetyl-glucosamine-6P; GLUAc: Acetyl-glutamate; GLY: Glyceraldehyde; ICI: Isocitrate; INO: Ino, Ade, Xan; LAC: Lactate; MANAc: N-acetyl-D-mannosamine; MIN: Myo inositol; MSE: Selenomethionine; NEUAc: N-Ac-neuraminate; ORN: Ornithine; ORNAc: Acetyl-ornithine; PEP: Phosphoenolpyruvic acid; PRO: Proline; RIB: D-ribose; SAM: S-adenosyl-methionine; SER: Homoserine; SUCSER: O-Succinyl-L-homoserine); TCA: Tricarboxylic acid; UDPGLR: UDP-glucuronate; UDPGLSAC: UDP-NAc-D-glucosamine; VAL: Valine.

Reprinted with permission from [124]. © American Chemical Society (2011).

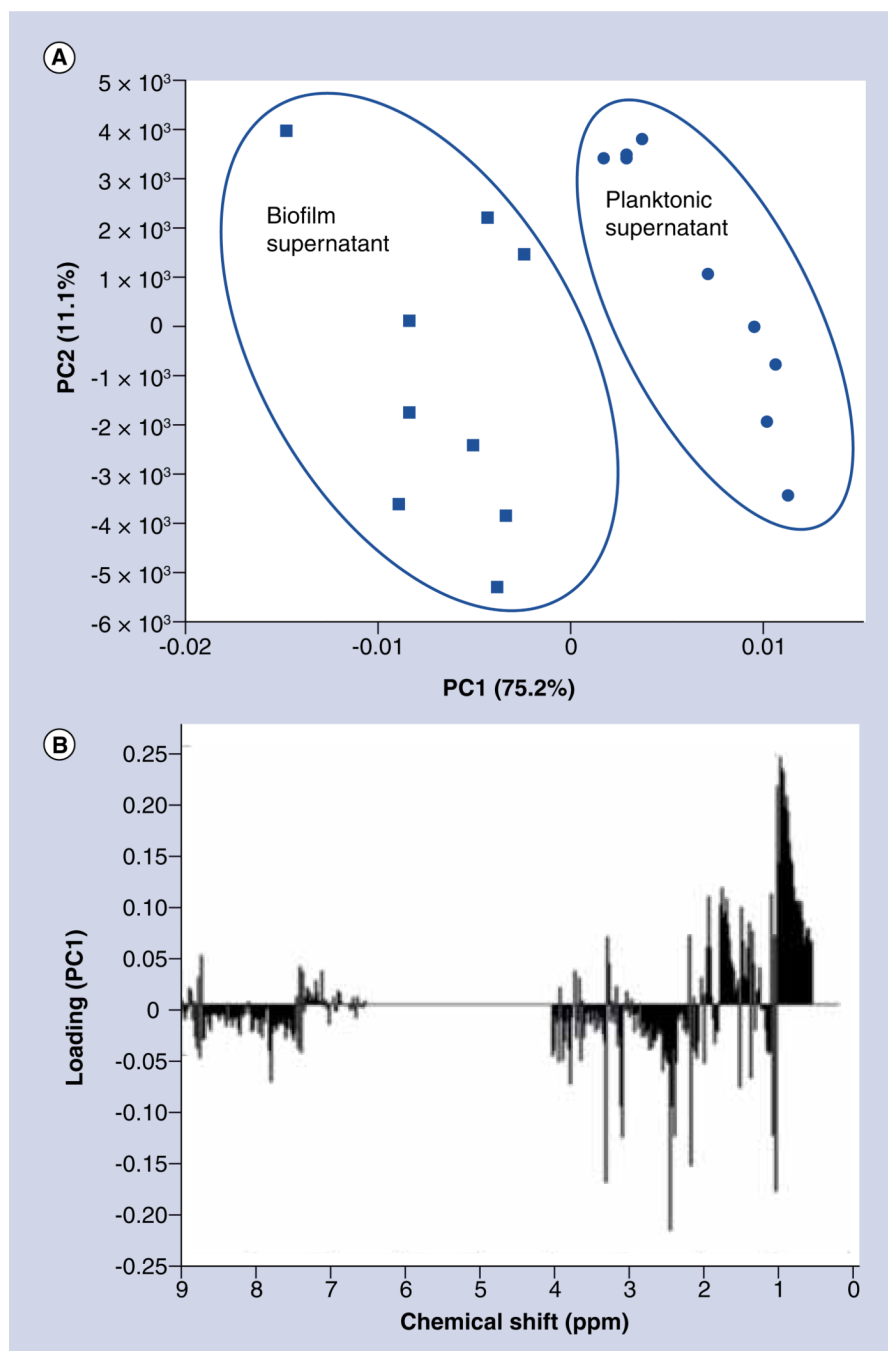


Figure 12. Metabolome difference between *Pseudomonas aeruginosa* planktonic and biofilm cells (A) PC analysis score plot for chemostat planktonic and biofilm cell samples for *Pseudomonas aeruginosa*. (B) Loading plot of the first PC.

PC: Principal component.

