Impact of L-arginine on *Streptococcus gordonii* gene expression and biofilm formation

Jill Christie Robinson

A thesis submitted for the degree of Doctor of Philosophy

School of Dental Sciences

Institute for Cell and Molecular Biosciences

265 Pages / 58,118 Words

February 2016

Acknowledgements

I would like to begin by expressing my utmost thanks to my main supervisor, Dr. Nick Jakubovics, for his constant patience and perseverance whilst dealing with me, my project and my thesis. This project has had some ups, some downs, and some moments that felt like running into a brick wall, but he has always been there to help me out and suggest the next course of action. His door is always open to all of us, and he always makes time for us, even when running short on it himself! He is a warm and friendly person, and an exceptional and inspirational scientist, and wherever my career takes me I will always remember the lessons I've learned under his guidance. I know that without him, this PhD project probably would have ground to a halt after the first dozen failed PCRs! I would also like to extend my deepest gratitude to my other two supervisors, Prof. Waldemar Vollmer and Dr. Matt German, for their constant guidance during our supervisory meetings, and their invaluable help and ideas, which have helped to shape my research during the past four years.

Of course, I don't know how I ever would have survived the past four years without the unfailing support of my lab mum, Mrs Lesley Old. She was always there to provide a shoulder to cry on, life guidance, advice on culturing bacteria or gram staining, or just asking for someone tall to reach a box off a high shelf for her! I don't know how I'll survive in my next workplace without her there to help me, but knowing that she will be helping other students in the same way makes it better.

When working in a research lab, you are constantly sharing ideas (and generally nattering) with those around you, and I feel I've made some excellent friends whilst working in the Oral Biology Lab. For the first year I worked there, my main company was Dr. Rob Shields, who is now pushing the boundaries of biofilm research in his new postdoctoral position in Florida. I hope people out there appreciate his antics as much as I did! Also, though I didn't always get to see much of her, Dr. Rachel Williams was always there to offer her help and friendship. I hope she finds luck with her postdoctoral position in London. After my first year, a sudden influx of new PhD students arrived – Syatirah, Suhair, Nieka, Sufian, Ahmed, Anas, Mustafa, Ali, Waleed, Halah, Martin, Andrea, Richard, and most recently, Nadia and Becky (I apologise if I've missed anyone off this list!). I may have known some of you better, and for longer than others, but you have all helped me to survive this PhD project in ways you'll never know. Also, thank you to Pam, John, Katrin, Wendy, Mark, Ruth, Luisa, Joy, Sandra, Ralf, Max and others for your help, suggestions, and good conversation over the years!

I would also like to thank all of my other friends for keeping me (not so) sane, with silly stories, pizza and films, and great memories over the years. Some of you – Steph, Rebecca, Frankie and Matthew – I've known since school, or those first days of undergraduate Microbiology, laughing and passing notes in lectures (when we probably should have been paying more attention!); others – Sarah, Daria, Andrew – I've met since starting my PhD, and have always been there to give me advice and support, and make my life that bit easier!

Finally, the people closest to me have, without a doubt, been the most supportive people in my life. My two wonderful parents, thank you so much for all of your never-ending love and support (and cups of tea!) over the years. You have always done everything you could to help me succeed, and supported me through every single up and down. I never would have got to this point without you both, and I am so deeply grateful to you for everything you've done for me. You always say that you're proud of me, and I hope very much that I can carry on making you proud.

To my Uncle Phil and Auntie Ann – thank you so much for all your kindness and support over the years. I will always appreciate it, and always remember it.

Last but not least, to my incredible partner Daniel. You've been my friend for ten years, and my boyfriend for a year and a half, and in that time you've unfailingly supported me through everything I've done, even from 5000 miles away! Thank you for your thesis-writing motivation, writing me scripts to sort my data, midnight chats and endless love and support. I can't wait to start my life with you (preferably in the same country, this time)!

Abstract

Streptococcus gordonii is an oral commensal bacterium, and an early coloniser of the acquired salivary pellicle that coats tooth surfaces. As such, it is a key organism in the establishment of dental plaque biofilms. The amino acid L-arginine has been previously shown to play a role in biofilm formation in other oral species, and depletion of L-arginine has a significant impact upon *S. gordonii* growth and gene regulation. Three L-arginine-dependent transcription regulators have been identified in *S. gordonii*, but it is currently not clear how these co-ordinate to sense and respond to changes in the exogenous L-arginine concentration. Therefore, the major aims of this work were to (i) further elucidate the impacts of L-arginine on *S. gordonii* growth and biofilm formation, (ii) investigate the roles of three putative arginine-dependent regulators in modulating arginine-responsive gene regulation, and (iii) assess the effects of L-arginine-dependent gene regulators on *S. gordonii* biofilm formation.

Initial growth experiments revealed that high concentrations (≥500 mM) of L-arginine retard *S. gordonii* planktonic growth in a chemically defined medium, resulting in lower growth yields than intermediate (0.5 mM) L-arginine. However, 500 mM L-arginine was not toxic to *S. gordonii* cells incubated in natural human saliva. *S. gordonii* has previously been shown to be conditionally auxotrophic for L-arginine, since it can biosynthesise L-arginine under strictly anaerobic conditions or during the gradual depletion of extracellular L-arginine, but it cannot grow following a rapid shift to medium lacking L-arginine. A similar lack of growth was also found following rapid depletion of L-histidine and the branched-chain amino acids. *S. gordonii* is predicted to encode all genes required for L-histidine and branched chain amino acids, and it is possible that this organism is conditionally auxotrophic for multiple amino acids.

Rapid depletion of L-arginine was shown previously to result in a change in expression of >20% of the *S. gordonii* genome. By comparing expression levels of some of the most strongly arginine-regulated genes when cells were challenged with depletion of L-histidine or branched chain amino acids, it was shown that some of the genes (for example, *argC*, SGO_1686, *asp5*) were specifically regulated by arginine depletion, whereas others (*bfbF*, SGO_1699) were similarly regulated following depletion of all amino acids. Therefore, it appears that depletion of L-arginine results in both an arginine-specific response and a more generalised stress response, presumably associated with growth arrest in this medium.

Investigation of the roles of three arginine-dependent regulators (ArcR, ArgR and AhrC) by gene expression microarrays identified a number of genes that were arginine-responsive and were differentially-regulated in the wild-type compared with the isogenic mutants Δ*arcR*, Δ*argR* or Δ*ahrC*. There was extensive overlap between the genes regulated by the ArgR and AhrC regulators, suggesting that these regulators perform similar and interdependent roles in *S. gordonii*. Regulatory responses following *arcR* disruption were distinct from those seen in the *argR* and *ahrC* mutants. In addition to three loci that have previously been described, one particular gene, SGO_0846, encoding a hypothetical protein, was highly up-regulated in response to *arcR* deletion. This thesis is the first holistic study of the three arginine-dependent regulators in *S. gordonii*, and shows that each one plays a key role in arginine-dependent gene regulation.

Finally, previous unpublished work from our group had demonstrated that *S. gordonii* Δ*arcR* displayed a defective biofilm phenotype, whereas the deletion strains of the other two regulators showed no such phenotype. To determine whether up-regulation of SGO_0846 is responsible for the biofilm attenuation in *S. gordonii* Δ*arcR*, a deletion mutant of SGO_0846 was constructed in both the wild-type and Δ*arcR* background. Disruption of the SGO_0846 gene showed no significant differences in biofilm formation levels in comparison to the wild-type background, and showed no effect on the biofilm defective phenotype of the Δ*arcR* mutant. This suggests that changes in expression of SGO_0846 are not responsible for the biofilm defects seen in the Δ*arcR* knockout, and that the ArcR regulator is affecting biofilm formation via another mechanism.

In conclusion, this thesis provides evidence that arginine has a clear impact on gene expression and biofilm formation in *S. gordonii*, and furthermore, that the ArcR regulator is critical for optimal biofilm formation. It is possible that in the future, this could be used as a target for controlling *S. gordonii* biofilm formation, and subsequent dental plaque development.

List of abbreviations

Δ	Gene deletion
µg	Microgram
µL	Microlitre
µm	Micrometre
µM	Micromolar
AD/ADI	Arginine deiminase
ADP	Adenosine diphosphate
ADS	Arginine deiminase system
Arg	Arginine
ATP	Adenosine triphosphate
A_X	Absorbance (at X nm)
BCAA	Branched-chain amino acids
BHY	Brain heart yeast extract broth
bp	Base pairs
°C	Celsius
CCR	Carbon catabolite repression
cDNA	Copy deoxyribonucleic acid
CDM	Chemically-defined medium
CFS	Cell-free saliva
CFU	Colony Forming Unit
CO_2	Carbon dioxide
COGFun	Functional clusters of orthologous genes
CSP	Competence signalling peptide
dH_2O	Deionised water
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DTT	Dithiothreitol
eDNA	Extracellular deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EPS	Extracellular polymeric substances

FDR	False discovery rate
x g	G-force
GAS	Group A streptococci
GCF	Gingival crevicular fluid
Gluc.	Glucose
gp340	Salivary glycoprotein 340
h	Hour
HCl	Hydrochloric acid
HF	High fidelity
Hist	Histidine
HS	Horse serum
Hz	Hertz
Ile	Isoleucine
kb	Kilobases
L	Litre
LB	Luria Bertani
Leu	Leucine
M	Molar
$MgCl_2$	Magnesium chloride
min	Minute
MIQE	Minimum Information for Publication of Quantitative Real-Time PCR Experiments
mL	Millilitre
mm	Millimetre
mM	Millimolar
MPa	Megapascals
Mpixel	Megapixels
mRNA	Messenger ribonucleic acid
MTP	Microtitre plate
NA	No annotation
NCBI	National Centre for Biotechnology Information
NS	No signal

nt	Nucleotide
nm	Nanometres
OD_X	Optical density (at X nm)
ovex	Overlap extension (PCR)
PBS	Phosphate buffered saline
PCA	Principal component analysis
PCR	Polymerase Chain Reaction
pfp	Percentage of false positives
PIA	Polysaccharide intercellular adhesin
PTS	Phosphotransferase System
qPCR	Quantitative polymerase chain reaction
qRT-PCR	Quantitative reverse transcriptase polymerase chain reaction
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
SAGP	Streptococcal acid glycoprotein
SD	Standard deviation
s	Second
SEM	Standard error of the mean
Spp.	Species
term	Terminator
TFBS	Transcription factor binding site
THYE	Todd Hewitt Yeast Extract broth
T_m	Melting temperature
tRNA	Transfer ribonucleic acid
U	Units
V	Volts
Val	Valine
v/v	Volume/volume
WHO	World Health Organisation
WT	Wild-type

List of Figures

List of Tables

Contents

1. Introduction

1.1. Bacterial biofilms

1.1.1. Prokaryotic organisms

Bacteria are arguably some of, if not the most prevalent life forms on the planet, estimated to number between 4-6 x 10^{30} cells (Whitman *et al.*, 1998) and to account for half of the "living protoplasm" on earth (Kluyver and van Niel, 1957). From the surfaces we touch, to the soil in which we cultivate our crops; from the water we drink to the air we breathe, bacteria are consistently present throughout. Even the most difficult conditions are home to these prokaryotic life forms. We have only to look to the abilities of the extremophilic microorganisms, from psychrophiles such as *Vibrio* spp. in surviving extreme cold, able to sustain replication and metabolic activity to such low temperatures as -20°C (D'Amico *et al.*, 2006); piezophiles such as *Moritella* spp. in surviving intense oceanic pressures of over 0.1 MPa (Lauro and Bartlett, 2008), salt-loving halophiles such as *Halococcus* spp. and heat-loving thermophiles like *Thermus* spp. to see that bacteria are able to adapt to not only survive, but thrive, in any environment. One such way that these prokaryotic organisms can adapt to survive, is by living in a sessile community.

From the first observations of bacteria by the Dutch scientist Antonie van Leeuwenhoek in 1684, within the dental plaque scraped from his teeth, right through most of the 19th century, the focus of research into these micro-organisms was almost exclusively on isolated strains in the planktonic phase of growth; something later termed the "pure culture period" of early microbiology (Atlas and Bartha, 1997). However, it is now estimated that less than 0.1% of all microbial life actually exists in this planktonic form, with the vast majority of cells surviving instead in a sessile, anchored form (Bjarnsholt, 2011). This sessile state is one way in which bacteria can adapt to aid their own survival in a multitude of different environments, and has become known as a biofilm.

1.1.2. Biofilms

The term "biofilm" was coined in 1978, by Costerton *et al.* (1978), and has been used widely since then. Many researchers have since published their own definitions of what a biofilm

consists of, and in 1999, Costerton followed suit and defined a biofilm as "a structured community of bacterial cells enclosed in a self-produced polymeric matrix and adherent to an inert or living surface" (Costerton *et al.*, 1999). A biofilm forms when bacterial (or fungal) cells find themselves at, or actively move to, an interface between two phases. This can consist of solid-air and liquid-air interfaces, but generally involves a solid, inert surface and an aqueous liquid phase (Donlan, 2002). Biofilms are able to form under any conditions such as these, for example, in marine and aquatic environments such as on ship hulls (Angst, 1923) and in ponds (Geesey *et al.*, 1977), and on surfaces in or on the human body. These include both artificial materials, such as catheters, pacemakers and other in-dwelling plastic and metal medical devices (Donlan, 2001), and natural surfaces, such as on skin, in chronic wound infections and on tooth surfaces, which will be discussed in more detail later.

1.1.3. Biofilm formation

It is important to note that this overview of biofilm formation is a general perspective, and that individual systems have their own specific issues. Generally, surface interfaces, particularly those between liquid and solid phases, are often covered and conditioned with a proteinaceous film consisting of organic or inorganic molecules derived from the aqueous phase above (Garrett *et al.*, 2008). Biofilm formation initiates when bacterial cells move close to this interface, and form a transitory attachment to it (Watnick and Kolter, 2000). This transient association is mediated by van der Waals forces, and electrostatic interactions between the cell surface and the solid substratum (Percival *et al.*, 2011), known as the DLVO theory (Hermansson, 1999). The cells then use protein adhesins and polysaccharides expressed on their surfaces to interact with the proteins within the conditioning film, and so form longer lasting connections with the solid surface (Garrett *et al.*, 2008).

The first bacterial cells to attach to a surface, the early coloniser species, settle on the conditioned surface to form a monolayer (Moorthy and Watnick, 2004) (Figure 1.1). This is a covering of single-cells, or cell clusters, that become transiently attached to the surface through non-specific binding mediated by electrostatic interactions and other forces (Flemming and Wingender, 2010). Cells are then able to initiate receptor-adhesin interactions between their cell surface, and the surface to which they are binding, to further adhere themselves to the

interface. It is thought that a mixture of factors, including conditioning of the solid surface, cell-surface hydrophobicity, solid surface roughness and the expression of fimbriae or flagella by the cells all influence and aid attachment of bacterial cells to a solid surface (Donlan, 2002; Vu *et al.*, 2009). Following initial binding of the cell monolayer to the surface by specific receptor-adhesin interactions, subsequent production of extracellular polymeric substances (EPS) further strengthens attachment of cells to the surface (Flemming and Wingender, 2010). As well as aiding attachment, EPS also helps to form the scaffold to which other cells can bind, and form a three-dimensional biofilm.

Following this irreversible attachment of cells to the inert surface, other cells are then recruited to the burgeoning biofilm. These may be more cells of the same species, as in monospecies biofilms such as *Staphylococcus aureus* colonisation of catheters; or cells of a different species, such as the multispecies biofilms commonly found in chronic wound infections (Dalton *et al.*, 2011). These cells bind to the existing cell monolayer by further receptor-adhesin interactions, whilst simultaneously secreting EPS to help protect and stabilise the biofilm structure (Flemming and Wingender, 2010).

Finally, following further recruitment of late coloniser cells and subsequent EPS production, a mature biofilm structure is formed. Cells may then naturally disperse from the biofilm, in order to target other environments for colonisation. In the case of motile cells, dispersal is a simple task of the cells actively transporting themselves to a new site via chemotaxis, as with *Pseudomonas aeruginosa*; however, non-motile cells require enzymes to release them from within the biofilm, such as the dispersin B enzyme of *Aggregatibacter actinomycetemcomitans* (Ramasubbu *et al.*, 2005). Natural dispersal of cells from within the biofilm allows perpetuation and propagation of the biofilm colonisation of the target interface.

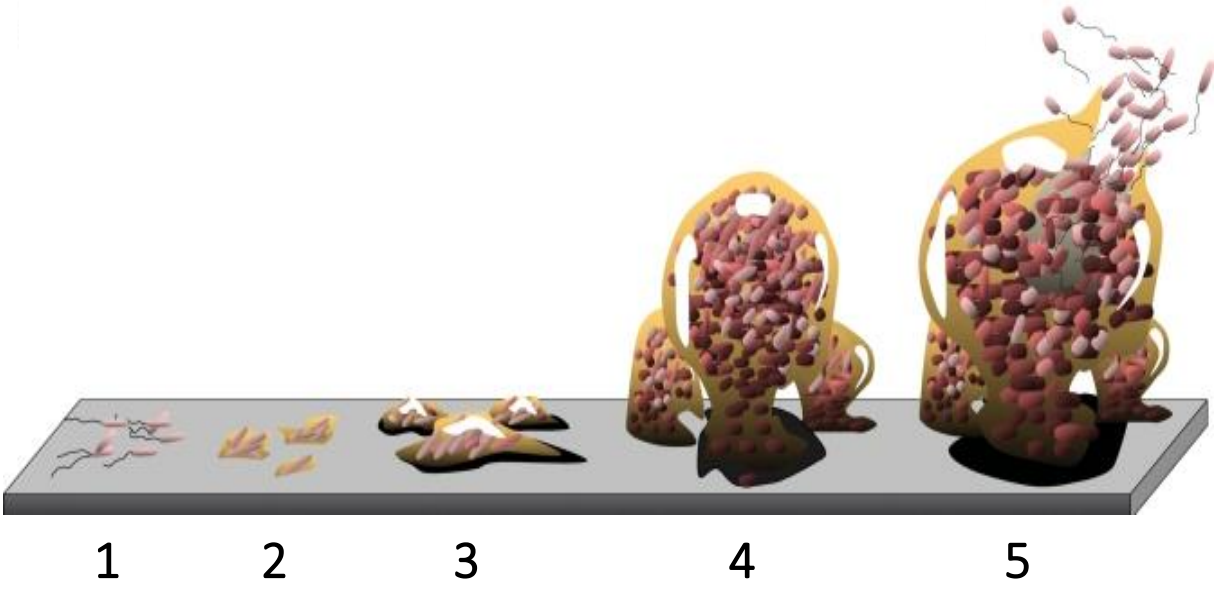


Figure 1.1. Diagrammatic representation of the stages of biofilm formation and development. Stages represent: 1) Initial attachment of cells to surface. 2) Irreversible attachment of monolayer following extracellular polymeric substance (EPS) production. 3) Recruitment of more cells to surface, resulting in early biofilm maturation. 4) Production of a mature three-dimensional biofilm structure, surrounded by the EPS matrix. 5) Release of cells from the surface of the biofilm to colonise elsewhere. Figure adapted from Monroe (2007).

1.1.4. Biofilm matrix

The mature biofilm consists of a heterogenous, channelled, three-dimensional network of cells (Wood *et al.*, 2002), between which EPS has been secreted to fill the matrix of the biofilm, which can account for up to 90% of the total biofilm weight (Flemming and Wingender, 2010). The biofilm matrix consists of many different components, including macromolecules such as extracellular DNA (eDNA), proteins, polysaccharides, in addition to smaller molecules such as lipids, enzymes, water, and organic and inorganic molecules.

The macromolecules of the matrix are thought to work together to form a scaffold upon which the cells can bind. This scaffold involves proteins such as amyloid fibres, produced by the FapC protein of many different *Pseudomonas* strains, including *P. fluorescens* and *P. aeruginosa* (Dueholm *et al.*, 2010); carbohydrates such as cellulose, produced by the *Escherichia coli bcs* genes (Zogaj *et al.*, 2001); and eDNA, released by oral streptococci *Streptococcus gordonii* and *Streptococcus sanguinis* in response to various environmental stimuli, including hydrogen peroxide (Kreth *et al.*, 2009b). These molecules are actively secreted across the cell envelope, via secretion systems in the case of proteins and polysaccharides, and by various different mechanisms in the case of eDNA, including secretion and autolysis (Bayles, 2007; Jakubovics, 2013). In addition to helping to structure the biofilm, eDNA can also thermodynamically aid in cell-cell and cell-surface adhesion (Das *et al.*, 2010), and mediate horizontal gene transfer between different species and cells within the biofilm (Flemming and Wingender, 2010). It is also thought to act as a source of carbon and nutrients for other bacterial cells within the biofilm (Finkel and Kolter, 2001).

The smaller molecular components of the biofilm matrix, unlike the macromolecules, serve non-structural functions. For example, enzymes within the biofilm matrix include polysaccharidases, hydrolases and lyases, with roles in degrading matrix components for nutrients and reshaping the architecture of the biofilm (Sutherland, 1999; Wingender *et al.*, 1999; Jakubovics, 2013), whilst water and other compounds within the biofilm serve to hydrate and nourish the cells.

The matrix of the biofilm serves many purposes to the bacterial cells. These include aiding adhesion (both between cells and to the substratum), and protecting cells from host defenses,

chemotherapy and environmental insults (Stewart and Costerton, 2001; Flemming and Wingender, 2010). In particular, the matrix is able to confer strong antimicrobial resistance upon the biofilm cells. One such example of this is in a β-lactamase-negative strain of the pulmonary pathogen *Klebsiella pneumoniae*, which had a minimum inhibitory concentration (MIC) in planktonic culture of 2 μg/mL ampicillin, but was able to survive treatment with up to 5000 μg/mL ampicillin when growing within a biofilm (Anderl *et al.*, 2000; Stewart and Costerton, 2001).

The exact mechanism by which the matrix (or biofilm) protects the bacterial cells from antibiotic treatment is not known, however, many mechanisms for resistance have been proposed. For example, resistance may be mediated by the presence of persister cells within the biofilm, a small population of cells that are highly resistant to antimicrobial compounds. One example of this is within *P. aeruginosa* biofilms, where cells were found to be resistant to the antibiotic ciprofloxacin, which was a substrate for a multi-drug resistance pump within the cells (Broun *et al.*, 2000). However, the bacterial cells were found not to be dependent on this pump for ciprofloxacin resistance, making the mechanism by which they were resistant unknown. These were designated persister cells. Another possible mechanism for antimicrobial resistance is retardation of antibiotic or chemical diffusion into the biofilm by the EPS in the matrix. Studies investigating tobramycin and gentamicin treatment of *P. aeruginosa* biofilms indicated that one EPS component, eDNA, is able to chelate cations within the matrix of the biofilm, and therefore sequester aminoglycosides and cationic antimicrobial peptides (Mulcahy *et al.*, 2008; Chiang *et al.*, 2013).

1.1.5. Significance of biofilms

Irrespective of the mechanism by which biofilm cells survive antibiotic treatment, the fact that they are able to survive is extremely significant for the management of infections. In fact, this ability to survive antimicrobial treatment was classified as one of the criteria for diagnosis of a biofilm infection in clinical specimens, alongside association with a surface, aggregation of cells, and localised infections that can result in culture-negative test results (Parsek and Singh, 2003; Hall-Stoodley and Stoodley, 2009).

Bacterial biofilms are estimated to be involved in "65% of human bacterial infections", according to the US Centre for Disease Control (Potera, 1999), and so the medical burden of antibiotic resistance in biofilms is significant. In particular, because of the types of infection often caused by biofilms, such as infections of in-dwelling medical devices such as pacemakers and prosthetic joints, treatment of these infections is often costly and impacts upon the patient's quality of life. It has been estimated that the cost of fitting, for example, a prosthetic hip joint is £3-4000, but the cost of treating an associated infection is £20-30,000 (Allison, 2000). This is because if antibiotic treatment of the infected area fails, the most reliable way of eradicating the resistant biofilm infection is by surgical or mechanical removal of the in-dwelling device causing the infection, such as an artificial joint. Part of the costs of treating a device-associated infection are often due to patient care, as they are unable to have the device replaced until the infection is cleared, which can be several months in the case of joint prostheses.

This is also the case with dental implant failure, due to oral inflammatory diseases such as peri-implantitis, mucositis, or gingivitis (Lang *et al.*, 2000). These diseases are all known to be triggered by the oral microbes colonising the surfaces of the nutrient-rich environment of the oral cavity, in the form of biofilms. The most widely-known example of these biofilms is the dental community known as plaque.

1.2. Dental plaque

1.2.1. Microbiology of the oral cavity

One of the common biofilms found within humans is the oral biofilm known as dental plaque. It is a polymicrobial community of bacterial and fungal cells, which occupies the surfaces of the teeth. In addition, oral biofilms also colonise other areas of the mouth, such as soft tissues.

Healthy individuals often have a wide range of microbial species in their oral cavity, with a number of taxa being particularly prevalent. Amongst these are the *Actinomyces, Haemophilus, Rothia* and *Streptococcus* genera, although different bacterial species prefer to inhabit different locations within the oral cavity (Eren *et al.*, 2014). According to the study by Eren *et al.* (2014), which used 16S rRNA oligotyping to characterise the prevalence of different bacterial species

on a number of surfaces within the mouth (sub- and supra-gingival plaque, tonsils, throat, saliva, hard palate, tongue dorsum, gingiva and buccal mucosa), different bacterial species within these genera preferentially colonised different areas of the mouth. For example, *Fusobacterium peridonticum* was found most abundantly on the tongue dorsum, alongside *Campylobacter concisus*, whereas other *Fusobacterium* and *Campylobacter* species were not as common on these sites, and found more often elsewhere. A number of *Streptococcus* species, including *S. mitis* and *S. infantis*, were found across all sites, in all individuals tested. Another study by Aas *et al.* (2005), using 16S rRNA sequencing, found the two most prevalent species within the entirety of the oral cavity were the Gram-positive organisms *Streptococcus mitis,* and *Gemella haemolysins,* a prevalent coloniser of human mucous membranes (Berger, 1985). These species occupied every one of the aforementioned sites within the mouth, in a number of different individuals. However, other species, such as the Gram-positive coccal species *S. gordonii*, *Rothia dentocariosa*, and members of the *Actinomyces* spp., were found mainly in one location – in this case, the surfaces of the tooth (Aas *et al.*, 2005). The microorganisms that bind to the tooth surface comprise the biofilm known as dental plaque.

1.2.2. Dental plaque biofilms

Dental plaque microorganisms bind not to the actual tooth surface, but to the acquired salivary pellicle that coats the tooth surfaces within minutes of brushing. This pellicle is acquired from whole saliva, which consists of secretions from three different pairs of glands – the sublingual, submandibular and parotid glands (Yao *et al.*, 2003). Whole saliva contains many different components, in particular, glycoproteins such as agglutinin, lactoferrin and fibronectin, and low-molecular-weight mucins, which are common targets for binding by oral bacteria looking to establish biofilms (Takamatsu *et al.*, 2006; Jakubovics *et al.*, 2009; Zijnge *et al.*, 2010; Nobbs *et al.*, 2011). These salivary components accumulate on the tooth surface, and form targets for different species of bacteria and fungi forming the plaque biofilm.

Dental plaque is widely reported to consist of between 100-200 different species of bacteria and fungi, in a healthy individual (Kolenbrander *et al.*, 2010), and these species come from a pool of between 500-700 different taxa that are able to live in the mouth (Aas *et al.*, 2005). However, more recent studies involving 16S rRNA pyrosequencing and other forms of high

throughput sequencing have indicated that these numbers could be considerably higher (Keijser *et al.*, 2008; Dewhirst *et al.*, 2010). Aas *et al.* (2005) demonstrated that *Streptococcus* spp. were one of the six main phyla in the oral cavity, and prevalent colonisers of both tooth surfaces and subgingival areas of the mouth. Subsequent studies using 16S rRNA sequencing or global sequence alignment analysis have also shown *Streptococcus* spp. to be among the most abundant species in the oral cavity (Lazarevic *et al.*, 2009; Dewhirst *et al.*, 2010; Ahn *et al.*, 2011), with one study demonstrating that streptococcal species were prevalent in the saliva, dental plaque and mucosa of human hosts of all ages, right from neonates to elderly adults who had permanent dentition (Xu *et al.*, 2015). The previously mentioned study by Eren *et al.* (2014) further demonstrated that the largest variety of *Streptococcus* species could be found within the sub-gingival and supra-gingival plaque regions.

These oral streptococci have been shown to express large numbers of surface receptors (Nobbs *et al.*, 2009), and dominate within the early coloniser species of tooth surfaces. In healthy individuals, streptococci can form up to 80% of the early plaque biofilm (Rosan and Lamont, 2000), with the most common primary colonisers being *S. mitis, S. oralis, S. sanguinis, S. gordonii* and *Actinomyces oris* (Figure 1.2) (Rickard *et al.*, 2003). These species all present many receptors on their surfaces, and have the ability to bind not only to the acquired pellicle on the tooth surface, but also to each other. They are the first species to initiate formation of the plaque biofilm, and by co-aggregating with each other, help to stabilise the foundations of the early biofilm.

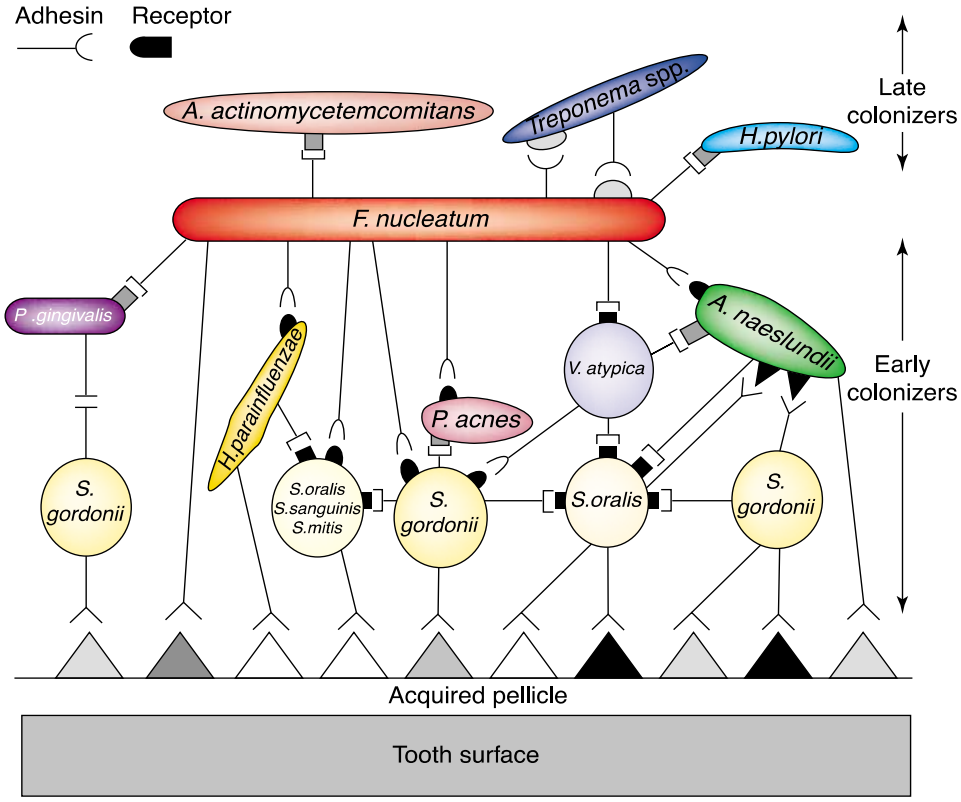


Figure 1.2. *Streptococcus* spp. colonisation of the acquired salivary pellicle. The early coloniser species, in particular *Streptococcus gordonii*, are shown to bind to many different bacterial species. Diagram taken from Rickard *et al.* (2003).

1.2.3. Plaque species interactions

Coaggregation between different species within the plaque biofilm is important for different reasons, including biofilm development, stability and structure. In particular, coaggregation between oral *Streptococcus* spp. and *Actinomyces* spp., and specifically between the organisms *Streptococcus gordonii* and *Actinomyces oris,* has been well-characterised, with the streptococcal surface protein SspB shown to be responsible for binding to *A. oris* (Rosan and Lamont, 2000; Jakubovics et al., 2005), and recently shown to specifically bind to a surface-associated polysaccharide produced by *A. oris* (Back *et al.*, 2015).

Table 1.1 shows some of the interactions made between *Streptococcus* spp. surface receptors, and host and microbial proteins, which aid in colonisation of the host surfaces and stabilization of the early biofilm. The early coloniser species bind to the pellicle by specific interactions between the adhesins on the bacterial surface, and salivary proteins such as agglutinin and fibronectin. More cells are then recruited to the developing biofilm through cell-cell communication via quorum sensing, using the molecules autoinducer-2 (AI-2) and streptococcal competence signalling peptide (CSP) (Loo *et al.*, 2000; Cvitkovitch *et al.*, 2003). Following early coloniser adherence, secondary coloniser species, such as *Porphyromonas gingivalis* and *Fusobacterium nucleatum*, are then recruited to bind the existing cell community, commencing formation of the three-dimensional structure of the biofilm.

Bacterial species	Adhesin	Target for binding	Reference
S. gordonii, *S. oralis*, *S. mitis*	SspA/B (antigen I/II)	*P. gingivalis*, *Streptococcus mutans*, *A. naeslundii*, *Candida albicans*, salivary gp340	(Rosan and Lamont, 2000; Jakubovics *et al.*, 2005)
S. gordonii	CshA/B	*A. naeslundii*, *C. albicans*, *Streptococcus oralis*, human fibronectin	(Rosan and Lamont, 2000; Jakubovics *et al.*, 2009)
S. gordonii	ScaA	*A. naeslundii*	(Rosan and Lamont, 2000)
S. gordonii	Hsa	Human fibronectin, platelets, sialic acid residues	(Jakubovics *et al.*, 2009)
S. salivarius	VBP	*Veillonella parvula*	(Rosan and Lamont, 2000)
S. cristatus	SrpA	*Fusobacterium nucleatum, Corynebacterium matruchotii*	(Handley *et al.*, 2005)
S. gordonii	GtfG	Endothelial cells	(Vacca-Smith *et al.*, 1994)
S. parasanguinis	Fap1	Salivary protein saccharide residues, platelets	(Ramboarina *et al.*, 2010)

Table 1.1. Examples of a wide-number of streptococcal surface adhesins, and their unique binding specificities.

In particular, *Fusobacterium* is an important secondary coloniser species as these cells are found in the middle layers of dental plaque (Zijnge *et al.*, 2010), able to bind to many different species within the biofilm, and are therefore often described as bridging organisms between the Gram-positive early colonisers and the Gram-negative late coloniser species (Kolenbrander and London, 1993). The ability of these cells to coaggregate is important for the formation of a structured multispecies biofilm (Kaplan *et al.*, 2009; Kolenbrander *et al.*, 2010). *F. nucleatum* expresses a number of surface adhesins that allow it to bind to different early and late coloniser species. For example, the arginine-inhibitable outer membrane protein RadD allows *F. nucleatum* to bind early coloniser streptococci, such as *S. cristatus, S. gordonii* and *S. sanguinis* (Edwards *et al.*, 2007), and also to the hyphae of *Candida albicans* (Wu *et al.*, 2015). However, another outer membrane protein, FomA, allows *F. nucleatum* to bind the late coloniser species, *P. gingivalis* (Kinder and Holt, 1993), a coaggregation reaction that is thought to enhance periodontal inflammation and disease (Liu *et al.*, 2010). Additionally, *F. nucleatum* can also be bound by the Msp protein of *Treponema denticola*, another late coloniser species linked to periodontal disease (Rosen *et al.*, 2008).

Finally, tertiary and late coloniser species can then bind to the top of the biofilm, becoming encased in extracellular polymeric substances that will protect them from the outside environment. Cells can later shed from the biofilm, in order to move elsewhere to colonise within the oral cavity (see section 1.1.3).

1.2.4. Biofilm-associated diseases

When the balance of species within the plaque biofilm shifts, diseases of the oral cavity can occur. Gingivitis, inflammation of the gingiva (or gums), and periodontitis, inflammation of the periodontal tissue which supports the teeth within the jaw, are both triggered by increased colonisation of subgingival sites by Gram-negative anaerobic microorganisms such as *P. gingivalis* and *T. denticola* (Pihlstrom *et al.*, 2005), and are found in up to 20% of middle-aged adults worldwide (World Health Organization, 2012). These are also the same organisms that are associated with dental implant-associated inflammation, such as peri-implantitis (Shibli *et al.*, 2008). Secretion of gingival crevicular fluid (GCF) from serum, which infiltrates periodontal pockets, is thought to be part of the inflammatory response that leads to periodontal diseases,

but also propagates survival of the organisms that cause it (Seneviratne *et al.*, 2011). GCF contains proteins such as albumin and fibrinogen, and can also contain passively transported immune cells such as neutrophils and leukocytes (Rahnama *et al.*, 2014). The recognition of bacterial colonisation of subgingival spaces by these immune cells triggers an inflammatory response, which in turn leads to an increase in GCF secretion, and can raise the pH of the subgingival spaces to pH 7.5 (Marsh, 1994). These neutral or alkaline pH are also thought to benefit periodontal pathogens, known as the "red complex" of bacterial species, consisting of *P. gingivalis*, *T. denticola* and *Tannerella forsythia* (Socransky et al., 1998), allowing them to better survive and can cause an over-abundance of these species in the plaque within periodontal spaces (Seneviratne *et al.*, 2011).

However, one new idea about the role of microorganisms in the initiation of periodontitis is the "keystone pathogen" hypothesis. This suggests that, rather than the presence of red complex bacteria being solely responsible for triggering periodontitis, certain species known as "keystone pathogens" (such as *P. gingivalis*) can modulate the host immune response, helping to tip the balance of microbial species within the biofilm over into a dysbiotic complex and resulting in the further accretion of species associated with disease (Hajishengallis *et al.*, 2011). This was demonstrated in mouse models of periodontitis, where low level colonisation of the oral cavity by *P. gingivalis* resulted in changes in microbiota composition, and additionally modulated the hose immune response by way of complement receptors C3a and C5a, leading to increased periodontal bone loss (Hajishengallis *et al.*, 2011). This was further supported by the lack of bone loss in C3a and C5a receptor-deficient mice following inoculation with *P. gingivalis*.

Furthermore, there is evidence that keystone pathogens depend on certain bacterial species, such as *S. gordonii*, to integrate into oral biofilms. *S. gordonii* is able to co-aggregate with *P. gingivalis* via interactions between the streptococcal SspB surface protein and the *Porphyromonas* FimA/Mfa1 fimbrial proteins, thus aiding *P. gingivalis* recruitment to the biofilm (Lamont *et al.*, 2002). Consequently, *S. gordonii* may be considered an "accessory pathogen". Whilst not necessarily a pathogen in itself, its ability to specifically interact with keystone pathogens or species associated with disease, results in accumulation of these pathogenic species within the biofilm and so increases the potential for subsequent dysbiosis

(Hajishengallis and Lamont, 2012). Accordingly, co-infections of periodontal mouse models with both *S. gordonii* and *P. gingivalis* resulted in increased alveolar bone loss in comparison to monospecies infections (Daep *et al.*, 2011). Therefore it seems that symbiotic and dysbiotic relationships within oral biofilms are responsible for oral health and disease, as opposed to one species being related to health, and another to disease.

As mentioned above, increased pH may have the ability to propagate the survival of keystone pathogens within the oral cavity, which could in turn lead to periodontitis. Conversely, acidic conditions are also damaging to oral health. Dental caries, also known as tooth decay, affects over 90% of adults worldwide to some degree (World Health Organization, 2012). Bacterial species such as *Streptococcus mutans* and *Lactobacillus acidophilus* are capable of rapidly metabolising carbohydrates such as glucose or sucrose, to form organic acids such as lactic and acetic acid (Featherstone, 2000; Featherstone, 2008). The acids they produce can erode and demineralise the enamel surfaces of the teeth, by dissolving calcium phosphate within the enamel, or the dentine layer beneath it (Featherstone, 2000), and eventually forming a cavity within the tooth structure. This can subsequently lead to pulpitis, an inflammation of the dental pulp within the tooth, which manifests as toothache for the sufferer.

The presence of acidogenic bacteria within the plaque biofilm subjects the oral cavity to prolonged periods of low pH, as opposed to the spikes of low pH that occur naturally following a meal in non-acidogenic plaque in a relationship described by the "ecological plaque hypothesis" (Marsh, 1994). This hypothesis states that changes in environmental conditions, such as pH or host inflammatory responses, causes a shift in the balance of species within the microbiome, predisposing the infection site to disease. Ordinarily, these short periods of acidity would result in demineralisation of tooth surfaces, and then be followed by an alkaline period of remineralisation, where calcium phosphate is restored to the tooth surfaces by the presence of saliva (Burne and Marquis, 2000). However, prolonged exposure of the oral cavity to acid promotes the growth of acid-tolerant (aciduric) organisms within the biofilm, and inhibits growth of the acid-sensitive organisms (Svensäter *et al.*, 1997). This helps aciduric and acidogenic (acid-producing) microorganisms to predominate within the biofilm community, further propagating low pH and consistent demineralisation of tooth surfaces.

However, some organisms are capable of counteracting the effects of the acidogenic bacteria when given exogenous amino acids such as L-arginine to metabolise into alkalis. One species such as this is the Gram-positive facultative anaerobe, *Streptococcus gordonii*.

1.3. *Streptococcus gordonii*

Streptococcus gordonii is a Gram-positive, facultative anaerobe of the Sanguinis group of streptococci (Facklam, 2002; Doern and Burnham, 2010). This group also contains the oral streptococci species *S. parasanguinis* and *S. sanguinis,* to which *S. gordonii* used to belong before reclassification as a separate species following extensive taxonomic analysis by Kilian *et al.* (1989). *S. gordonii* is a commensal of the oral cavity, and was generally considered to be a species associated with oral health, partly due to its ability to synthesise hydrogen peroxide which prevents growth of oral pathogens (Hillman *et al.*, 1985; Barnard and Stinson, 1999). However, recent research has indicated that its role is more as an accessory pathogen, despite not being a pathogenic species in itself (see section 1.2.4). It is also one of the early colonisers of surfaces within the mouth, particularly of the acquired salivary pellicle coating tooth surfaces (Kreth *et al.*, 2009a), in addition to the mucosa and within saliva (Xu *et al.*, 2015), and its colonisation is facilitated by the large number of surface proteins it expresses.

1.3.1. Infective endocarditis

Despite *S. gordonii* not playing a main role in pathogenesis in the oral cavity, it is capable of acting as an opportunistic pathogen in humans. If it can gain access to the bloodstream, and travel to the cardiovascular system, *S. gordonii* has the ability to colonise damaged heart valves and cause infective bacterial endocarditis. This is mediated in part by two allelic surface protein variants, Hsa and GspB, serine-rich repeat proteins that bind to both fibronectin on endothelial cells and platelets, aiding bacterial pathogenicity and colonisation of the valve (Takahashi *et al.*, 2006; Xiong *et al.*, 2008). Hsa and GspB are homologous proteins, both with the ability to bind to sialylated glycoproteins, particularly those found on platelet membranes such as glycoprotein Ibα (Bensing *et al.*, 2004; Takamatsu *et al.*, 2006). This allows formation of a thrombus, in addition to bacterial colonisation of the valve, causing endocarditis. Futhermore, a second protein on the surface of *S. gordonii* cells, PadA, was found to bind to platelets during

endocarditis via the fibrinogen receptor GPIIbIIIa on platelet surfaces, triggering further platelet adhesion and aggregation with *S. gordonii* (Petersen *et al.*, 2010).

The Hsa and GspB proteins also play an important role in the ability of *S. gordonii* cells to colonise the oral cavity as a commensal organism. Fibronectin is found on oral surfaces, deposited from parotid saliva by the acquisition of the salivary pellicle on the surfaces of the teeth (Babu and Dabbous, 1986). Additional binding to other sialylated glycoproteins is also mediated by Hsa and GspB, such as binding to salivary agglutinin (Takamatsu *et al.*, 2006). The sialic acid-binding capabilities of both Hsa and GspB, along with two other fibronectin-binding proteins, the hydrophobic fibrillar proteins CshA and CshB, all contribute to *S. gordonii* binding to the salivary pellicle, in turn initiating the formation of the plaque biofilm (McNab *et al.*, 1994; Bensing *et al.*, 2004; Nobbs *et al.*, 2009).

1.3.2. Arginine deiminase system

As mentioned above, *S. gordonii* is a commonly-found microorganism within the oral cavity. However, if the distribution of species within the plaque biofilm becomes skewed, for example, due to an over-abundance of *S. mutans*, which in turn causes high acid concentrations within the biofilm matrix, species that are acid-sensitive can be inhibited. *S. gordonii*, however, has some degree of acid tolerance as a result of a certain catabolism system within the cells, and it is this system which caused *S. gordonii* to be labelled as a "cariostatic" organism (Burne and Marquis, 2000), suggesting that it can play a role in alleviating the effects of the acidogenic bacteria, and the formation of dental caries. This catabolic system is called the arginine deiminase system, or ADS, and has the ability to convert the amino acid L-arginine to ammonia. The ADS of *S. gordonii* was first described in 1983, by Ferro *et al.* (1983), when *S. gordonii* was still classified as a member of the *S. sanguis* species, prior to reclassification as a novel species (Kilian *et al.*, 1989).

The ADS of *S. gordonii* consists of three enzymes, which catalyse the breakdown of exogenous L-arginine to produce ammonia, carbon dioxide, ornithine (an amino acid, and L-arginine precursor) and ATP (Cunin *et al.*, 1986). The three enzymes involved in the pathway are termed ArcA or arginine deiminase (encoded by the *arcA* gene), which converts L-arginine to citrulline

and ammonia by hydrolysis; ArcB, or ornithine carbamoyltransferase (encoded by *arcB*), which converts citrulline to carbamoylphosphate and ornithine; and ArcC, carbamate kinase (encoded by *arcC*), which dephosphorylates carbamoylphosphate, in order to form ATP, carbon dioxide and ammonia (Cunin *et al.*, 1986; Dong *et al.*, 2002). Together, these genes form part of the *arc* operon (Figures 1.3 and 1.4).

1.3.3. The arc *operon*

In *S. gordonii*, the *arcABC* locus also contains the genes *arcD*, encoding an L-arginine-ornithine antiporter, and *arcT*, a hypothetical protein thought to encode a dipeptidase enzyme, able to cleave L-arginine residues from other proteins (Dong *et al.*, 2002; Liu and Burne, 2009). These genes were thought to be co-transcribed with *arcABC* as part of an operon, but recently *arcD* was shown to have differential expression to the *arcABC* genes under changing L-arginine concentrations, and discovery of a predicted promoter upstream of *arcD* suggested that it may be transcribed from both the *arcA* promoter, and its own promoter (Jakubovics *et al.*, 2015). Additionally, ArcD was found to play a role in recruiting *Fusobacterium nucleatum* to the plaque biofilm, due to its function in transporting the amino acid ornithine out of *S. gordonii* cells (Sakanaka *et al.*, 2015). *F. nucleatum* cells utilise ornithine as an energy source, and deletion of the *arcD* gene of *S. gordonii* within a mixed-species biofilm model was found not only to reduce the levels of *F. nucleatum* within the biofilm, but also to increase sensitivity of *S. gordonii* cells to low pH. This is likely due to the activity of ArcD in transporting L-arginine inside the cells for use by the ADS, and subsequent ammonia production contributing to acid-tolerance within the biofilm.

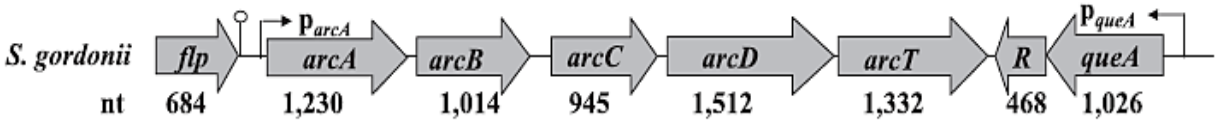


Figure 1.3. Structure of the *arc* operon of *S. gordonii*. The *arcR* gene, labelled "R", is shown at the 3' end of the *arcT* gene, and co-transcribed with *queA* under the control of the p_{queA} promoter. The *flp* gene, which also encodes an activator of *arc* operon expression, is shown upstream of the *arcA* gene, with a predicted terminator (circle and line) and promoter (p_{arcA}) shown also. Numbers represent the gene length in nucleotides (nt). Diagram taken from Liu *et al.* (2008).

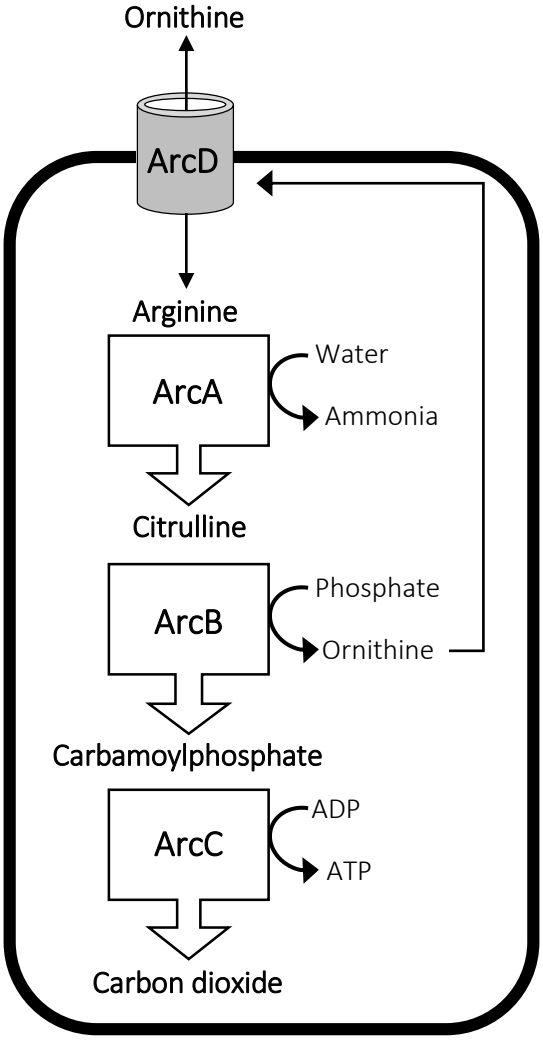


Figure 1.4. Conversion of L-arginine to ammonia, carbon dioxide and ATP by the *arc* operon of *Streptococcus gordonii*. The ArcD antiporter transports ornithine out and arginine into the *S. gordonii* cells. The arginine is taken up by the arginine deiminase (ArcA) protein, and converted into citrulline, with ammonia as a by-product. The citrulline is converted to carbamoylphosphate and ornithine by the ornithine carbamoyltransferase (ArcB) enzyme. The carbamoylphosphate can either be used in the biosynthesis of arginine by the Arg genes, or be converted to carbon dioxide, with the concurrent phosphorylation of ADP to ATP by the carbamate kinase (ArcC) gene. Excess ornithine from ArcB activity is either used alongside carbamoylphosphate in arginine biosynthesis, or transported out of the cell by the ArcD antiporter. Figure adapted from Cusumano *et al.* (2014).

Finally, the gene *arcR* encodes the ArcR transcriptional activator, which controls the expression of the ADS in response to L-arginine concentration (Dong *et al.*, 2002), and is co-transcribed alongside another gene, *queA*, encoding an *S*-adenosylmethionine:tRNA ribosyltransferase-isomerase enzyme (Liu *et al.*, 2008).

Interestingly, the study by Liu *et al.* (2008) also suggested that *S. gordonii* appeared to be unique in its linkage of the *queA* and *arcR* genes, as subsequent sequence analysis of other bacterial genomes demonstrated the presence of both genes separately, but not linked together or co-transcribed. The *queA* gene is involved in incorporating the nucleoside queosine into tRNA molecules, particularly those with an anticodon specific to the amino acids aspartic acid, asparagine, tyrosine and histidine (Slany and Kersten, 1994; Dineshkumar *et al.*, 2002). It has also been linked previously to virulence in *Shigella flexneri* (Durand *et al.*, 1994), and stationary growth phase survival in *E. coli* K12 (Noguchi *et al.*, 1982), although neither of these associations have been found with the QueA protein in *S. gordonii*.

In addition to the ArcR regulation of the *arc* operon, expression of the ADS is also regulated by other mechanisms. Dong *et al.* (2002) demonstrated that the ADS pathway of *S. gordonii* is subject to carbon catabolite repression (CCR) by the CcpA protein, which caused repression of arginine deiminase activity in the presence of glucose, but not galactose. Additionally, the same group showed induction of the ADS by another protein, Flp (shown immediately upstream of the *arcA* gene in Figure 1.3), under aerobic conditions (Dong *et al.*, 2004).

1.3.4. Additional roles of the arc *genes*

The *arcA* gene itself may also act as a surface protein in oral streptococci species. In *S. gordonii*, *S. intermedius* and *S. cristatus*, ArcA has been shown to inhibit biofilm formation by the gingival pathogen *Porphyromonas gingivalis* (Christopher *et al.*, 2010; Wu and Xie, 2010). In the case of *S. gordonii*, this was contradictory to previous evidence that suggested *S. gordonii* ArcA actually recruited *P. gingivalis* to the biofilm (Lin *et al.*, 2008). However, this apparent recruitment of *P. gingivalis* by *S. gordonii* may be due to *S. gordonii* producing lower levels of ArcA on its surface, which may not have been high enough to prevent the *P. gingivalis* cells from binding to *S. gordonii* biofilms (Wu and Xie, 2010). In all three of these streptococcal species, ArcA has been

shown to signal to *P. gingivalis* to down-regulate expression of its surface fimbrial subunit protein, FimA, and furthermore it was shown that addition of purified ArcA from *S. intermedius* to the supernatant of *P. gingivalis* cells, caused active down-regulation of the expression of fimbriae by *P. gingivalis* itself (Christopher *et al.*, 2010).

Other studies of cell wall-associated and extracellular proteins in *Streptococcus pyogenes* demonstrated that all three enzymes of the ADS pathway, arginine deiminase (ArcA), ornithine carbamoyltransferase (ArcB) and carbamate kinase (ArcC), are wall-associated in GAS (Cole *et al.*, 2005). In fact, arginine deiminase may even be released from the cell surface (Lei *et al.*, 2000), possibly contributing to pathogenicity in this species. In *S. pyogenes* strains, deletion of the arginine deiminase gene, known as streptococcal acid glycoprotein in this species, made the cells less able to invade human epithelial cells (Degnan *et al.*, 2000; Marouni *et al.*, 2003).

S. gordonii also has the ability to biosynthesise L-arginine, and therefore it is important for it to tightly regulate biosynthetic and catabolic gene expression in order to avoid a futile cycle of producing and degrading L-arginine. This coordinated regulation is achieved by three homologous arginine-dependent regulators, ArcR, ArgR and AhrC. Whilst ArcR plays a role in regulating the expression of mainly the catabolism genes, the primary role of ArgR and AhrC appears to be in regulating L-arginine biosynthesis, though they also seem to regulate some genes involved in catabolism (such as *arcDT*) which indicates that their regulation overlaps slightly with that of ArcR (Jakubovics *et al.*, 2015).

1.3.5. ArgR and AhrC regulation of arginine biosynthesis

The ArgR and AhrC proteins of *S. gordonii* were identified by comparison of the *S. gordonii* DL1 (wild-type strain) genome with the sequences of known L-arginine-dependent regulators in other bacterial species. ArgR was identified through sequence homology to the ArgR protein of *E. coli*, and the AhrC strain through homology to the *Bacillus subtilis* AhrC protein (Jakubovics *et al.*, 2015). ArcR was identified earlier than this, by homology to both *B. subtilis* AhrC, and the ArgR proteins of *E. coli* and *Bacillus licheniformis* (Dong *et al.*, 2002; Zeng *et al.*, 2006). Whilst ArcR is thought to work independently of the other two regulators, ArgR and AhrC are hypothesised by one model to form a protein complex, similar to a proposed model of

ArgR/AhrC coordination in *Lactococcus lactis* (discussed in the next section), in order to bind and regulate the arginine biosynthesis genes (Jakubovics *et al.*, 2015).

Further analysis of the role of the ArgR and AhrC regulators, by qRT-PCR analysis of gene expression in isogenic mutant strains of *S. gordonii*, revealed that these proteins transcriptionally repress the L-arginine biosynthesis and transport genes of *S. gordonii* (Jakubovics *et al.*, 2015). This includes the genes *argC* (encoding an *n*-acetyl-gamma-glutamyl-phosphate reductase enzyme) and *argG* (argininosuccinate synthase), both involved in L-arginine biosynthesis; and $pyrA_b$ (also known as *carB*, carbamoyl phosphate synthase), which is involved in both L-arginine and pyrimidine nucleotide biosynthesis. ArcD was also shown to be regulated in response to L-arginine concentration, and plays a role in transporting L-arginine into the cell. ArcB was shown to have high sequence homology to an anabolic ornithine carbamoyltransferase enzyme called ArgF, and deletion of the *arcB* gene resulted in a strain of *S. gordonii* unable to grow under anaerobic conditions without exogenous L-arginine (Jakubovics *et al.*, 2008), suggesting that it plays a role in both arginine catabolism and anabolism. This would mean that during arginine biosynthesis, it had the ability to act opposite to its role in arginine catabolism by catalysing the conversion of carbamoylphosphate to citrulline (Jakubovics *et al.*, 2015). In all, the *S. gordonii* L-arginine biosynthesis genes are considered to consist of *argCJBD*, *argGH*, $pyrA_a$ and $pyrA_b$ (also known as *carA* and *carB*), and *arcB*. The catabolism genes consist only of *arcABC* of the ADS, and *arcDT*, although the latter have not been proven experimentally.

All genes involved in arginine biosynthesis and transport, are down-regulated in response to high exogenous L-arginine concentrations by the ArgR and AhrC regulators. ArcR was also shown to regulate arginine biosynthesis genes (Jakubovics *et al.*, 2015), as disruption of the *arcR* gene resulted in increased levels of *argG* expression and a lack of up-regulation of *arcABC* under high arginine. ArgR and AhrC were not shown to affect expression of the *arc* catabolism genes.

Aside from control of L-arginine metabolism genes by regulatory proteins such as ArcR, ArgR and AhrC, coaggregation between *S. gordonii* and *Actinomyces oris* has also been shown to affect expression of these genes. Coaggregation between these two species caused an up-

regulation of expression of different genes involved in L-arginine biosynthesis and transport in *S. gordonii*. Furthermore, coaggregation with *A. oris* allowed *S. gordonii* to grow in low exogenous L-arginine concentrations under aerobic conditions (Jakubovics *et al.*, 2008).

1.4. L-arginine regulation in other bacterial species

As mentioned previously, the L-arginine regulators of *S. gordonii* were discovered through homology to proteins in other species – specifically, *B. subtilis* AhrC*,* and *E. coli* and *B. licheniformis* ArgR (Dong *et al.*, 2002; Zeng *et al.*, 2006; Jakubovics *et al.*, 2015).

In *B. subtilis*, AhrC is the main regulator of both L-arginine biosynthesis and catabolism. The L-arginine biosynthetic operons of *B. subtilis* consist of *argCJBD*, *carAB* and another gene, *argF*, which converts ornithine to citrulline for use in L-arginine biosynthesis, a function predicted to be performed by *arcB* in *S. gordonii* (North *et al.*, 1989; Gardan *et al.*, 1997; Jakubovics *et al.*, 2008). These genes are all repressed by AhrC under high L-arginine conditions, whilst the AhrC protein simultaneously up-regulates expression of L-arginine catabolism genes *rocABC* and *rocDEF*, which are also positively regulated by another protein called RocR (Gardan *et al.*, 1997). Both RocR and AhrC up-regulate these catabolism genes under high exogenous L-arginine (or citrulline, proline or ornithine) conditions, through direct binding of AhrC to RocR, and RocR to the promoter upstream of *rocA*. However, the *rocABC* and *rocDEF* genes encode an arginase pathway, catabolising arginine to glutamate, rather than an arginine deiminase pathway such as those found in *S. gordonii*, or *B. licheniformis*.

The environmental bacterium *B. licheniformis* expresses its own arginine deiminase system, controlled by the regulator protein ArgR. This protein acts as both an activator of the arginine deiminase pathway, and a repressor of L-arginine biosynthesis (Maghnouj *et al.*, 1998). In this species, two copies of the ornithine carbamoyltransferase enzyme are present – one anabolic and one catabolic (Broman *et al.*, 1975). Ornithine carbamoyltransferase enzymes have the ability to catalyse a reaction to form citrulline and phosphate from ornithine and carbamoylphosphate, or vice versa. In *B. licheniformis,* the catabolic copy is present within the arginine deiminase system, playing a similar role to the *arcB* gene in *S. gordonii*; whereas the

anabolic copy is involved in biosynthesising L-arginine from ornithine and carbamoylphosphate. Both are controlled in their expression by the same ArgR protein.

In some species, particularly within the *Lactobacillales* order of bacteria, there are two or more L-arginine regulators present, similar to the systems operating in *S. gordonii*. For example, *Lactococcus lactis* contains two regulators, ArgR and AhrC, which are thought to either work together as a complex, or one protein regulates the other, in order to control expression of L-arginine biosynthesis and catabolism genes (Larsen *et al.*, 2004). AhrC was later shown by Larsen *et al.* (2005) to aid the ArgR regulator in binding the regulatory ARG box sequences upstream of the L-arginine biosynthesis operons, to repress their expression under high exogenous L-arginine conditions; and to additionally prevent ArgR from binding to the ARC box sequences upstream of the L-arginine catabolism genes, which would prevent activation of their expression under high L-arginine. A theory was proposed whereby they were able to do this by existing as homohexameric protein complexes of ArgR and AhrC separately, in the absence of L-arginine, and that the presence of L-arginine caused the proteins to form a heterohexameric subunit consisting of one ArgR and one AhrC trimer (Larsen *et al.*, 2005). This would in turn allow them to regulate both biosynthetic and catabolic L-arginine operons, by way of binding distinct ARG and ARC operator binding sites respectively.

1.4.1. *Regulatory binding sites*

Regulation of gene expression by interaction of a regulatory protein with an operator binding site upstream of the target gene is a well-known mechanism, with those sites often highly conserved between bacterial species due to families of regulator orthologs. However, a study by Makarova *et al.* (2001) indicated that whilst these types of L-arginine-specific ARG operator binding site are universally found throughout the genomes of many different bacterial species, they are poorly conserved. Despite strong conservation of the amino acid sequences of the actual L-arginine regulator orthologs in different species, the ARG box recognition sites for the same proteins were weakly conserved across different bacterial species. This is possibly due to the many versions of these regulators and L-arginine catabolism pathways found in different bacterial lineages, leading to divergence in mechanisms of binding and therefore the recognition sites themselves.

Additionally, different binding sites have been reported for different regulators even within the same bacterial species. For example, the aforementioned ARG and ARC boxes in *L. lactis*, direct the ArgR regulator to the promoters of either L-arginine biosynthetic genes (ARG boxes) or L-arginine catabolism genes (ARC boxes) dependent on whether or not arginine is bound to the regulators (Larsen *et al.*, 2005). Another example of this divergence in binding sequence within the same species is found within *S. gordonii* itself. In this case, the ArgR and AhrC regulators bind ARG box sequences upstream of the L-arginine biosynthetic genes, but ArcR binds a different recognition sequence upstream of the *arcABC* catabolism genes (Zeng *et al.*, 2006).

1.4.2. *Arginine regulation and virulence*

The presence of three or more L-arginine regulator proteins in one bacterial species, such as is found in *S. gordonii*, is not uncommon amongst the *Lactobacillales* order. Another species with a similar set of three regulators such as this is the pathogenic Gram-positive species, *Streptococcus pneumoniae*. *S. pneumoniae* encodes three separate L-arginine-dependent regulatory proteins, ArgR1, ArgR2 and AhrC, which between them control L-arginine catabolism and transport within its cells. *S. pneumoniae* is auxotrophic for arginine due to an inability to synthesise glutamate, and is therefore unable to biosynthesise it, relying upon exogenous arginine for growth (Hoskins *et al.*, 2001; Härtel *et al.*, 2012). The protein ArgR2 is orthologous to the ArcR protein of *S. gordonii*, controlling expression of the ADS and *arcD* genes in *S. pneumoniae* (Schulz *et al.*, 2014).

However, whilst the *S. pneumoniae* ArgR1 and AhrC proteins have been shown to regulate the *argGH* L-arginine biosynthesis genes, and the transport genes *artPQ, abpAB* and *aapA,* as expected, they also were found to regulate genes involved in bacterial virulence, including genes for L-arginine uptake from the environment and nasopharyngeal colonisation (Kloosterman and Kuipers, 2011). The cited study also proposed that the ArgR1 and AhrC proteins of *S. pneumoniae* may act in a similar way to the suggested mechanism in *L. lactis*, namely by forming a heterohexameric protein complex to regulate expression of these different genes (Figure 1.5).

This linkage of virulence with L-arginine metabolism in *S. pneumoniae* is not the first time that exogenous or endogenous L-arginine concentrations have been linked to something other than metabolism or growth in bacterial cells.

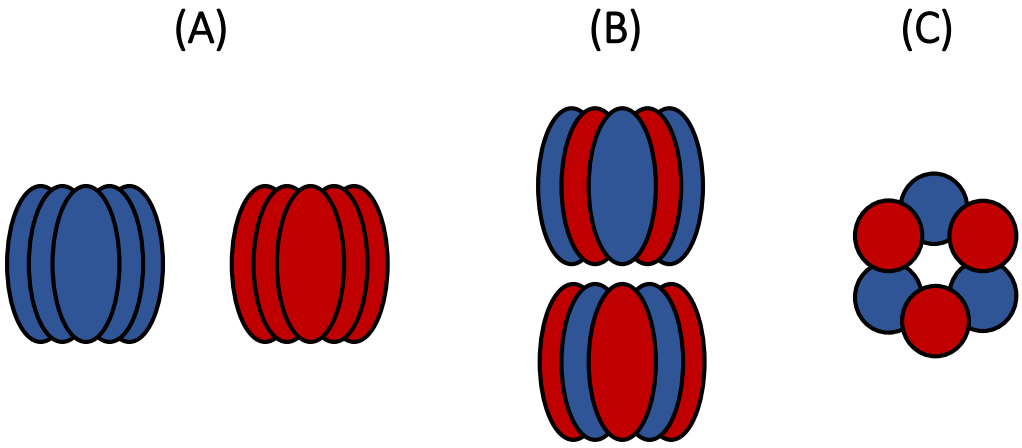


Figure 1.5. Visual representation of the homohexameric and heterohexameric protein complexes that may regulate arginine metabolism. (A) represents the homohexameric complexes of the two separate arginine regulator proteins, for example, ArgR and AhrC of *Lactococcus lactis* or *Streptococcus gordonii*, or ArgR1 and AhrC of *Streptococcus pneumoniae*. (B) represents the heterohexameric arrangements of the trimers of these two regulatory proteins. (C) shows the alternating subunit arrangement of the proteins from above.

1.5. Roles of L-arginine in other bacterial species

Even within *S. pneumoniae*, the other L-arginine-dependent regulator, ArgR2, is linked to yet another differential phenotype. An isogenic *S. pneumoniae* TIGR4 mutant of this regulator showed an increase in fitness in comparison to the wild-type strain in a nasopharyngeal infection model (Schulz *et al.*, 2014), and indicated that in this species, ArgR2 may have further regulatory effects on genes involved in fitness and virulence, which were subsequently up-regulated to compensate for the loss of ADS and ArcD expression that came with deletion of the *argR2* gene.

It is not only in *S. pneumoniae* that L-arginine metabolism or regulation has been linked to virulence. Random transposon mutagenesis of the opportunistic pathogen *Enterococcus faecalis* demonstrated a number of genes that impacted upon biofilm formation in this species (Kristich *et al.*, 2008), including *argR*, *ahrC*, and the pyrimidine/L-arginine biosynthesis gene *pyrC*. Further analysis of these genes revealed that interruption of the *ahrC* gene not only affected biofilm formation, but also the virulence of the *E. faecalis* strain in an endocarditis infection model (Frank *et al.*, 2013). In contrast, an *argR* mutant was not affected in endocarditis virulence. Investigation into the role of AhrC in *E. faecalis* infections demonstrated that an *ahrC* transposon mutant strain of *E. faecalis* OG1RF was impaired in both initial attachment of cells to a surface, and further accumulation of cells to the developing biofilm, although its specific role in virulence was not clear (Frank *et al.*, 2013). It may be that, as the endocarditis model used to assess these strains demonstrated a biofilm-associated infection, the defect in initial biofilm formation translated into a decrease in virulence in this *E. faecalis* strain.

Within the broadly pathogenic species *Streptococcus pyogenes*, the arginine deiminase system and its activity in catabolising exogenous L-arginine have been linked to more than one virulence phenotype. The streptococcal acid glycoprotein (SAGP) of *S. pyogenes*, found in cell extracts from this bacterium, was able to inhibit the proliferation of human peripheral blood mononucleocytes in proliferation assays (Degnan *et al.*, 1998). This protein was also found to have arginine deiminase activity, and it was theorised that its ability to prevent proliferation of host immune cells may be due to the activity of this enzyme in the catabolism of exogenous L-arginine. Depletion of L-arginine from the growth medium by SAGP may prevent the

mononuclear cells from growing and proliferating, or may be related to their inability to synthesise nitric oxide in the absence of L-arginine, which is a necessary requirement for DNA synthesis within these cells (Efron *et al.*, 1991). Irrespective of the mechanism by which SAGP prevents mononuclear cell proliferation, this activity may down-regulate the host immune response to *S. pyogenes* infection, thereby increasing the pathogenicity of this species.

Another instance of the ADS of *S. pyogenes* affecting virulence and fitness in the cells, is the apparent necessity of the ArcA and ArcB enzymes for survival of *S. pyogenes* within host tissues. Cusumano *et al.* (2014) showed that deletion of the *arcA* gene affected *S. pyogenes* colonisation of soft tissues and mucosa, and *arcB* (and *arcC*) deletion caused poor colonisation of cutaneous tissue. It was determined that, whilst the ability to metabolise L-arginine was linked to mucosal invasion in *S. pyogenes,* the ability to metabolise both L-arginine and L-citrulline was important in allowing colonisation of soft tissue and manipulation of the immune system in a similar way to that mentioned above for *E. faecalis* AhrC. Additionally, catabolism of L-arginine and citrulline by the ADS system confers increased acid tolerance on *S. pyogenes* cells, which was deemed important for bacterial colonisation of human soft tissues, as high bacterial densities would cause a build-up of acid within the colonised tissues (Cusumano *et al.*, 2014).

In the periodontal pathogen *F. nucleatum*, the adhesive functions of the outer membrane protein RadD are inhibited by L-arginine. This protein mediates the coaggregation of *F. nucleatum* with Gram-positive early coloniser species of the dental plaque biofilm, such as *A. oris*, *S. sanguinis* and *C. albicans* (Kolenbrander *et al.*, 1989; Kaplan *et al.*, 2009; Wu *et al.*, 2015). However, RadD has also been shown to cause *F. nucleatum* to induce human lymphocytic cell death *in vitro*, in conjunction with another surface protein, Fap2, via a contact-based mechanism (Kaplan *et al.*, 2010). This helps *F. nucleatum* to suppress the immune system *in vivo*, and endure within the plaque biofilm. In addition, RadD-mediated binding of *F. nucleatum* to *S. sanguinis* has also been shown to allow *F. nucleatum* cells to better survive exposure to hydrogen peroxide in a mixed species microbiota model (He *et al.*, 2012), showing that this surface protein, whilst inhibitable by L-arginine, is also able to aid *F. nucleatum* survival against killing by both the host immune system and other bacterial by-products within the plaque biofilm.

Not all roles that L-arginine plays in other bacterial species are related to virulence or survival, although many have been shown to be linked to biofilm formation, including the examples above. The *E. faecalis* AhrC regulator affected virulence in a biofilm-associated infection model; *S. pyogenes* infection of soft tissues within a human host, aided by the metabolism of L-arginine, is often caused by a biofilm community; and *F. nucleatum* binding to early colonisers within the plaque biofilm was mediated and inhibited by L-arginine.

One further example of a link between L-arginine transport and biofilm formation and pathogenesis, is in the human pathogen *Staphylococcus aureus*. Here, the ArcD protein, which encodes an arginine-ornithine antiporter, has been shown to be important in biofilm formation due to its ability to enhance accumulation of the polysaccharide intercellular adhesin (PIA) (Zhu *et al.*, 2007), which is necessary for formation of a multi-layered biofilm by *S. aureus* (Cramton *et al.*, 1999).

In a microfluidics system model of *S. gordonii* biofilms, different concentrations of L-arginine were found to affect biofilm formation in different ways (Jakubovics *et al.*, 2015). For example, in saliva, the addition of low concentrations of L-arginine (0.5-500 µM) enhanced biofilm formation when compared to saliva without added exogenous arginine, whereas adding high concentrations (5-500 mM) reduced the levels of biofilm formation. This observation forms the basis for part of the work in the first results chapter (Chapter 4) of this thesis.

1.6. Thesis outline and programme of work

As mentioned above, exogenous L-arginine concentration was shown to have varying effects on the biofilm formation ability of *S. gordonii* DL1. However, the impact of L-arginine concentration on planktonic growth of *S. gordonii* has never been directly assessed, and so forms part of the experimental work in Chapter 4. Leading on from this work, an assessment of gene regulation in response to arginine depletion in *S. gordonii* planktonic cells was made, and then further investigated to elucidate which of these gene expression responses were specific to L-arginine depletion, and which appeared to be part of a general stress response.

Chapter 5 focuses on looking at the role the three L-arginine-dependent regulators of *S. gordonii*, ArcR, ArgR and AhrC, play within the *S. gordonii* cells. This will particularly focus on the regulation of different genes throughout the chromosome, as measured by genome-wide microarray analysis in response to these three regulatory proteins, and aiming to build a picture of the specific regulons of each of these proteins.

Chapter 6 concentrates on the ArcR regulator, and in particular, an *S. gordonii* isogenic mutant strain of the *arcR* gene, which was found to have a biofilm-defective phenotype. The *S. gordonii argR* and *ahrC* mutant strains were not shown to have a similar attenuation in biofilm formation, and analysis of the ArcR-specific microarray experiment showed the up-regulation of one particular uncharacterised gene, SGO_0846, within the Δ*arcR* mutant strain. Therefore, investigations were made into whether SGO_0846 was responsible for the *S. gordonii arcR* mutant biofilm defect, what role it may be playing within *S. gordonii* cells, and whether complementation of an intact copy of the *arcR* gene would reverse the biofilm attenuation observed in the *S. gordonii* Δ*arcR* strain.

1.7. Thesis aims and objectives

In all, the primary aim of this work is to investigate the role that L-arginine plays in the expression of different genes, and the biofilm formation of *S. gordonii*. Furthermore, it also aims to examine the role of the ArcR regulator in mediating these effects. The objectives of this work are as follows:

1. Does L-arginine affect *S. gordonii* planktonic growth?
    - Planktonic growth in *S. gordonii* will be assessed for differences in response to varying concentrations of L-arginine.

2. Are gene expression responses to arginine depletion L-arginine-specific, or part of a general stress response?
    - Genes of interest will be analysed for changes in expression in response to arginine depletion and depletion of other amino acids, in order to determine L-arginine-specific and general depletion-based stress responses.

3. What roles do the three L-arginine-dependent regulator proteins (ArcR, ArgR and AhrC) play in gene regulation in *S. gordonii?*

- Isogenic mutant strains of the three arginine regulators will be assessed by genome-wide microarray, in order to determine the regulons for each arginine-dependent regulator.

4. What causes the defective biofilm phenotype of the *S. gordonii* Δ*arcR* strain?

- Complementation of the *S. gordonii* Δ*arcR* strain, and subsequent analysis of biofilm formation levels of this and the deletion strain, will determine whether ArcR is responsible for the biofilm defective phenotype seen in that mutant.

5. What roles are ArcR and SGO_0846 playing within *S. gordonii* biofilms?

Deletion of the SGO_0846 gene, and analysis of the biofilm formation levels of that strain alongside the *S. gordonii arcR* strains will determine whether this protein is responsible for the Δ*arcR* biofilm defect, in addition to characterisation of the SGO_0846 protein by *in silico* analysis

.

2. Materials and Methods

2.1. Bacterial strains and culture techniques

2.1.1. Routine culture of microorganisms

Streptococcus gordonii strains (Table 2.1) were routinely cultured for 18 h overnight in THYE (30 g/L Bacto™ Todd Hewitt Broth [Becton Dickinson and Co.], 5 g/L yeast extract [Melford Laboratories Ltd.]) or BHY (37 g/L Brain Heart Infusion [Melford Laboratories Ltd.], 5 g/L yeast extract) medium, at 37°C anaerobically (80% N_2, 10% H_2, 10% CO_2 [Bugbox Plus, Ruskin]).
S. gordonii cells were alternatively cultured on BHY or THYE medium solidified by the inclusion of 1.5% w/v agar prior to autoclaving. Inoculated plates were grown at 37°C for 48 h in a candle jar.
Chemically-defined FMC medium was also used as growth medium for both planktonic and biofilm cultures, made to a modified recipe outlined in Jakubovics *et al.* (2008). On occasion, L-arginine, L-histidine, or the branched-chain amino acids (L-leucine, L-isoleucine, L-valine) were omitted from FMC, or were added at different concentrations. Where necessary, media for *S. gordonii* were supplemented with antibiotics at the following concentrations: spectinomycin 250 µg/mL, erythromycin 2 µg/mL, kanamycin 250 µg/mL.

Escherichia coli strains, used for cloning experiments, were grown in Luria Bertani (LB) broth (Melford Laboratories Ltd.), or on LB solid medium containing 1.5% w/v agar (added prior to autoclaving), at 37°C in aerobic conditions with shaking at 180 rpm. Media were supplemented with antibiotics if necessary, at the following concentrations: erythromycin 400 µg/mL, kanamycin 25 µg/mL, ampicillin 100 µg/mL.

2.1.2. Bacterial frozen glycerol stocks

Bacterial planktonic cultures were grown for 18 h overnight, at 37°C, in liquid broth (LB for *E. coli* stocks, THYE or BHY for *S. gordonii* stocks). Cells were harvested at 3800 x g for 10 min (Sigma Laborzentrifugen model 3K10), and resuspended in 1 mL fresh medium containing sterile 50% v/v glycerol (VWR), in 1.5 mL sterile screw-cap tubes. Cells were then frozen at - 80°C for use as inoculum in other experiments.

Strain name	Genotype	Source/Reference
Streptococcus gordonii DL1 (Challis)	WT* strain	H. Jenkinson, Bristol University, UK
Streptococcus gordonii PK3346	*argR*::*aphA3* (Kan^R)	N. Jakubovics, Newcastle University, UK
Streptococcus gordonii PK3347	*argR*::*ermAM* (Erm^R)	N. Jakubovics, Newcastle University, UK
Streptococcus gordonii PK3348	*argR*::*aphA3* / *arcR*::*ermAM* (Kan^R, Erm^R)	N. Jakubovics, Newcastle University, UK
Streptococcus gordonii PK3349	*argR*::*aphA3* / *ahrC*::*ermAM* (Kan^R, Erm^R)	N. Jakubovics, Newcastle University, UK
Streptococcus gordonii PK3350	*ahrC*::*ermAM* (Erm^R)	N. Jakubovics, Newcastle University, UK
Streptococcus gordonii PK3351	*arcR*::*ermAM* (Erm^R)	N. Jakubovics, Newcastle University, UK
Streptococcus gordonii PK3354	*arcR*::*aad9* ($Spec^R$)	N. Jakubovics, Newcastle University, UK
Streptococcus gordonii PK3355	*arcR*::*aad9* / *argR*::*aphA3* ($Spec^R$, Kan^R)	N. Jakubovics, Newcastle University, UK
Streptococcus gordonii PK3356	*arcR*::*aad9* / *ahrC*::*ermAM* ($Spec^R$, Erm^R)	N. Jakubovics, Newcastle University, UK
Streptococcus gordonii PK3357	*arcR*::*aad9* / *argR*::*aphA3* / *ahrC*::*ermAM* ($Spec^R$, Kan^R, Erm^R)	N. Jakubovics, Newcastle University, UK
Streptococcus gordonii NU05	SGO_0846::*ermAM* (Erm^R)	This thesis
Streptococcus gordonii NU06	SGO_0846::*aphA3* (Kan^R)	This thesis

Streptococcus gordonii NU07	*arcR*::*ermAM* / SGO_0846::*aphA3* (Erm^R, Kan^R)	This thesis
Streptococcus gordonii NU08	*arcR*::*aad9* / SGO_0846::*ermAM* ($Spec^R$, Erm^R)	This thesis
Streptococcus gordonii NU09	*arcR*::*aad9* / SGO_0846::*aphA3* ($Spec^R$, Kan^R)	This thesis
Streptococcus gordonii NU10	PK3354 (p$arcR_{comp}$) (Erm^R)	This thesis
Streptococcus gordonii NU11	NU09 (p$arcR_{comp}$) (Erm^R)	This thesis
Escherichia coli TOP10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen
Escherichia coli HST08	F-, endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Φ80d lacZΔ M15, Δ (lacZYA - argF) U169, Δ (mrr - hsdRMS - mcrBC), ΔmcrA, λ–	Clontech
Escherichia coli JM83	F– ara Δ(lac-proAB) rpsL (Strr)[ϕ80 dlacΔ(lacZ)M15] thi (pVA838)	Invitrogen

*WT = wild-type

Table 2.1. List of bacterial strains used in this project.

2.1.3. Saliva sample collection

Whole stimulated saliva was collected from ≥5 healthy individuals, not on courses of antibiotic treatment, who had not eaten for 2 h prior to donation. Saliva production was stimulated by chewing unflavoured chewing gum base. Saliva was kept on ice, dithiothreitol (DTT) was added to a final concentration of 2.5 mM, and the mixture was then centrifuged at 15,000 x g for 30 min, 4°C (using a Beckman model J2-21 centrifuge). The supernatant was collected, diluted 1:2 or 1:4 in sterile dH_2O, and filter sterilised through a 0.2 µm membrane (Acrodisc Syringe Filters, Pall Life Sciences). This was then aliquoted and stored at -20°C, for use as growth medium.

2.1.4. Growth curves and spectrophotometric measurements

Bacterial cells were cultured for 18 h in sterile medium, harvested at 3800 x g (using Sigma Laborzentrifugen model 3K10) and resuspended in fresh medium. These cells were then used as an inoculum for fresh growth medium, in order to follow the growth rate and yield of the bacterial strain. Sterile medium was inoculated with a 1:20 dilution of bacterial overnight culture. Growth was quantified by taking optical density measurements of the bacterial cultures until cells reached stationary growth phase. Optical density of planktonic cultures was measured using plastic screw-cap tubes (16 x 125 mm Culture Tubes, Corning), in a Biochrom Libra S11 spectrophotometer at 600 nm wavelength.

2.1.5. Phase contrast microscopy

Microscopy was performed on planktonic bacterial cultures, or smeared bacterial colonies taken from agar plates, in order to check cell morphology and ensure no contamination was present in the samples. Light microscopy was carried out on a Zeiss Universal microscope, using a Zeiss Neofluar 25X (25/0.60 Ph2 160/0.17) objective.

2.2. Biofilm growth and quantification

2.2.1. Routine biofilm culture of microorganisms

S. gordonii cells were cultured anaerobically in 6-well or 96-well microtitre plates (MTPs) (CELLSTAR, Greiner Bio-One), with 50% BHY (50% v/v in sterile dH_2O), or 100% FMC medium, to encourage biofilm formation.

Alternatively, biofilms grown for quantification assays shown in Figure 5.1 were cultured aerobically in 96-well MTPs, with 100% TYEG medium (10 g/L Bacto™ Tryptone [Becton Dickinson and Co.], 5 g/L yeast extract, 3 g/L dipotassium phosphate [VWR], 2 g/L D-glucose [Melford Laboratories Ltd.]; pH 7.5), to encourage biofilm formation.

MTPs were placed in a humid, anaerobic (or aerobic) environment at 37°C for 24 h, to allow biofilms to form on the plate surface. Biofilms were then quantified downstream using crystal violet assays.

2.2.2. Crystal violet biofilm assay

S. gordonii cells were cultured to form biofilms as described in section 2.2.1 above. Levels of biofilm formation were quantified by the crystal violet method (Christensen *et al.*, 1985; Stepanović *et al.*, 2000). Blank wells were included as a control, containing only sterile medium with no cells. Prior to staining, the optical density of the wells was measured at 600 nm (Synergy™ HT Microplate Reader, BioTek) in order to estimate the total bacterial growth for each well, then the planktonic cell phase was removed from each well. Biofilms were stained with 1 volume 0.1% crystal violet dye, and then washed three times in 0.5 volumes PBS to remove loosely-bound cells and unbound dye. Following each wash, the plate was tapped to remove any loosely-bound cells or excess liquid. The crystal violet stain on the cells was then dissolved in 1 volume 7% acetic acid, and the OD_{570} taken, using the Synergy microplate reader, as a measure of biofilm formation. 570 nm was used as the best wavelength for measuring crystal violet dye. For each well, the average optical density for the blank wells was subtracted as background before quantifying levels of biofilm formation.

2.2.3. Statistical analysis of biofilm quantification assays

Differences in the levels of biofilm formation between different *S. gordonii* strains were assessed for statistical significance at a 95% confidence level, using one-way ANOVAs with Tukey's pairwise comparisons (assuming equal variances) and two-sample T-tests, performed with the Minitab 17 software (Minitab, Inc.).

2.3. Bacterial cloning techniques

2.3.1. Agarose gel electrophoresis

Gel electrophoresis was carried out on DNA products from PCR, qRT-PCR or restriction digest reactions. qRT-PCR products were analysed using 3% agarose gels, whereas standard PCR and digest products were analysed on a gel containing between 0.8-3% agarose (Melford Laboratories), depending on the size of the products being analysed. DNA products were mixed 1:5 with 5x DNA Loading Buffer (Bioline) for loading on the gel, unless already stained with ReddyMix from a previous PCR reaction (see section 2.3.3d).

GelRed Nucleic Acid Gel Stain (10,000x in water; Biotium) was added to the gel in order to visualise DNA. The following DNA molecular weight markers were used for product size reference: HyperLadder 25 bp (25-500 bp; Bioline); HyperLadder 100 bp (100-1013 bp; Bioline), HyperLadder 1 kb Plus (250-12,007 bp; Bioline).
Gels were run at 70-80 V, for 90-120 min, using a Bio-Rad Power Pac 300. Once run, they were visualised on the G:BOX Transilluminator (Syngene) on transilluminator setting, on live capture at 5.51 Mpixel.

2.3.2. DNA extraction and purification

2.3.2a. Genomic DNA extraction from whole cells

Extraction of DNA from *S. gordonii* cells was carried out using a modified version of the MasterPure DNA Purification kit (Epicentre) protocol.

S. gordonii cells were cultured for 18 h in BHY medium at 37°C. Cells were harvested from this culture, and resuspended in 37°C pre-warmed spheroplasting buffer (20 mM Tris-HCl, pH 6.8; 10 mM $MgCl_2$; 26% w/v raffinose.$5H_2O$). Lysozyme (250 µg/mL) and 50 U mutanolysin (reconstituted to 10,000 U/mL) were added to the cells prior to incubation at 37°C for 30 min. Following incubation, cells were placed into screw-cap tubes with 2x T&C Lysis Solution (Epicentre) and acid-washed glass beads, and placed into a bead-lysis machine (Qiagen TissueLyser LT) at 50 Hz for 5 min. Tubes were transferred to ice, and 50 µg/µl Proteinase K (Epicentre) was added.

Cells were incubated at 65°C for 30 min, with brief vortexing every 5 m, and cooled to 37°C. RNase A (5 µg/µl, Epicentre) was added before further incubation at 37°C for 30 min. Samples were placed on ice, and the protocol continued from the DNA precipitation stage of the MasterPure™ Gram Positive DNA Purification Kit (Epicentre) protocol onwards, with the exclusion of the later RNase A step. DNA was eluted into TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). DNA concentrations were determined using a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific), using the DNA-50 setting. Concentration was measured in ng/µl, and purity and quality determined by assessment of the 260/280 and 260/230 readings. Following extraction, DNA was stored at - 20°C for use in downstream applications.

2.3.3. Molecular methods

2.3.3a. Primer design

Primers for use in quantitative reverse transcription PCR (qRT-PCR) were designed using Primer3 software (Rozen and Skaletsky, 2000). qRT-PCR primers were designed to be between 15-25 bp in length, with a T_m of approximately 60°C, and to generate a product between 50-200 bp in length.

Primers for other PCR applications were designed using SnapGene software (GSL Biotech, 2015), and NCBI Primer-BLAST (Ye *et al.*, 2012). All nucleotide sequences for primer design were obtained from NCBI Gene (NCBI, 2011-2012). Primer sequences can be found in Table 2.2 and Table 2.3.

2.3.3b. Plasmid design and mapping

Recombinant plasmids were designed and mapped using the SnapGene software (GSL Biotech, 2015). A description of plasmids used in this project can be found in Table 2.4.

qRT-PCR primers				
Primer name	Primer sequence	Product length	qRT-PCR target gene	Source/Reference
16SSg	F: AGACACGGCCCAGACTCCTAC R: CTCACACCCGTTCTTCTCTTACAA	138 bp	*S. gordonii* 16S rRNA gene	(Jakubovics *et al.*, 2008)
q0175	F: GCAAGGACACACCTTCTATGAACAA R: CCGTCTTGTGGGCAATTTCA	98 bp	SGO_0175 (*argG*) gene	(Jakubovics *et al.*, 2008)
q0698	F: ATGAGGATTTCGTTGCCTTG R: CGAGCCTTATGCTCCTCTTG	112 bp	SGO_0698 (*recN*) gene	This thesis
q0846	F: TTACCCACCAGGAAAACCAG R: ACCTGGATCGTTTGGATCTG	181 bp	SGO_0846 gene	This thesis
q0966	F: GACCCTTCAGGAAATGCAAC R: TCGGCAGGGTCATACTTTTC	107 bp	SGO_0966 (*hsa*) gene	This thesis
q0978	F: CAATTTTTCAGCGACACTCG R: TCTTTTTCCCAAAGTTTCTTTCC	70 bp	SGO_0978 gene	This thesis
q1401	F: ATTACGAGGGCGAGATTGTC R: GGTTGCCATATTGCTGGTTC	144 bp	SGO_1401 gene	This thesis
q1411	F: GCAGTTGGAGGAAATTCTGG R: CAGCATCCTTGGAAAAGGTC	146 bp	SGO_1411 (*hisC*) gene	This thesis
q1576	F: ATTTTGGCGCCTATGACATC R: CCCAAGAAGGCTCCTATTCC	119 bp	SGO_1576 (*bfbC*) gene	This thesis

q1582	F: TATCCGGCTACTTGCAATCC R: GCTCGCTAAAGTCCACCTTG	154 bp	SGO_1582 (*bfbF*) gene	This thesis
q1587	F: GCGACGCCCTTGTCATGAAT R: CGTGACCACCAGTTCCAGGT	78 bp	SGO_1587 (*queA*) gene	This thesis
q1588	F: ATGGAAGACGCCCTCATCAT R: CAGTGCATCCAGAATCGCTC	102 bp	SGO_1588 (*arcR*) gene	This thesis
q1592	F: AGTTTTGGGCCGTATGTTTG R: TCGTCAGTCAAACCATTCCA	111 bp	SGO_1592 (*arcB*) gene	(Jakubovics *et al.*, 2015) [Known as 1447F/R]
q1686	F: AAGATCGCACTCAGCCTTTG R: TATCGCCTAAGCGAAACAGG	105 bp	SGO_1686 gene	This thesis
q1699	F: GAAACTCCCAATGCAACTCC R: CCACACGACGATCAATATCAG	136 bp	SGO_1699 gene	This thesis
q2015	F: TCCTTGACCCTGAGCATTTC R: CCATAAGGACATTCCGCAAC	136 bp	SGO_2015 gene	This thesis
q2028	F: GACCCATGCGGGTATTTATG R: CACCACCCAAAAGGTCAATC	131 bp	SGO_2028 (*wzg*) gene	This thesis
q2056	F: AGCCATCTCAAGAAGGTCCA R: CCTGTACGGGCATTTTCAGT	101 bp	SGO_2056 (*mutS*) gene	This thesis

Table 2.2. List of qRT-PCR primers used throughout this project.

Standard PCR primers				
Primer name	**Primer sequence***	**Product length**	**Use in project**	**Source**
SGO_0846-*Bsu*361	**F1:** TGCACGTAACTCTCGTTACAACAAC	1562 bp	SGO_0846 knock-out	This thesis
	R1: TGCATCTGACCAAGTCAACTCTTTC			
*aphA3-Bsu*361	**F3:** TGCA***CCTGAGG***AAGGAACAGTGAATTGGA	910 bp	SGO_0846 knock-out	This thesis
	R3: TGCA***CCTGAGG***AAGCTTTTTAGACATCTAAATC			
*ermAM-Bsu*361	**F1:** TGCA***CCTGAGG***GGAATTGAAGTTAAATTAAATGCT	816 bp	SGO_0846 knock-out	This thesis
	R1: TGCA***CCTGAGG***GGAATTTACAAAAGCGACTCATAGA			
0846 F1 ovex	TGCATTATCTTGTTCAGTATTAGCTGCAG	869 bp	SGO_0846 knock-out	This thesis
0846 R1 Kan ovex	ctccaattcactgttccttgcTCTGGTAAATCAACACTTGGAGTT			
0846 R1 Erm ovex	gttcatgtaatcactccTCTGGTAAATCAACACTTGGAGTT			
0846 F2 Kan ovex	agatttagatatctaaaaagcttatCAACACCAGGTTTGGTAACAAC	903 bp	SGO_0846 knock-out	This thesis
0846 F2 Erm ovex	acgggaggaaataattcCAACACCAGGTTTGGTAACAAC			
0846 R2 ovex	TGCACACTGTTCGCCCATTTCTTA			
Kan ovex	**F1:** gcaaggaacagtgaattggag	914 bp	SGO_0846 knock-out	This thesis
	R1: ataagctttttagatatctaaatct			

Erm ovex	**F1:** ggagtgattacatgaacAAA	752 bp	SGO_0846 knock-out	This thesis
	R1: gaattatttcctcccgtTAA			
arcR clo	**F1:** CTAG***GCATGC***ATGAATAAAATAGAAAGTAGACATCGT	554 bp	arcR complementation	P. Bateson (unpublished)
	R1: TGCA***GTCGAC***CCGTTTGGCTGGAGAATAAA			This thesis
pPE1010	**F2:** GAGGTGCTCCAGTGGCTTCT	5652 bp	arcR complementation	This thesis
	R2: CGCCTGGGGTAATGACTCTCT			
CP25	**F3:** ccactggagcacctcATGTTGTGTGGAATTGTGAGCG	181 bp	arcR complementation	This thesis
	R2: tattcataacagtacTATTTTATTATACCAGCCCCCT			
arcR comp	**F2:** gtactgttatgaataAAATAGAAAGTAGACATCGTTTAATTCGTTCCCT	494 bp	arcR complementation	This thesis
	R2: tcattaccccaggcgTTATTTACTAAAGAAAAATGGTGGGGCAAATTCTTTAAGCTTAT			

*Regions within primer sequences correspond to: ***RESTRICTION ENZYME SITE***; sequence complementary to another primer.

Table 2.3. List of standard PCR primers used throughout this project.

Plasmid name	Genotype	Source/Reference
p*arcR*$_{comp}$	pPE1010 vector, CP25^{+} *arcR*$^{+}$ (Erm^{R})	This thesis
pPE1010	*ermAM* cassette, *gfp* reporter plasmid (Erm^{R})	(Egland *et al.*, 2004)
pCM18	*ermAM* cassette, CP25 promoter, *gfp*mut3* gene (Erm^{R})	(Hansen *et al.*, 2001)
pSF151	Suicide vector, *aph*A3 cassette (Kan^{R})	(Lin *et al.*, 1992)
pVA838	Shuttle vector, *ermAM* cassette (Erm^{R})	(Macrina *et al.*, 1982)
pCR2.1	*amp* cassette, *kan* cassette, *lacZ*α reporter gene (Amp^{R}, Kan^{R})	Life Technologies TA Cloning Kit

Table 2.4. List of plasmids used in this project.

2.3.3c. Release of DNA from whole cells using GeneReleaser

Colony PCR was carried out on 1 µl whole bacterial cells from 18 h planktonic cultures, using 20 µl resuspended GeneReleaser (BioVentures, Inc.) to break open cells and inactivate PCR inhibitors before performing a normal PCR reaction, using the newly-released DNA as a template. The GeneReleaser reaction was performed to manufacturer's instructions, using the following thermocycler programme: 1) 65°C for 30 s; 2) 8°C for 30 s; 3) 65°C for 90 s; 4) 97°C for 180 s; 5) 8°C for 60 s; 6) 65°C for 180 s; 7) 97°C for 60 s; 8) 65°C for 60 s; 9) 80°C ∞.

Once reactions were held at 80°C, PCR reaction mixture was added, and PCR was performed as detailed below.

2.3.3d. Polymerase chain reaction

All thermocycler reactions were carried out on T100 Thermal Cycler (Bio-Rad) or DNA Engine PTC-200 (MJ Research) thermocycler machines.

Standard polymerase chain reaction (PCR) was performed on DNA samples using the following reaction: 1) 94°C for 2 min; 2) 94°C for 10 s; 3) 55°C for 30 s; 4) 68°C for 60 s; repeat from step 2 35x; 5) 68°C for 7 min; 6) 4°C ∞.

DNA template was used at around 50 ng/µl for *S. gordonii* chromosomal DNA, and around 30 ng/µl concentrations for plasmid DNA. Primers were used at 2.5 µM concentrations, diluted in sterile dH_2O. ThermoPrime 2x ReddyMix PCR Master Mix (Thermo Scientific), containing *Taq* polymerase and 1.5 mM $MgCl_2$, was used as a polymerase enzyme for standard PCR reactions. This also contained an inert red dye, which bound directly to PCR products for downstream analysis using agarose gel electrophoresis.

Step 3) of the PCR protocol described above constituted the annealing step of the PCR programme, the temperature of which depended upon the melting temperature (T_m) of the primers used. T_m was calculated using the primer sequence, by the 2(AT) + 4(GC) method. The annealing temperature of step 3) of the PCR protocol was then determined to be 5°C lower than the primer T_m.

For each PCR reaction, one negative (no template) control was also included for each primer set, for comparison with experimental samples. In some reactions (i.e. checking gene knockout or complementation by PCR), a positive control reaction containing chromosomal or plasmid DNA was also included.

Long-range overlap extension (ovex) PCR was carried out on DNA samples using the following reactions: 1) 95°C for 5 min; 2) 92°C for 10 s; 3) 55°C for 30 s; 4) 68°C for 2 min; repeat from step 2 10x; 5) 92°C for 15 s; 6) 55°C for 30 s; 7) 68°C for 2 min plus 20 s/cycle; repeat from step 5 25x; 8) 68°C for 7 min; 9) 4°C ∞.