Reprinted with permission from [149] © American Chemical Society (2007).

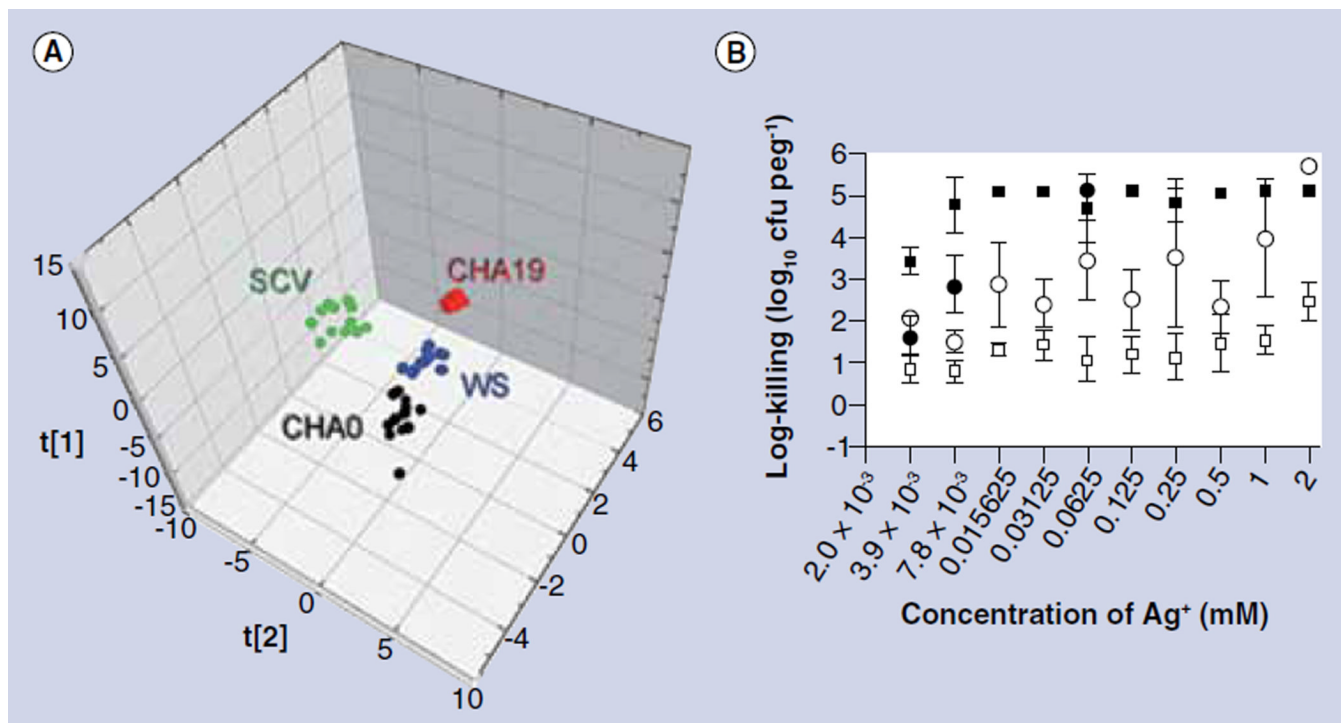


Figure 13. NMR characterization of different *Pseudomonas fluorescens* biofilm cultures
(A) Partial least squares discriminant analysis (PLS-DA) scores plot of the metabolite concentrations of *Pseudomonas fluorescens* CHA0 (black), CHA19 (red), SCV (green) and WS (blue). Each data point represents a single extract and the position determined as a linear combination of 32 metabolite concentrations obtained from the ¹H NMR spectra. The four strains could be separated along three components. **(B)** Killing curves of *P. fluorescens* CHA0 (open squares), CHA19 (filled squares), SCV (closed circles) and WS (open circles). Biofilms of each of the strains were exposed to a series of metal concentrations for 4 h followed by viable cell counting. Shown are the log killing values, which are the number of cells killed following the exposure to metal. Error bars represent standard deviation calculated from four replicates. Average cell counts for the initial unexposed controls were 5.38 ± 0.47 , 4.89 ± 0.45 , 5.23 ± 0.39 , 5.17 ± 0.47 log₁₀ cfu peg⁻¹ for CHA0, CHA19, SCV and WS, respectively.

SCV: Small colony variants; WS: Wrinkly spreaders.

Reprinted with permission from reference [246]. © Wiley-Blackwell (2010).