DNA templates and primers were used at the same concentrations as detailed for standard PCR above. CloneAmp HiFi PCR Premix (Clontech), containing 1.5 mM $MgCl_2$, was used as a polymerase enzyme for long-range PCR reactions for use in downstream In-Fusion cloning reactions; Expand High Fidelity PCR System (Roche), containing a mix of *Taq* and *Tgo* polymerases (with 1.5 mM $MgCl_2$), was used as a polymerase enzyme for long-range PCR reactions for downstream restriction digests or overlap extension PCR.

As for standard PCR reactions, the annealing temperatures of steps 3) and 6) of the above PCR protocol varied according to the T_m of the primer set used. The extension time of steps 4) and 7) also varied according to the expected length of the amplified product. A negative control was also included in every PCR reaction, as for standard PCR.

2.3.3e. PCR product and plasmid purification

PCR product and plasmid purification reactions were carried out using the QIAquick PCR purification kit and QIAprep Spin Miniprep plasmid kit (QIAGEN), in accordance with the manufacturer's protocols.

2.3.4. Bacterial transformation

2.3.4a. Restriction digest and ligation

Restriction digests were carried out using the relevant buffer and inactivation conditions for each enzyme, with 20 U enzyme (at concentration 20,000 U/mL) used per 10 µl reaction. For

*Sal*I-HF (high-fidelity *Sal*I enzyme) and *Sph*I-HF (New England Biolabs), a double digest reaction required the use of Buffer 4 (New England Biolabs), with an incubation at 37°C for 1 h followed by heat inactivation of the enzyme at 65°C for 20 min. For *Bsu*361 digestion, Cutsmart buffer (New England Biolabs) was required, with a 1 h incubation at 37°C followed by inactivation at 80°C for 20 min. In reactions where two enzymes were used, single digests were run alongside the double digests as controls.

Ligation reactions were performed using 400 U T4 DNA Ligase (400,000 U/mL, New England Biolabs), in a 10 µl reaction, with 1:1 and 3:1 ratio reactions of insert to vector. Reactions were performed to the manufacturer's instructions, with a 10 min incubation at 20°C, followed by heat inactivation of the enzyme at 65°C for 10 min. Ligated products were then used in downstream transformation reactions.

2.3.4b. Ligation-independent cloning

Ligation-independent cloning was carried out using the In-Fusion HD Cloning kit (Clontech), according to manufacturer's instructions. PCR products were designed to contain ends that were overlapping and complementary to both each other and the ends of the vector, using the SnapGene software (GSL Biotech, 2015). The vector was linearised, and the PCR products amplified, using the 2x CloneAmp HiFi PCR Premix (see section 2.3.3d). Both the linearised vector and the PCR products were then incubated with the 5x In-Fusion HD Enzyme Premix (Clontech), and transformed into *E. coli* according to the protocol below (see section 2.3.4c). Following successful *E. coli* transformation, the recombinant vector was extracted and transformed into *S. gordonii* (see section 2.3.4d). Transformant colonies were checked for integrity by DNA sequencing (see section 2.3.4e).

2.3.4c. *Escherichia coli* transformation

E. coli transformations were carried out according to TA Cloning kit (Invitrogen) or In-Fusion HD Cloning kit (Clontech) protocols, with no modifications. Transformed cells were plated onto selective LB agar, containing the antibiotic concentrations detailed in section 2.1.1. Plates were inoculated for 24 h at 37°C, aerobically, and checked for transformants after this time. Transformant colonies were then checked for integrity by DNA sequencing (see section 2.3.4e).

2.3.4d. *Streptococcus gordonii* transformation

S. gordonii cells were cultured in BHY medium for 18 h at 37°C, then sub-cultured into 5 mL pre-warmed BHY broth containing 1% v/v heat-inactivated horse serum (HS) and 10% v/v glucose (BHY/HS/gluc). Cells were grown at 37°C in a candle jar until they reached an OD_{600} of around 0.25-0.35, and then diluted 1:100 into 5 mL fresh pre-warmed BHY/HS/gluc. After a further 60 min incubation at 37°C in a candle jar, cells were aliquoted to 0.8 mL portions and incubated with up to 10 µg DNA for 3-4 h. Following this, 100 µl *S. gordonii* cells were plated onto selective BHY or THYE agar, containing the antibiotic concentrations detailed in section 2.1.1. Plates were incubated for 48 h at 37°C in a candle jar, and checked for transformants after this time. Transformant colonies were then checked for integrity by DNA sequencing (see section 2.3.4e below).

2.3.4e. DNA sequencing

Sequencing was carried out on DNA products - which had been pre-prepared in accordance with the recommended protocols - by Eurofins Genomics, using prepaid barcodes or Mix2Seq tubes. Sequencing was performed on samples using custom primers, and the reactions were implemented using dideoxy chain termination/cycle sequencing on ABI 3730XL sequencing machines.

Received sample sequences were then aligned against expected sequences using MEGA6 software (Tamura *et al.*, 2013), with alignment by Clustal W. Sequence trace data were manually screened to determine and correct any inaccuracies or inconsistencies in the sequences.

2.4. cDNA analysis and microarray

2.4.1. mRNA extraction from whole cells

S. gordonii cells were cultured in BHY medium for 18 h at 37°C, and then sub-cultured into fresh media, pre-warmed to 37°C. Cells from which RNA was being extracted for downstream qRT-PCR analysis of *recN* and *mutS* genes were grown to OD_{600} 0.5 in 5 mL BHY medium, at which point 1 volume of RNALater (Invitrogen) was added, and the tubes were incubated at room

temperature for 5 min. Following incubation, cells were harvested at 3000 x g for 20 min, the supernatant was removed, and the pellets were frozen at -80°C for RNA extraction at a later date.

From samples where RNA was being extracted for downstream microarray or further qRT-PCR analysis, cells were harvested from an 18 h overnight culture, washed in FMC medium, and sub-cultured into 20 mL replete FMC medium, where they were grown at 37°C until reaching an OD_{600} of 0.3-0.4. At this point, cultures were split into four, and 5 mL cells were harvested and resuspended in 5 mL of either replete FMC, or FMC depleted of arginine (FMC -arg.), histidine (FMC -hist.) or the branched-chain amino acids (FMC -BCAA). Following incubation at 37°C for 30 min, 1 volume of RNALater was added, tubes were incubated at room temperature for 5 min, and then cells were harvested at 3000 x g for 20 min. After harvesting, the supernatant was removed, and the cell pellet frozen at -80°C for RNA extraction within 5 days.

RNA was extracted using the Ambion RiboPure Bacteria RNA Purification kit (Life Technologies), with the following modifications to the cell disruption stage of the protocol: cell pellets were thawed on ice, and resuspended in 100 µl spheroplasting buffer containing 0.1 mg/mL chloramphenicol or spectinomycin. Mutanolysin was added to cells at a concentration of 500 U/mL, and the cells were then incubated at 37°C for 5 min. 350 µl RNAWiz solution (Life Technologies) was added to cells, and the mixture was vortexed vigorously for 15 s. From this point onwards, the original protocol was followed. RNA concentrations were determined using a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific), using the RNA-40 setting. Concentration was measured in ng/µl, and purity and quality determined by assessment of the 260/280 and 260/230 readings. Following extraction, RNA was stored at -80°C for use in downstream applications.

2.4.2. Reverse transcription of RNA

Reverse transcription was carried out on RNA samples extracted from *S. gordonii* cells (see section 2.4.1), using a QuantiTect Reverse Transcription kit (QIAGEN), in conjunction with 3 µg/mL random hexamer primers (Bioline). Reactions were performed according to manufacturer's instructions, with the modification of the Bioline hexamers being used in place

of the QuantiTect oligo-dT primers. Subsequent cDNA was then stored at -20°C for use in qRT-PCR and microarray experiments.

2.4.3. Quantitative reverse transcription PCR

Quantitative reverse transcription PCR was carried out on cDNA samples that were reverse transcribed (see section 2.4.2) from extracted transcriptomic RNA. Samples were placed in triplicate into frosted 96-well (non-skirted, low profile) qPCR plates (Eurogentec, Ltd.), and sealed with an optical Microseal 'B' Adhesive seal (Bio-Rad). Reactions were performed on a DNA Engine Opticon 2 (MJ Research), on the following reaction: 1) 95°C for 10 min; 2) 95°C for 15 s; 3) 60°C for 1 min; 4) Plate read; repeat from step 2 39x; 5) Melting curve from 55-90°C, read every 1°C, hold for 5 s.

Primers were used at 2.5 µM concentrations, and SYBR Green dye from the SensiMix SYBR No-ROX kit (Bioline) was used as both a polymerase and fluorescent dye label for the qPCR reaction. Every reaction plate contained a standard curve, consisting of dilutions of *S. gordonii* DL1 chromosomal DNA, from a 5 x 10^{1-2} to 5 x 10^{8} dilution; no template (blank) control wells; and "no RT" control samples, for control comparison against the experimental samples.

Standard curves, melting curves (to ensure purity of the product), and agarose gel electrophoresis (to check the size of the DNA product), were all performed to validate the qRT-PCR experiments, and were carried out in accordance with MIQE guidelines (Bustin *et al.*, 2009). 16S gene expression was measured for each sample as a reference, and target gene expression was standardised and normalised against 16S expression levels for that sample, to give a fold-change value for the target gene.

2.4.4. Microarray experiments

S. gordonii cDNA was tested by custom single-colour microarray (Agilent Technologies) for the differential expression of 2051 different genes, in two different bacterial strains grown in both high and no arginine conditions, throughout the *S. gordonii* genome.

Samples of RNA (from *S. gordonii* PK3347 and PK3350, in high and no arginine conditions – see section 2.4.1) from four independent experiments, of at least 30 µg/mL concentration, were sent to the Functional Genomics and Proteomics Facility of Birmingham University (UK). There, the samples were reverse transcribed into target cDNA sequences and fluorescently labelled with Cy3 dye, before being hybridised to 60-mer oligonucleotide probes (optimised for hybridisation at 65°C). Two chips were used for the microarray, with each chip capable of holding 8 samples.

Data was analysed with the assistance of Mr Matthew Bashton and Mr John Casement (Bioinformatics Core Facility, Newcastle University), using GeneSpring software (Agilent). Probe annotation was obtained from the most recent GenBank records and associated FASTA file (http://www.ncbi.nlm.nih.gov/nuccore/NC_009785.1), last modified 11th February 2015. A sensitive Bowtie alignment (using Bowtie2 (Langmead and Salzberg, 2012)) was used to align probe sequences against FASTA sequences from the GenBank records, and corresponding annotations were extracted in order to obtain the probe annotation. A small number of probes aligned against positions without annotation, and were therefore labelled "NA".

2.4.5. Bioinformatic analyses of microarray data

Comparisons were made between the new microarray detailed above, and a previous microarray carried out between *S. gordonii* DL1 and PK3354 (as detailed in Jakubovics *et al.* (2015)). Microarray data and probe sequences for the previous microarray can be found in the Gene Expression Omnibus (GEO) database under accession numbers GSE51346 and GPL17786. Statistical analyses were again performed with the assistance of Mr Matthew Bashton and Mr John Casement (Bioinformatics Core Facility, Newcastle University).

For comparison between samples within the same microarray (i.e. DL1 vs PK3354, PK3347 vs PK3350), data was normalised using the 75th percentile normalisation with baseline to median. Any significant differences between gene expression levels in high and no arginine conditions were assessed using T-tests, with *P*-values corrected for multiple comparisons using Benjamini-Hochberg false discovery rate (FDR) corrections (Reiner *et al.*, 2003). Genes were deemed significantly regulated if they had an FDR-corrected *P*-value of <0.05, and a fold-change value of >2. Principal component analysis (PCA) was performed on all samples from both microarrays

using the Genespring software (Agilent), in order to analyse the variance of the data and identify any anomalous results. In one microarray experiment, one sample (*ahcR+*, slide 3 sample 1_2) was removed after being determined to be an outlier following PCA, and so differential expression was carried out only on the remaining 15 samples.

COGFun (Functional Clusters of Orthologous Genes) designations, for labelling of gene descriptions, were taken from the MicrobesOnline database (http://meta.microbesonline.org/operons/ gnc467705.htmL).

Data from both microarrays were subject to cross-comparisons, to determine which genes showed differential regulation between *S. gordonii* DL1 and the arginine regulator mutant strains under different arginine conditions. As no mixing of samples occurred between the two microarrays - due to restraint on the sample size for each microarray chip - and a significant batch effect was shown between the two microarrays following PCA analysis, a simple batch correction was not sufficient as it could not then be determined whether differences in gene expression were biologically significant or due to the batch effect. Therefore, preliminary Rank Product analysis was performed to determine which genes showed significant regulation between the two microarrays. A gene was listed as significantly differentially expressed if its pfp (percentage of false-positive predictions) was less than 0.05 (95% confidence). Data analysis was then performed upon the subsequent data.

For comparisons of microarray vs qRT-PCR data, the correlation of $\log_2$ data from the two datasets was analysed using a scatter graph with a linear regression analysis, and the r^2 value and slope of the line measurements from the linear regression line were used as indicators of levels of correlation between the datasets. These analyses were performed using the SigmaPlot 12.5 software (Systat Software, Inc.).

2.5. *In silico* analysis of *S. gordonii* genes

2.5.1. Promoter prediction

Streptococcus gordonii genes were assessed for predicted promoters using the PromBase website (Rangannan and Bansal, 2011). Promoters were judged by their average free energy

over the length of the predicted promoter sequence, and then ranked by strength and likelihood of being a true positive result. The most reliable promoter was chosen as the predicted promoter during the *in silico* analysis.

2.5.2. Operon prediction

S. gordonii genes were analysed for operon formation and co-transcription with other genes using the DOOR^2 database (Database of prOkaryotic OpeRons v2) (Dam *et al.*, 2007; Mao *et al.*, 2009).

2.5.3. Transcription factor binding site prediction

Transcription factor binding sites (TFBS) were predicted within the *S. gordonii* genome using the PePPER TFBS Search function (de Jong *et al.*, 2012). Homologs of the arginine regulators in other bacterial species (e.g. *Lactococcus lactis*), and their binding sites, were used to scan for similar sites within *S. gordonii*. Each predicted site was then scored on their strength, and the strongest binding sites were deemed the most likely to be regulated by one of the arginine regulatory genes.

2.5.4. Prokaryotic terminator prediction

Terminators were predicted using the TransTermHP database (Kingsford *et al.*, 2007), which scanned the *S. gordonii* genome for predicted Rho-independent bacterial terminators. Gene sequences were then analysed for the position of the terminator, nearby gene promoters, whether the gene was part of an operon, and therefore how likely it was that the predicted terminator was a true positive result. Predicted terminators were scored by strength based on the potential energy of the hairpin in the terminator sequence.

3. L-arginine-specific regulation of gene expression

3.1. Introduction

Previous work on growth in *S. gordonii* in response to L-arginine concentration has shown that, whilst L-arginine appears to have a large effect on biofilm formation, it seems to have little effect on the planktonic growth of the cells. Specifically, whilst low concentrations (up to 500 µM) of L-arginine encouraged *S. gordonii* biofilm growth, very high concentrations of around 500 mM - the same levels found in oral healthcare products - were detrimental to the formation of biofilms when grown in cell-free saliva in a BioFlux microfluidics system (Jakubovics *et al.*, 2015). These experiments also demonstrated that the same 500 µM and 500 mM L-arginine concentrations had no effect on the planktonic growth of the cells, in both saliva and chemically-defined FMC medium.

Additional unpublished data from our lab (Jakubovics, unpublished) showed that growing *S. gordonii* cells in a microtitre plate in FMC medium, with decreasing concentrations of L-arginine from 500 mM to 5 µM, caused no differences in the levels of total biomass or cell yield of either planktonic cells or biofilms across all L-arginine concentrations. However, both this work and the previous experiment were measuring planktonic growth in biofilm models – to date, no data have been obtained measuring growth in strictly planktonic cultures. Furthermore, these experiments were performed under aerobic conditions, and did not assess the growth rate of *S. gordonii*. It is possible that L-arginine could have an effect on planktonic growth, similar to the effect it demonstrates in biofilms, but that this effect was masked in these experiments by the concurrent biofilm growth. Additionally, anaerobic growth of planktonic cells could demonstrate differences to the aerobic conditions used in the experiments mentioned above. A number of *Streptococcus* strains, including *S. salivarius* and *S. mutans*, are known to require L-arginine in order to grow aerobically, despite being able to grow anaerobically without it (Terleckyj and Shockman, 1975).

As L-arginine is added to some oral health products in relatively high concentrations of up to 460 mM (Sullivan *et al.*, 2014), it is important to understand how oral bacteria generally respond to these concentrations. As a start, this work focuses on one important species, *Streptococcus gordonii*, known to be a widespread coloniser of the oral cavity and a key

organism in the establishment of the dental plaque biofilm (Rickard *et al.*, 2003). It is likely that *S. gordonii* is exposed to a wide range of different L-arginine concentrations whilst growing *in vivo* in a plaque biofilm, as the concentration of L-arginine naturally occurring in saliva is low (5-10 µM (Brand *et al.*, 1997)), and the concentration in some oral healthcare products significantly higher. The effect of these differing concentrations on planktonic growth could vary in response to whether bacterial cells are grown *in vitro* in chemically-defined medium, or in a more biologically-relevant medium such as saliva, and this will be investigated within this chapter.

It is known that arginine depletion in *S. gordonii* plays a large role in gene expression, as a genome-wide microarray performed previously by our group demonstrated a large number of genes regulated in response to arginine depletion, which is discussed in more detail below (Jakubovics *et al.*, 2015). However, arginine depletion is also known to cause growth arrest in *S. gordonii*, as previous work showed that depletion of arginine within a planktonic monoculture of *S. gordonii* in FMC medium resulted in immediate entry of the cells into stationary phase, and a lower growth yield than cells cultured in arginine-replete FMC medium (Jakubovics *et al.*, 2008). Therefore, in order to distinguish between genes that are regulated in response to arginine depletion, and genes that are regulated by a global stress response caused by amino acid depletion, a control consisting of samples depleted of other amino acids was needed. Therefore, both growth and gene expression were analysed in samples that were depleted of L-arginine, L-histidine and the branched-chain amino acids (L-leucine, L-isoleucine and L-valine) in this chapter.

Histidine was chosen for comparison with L-arginine, as there is an overlap in gene regulation pathways between L-arginine and L-histidine, with the arginine-dependent regulator ArgR repressing biosynthesis of both amino acids under high arginine conditions (Jakubovics *et al.*, 2015). The branched-chain amino acids have been shown to be necessary for planktonic growth of *S. sanguinis*, with cells inhibited in growth without them (Cowman *et al.*, 1975). Preliminary growth experiments indicated this is also the case for *S. gordonii*, with the same growth pattern observed after depletion of the branched-chain amino acids and L-histidine as with depletion of L-arginine – namely, arrest of growth and entry to stationary phase when cells are placed into deplete medium. *S. gordonii* gene expression analysis following the depletion of L-histidine or the branched-chain amino acids, when compared to L-arginine, would demonstrate whether

any differences in gene expression observed between these samples were arginine-specific, or a general stress response.

As mentioned, a previous study on arginine-specific gene regulation by our group, as assessed by genome-wide microarray analysis, identified major responses to arginine depletion. Some of these responses appeared to be linked to changes in arginine requirements, and others seemed more characteristic of a general stress response or cessation of metabolic activity (Jakubovics *et al.*, 2015). In all, 464 genes, which represented a large percentage (22.6%) of the *S. gordonii* chromosome, were differentially-regulated in response to arginine depletion. Of these 464 genes, a number of interesting genes that may be linked to biofilm formation showed regulation in response to L-arginine depletion. These formed the basis for the genes chosen for qRT-PCR analysis in this chapter. The 12 genes of interest used here were chosen as they either showed interesting results within the microarray, or were known to be linked to arginine biosynthesis or biofilm formation within *S. gordonii* in some way. These genes are described in detail below.

SGO_0966 (*hsa*) and SGO_0978 (*asp5*) are part of the *hsa* locus. Hsa is a streptococcal surface adhesin that mediates *S. gordonii* binding to sialic acid residues on human fibronectin (Jakubovics *et al.*, 2009). SGO_1401 and 1411 (*hisC*) are part of the histidine biosynthesis locus, which is known to share regulation with the arginine biosynthesis pathways (Jakubovics *et al.*, 2015). SGO_1569 (*argC*) was chosen as a highly up-regulated gene under arginine-deplete conditions in the microarray, which as an arginine biosynthesis gene might be expected to be up-regulated specifically in response to arginine depletion. A similar rationale was also used for selecting SGO_1592 (*arcB*), which as an important arginine biosynthesis gene (despite being found within the *S. gordonii* ADS locus) may be expected to be up-regulated under no arginine conditions.

SGO_1576 (*bfbC*) and SGO_1582 (*bfbF*) were chosen for qRT-PCR analysis as members of the *bfb* locus, which encodes the cellobiose PTS system, and which showed interesting down-regulation in response to arginine depletion in the microarray. This locus has also been previously linked to biofilm formation and adhesion in *S. gordonii* (Kiliç et al., 2004). The final two operons which were chosen also showed unexpected regulation in response to arginine

depletion – those being the fatty acid biosynthesis genes SGO_1686 and 1699, which showed strong down-regulation following arginine depletion in the previous microarray analysis; and the receptor polysaccharide genes SGO_2015 (*wefE*) and 2028 (*wzg*). Receptor polysaccharide also facilitates coaggregation between *S. gordonii* and the type II fimbriae of *Actinomyces oris* (Cisar et al., 1995; Mishra et al., 2010), another primary coloniser at the base of the biofilm, so the arginine-specific regulation that that locus appeared to show within the microarray was unexpected.

Analysis of the expression of these genes in response to not only L-arginine, but also L-histidine and branched-chain amino acid depletion was performed to indicate which genes showed an arginine-specific regulation, and which displayed general responses to amino acid depletion, which may be indicative of a larger depletion-based stress response within the cells. In addition, comparison of the expression levels of these genes as assessed by qRT-PCR, against the findings of the genome-wide microarray, may allow validation of the findings of the microarray.

Therefore, the overall aims of this chapter were to assess the impact of different concentrations of L-arginine on planktonic growth of *S. gordonii* cells under anaerobic conditions, in both defined medium and cell-free saliva; and to analyse any differences in gene expression of selected genes of interest in response to arginine depletion in planktonic growth conditions. Furthermore, this chapter aims to establish any arginine-specific responses in gene expression, and distinguish them from general amino acid-depletion responses.

3.2. Effects of L-arginine on *S. gordonii* planktonic growth

3.2.1. S. gordonii *DL1 growth in chemically-defined media with differing L-arginine concentrations*

S. gordonii has previously been shown to cease growth following transfer to no-L-arginine medium under aerobic conditions (Jakubovics *et al.*, 2008). To investigate the impact of L-arginine on *S. gordonii* growth in more detail, *S. gordonii* DL1 cells were grown to stationary phase in FMC supplemented with between 0 µM and 500 mM L-arginine.

Cells were cultured overnight in nutrient-rich BHY medium, then used to inoculate FMC, or FMC amended with different L-arginine concentrations. The concentrations of L-arginine present in

the different FMC samples were as follows: 0 µM, 10 µM, 50 µM, 100 µM, 1 mM, 50 mM, and 500 mM. Cells were grown strictly anaerobically, with timepoints taken every hour over a 7-hour time period, in order to follow the growth of the cells to stationary phase. Growth was followed by measuring OD_{600} of the different samples as described in Materials and Methods, with additional OD_{600} measurements also taken at 24 h and 48 h timepoints.

Cells grown in 1 mM L-arginine grew rapidly, with a maximum specific growth rate of 1.18 h^{-1} (Figure 3.1). There was a slight decrease in exponential growth rate at 100 µM L-arginine (0.98 h^{-1}), and a more pronounced decrease at 50 µM L-arginine (0.63 h^{-1}). All other samples showed extended lag phases, although by 24 h all cultures had reached OD_{600} >1, with the exception of cells grown in 500 mM L-arginine. The 500 mM sample showed some small increase in optical density over all timepoints (with a mean specific growth rate of 0.31 h^{-1}), however it reached stationary phase at a lower OD_{600} (and therefore a gave a lower growth yield) than the other cultures (OD_{600} 0.28).

This poor growth may have been due to *S. gordonii* cells rapidly metabolising the high concentrations of L-arginine into ammonia under anaerobic conditions, resulting in an extremely basic solution within the tube that was retarding growth. The starting pH of every culture was the same, however, measurement of the final pH of the samples showed that the pH of the 500 mM L-arginine sample was higher (pH 5.9) than for 1 mM L-arginine (pH 4.5). However, it is not clear whether this was due to increased alkali production through the ADS, or decreased acid production from lower growth yield. Regardless, the pH of the 500 mM L-arginine sample was still not high enough to inhibit growth. It is possible that the intracellular pH may have been sufficiently high to impede growth.

Similarly, in preliminary experiments looking at aerobic growth, *S. gordonii* DL1 yield was much lower at 500 mM L-arginine concentration compared with 1 mM concentration (OD_{600} 0.052 compared to 2.015 after 48 h). Microscopic analysis of these samples showed normal morphology, with no signs of cell lysis (data not shown). Cells incubated aerobically in 0 µM L-arginine as a control did not grow, as has been previously shown (Jakubovics *et al.*, 2008). Therefore it is not clear why 500 mM L-arginine had an apparent bacteriostatic effect on the

cells, despite previous work showing no effect on cell growth or yield (under aerobic conditions).

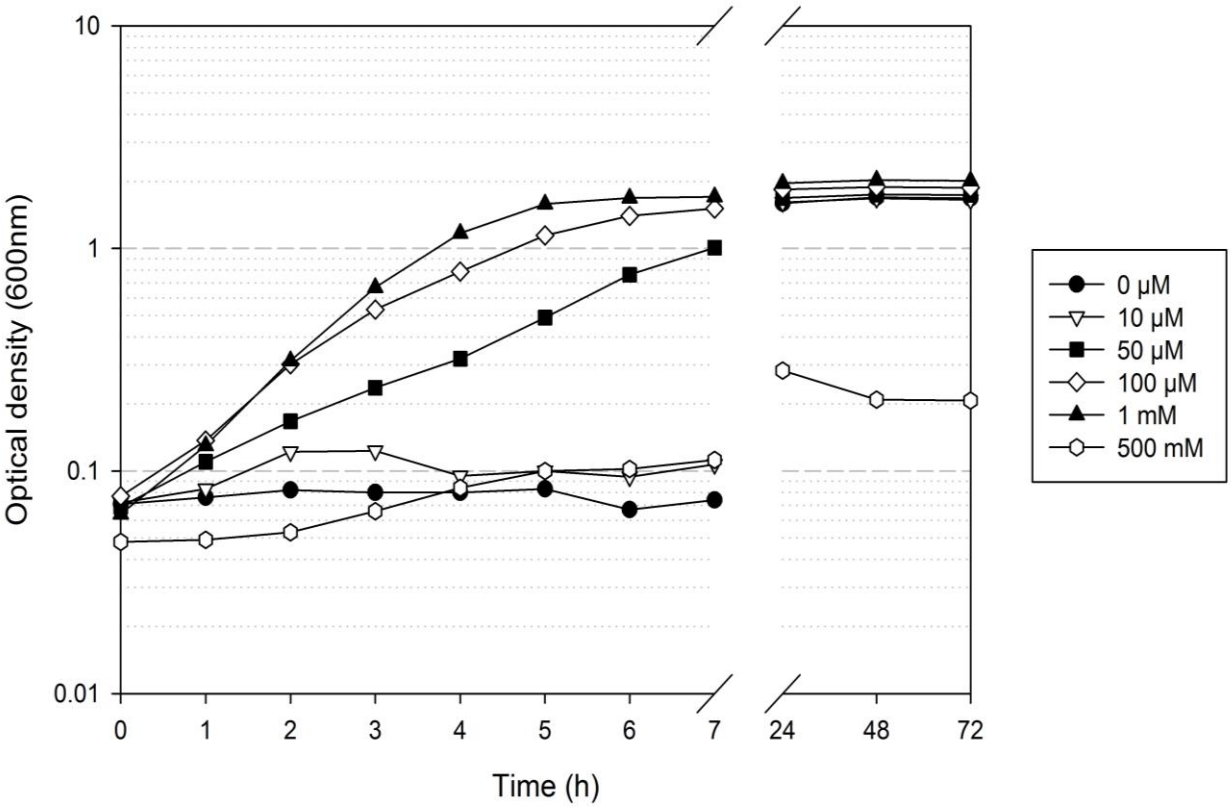


Figure 3.1. Representative growth curve of *S. gordonii* DL1 cells in CDM supplemented with differing L-arginine concentrations. Cells were grown in FMC medium with different L-arginine concentrations, and spectrophotometric measurements were taken at various timepoints as a measure of cell growth. Graph shows one of four independent experiments, all of which showed similar results. All cells grew to an OD_{600} >1 within 24 h, with the exception of *S. gordonii* DL1 cultured in 500 mM L-arginine, which showed a lower growth yield than other samples.

3.2.2. S. gordonii *DL1 growth in saliva with differing L-arginine concentrations*

Having shown that L-arginine concentration had a major impact on planktonic cell growth in FMC, the effects of L-arginine concentration on *S. gordonii* growth in saliva were assessed. Saliva naturally has a low concentration of free L-arginine, with an average concentration of around 5-10 µM (Brand *et al.*, 1997). However, it is important to note that saliva is a complex fluid, and that L-arginine residues could potentially also be released from salivary polypeptides by proteolytic cleavage. Here, *S. gordonii* was cultured in saliva supplemented with L-arginine at concentrations up to 500 mM.

It was found (Figure 3.2) that viability of the cells decreased slightly over the 48 h time period (7.39-fold change in 500 mM L-arg, 5.77-fold change in 0 mM L-arg), and no growth was detected over this period, regardless of L-arginine concentration. These data indicate that high L-arginine concentrations are not toxic to *S. gordonii* cells. There was no clear difference between the concentration of viable cells in 500 mM or 0 µM L-arginine samples, and whilst there was some decline in cell viability towards the later timepoints, this was likely due to natural cell death. Phase contrast microscopy showed no difference in cell morphology between high and low arginine concentrations, although it was noted that whilst cells appeared to form aggregates in lower concentrations of L-arginine, these aggregates did not form in the 500 mM samples (Figure 3.3).

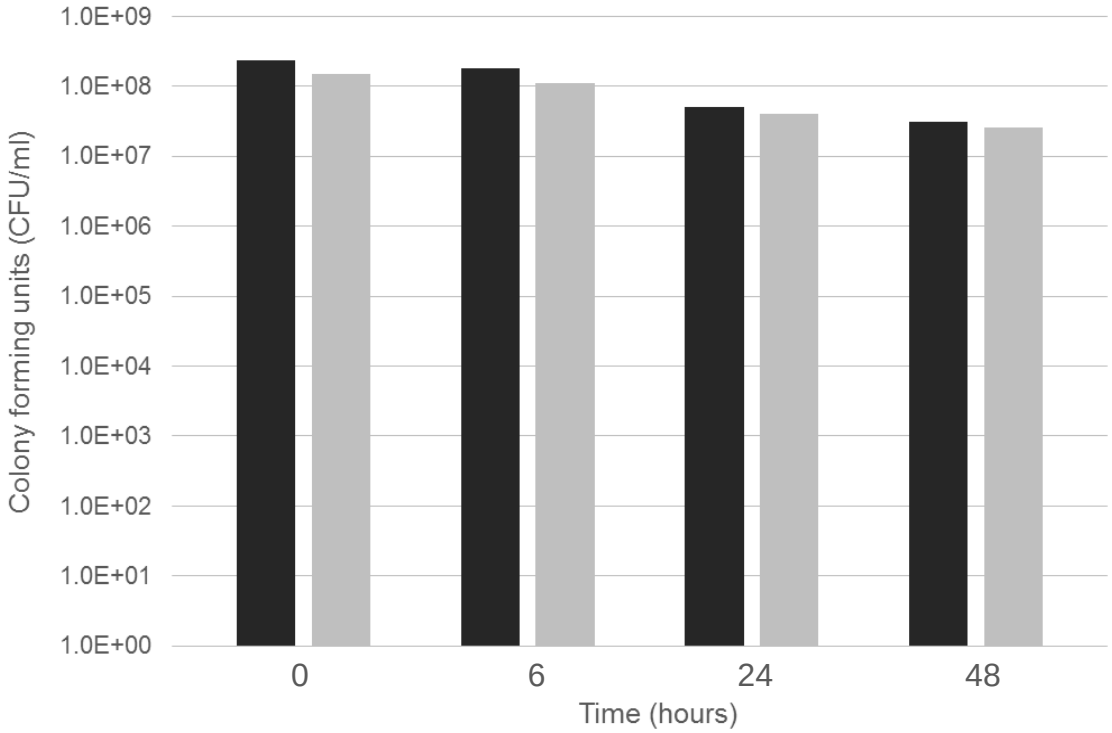


Figure 3.2. Representative graph showing the growth of *S. gordonii* DL1 cells in high and no arginine conditions, as measured by total viable count (CFU/mL). Cells were incubated in 25% saliva and 500 mM (dark grey) or 0 mM arginine (light grey), and viable counts were taken at different timepoints as a measure of cell growth. Graph shows one biological replicate of four independent experiments, all of which showed similar results. Decrease in CFU/mL in the 24 h and 48 h timepoints is likely due to natural cell death.

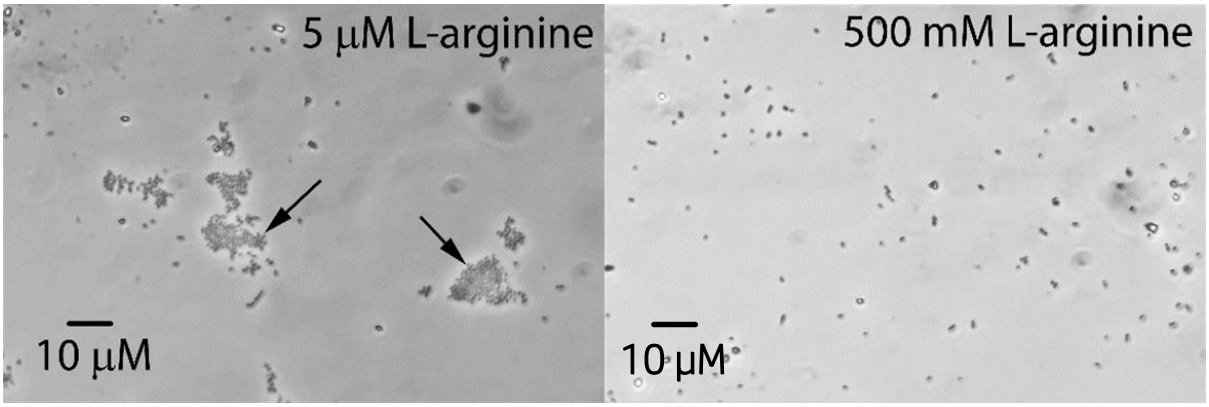


Figure 3.3. Phase contrast microscopy images of *S. gordonii* DL1 cells incubated in differing arginine concentrations in saliva. The arrows present in the 5 μM L-arginine image (left) indicate clusters of *S. gordonii* cells, caused by auto-aggregation of the cells in saliva. The same auto-aggregates were not present in the 500 mM sample (right). Despite the lack of aggregation, cell morphology was found to not be affected by the L-arginine concentration.

3.2.3. S. gordonii *DL1 growth in the absence of different amino acids*

S. gordonii contains all necessary genes for L-arginine biosynthesis, and has been shown above to have the ability to grow in medium that is entirely depleted of L-arginine, under strictly anaerobic conditions (Terleckyj and Shockman, 1975; Jakubovics *et al.*, 2008). Previous work in our group also indicated that when actively growing *S. gordonii* cells were sub-cultured from L-arginine-replete medium into fresh medium lacking L-arginine, growth arrested rapidly (Jakubovics *et al.*, 2008). In order to investigate whether this functional auxotrophy was an L-arginine-specific response or whether similar responses would be seen following rapid depletion of other amino acids, the experiment was repeated with cells sub-cultured into fresh medium lacking either L-arginine, L-histidine, or the branched-chain amino acids (L-leucine, L-isoleucine and L-valine). *S. gordonii* carries the essential biosynthetic genes for all of these amino acids according to the KEGG database entry for amino acid biosynthesis (http://www.genome.jp/kegg-bin/show_pathway?sgo01230), and theoretically should have the ability to grow in defined media without them.

S. gordonii cells were cultured in an amino acid replete medium (FMC), and aliquots were harvested and resuspended in FMC medium, or FMC depleted of L-arginine (FMC –arg), L-histidine (FMC –hist), or the branched-chain amino acids (FMC –BCAA). Growth of the cells was followed from the point of sub-culture to stationary phase.

Whilst cells that were placed back into amino acid-replete medium continued to grow rapidly (Figure 3.4), and reached a high yield (OD_{600} >1), the cells that were placed into any of the depleted media all entered stationary phase soon after sub-culture. None of these samples resumed exponential growth, and none achieved a growth yield of OD_{600} >1. Therefore, *S. gordonii* appears to have multiple conditional auxotrophies.

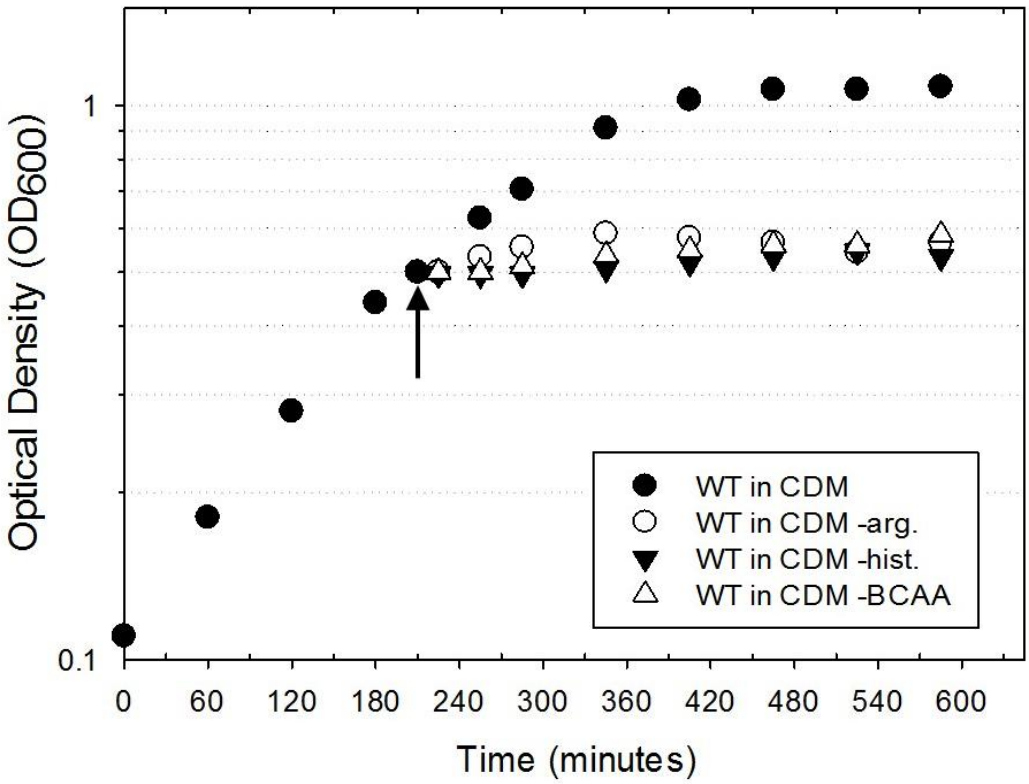


Figure 3.4. Effect of amino acid depletion on the growth of *S. gordonii* DL1 cells in chemically-defined medium (CDM). Cells were grown in replete FMC medium (CDM) to mid-exponential phase, and then split into four parts (point of sub-culture indicated by arrow). One part was washed and resuspended in a fresh tube of the amino acid-replete FMC; and the other three were washed and resuspended in FMC lacking L-arginine, L-histidine or the branched-chain amino acids (L-valine, L-leucine and L-isoleucine). The cells placed into the replete FMC continued growing to stationary phase, whereas the cells grown in any of the depleted media showed growth arrest shortly after the point of sub-culturing.

As *S. gordonii* contains all necessary biosynthetic genes and pathways for these amino acids, and is able to grow in medium lacking L-arginine under anaerobic conditions, the reason is unclear as to why these cells are apparently unable to synthesise the depleted amino acids and resume aerobic growth upon rapid depletion of L-arginine. The lack of growth in depleted media may be due to an over-accumulation of an intermediary molecule, or failure to synthesise key biosynthetic enzymes (which are repressed in their expression during growth in amino acid-replete medium).

Identification of similar growth responses to depletion of different amino acids allowed comparison of gene regulation responses, aimed at distinguishing between arginine-specific regulation, and gene regulation in response to general stress and growth arrest (see section 3.3.2). Therefore, the expression levels of different genes in response to depletion of L-arginine was next analysed.

3.3. Effects of arginine on gene expression

A genome-wide gene expression microarray was previously carried out on *S. gordonii* DL1 cDNA (Jakubovics *et al.*, 2015), in high and no arginine conditions, in order to investigate which genes were being specifically-regulated in response to arginine depletion in planktonic cells. The details of the genes that were regulated within this microarray can be found in the introduction to this chapter.

3.3.1. qRT-PCR of gene expression following L-arginine depletion

In order to validate the findings of the *S. gordonii* DL1 gene expression microarray (Jakubovics *et al.*, 2015), qRT-PCR analysis was performed on the same cDNA samples that were used for the microarray analysis. qRT-PCR was performed on 12 genes of interest (detailed in the chapter introduction), contained within 6 different operons. The 16S gene was also tested as a reference gene. qRT-PCR was validated using melt curves, standard curves, and agarose gel electrophoresis (Figure 3.5). Reaction efficiency for each experiment was also calculated, and the data from any reactions that had an efficiency of between 80-120% was used for further analysis, provided gel electrophoresis analysis showed a DNA product of the expected size.

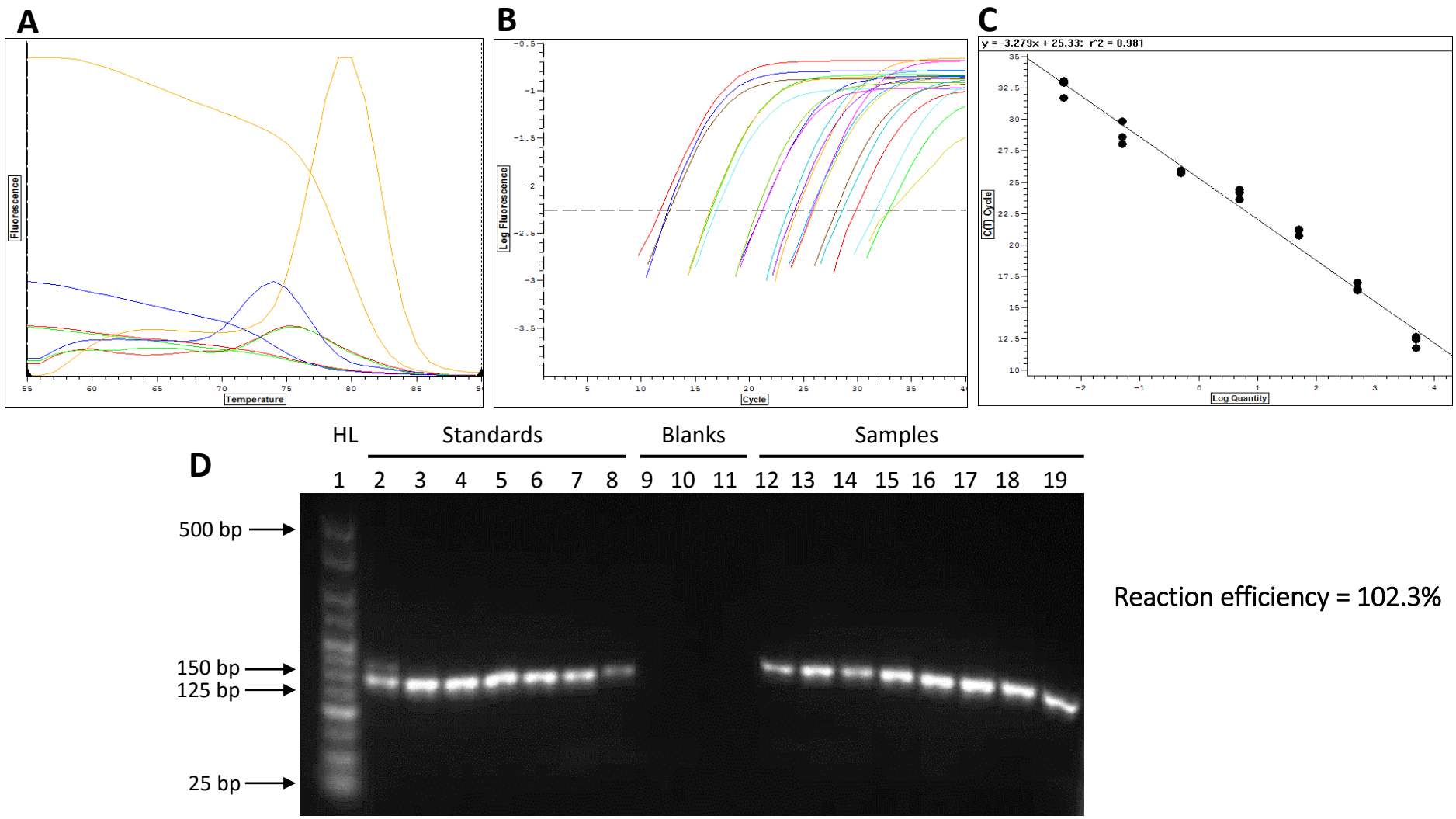


Figure 3.5. Example of assay development for qRT-PCR: analysis of the SGO_2028 (*wzg*) gene. Melt curves (A) and DNA standard curves (B and C) were run alongside samples during qRT-PCR reactions as controls. The standard curve contained known ten-fold dilutions of *S. gordonii* DL1 chromosomal DNA (from 1:500 to 1:5x10^8), which covered the range of dilutions for the samples tested. The melt curve was used to ensure that the product formed was pure, and not a result of primer dimers. The red, blue and green lines represent triplicate blank (no template) wells, with the yellow line (from one 1:500 DNA standard) for comparison. (D) shows a 3% agarose gel electrophoresis, used to check the size of the qPCR product (131 bp expected). HL represents Hyperladder V (Bioline; 25-500 bp). There were no bands present in the blank wells (9-11), indicating no contamination of the reaction. All bands in the standard (wells 2-8, decreasing from 1:500 to 1:5x10^8 conc.) and sample wells (12-19, replicates of cDNA from *S. gordonii* DL1 and Δ*arcR* in high and no arg) were the expected size.

The genes that were chosen for qRT-PCR analysis are listed in Table 3.1. Target gene expression was measured by qRT-PCR, with the resultant values standardised and normalised against 16S reference data to give a $\log_2$ fold-change value for expression of each of the target genes measured (Figure 3.6). The absolute gene expression levels (yielded by multiplying 2 to the power of the $\log_2$ values) ranged from strongly down-regulated (e.g. the *bfbF* gene, down-regulated 85-fold), to highly up-regulated (e.g. *argC*, up-regulated 420-fold in response to arginine depletion).

The changes in gene expression for qRT-PCR were then compared with the microarray, for each of the 12 target genes (Table 3.1 and Figure 3.7). It was found that the levels of gene expression measured for target genes by both qRT-PCR and microarray correlated very closely, with a direct comparison between the two sets of values yielding a graph with a linear regression line, with a slope of 0.995 and an r^2 value of 0.947. This indicated a strong positive correlation between the two sets of data, validating the findings of the microarray.

Target gene name	Gene/Locus function	Microarray expression level (absolute fold-change)	qRT-PCR expression level (absolute fold-change)
SGO_0966 (*hsa*)	*hsa* locus, binding to oral cavity	-8.22	-2.86
SGO_0978 (*asp*5)		-8.31	-8.28
SGO_1401	Histidine biosynthesis locus	4.36	3.22
SGO_1411 (*hisC*)		13.42	7.54
SGO_1569 (*argC*)	Arginine biosynthesis	520.12	420.22
SGO_1576 (*bfbC*)	Cellobiose PTS system locus	-21.10	-20.65
SGO_1582 (*bfbF*)		-35.01	-85.04
SGO_1592 (*arcB*)	ADS system locus	-2.81	-3.41
SGO_1686	Fatty acid biosynthesis locus	-17.15	-51.48
SGO_1699		-12.55	-11.00
SGO_2015 (*wefE*)	Receptor polysaccharide locus	-12.31	-12.91
SGO_2028 (*wzg*)		-4.03	-8.07

Table 3.1. Comparison between gene expression levels for target genes as measured by microarray or qRT-PCR. Negative numbers represent a down-regulation in gene expression in high vs no arginine conditions.

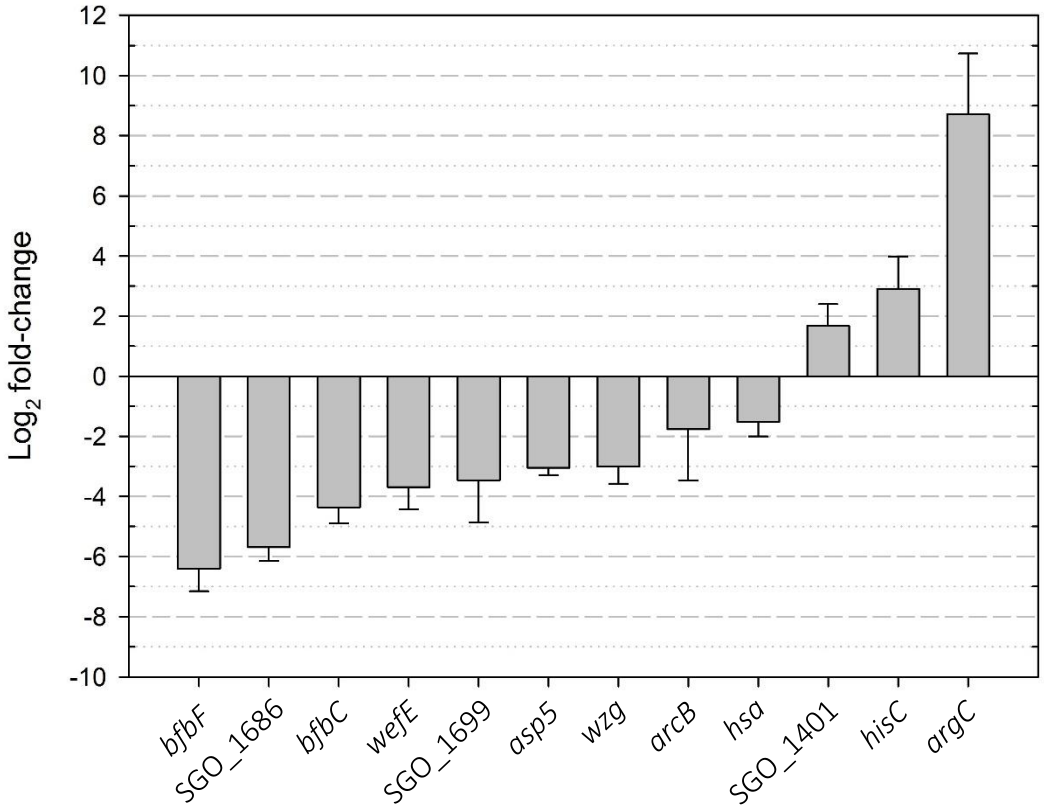


Figure 3.6. Log_2 gene expression fold-change values for genes assessed by qRT-PCR analysis. Target genes were measured in their change in expression in response to high vs no arginine conditions, using relative quantification by qRT-PCR. After normalisation against 16S background gene expression, target gene expression was expressed as a log_2 fold-change value, representing the change in gene expression level within DL1 cells when moving from high to no arginine conditions.

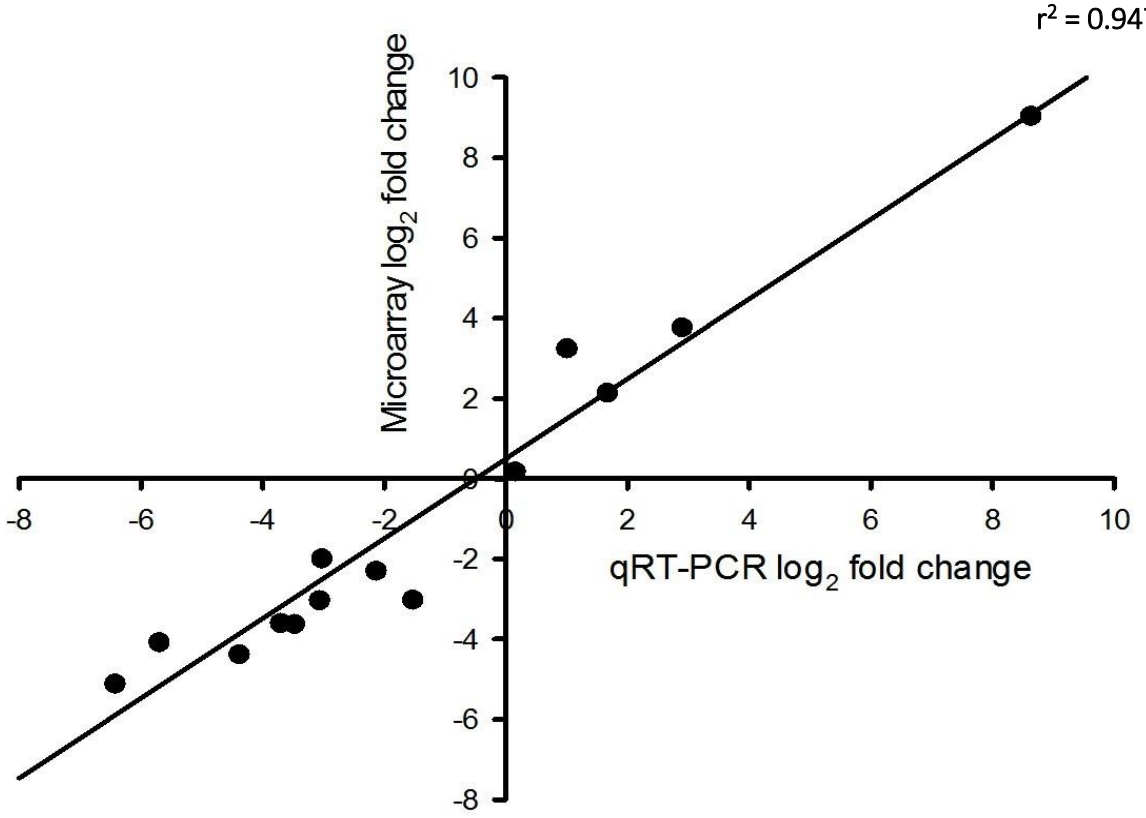


Figure 3.7. Correlation between gene expression levels measured by microarray or qRT-PCR. Levels of gene expression were measured as $\log_2$ fold-change when moving from high to no arginine conditions – comparison between these values from the *S. gordonii* DL1 microarray data and the qRT-PCR reactions (run on the same samples as used for the microarray) shows a strong correlation (r^2=0.947, slope=0.995), indicating that qRT-PCR accurately validates the findings of the microarrays.

3.3.2. qRT-PCR of gene expression following L-histidine and branched-chain amino acid depletion

The microarray analysis carried out in the paper by Jakubovics *et al.* (2015) demonstrated 464 genes responding to arginine depletion by changing their expression level. It was possible that at least part of the regulatory response was due to general stress, linked to growth arrest. Therefore, in order to test which of the differences in gene expression were linked specifically to arginine depletion, qRT-PCR analysis was performed on the same target genes tested in section 3.3.1, this time on cDNA samples taken from *S. gordonii* DL1 cells that were deprived of L-arginine, L-histidine or the branched-chain amino acids (L-valine, L-leucine and L-isoleucine).

S. gordonii DL1 cells were grown in the same way as for the original microarray, to mid-exponential phase in replete chemically-defined FMC medium. At this stage of growth, cells were split into four different cultures, harvested and then resuspended in either replete FMC, FMC without arginine (-arg), FMC without histidine (-hist) or FMC without the branched-chain amino acids (-BCAA). They were then incubated for 30 min, and the cells were harvested and RNA was extracted. Following reverse transcription to cDNA, the samples were analysed by qRT-PCR. Four independent biological experiments were used for each sample.

This analysis showed that whilst some of the target genes assessed appeared to be differentially-regulated due to a generalised stress response, others were specifically-regulated by depleting a particular amino acid (Table 3.2).

For example, *argC* was strongly regulated in response to arginine depletion and not by depletion of other amino acids, as expected for an arginine biosynthesis gene (Figure 3.8). This suggests that this is an arginine-specific response. SGO_1686 (fatty acid biosynthesis) and *asp*5 (involved in *hsa* surface protein biosynthesis), also appeared to show arginine-specific regulation, with much less regulation of these genes in response to histidine or branched-chain amino acid depletion.

However, other genes, such as *hsa* or *bfbF*, appeared to show equal levels of expression across all amino acids, suggesting that they may be regulated as part of a general stress response. The

qRT-PCR values for the *hsa* gene showed much lower levels of expression in response to depletion of not only L-arginine, but also L-histidine and the branched-chain amino acids, in comparison to the value seen in the original microarray analysis (1-fold down-regulation vs 8-fold down-regulation). This may be due to microarray technology having a broader, genome-wide focus, and as such generally being less sensitive to changes in lowly-expressed genes (Wurmbach *et al.*, 2003). In addition to this, some evidence suggests that differences in the location of the microarray probe, and qRT-PCR primers within the target gene can result in differences in apparent gene expression (Etienne *et al.*, 2004).

These analyses indicated that other genes, such as *hisC* and *wefE*, appeared to show specificity to both arginine and histidine, but not the branched-chain amino acids. This may be due to the fact that arginine and histidine share overlapping regulatory pathways, with the protein ArgR repressing both histidine and arginine biosynthesis under high arginine conditions (Jakubovics *et al.*, 2015).

It is also important to note that some large differences were observed between the original qRT-PCR analysis on samples deprived of L-arginine (Table 3.1), and the qRT-PCR analysis performed in this section (Table 3.2). These are likely due to the use of different biological samples for the qRT-PCR analyses. Any slight differences in growth phase or yields of these replicates prior to RNA extraction (and subsequent cDNA synthesis) may possibly result in changes in gene expression across the genome, leading to large differences in fold-change expression of different genes as observed here (such as *bfbF*, 9-fold down-regulation following arginine depletion in these experiments vs 85-fold down-regulation following arginine depletion in the original qRT-PCR analysis). This could be controlled for in future qRT-PCR analyses by measuring the optical density of the cells prior to RNA extraction, to ensure similar growth yields. However, this was not called for in the method used in these analyses and so was not performed, potentially leading to the differences observed between the two datasets.

Overall, some of the genes shown to be regulated in response to arginine appeared to be specific for arginine, others appeared to be regulated as part of a general response to amino acid depletion, suggestive of a more general stress response following growth arrest.

Target gene name	-Arg absolute fold-change	-Hist absolute fold-change	-BCAA absolute fold-change
SGO_0966 (*hsa*)	-1.45	-1.23	-1.06
SGO_0978 (*asp5*)	-21.65	-2.71	-2.99
SGO_1401	4.77	13.61	1.19
SGO_1411 (*hisC*)	19.28	17.88	5.85
SGO_1569 (*argC*)	33.92	1.28	-1.16
SGO_1576 (*bfbC*)	-19.59	-19.00	-13.57
SGO_1582 (*bfbF*)	-9.59	-10.15	-7.35
SGO_1592 (*arcB*)	-2.31	-2.22	1.02
SGO_1686	-35.33	-2.14	-4.30
SGO_1699	-2.78	-1.24	-1.63
SGO_2015 (*wefE*)	-17.81	-14.30	-4.52
SGO_2028 (*wzg*)	-4.28	-2.08	-1.10

Table 3.2. Comparison of absolute fold-change values for qRT-PCR samples deprived of arginine, histidine or the branched-chain amino acids. Fold-change represents change in gene expression when moving from replete to depleted medium. Negative values represent down-regulation of gene expression.

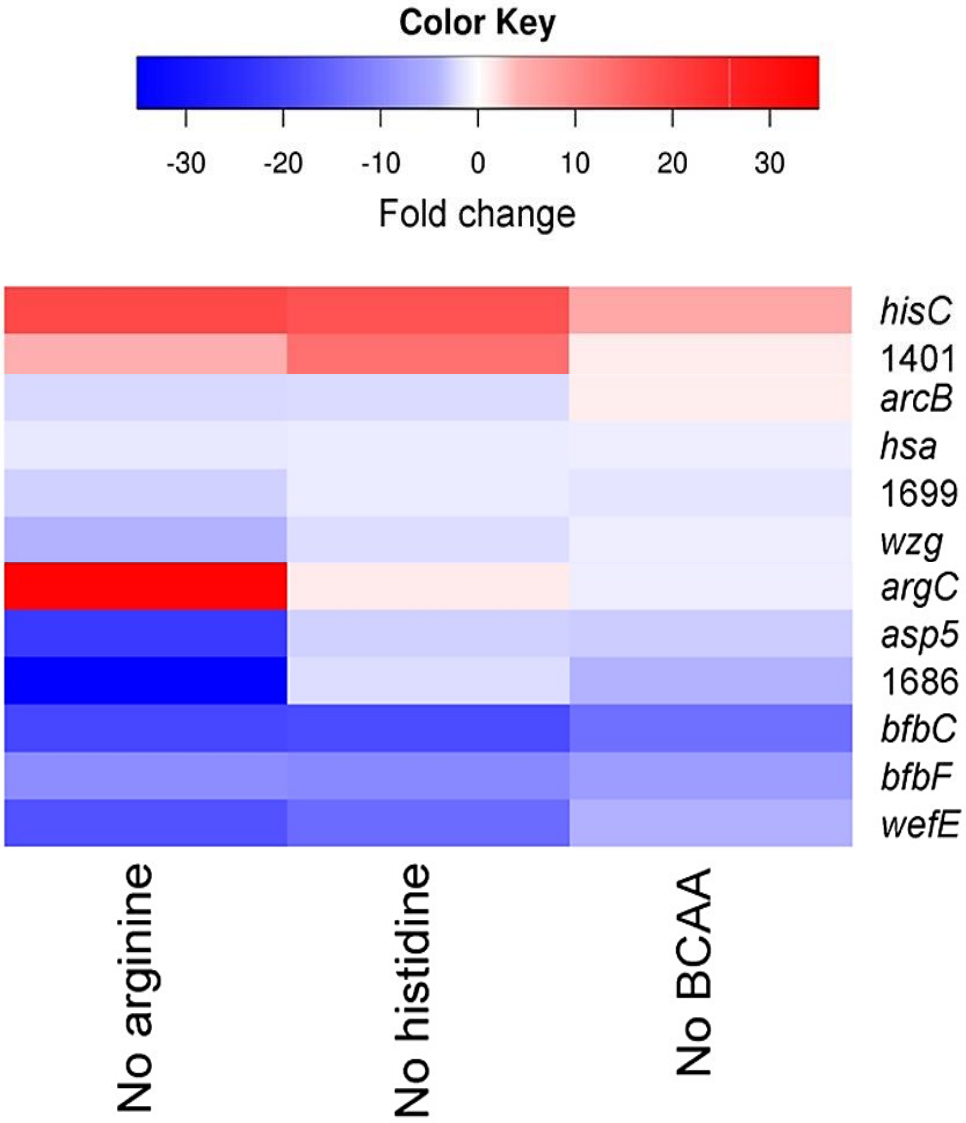


Figure 3.8. Heat map showing visual representation of differential gene regulation in response to amino acid deprivation. qRT-PCR analysis was carried out on cDNA samples from *S. gordonii* DL1, resuspended in replete media or media depleted of arginine/histidine/branched-chain amino acids. The differences in $\log_2$ fold-change are represented here visually, with those genes that were up-regulated (during a movement from replete to amino acid-depleted medium) shown in red, and those that were down-regulated (under the same conditions) shown in blue. It appears that gene regulation consists of amino acid-specific responses (e.g. *argC*), and general stress responses (e.g. *bfbF*).

3.4. Discussion

The work carried out in this chapter showed that whilst high concentrations of L-arginine do not kill *S. gordonii* planktonic cells (as demonstrated by the growth in saliva experiment, where cells retained viability), they do inhibit growth, causing lower growth yields and preventing auto-aggregation of *S. gordonii* cells.

As mentioned in the chapter introduction, *S. gordonii* has not been shown before to have difficulty growing in 500 mM L-arginine. However, these experiments (Jakubovics *et al.* (2015); Jakubovics, unpublished) were performed under aerobic conditions, in a model that was designed for assessing biofilm growth more than planktonic growth. The research performed directly on planktonic growth in this chapter contradicted the findings of these previous experiments, showing that L-arginine does in fact appear to inhibit planktonic growth in *S. gordonii*. This suggests that in past work, the concurrent growth of the biofilm in the model may have been masking any effect L-arginine could have been having on planktonic cells. The larger surface area in these models could have enhanced the planktonic growth yield of the cells by allowing *S. gordonii* cells to detach from the biofilm and move back into the planktonic phase, increasing the apparent planktonic cell yield. As the experiments performed in this chapter were carried out in test tubes, this reduced the surface area available to the *S. gordonii* cells to form biofilms, and allowed a more direct measurement of planktonic cell yield and growth. Here, it has been shown that 500 mM L-arginine concentrations reduce cell growth and yield in defined medium under planktonic conditions.

It is possible that this lower growth yield in high L-arginine is due to the pH of the culture supernatant during growth. *S. gordonii* cells are known to produce large amounts of ammonia as a by-product of arginine catabolism by the ADS during long-term growth (Cunin *et al.*, 1986; Liu and Burne, 2009). This ammonia can build up and potentially could cause retardation of growth and cell toxicity. As mentioned in section 3.2.1, the pH of the bulk fluid of the bacterial culture was measured for alkalinity – a large difference was found between the 500 mM and 1 mM samples (pH 5.9 vs pH 4.5). However, it was still not strongly alkaline pH, and was unlikely to be high enough in the bulk fluid to inhibit growth. It may be that the intracellular pH of the cells became strongly alkaline, as opposed to the bulk fluid, and this could have caused suppression of growth in high arginine. Intracellular pH is known to fluctuate widely in different

types of bacteria, in response to extracellular pH, sometimes across as wide a range as from pH 6 to 9 (Booth, 1985; Breeuwer *et al.*, 1996). It would be interesting to assess the intracellular pH of *S. gordonii* whilst growing in the above conditions, using a fluorescent pH-sensitive dye such as carboxyfluorescein succinimidyl ester (cFSE) (Breeuwer *et al.*, 1996), or a pH-sensitive variant of a fluorescent protein such as GFP or YFP (Olsen *et al.*, 2002; Wilks and Slonczewski, 2007).

When incubated in high arginine concentrations in cell-free 25% saliva, the *S. gordonii* cells did not appear to grow. The cells remained viable up to 48 h after initial incubation, as demonstrated by the viable counts shown in Figure 3.2, but the cell yield did not increase over this time. The slight decrease in viability over this time period was likely due to natural cell death. The reason as to why the *S. gordonii* cells did not grow in saliva was likely due to the initial inoculum being too high – at high cell densities, e.g. 1 x 10^7 cells mL^{-1}, the nutrient concentration in the saliva is too low to support multiple rounds of cell division, and so cells do not grow (Palmer *et al.*, 2001). Lower starting inoculi allow enough nutrients to support repeated cell division. *S. gordonii* has been shown to have the ability to grow both planktonically and in biofilms, and the paper by Palmer *et al.* (2001) indicates that a starting inoculum of around 1 x 10^4 cells mL^{-1} would be a suitable starting inoculum in order to assess the impact of 500 mM L-arginine on growth in saliva. The high arginine concentration also prevented aggregation between the bacterial cells. This could potentially cause an impact on the viable cell counts, which may increase when bacterial auto-aggregates are dispersed.

S. gordonii is known to auto-aggregate when grown in saliva as a medium, due to the adherence of the cells to certain salivary glycoproteins and low-molecular-weight mucins (Ligtenberg *et al.*, 1992; Murray *et al.*, 1992). The inhibition of *S. gordonii* auto-aggregation in high L-arginine conditions has not been noted before. However, L-arginine has been shown to affect co-aggregation in *Fusobacterium nucleatum,* with the arginine-inhibitable RadD adhesin being responsible for *F. nucleatum* binding to Gram-positive early-coloniser species, such as *S. gordonii* (Kaplan *et al.*, 2009). Additionally, L-arginine is regularly used in industry and bioscience at a high concentration to prevent protein aggregation, and aid in denaturation and solubilisation of proteins (Tsumoto *et al.*, 2004; Ishibashi *et al.*, 2005; Arakawa *et al.*, 2007). The

lack of auto-aggregation between cells in the 500 mM samples is therefore likely due to prevention of protein interactions on the bacterial cell surface, mediated by L-arginine.

The conditional auxotrophy for L-arginine, L-histidine and the branched-chain amino acids displayed by *S. gordonii* cells in section 3.2.3 is an interesting result due to the fact that *S. gordonii* apparently has all necessary genes in the biosynthetic pathways for all of these amino acids, according to the KEGG database (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg-bin/show_pathway?sgo01230 (Kanehisa and Goto, 2000)). However, it is possible that *S. gordonii* is in fact fully auxotrophic for L-histidine and the branched-chain amino acids. Therefore, future experiments would need to test whether *S. gordonii* can biosynthesise these amino acids when they are absent from the growth medium, and under which conditions it is able to do this.

It has been shown in *Staphylococcus aureus* isolates, that carbon catabolite repression (CCR, via the *ccpA* gene) contributed to repression of L-arginine biosynthesis in chemically-defined medium depleted of L-arginine, causing conditional auxotrophy (Nuxoll *et al.*, 2012). Addition of secondary, non-preferred sources of carbon in the growth medium alleviated the effects of CCR, and allowed growth of *S. aureus* cells in L-arginine-depleted CDM. It is possible that CCR is contributing to the arrest of *S. gordonii* growth in L-arginine-depleted medium also, since the CcpA protein is known to co-regulate L-arginine catabolism in *S. gordonii* (Dong et al., 2002). Therefore, it would be useful for future studies to focus on investigating the CCR mechanism in *S. gordonii* – for example, by addition of a secondary carbon source in the L-arginine-deplete growth medium, in order to see whether that alleviates the auxotrophic phenotype as with *S. aureus*. Alternatively, by engineering a *ccpA*-knockout strain of *S. gordonii*, and assessing whether this strain has the ability to grow in arginine-deplete medium, where the wild-type *S. gordonii* strain does not.

The impact of amino acid depletion on gene regulation appeared to show some quite specific effects, with some of the genes assessed being clearly regulated in response to arginine only, and others being equally regulated across all three amino acid depletion samples. For example, the *hsa* gene in particular was equally lowly-regulated across all the –arg, -hist and –BCAA samples, as were *bfbC* and *bfbF*. The *hsa* gene encodes part of the Hsa surface protein, and low

levels of regulation of this gene observed in the qRT-PCR reactions in this chapter may be due to natural variation in gene expression levels. However, it is also possible that any differences in correlation between the microarray analysis and qRT-PCR analysis, as for this gene, are due to small variations in microarray specificity as discussed in section 3.3.2 above.

The *bfb* genes are involved in the cellobiose phosphotransferase system, which imports cellobiose into the bacterial cell. According to the qRT-PCR data (shown in Table 3.2 and Figure 3.8), the *hsa* gene is slightly down-regulated in response to amino acid depletion. The *bfb* genes, however, are strongly down-regulated. This may be in response to the bacterial cells undergoing a stress response in reaction to the depletion of any of the amino acids tested.

The *bfb* genes are involved in actively transporting nutrients – in this case, the sugar cellobiose – into the *S. gordonii* cells. However, they have also been shown to be switched on during biofilm formation in *S. gordonii* (Kiliç et al., 2004). As the depletion of amino acids in the growth medium puts the cells into a global stress response, any genes involved in carbohydrate transport would likely be down-regulated in order to conserve energy within the cells. A similar response can be observed in *E. coli* in response to amino acid starvation – namely, a global stress response mediated by the alarmones (p)ppGpp (guanosine 5′,3′ bispyrophosphate and guanosine pentaphosphate) (Traxler *et al.*, 2008). The same (p)ppGpp-mediated response can also be observed in *Bacillus subtilis* in response to iron-deprivation, causing up-regulation of amino acid biosynthesis pathways – in this case, in order to synthesise the iron chelating molecule bacillibactin, so more iron can be taken up from the external environment (Miethke *et al.*, 2006).

It is possible that a number of the genes seen to be regulated equally across all strains here may be part of a global stress response, potentially mediated by (p)ppGpp in *S. gordonii*. Future work to determine if this is the case, and which genes may be affected in their expression by this global stress response, could focus on deletion of the genes which synthesise (p)ppGpp – in this case, the *relAPQ* genes, well-known in *S. mutans* (Lemos et al., 2007), and also present in *S. gordonii*. Knocking out these genes would prevent the production of the alarmone molecules, and would allow analysis of which genes are still affected in their expression in response to amino acid depletion, and which remain unchanged. This would help to establish

the regulatory circuitry for the global stress response in *S. gordonii*. Another alternative would be to test the expression of these genes in response to the depletion of other amino acids – if the expression levels remain constant irrespective of which amino acid has been depleted, that would indicate a global stress response.

Of the genes that are apparently specific to arginine, some were expected – such as *argC* – and some were not, for example, in the case of *wzg*, *asp5* and SGO_1686. These genes were unexpectedly regulated specifically in response to arginine. The *asp5* gene was particularly unexpected, as that encodes part of the *hsa* operon. As Hsa is a surface adhesin, involved in binding to fibronectin residues in the human oral cavity, the genes in this operon may not need to be expressed whilst the bacterial cells were growing in planktonic culture in defined medium, and so expression of these genes may be effectively switched off in response to amino acid deprivation. In the case of *asp5*, the expression of this gene was strongly down-regulated, but only in response to arginine depletion. This was in contrast to the *hsa* gene itself, which was only slightly down-regulated. It is possible, therefore, that instead of these genes being part of the same operon, they are part of different operons and as such are regulated and expressed differently. Alternatively, the presence of internal promoters that are regulated by different sigma factors, such as those found in *E. coli* (Shimada et al., 2014; Peano et al., 2015), could be responsible for differential expression of these genes.

No evidence has been found as yet to suggest that L-arginine, or in fact any other amino acid, has any bearing on the biosynthesis or expression of the Hsa protein in *S. gordonii*, however the strong down-regulation of this gene in the arginine-deplete sample suggests that arginine, or one of its regulators, is playing an indirect role in the regulation of the *asp5* gene. It may be that one of the arginine regulatory molecules – ArcR, ArgR or AhrC – may be affecting the expression of a gene that in turn regulates *asp5* expression, in response to arginine depletion. However, this would have to be investigated further before any definite regulatory mechanisms could be proposed.

Specific regulation in response to arginine depletion was also observed in the *wzg* and SGO_1686 genes. SGO_1686 is involved in fatty acid biosynthesis, and whilst down-regulation of this gene might be expected in a bacterial stress response to nutrient depletion, in order to

conserve energy by preventing unnecessary biosynthesis, the down-regulation of this gene specifically in response to arginine depletion is unexpected. The *wzg* gene is part of an operon that encodes biosynthesis of the receptor polysaccharide (RPS) surface protein, which mediates binding between *S. gordonii* and *A. oris* (Cisar et al., 1995; Mishra et al., 2010). This gene was more strongly down-regulated in response to arginine depletion than to histidine or branched-chain amino acid depletion – a four-fold down-regulation in arginine versus a two-fold down-regulation in histidine. However, this still shows a slight degree of arginine-specific down-regulation. As for the *asp5* gene above, it is possible that for any of these genes, rather than arginine having a direct effect on expression of these genes, it is instead playing an indirect role in the regulation of these genes through a downstream signalling cascade of some kind. Direct signalling was not assessed in these experiments, and could form the basis for future work. However, the fact that both *asp5* and *wzg* are part of operons encoding major surface proteins in *S. gordonii* should not be overlooked, as this could play a role in affecting biofilm formation in *S. gordonii*.

Therefore, the work in this chapter has demonstrated that, whilst high concentrations (500 mM) of L-arginine do not kill *S. gordonii* cells, they inhibit planktonic growth. In addition to this, planktonic cultures of *S. gordonii* have been shown to exhibit a functional auxotrophy for L-arginine. Analysis of the regulation of certain genes in response to the depletion of these amino acids demonstrated the presence of some arginine-specific responses in gene regulation, and some genes that showed changes in expression in response to depletion of any amino acid. The next chapter of this thesis will take this analysis of gene regulation in response to L-arginine one step further, by looking at the role that the different putative arginine-dependent regulators play in controlling gene expression in *S. gordonii*.

4. Role of transcription regulators in arginine-dependent gene regulation

4.1. Introduction

Chapter 3 looked in detail at the arginine-dependent regulation of 12 genes of interest, specifically chosen for their roles in arginine biosynthesis and biofilm formation, and described in detail in the previous chapter's introduction.

The microarray analysis which formed the rationale for choosing those 12 genes was recently analysed and published in a paper from our group (Jakubovics *et al.*, 2015). The analysis within this paper focused on the responses of wild-type *S. gordonii* DL1 to a shift from high to no arginine conditions. This microarray analysis was performed alongside an equivalent analysis of *S. gordonii* Δ*arcR*, which carries a deletion of the gene which encodes one of the three *S. gordonii* arginine-dependent regulators, ArcR. The expression of 2051 genes was also measured in this strain in high arginine and in a shift from high to no arginine conditions. Unlike the *S. gordonii* DL1 data, the data from *S. gordonii* Δ*arcR* have not yet been published, nor analysed in detail. Therefore, this chapter will focus on discussing the results of this *S. gordonii* Δ*arcR* microarray. In addition, new microarray analyses were performed here to assess the impact of disrupting the *argR* or *ahrC* genes, encoding the putative arginine-dependent regulators ArgR and AhrC, on *S. gordonii* gene regulatory responses to shifts from high to no arginine.

The original *S. gordonii* DL1 microarray found that, of 2051 genes assessed, covering >95% of the total predicted genome, 464 were differentially expressed in response to L-arginine depletion. These genes represent nearly 23% of the *S. gordonii* genome, and a majority of the genes that can be functionally characterised are associated with amino acid metabolism and transport functions, according to COGFun classifications (functional classifications of clusters of orthologous groups; an approach that classifies genes into general classes based on their function within the bacterial cell). Approximately equal numbers of genes appeared to be up-regulated and down-regulated in response to arginine depletion, with 247 genes increasing in expression (up-regulated) vs 217 decreasing in expression (down-regulated).

Some genes were very strongly regulated in response to shifts in arginine concentration, with 18 loci containing single or multiple genes that were regulated >10-fold. These include up-regulated operons involved in arginine and pyrimidine biosynthesis, histidine biosynthesis, and a number of hypothetical genes that are uncharacterised. The extent of regulation of genes within these operons extended from around 10-fold up-regulation, to as high as 339-fold. The operons that were down-regulated include the Hsa and receptor polysaccharide (RPS) surface protein biosynthesis operons, in addition to the fatty acid biosynthesis and cellobiose phosphotransferase (PTS) system genes. The extent of down-regulation in these operons was between 5-fold to 30-fold down-regulation.

The work in Chapter 3 aimed to differentiate between arginine-specific gene regulation and gene regulation that was associated with general stress induced by growth arrest following shifts to no arginine. In many bacteria, direct arginine-dependent regulation is mediated by regulators of the ArgR/AhrC family (Cunin *et al.*, 1986). The *S. gordonii* genome encodes three putative arginine-dependent regulators and it has been shown that these each participate in regulation of arginine metabolism genes (Jakubovics *et al.*, 2015). The microarray analysis in this chapter aimed to determine which genes are regulated by each of these arginine-dependent regulators, and whether any genes are regulated specifically in response to arginine in an ArcR/ArgR/AhrC-independent manner.

4.2. Confirmation of the non-polarity of arginine regulator mutants

4.2.1. qRT-PCR analysis of downstream gene expression

Arginine was seen to have an effect on expression of a large number of genes, including genes associated with biofilm formation (see previous chapter). Here, the role of three arginine-dependent regulators in gene regulation was analysed using isogenic mutants. Previous work in our lab had resulted in construction of deletion strains of all three arginine regulators, encoded by the genes *argR*, *arcR* and *ahrC*, engineered together in combinations of single, double and triple mutant strains. However, the possible polar effects of these gene disruptions had not previously been investigated, and the work here aimed to assess downstream gene expression to ensure that the mutations did not have polar effects on surrounding genes. The work described below was included in the publication by Jakubovics *et al.* (2015).

Deletions of *S. gordonii argR*, *arcR* and *ahrC* genes were previously engineered by replacement of the regulator gene with an antibiotic resistance cassette, thereby knocking out the function. However, all of these regulators are found <200 bp upstream of other genes, and the surrounding regions for each are shown in Figure 4.1. The *argR* gene lies 178 bp upstream of *mutS*, the 3' end of the *arcR* gene is just 34 bp away from the oppositely oriented *arcT* gene, and *ahrC* is just 7 bp upstream of *recN*.

In order to ensure that downstream gene expression was not being affected by insertion of antibiotic-resistance cassettes, qRT-PCR was carried out on three replicates of cDNA samples taken from the *argR* and *ahrC* mutant strains (PK3346 and PK3350 respectively), grown anaerobically in BHY rich medium. Primers were designed to target the gene downstream of each regulator gene, namely the DNA mismatch repair gene *mutS* in the *S. gordonii* Δ*argR* strain, and the DNA repair gene *recN* in the *S. gordonii* Δ*ahrC* strain. Microarray analysis was also performed on these genes, within the relevant mutant strains, as part of the work in this chapter. In the case of *arcR* (strain PK3354), expression of the downstream hypothetical *arcT* gene had already been measured as part of the previous microarray analysis (see section 4.3.1 below).

The absolute fold-change values for the genes of interest were assessed by qRT-PCR, and standardised against the levels of gene expression in the *S. gordonii* DL1 strain, to give a fold-change value of mutant vs wild-type. It was found that gene expression for the *mutS* (1.26-fold up-regulation in comparison to DL1) and *recN* (1.09-fold down-regulation in comparison to DL1) genes was not significantly affected by deletion of the preceding regulatory gene. Analysis of the expression levels of these genes from the *S. gordonii* Δ*argR* and Δ*ahrC* microarrays carried out in this chapter also showed that each of these genes were regulated <2-fold. The expression of the *arcT* gene was significantly different, with a 5.21-fold up-regulation observed. ArcR has not previously been shown to regulate *arcT*, which is predicted to encode a transpeptidase enzyme that can cleave arginine residues from internalised peptides (Zúñiga *et al.*, 1998; Liu *et al.*, 2008). However, due to the putative role of ArcT and its position directly downstream of the *arcD* gene, it is likely that this gene is co-transcribed with *arcD* and linked to the arginine catabolism genes, cleaving arginine residues for conversion to ammonia.

Therefore, it is possible that ArcR would also regulate expression of the *arcT* gene, in addition to *arcABCD,* and so it is likely that the up-regulation of the *arcT* gene is due to the loss of regulation of the *arc* operon by the ArcR transcriptional regulator, rather than due to a downstream polar effects.

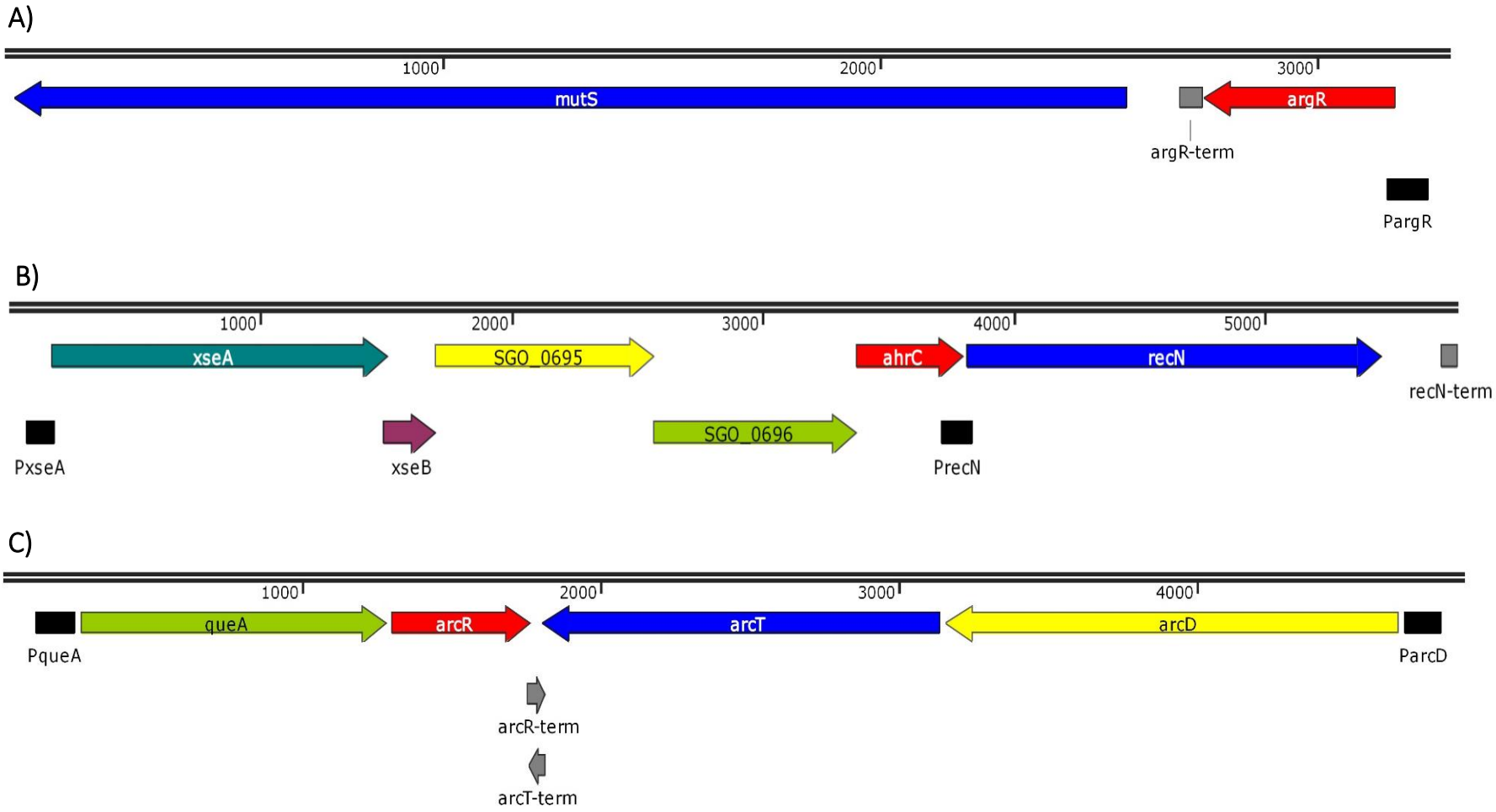


Figure 4.1. Positions of regulatory genes within their operons, and downstream genes. A) shows the *argR* gene, with its putative promoter (labelled P_{argR}), and terminator (*argR*-term). The *mutS* gene lies downstream of this, and no predicted promoter was found between *argR* and *mutS*. B) shows the *ahrC* gene, which is apparently co-transcribed with the genes *xseA*, *xseB*, SGO_0695 and SGO_0696, all from the promoter P_{xseA}. The *recN* gene is immediately downstream, with its own predicted promoter (P_{recN}) lying within the *ahrC* gene itself, and a terminator lying downstream. C) shows the *arcR* gene, co-transcribed from P_{queA} with the *queA* gene. The genes *arcDT*, transcribed together from another promoter, lie in the opposite direction, with both *arcR* and *arcT* sharing a terminator sequence. (Promoters predicted using PromBase, terminators predicted using DOOR[2]). Map created using the SnapGene software, numbers on DNA strand represent the scale in base pairs.

4.3. Roles of ArcR, ArgR and AhrC in arginine-dependent gene regulation

After establishing that the arginine regulator knock-out strains were not having polar effects on downstream gene expression, the impact of these mutations on genome-wide gene expression in response to changes in exogenous arginine was assessed.

Genome-wide microarray analysis was performed on single knock-out strains of *argR* and *ahrC*. *S. gordonii* Δ*argR* (PK3346) and Δ*ahrC* (PK3350) cells were grown in arginine-replete chemically-defined FMC medium, then transferred either to medium replete in or depleted of arginine. RNA was extracted after 30 min, and gene expression was assessed by microarray. Data were then compared with microarray analyses of *S. gordonii* DL1 and Δ*arcR* (PK3354), subjected to the same arginine depletion protocols. It is important to note that the data between the sets of arrays may not be comparable, as the wild-type was not included on the *S. gordonii* Δ*argR* and Δ*ahrC* arrays due to cost, and so it was hoped that the new arrays would be comparable with the previous ones. However, an important part of the bioinformatics analysis was to determine whether a comparison between the experiments that were done at different times was reasonable.

4.3.1. Genome-wide microarray of S. gordonii *Δ*arcR *gene expression in high and no arginine*

As mentioned in section 4.1, the *S. gordonii* Δ*arcR* microarray analysis was carried out at the same time as the original *S. gordonii* DL1 microarray, published in Jakubovics *et al.* (2015), however the Δ*arcR* microarray data has not yet been published or analysed in detail until now.

In the *S. gordonii* Δ*arcR* strain, 576 genes were differentially-regulated in response to a shift from high to no arginine conditions. Of these genes, 385 were up- or down-regulated >2.5-fold in response to arginine depletion. Amongst these genes were amino acid (including arginine) biosynthesis genes, receptor polysaccharide and Hsa biosynthesis genes, and genes from the *bfb* biofilm-associated locus. However, only direct comparison of the data from the *S. gordonii* DL1 and Δ*arcR* microarrays would indicate which genes were being directly regulated by the ArcR protein.

Direct comparisons between the *S. gordonii* DL1 and Δ*arcR* strains under both high, and no arginine conditions, were then made. Under no arginine conditions, no genes were significantly differentially expressed between the two strains. However, under high arginine conditions, five genes (excluding *arcR* itself) were shown to be differentially regulated in the *S. gordonii* Δ*arcR* strain vs DL1 (Table 4.1). These genes are therefore thought to be directly regulated by the arginine-dependent ArcR regulator.

These genes consisted of *queA*, *argG, argH*, SGO_0177 and SGO_0846. The *queA* gene is found immediately upstream of *arcR*, and co-transcribed with the *arcR* gene from its own promoter. This gene was up-regulated in the *S. gordonii* Δ*arcR* strain under high (and no) arginine conditions when compared to the wild-type (11-fold up-regulated in the mutant vs wild-type). It has been suggested that QueA is responsible for negatively regulating the expression of the *arc* operon genes (Liu *et al.*, 2008; Liu *et al.*, 2012), and so may be controlled in its expression by the ArcR regulator via the p_{queA} promoter. This would also be supported by the fact that deletion of the *arcR* gene results in up-regulation of this gene, under both high and no arginine conditions. The arginine biosynthesis genes *argGH*, usually repressed slightly by the ArcR protein under high concentrations of exogenous arginine, were also up-regulated within the Δ*arcR* mutant, as was the SGO_0177 gene, thought to be co-transcribed with *argGH* (all 5 to 6-fold up-regulated in mutant vs wild-type).

However, the most surprising find was the strong up-regulation of the SGO_0846 gene. This gene was 148-fold up-regulated in the *S. gordonii* Δ*arcR* strain in comparison to the DL1 strain, under high arginine conditions, and was not significantly regulated in either the *S. gordonii* Δ*argR* or Δ*ahrC* strains under high arginine conditions. It is labelled as a putative cell wall protein by NCBI Gene (NCBI, 2015), but is otherwise wholly uncharacterised. Analysis of this gene, and its possible role within *S. gordonii* cells, will form the basis of the work undertaken in Chapter 5.

Gene name	Function	Absolute fold-change
SGO_0846	Hypothetical	148.66
queA	Queosine modification of tRNA	11.05
argH	Arginine biosynthesis	6.54
argG	Arginine biosynthesis	5.97
SGO_0177	Arginine biosynthesis (hypothetical)	5.58
arcR	Transcriptional regulator	-73.54

Table 4.1. Genes that were differentially expressed between the *S. gordonii* Δ*arcR* vs wild-type strains under high arginine conditions. Negative numbers represent down-regulated genes. The *arcR* gene showed down-regulation in the Δ*arcR* mutant due to loss of its expression in high arginine conditions.

4.3.2. Principal component analysis of microarray datasets

As all four microarray analyses – on wild-type *S. gordonii* DL1, Δ*arcR*, Δ*argR* and Δ*ahrC* strains – were carried out at different times, with no mixing of samples between arrays, it was prudent to check the data for a batch effect - a form of technical variance within an experiment - between the microarrays before carrying out any direct analysis between the datasets. The wild-type DL1 and Δ*arcR* microarrays were performed together, and the Δ*argR* and Δ*ahrC* microarrays were performed together. Principal component analysis was then performed to assess whether the different sets of arrays were comparable, in order to perform direct comparisons and statistical tests across the two sets of array data.

The principal component analysis, performed by Mr John Casement of the Bioinformatics Core Facility (Newcastle University), measured the amounts of variability between the two microarray datasets (Figure 4.2).

The samples analysed within the same microarray experiments (i.e. samples on slides 1 and 2, and slides 3 and 4), are mixed well with each other. However, there is no mixing of samples between the two different microarrays, which indicates a significant batch effect is present within the data. It has not been possible to correct for this using parametric statistical tests, and therefore direct comparisons of gene expression between the *argR* and *ahrC* mutant samples, and the DL1 wild-type, are not feasible. Even so, a limited comparison of data across different arrays is presented in section 4.3.5 below.

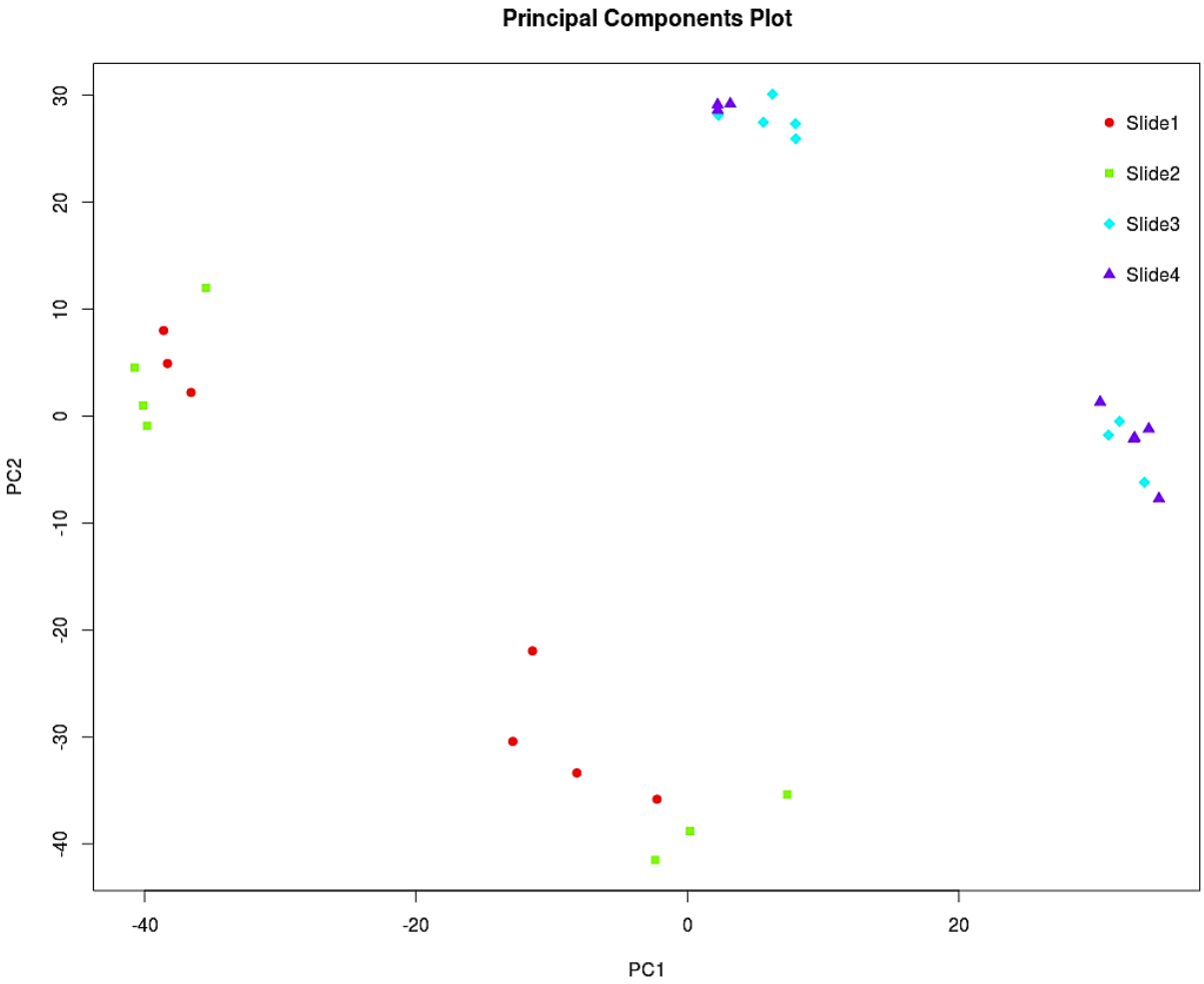


Figure 4.2. Principal component analysis (PCA) plot of data from all four microarray experiments. The axes of the graph – labelled PC1 and PC2 – represent the two greatest directions of variability within the data, with PC1 showing the direction of greatest variability, and PC2 the direction of greatest variability within the constraints of the first direction. Each data point on the graph represents one of the samples on one of the microarrays, thirty-two samples in total (four biological replicates of four strains, each under two conditions), although two samples were removed as outliers and were not included in any of the analyses. Slides 1 and 2 represent the original wild-type DL1 and Δ*arcR* microarrays, and Slides 3 and 4 represent the Δ*argR* and Δ*ahrC* microarrays. No mixing of samples occurs between the first and second microarrays, indicating a significant batch effect between the two sets of data. (PCA analysis performed by Mr John Casement, of the Bioinformatics Core Facility, Newcastle University).

4.3.3. Comparison of microarray with qRT-PCR analysis of gene expression in the arginine-dependent regulator mutants

Following discovery of the batch effect between the two microarrays, it was important to validate the results of each microarray, in order to ensure the datasets were similar, if not directly comparable. In the paper by Jakubovics *et al.* (2015), qRT-PCR analysis was performed in the wild-type strain on seven different genes – six of which were involved in arginine metabolism and transport (*argC, argG, pyrA$_b$, arcA, arcB,* and *arcD*), and the seventh as a control gene (*amyB*). However, qRT-PCR analysis was also performed at the same time on the same genes in the Δ*arcR*, Δ*argR* and Δ*ahrC* strains (Jakubovics *et al*., unpublished work). Therefore, in order to validate the findings of the Δ*arcR,* Δ*argR* and Δ*ahrC* microarrays, the qRT-PCR data for the above-mentioned genes was compared against the microarray values for each strain. Each microarray was validated separately using this technique, as the variability of the data due to the batch effect meant they could not be validated together.