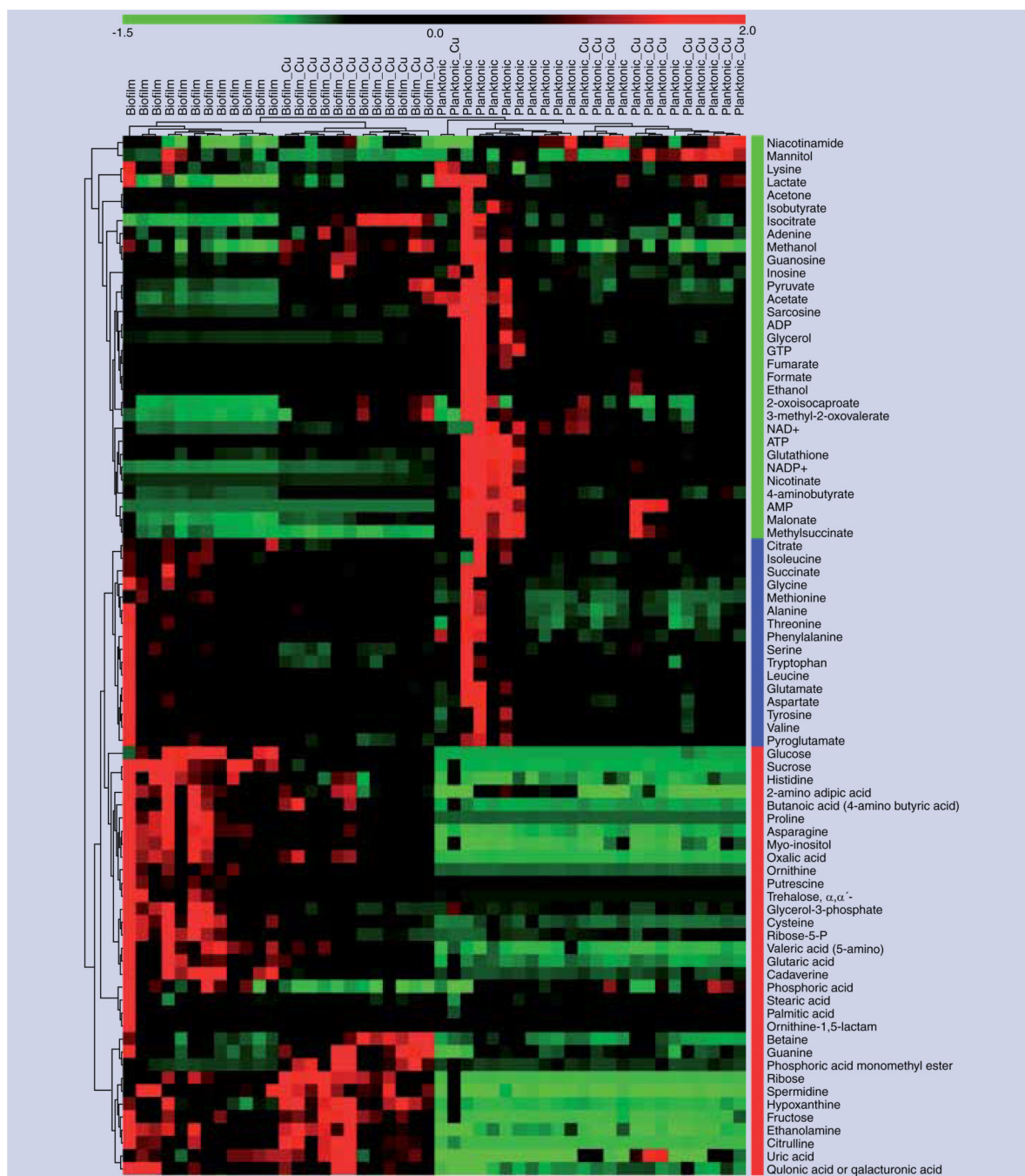


Figure 14. Hierarchical clustering analysis of control and copper-exposed biofilm and planktonic culture metabolite concentrations

Identified metabolites were analyzed for any apparent patterns. Clustering was performed using Pearson correlation as the distance metric. The cluster tree shows how the samples and metabolites divide. The samples are shown across the top of the figure, labeled by their class, and the metabolites are shown on the side. These were colored according to the grouping pattern that they showed: green metabolites were only changing in the planktonic cultures, red only in biofilms and blue were changing in both cultures in response to copper exposure.

Reprinted with permission from reference [148]. © American Chemical Society (2011).

Table 1

NMR software packages.

Software	Operating system	Source	Dimension	Ref.
ACDLab	Win	Commercial	1D and 2D	[302]
CCPNmr	Win/Mac/Unix	Academic	2D	[269,303]
Felix	Win/Unix	Commercial	2D	[304]
FERCH	Win	Commercial	1D	[305]
Gifa	Unix	Academic	1D and 2D	[270,306]
iNMR	Mac	Commercial	1D and 2D	[307]
matNMR	Win/Mac/Unix	Academic	1D and 2D	[271,308]
MestreNova	Win/Mac/Unix	Commercial	1D and 2D	[309]
NMRPipe	Win/Mac/Unix	Academic	2D	[272,310]
NMRViewJ	Win/Mac/Unix	Commercial	2D	[311]
NUTS	Win	Commercial	1D and 2D	[312]
RMN	Mac	Academic	1D and 2D	[313]
rNMR	Win/Mac/Unix	Academic	2D	[273,314]
Sparky	Win/Mac/Unix	Academic	2D	[315]
SpinWorks	Win/Mac/Unix	Academic	1D and 2D	[316]
TARQUIN	Win/Mac/Unix	Academic	Solid state	[274,317]
WIN-NMR	Win	Commercial	1D	[318]