The $\log_2$ fold-change values for expression of each of these seven genes, in wild-type *S. gordonii* DL1 and each of the arginine-dependent regulator mutant strains, were compared between the qRT-PCR and microarray, for a shift from high to no arginine conditions (Figure 4.3 and Figure 4.4).

The comparison showed a strong correlation between the qRT-PCR and microarray analysis for these genes in all four strains. The r^2 values of the linear regression lines were over 0.9 for each strain, with the exception of the *S. gordonii* Δ*arcR* strain, where the r^2 value was only 0.777. For the wild-type DL1 strain, the slope of the line was 0.985, and for the other three strains the value of the slope was over 1.1. On the whole, the values for gene expression in the microarray and qRT-PCR correlated well for all strains, despite some slight differences in the extent of regulation. Therefore, the microarray data were comparable with the qRT-PCR analyses.

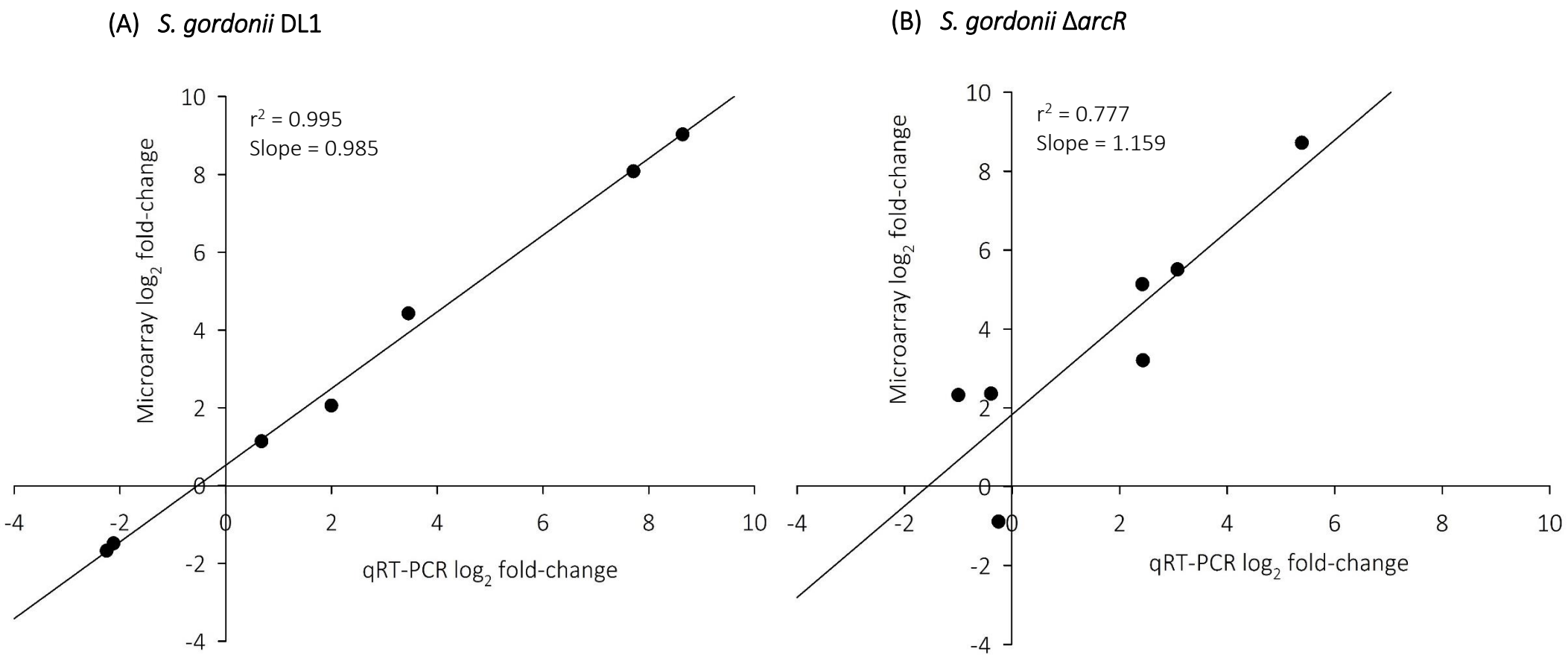


Figure 4.3. Validation of the *S. gordonii* DL1 and Δ*arcR* microarrays by comparison with qRT-PCR data. Gene expression was measured as a $\log_2$ fold-change value when moving from high to no arginine conditions, in both the wild--type *S. gordonii* DL1 (A) and Δ*arcR* (B) strains. Seven genes were analysed by qRT-PCR analysis, and correlated against the fold-change values for the same genes within each individual microarray. The linear regression line for each correlation is shown. Values on the graphs indicate the r^2 value and the slope of the line.

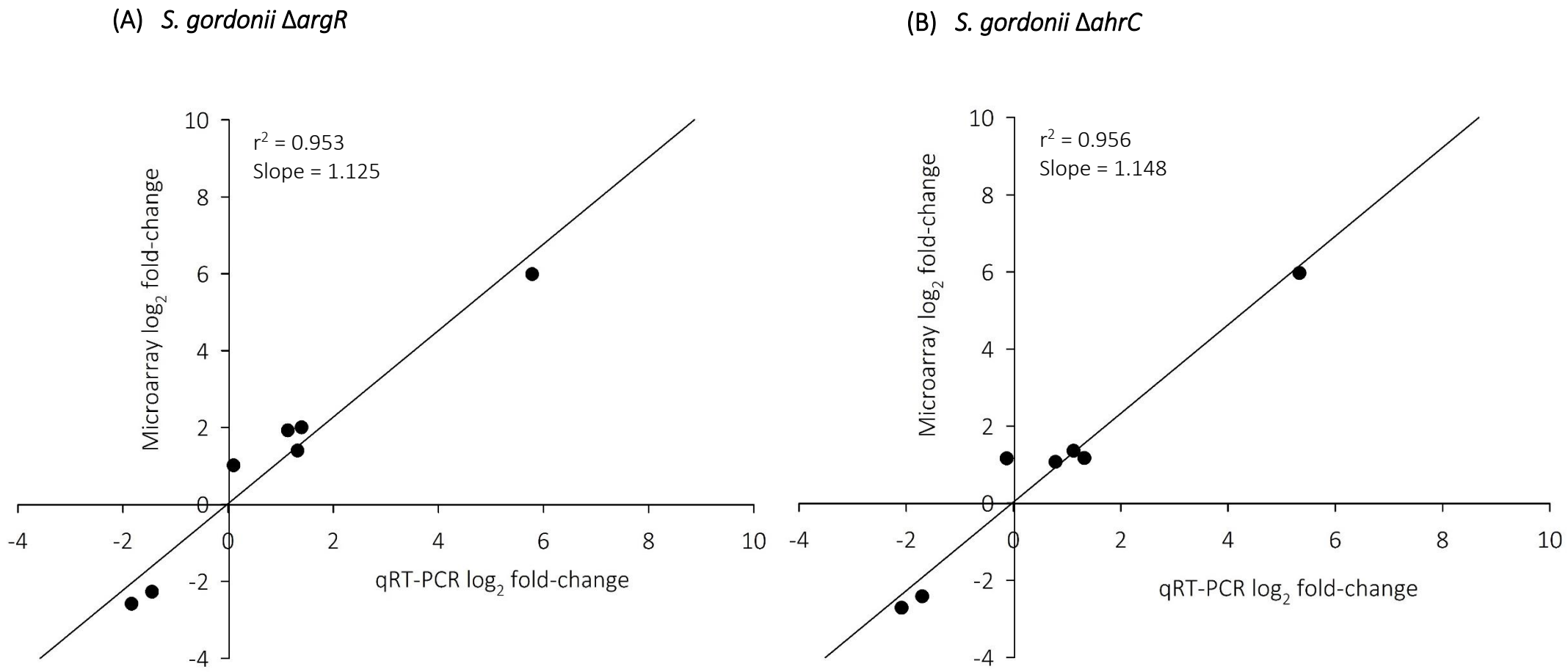


Figure 4.4. Validation of the *S. gordonii* Δ*argR* and Δ*ahrC* microarrays by comparison with qRT-PCR data. Gene expression was measured as a $\log_2$ fold-change value when moving from high to no arginine conditions, in both the *S. gordonii* Δ*argR* (A) and Δ*ahrC* (B) strains. Seven genes were analysed by qRT-PCR analysis, and correlated against the fold-change values for the same genes within each individual microarray. The linear regression line for each correlation is shown. Values on the graphs indicate the r^2 value and the slope of the line.

4.3.4. Genome-wide microarrays of gene expression in S. gordonii *Δ*argR *and Δ*ahrC *in high and no arginine*

As the data from the *S. gordonii* Δ*argR* and Δ*ahrC* strains were shown to be comparable to the qRT-PCR data for seven genes of choice, further analysis of the data from these microarrays can now be undertaken. However, no direct comparisons can be drawn between these two mutants and the wild-type strain, as with the *S. gordonii* Δ*arcR* strain.

In *S. gordonii* Δ*argR*, 347 genes were regulated in response to a shift from high to no arginine growth medium. By looking at genes that were strongly regulated in response to arginine in *S. gordonii* Δ*argR*, it was possible to identify loci that were regulated in response to arginine by factors other than ArgR (Table 4.2).

Three of the most highly up-regulated genes were the arginine biosynthesis genes, *argG* (63-fold up-regulated in response to arginine depletion), *argH* (53-fold up-regulated) and the hypothetical protein SGO_0177 (up-regulated 78-fold). These genes are predicted to be co-transcribed, and were apparently regulated in response to arginine by the ArcR protein, which is still encoded by the intact *arcR* gene in the Δ*argR* strain. However, other arginine biosynthesis genes, including *argC* and *argJ*, which were the most highly up-regulated genes in response to arginine depletion in the wild-type (approx. 500-fold and 300-fold up-regulation respectively), were only two-fold up-regulated in response to arginine depletion in *S. gordonii* Δ*argR*. All of these genes, *argC, argJ,* and *argGH*, have been shown to be regulated by the ArgR and AhrC arginine-dependent regulators (Jakubovics *et al.*, 2015).

Up-regulated genes	Absolute fold-change	Down-regulated genes	Absolute fold-change
SGO_0177	78.53	*bfbB*	-101.81
argG	63.27	*bfbG*	-81.68
argH	53.09	*bfbF*	-54.98
SGO_0648	9.75	*bfbR*	-29.56
SGO_0646	8.01	*bfbD*	-27.21
SGO_0647	7.77	*bfbA*	-27.16
SGO_0645	7.40	*bfbC*	-27.01
SGO_1277	6.10	SGO_1575	-13.07
nadR	4.69	SGO_1700	-8.95
SGO_0874	4.51	SGO_1686	-8.49

Table 4.2. Most highly up- and down-regulated genes in the *S. gordonii* Δ*argR* strain, in high vs no arginine conditions. Negative numbers represent down-regulated genes.

Other highly up-regulated genes include a number of hypothetical genes, such as SGO_0645 (seven-fold up-regulated) and SGO_0646 (eight-fold up-regulated). The neighbouring genes SGO_0647 and SGO_0648 encode a DNA-directed DNA polymerase IV enzyme, and a DNA binding protein respectively. They are predicted by the DOOR^2 database (Mao *et al.*, 2009) to form an operon, along with the aforementioned SGO_0645 and SGO_0646, and so it is likely that the roles that the hypothetical genes in this operon play may be related to the roles performed by the two characterised genes within the operon, i.e. in DNA replication and mutagenesis, akin to the known roles of DNA polymerase IV (Wagner *et al.*; Ohmori *et al.*, 2001). The DNA polymerase IV protein is also known to be regulated in response to stress in *E. coli* (Layton and Foster, 2003), and so may be regulated as part of a general stress response to arginine depletion in these samples.

The remaining up-regulated genes shown in Table 4.2 are SGO_0874, a putative lipoprotein-encoding gene, and *nadR*, a nicotinamide-nucleotide adenylyltransferase transcriptional regulator (NCBI, 2014), which were both up-regulated four-fold in response to arginine depletion. The *nadR* gene was regulated to a similar levels across all four microarrays (wild-type *S. gordonii* DL1, Δ*arcR*, Δ*argR*, and Δ*ahrC*), so seems likely to be a general stress response. However, the SGO_0874 gene was regulated to a similar level in all three arginine-dependent regulator mutants, but was more strongly up-regulated in the wild-type strain. This may suggest a degree of control, either direct or indirect, over expression of this gene by the arginine regulators.

The most strongly down-regulated genes include the *bfb* operon genes, which encode components of a cellobiose phosphotransferase system, previously linked to biofilm formation in *S. gordonii* (Kiliç *et al.*, 2004; Jakubovics *et al.*, 2008). The *bfb* locus consists of the genes SGO_1575, and *bfbABCDFGR*, which were down-regulated in response to arginine depletion between 13-fold and 101-fold. These genes appeared to be more strongly down-regulated in both the *S. gordonii* Δ*argR* and Δ*ahrC* mutants, and less strongly down-regulated in the Δ*arcR* strain, in comparison to the levels of regulation in the wild-type. This may suggest that the ArgR/AhrC proteins play a role in limiting the down-regulation of these genes, and once those regulators are deleted the expression level drops further. By contrast, ArcR may directly down-regulate these genes, as deletion of *arcR* relieved down-regulation to some extent.

Finally, the remaining two down-regulated genes, SGO_1686 and SGO_1700, both encode hypothetical proteins.

Of the 374 genes regulated in response to arginine depletion in *S. gordonii* Δ*argR*, twelve were found to be unique to the Δ*argR* strain (Table 4.3). This list included the chromosomal segregation protein *spo*0J (2.11-fold up-regulated in a shift from high to no arginine), and a predicted membrane protein (SGO_0917, 2.21-fold up-regulated). All 12 of these genes appeared to be weakly-regulated, with absolute fold-change values consisting of <2.5-fold changes up or down in gene expression.

In the *S. gordonii* Δ*ahrC* strain, 366 genes were regulated in response to a shift from high to no arginine conditions. The most strongly up- and down-regulated genes in this strain were the same as those found in the *S. gordonii* Δ*argR* strain (Table 4.4). The Δ*ahrC* strain also had 12 genes that were found to be uniquely regulated in that strain (Table 4.5).

Gene name	Gene function	Δ*argR* absolute fold-change values
cadD	Cadmium transporter	2.24
SGO_0917	Membrane protein	2.21
spo0J	Chromosome segregation protein	2.11
dnaI	Primosomal protein	2.05
SGO_1507	Hypothetical	2.03
pfl	Formate acetyltransferase	2.02
fruR	Phosphotransferase system repressor	2.00
SGO_1571	UDP-N-acetylenolpyruvoyl-glucosamine reductase	-2.02
rplM	50S ribosomal protein L13	-2.12
SGO_0061	Type VII secretion protein	-2.17
SGO_0062	Hypothetical	-2.17
rpsJ	30S ribosomal protein S10	-2.18

Table 4.3. Genes regulated in the *S. gordonii* Δ*argR* strain only, in response to a shift from high to no arginine conditions. Negative numbers represent down-regulated genes.

Up-regulated genes	Absolute fold-change	Down-regulated genes	Absolute fold-change
SGO_0177	71.72	*bfbG*	-121.43
argG	62.38	*bfbB*	-113.64
argH	55.01	*bfbF*	-68.89
SGO_0648	9.80	*bfbR*	-34.01
SGO_0645	8.93	*bfbA*	-32.03
SGO_1277	8.52	*bfbC*	-30.83
SGO_0647	8.35	*bfbD*	-29.81
SGO_0646	8.32	SGO_1700	-10.95
nadR	5.83	SGO_1575	-10.88
SGO_0874	5.10	SGO_1686	-8.97

Table 4.4. Most highly up- and down-regulated genes in the *S. gordonii* Δ*ahrC* strain, in high vs no arginine conditions. Negative numbers represent down-regulated genes.

Gene name	Gene function	Δ*ahrC* absolute fold-change values
SGO_1664	Membrane protein	2.20
SGO_1457	Membrane protein	2.13
SGO_0050	Hypothetical	2.10
SGO_0479	Transcriptional regulator (AraC family)	2.07
SGO_0894	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	2.03
proV	Proline/glycine transporter	2.02
SGO_0756	Membrane protein	2.01
SGO_0252	Transcriptional regulator (TetR family)	2.01
SGO_0747	5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase	2.01
SGO_1083	Heme ABC transporter ATP-binding protein	-2.05
SGO_1645	Cellobiose phosphotransferase system subunit IIA	-2.08
SGO_0454	Transcriptional regulator	-2.09

Table 4.5. Genes regulated in the *S. gordonii* Δ*ahrC* strain only, in response to a shift from high to no arginine conditions. Negative numbers represent down-regulated genes.

This list of 12 genes included three predicted membrane proteins - SGO_0756, SGO_1457 and SGO_1664, all of which had two-fold changes in response to arginine depletion. However, as with the *S. gordonii* Δ*argR* strain, all of these genes were expressed at levels of <2.5-fold change only. It is likely that in both this strain, and the Δ*argR* strain, the genes that were found to be "uniquely regulated" within each strain (Table 4.3 and Table 4.5) were regulated to a low level in both the Δ*argR* and Δ*ahrC* strains, however, only just reach over the cut-off values for regulation (>2-fold regulation, $p < 0.05$) in one of the strains.

ArgR and AhrC have been posited to work coordinately with each other in *S. gordonii* (Jakubovics et al., 2015), and in other species such as *Lactococcus lactis* (Larsen *et al.*, 2005) and *Streptococcus pneumoniae* (Kloosterman and Kuipers, 2011), although in these species, the regulons of the genes do not overlap precisely, and some genes have been found to be differentially expressed between these two regulators. In this analysis, genes that were found to be regulated to the same levels in both *S. gordonii* Δ*argR* and Δ*ahrC* in response to arginine depletion, but were not regulated to the same degree in the DL1 or Δ*arcR* strains, were also analysed. Thirty-six genes were found, 21 up-regulated and 15 down-regulated in response to a shift from high to no arginine conditions. Again, most (but not all) of these genes were <2.5-fold up- or down- regulated (Table 4.6).

Gene name	Description	Δ*argR* absolute FC	Δ*ahrC* absolute FC
glgP-2	Maltodextrin phosphorylase	3.30	3.21
malQ	4-alpha-glucanotransferase	3.13	3.35
SGO_0579	Amino acid ABC transporter ATP-binding protein	2.52	2.19
msrA	Peptide methionine sulfoxide reductase	2.45	2.63
SGO_1499	Bacteriocin ABC transporter ATP-binding protein	2.42	2.27
hisS	Histidyl-tRNA ligase	2.31	2.37
SGO_1662	DNA methyltransferase	2.29	2.10
SGO_0984	Amino acid permease	2.28	2.04
SGO_2101	Nitrate ABC transporter permease	2.24	2.50
SGO_0578	Amino acid ABC transporter permease	2.23	2.03
pflC	Formate acetyltransferase activating enzyme	2.22	2.44
SGO_1181	Transcriptional regulator (OmpR family)	2.21	2.53
SGO_0291	Copper-translocating P-type ATPase	2.16	2.48
SGO_1933	Copper transporter CopZ-related protein	2.15	2.24
SGO_0269	Hypothetical	2.12	2.35
SGO_0289	Uracil phosphoribosyltransferase	2.11	2.61
SGO_1934	Copper-translocating P-type ATPase	2.11	2.15
SGO_0290	ATPase	2.10	2.36
SGO_0511	Nitroreductase	2.07	2.07
SGO_1180	Histidine kinase	2.07	2.04
SGO_0052	Transcriptional regulator (PadR family)	2.03	2.15
tatA	Twin arginine-targeting protein translocase	-2.02	-2.01
SGO_1084	BCAA ABC transporter permease	-2.04	-2.04
SGO_2030	Acetyltransferase	-2.11	-2.51
gyrA	DNA gyrase subunit A	-2.12	-2.25
SGO_1006	Nif3-like dinuclear metal center hexameric protein	-2.14	-2.04
rnc	Ribonuclease III	-2.18	-2.19
SGO_0063	Hypothetical	-2.22	-2.00
gatC	Aspartyl/glutamyl-tRNA amidotransferase subunit C	-2.28	-2.28
atpC	ATP synthase epsilon chain	-2.36	-2.20
nrdG	Anaerobic ribonucleotide triphosphate reductase activator protein	-2.64	-2.72
rpmG	50S ribosomal protein L33	-2.76	-2.42
adk	Adenylate kinase	-2.83	-2.93
rpmF	50S ribosomal protein L32	-2.94	-2.53
SGO_1243	Pseudo	-4.98	-4.44
SGO_0832	Hypothetical	-6.09	-6.75

Table 4.6. Genes regulated only in the *S. gordonii* Δ*argR* and Δ*ahrC* strains in response to a shift from high to no arginine conditions. Negative numbers represent down-regulated genes. "Absolute FC" = absolute fold-change value for that gene.

This list included genes encoding transcriptional regulators, permease proteins, and a histidine kinase gene. However, the most highly up-regulated genes in a shift from high to no arginine were a maltodextrin phosphorylase gene, *glgP*-2 (3.30-fold vs 3.21-fold up-regulated in *S. gordonii* Δ*argR* vs Δ*ahrC*) and *malQ*, a 4-alpha-glucanotransferase enzyme (3.13-fold vs 3.35-fold up-regulated in *S. gordonii* Δ*argR* vs Δ*ahrC*). The most down-regulated genes were SGO_1243, annotated as a pseudogene by NCBI Gene (4.98-fold vs 4.44-fold down-regulated in *S. gordonii* Δ*argR* vs Δ*ahrC*); and SGO_0832, a hypothetical gene (6.09-fold vs 6.75-fold down-regulated in *S. gordonii* Δ*argR* vs Δ*ahrC*). However, none of these genes appeared to be significantly more strongly regulated in these mutant strains in comparison to the wild-type, although a direct comparison between the two sets of data cannot be made. See section 4.3.5 for an analysis of between-array data.

A number of these 36 regulated genes appear to have generic metabolism or transport roles, or play a part in transcription or translation. It is possible that the changes in regulation of these genes in response to arginine depletion are due to a general stress response, rather than the arginine regulatory genes having a direct effect on their gene expression levels. However, it is clear that these genes are all regulated to a similar level between the *S. gordonii argR* and *ahrC* mutants. Therefore, direct comparison of the *S. gordonii* Δ*argR* and Δ*ahrC* strains identified no genes that were differentially regulated between the two strains under high or no arginine conditions, indicating that in *S. gordonii* the ArgR and AhrC regulons overlap precisely.

4.3.5. Comparison between microarray data from all S. gordonii *strains*

Following the PCA analysis performed in section 4.3.2, no direct statistical analysis of data between the two microarray experiments can be performed. However, fold-change values can still be analysed across the four *S. gordonii* strains. Nevertheless, it is important to note that these fold-change values cannot be directly or statistically compared, as the batch effect may be skewing the data to seem more or less significant than it is.

The fold-change values for the most highly up-regulated (Table 4.7) and down-regulated (Table 4.8) genes (following arginine depletion) from the wild-type *S. gordonii* DL1 microarray were looked at alongside the fold-change values for the same genes within the Δ*arcR*, Δ*argR* and Δ*ahrC* strains.

Gene name	Function	DL1 absolute FC	Δ*arcR* absolute FC	Δ*argR* absolute FC	Δ*ahrC* absolute FC
argC	Arginine biosynthesis	520.12	420.15	2.62	2.55
SGO_0177	Hypothetical	342.42	77.38	78.53	71.72
argJ	Arginine biosynthesis	319.95	325.64	2.13	2.13
argB	Arginine biosynthesis	306.44	340.62	2.25	1.95
argG	Arginine biosynthesis	269.33	45.37	63.27	62.38
argD	Arginine biosynthesis	260.85	315.85	2.16	1.95
argH	Arginine biosynthesis	207.89	39.66	53.09	55.01
SGO_0648	DNA-binding protein	54.12	18.49	9.75	9.80
SGO_1656	Phosphoenolpyruvate carboxykinase	44.43	49.98	3.22	3.29
SGO_0647	DNA polymerase IV	42.30	16.99	7.77	8.35
SGO_0091	Membrane protein	36.62	17.27	NS	NS
SGO_0646	Hypothetical	36.26	14.16	8.01	8.32
SGO_0645	Hypothetical	32.39	16.55	7.40	8.93
SGO_1105	Pyrimidine/arginine biosynthesis	28.94	56.81	3.06	3.53
SGO_1106	Pyrimidine/arginine biosynthesis	28.46	53.12	NS	NS
SGO_0021	Hypothetical	26.67	12.62	4.10	4.83
pyrA$_a$	Pyrimidine/arginine biosynthesis	23.70	28.59	1.99	2.03
pyrA$_b$	Pyrimidine/arginine biosynthesis	21.43	35.02	2.00	2.24
SGO_1102	Pyrimidine/arginine biosynthesis	18.72	24.22	1.83	1.96
SGO_1279	Hypothetical	17.98	6.00	NS	NS
SGO_1410	Histidine biosynthesis	17.01	20.81	3.84	2.89
hisD	Histidine biosynthesis	15.18	16.18	3.56	2.66
hisG	Histidine biosynthesis	15.14	20.38	4.18	3.02
SGO_1278	Hypothetical	14.26	5.42	NS	NS
hisC	Histidine biosynthesis	13.42	16.83	2.89	2.58
hisB	Histidine biosynthesis	12.45	16.95	3.40	2.57
SGO_0092	Molecular chaperone	12.28	5.10	NS	NS
hisH	Histidine biosynthesis	11.15	12.15	2.86	2.13
SGO_0093	Hypothetical	10.64	6.48	NS	NS
SGO_1835	Hypothetical	10.55	4.42	2.03	2.02
SGO_0874	Lipoprotein	10.11	4.56	4.51	5.10

Table 4.7. Strongly up-regulated genes (>10-fold) in the wild-type *S. gordonii* DL1 strain, in response to arginine depletion, compared to the absolute fold-change (FC) values for the same genes in each of the three arginine-dependent regulator mutant strains. Shaded cells indicate genes where regulation is likely within that mutant strain. NS means no signal given for that gene in that microarray.

Gene name	Function	DL1 absolute FC	Δ*arcR* absolute FC	Δ*argR* absolute FC	Δ*ahrC* absolute FC
bfbG	Cellobiose PTS	-81.92	**-33.28**	-81.68	-121.43
bfbB	Cellobiose PTS	-51.14	**-22.44**	-101.82	-113.64
bfbF	Cellobiose PTS	-35.01	**-13.41**	-54.98	-68.89
SGO_1700	Enoyl-CoA hydratase	-30.82	-20.71	**-8.95**	**-10.95**
bfbR	Cellobiose PTS	-24.98	**-8.98**	-29.56	-34.01
bfbA	Cellobiose PTS	-24.43	**-8.96**	-27.16	-32.03
bfbD	Cellobiose PTS	-23.63	**-6.17**	-27.21	-29.81
bfbC	Cellobiose PTS	-21.10	**-6.31**	-27.01	-30.83
SGO_1686	Fatty acid biosynthesis	-17.15	-17.40	**-8.49**	**-8.97**
SGO_1575	Cellobiose PTS	-15.59	**-2.62**	-13.07	-10.88
accC	Fatty acid biosynthesis	-15.18	-9.29	**-4.42**	**-5.44**
SGO_0831	Hypothetical	-13.73	-10.53	NS	NS
SGO_0832	Hypothetical	-13.26	-14.75	NS	NS
wzx	RPS biosynthesis	-12.93	-14.85	-5.80	-6.29
accD	Fatty acid biosynthesis	-12.75	-7.51	**-4.51**	**-5.43**
SGO_1699	Fatty acid biosynthesis	-12.55	-9.27	**-3.54**	**-3.99**
rpsD	Ribosomal protein S4	-12.54	-9.99	**-4.72**	**-3.90**
accA	Fatty acid biosynthesis	-12.33	-8.62	**-4.54**	**-4.93**
wefE	RPS biosynthesis	-12.31	**-23.94**	-6.05	-6.79
SGO_1698	Fatty acid biosynthesis	-11.93	-6.71	-2.85	-3.41
SGO_1692	Fatty acid biosynthesis	-11.32	-5.12	-3.45	-3.96
ileS	Isoleucyl tRNA-synthetase	-11.12	-6.66	-5.04	-4.51
SGO_0976	Hypothetical	-10.81	-8.31	-3.62	-3.53
glf	RPS biosynthesis	-10.69	**-14.56**	-5.50	-5.90
SGO_0977	Hypothetical	-10.50	-8.88	-3.77	-3.86
fabZ	Fatty acid biosynthesis	-10.47	-5.68	-3.30	-3.65
wefA	RPS biosynthesis	-10.44	-8.55	-4.55	-5.69
accB	Fatty acid biosynthesis	-10.16	-5.16	-3.28	-3.56

Table 4.8. Strongly down-regulated genes (>10-fold) in the wild-type *S. gordonii* DL1 strain, in response to arginine depletion, compared to the absolute fold-change (FC) values for the same genes in each of the three arginine-dependent regulator mutant strains. Shaded cells indicate genes where regulation is likely within that mutant strain. Negative values represent down-regulation in gene expression, NS means no signal given for that gene in that microarray.

Comparison of the fold-changes for the most strongly up- and down-regulated genes in the *S. gordonii* DL1 microarray showed some genes with interesting differential expression between the four strains, shaded grey in the tables above. These comparisons are not statistically valid, but do highlight genes of interest. For example, the arginine biosynthesis genes *argBCDJ* all dropped significantly in expression in the Δ*argR* and Δ*ahrC* mutant strains, and comparison of this data with a previous qRT-PCR analysis indicated that this is indeed differential regulation between the strains (see section 4.3.3). The ArgR and AhrC proteins have previously been shown to regulate the expression of these biosynthesis genes in response to low arginine conditions (Jakubovics *et al.*, 2015). In addition, the SGO_0177 and *argGH* genes all appear to be less strongly up-regulated to differing extents in all three mutant strains, when compared to the levels of expression within the wild-type *S. gordonii* DL1 strain. It is likely that all three of these genes are regulated by the ArcR, ArgR and AhrC regulators.

Amongst the strongly down-regulated genes, the *bfb* operon (*bfbABCDFGR* and SGO_1575) in particular appears to be less strongly down-regulated within the *S. gordonii* Δ*arcR* strain, than in any of the other three strains. Further comparison of the fold-change values for the *bfb* genes in the wild-type *S. gordonii* DL1 vs Δ*arcR* strains, under high and then low arginine conditions, showed that the *bfb* genes were more strongly down-regulated in the Δ*arcR* strain in high arginine conditions than low. This may be indicative of direct regulation and activation of these genes by the ArcR protein, in response to high exogenous arginine, however none of these genes were predicted to have a TFBS (transcription factor binding site) in their intergenic regions. Still, as mentioned previously, the ArcR binding footprint has not been clearly defined, and so it may be binding to a different target sequence to the ArgR/AhrC regulators. Other genes that appear to be directly affected in their expression within the Δ*arcR* strain are the *wefE* and *glf* genes of the RPS (receptor polysaccharide) biosynthesis operon, which were more strongly down-regulated within the *S. gordonii* Δ*arcR* strain, when compared to the wild-type.

Analysis of the genes found in Table 4.7 and Table 4.8 indicated that a number of genes appear to be differentially expressed within the *S. gordonii* Δ*argR* and Δ*ahrC* strains only, which may suggest direct regulation of the expression of those genes by the ArgR and AhrC regulatory proteins. These genes include $pyrA_a$ and $pyrA_b$ and *argBCDJ*, known to be regulated by these arginine-dependent regulators, and additionally SGO_1102 and 1105 (both within the

pyrA$_a$/*pyrA*$_b$ operon); SGO_1406 to 1411, all genes within the histidine biosynthesis pathway; SGO_1656 and SGO_0021. All of these genes are up-regulated in response to arginine depletion within the wild-type strain, but less so within the *S. gordonii* Δ*argR* and Δ*ahrC* strains. Additionally, genes that appear to be specifically down-regulated by ArgR and AhrC include the fatty acid biosynthesis genes SGO_1686, SGO_1699 and *accACD*; the ribosomal subunit protein *rpsD*; and the enoyl-CoA catabolism gene SGO_1700.

4.3.6. Prediction of arginine regulator-specific transcription factor binding sites

In order to predict whether any of the genes apparently regulated in response to arginine depletion by the *S. gordonii* Δ*arcR*, Δ*argR* and Δ*ahrC* strains, may be directly regulated in their expression by these arginine-dependent regulators, the *S. gordonii* genome was analysed for putative TFBSs for ArgR. ArgR is the name given to the regulator family of *S. gordonii* by the RegPrecise database (RegPrecise, 2009-2015), and consists of the three regulatory proteins ArgR, ArcR and AhrC. The transcription factor binding site of the ArcR regulator at the *arcA* promoter has been assessed by DNase I footprinting (Zeng *et al.*, 2006). However, the binding site is longer than usual, at 27 bp, and has not been widely analysed nor clearly defined. A BLASTN search of the *S. gordonii* genome for similar sequences or binding sites only shows the sequence upstream of *arcA*, which is referenced in the aforementioned paper, and no others within the *S. gordonii* chromosome. Therefore, for further *in silico* analysis of putative arginine-dependent regulator binding sites in *S. gordonii*, the ArgR regulon sites listed on the RegPrecise database were used.

The TFBS sequences for the ArgR regulator in all Streptococcaceae species (125 sequences in total) were downloaded from the RegPrecise database (RegPrecise, 2009-2015), and compiled into a position frequency matrix (PFM) for the ArgR regulator. This listed the number of times (out of 125 sequences) that a particular nucleotide (e.g. A, G, T and C) was found at each position along an 18 bp sequence. This PFM was then used to analyse the *S. gordonii* DL1 genome, throughout both the whole chromosome and the intergenic regions, using the PePPER server (de Jong *et al.*, 2012). Unfortunately, the output for the website does not give the sequence of the predicted TFBS, however, the PFM used to search for these sites is shown in Figure 4.5.

(A)

A	[91	53	28	12	50	119	3	118	116	71	38	118	2	79	29	89	68	56]
C	[0	3	2	4	6	1	0	1	2	4	6	0	4	7	78	5	13	2]
G	[2	16	16	79	5	1	1	0	1	2	2	0	0	22	8	3	8	4]
T	[32	53	79	30	64	4	121	6	6	48	79	7	119	17	10	28	36	63]

(B)

Figure 4.5. Position frequency matrix and sequence logo of the ARG box transcription factor binding sequence. (A) shows the position frequency matrix used to search for TFBSs – the sequences are 18 bp long, and each position indicates how often each nucleotide appears at that position within the sequence, out of a total of 125 analysed sequences. (B) shows a sequence logo for the same PFM, where the larger the letter, the more likely it is to be found at that position. Sequence logo created using the WebLogo website (Schneider and Stephens, 1990; Crooks *et al.*, 2004).

In total, 54 genes were predicted to have ArgR TFBSs in their intergenic regions. Of these putative intergenic sites, 26 were found within genes regulated (to any extent, not just a significant extent) by one or more of the arginine-dependent regulators. The most likely TFBSs are shown in Table 4.9, and were decided to be likely sites as these genes showed clear evidence of regulation by the arginine-dependent regulators. Within this list appeared the promoter regions upstream of *argG* and *argC* (regulated to some extent by all three arginine-dependent regulators), and SGO_1102, a gene within the same operon as the pyrimidine/arginine biosynthesis genes $pyrA_a$ and $pyrA_b$, encoding the PyrR bifunctional gene, and shown in Table 4.7 (section 4.3.5) to be regulated by the *S. gordonii* ArgR and AhrC regulators. Other genes that are apparently regulated by ArgR/AhrC (according to the values in Table 4.7), that have predicted TFBSs within the intergenic region before the gene are the histidine biosynthesis gene *hisC*, and SGO_0021, a hypothetical protein-encoding gene (Table 4.9).

However, more work would be needed on these putative binding sites, such as DNA electrophoretic mobility shift assays to analyse regulator binding to promoter regions, and therefore determine whether any of these genes are in fact directly regulated by any of the three arginine-dependent regulatory proteins, ArcR, ArgR or AhrC.

Gene name	Position of TFBS (base pairs upstream)
argC	-87 bp
argG	-43 bp
SGO_1102	-222 bp
hisC	-471 bp
SGO_0021	-260 bp

Table 4.9. Putative transcription factor binding sites found in intergenic regions before genes, and their binding position. Sites were searched for and predicted using a position frequency matrix. Number dictates how far upstream of the start of the gene the TFBS begins.

4.4. Discussion

Nucleic acid insertions have long been known to have the ability to affect expression of downstream genes, causing what is known as a polar effect (Franklin and Luria, 1961; Ciampi and Roth, 1988). Altering genetic information, by way of either insertion or deletion of DNA within the bacterial chromosome, has the ability to shift the translational frame of downstream genes, and cause frameshift mutations including both nonsense and missense mutations (Starlinger and Saedler, 1972). Analysis of three *S. gordonii* strains carrying replacement knockouts of the *arcR*, *argR* and *ahrC* genes in this chapter required assurance that any phenotypes or differential gene expression observed in these strains was not due to polar effects on downstream genes, but due to knockout of those specific genes and loss of their activity.

The qRT-PCR analysis performed on expression of genes downstream of these three regulator knockouts showed that, with the exception of the *arcT* gene in the Δ*arcR* strain, neither *mutS* nor *recN* were significantly altered in their expression following deletion of the preceding gene. In the case of *arcT*, however, there was a small difference observed. ArcR has never been shown to have a regulatory effect on the expression of the *arcT* gene, however, *arcT* is predicted to encode a transpeptidase enzyme, in order to cleave arginine residues from internalised peptides (Zúñiga *et al.*, 1998; Liu *et al.*, 2008). It is likely that this gene is linked to the *arc* operon, by providing an additional source of L-arginine for catabolism by the ADS system, and so this would suggest that it may also be regulated by ArcR. ArcT has been previously shown to be linked to the *arc* operon in other species, such as *Lactobacillus sakei*, where it appears to be co-transcribed with the *arcD* gene, and regulated by ArcR (Zúñiga *et al.*, 1998; Zúñiga *et al.*, 2002).

In the gene expression microarray experiments, comparison of the *S. gordonii* Δ*arcR* strain with the wild-type, under high arginine conditions, showed that *queA*, *argGH*, SGO_0177 and SGO_0846 were all up-regulated in the Δ*arcR* strain. In addition, ArcR also appears to be involved in the regulation of the *bfb* locus. The *bfb* operon has not been linked to regulation by arginine, or an arginine-dependent regulatory protein, previously. However, the expression of this operon was between 15 to 50-fold less strongly down-regulated within the *S. gordonii* Δ*arcR* strain following arginine depletion, in comparison to the wild-type strain. The SGO_1575

gene, which encodes another subunit of the cellobiose PTS, was also 13-fold less strongly down-regulated in the Δ*arcR* strain, indicating it too is regulated by ArcR. These genes were also found to be more differentially regulated between the Δ*arcR* and wild-type DL1 strains under high exogenous arginine conditions, as opposed to low, suggesting that ArcR may play a role in activating the *bfb* locus under high arginine conditions. As the *bfb* locus is also known to be associated with biofilm formation in *S. gordonii*, it is possible that ArcR may play a role in biofilm formation in this species (Kiliç *et al.*, 2004; Jakubovics *et al.*, 2008).

The *wefE* and *glf* genes of the RPS biosynthesis operon were found to actually be more strongly down-regulated following deletion of *arcR* (12-fold and 5-fold respectively), when compared to the levels of expression within the *S. gordonii* wild-type. This suggests that ArcR may be playing a role in alleviating the down-regulation of these genes following arginine depletion. Neither these genes, nor the *bfb* locus genes were predicted to contain an ArgR TFBS within their intergenic regions. However, as mentioned before, the ArcR protein appears to bind to a different sequence to ArgR and AhrC (Zeng *et al.*, 2006). Given the extent of differential regulation of the *bfb* locus, SGO_1575 and *wefE* within the *S. gordonii* Δ*arcR* strain, it seems likely that the ArcR protein is involved in their regulation to some extent. Additionally, the discovery of the strong up-regulation of the SGO_0846 gene within *S. gordonii* Δ*arcR* suggests that the ArcR protein is responsible for regulating the expression of this gene also, and this forms the basis for the work in Chapter 5.

In the microarray analyses of the *S. gordonii* Δ*argR* and Δ*ahrC* strains, many of the genes that were differentially regulated in response to arginine depletion appear to have functions in metabolism, transport and transcription/translation. This is suggestive of a more general stress response in reaction to arginine depletion within the growth medium. In particular, up-regulation of transport protein-encoding genes, and down-regulation of genes encoding ribosomal subunits is indicative of a stress response, triggered by amino acid starvation (Farr and Kogoma, 1991). This would also appear similar to a nutrient stress response – a paper by Betts *et al.* (2002), investigating genetic responses to nutrient stress in *Mycobacterium tuberculosis*, showed a similar regulation in different functional classes of genes to those seen in response to arginine depletion in these microarrays.

Initial comparisons of the data from the microarrays carried out in this chapter has already shown some genes that appeared to be differentially-expressed in one of the arginine-dependent regulator mutant strains, when compared to *S. gordonii* DL1 in a shift from high to no arginine conditions. Due to the batch effect shown in section 4.3.2, no direct statistical comparisons between the microarray data sets could be undertaken at this time. However, in collaboration with the Bioinformatics Support Unit at Newcastle University, the possibility of comparing between arrays using rank product analysis is being explored. This is a method of analysing microarray data, by ranking genes in order of how consistently they are highly expressed within a particular strain or under certain conditions (Breitling *et al.*, 2004). Preliminary data analysis using rank product analysis to compare across the microarray datasets, however, appeared to be missing key genes (such as *argGH*), that were shown in section 4.3.5 to be differentially-regulated between the *S. gordonii* DL1 and Δ*argR*/Δ*ahrC* strains following the loss of regulation by the ArgR and AhrC proteins. However this may be due to the system used to filter out genes that were not significantly expressed in the microarray prior to the rank product analysis, and so analysis is ongoing to determine whether rank product analysis is the best way to directly compare data between the microarrays.

Preliminary analysis of the levels of expression of different genes between all four *S. gordonii* strains indicates that the *argBCDJ* arginine biosynthesis genes appear to be specifically regulated by the ArgR/AhrC transcriptional regulators, in addition to genes from the pyrimidine/arginine biosynthesis operon (SGO_1102, $pyrA_a$, $pyrA_b$) and SGO_1656. The qRT-PCR analysis undertaken in section 4.3.3 corroborated these findings. Furthermore, the biofilm-associated *bfb* operon appears to be specifically regulated by the ArcR protein, as these genes were less strongly down-regulated in the absence of arginine following deletion of the *arcR* gene than they were in the *S. gordonii* wild-type, Δ*argR* or Δ*ahrC* strains.

The arginine and pyrimidine biosynthesis pathways are known to be linked by the carbamoylphosphate synthetase enzyme, encoded by *arcB* in *S. gordonii* (Larsen *et al.*, 2008; Jakubovics *et al.*, 2015), which biosynthesises the molecule carbamoylphosphate, a precursor to both the arginine and pyrimidine biosynthetic pathways (Raushel *et al.*, 1999). Additionally, the *carAB* genes of *E. coli* (homologs of the $pyrA_a$ and $pyrA_b$ genes in *S. gordonii*) have been previously shown to be regulated by the ArgR transcriptional regulator (Devroede *et al.*, 2004).

Therefore it is not surprising that the arginine and pyrimidine biosynthesis genes of *S. gordonii* appear to be under the control of the same regulator proteins, in this case, ArgR and AhrC. Within the *S. gordonii* Δ*argR* and Δ*ahrC* strains, when compared to the levels of expression for these genes within the wild-type strain, the *argC* gene was approximately 500-fold less up-regulated; the *argGH* genes were around 200-fold less up-regulated; *argBDJ* were 250 to 300-fold less up-regulated; and the $pyrA_a$ and $pyrA_b$ genes were 20-fold less up-regulated under no arginine conditions. Additionally, the SGO_1102 gene (within the pyrimidine/arginine biosynthesis operon) was 16-fold more up-regulated within the wild-type; SGO_1656 was 40-fold more highly regulated; and SGO_0021 was 20-fold more highly regulated in the wild-type strain than the *argR* and *ahrC* mutants.

The differences in expression of these genes strongly suggest that they are up-regulated in their expression in low arginine conditions by the ArgR and AhrC regulators *in vivo*. Furthermore, some of these genes, namely *argC*, *argG*, SGO_1102 and SGO_0021, have putative TFBSs for ArgR in the intergenic regions before these genes. These genes, in particular SGO_1656, 1102 and 0021, would therefore be good targets for following up with qRT-PCR analysis to confirm regulation by the ArgR and AhrC regulatory proteins.

Therefore, the work in this chapter has resulted in a preliminary designation of the regulons of the ArcR, ArgR and AhrC regulators, which can be further investigated if a statistical method for directly comparing the data between microarrays is found. In any case, a number of genes have been identified that are putatively regulated by ArgR and AhrC, and these can be further investigated by qRT-PCR or other transcriptional analysis methods. These initial regulons suggest that ArgR and AhrC overlap precisely in their regulation, which has not been shown in any other species before. In the paper by Jakubovics *et al.* (2015), models for regulation of arginine catabolism and biosynthesis genes by ArcR, ArgR and AhrC were proposed. ArgR and AhrC were proposed to form a heteromeric complex in order to regulate gene function, as with the ArgR and AhrC homologs of *Lactococcus lactis* (Larsen *et al.*, 2005). In this species, both of these regulators interact with each other to form a heteromeric complex, necessary to regulate expression of arginine biosynthesis genes such as *argC*. It is possible that, as the regulons appear to overlap precisely between these proteins, this is the same method that *S. gordonii* ArgR and AhrC use to regulate gene expression.

The models of arginine-dependent regulation proposed in the paper by Jakubovics *et al.* (2015) suggested that in high exogenous arginine conditions, ArcR up-regulates expression of arginine catabolism genes such as *arcABC*, and down-regulates *argGH* biosynthesis gene expression, whilst the ArgR/AhrC complex represses transcription of the biosynthesis genes and accessory arginine-related genes, such as the *arcD* arginine transporter. In low or no arginine conditions, ArcR is responsible for slight up-regulation of the *argC* and *pyrR* genes (involved in arginine and pyrimidine biosynthesis), at the same time as the ArgR/AhrC complex is up-regulating arginine biosynthesis genes. The microarray analyses performed within this chapter corroborates these hypotheses, and in addition, has identified a number of other genes that appear to be regulated by ArcR, ArgR and/or AhrC.

Therefore, future work is needed to follow up the preliminary findings of the microarray, and should focus on direct comparisons between the data from all four *S. gordonii* strains in response to arginine depletion, qRT-PCR analysis on the genes which comprise the regulons of ArcR, ArgR and AhrC, and elucidation of the mechanisms by which these proteins regulate gene expression in response to arginine depletion.

The work performed in Chapter 5 aims to further investigate the *S. gordonii* Δ*arcR* strain, and the SGO_0846 gene that was up-regulated within this strain. It also aims to characterise the SGO_0846 protein, and determine the role it is playing within the *S. gordonii* cells.

5. Further analysis of the *S. gordonii* Δ*arcR* mutant

5.1. Introduction

In the previous chapter, it was found that many genes were regulated in response to arginine, and that some were also regulated in response to the deletion of the arginine regulatory genes. One such gene that was specifically regulated in response to deletion of the *arcR* regulator was an uncharacterised gene, SGO_0846. This gene is annotated as encoding a possible cell wall protein by NCBI Gene (NCBI, 2015), although the analysis undertaken in this chapter suggests that it is more likely a released protein. SGO_0846 was found to be up-regulated 148-fold in *S. gordonii* Δ*arcR* under high arginine conditions compared with the isogenic wild-type. It is hypothesised that, as a putative released protein, it may be involved either in cell signalling and recruitment of cells to the biofilm, or in structuring of the biofilm and dissemination of cells to other sites *in vivo*, roles that both contribute to biofilm formation in different ways. Additionally, SGO_0846 was not up-regulated in the other two regulator mutant strains, *S. gordonii* Δ*argR* or Δ*ahrC*.

Previous work undertaken in our lab (Jakubovics, unpublished work), investigated the ability of the three arginine regulator mutant strains to form biofilms under aerobic conditions, and showed that whilst the *S. gordonii* Δ*argR* and Δ*ahrC* strains (PK3346, PK3349, PK3350) formed biofilms that were not impaired in growth, the Δ*arcR* strains (PK3354, PK3355, PK3356) appeared to be defective in biofilm formation (Figure 5.1). These defective biofilms were found to be statistically significantly different in their levels of formation from wild-type *S. gordonii* DL1 (DL1 vs Δ*arcR* strain and DL1 vs Δ*arcR*/Δ*argR* strain $p<0.01$, DL1 vs Δ*arcR*/Δ*ahrC* strain $p<0.05$).

The ArcR protein of *S. gordonii* is known to regulate a number of different genes – namely, the *arcABC* and *arcDT* genes, some apparent regulation on its upstream gene *queA* (as observed in the previous microarray experiments), and minor regulation on the arginine biosynthesis genes (*argC, argJ, argGH*). Work from our group (Jakubovics *et al.*, 2015) showed that strains of *S. gordonii* carrying respective deletions of the genes *arcA*, *arcB* and *argH* showed no defective biofilm phenotypes. The *queA, arcD* and *argG* genes have not been tested, however, none of these genes have previously been shown to impact on biofilm formation in *S. gordonii*. The

queA gene, which is involved in queosine modification of tRNA molecules, has not been linked to biofilm formation, however, QueA has been shown to be linked to virulence and stationary phase survival in other bacterial species (Liu *et al.*, 2008).

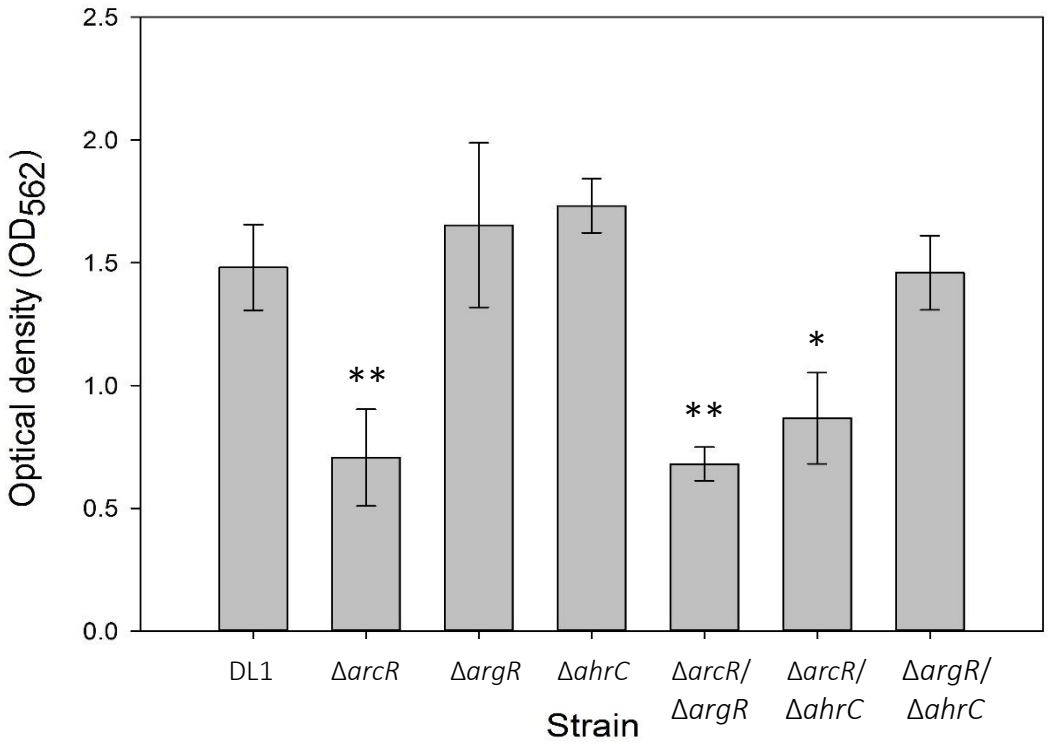


Figure 5.1. Biofilm formation by arginine regulator mutant strains, drawn based on unpublished data from the Jakubovics lab. Biofilms were grown in TYEG medium, for 18 h under aerobic conditions, and quantified by measuring absorbance of crystal violet-stained cells at 562 nm. The levels of biofilm formation by Δ*arcR* single or double mutant strains were statistically significantly lower than those of the other mutant strains (as measured by two-tail unpaired T-test, assuming unequal variance). A triple mutant strain (PK3357) containing deletions of all three regulators also showed a similar impairment in biofilm formation (data not shown). Bars represent arithmetic means of three independent biological replicates, and error bars show standard deviation. * p = <0.05, ** p = <0.01. The Δ*arcR* mutant strains show a significant biofilm defective phenotype, not displayed by the other regulator mutant

Due to this, there was no clear explanation for the observed biofilm attenuation in the Δ*arcR* gene. Therefore, other genes needed to be investigated to find the basis of this defect, and when the dramatic up-regulation of the SGO_0846 gene was found, it was hypothesised that the SGO_0846 protein may be involved in the attenuated biofilm phenotype of *S. gordonii* Δ*arcR*. This chapter aims to test this hypothesis, and to further characterise the SGO_0846 gene.

In addition, in order to confirm that the biofilm phenotype observed in the Δ*arcR* strain was related to disruption of *arcR* rather than the insertion of an antibiotic resistance cassette or a second site mutation, in this chapter complementation of the Δ*arcR* strain was achieved by placing a copy of the native *arcR* gene back into the bacterial cells.

Overall, this chapter of the thesis aimed to oversee complementation of the Δ*arcR* strain, and deletion of the SGO_0846 gene in both the wild-type and Δ*arcR* backgrounds. Furthermore, in depth analysis of SGO_0846 itself was performed, and the biofilm-forming abilities of both the *arcR*-complemented and SGO_0846 knockout strains were investigated.

5.2. Confirmation of *S. gordonii* Δ*arcR* biofilm defect

5.2.1. Biofilm growth under anaerobic conditions

Previous unpublished work in our group demonstrated that Δ*arcR* single mutants, and strains carrying a deletion of *arcR* in addition to one of the other arginine-dependent regulators (Δ*arcR* double mutants) displayed a clear defect in biofilm formation under aerobic conditions. In order to investigate whether the Δ*arcR* biofilm defective phenotype also occurs under anaerobic conditions, the *S. gordonii* DL1, PK3351 (Δ*arcR*::*ermAM*) and PK3354 (Δ*arcR*::*aad9*) strains were cultured anaerobically in 50% BHY medium for 18 h overnight to form biofilms, stained with crystal violet dye and quantified (Figure 5.2).

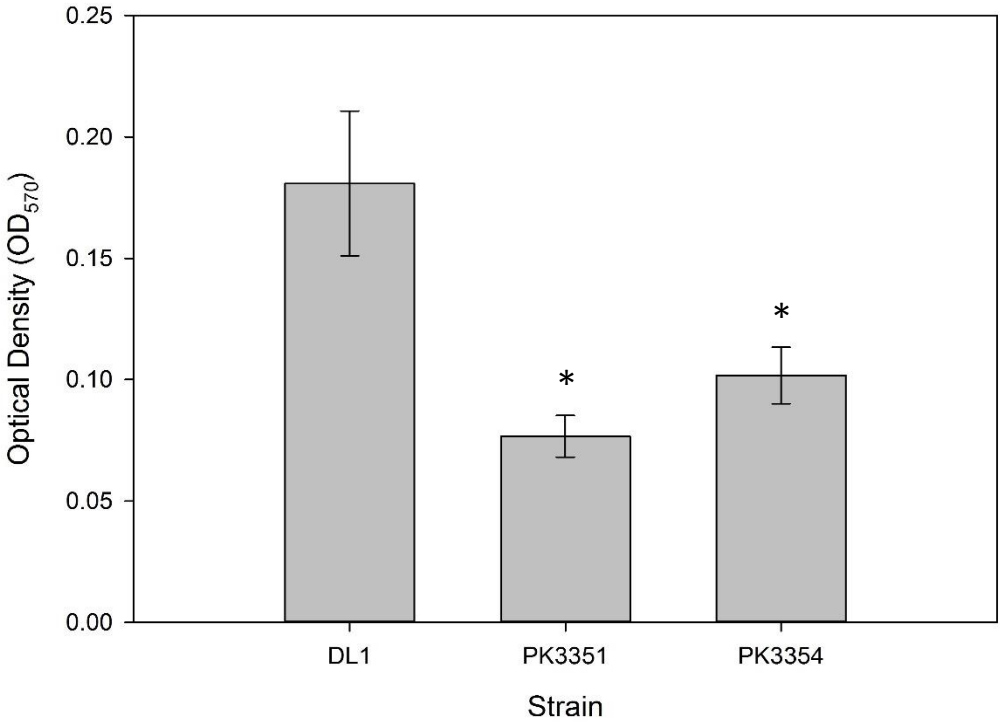


Figure 5.2. Role of ArcR in anaerobic biofilm formation. To assess whether ArcR was important for biofilm formation under anaerobic conditions, biofilms were grown from wild-type *S. gordonii* DL1, PK3351 (Δ*arcR*::*ermAM*) and PK3354 (Δ*arcR*::*aad9*) strains. The biofilms were stained with crystal violet, and quantified by measuring optical density (at 570 nm). Bars represent arithmetic means of six replicates, and error bars show standard error of mean. (One-way ANOVA, two sample T-test, *p <0.05). Both Δ*arcR* strains were significantly impaired in biofilm formation in comparison to the wild-type DL1 strain.

It was found that the Δ*arcR* mutant strains displayed a defective biofilm phenotype under anaerobic conditions, with both mutant strains showing statistically significantly reduced biofilm formation compared with the wild-type DL1 strain (DL1 vs PK3351 p=0.021, DL1 vs PK3354 p=0.048). The consistent observation of this defect under both aerobic and anaerobic conditions, and the knowledge that neither *arcR* nor its co-transcribed gene *queA* are subject to regulation by oxygen, suggests that this is not an oxygen-sensitive phenotype. Also, the lack of observed defect in the Δ*argR* and Δ*ahrC* strains suggests that this is a phenotype unique to ArcR.

It is worth mentioning that, despite the obvious biofilm attenuation shown by the Δ*arcR* strains, the growth yields of the planktonic cultures (assessed by measurement of the OD_{600}) for all strains were comparable across all strains and replicates (data not shown). When examined qualitatively during staining, biofilms in the Δ*arcR* strains appeared to grow normally and to the same levels as the DL1 strain, and it was only upon washing to remove the loosely-bound dye and cells during the protocol that the biofilm began to disintegrate. Therefore, it appears that the biofilm-defective phenotype in *S. gordonii* Δ*arcR* is not linked to poor biofilm growth, but rather is due to poor binding between the cells, causing them to form weak biofilms that are easily removed. Whether this phenotype is a result of the apparent over-expression of the SGO_0846 gene remains to be seen, and was investigated with *in silico* analysis of the gene and protein, and knock-out of the gene.

5.3. *In silico* analysis of the SGO_0846 gene

In order to find out more information about the SGO_0846 gene, and the possible function of the encoded protein product, *in silico* analysis was performed on both the nucleotide and amino acid sequences. Firstly, the 2004 bp nucleotide sequence was used to create a map of the gene and operon structure, and to investigate putative regulatory sequences (Figure 5.3). The DOOR^2 database (Mao *et al.*, 2009) was first used to help predict whether the SGO_0846 gene was part of an operon with the surrounding genes. The SGO_0845 gene lies around 800 bp upstream of SGO_0846, and the SGO_0847 gene is 223 bp downstream. Each of these genes, SGO_0845, 0846 and 0847, are predicted to be transcribed separately since each gene has its own predicted terminator (predicted by the DOOR^2 database), and each has its own

promoter, predicted by the PromBase database (Rangannan and Bansal, 2011). These putative terminators and promoters are shown in Figure 5.3.

Analysis of the nucleotide sequences around SGO_0846 with the PePPER transcription factor binding site (TFBS) prediction tool (de Jong *et al.*, 2012), identified one putative TFBS around 614 bp upstream of SGO_0846. PePPER uses known bacterial regulons and regulator binding sites to predict TFBSs in a target bacterial genome. In this case, analysis of the nucleotide sequences from the start of SGO_0845 to the end of SGO_0847 for the presence of any AhrC or ArgR homolog binding sites in the species *Bacillus subtilis, Lactococcus lactis* and *Escherichia coli* showed the presence of a predicted AhrC binding site, homologous to *B. subtilis* AhrC binding sites, upstream of the SGO_0846 gene and running in the same direction. It is possible that this could be the site used by an arginine regulator, such as ArcR, to control expression of the SGO_0846 gene.

Further analysis was then performed upon the amino acid sequence of the SGO_0846 protein. The sequence is 667 residues in length, and 76.1 kDa in mass (UniProtKB, 2007). The LocateP database (Zhou *et al.*, 2008) predicted SGO_0846 to be a secreted molecule, recognised by a type I signal peptidase enzyme (SPI). The protein itself is predicted to have a signal peptide cleavage site between residues 29 and 30, with the sequence NTVQAADY. Cleavage was predicted to occur between the double alanine residues, by the SignalP database (Petersen *et al.*, 2011).

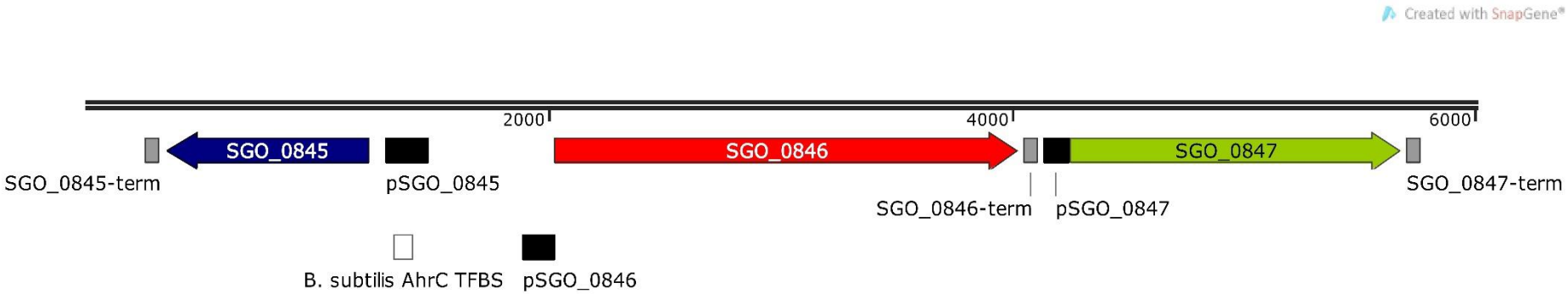


Figure 5.3. Map of the SGO_0846 gene and neighbouring genes. The SGO_0846 gene is marked in red, with the SGO_0845 gene 800 bp upstream (blue) and the SGO_0847 gene 223 bp downstream (green). The predicted promoters (black) for each gene are shown, as predicted by PromBase, as are the predicted terminators (grey), predicted by DOOR[2]. A predicted transcription factor binding site is also shown 614 bp upstream of the start of SGO_0846 (white), for a 77 bp sequence running 5'-3' that shows homology to the binding sites of the *Bacillus subtilis* 168 AhrC transcriptional arginine metabolism regulator (predicted by the PePPER TFBS tool). (Map created using the SnapGene software, and drawn to scale. Numbers on the DNA strand represent the scale in base pairs).

According to the annotation in the NCBI Gene database (NCBI, 2015), SGO_0846 is predicted to be an uncharacterised 'cell wall protein'. However, upon examination of the polypeptide sequence for wall anchor motifs using the CW-PRED server (Litou *et al.*, 2008), no LPxTG or non-LPxTG (e.g. NPxTG, LPxTA, LAxTG) anchoring motifs were found. The peptide was also found not to be a lipoprotein, or lipid-anchored.

Secondary structure prediction using the PSIPRED server (Buchan *et al.*, 2013) revealed a number of putative helices and sheets, along with a large amount of disorder present in the protein structure, predicted by the DISOPRED3 tool. This tool uses x-ray crystal structures of other disordered proteins as a basis for comparison with the user's target amino acid sequence, in order to predict regions of disorder (Ward *et al.*, 2004a). Some disordered regions were predicted to be protein binding. Figure 5.4 shows a map of the amino acid sequence, with predicted helices and sheets shown, as well as the region of disorder. The Phyre^2 analysis software also confirmed this finding, predicting that around 53% of the amino acid sequence was disordered (Kelley *et al.*, 2015). The PSIPRED Bioserf 3D modeller was used to create a predicted tertiary structure model for SGO_0846 (Buchan *et al.*, 2013) (Figure 5.5).

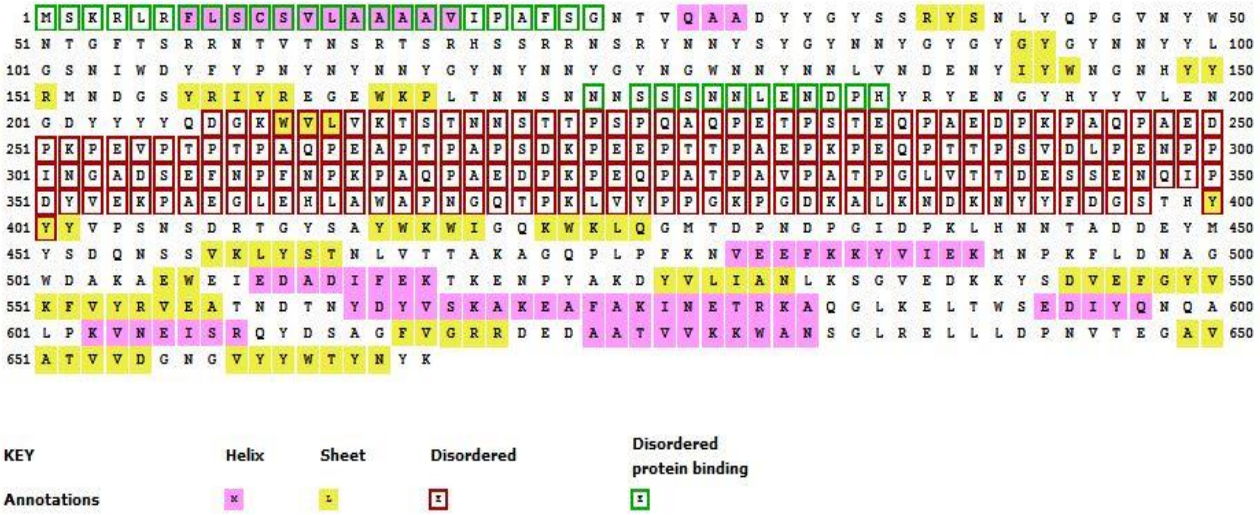


Figure 5.4. Map of the amino acid sequence of SGO_0846, as predicted by the PSIPRED software. Residues highlighted in pink form part of a helix, those highlighted yellow form part of a sheet structure. Regions outlined in red form an intrinsic disordered region (IDR), and those highlighted in green are disordered but predicted to be protein-binding. Analysis of the secondary structure was performed by PSIPRED, and disorder prediction by DISOPRED3. The SGO_0846 protein is predicted to be around 50% disordered.

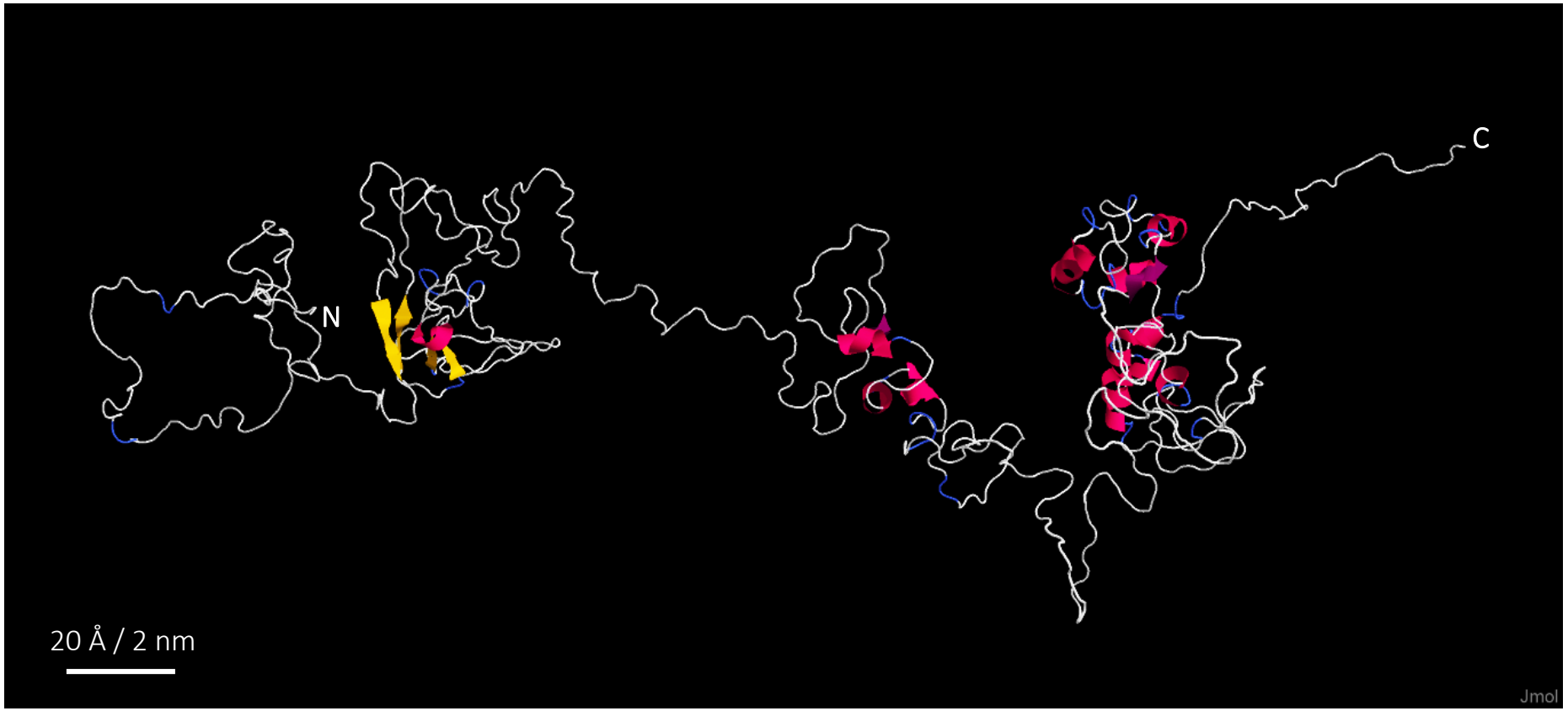


Figure 5.5. Predicted tertiary structure of the SGO_0846 protein, as predicted by the PSIPRED Bioserf server. Polypeptide structure is depicted from N-terminus to C-terminus, left to right (labelled as "N" and "C"). Yellow regions depict β-sheets, pink depicts α-helices, purple depicts 3_{10} helices, and blue regions of the chain show β-turns. Size bar shows a 20 angstrom (Å), or 2 nm distance. Tertiary structure was predicted by the Bioserf server, and modelled using JMol software.

PSIPRED MEMSAT-SVM analysis (Jones *et al.*, 1994; Jones, 2007; Nugent and Jones, 2009) was used to predict transmembrane protein topology, and predicted a pore-lining helix region between residues 648-663, at the C-terminal end of the polypeptide. Aside from this region, and the signal peptide region at the N-terminal end of the peptide, the remainder of the protein was predicted to be extracellular.

The primary amino acid sequence is asparagine-rich, with 71/667 (10.6%) of the sequence consisting of N residues. Tyrosine (Y) and proline (P) residues also appear frequently throughout the sequence, with the intrinsically disordered region (IDR) (residues 208-397) in particular appearing to be where the majority of the proline-residues were concentrated.

Searches for proteins with identity to the SGO_0846 protein using NCBI BLASTP (Altschul *et al.*, 1997; Altschul *et al.*, 2005) showed seven different *Streptococcus* species that contain genes encoding proteins with over 90% identity to the full sequence of SGO_0846. These results were all for annotated cell wall proteins, and included proteins from three species of *S. gordonii* (strains KCOM 1506, isolated from acute pulpitis infections, 95% identity; IE35, from infective endocarditis patients, 95% identity; and G9B, a dental plaque isolate, 98% identity), as well as *S. mitis* (99% identity), *S. oligofermentans* (98% identity), *S. sanguinis* (97% identity), and *Streptococcus* sp. 2_1_36FAA (isolated by the Human Microbiome Project, 97% identity). A number of other results were found for apparent cell wall protein in other *Streptococcus* species also, however these contained percentage identities lower than 90%, for less than 100% of the full SGO_0846 amino acid sequence, and so will not be discussed here.

In addition to searching for homologs using NCBI, an additional search was made using the StreptoBase database (Wenning Zheng, unpublished data). The Jakubovics group was granted pre-publication access to this database, which contains the genomes of many *Streptococcus* spp. bacteria, some of which have not previously had their genomes sequenced or published. Searching this database allowed us to find other proteins with high identity to SGO_0846, specifically within other *S. gordonii* strains. This search found that, out of 15 *S. gordonii* isolate genomes in the database, the SGO_0846 protein appears to be highly conserved between all of them, including *S. gordonii* Challis, FSS2, FSS3, M99, 40, PK488 and Blackburn. These results,

taken along with the NCBI BLASTP results, suggest that SGO_0846 may be a protein that is specific to streptococcal species, and oral streptococci in particular

5.4. Genetic complementation of *S. gordonii* Δ*arcR* mutant

As the Δ*arcR* strain has now been shown to form abnormal biofilms irrespective of growth conditions, it was important to complement the Δ*arcR* deletion strain with an intact copy of *arcR* and observe the effects upon growth.

5.4.1. Plasmid cloning strategy

A strategy was devised to complement the *arcR* gene using ligation-dependent cloning. The pCM18 plasmid (Hansen *et al.*, 2001), which contains a *gfpmut3** gene – specifically designed to allow for GFP production under anaerobic conditions in *S. gordonii* – and the strong constitutive lactococcal promoter CP25, was chosen as the vector for insertion of the *arcR* gene. The pCM18 plasmid contains a multiple cloning site (MCS), and many other restriction sites throughout the plasmid. Analysis of both the plasmid, and the *arcR* gene, found two restriction enzymes (*Sal*I and *Sph*I) that would cleave the *gfpmut3** gene from the plasmid without cleaving anywhere else in the plasmid or *arcR* gene. Amplification of the *arcR* gene with primers that contained the restriction sites for these two enzymes (see Materials and Methods section for primer sequences) would allow subsequent digestion, and ligation of both the newly-linearised vector with the *arcR* gene, forming a recombinant plasmid which could then be transformed directly into *S. gordonii*. Unfortunately, multiple attempts with this strategy yielded no transformants, and so a new strategy for *arcR* complementation using ligation-independent cloning was designed.

5.4.2. In-Fusion cloning strategy

Complementation of Δ*arcR* was successfully performed using a ligation-independent cloning (LIC) strategy – namely, the Clontech "In-Fusion" kit. This LIC method allows multiple fragments of DNA to be stitched together by use of primers that contain tails that are complementary to the adjoining section of DNA (Figure 5.6). The use of these primers to amplify the desired regions of DNA results in overlapping fragments of target DNA sequences that can be joined using the "In-Fusion" enzyme – a Vaccinia virus proofreading exonuclease enzyme that detects

complementary regions of double-stranded DNA and strips off the ends to make single-stranded overhangs (Irwin *et al.*, 2012). These DNA products can then line up with each other, anneal by hydrogen bonding via DNA base pairs, and once transformed into *E. coli*, become covalently-bonded recombinant molecules. In the case of ArcR, the *arcR* gene was attached to CP25, the constitutive promoter from the pCM18 plasmid (Hansen *et al.*, 2001); and the pPE1010 vector, which contains an erythromycin-resistance cassette, and a strong terminator (*rrnB* T1 terminator) (Egland *et al.*, 2004). This created the recombinant p*arcR*$_{comp}$ plasmid (Figure 5.7), which was transformed into *S. gordonii* Δ*arcR*::*aad9*. The plasmid could not be transformed into the Δ*arcR*::*ermAM* strain, as both the plasmid and that strain carry erythromycin-resistance cassettes, which would make selection for the uptake of the plasmid impossible.

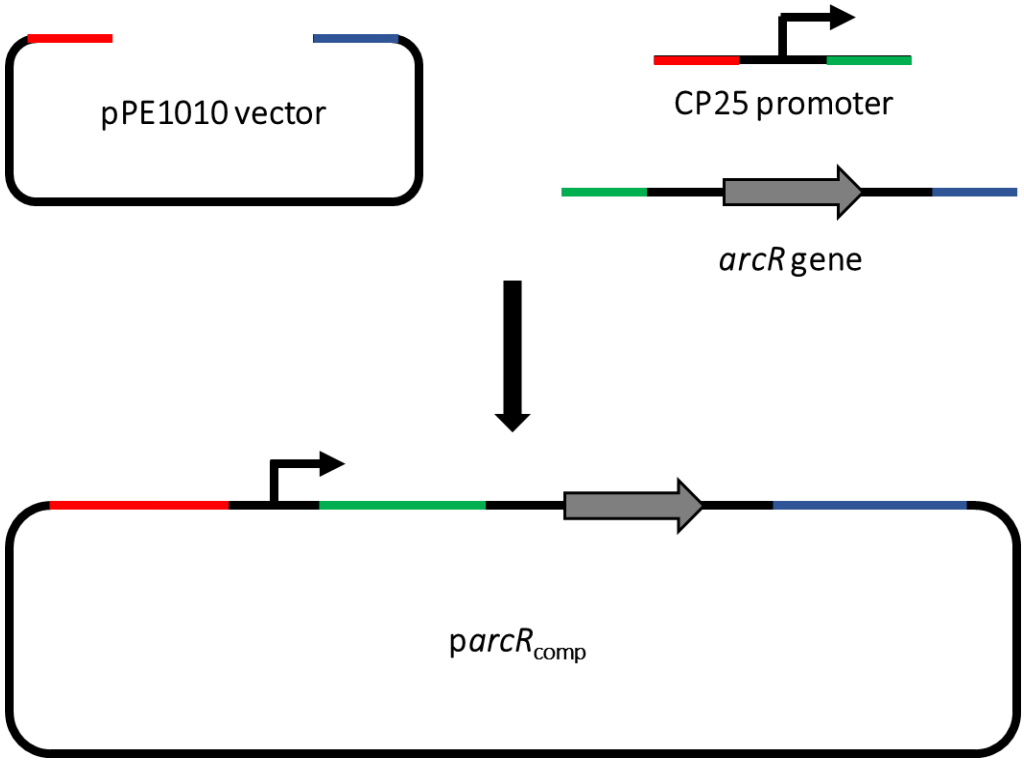


Figure 5.6. Visual representation of the overlapping DNA fragments required for ligation-independent cloning (LIC). The CP25 promoter and *arcR* gene were engineered to be complementary to both each other, and the pPE1010 vector. These products were then joined to form the recombinant p*arcR*$_{comp}$ plasmid. Diagram not drawn to scale.

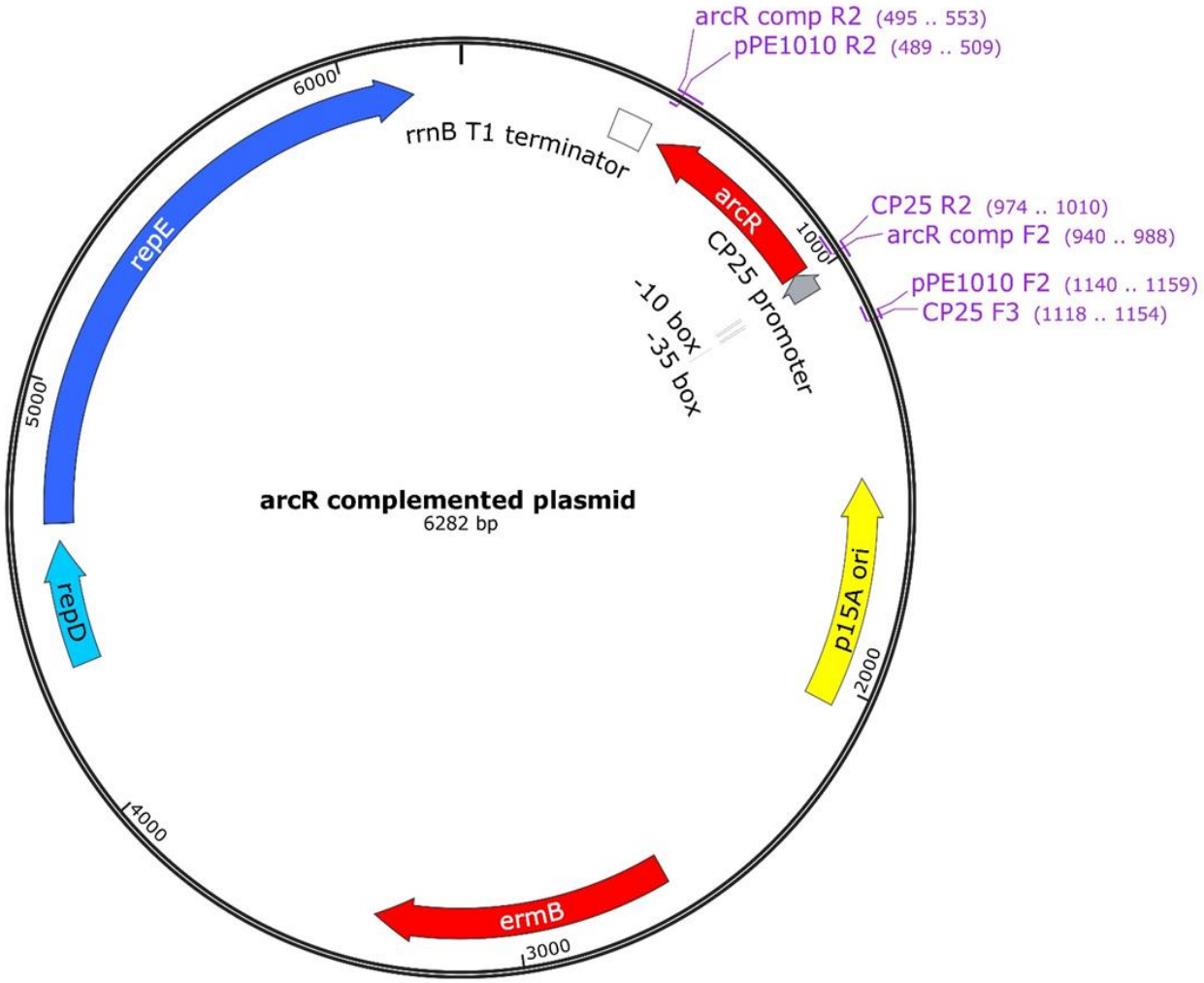


Figure 5.7. Final map of recombinant *arcR*$_{comp}$ plasmid. The plasmid contains the CP25 constitutive lactococcal promoter (from the pCM18 plasmid), the *arcR* gene, and the *rrnB* T1 terminator. An erythromycin resistance gene, *ermB*, is present to allow for antibiotic selection of transformants containing the plasmid. The binding locations of the primers used to copy the pPE1010 vector, CP25 promoter and *arcR* gene are also shown. Diagram drawn to scale, numbers around plasmid represent sequence size in base pairs. Full plasmid size was 6282 bp. (Map created using the SnapGene software).

The sequences of the primers that were used for amplifying DNA fragments with complementary ends can be found in the Methods chapter, along with details of their complementary regions. Once the plasmid had been transformed into *E. coli*, transformant colonies were picked and the plasmid removed from the cells. Eight of these were then sequenced to check that the expected sequences were present. Of the eight plasmids sequenced, two had no errors, and one of these was chosen for further work.

5.4.3. Confirmation of arcR *complementation*

Plasmids were extracted from eight transformant *E. coli* colonies, and sent to Eurofins MWG for dideoxy sequencing. Each sequence was aligned by Clustal W against the expected *arcR* and CP25 sequences using the MEGA6 software (Tamura *et al.*, 2013), in order to screen for any mutations (Figure 5.8). The transformant plasmid with the best alignment (in this case, 100% alignment) was then used for transformation into *S. gordonii* PK3354. The subsequent *S. gordonii arcR*$_{comp}$ strain was then used for further analysis into the role of the ArcR regulator in *S. gordonii* biofilm formation (see section 5.6.2).

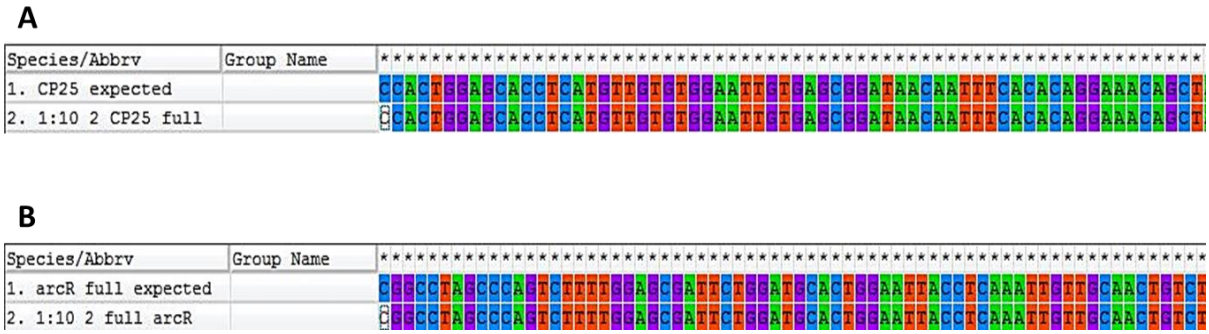


Figure 5.8. Sequence alignment between expected and received sequences for *arcR* complemented strain. The complemented *arcR* plasmid was sent to Eurofins MWG for dideoxy sequencing, using the CP25F/R and arcRF/R primers. A) shows alignment between the expected CP25 promoter sequence, and the plasmid sequence from one of the colonies picked from the *E. coli* transformation. B) shows the alignment between the expected *arcR* nucleotide sequence, and the plasmid sequence from the same transformant. This plasmid was then used for further transformation into *S. gordonii*. (Sequences aligned by ClustalW using the MEGA6 software).

5.5. Knock-out of *S. gordonii* SGO_0846 gene

Having engineered an *S. gordonii* complemented *arcR* strain, the next stage was to engineer a strain carrying a knock-out of the SGO_0846 strain, in the *S. gordonii* DL1 (wild-type), Δ*arcR* and $arcR_{comp}$ backgrounds. This would allow side-by-side comparison of the biofilm formation abilities of the *S. gordonii* Δ*arcR*, $arcR_{comp}$ and ΔSGO_0846 strains, as well as the DL1 wild-type strain.

5.5.1. Restriction digest strategy

A restriction digest strategy was devised to enable knock-out of the SGO_0846 gene. The SGO_0846 and pCR2.1 vector sequences were analysed for a restriction enzyme that cleaved once inside the SGO_0846 gene, but not within the pCR2.1 sequence. The enzyme *Bsu*361 was found to cut approximately half-way through the SGO_0846 gene. Primers were designed to amplify both erythromycin-resistance and kanamycin-resistance cassettes from other plasmids (pVA838 and pSF151 respectively, see Materials and Methods), and the restriction site for *Bsu*361 was placed within the 5' end of each primer's sequence.

The strategy was to insert the SGO_0846 gene into the linearised pCR2.1 vector, using a TA cloning kit (Life Technologies). The erythromycin and kanamycin-resistance cassettes could then be amplified using primers containing restriction sites for the *Bsu*361 enzyme. Finally, both the SGO_0846 plasmid and the antibiotic resistance cassettes could be cleaved with *Bsu*361, and ligated together to form recombinant SGO_0846::*ermAM* and SGO_0846::*aphA3* plasmids for insertion into the *S. gordonii* DL1 and Δ*arcR* strains. However, a flaw was later discovered in the primer design, and rather than continue with this strategy, another was developed allowing the transformation step through *E. coli* to be by-passed.

5.5.2. Overlap extension PCR strategy

Having had problems with a ligation-dependent strategy, it was decided upon to try a different approach, and an overlap extension PCR reaction was selected to knock-out the SGO_0846 gene. Primers were designed that were specific to flanking regions of the SGO_0846 gene, and to the same kanamycin and erythromycin-resistance cassettes used above. These primers also contained regions that were complementary between the antibiotic cassettes and the

SGO_0846 flanking regions, allowing overlap between the fragments, which could be stitched together in a long-range PCR reaction (primer sequences can be found in the Materials and Methods section). This would form a single PCR product, containing a copy of the SGO_0846 gene with an antibiotic resistance cassette inserted in the middle, which could be transformed into wild-type *S. gordonii* DL1 and Δ*arcR* strains (Figure 5.9). Theoretically, the SGO_0846::*abx*^{R} product (the SGO_0846 gene containing an inserted antibiotic-resistance cassette) would then homologously recombine onto the bacterial chromosome, replacing the intact copy of the SGO_0846 gene.

The overlap extension PCR products were transformed directly into *S. gordonii* DL1, Δ*arcR*::*ermAM* and Δ*arcR*::*aad9* strains. This resulted in strains that were both single (ΔSGO_0846::*ermAM*, ΔSGO_0846::*aphA3*) and double mutant strains (Δ*arcR*::*ermAM*/ΔSGO_0846::*aphA3*, Δ*arcR*::*aad9*/ΔSGO_0846::*ermAM*, Δ*arcR*::*aad9*/ΔSGO_0846::*aphA3*). These strains were confirmed by sequencing, as described in section 5.4.3 above, and found to contain the expected sequences. These strains were then used for further analysis into the role of SGO_0846 and ArcR in *S. gordonii* biofilm formation (see section 5.6).

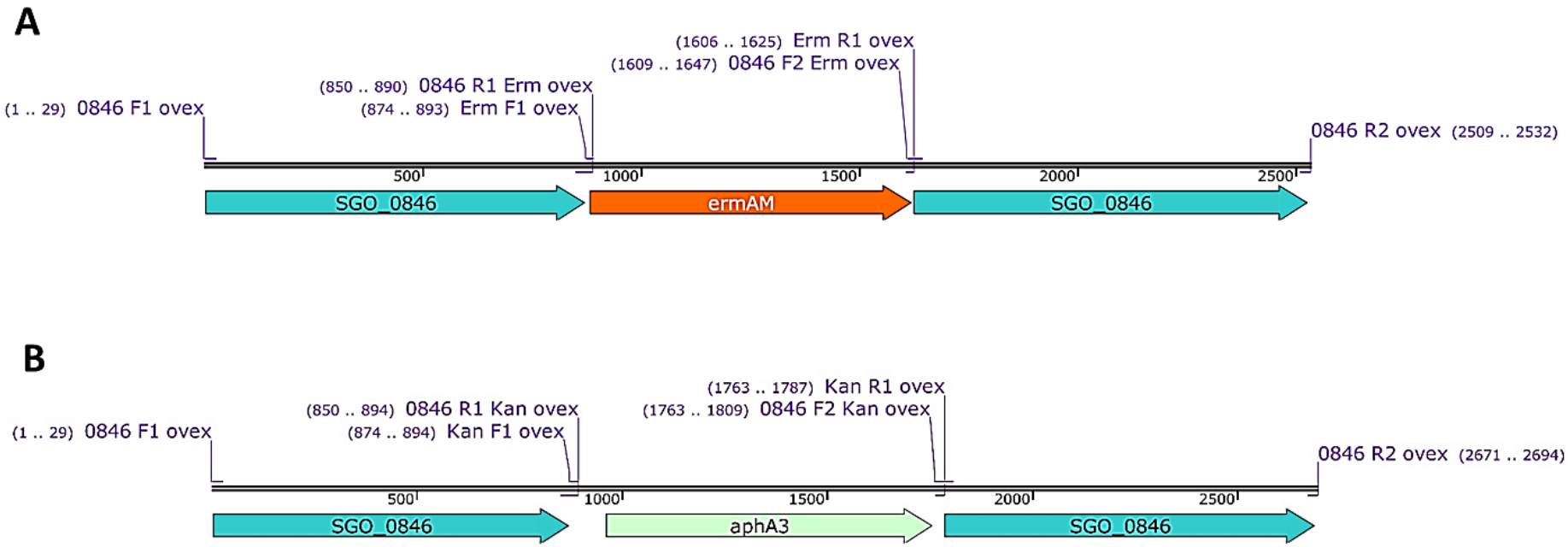


Figure 5.9. Overlap-extension (ovex) PCR products for *S. gordonii* SGO_0846 knock-out. Flanking regions of the SGO_0846 gene were amplified, using primers containing regions complementary to the ends of either the erythromycin-resistance *ermAM* cassette, or the kanamycin-resistance *aphA3* cassette. PCR products for the SGO_0846 flanking regions and the antibiotic resistance cassettes were then stitched together using a long-range PCR reaction. Maps were created to scale using the SnapGene software. Numbers on the DNA strand indicate the scale in base pairs.

5.6. Characterisation of *S. gordonii* SGO_0846 knock-out and *arcR* complemented strains

5.6.1. Growth of S. gordonii *ΔSGO_0846 knock-out strains*

To investigate the impact of SGO_0846 gene deletion on *S. gordonii* growth, wild-type *S. gordonii* DL1, Δ*arcR* (PK3351 and PK3354), ΔSGO_0846 (NU06) and Δ*arcR*/ΔSGO_0846 (NU09) cells were grown to stationary phase in replete FMC medium under anaerobic conditions. Complemented *arcR* strains did not form part of this experiment.

Cells were cultured overnight in nutrient-rich BHY medium, then used to inoculate pre-reduced replete FMC medium. Cells were grown strictly anaerobically, with timepoints taken every hour over an 11-hour time period, to follow the growth of all strains to stationary phase. Growth was followed by measuring OD_{600} of the different strains as described in Materials and Methods, with a final OD_{600} measurement also taken at the 24 h timepoint.

With the exception of the PK3351 (Δ*arcR*) strain, all strains had reached stationary phase, and a growth yield of OD_{600} >1, within six hours of inoculation (Figure 5.10). The PK3351 sample displayed a slightly extended lag phase, resulting in this strain taking an additional two hours to reach stationary phase following inoculation. This difference may have been due to a lower starting inoculum for the PK3351 strain, following poor growth in the overnight culture, however the mean specific growth rate for this strain was still comparable to that of the *S. gordonii* wild-type strain (PK3351 1.11 h^{-1} vs DL1 1.09 h^{-1}). The other Δ*arcR* strain, PK3354, exhibited normal growth with a mean specific growth rate of 1.15 h^{-1}.

However, the NU06 strain (ΔSGO_0846), despite appearing to grow normally and reaching the same growth yield (OD_{600} >1) within the same timeframe as the other strains, exhibited a lower mean specific growth rate (0.67 h^{-1}) than any other strain. It is possible that deletion of the SGO_0846 gene contributed to a lower growth rate for this strain, however, the NU09 strain carrying a double deletion of both the *arcR* and SGO_0846 genes did not display a similar phenotype, with a mean specific growth rate comparable to that of the wild-type strain (1.07 h^{-1}).

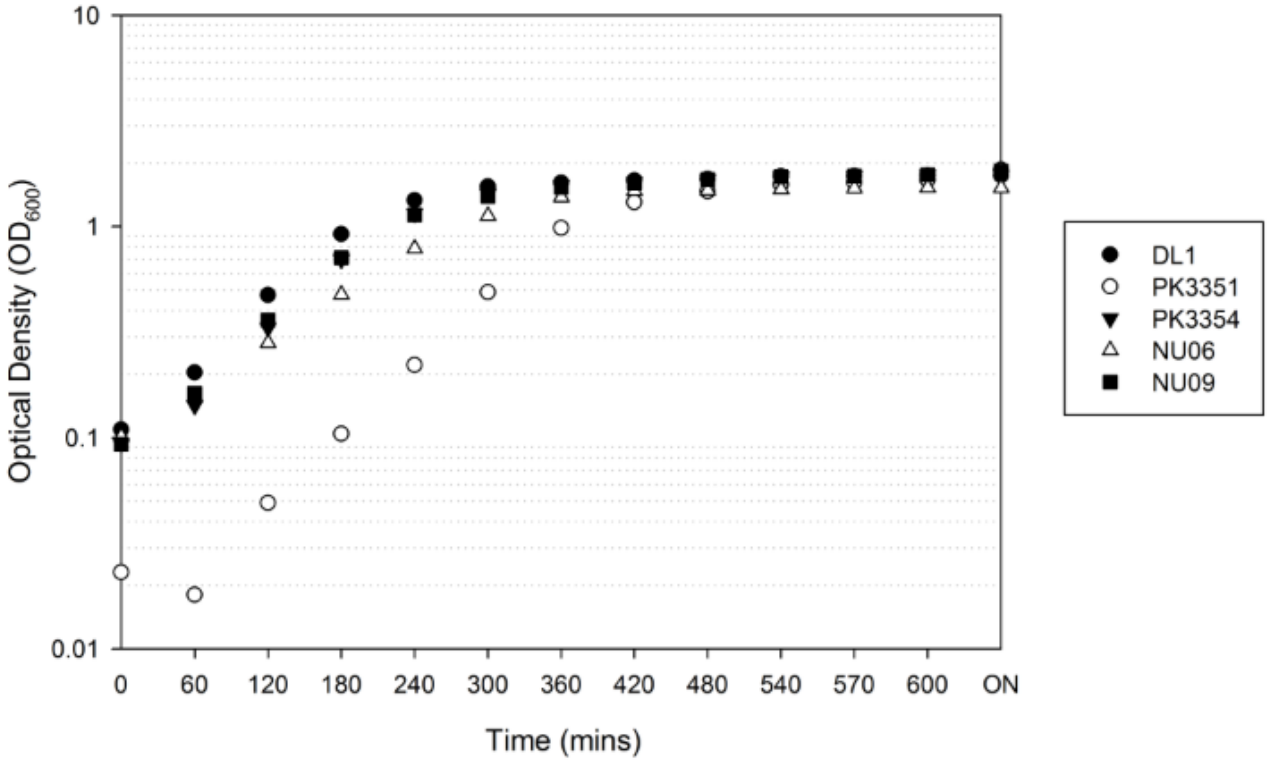


Figure 5.10. Representative growth curve of *S. gordonii* strains in chemically-defined medium. Wild-type *S. gordonii* DL1, PK3351, PK3354, NU06 and NU09 cells were grown anaerobically in full FMC medium, and spectrophotometric measurements were taken at various timepoints as a measure of cell growth. Graph shows one of two independent experiments, which both showed similar results. All cells grew to an OD_{600} >1 within 8 hours of inoculation. The NU06 strain displayed a lower mean specific growth rate than other strains.

Additionally biofilm experiments performed on the NU06 strain in section 5.6.2 showed no significant differences between the planktonic growth yields of either the NU06 or the DL1 strains. Therefore, the apparent slower growth rate of the NU06 strain was not subject to further investigation.

5.6.2. Biofilm formation in S. gordonii *complemented and knock-out strains*

Following construction of the *S. gordonii* ΔSGO_0846 (ΔSGO_0846::*aphA3*) and complemented *arcR* strains (Δ*arcR* p*arcR*$_{comp}$, and Δ*arcR*/ΔSGO_0846 p*arcR*$_{comp}$), the next step was to analyse their biofilm phenotypes in order to assess the potential impact of both ArcR and SGO_0846 on biofilm formation.

The strains wild-type *S. gordonii* DL1, Δ*arcR* (PK3354), ΔSGO_0846 (NU06), Δ*arcR*/ΔSGO_0846 (NU09), Δ*arcR* p*arcR*$_{comp}$ (NU10) and Δ*arcR*/ΔSGO_0846 p*arcR*$_{comp}$ (NU11) were cultured for 18 h overnight in 96-well MTPs to form biofilms, in 50 % BHY medium under humid anaerobic conditions. Levels of planktonic cell growth across all strains were compared, to ensure that all strains had grown to a similar cell yield, and any differences in biofilm formation levels were therefore not due to differences in planktonic growth (Figure 5.11). Biofilms were stained with crystal violet dye, washed to remove any loosely-bound cells, and the optical density at 570 nm taken as a quantification of biofilm formation levels (Figure 5.12).

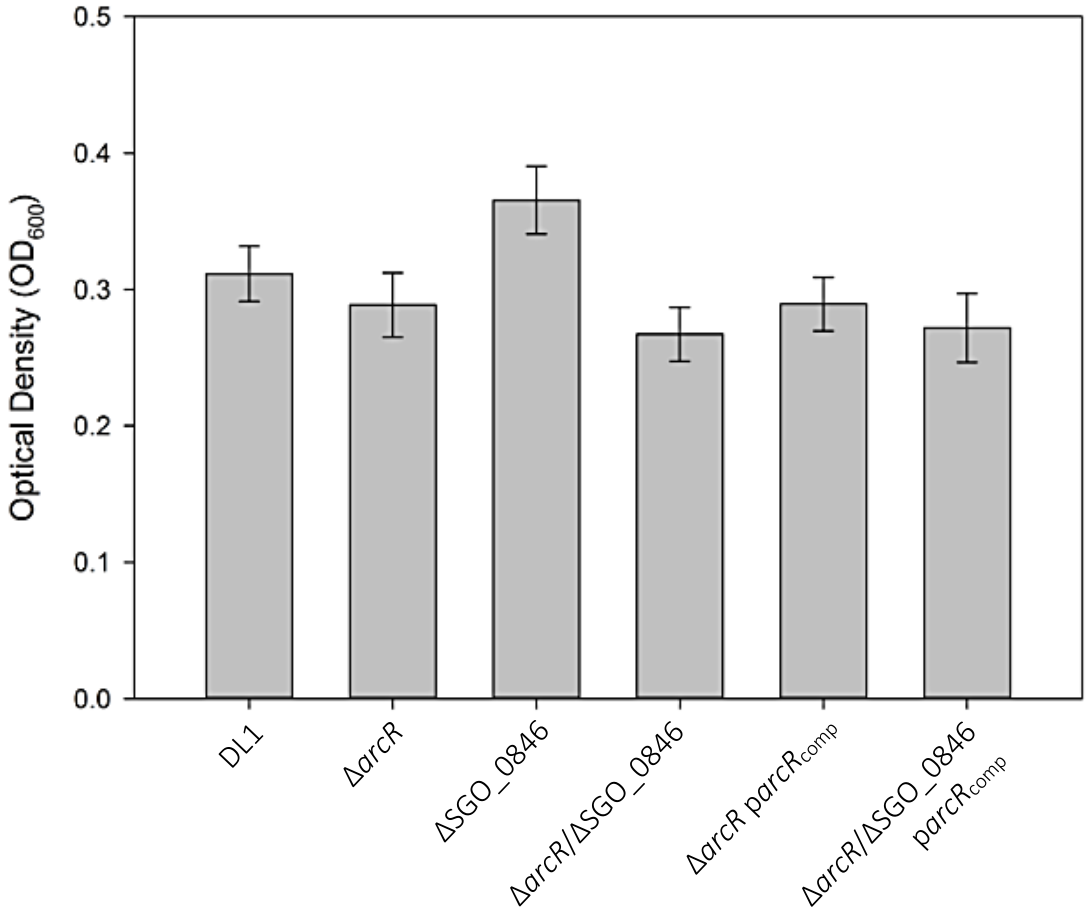


Figure 5.11. Quantification of levels of planktonic growth associated with biofilms formed by different *S. gordonii* strains. Biofilms were grown from wild-type *S. gordonii* DL1, Δ*arcR*, ΔSGO_0846, Δ*arcR*/ΔSGO_0846, Δ*arcR* p*arcR*$_{comp}$ and Δ*arcR*/ΔSGO_0846 p*arcR*$_{comp}$. The levels of planktonic cell growth in the medium above the biofilms were quantified for each strain by measuring optical density at 600 nm. Bars show arithmetic means of six replicates, error bars show SEM. Statistical tests performed, using a one-way ANOVA and two-sample T-tests comparing wild-type DL1 vs mutants, showed no significant differences in planktonic growth between the DL1 strain or any of the mutant strains. Planktonic growth was found to be similar across all strains tested, when grown in the context of a biofilm assay.

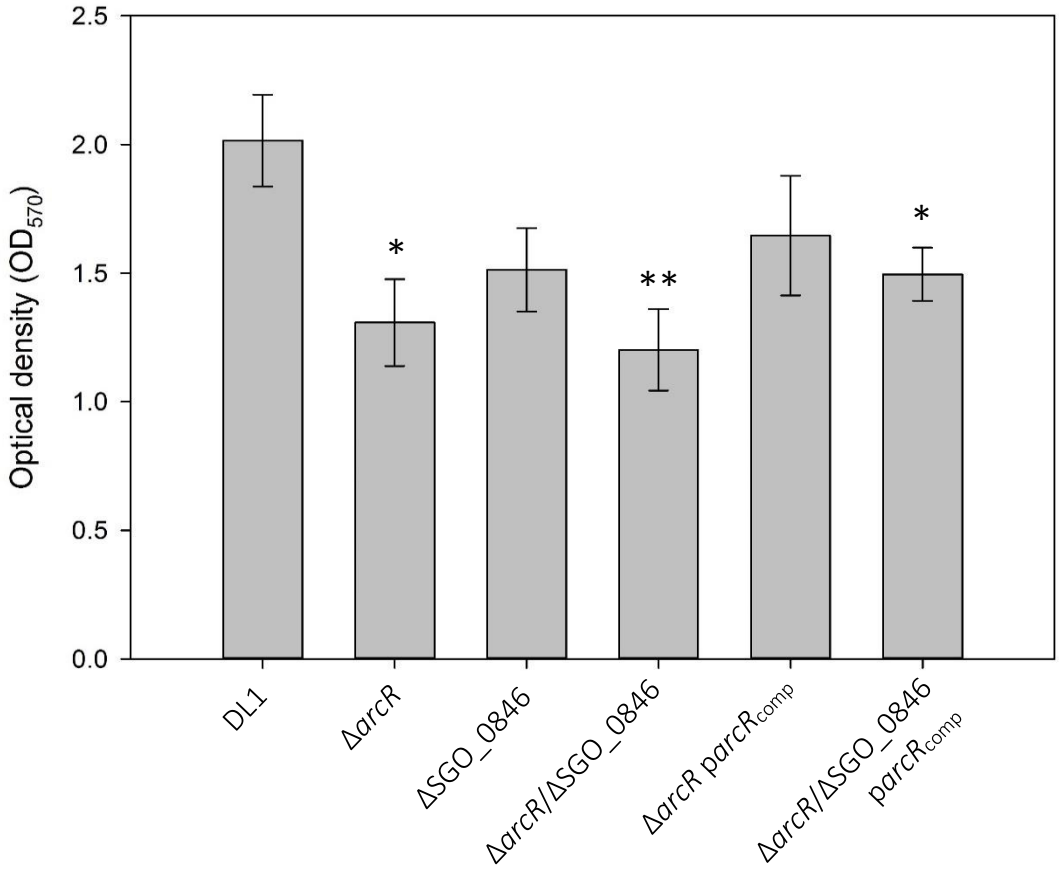


Figure 5.12. Quantification of levels of biofilm formation by *S. gordonii* strains. Biofilms were grown from wild-type *S. gordonii* DL1, Δ*arcR*, ΔSGO_0846, Δ*arcR*/ΔSGO_0846, Δ*arcR* p*arcR*$_{comp}$ and Δ*arcR*/ΔSGO_0846 p*arcR*$_{comp}$. The biofilms were stained with crystal violet, and quantified by measuring optical density at 570 nm. The Δ*arcR*, Δ*arcR*/ΔSGO_0846 and Δ*arcR*/ΔSGO_0846 p*arcR*$_{comp}$ strains all show significant impairment in biofilm formation when compared to the wild-type strain DL1. Bars show arithmetic means of six replicates, error bars show SEM. Statistical tests performed were one-way ANOVA and two-sample T-tests between DL1 vs mutants, * p <0.05, ** p <0.01. Complementation of the *arcR* gene back into the Δ*arcR* backgrounds shows a return of biofilm levels to a non-significantly different level.

Quantification of planktonic growth across all strains showed no significant differences ($p > 0.05$) between planktonic growth levels for the *S. gordonii* DL1 strain vs any of the mutant strains, which was consistent with the results of the planktonic growth experiments shown in section 5.6.1. Biofilm assays showed that although there was a small reduction in biofilm formation in *S. gordonii arcR*$_{comp}$ compared with *S. gordonii* DL1 (OD_{570} 2.02±0.18 vs 1.65±0.23), it was not statistically significant ($p > 0.05$). This indicates that complementation with plasmid-borne *arcR* restored the capacity of the Δ*arcR* strain to form biofilms. Biofilm formation by *S. gordonii* ΔSGO_0846 was also not significantly different from the wild-type (OD_{570} 2.02±0.18 vs 1.51±0.16, $p > 0.05$), although the mean level of biofilm formation was lower in comparison to the DL1 wild-type strain.

Interestingly, the Δ*arcR*/ΔSGO_0846 double mutant was the strain that had the most significantly different level of biofilm formation when compared to the wild-type ($p < 0.01$), although this was not significantly different from either the Δ*arcR* or ΔSGO_0846 single mutants (OD_{570} 1.20±0.16 for Δ*arcR*/ΔSGO_0846, 1.31±0.17 for Δ*arcR*, 1.51±0.16 for ΔSGO_0846). In addition to this, the Δ*arcR*/ΔSGO_0846 p*arcR*$_{comp}$ strain was also significantly lower in biofilm formation in comparison to *S. gordonii* DL1 (OD_{570} 1.49±0.10, $p < 0.05$). Statistical comparisons between all other combinations of strains showed no other significant differences in biofilm formation levels.

Therefore, complementation with plasmid-borne *arcR* restored biofilm-forming capacity in the Δ*arcR* background. However, deletion of the SGO_0846 gene appears to have no significant effect on the biofilm formation of *S. gordonii.* The levels of biofilm formation within the ΔSGO_0846 strain were lower than that of the wild-type, however, and the Δ*arcR*/ΔSGO_0846 strain is significantly impaired in biofilm formation. It is possible that both SGO_0846 and ArcR play a role in affecting biofilm formation to some degree, and that the effect is seen additively in the Δ*arcR*/ΔSGO_0846 strain, but further work would be needed to elucidate the exact role – if any – that SGO_0846 plays within the *S. gordonii* biofilm.

5.7. Discussion

This chapter has shown that the ArcR regulatory protein has a marked effect on biofilm formation in *S. gordonii*. A strain carrying a deletion of the *arcR* gene was shown to be significantly impaired in biofilm formation in comparison to the *S. gordonii* DL1 wild-type strain, under both aerobic and anaerobic conditions, and complementation of the Δ*arcR* background with an intact copy of the *arcR* gene resulted in a reversal of the observed attenuated biofilm phenotype.

ArcR has not been previously shown to be associated with biofilm formation, but work performed in Chapter 4 showed that it may play a role in activating expression of the biofilm-associated *bfb* locus under high arginine conditions. There is evidence that L-arginine-dependent regulators have roles in biofilm formation in Gram-positive cocci. For example, disruption of the arginine metabolism regulator AhrC of *E. faecalis* by transposon mutagenesis resulted in defective biofilms (Kristich *et al.*, 2008). It was subsequently found that *E. faecalis* AhrC promotes early cell attachment and aids in accumulation of bacterial cells within the developing biofilm, as does the ArgR protein to a lesser degree. However, it was also found that *E. faecalis* AhrC plays an important role in bacterial pathogenesis, as an Ω*ahrC* strain (carrying a transposon insertion within the *ahrC* gene) had significantly attenuated virulence in a rabbit endocarditis model (Frank *et al.*, 2013). It was postulated that AhrC expression was important early on in the establishment of an endocarditis infection, in order to promote biofilm attachment.

There is now a body of evidence indicating that arginine-dependent regulators and their target genes are important in the virulence of a range of Gram-positive cocci. For example, the arginine deiminase system (ADS), which is controlled by ArgR and ArcR homologs, has been linked to virulence in *Streptococcus pyogenes*. Cell extracts from a strain of *S. pyogenes* that was carrying a deletion of the streptococcal acid glycoprotein (SAGP), thought to have ADS activity, were unable to inhibit human peripheral blood mononucleocyte proliferation, and this strain was also less able to invade epithelial cells *in vitro* (Degnan *et al.*, 1998; Degnan *et al.*, 2000). Additionally, ADS activity has been suggested to be linked to virulence not only in *Staphylococcus aureus*, but also *Staphylococcus epidermidis* (Diep *et al.*, 2006; Fey and Olson, 2010). In these bacterial strains, it is thought that the ADS may aid bacterial virulence by

allowing survival of the cells under low pH conditions, often encountered intracellularly by invading bacteria due to the activity of lysosomes within human host cells (Myers *et al.*, 1995); and additionally by depleting free L-arginine so that it cannot be used by the immune system, alongside oxygen and NADPH, to biosynthesise nitric oxide which is released as a bacteriocidal free radical in response to bacterial infection.

It is possible that the ArcR protein of *S. gordonii* may be important for virulence in infective endocarditis, by modulating biofilm formation in a similar way to AhrC in *E. faecalis*. This would need further investigation in an appropriate animal model. The importance of *S. gordonii* ArcR for oral biofilm formation and persistence within the host oral cavity is also not clear. It would be interesting to measure expression of the *arcR* gene by qRT-PCR at different stages of growth, to see when it was most highly expressed, for example, during early cell growth and initial attachment.

The SGO_0846 gene was annotated as encoding a putative cell wall protein by NCBI, although further *in silico* analysis in this chapter contradicted this, and predicted that it is likely to be released through the surface of the cells. SGO_0846 was also predicted to form a pore-lining helix at the C-terminal end of the protein. These helices would normally line the channel of a pore through the bacterial outer membrane, however, Gram-positive bacteria such as *S. gordonii* do not have an outer membrane. Despite that, however, many pore-lining helical domains of proteins have been characterised in Gram-positive bacteria. These include the α4 helices of the insecticidal Cry toxin of *Bacillus thuringiensis*, which is released in crystal form as a protoxin around the bacterial spores (Girard *et al.*, 2008); the M1 and M2 helices of KcsA, a potassium channel of *Streptomyces lividans* (Shrivastava *et al.*, 2000); and transmembrane segment 8 of the GltT glutamate transporter of *Bacillus stearothermophilus* (Slotboom et al., 2001).

It is possible that, as the PSIPRED server which predicted the pore-lining helical region will use other proteins as templates for modelling and predicting topology of the target protein, it may have incorrectly predicted this pore-lining helical region. This would also be supported by the fact that no wall-anchoring motifs were predicted anywhere in the SGO_0846 amino acid sequence. However, the motifs for wall-anchoring of a protein are not always known, and there

is precedence for pore-lining helices in Gram-positive proteins, as mentioned above. Cysteine-scanning mutagenesis (Frillingos *et al.*, 1998), to create mutants of the SGO_0846 protein, would allow us to test whether the predicted pore-lining helical region of SGO_0846 was actually involved in the activity of the protein, as it would be if the protein was acting as a porin molecule, possibly involved in aiding virulence during opportunistic infection. However the role of SGO_0846 within *S. gordonii* cells would have to be known before this could be tested, in order to observe any differences in phenotype due to altered activity of a mutant form of the protein.

The predicted disordered region in the central domain of the SGO_0846 protein is indicative of an ability to change the conformation of the protein structure. Disordered proteins have been suggested to play an important role in cell signalling amongst other things, but are relatively uncommon in prokaryotic organisms (Ward *et al.*, 2004b). Using the same DISOPRED disorder prediction analysis as performed in this chapter, Ward *et al.* (2004) found that out of 35,000 bacterial protein sequences analysed, the percentage of frequency for finding regions of disorder longer than 50 amino acid residues was only 1.6%. The SGO_0846 protein was predicted by DISOPRED to contain a region of disorder consisting of 190 contiguous residues, in addition to other shorter regions predicted to be protein-binding disordered regions.

It is thought that the ability of disordered proteins to change their conformational shape makes them more adept at binding molecules such as other proteins, whilst avoiding any changes in binding affinity; resulting in high specificity and low affinity interactions (Dogan *et al.*, 2014). Therefore it is possible that, if SGO_0846 is indeed being secreted from the *S. gordonii* cells into the cell supernatant, it could play a role in signalling by binding other proteins, or even surface-associated molecules. This would be contrary to the suggestion above that the pore-lining helical region may be contributing to the activity of the SGO_0846 protein. If SGO_0846 was playing a role in cell signalling, this could in turn aid biofilm formation, although this role is clearly not mediated by the ArcR protein as deletion of the SGO_0846 gene in the *S. gordonii* wild-type background, containing an intact copy of *arcR*, did not cause biofilm attenuation.

Analysis of the biofilm phenotype of the ΔSGO_0846 mutant showed that deletion of the SGO_0846 gene did not cause any significant differences in levels of biofilm formation between

the wild-type DL1 strain and the ΔSGO_0846 strains, although the average level of biofilm formation in the ΔSGO_0846 strain was found to be slightly lower than the wild-type. A double mutant of *arcR* and SGO_0846 showed a greater reduction in biofilm formation than the ΔSGO_0846 single mutant. Therefore, it is possible that, following further analysis, the defective biofilm phenotype of the *S. gordonii* Δ*arcR* strain may instead be due to the apparent role it plays in regulating the *bfb* genes, shown in Chapter 6, which are known to be necessary for *S. gordonii* biofilm formation (Kiliç *et al.*, 2004; Jakubovics *et al.*, 2008), as it is clearly not mediated by overexpression of the SGO_0846 protein.

If SGO_0846 is indeed a released molecule, as predicted by the LocateP server, it could potentially act as a signalling molecule between the bacterial cells in order to promote biofilm formation independently of ArcR, or recruit other cells to join the developing biofilm. This hypothesis could further be supported by the discovery of the disordered region within the SGO_0846 protein, often prevalent in extracellular signalling proteins. Other bacterial species have been shown to secrete proteins in order to promote or enhance biofilm formation before – both the LapA protein of *Pseudomonas fluorescens*, and the BapA protein of the *Salmonella enterica* serovar Enteritidis are secreted, surface-associated proteins that promote increased bacterial biofilm biomass and adhesion to the colonising surface (Hinsa *et al.*, 2003; Latasa *et al.*, 2005). SGO_0846 may possibly play a role within *S. gordonii* biofilms analagous to the ones that the BapA and LapA proteins play in these other species, although the single ΔSGO_0846 mutant did not show a significant decrease in biofilm formation levels in comparison to the DL1 wild-type strain. Furthermore, it seems unlikely that overexpression of a protein involved in recruiting cells to the biofilm would result in lower levels of biofilm formation, as were seen in the ΔSGO_0846 mutant. Additionally, it is unlikely that a protein released from the cell surface plays a major role in adhesion of cells within the biofilm.

It is possible that SGO_0846 may instead play a role in biofilm structuring. Enzymes involved in shaping the structure of biofilms are responsible for dispersing cells in order to form channels, allowing nutrients and water to reach cells at the base of the biofilm. They can also be responsible for detachment of cells from the surface of the biofilm to allow dissemination of cells to other sites, as is the case with the phenol-soluble modulin peptides (PSMs) of *Staphylococcus aureus*. These toxins are controlled by the Agr quorum sensing system, and

have been shown to be necessary for the formation of channels within the biofilm, and dissemination of cells to other sites *in vivo*, with mutants displaying defective biofilm structures, decreased cell detachment and increased biofilm thickness (Periasamy *et al.*, 2012). Additionally, many other bacteria also produce peptides that play roles in cell detachment, such as the dispersin B protein DspB of the oral pathogen *Aggregatibacter actinomycetemcomitans*, which hydrolyses *N*-acetylglucosamine residues, or the SPRE enzyme of *S. mutans* NG8, which releases proteins from the surface of the cells, thereby aiding biofilm dissemination (Lee *et al.*, 1996; Kaplan *et al.*, 2003). If SGO-0846 is indeed a secreted molecule, rather than being involved in recruiting cells to the biofilm, it may be involved in dispersing cells from the surface of the biofilm, although again, this would not explain the reduced levels of biofilm formation in the ΔSGO_0846 mutant.

It is also important to note that further analysis of the microarray data displayed in Chapter 4 indicated that SGO_0846 was not expressed highly enough to produce a significant signal within either the Δ*argR* or Δ*ahrC* microarrays, indicating that it may be a lowly-expressed gene. Any small up- or down-regulation of a lowly-expressed gene would result in a large fold-change, which may explain the 148-fold up-regulation observed in SGO_0846 in the Δ*arcR* strain. However, levels of gene expression are not an indication of absolute protein levels, and so until more is known about SGO-0846 it is hard to say what effect a large up-regulation in gene expression may have on the levels of protein expression. Levels of SGO_0846 protein expression in the *S. gordonii* wild-type strain could be analysed using tandem mass spectrometry with selected reaction monitoring, and then compared to levels found in the Δ*arcR* strain.

Clearly, despite the *in silico* analysis of the SGO_0846 protein thus far, more work would be needed to elucidate the exact role that it plays within *S. gordonii* cells. Experiments such as immunolocalisation (Chikwamba *et al.*, 2003) using anti-SGO_0846 antibodies, or fluorescent protein fusions to SGO_0846 (Beilharz *et al.*, 2015), could be used to determine whether the protein is localised within the cell wall, cell membrane or cell supernatant, as with the *P. fluorescens* LapA protein or the DspB enzyme of *A. actinomycetemcomitans*. Further work could also be performed on adherence and aggregation and biofilm structure in *S. gordonii*, in order to determine whether its role is focused more on cell binding or cell dispersal. This work could

include testing auto-aggregation (between *S. gordonii* cells), and co-aggregation with other species (such as *Actinomyces oris*), to determine whether SGO_0846 may play a role in cell-cell binding or signalling; or using confocal laser scanning microscopy, in conjunction with a microfluidics system, to determine whether deletion of SGO_0846 is affecting the formation of the biofilm structure. It is also important to determine the mechanism by which ArcR affects biofilm formation, for example, by using DNA electrophoretic mobility shift assays (EMSA) or chromatin immunoprecipitation (ChIP) assays to find where the ArcR protein may be binding within the *S. gordonii* genome.

In summary, the work in this chapter of the thesis showed clearly that the ArcR protein is critical for biofilm formation within *S. gordonii* cells. Additionally, characterisation was begun upon a new protein, SGO_0846, which is strongly regulated by the ArcR regulator of *S. gordonii*.

6. General discussion

The L-arginine deiminase system (ADS) of *S. gordonii* was first described in 1983, by Ferro *et al.* (1983). Subsequent analysis determined that ADS expression was controlled by an L-arginine-dependent transcriptional activator, called ArcR (Dong *et al.*, 2002). ArcR is paralogous to two other L-arginine-dependent regulators of *S. gordonii*, ArgR and AhrC, which have been shown to control expression of L-arginine biosynthesis genes in this species (Jakubovics *et al.*, 2015). Before the work carried out in this thesis, a biofilm-defective phenotype had been observed in the Δ*arcR* mutant strain, and so, in order to determine the basis for the defect, global gene regulation had already been analysed using microarray experiments in this strain. However, global gene expression had not yet been analysed in the Δ*argR* or Δ*ahrC* strains. Therefore, it was not clear how arginine metabolism may be linked to biofilm formation in *S. gordonii*. Finally, the impact of L-arginine depletion upon global gene expression in *S. gordonii* cells had not previously been investigated.

6.1. Summary of thesis

The thesis began with an investigation into the effects of L-arginine on the growth of *S. gordonii* planktonic cells (Chapter 3). Prior to this work, the direct impact of L-arginine on planktonic growth of cells had not been studied in detail. *S. gordonii* may be subjected to large fluctuations in the external L-arginine concentration during growth in the oral cavity, as although the concentration of free L-arginine within saliva is 5-10 µM (Brand *et al.*, 1997), L-arginine is also currently added to oral healthcare products such as toothpastes at concentrations of up to 460 mM (8%) (Sullivan *et al.*, 2014).

Analysis of planktonic growth in chemically-defined medium showed no differences in growth of *S. gordonii* cells under all L-arginine concentrations, except for the cells in 500 mM L-arginine, which had a lower overall growth yield in comparison to the other samples. The survival of *S. gordonii* planktonic cells in saliva was not affected by prolonged exposure to high 500 mM L-arginine concentrations, indicating that cells retained their viability and that L-arginine is not toxic to *S. gordonii* at high concentrations. Further investigations into growth in *S. gordonii* focused on amino acid depletion, with planktonic cultures depleted of L-arginine, L-histidine and the branched-chain amino acids showing an arrest in growth following depletion. This is

likely due to cells entering a period of hibernation in response to nutrient starvation. These types of responses are likely formed in order to allow the cells to survive longer periods of time in nutrient-deprived conditions – for example, cells found in the interior and at the base of dental plaque biofilms (such as *S. gordonii*) often go through periods of localised nutrient deprivation due to limited diffusion of nutrients through to the base of the biofilm (Stewart, 2003). A hibernation-type response under starvation conditions would allow the cells to survive for longer periods of time. A similar response has also been observed in *E. coli* cells in response to isoleucine deprivation (Traxler *et al.*, 2008). Therefore, the findings from this chapter suggested a form of functional auxotrophy in response to L-arginine depletion in *S. gordonii*, which was then further investigated by global gene expression analysis in Chapter 4.

Investigation into expression of different genes under arginine depletion conditions using qRT-PCR analysis showed that a number of genes involved in adhesion and biofilm formation were differentially expressed in response to arginine depletion, and that whilst some of these responses appeared to be part of a general stress response, others were apparently arginine-specific responses in gene expression. General stress responses in bacteria are triggered by unfavourable changes in the exogenous growth environment, such as oxidative stress, heat shock or nutrient starvation. These responses cause a switch of gene expression profile within the cells from one associated with cell growth and replication, to one that allows cells to enter a prolonged stationary phase and thereby increases their chance of surviving the unfavourable conditions (Traxler *et al.*, 2008). The regulation of genes as part of an amino acid starvation response is likely caused by the alarmone molecule (p)ppGpp, which is known to specifically change expression of genes in response to stress or starvation conditions, in order to aid cell survival (Betts *et al.*, 2002; Miethke *et al.*, 2006). It is particularly known for downregulating the expression of genes involved in metabolism and replication, and upregulating genes involved in biosynthesis and stationary phase survival.

As for the arginine-specific responses in gene expression, whilst arginine is known to functionally affect binding, adhesion and biofilm formation within other species, it has not previously been shown to affect these functions within *S. gordonii* (Kaplan *et al.*, 2009; Frank *et al.*, 2013; Cusumano *et al.*, 2014; Zhuo *et al.*, 2015). The arginine-specific differential expression of these genes is likely due to regulation of these genes either directly by one of the

arginine-dependent regulators, or indirectly by a downstream signalling cascade. It is possible that *S. gordonii* cells may see exogenous arginine depletion as a sign of subsequent nutrient deprivation, and so switch off expression of genes involved in adhesion and biofilm formation in order to conserve resources for the next period of nutrient repletion and growth. This would be consistent with findings in other bacterial species, such as *P. aeruginosa*, which indicate that nutrient starvation and metabolic substrate depletion are linked to decreased biofilm formation, and to detachment of cells from the biofilm (Hunt *et al.*, 2004). Additionally, L-arginine has also been found to act as an energy source for species such as *Lactococcus lactis* and *Streptococcus anginosus* following carbohydrate depletion, so depletion of L-arginine and other amino acids from the chemically-defined minimal growth medium (as seen in experiments performed in this chapter) may trigger a hibernation-type stress response from the cells (Moat and Foster, 1995; Stuart *et al.*, 1999; Chávez de Paz *et al.*, 2008). Taken together, the experiments performed in this chapter all suggest that L-arginine has a clear effect on both planktonic growth, and biofilm formation ability in *S. gordonii*, which could potentially impact upon dental plaque formation *in vivo*.

Chapter 4 aimed to look into the extent of regulation by each of the three arginine-dependent regulator proteins of *S. gordonii*, by genome-wide microarray analysis. With the exception of ArcR (Dong *et al.*, 2002; Zeng *et al.*, 2006), little was known about these regulators and their effects on genes other than those involved in arginine metabolic systems. Comparison of these microarrays aimed to determine the subset of genes that were regulated by the ArcR, ArgR and AhrC regulators. However, the arrays were run separately, with only two strains on each array (i.e. DL1 and Δ*arcR* in the first, Δ*argR* and Δ*ahrC* in the second), and no mixing of samples between the two. Unfortunately, there was a significant batch effect observed. This meant that due to technical variation between the experiments, the data could not be directly compared between the first and second sets of arrays. Despite this, certain genes were found to be either consistently regulated in response to arginine depletion, or in response to deletion of one of the three regulator proteins. These included the apparent regulation of the *argBCDJ*, *argGH*, $pyrA_a$ and $pyrA_b$, SGO_1102, 1656 and 0021 genes by the ArgR and AhrC regulators; and the activation of the biofilm-associated *bfb* operon and SGO_1575 genes, and the *wefE* and *glf* RPS biosynthesis genes, by the ArcR protein.

Direct comparisons between the strains that were assessed within the same microarray experiments were also made, showing that between the ArgR and AhrC regulators, there were no differences in gene regulation. This suggests that these two proteins either work together as a regulatory complex, or that one is necessary for regulating the other, as suggested by Jakubovics *et al.* (2015). Other bacterial species have been shown to also express more than one arginine-dependent regulator, and, in the case of species where there are two ArgR/AhrC-like regulators working together, it is often found that one is responsible for activating or limiting the effects of the other, or that they work partially independently of each other and regulate different genes in response to the same stimuli (Barcelona-Andrés *et al.*, 2002; Larsen *et al.*, 2004). *S. gordonii* appears to be the first instance of a bacterial species where the ArgR and AhrC regulators overlap exactly in their regulation. In order to confirm whether these proteins do indeed work together in a complex, further work on protein structure and interactions, such as protein crystallography, would have to be performed. Furthermore, comparison of the *S. gordonii* wild-type DL1 and Δ*arcR* strains showed that one gene, SGO_0846, was highly up-regulated under high arginine conditions within the Δ*arcR* mutant strain in comparison to wild-type DL1, but was not highly expressed in either the Δ*argR* or Δ*ahrC* strains.

Chapter 5 aimed to investigate the SGO_0846 gene, and determine whether this was the basis for a biofilm defect that was found in the *S. gordonii* Δ*arcR* mutant strain, but not in either the Δ*argR* or Δ*ahrC* strains. It was hypothesised that the up-regulation of the SGO_0846 gene, seen in the Δ*arcR* microarray analysis in the previous chapter, could be responsible for the defective biofilm phenotype. As SGO_0846 was an uncharacterised protein-encoding gene, an *in silico* analysis of both the gene and predicted protein structure was carried out in order to make some preliminary predictions about gene expression and protein structure and localisation, which might help in the design of experiments to test the role of SGO_0846 in the biofilm defect of *S. gordonii* Δ*arcR*.

An isogenic *S. gordonii* mutant strain of SGO_0846 was constructed, and *S. gordonii* Δ*arcR* was genetically complemented with a plasmid-borne copy of the *arcR* gene under a strong constitutive promoter. No significant differences in biofilm formation were observed between *S. gordonii* DL1 (wild-type) and ΔSGO_0846 strains, indicating that SGO_0846 was not

responsible for the defective phenotype found in the Δ*arcR* mutant. Complementation of the Δ*arcR* gene reversed the biofilm-defective phenotype observed in its isogenic mutant, indicating that ArcR was not operating through SGO_0846 to control biofilm formation in *S. gordonii*. However, it is possible that ArcR may be affecting biofilm formation through the apparent role it plays in regulating the *bfb* locus, shown by previous research to contribute to biofilm formation in *S. gordonii*. Therefore, further investigation, such as studying the subcellular localisation of ArcR, and searching for interacting proteins using a two-hybrid system would be needed to elucidate the mechanism by which ArcR is affecting biofilm formation in *S. gordonii*, and also to determine what role the SGO_0846 protein may be playing within the cells. At present, it is unclear why SGO_0846 is up-regulated under high arginine in *S. gordonii* Δ*arcR*.

6.2. Thesis findings and future work

In conclusion, this thesis has demonstrated that arginine has a clear effect on gene expression, growth and biofilm development in *S. gordonii*, mediated both directly by arginine itself, and also through the actions of the arginine-dependent regulatory proteins.

S. gordonii was shown to have a clear functional auxotrophy for L-arginine, as cells appeared to be unable to synthesise their own arginine under aerobic conditions, whilst growing in monoculture. This may potentially be mediated by carbon-catabolite repression of the anabolic functions of the ArcB protein under aerobic, no arginine conditions. To elucidate whether the foundation of this functional auxotrophy was indeed repression of ArcB function by the CcpA regulator, growth experiments could be performed on *S. gordonii* cells growing in aerobic, arginine-deplete conditions, with primary carbon sources (e.g. glucose) also depleted from the growth medium. The supply of only secondary sources of carbon (e.g. galactose) should alleviate the effects of CCR, and may in turn allow aerobic growth in the absence of arginine to resume. Alternatively, the growth of *S. gordonii* in low arginine could be studied in a *ccpA* knockout mutant, in which carbon catabolite repression is not present.

Depletion of different amino acids was found to have effects on the expression of different genes of interest, some of which are involved in the cell surface, adhesion and biofilm formation. Some of these effects were found to be equal following depletion of all three sets

of amino acids, however, some were specific to a particular amino acid. In the case of the genes that were equally-regulated in response to depletion, these are likely to be part of a general stringent response by the cells. This could be tested experimentally either by depletion of other amino acids from the growth medium, to see if the changes in expression of these genes remained the same irrespective of which amino acid was depleted; or further by creating a strain of *S. gordonii* where the *relAPQ* genes responsible for the production of the stringent response molecule (p)ppGpp were knocked-out, and analysing whether the regulation of these genes of interest in response to amino acid depletion was subsequently lost.

The ArgR, ArcR and AhrC arginine-dependent regulators were shown to specifically regulate the expression of different genes in response to arginine concentration, and potential regulons for each regulator were compiled from the data obtained in this thesis. However, more work is needed to confirm the findings of the gene expression microarrays – for example, potential cross-comparison between the microarrays using rank product analysis (Breitling *et al.*, 2004), in order to confirm the predicted regulons for ArgR, ArcR and AhrC that were laid out. Furthermore, qRT-PCR analysis could be performed on the genes that appear to form these regulons in the different regulator deletion strains, to see whether qRT-PCR analysis confirms the same changes in expression of these genes in response to loss of one of the arginine-dependent regulators as seen in the microarray. The question of whether these genes are in fact directly regulated by the arginine-dependent regulators could also be addressed, by electrophoretic mobility shift assays, chromatin immunoprecipitation assays and DNase footprinting assays, to show direct binding of the regulatory proteins to the DNA sequences for the genes composing the putative regulons.

Further evidence was also found to support the theory that the ArgR and AhrC proteins are orthologs that work together as a complex in order to regulate gene expression in response to exogenous arginine concentrations, as the apparent regulons of these two genes overlapped perfectly. To test whether this was indeed the case, co-immunoprecipitation assays could indicate whether these two proteins interact with each other, and subsequent x-ray crystallography analysis could confirm this, and show the specific conformation of the proteins. This could also be used to test for protein conformation and interaction under different arginine

conditions, to see whether ArgR and AhrC only interact under certain situations or conditions in order to regulate gene expression.

The ArcR regulatory protein has been shown to not only control expression of metabolic genes, but also biofilm formation in *S. gordonii*, likely through secondary regulation mechanisms which are yet to be determined. These may include regulation of the *bfb* cellobiose PTS genes, which have been linked to biofilm formation in *S. gordonii* before. In order to test for direct interaction between the ArcR protein and the *bfb* genes, amongst others, DNase footprinting assays, electrophoretic mobility shift assays (EMSA) or chromatin immunoprecipitation (ChIP) assays could be employed to map the ArcR recognition sequences in the *S. gordonii* genome, as the binding footprint of ArcR is not well-characterised.. If one such binding site is found within the *bfb* genes, it seems likely that their differential regulation would be responsible for the Δ*arcR* strain defective biofilm phenotype.

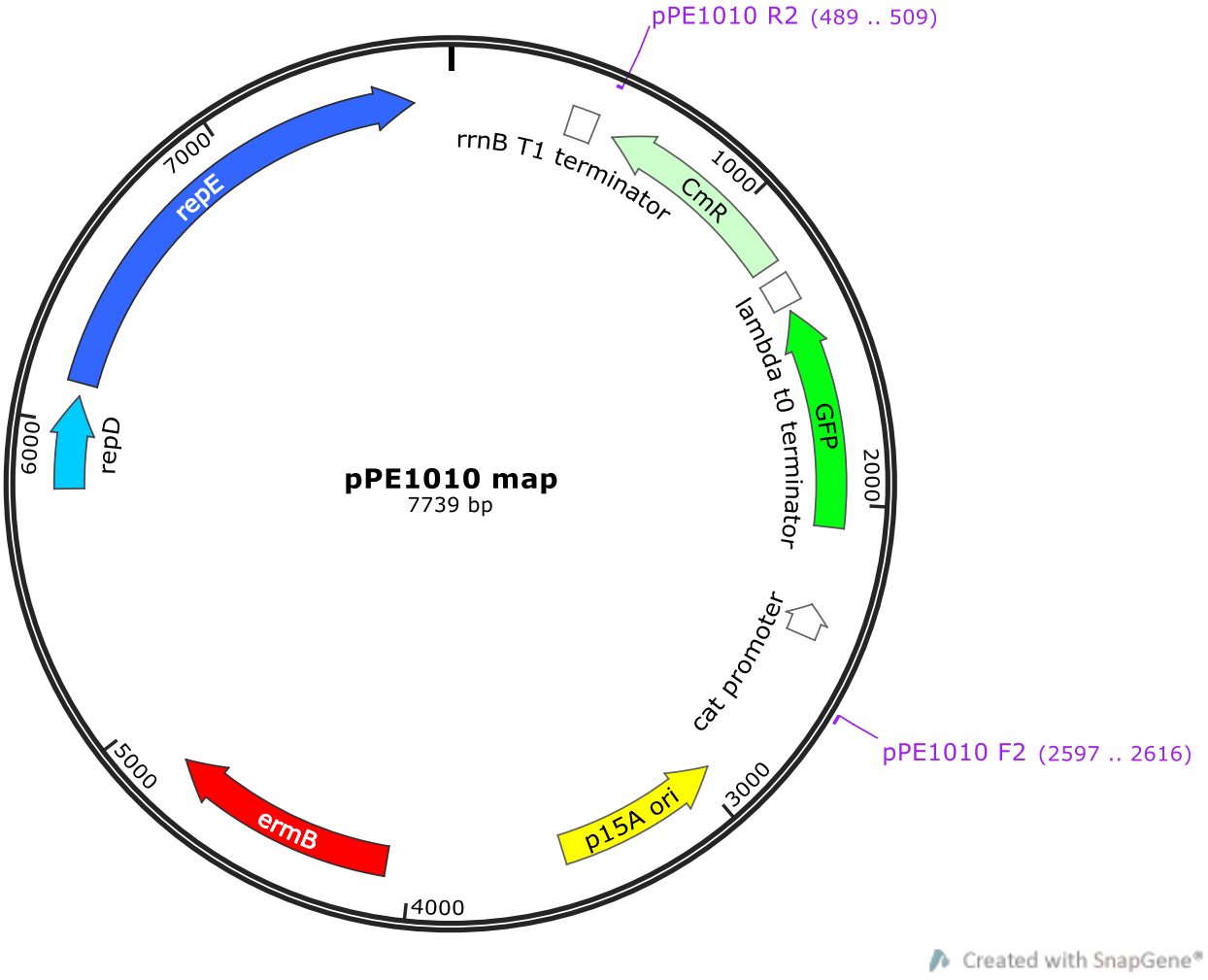
Finally, a previously uncharacterised gene of *S. gordonii* (SGO_0846) was partially characterised through *in silico* analysis and biofilm assays using an isogenic mutant. The *in silico* analysis indicated that SGO_0846 is likely to be released from the surface of the cells, and biofilm quantification assays demonstrated that it does not play a major role in ArcR-mediated biofilm formation. These observations could form a basis for future work, determining the subcellular localisation of SGO_0846 in *S. gordonii* cells, using immunolocalisation or fluorescent protein fusion experiments. Moreover, the contribution, if any, of SGO_0846 to either biofilm formation (through accumulation of cells to the biofilm), or biofilm structuring and dissemination could be tested through experiments that look at aggregation of *S. gordonii* cells with itself and other bacterial species, and the use of microfluidics systems alongside confocal microscopy and quantitative image analysis to observe the differences between biofilms formed by wild-type and ΔSGO_0846 strains.

It is possible that, in the future, arginine sensing by *S. gordonii* could be utilised as a tool to control bacterial biofilm formation within the oral cavity – for example, by targeting ArcR in order to inhibit biofilm formation, or utilising high concentrations of arginine, shown to reduce biomass when applied to a biofilm model (Kolderman *et al.*, 2015), to promote a non-biofilm form of growth for *S. gordonii*, allowing planktonic cells to be cleared from the oral cavity in

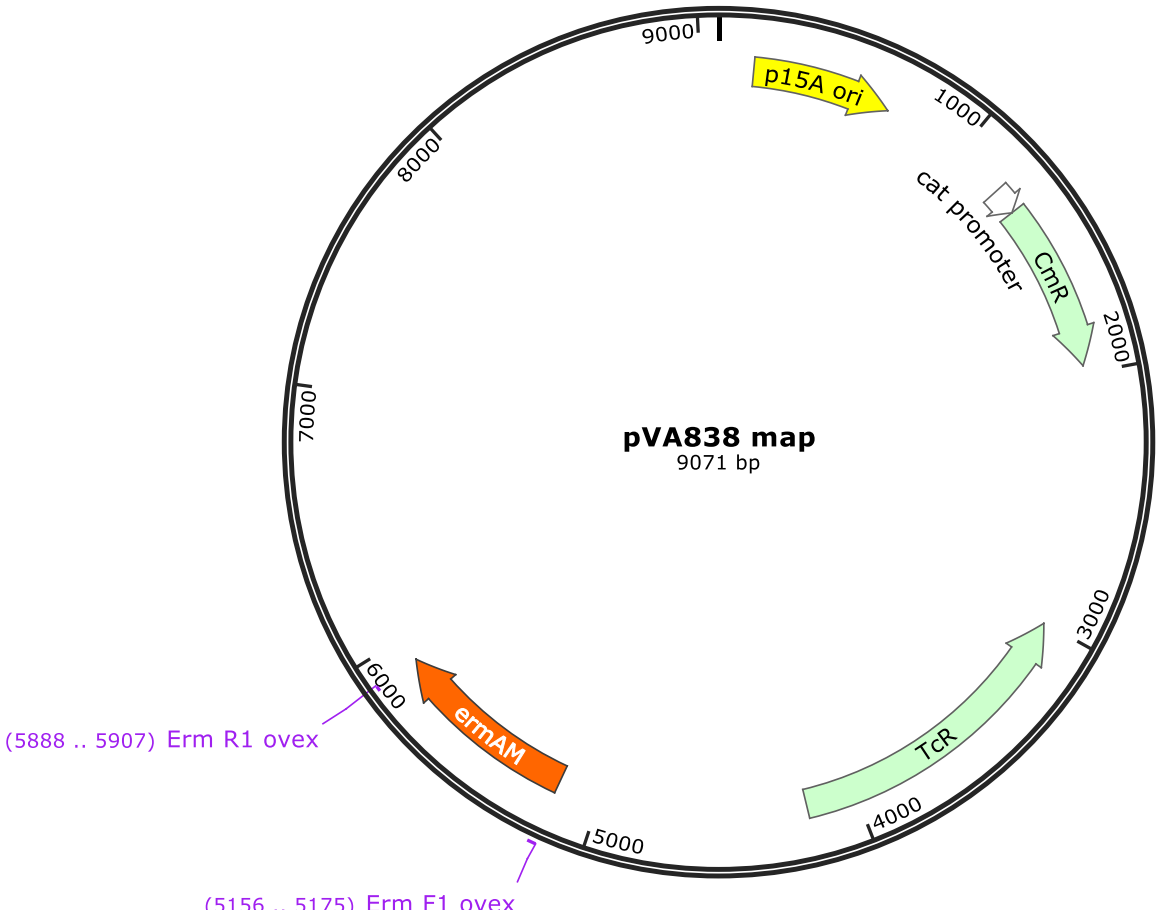
saliva. Alternatively, if SGO_0846 is indeed a biofilm restructuring molecule, overexpression of this protein may encourage excess detachment of cells from the surface of the biofilm, allowing planktonic cell clearance from the oral cavity by salivary flow. Together, these strategies all may help to control initiation of biofilm formation on the tooth surfaces, and help us to control the incidence of oral diseases such as dental caries or periodontitis.

Appendix A: Vector maps

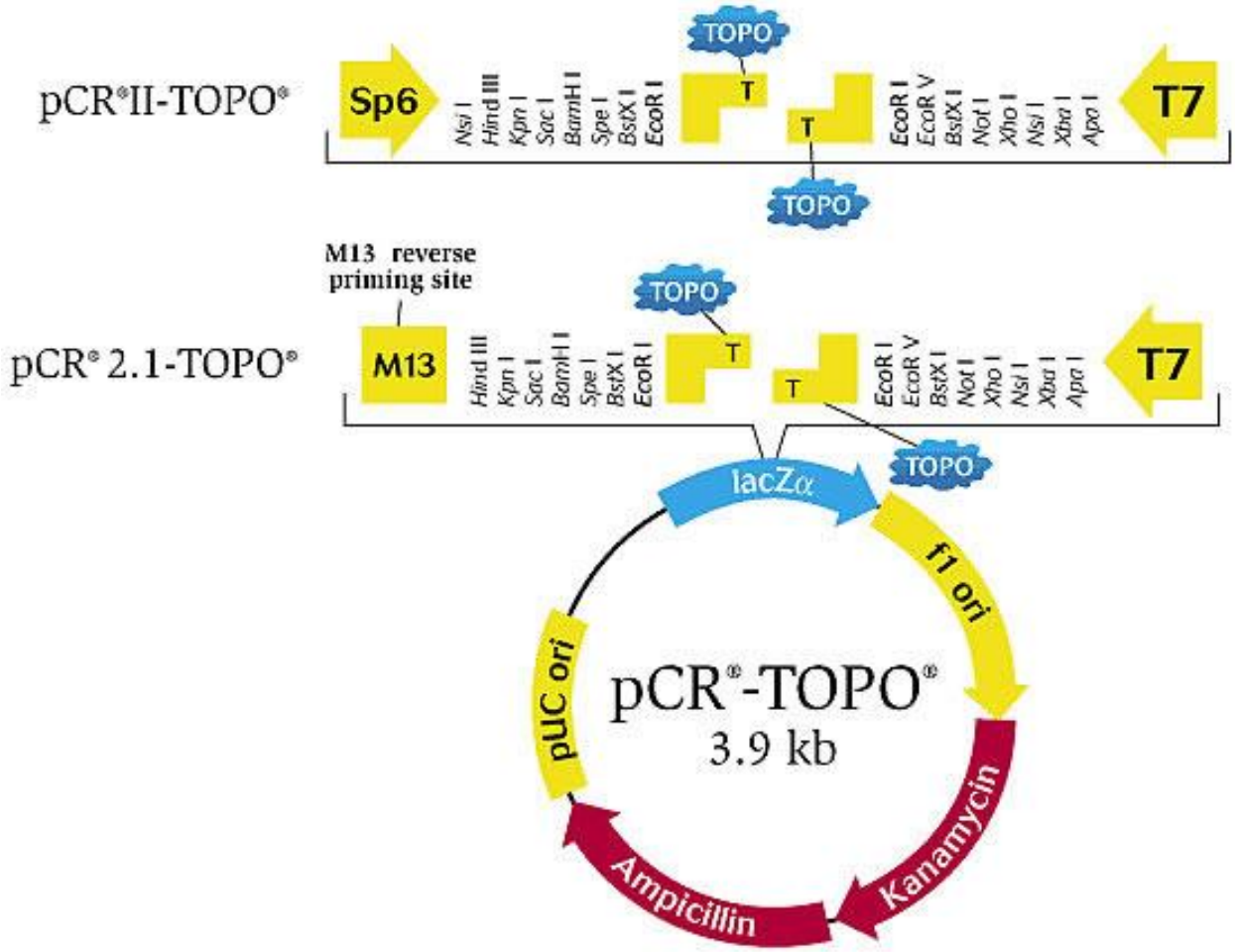
pPE1010



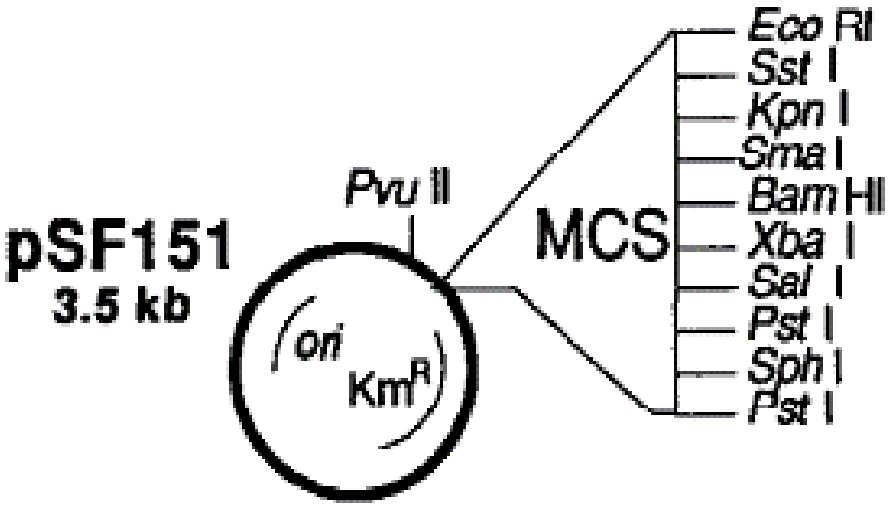
pVA838


9000
1000
2000
3000
4000
5000
6000
7000
8000
p15A ori
cat promoter
CmR
pVA838 map
9071 bp
TcR
ermAM
(5888 .. 5907) Erm R1 ovex
(5156 .. 5175) Erm F1 ovex

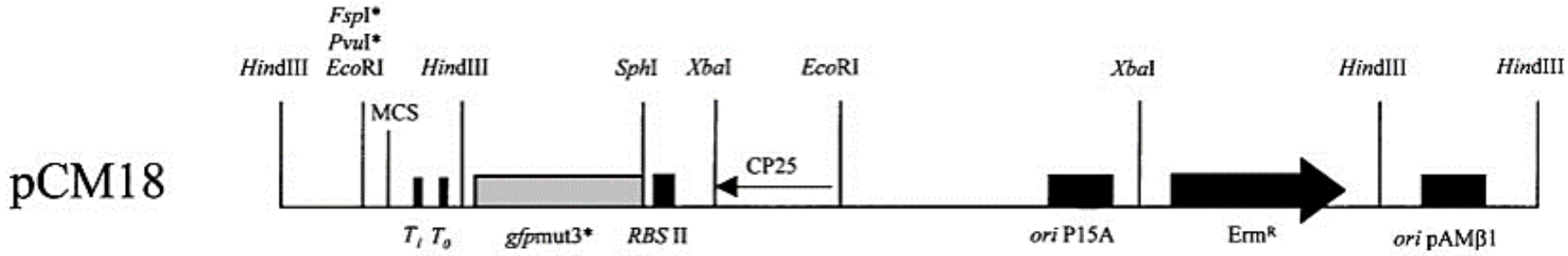
pCR2.1


pCR®II-TOPO®
Sp6
T7
TOPO
T
EcoR I
EcoR V
BstX I
Not I
Xho I
Nsi I
Xba I
Apa I
Nsi I
Hind III
Kpn I
Sac I
BamH I
Spe I
BstX I
EcoR I
pCR® 2.1-TOPO®
M13 reverse priming site
M13
lacZα
f1 ori
pUC ori
pCR®-TOPO®
3.9 kb
Ampicillin
Kanamycin

pSF151


pSF151
3.5 kb
Pvu II
ori
KmR
MCS
Eco RI
Sst I
Kpn I
Sma I
Bam HI
Xba I
Sal I
Pst I
Sph I
Pst I

pCM18


pCM18
FspI*
PvuI*
HindIII
EcoRI
HindIII
SphI
XbaI
EcoRI
XbaI
HindIII
HindIII
MCS
CP25
T1 T0
gfpmut3*
RBS II
ori P15A
ErmR
ori pAMβ1

Appendix B: Microarray data

S. gordonii DL1 microarray data, regulation of genes in response to arginine depletion

Locus tag	Gene name	FC (abs)
SGO_1569	*argC*	520.12
SGO_0177	SGO_0177	342.42
SGO_1568	*argJ*	319.95
SGO_1567	*argB*	306.44
SGO_0175	*argG*	269.33
SGO_1566	*argD*	260.85
SGO_0176	*argH*	207.89
SGO_0648	SGO_0648	54.12
SGO_1656	SGO_1656	44.43
SGO_0647	SGO_0647	42.30
SGO_0091	SGO_0091	36.62
SGO_0646	SGO_0646	36.26
SGO_0645	SGO_0645	32.39
SGO_1105	SGO_1105	28.94
SGO_1106	SGO_1106	28.46
SGO_0021	SGO_0021	26.67
SGO_1103	*carA*	23.70
SGO_1104	*carB*	21.43
SGO_1102	SGO_1102	18.72
SGO_1279	SGO_1279	17.98
SGO_1410	SGO_1410	17.01
SGO_1408	*hisD*	15.18
SGO_1409	*hisG*	15.14
SGO_1278	SGO_1278	14.26
SGO_1411	*hisC*	13.42
SGO_1407	*hisB*	12.45
SGO_0092	SGO_0092	12.28
SGO_1406	*hisH*	11.15
SGO_0093	SGO_0093	10.64
SGO_1835	SGO_1835	10.55
SGO_0874	SGO_0874	10.11
SGO_1832	SGO_1832	9.83
SGO_0090	SGO_0090	9.79
SGO_0301	SGO_0301	9.60
SGO_1833	SGO_1833	9.40
SGO_0846	SGO_0846	9.35
SGO_1831	SGO_1831	9.20
SGO_1137	SGO_1137	9.13
SGO_0624	SGO_0624	8.86
SGO_1282	SGO_1282	8.14
SGO_1405	*hisA*	7.95

SGO_1317	SGO_1317	6.90
SGO_0427	SGO_0427	6.77
SGO_1716	SGO_1716	6.21
SGO_1403	*hisIE*	6.14
SGO_0845	SGO_0845	6.11
SGO_1589	*arcT*	6.10
SGO_1404	*hisF*	5.87
SGO_1752	SGO_1752	5.83
SGO_1485	SGO_1485	5.81
SGO_0496	*rgg*	5.55
SGO_0089	SGO_0089	5.49
SGO_1402	*hisE*	5.35
SGO_1830	*mga*	5.31
SGO_1126	*xerS*	5.30
SGO_0337	*trx*	5.27
SGO_0572	SGO_0572	4.99
SGO_1107	SGO_1107	4.91
SGO_0682	SGO_0682	4.91
SGO_0334	SGO_0334	4.91
SGO_0602	SGO_0602	4.88
SGO_1194	SGO_1194	4.81
SGO_1729	SGO_1729	4.62
SGO_1196	*satD*	4.53
SGO_1110	SGO_1110	4.52
SGO_1564	SGO_1564	4.51
SGO_0094	SGO_0094	4.48
SGO_0709	SGO_0709	4.37
SGO_1401	SGO_1401	4.36
SGO_0178	*rnpA*	4.22
SGO_0687	SGO_0687	4.20
SGO_1195	SGO_1195	4.20
SGO_1563	SGO_1563	4.19
SGO_1590	*arcD*	4.13
SGO_0571	SGO_0571	4.12
SGO_0179	SGO_0179	4.08
SGO_1134	SGO_1134	4.08
SGO_0157	SGO_0157	4.04
SGO_1751	SGO_1751	4.02
SGO_0883	SGO_0883	4.01
SGO_0562	SGO_0562	3.98
SGO_0397	SGO_0397	3.98
SGO_1138	SGO_1138	3.96
SGO_0665	SGO_0665	3.92
SGO_0589	SGO_0589	3.90
SGO_0601	SGO_0601	3.90
SGO_1280	SGO_1280	3.89

SGO_1529	SGO_1529	3.85
SGO_0957	SGO_0957	3.85
SGO_1281	SGO_1281	3.83
SGO_0360	SGO_0360	3.82
SGO_0688	SGO_0688	3.77
SGO_2084	SGO_2084	3.73
SGO_1655	SGO_1655	3.73
SGO_0561	SGO_0561	3.71
SGO_0563	SGO_0563	3.70
SGO_0396	SGO_0396	3.67
SGO_1210	*fhs*	3.64
SGO_1715	*hppH*	3.62
SGO_0508	*nrdR*	3.59
SGO_1135	SGO_1135	3.58
SGO_0138	SGO_0138	3.57
SGO_1337	SGO_1337	3.49
SGO_2043	SGO_2043	3.45
SGO_0339	SGO_0339	3.45
SGO_2083	SGO_2083	3.43
SGO_0564	SGO_0564	3.41
SGO_1108	SGO_1108	3.39
SGO_0734	SGO_0734	3.38
SGO_0185	SGO_0185	3.33
SGO_0964	*pdxK*	3.32
SGO_0410	SGO_0410	3.32
SGO_0368	*merA*	3.32
SGO_0523	SGO_0523	3.31
SGO_1806	*ftcD*	3.30
SGO_1256	SGO_1256	3.26
SGO_0758	SGO_0758	3.25
SGO_0912	SGO_0912	3.24
SGO_0430	SGO_0430	3.21
SGO_0417	*acpS*	3.18
SGO_0095	*mccF*	3.17
SGO_1750	SGO_1750	3.11
SGO_0666	SGO_0666	3.11
SGO_0524	SGO_0524	3.11
SGO_2102	SGO_2102	3.10
SGO_1386	SGO_1386	3.09
SGO_1109	*pyrB*	3.03
SGO_0180	*jag*	3.03
SGO_0280	*trzA*	2.99
SGO_1620	SGO_1620	2.98
SGO_2012	SGO_2012	2.98
SGO_2040	SGO_2040	2.96
SGO_1625	SGO_1625	2.92

SGO_0684	SGO_0684	2.91
SGO_2042	SGO_2042	2.91
SGO_0303	SGO_0303	2.88
SGO_1136	SGO_1136	2.87
SGO_0053	SGO_0053	2.87
SGO_0181	SGO_0181	2.87
SGO_2041	SGO_2041	2.85
SGO_1065	SGO_1065	2.84
SGO_0865	SGO_0865	2.82
SGO_1856	SGO_1856	2.82
SGO_0872	SGO_0872	2.79
SGO_1500	SGO_1500	2.78
SGO_0726	SGO_0726	2.78
SGO_0414	SGO_0414	2.77
SGO_1498	SGO_1498	2.76
SGO_0840	SGO_0840	2.76
SGO_0667	SGO_0667	2.76
SGO_0405	SGO_0405	2.74
SGO_2103	SGO_2103	2.73
SGO_1728	*glnQ*	2.73
SGO_1597	SGO_1597	2.72
SGO_0965	SGO_0965	2.72
SGO_0842	SGO_0842	2.69
SGO_0042	SGO_0042	2.65
SGO_0873	SGO_0873	2.63
SGO_0297	*arb*	2.62
SGO_1791	SGO_1791	2.61
SGO_1384	SGO_1384	2.60
SGO_0182	*sapR*	2.59
SGO_1999	*ctsR*	2.59
SGO_1385	SGO_1385	2.56
SGO_0398	SGO_0398	2.55
SGO_1734	SGO_1734	2.55
SGO_0428	SGO_0428	2.53
SGO_1840	SGO_1840	2.53
SGO_0509	SGO_0509	2.52
SGO_0415	*secA*	2.51
SGO_0685	SGO_0685	2.49
SGO_1035	*gloA*	2.49
SGO_0841	SGO_0841	2.46
SGO_0596	*priA*	2.46
SGO_1501	SGO_1501	2.45
SGO_1611	*icd*	2.45
SGO_1122	SGO_1122	2.43
SGO_0079	SGO_0079	2.41
SGO_0870	SGO_0870	2.38

SGO_1820	SGO_1820	2.38
SGO_0117	SGO_0117	2.37
SGO_0228	SGO_0228	2.37
SGO_1940	SGO_1940	2.37
SGO_1839	SGO_1839	2.36
SGO_1527	SGO_1527	2.36
SGO_0935	SGO_0935	2.35
SGO_1815	SGO_1815	2.34
SGO_1838	SGO_1838	2.33
SGO_1595	SGO_1595	2.32
SGO_0186	SGO_0186	2.32
SGO_0948	*pgdA*	2.31
SGO_1483	SGO_1483	2.30
SGO_0418	*alr*	2.26
SGO_1736	SGO_1736	2.26
SGO_1727	SGO_1727	2.25
SGO_0990	SGO_0990	2.24
SGO_1998	*clpB*	2.23
SGO_1594	SGO_1594	2.21
SGO_1075	SGO_1075	2.19
SGO_2044	*spxA*	2.17
SGO_2051	SGO_2051	2.16
SGO_1215	*manB*	2.16
SGO_0553	SGO_0553	2.15
SGO_0943	SGO_0943	2.15
SGO_0958	SGO_0958	2.14
SGO_0500	*rggD*	2.14
SGO_0493	SGO_0493	2.12
SGO_2061	SGO_2061	2.10
SGO_0395	SGO_0395	2.10
SGO_0354	SGO_0354	2.09
SGO_1738	SGO_1738	2.08
SGO_0839	SGO_0839	2.08
SGO_0340	SGO_0340	2.06
SGO_0477	SGO_0477	2.01
SGO_0597	*fmt*	2.01
SGO_1434	*thiJ*	-2.01
SGO_0288	SGO_0288	-2.01
SGO_0447	*nadD*	-2.03
SGO_0995	SGO_0995	-2.05
SGO_1183	*brnQ*	-2.06
SGO_1683	*serS*	-2.07
SGO_0558	SGO_0558	-2.07
SGO_0697	SGO_0697	-2.11
SGO_1659	SGO_1659	-2.13
SGO_1472	SGO_1472	-2.13

SGO_0391	SGO_0391	-2.13
SGO_1361	SGO_1361	-2.18
SGO_0449	SGO_0449	-2.19
SGO_1375	*aroD*	-2.20
SGO_1642	SGO_1642	-2.20
SGO_0985	SGO_0985	-2.21
SGO_0171	*radA*	-2.21
SGO_1753	SGO_1753	-2.25
SGO_1559	*nrdF*	-2.26
SGO_1641	SGO_1641	-2.27
SGO_1056	*pstC*	-2.28
SGO_0858	SGO_0858	-2.28
SGO_2106	SGO_2106	-2.29
SGO_0490	SGO_0490	-2.31
SGO_0465	SGO_0465	-2.32
SGO_0560	*hsdM*	-2.32
SGO_0236	SGO_0236	-2.33
SGO_1024	SGO_1024	-2.34
SGO_1230	*srtA*	-2.35
SGO_1824	*prmA*	-2.36
SGO_1377	SGO_1377	-2.39
SGO_1737	*rpmB*	-2.40
SGO_0133	SGO_0133	-2.42
SGO_0438	SGO_0438	-2.43
SGO_1643	SGO_1643	-2.43
SGO_1912	SGO_1912	-2.44
SGO_0814	SGO_0814	-2.44
SGO_0173	SGO_0173	-2.44
SGO_1053	*app*	-2.46
SGO_1858	SGO_1858	-2.46
SGO_1057	*pstA*	-2.47
SGO_0515	*murC*	-2.47
SGO_1979	*rpsC*	-2.49
SGO_1880	*ssb*	-2.53
SGO_1959	*rpoA*	-2.53
SGO_1878	SGO_1878	-2.54
SGO_1234	*rpsA*	-2.55
SGO_1058	*pstB2*	-2.56
SGO_0970	*secY*	-2.56
SGO_1745	*fba*	-2.58
SGO_1584	SGO_1584	-2.59
SGO_1754	SGO_1754	-2.59
SGO_0489	SGO_0489	-2.60
SGO_1389	SGO_1389	-2.63
SGO_1390	*ligA*	-2.64
SGO_2036	SGO_2036	-2.65

SGO_0172	SGO_0172	-2.65
SGO_0763	*murA*	-2.66
SGO_0637	SGO_0637	-2.67
SGO_1435	*mreB*	-2.67
SGO_0765	*endA*	-2.68
SGO_0817	SGO_0817	-2.68
SGO_0235	SGO_0235	-2.72
SGO_1876	SGO_1876	-2.72
SGO_0816	SGO_0816	-2.75
SGO_1025	*rgp*	-2.76
SGO_0238	SGO_0238	-2.77
SGO_1630	SGO_1630	-2.77
SGO_0528	*ilvC*	-2.79
SGO_0450	SGO_0450	-2.81
SGO_1383	*rplS*	-2.83
SGO_0254	SGO_0254	-2.84
SGO_1007	SGO_1007	-2.87
SGO_1186	SGO_1186	-2.90
SGO_0643	SGO_0643	-2.90
SGO_0453	SGO_0453	-2.91
SGO_0014	*pgsA*	-2.92
SGO_1446	*murF*	-2.95
SGO_0527	*ilvH*	-2.99
SGO_0698	*recN*	-3.01
SGO_1074	SGO_1074	-3.01
SGO_1482	SGO_1482	-3.02
SGO_1672	*xerD*	-3.02
SGO_1685	SGO_1685	-3.08
SGO_0639	*valS*	-3.10
SGO_0434	*aspS*	-3.13
SGO_0570	SGO_0570	-3.15
SGO_0804	*murB*	-3.16
SGO_1593	*arcA*	-3.21
SGO_1187	SGO_1187	-3.23
SGO_1984	*rplD*	-3.26
SGO_0699	SGO_0699	-3.26
SGO_1536	SGO_1536	-3.34
SGO_1060	SGO_1060	-3.34
SGO_1936	*adcA*	-3.34
SGO_0526	*ilvB*	-3.37
SGO_1413	SGO_1413	-3.38
SGO_1444	SGO_1444	-3.39
SGO_0205	*rpsG*	-3.40
SGO_1455	*rplA*	-3.40
SGO_0242	SGO_0242	-3.43
SGO_1626	SGO_1626	-3.43

SGO_1059	*pstB1*	-3.43
SGO_0141	SGO_0141	-3.44
SGO_1443	SGO_1443	-3.52
SGO_0467	SGO_0467	-3.53
SGO_0243	SGO_0243	-3.58
SGO_1371	SGO_1371	-3.59
SGO_0136	SGO_0136	-3.62
SGO_1373	*aroB*	-3.65
SGO_1627	SGO_1627	-3.67
SGO_1629	*livH*	-3.68
SGO_0322	SGO_0322	-3.69
SGO_1465	SGO_1465	-3.69
SGO_1857	SGO_1857	-3.73
SGO_0575	*pbp2X*	-3.74
SGO_0204	*rpsL*	-3.74
SGO_0437	*gatB*	-3.74
SGO_1628	*braE*	-3.75
SGO_0529	*ilvA*	-3.77
SGO_0015	*cbiO*	-3.78
SGO_1591	*arcC*	-3.84
SGO_0452	SGO_0452	-3.86
SGO_1372	*aroC*	-3.89
SGO_1534	SGO_1534	-3.90
SGO_0792	SGO_0792	-3.90
SGO_1558	*nrdE*	-3.90
SGO_1506	SGO_1506	-3.93
SGO_1026	SGO_1026	-3.95
SGO_1206	*asd*	-4.01
SGO_0973	*asp3*	-4.03
SGO_2028	*wzg*	-4.03
SGO_0253	SGO_0253	-4.05
SGO_1937	*adcB*	-4.06
SGO_1456	*rplK*	-4.10
SGO_1671	*scpA*	-4.15
SGO_1748	*pyrG*	-4.18
SGO_1684	SGO_1684	-4.18
SGO_1314	SGO_1314	-4.22
SGO_1008	SGO_1008	-4.23
SGO_1082	SGO_1082	-4.25
SGO_1535	*ogt*	-4.27
SGO_1397	*map*	-4.27
SGO_1846	SGO_1846	-4.37
SGO_0451	SGO_0451	-4.39
SGO_0576	*mraY*	-4.41
SGO_0713	*era*	-4.41
SGO_1670	*scpB*	-4.42

SGO_0764	SGO_0764	-4.51
SGO_0432	*entB*	-4.52
SGO_1015	SGO_1015	-4.53
SGO_1669	SGO_1669	-4.57
SGO_0714	SGO_0714	-4.58
SGO_1388	*pulA*	-4.61
SGO_1396	SGO_1396	-4.62
SGO_0408	*zmpB*	-4.72
SGO_0431	SGO_0431	-4.79
SGO_0568	*glyQ*	-4.79
SGO_2027	*wzh*	-4.87
SGO_0580	SGO_0580	-4.92
SGO_0793	SGO_0793	-4.99
SGO_1118	SGO_1118	-5.12
SGO_0715	*mutM*	-5.20
SGO_0784	*smc*	-5.23
SGO_0972	*asp2*	-5.26
SGO_1847	*polC*	-5.28
SGO_0805	*potA*	-5.41
SGO_1374	*aroE*	-5.59
SGO_0567	SGO_0567	-5.64
SGO_0016	*cbiO*	-5.69
SGO_1244	SGO_1244	-5.79
SGO_0860	SGO_0860	-5.90
SGO_0569	*glyS*	-5.94
SGO_1258	SGO_1258	-5.99
SGO_1239	*parC*	-6.02
SGO_1391	SGO_1391	-6.04
SGO_1785	SGO_1785	-6.05
SGO_1016	SGO_1016	-6.10
SGO_0017	SGO_0017	-6.10
SGO_0787	*ftsY*	-6.21
SGO_1243	SGO_1243	-6.31
SGO_0859	*pheS*	-6.31
SGO_1696	SGO_1696	-6.41
SGO_1609	SGO_1609	-6.48
SGO_0786	SGO_0786	-6.49
SGO_0806	*potB*	-6.59
SGO_1242	SGO_1242	-6.79
SGO_0778	*thrS*	-6.84
SGO_0777	SGO_0777	-6.84
SGO_0323	SGO_0323	-6.93
SGO_0975	*gtaA*	-7.16
SGO_1245	*parE*	-7.42
SGO_0785	SGO_0785	-7.51
SGO_0717	SGO_0717	-7.52

SGO_1027	SGO_1027	-7.58
SGO_2025	*wze*	-7.63
SGO_1697	*acpP*	-7.69
SGO_1668	SGO_1668	-7.85
SGO_1017	SGO_1017	-8.06
SGO_1695	SGO_1695	-8.07
SGO_0716	*coaE*	-8.18
SGO_0966	*hsa*	-8.22
SGO_0978	SGO_0978	-8.31
SGO_2026	*wzd*	-8.39
SGO_2019	*licD*	-8.42
SGO_2024	SGO_2024	-8.48
SGO_2021	SGO_2021	-8.65
SGO_1694	*fabD*	-8.80
SGO_0861	*pheT*	-8.91
SGO_1240	SGO_1240	-8.98
SGO_0807	*potC*	-8.99
SGO_2020	SGO_2020	-8.99
SGO_0506	*rgfB*	-9.08
SGO_1241	SGO_1241	-9.11
SGO_1693	*fabG*	-9.30
SGO_2023	SGO_2023	-9.59
SGO_2018	SGO_2018	-9.75
SGO_0808	*potD*	-9.92
SGO_1691	*accB*	-10.16
SGO_2022	SGO_2022	-10.44
SGO_1690	*fabZ*	-10.47
SGO_0977	SGO_0977	-10.50
SGO_2016	SGO_2016	-10.69
SGO_0976	SGO_0976	-10.81
SGO_0681	*ileS*	-11.12
SGO_1692	SGO_1692	-11.32
SGO_1698	SGO_1698	-11.93
SGO_2015	SGO_2015	-12.31
SGO_1687	*accA*	-12.33
SGO_2098	*rpsD*	-12.54
SGO_1699	SGO_1699	-12.55
SGO_1688	*accD*	-12.75
SGO_2017	SGO_2017	-12.93
SGO_0832	SGO_0832	-13.26
SGO_0831	SGO_0831	-13.73
SGO_1689	*accC*	-15.18
SGO_1575	SGO_1575	-15.59
SGO_1686	SGO_1686	-17.15
SGO_1576	*celD*	-21.10
SGO_1577	SGO_1577	-23.63

SGO_1578	*celC*	-24.43
SGO_1579	SGO_1579	-24.98
SGO_1700	SGO_1700	-30.82
SGO_1582	*celA*	-35.01
SGO_1580	*celB*	-51.14
SGO_1581	SGO_1581	-81.92

S. gordonii Δ*arcR* microarray data, regulation of genes in response to arginine depletion

Locus tag	Gene name	FC (abs)
SGO_1569	*argC*	420.15
SGO_1567	*argB*	340.62
SGO_1568	*argJ*	325.64
SGO_1566	*argD*	315.85
SGO_0177	SGO_0177	77.38
SGO_1105	SGO_1105	56.81
SGO_1106	SGO_1106	53.12
SGO_1656	SGO_1656	49.98
SGO_0175	*argG*	45.37
SGO_0176	*argH*	39.66
SGO_1104	*carB*	35.02
SGO_1103	*carA*	28.59
SGO_1102	SGO_1102	24.22
SGO_1410	SGO_1410	20.81
SGO_1409	*hisG*	20.38
SGO_0648	SGO_0648	18.50
SGO_0091	SGO_0091	17.27
SGO_0647	SGO_0647	16.99
SGO_1407	*hisB*	16.95
SGO_1411	*hisC*	16.83
SGO_0645	SGO_0645	16.55
SGO_1408	*hisD*	16.18
SGO_0646	SGO_0646	14.16
SGO_0021	SGO_0021	12.62
SGO_1406	*hisH*	12.15
SGO_1405	*hisA*	11.58
SGO_1404	*hisF*	10.43
SGO_1590	*arcD*	9.20
SGO_1716	SGO_1716	9.06
SGO_1403	*hisIE*	8.83
SGO_1137	SGO_1137	8.23
SGO_1402	*hisE*	7.48
SGO_1832	SGO_1832	6.53
SGO_1317	SGO_1317	6.51
SGO_0427	SGO_0427	6.51
SGO_1401	SGO_1401	6.48
SGO_0093	SGO_0093	6.48
SGO_1833	SGO_1833	6.12
SGO_1279	SGO_1279	6.00
SGO_1194	SGO_1194	5.70
SGO_0179	SGO_0179	5.63
SGO_1805	*hutU*	5.55

SGO_1278	SGO_1278	5.42
SGO_1277	SGO_1277	5.28
SGO_1589	*arcT*	5.21
SGO_0092	SGO_0092	5.10
SGO_1256	SGO_1256	4.96
SGO_1564	SGO_1564	4.93
SGO_1107	SGO_1107	4.84
SGO_0496	*rgg*	4.68
SGO_1195	SGO_1195	4.68
SGO_0178	*rnpA*	4.63
SGO_0086	SGO_0086	4.60
SGO_0874	SGO_0874	4.56
SGO_0301	SGO_0301	4.51
SGO_1835	SGO_1835	4.42
SGO_0094	SGO_0094	4.40
SGO_1196	*satD*	4.37
SGO_0883	SGO_0883	4.33
SGO_1655	SGO_1655	4.30
SGO_1109	*pyrB*	4.27
SGO_1563	SGO_1563	4.21
SGO_1834	SGO_1834	4.09
SGO_0571	SGO_0571	4.08
SGO_1562	SGO_1562	4.01
SGO_0090	SGO_0090	3.94
SGO_1108	SGO_1108	3.93
SGO_1110	SGO_1110	3.71
SGO_1126	*xerS*	3.67
SGO_0089	SGO_0089	3.59
SGO_0157	SGO_0157	3.58
SGO_0687	SGO_0687	3.57
SGO_0508	*nrdR*	3.56
SGO_1485	SGO_1485	3.55
SGO_0180	*jag*	3.49
SGO_0302	SGO_0302	3.45
SGO_1561	SGO_1561	3.42
SGO_2084	SGO_2084	3.42
SGO_1618	SGO_1618	3.41
SGO_0185	SGO_0185	3.35
SGO_1597	SGO_1597	3.33
SGO_0682	SGO_0682	3.33
SGO_1135	SGO_1135	3.32
SGO_1138	SGO_1138	3.32
SGO_0368	*merA*	3.28
SGO_0572	SGO_0572	3.24
SGO_0417	*acpS*	3.23
SGO_1997	*nadR*	3.22

SGO_0280	*trzA*	3.19
SGO_1750	SGO_1750	3.19
SGO_1280	SGO_1280	3.18
SGO_0251	SGO_0251	3.16
SGO_1134	SGO_1134	3.16
SGO_2043	SGO_2043	3.15
SGO_0297	*arb*	3.15
SGO_1386	SGO_1386	3.14
SGO_0957	SGO_0957	3.14
SGO_1922	*comYC*	3.13
SGO_0334	SGO_0334	3.11
SGO_0845	SGO_0845	3.09
SGO_0665	SGO_0665	3.09
SGO_2012	SGO_2012	3.08
SGO_1902	*ssb*	3.05
SGO_1920	SGO_1920	3.04
SGO_1625	SGO_1625	3.03
SGO_1751	SGO_1751	3.02
SGO_1923	*comYB*	3.00
SGO_1337	SGO_1337	2.99
SGO_1065	SGO_1065	2.97
SGO_0709	SGO_0709	2.95
SGO_1560	SGO_1560	2.92
SGO_1729	SGO_1729	2.91
SGO_1877	SGO_1877	2.91
SGO_0181	SGO_0181	2.90
SGO_1830	*mga*	2.88
SGO_0758	SGO_0758	2.87
SGO_0964	*pdxK*	2.87
SGO_1255	*pyrd*	2.87
SGO_0667	SGO_0667	2.86
SGO_0666	SGO_0666	2.85
SGO_0602	SGO_0602	2.85
SGO_1385	SGO_1385	2.84
SGO_0046	SGO_0046	2.84
SGO_2083	SGO_2083	2.84
SGO_1384	SGO_1384	2.83
SGO_0138	SGO_0138	2.82
SGO_0411	SGO_0411	2.80
SGO_0561	SGO_0561	2.79
SGO_0095	*mccF*	2.79
SGO_0523	SGO_0523	2.77
SGO_1501	SGO_1501	2.77
SGO_0562	SGO_0562	2.77
SGO_2103	SGO_2103	2.76
SGO_0416	SGO_0416	2.75

SGO_0339	SGO_0339	2.74
SGO_0397	SGO_0397	2.74
SGO_1498	SGO_1498	2.73
SGO_0184	SGO_0184	2.71
SGO_0601	SGO_0601	2.70
SGO_0118	SGO_0118	2.70
SGO_1733	SGO_1733	2.70
SGO_0734	SGO_0734	2.70
SGO_2033	*nrdD*	2.69
SGO_0182	*sapR*	2.68
SGO_1856	SGO_1856	2.68
SGO_0430	SGO_0430	2.65
SGO_1529	SGO_1529	2.65
SGO_1521	SGO_1521	2.62
SGO_1752	SGO_1752	2.61
SGO_2031	SGO_2031	2.61
SGO_1728	*glnQ*	2.61
SGO_2041	SGO_2041	2.60
SGO_0186	SGO_0186	2.60
SGO_0418	*alr*	2.60
SGO_1525	*lacB*	2.60
SGO_1484	SGO_1484	2.60
SGO_0596	*priA*	2.58
SGO_0047	SGO_0047	2.56
SGO_1596	SGO_1596	2.56
SGO_1520	SGO_1520	2.55
SGO_0409	SGO_0409	2.55
SGO_1522	SGO_1522	2.54
SGO_0410	SGO_0410	2.52
SGO_0912	SGO_0912	2.52
SGO_1064	SGO_1064	2.52
SGO_1523	*lacD*	2.49
SGO_0563	SGO_0563	2.46
SGO_0842	SGO_0842	2.46
SGO_0727	SGO_0727	2.46
SGO_0740	SGO_0740	2.45
SGO_0270	*gltP*	2.45
SGO_1518	*lacB*	2.44
SGO_1527	SGO_1527	2.44
SGO_1500	SGO_1500	2.44
SGO_0158	SGO_0158	2.42
SGO_0688	SGO_0688	2.40
SGO_1610	SGO_1610	2.40
SGO_1524	*lacC*	2.39
SGO_0990	SGO_0990	2.38
SGO_0296	SGO_0296	2.38

SGO_2044	*spxA*	2.36
SGO_2040	SGO_2040	2.36
SGO_0748	SGO_0748	2.35
SGO_0726	SGO_0726	2.35
SGO_1517	*lacC*	2.31
SGO_1795	SGO_1795	2.30
SGO_0729	SGO_0729	2.29
SGO_0564	SGO_0564	2.29
SGO_1998	*clpB*	2.29
SGO_0865	SGO_0865	2.28
SGO_0415	*secA*	2.27
SGO_0398	SGO_0398	2.27
SGO_1804	*hutI*	2.25
SGO_0078	SGO_0078	2.25
SGO_1526	*lacA*	2.24
SGO_0394	SGO_0394	2.23
SGO_1210	*fhs*	2.23
SGO_1281	SGO_1281	2.23
SGO_1136	SGO_1136	2.23
SGO_2057	*argR*	2.22
SGO_0367	*cadX*	2.21
SGO_0396	SGO_0396	2.21
SGO_2058	*argS*	2.21
SGO_0684	SGO_0684	2.19
SGO_0043	SGO_0043	2.19
SGO_1519	*lacA*	2.19
SGO_0268	SGO_0268	2.19
SGO_0048	SGO_0048	2.18
SGO_0211	*sspB*	2.17
SGO_0685	SGO_0685	2.17
SGO_2011	SGO_2011	2.17
SGO_2051	SGO_2051	2.17
SGO_0420	SGO_0420	2.17
SGO_0044	SGO_0044	2.17
SGO_1070	*bioY*	2.16
SGO_2042	SGO_2042	2.16
SGO_1394	SGO_1394	2.16
SGO_0395	SGO_0395	2.13
SGO_1786	*gldA*	2.12
SGO_0500	*rggD*	2.11
SGO_0478	SGO_0478	2.11
SGO_1996	SGO_1996	2.09
SGO_1791	SGO_1791	2.09
SGO_0839	SGO_0839	2.09
SGO_0943	SGO_0943	2.08
SGO_0119	SGO_0119	2.08

SGO_0419	*recG*	2.08
SGO_1066	SGO_1066	2.07
SGO_1653	SGO_1653	2.07
SGO_0870	SGO_0870	2.06
SGO_0608	*comFA*	2.06
SGO_2061	SGO_2061	2.06
SGO_0603	SGO_0603	2.06
SGO_0840	SGO_0840	2.05
SGO_0399	SGO_0399	2.05
SGO_0872	SGO_0872	2.05
SGO_0525	SGO_0525	2.04
SGO_0873	SGO_0873	2.03
SGO_0405	SGO_0405	2.03
SGO_1215	*manB*	2.03
SGO_1123	*ffh*	2.03
SGO_0731	SGO_0731	2.03
SGO_0589	SGO_0589	2.02
SGO_1855	*yajC*	2.02
SGO_1061	SGO_1061	2.01
SGO_0728	SGO_0728	2.01
SGO_0676	SGO_0676	2.01
SGO_1738	SGO_1738	2.01
SGO_1058	*pstB2*	-2.00
SGO_1434	*thiJ*	-2.01
SGO_1055	SGO_1055	-2.02
SGO_0722	*tehB*	-2.02
SGO_1310	SGO_1310	-2.03
SGO_1077	*coaA*	-2.03
SGO_1760	SGO_1760	-2.05
SGO_1960	*rpsK*	-2.06
SGO_1544	*atpA*	-2.07
SGO_1238	*ilvE*	-2.09
SGO_0240	*mvaD*	-2.10
SGO_0581	*trxB*	-2.11
SGO_1005	SGO_1005	-2.12
SGO_0817	SGO_0817	-2.12
SGO_0235	SGO_0235	-2.12
SGO_0858	SGO_0858	-2.12
SGO_0642	SGO_0642	-2.12
SGO_0544	SGO_0544	-2.12
SGO_0136	SGO_0136	-2.12
SGO_1353	SGO_1353	-2.12
SGO_2063	SGO_2063	-2.13
SGO_1470	SGO_1470	-2.13
SGO_1400	*murA*	-2.13
SGO_1547	*atpB*	-2.16

SGO_1961	*rpsM*	-2.16
SGO_0814	SGO_0814	-2.16
SGO_1745	*fba*	-2.17
SGO_1364	*rumA*	-2.17
SGO_0517	SGO_0517	-2.17
SGO_1546	*atpF*	-2.18
SGO_2070	SGO_2070	-2.18
SGO_0010	SGO_0010	-2.18
SGO_1967	*rpmD*	-2.19
SGO_1939	*adcR*	-2.19
SGO_0013	SGO_0013	-2.20
SGO_0305	SGO_0305	-2.20
SGO_2036	SGO_2036	-2.20
SGO_0277	*pyrA*	-2.21
SGO_0391	SGO_0391	-2.21
SGO_0065	SGO_0065	-2.21
SGO_1824	*prmA*	-2.22
SGO_1383	*rplS*	-2.23
SGO_0701	*hup*	-2.24
SGO_1370	SGO_1370	-2.25
SGO_1024	SGO_1024	-2.25
SGO_1054	SGO_1054	-2.26
SGO_0815	*thiI*	-2.27
SGO_0969	*nss*	-2.27
SGO_1683	*serS*	-2.28
SGO_0826	SGO_0826	-2.28
SGO_0515	*murC*	-2.29
SGO_0744	SGO_0744	-2.29
SGO_0173	SGO_0173	-2.30
SGO_0643	SGO_0643	-2.30
SGO_0132	SGO_0132	-2.30
SGO_1971	*rpsH*	-2.31
SGO_0070	SGO_0070	-2.31
SGO_1972	*rpsN*	-2.31
SGO_1959	*rpoA*	-2.32
SGO_1548	*atpE*	-2.33
SGO_1375	*aroD*	-2.33
SGO_0566	*sgc*	-2.34
SGO_1435	*mreB*	-2.35
SGO_1025	*rgp*	-2.36
SGO_0066	SGO_0066	-2.37
SGO_1993	SGO_1993	-2.37
SGO_1659	SGO_1659	-2.38
SGO_1973	*rplE*	-2.40
SGO_0970	*secY*	-2.40
SGO_1991	*hslO*	-2.40

SGO_0012	*mpp*	-2.40
SGO_0244	SGO_0244	-2.41
SGO_1845	SGO_1845	-2.41
SGO_1970	*rplF*	-2.46
SGO_0527	*ilvH*	-2.46
SGO_1390	*ligA*	-2.48
SGO_0637	SGO_0637	-2.48
SGO_0438	SGO_0438	-2.49
SGO_1538	SGO_1538	-2.49
SGO_0067	SGO_0067	-2.50
SGO_1361	SGO_1361	-2.50
SGO_1965	*secY*	-2.50
SGO_0134	SGO_0134	-2.51
SGO_1117	*pcnA*	-2.51
SGO_1543	*atpG*	-2.52
SGO_0689	SGO_0689	-2.52
SGO_1968	*rpsE*	-2.52
SGO_1639	SGO_1639	-2.52
SGO_1977	*rpmC*	-2.52
SGO_1204	SGO_1204	-2.53
SGO_0763	*murA*	-2.53
SGO_0135	SGO_0135	-2.53
SGO_1754	SGO_1754	-2.54
SGO_1053	*app*	-2.56
SGO_1969	*rplR*	-2.57
SGO_0560	*hsdM*	-2.58
SGO_1482	SGO_1482	-2.58
SGO_1938	*adcC*	-2.59
SGO_1059	*pstB1*	-2.59
SGO_0064	SGO_0064	-2.61
SGO_0241	SGO_0241	-2.62
SGO_1978	*rplP*	-2.64
SGO_0698	*recN*	-2.65
SGO_1963	*infA*	-2.66
SGO_2106	SGO_2106	-2.66
SGO_1352	SGO_1352	-2.67
SGO_0452	SGO_0452	-2.68
SGO_0429	SGO_0429	-2.68
SGO_0216	SGO_0216	-2.69
SGO_0520	SGO_0520	-2.70
SGO_1283	SGO_1283	-2.70
SGO_1903	SGO_1903	-2.71
SGO_0328	SGO_0328	-2.72
SGO_0639	*valS*	-2.72
SGO_1443	SGO_1443	-2.72
SGO_1260	*deoD*	-2.75

SGO_1186	SGO_1186	-2.75
SGO_1060	SGO_1060	-2.75
SGO_0068	SGO_0068	-2.77
SGO_1230	*srtA*	-2.77
SGO_0144	SGO_0144	-2.79
SGO_0942	*zmpC*	-2.80
SGO_0699	SGO_0699	-2.80
SGO_0069	SGO_0069	-2.83
SGO_1962	*rpmJ*	-2.83
SGO_1073	SGO_1073	-2.83
SGO_1823	SGO_1823	-2.84
SGO_1880	*ssb*	-2.85
SGO_1413	SGO_1413	-2.86
SGO_1143	*thyA*	-2.87
SGO_1322	SGO_1322	-2.88
SGO_0961	SGO_0961	-2.89
SGO_1672	*xerD*	-2.90
SGO_1205	*dapA*	-2.90
SGO_1397	*map*	-2.90
SGO_1313	SGO_1313	-2.92
SGO_1937	*adcB*	-2.92
SGO_1979	*rpsC*	-2.95
SGO_1082	SGO_1082	-2.95
SGO_1371	SGO_1371	-2.95
SGO_1377	SGO_1377	-2.97
SGO_1323	*rpsP*	-2.98
SGO_0962	SGO_0962	-2.98
SGO_1412	SGO_1412	-3.00
SGO_0437	*gatB*	-3.01
SGO_1626	SGO_1626	-3.02
SGO_1373	*aroB*	-3.04
SGO_0205	*rpsG*	-3.04
SGO_0014	*pgsA*	-3.10
SGO_0528	*ilvC*	-3.10
SGO_1936	*adcA*	-3.11
SGO_1389	SGO_1389	-3.13
SGO_0526	*ilvB*	-3.15
SGO_1630	SGO_1630	-3.15
SGO_0238	SGO_0238	-3.15
SGO_0243	SGO_0243	-3.17
SGO_1314	SGO_1314	-3.19
SGO_0242	SGO_0242	-3.19
SGO_0434	*aspS*	-3.20
SGO_1007	SGO_1007	-3.21
SGO_0570	SGO_0570	-3.24
SGO_1537	SGO_1537	-3.27

SGO_1536	SGO_1536	-3.30
SGO_1118	SGO_1118	-3.31
SGO_1559	*nrdF*	-3.32
SGO_1465	SGO_1465	-3.32
SGO_0466	SGO_0466	-3.32
SGO_1444	SGO_1444	-3.34
SGO_1640	SGO_1640	-3.36
SGO_0467	SGO_0467	-3.38
SGO_0436	*gatA*	-3.38
SGO_1627	SGO_1627	-3.39
SGO_1879	*rpsR*	-3.40
SGO_0805	*potA*	-3.42
SGO_0254	SGO_0254	-3.43
SGO_2048	SGO_2048	-3.44
SGO_0575	*pbp2X*	-3.44
SGO_0576	*mraY*	-3.45
SGO_0453	SGO_0453	-3.47
SGO_1557	SGO_1557	-3.47
SGO_1878	SGO_1878	-3.51
SGO_0529	*ilvA*	-3.51
SGO_0204	*rpsL*	-3.52
SGO_1542	*atpD*	-3.52
SGO_1372	*aroC*	-3.55
SGO_0253	SGO_0253	-3.57
SGO_1917	SGO_1917	-3.59
SGO_0359	*rpsI*	-3.59
SGO_1534	SGO_1534	-3.62
SGO_0764	SGO_0764	-3.63
SGO_0232	SGO_0232	-3.63
SGO_0765	*endA*	-3.65
SGO_0860	SGO_0860	-3.66
SGO_1446	*murF*	-3.70
SGO_1455	*rplA*	-3.72
SGO_1008	SGO_1008	-3.74
SGO_1617	*prfC*	-3.74
SGO_1737	*rpmB*	-3.76
SGO_2135	*tilS*	-3.80
SGO_0275	SGO_0275	-3.83
SGO_1140	*clpX*	-3.85
SGO_1206	*asd*	-3.89
SGO_1026	SGO_1026	-3.90
SGO_1615	SGO_1615	-3.90
SGO_1628	*braE*	-3.93
SGO_1748	*pyrG*	-3.95
SGO_1187	SGO_1187	-3.96
SGO_1881	*rpsF*	-3.98

SGO_1456	*rplK*	-3.99
SGO_1847	*polC*	-4.00
SGO_1506	SGO_1506	-4.03
SGO_1139	*engB*	-4.04
SGO_0793	SGO_0793	-4.04
SGO_0714	SGO_0714	-4.06
SGO_1374	*aroE*	-4.07
SGO_0859	*pheS*	-4.08
SGO_1074	SGO_1074	-4.08
SGO_0140	SGO_0140	-4.12
SGO_1846	SGO_1846	-4.13
SGO_1696	SGO_1696	-4.16
SGO_0972	*asp2*	-4.17
SGO_1616	SGO_1616	-4.17
SGO_1229	*radC*	-4.18
SGO_0792	SGO_0792	-4.19
SGO_0806	*potB*	-4.23
SGO_0015	*cbiO*	-4.24
SGO_1535	*ogt*	-4.27
SGO_0506	*rgfB*	-4.29
SGO_0568	*glyQ*	-4.34
SGO_1015	SGO_1015	-4.34
SGO_1558	*nrdE*	-4.35
SGO_0966	*hsa*	-4.40
SGO_1396	SGO_1396	-4.42
SGO_0276	*gdhA*	-4.42
SGO_0973	*asp3*	-4.48
SGO_1614	SGO_1614	-4.50
SGO_1388	*pulA*	-4.53
SGO_0715	*mutM*	-4.54
SGO_0432	*entB*	-4.64
SGO_0497	*gtfG*	-4.67
SGO_0580	SGO_0580	-4.87
SGO_1629	*livH*	-4.89
SGO_0713	*era*	-4.90
SGO_0807	*potC*	-4.90
SGO_1984	*rplD*	-4.91
SGO_0786	SGO_0786	-4.94
SGO_0431	SGO_0431	-4.98
SGO_1592	*arcB*	-5.00
SGO_0322	SGO_0322	-5.02
SGO_1591	*arcC*	-5.05
SGO_1694	*fabD*	-5.05
SGO_0567	SGO_0567	-5.06
SGO_0861	*pheT*	-5.07
SGO_2028	*wzg*	-5.09

SGO_1593	*arcA*	-5.12
SGO_1692	SGO_1692	-5.12
SGO_1691	*accB*	-5.16
SGO_1259	SGO_1259	-5.16
SGO_0141	SGO_0141	-5.18
SGO_1245	*parE*	-5.27
SGO_0784	*smc*	-5.29
SGO_2027	*wzh*	-5.35
SGO_0717	SGO_0717	-5.40
SGO_1695	SGO_1695	-5.65
SGO_1690	*fabZ*	-5.68
SGO_1671	*scpA*	-5.69
SGO_1258	SGO_1258	-5.73
SGO_0778	*thrS*	-5.75
SGO_0787	*ftsY*	-5.79
SGO_0808	*potD*	-5.87
SGO_1670	*scpB*	-5.89
SGO_1693	*fabG*	-5.95
SGO_0716	*coaE*	-5.96
SGO_0408	*zmpB*	-6.16
SGO_1669	SGO_1669	-6.17
SGO_1577	SGO_1577	-6.17
SGO_0785	SGO_0785	-6.24
SGO_1391	SGO_1391	-6.31
SGO_0777	SGO_0777	-6.37
SGO_1192	*rplJ*	-6.63
SGO_1244	SGO_1244	-6.65
SGO_0681	*ileS*	-6.66
SGO_1697	*acpP*	-6.69
SGO_1698	SGO_1698	-6.71
SGO_1668	SGO_1668	-7.26
SGO_1243	SGO_1243	-7.26
SGO_0017	SGO_0017	-7.28
SGO_1017	SGO_1017	-7.43
SGO_1242	SGO_1242	-7.47
SGO_1027	SGO_1027	-7.50
SGO_1688	*accD*	-7.51
SGO_1240	SGO_1240	-7.56
SGO_0016	*cbiO*	-7.57
SGO_2024	SGO_2024	-7.60
SGO_2025	*wze*	-7.66
SGO_2026	*wzd*	-7.85
SGO_0323	SGO_0323	-7.92
SGO_1016	SGO_1016	-8.04
SGO_0975	*gtaA*	-8.23
SGO_0976	SGO_0976	-8.31

SGO_1191	*rplL*	-8.49
SGO_2022	SGO_2022	-8.55
SGO_1785	SGO_1785	-8.62
SGO_1687	*accA*	-8.62
SGO_1241	SGO_1241	-8.66
SGO_0977	SGO_0977	-8.88
SGO_1578	*celC*	-8.96
SGO_1579	SGO_1579	-8.98
SGO_2021	SGO_2021	-9.01
SGO_0978	SGO_0978	-9.07
SGO_1699	SGO_1699	-9.27
SGO_1689	*accC*	-9.29
SGO_2020	SGO_2020	-9.78
SGO_2023	SGO_2023	-9.88
SGO_2098	*rpsD*	-9.99
SGO_0831	SGO_0831	-10.53
SGO_2019	*licD*	-11.38
SGO_2018	SGO_2018	-12.03
SGO_1239	*parC*	-13.22
SGO_1582	*celA*	-13.41
SGO_2016	SGO_2016	-14.56
SGO_0832	SGO_0832	-14.75
SGO_2017	SGO_2017	-14.85
SGO_1686	SGO_1686	-17.40
SGO_1700	SGO_1700	-20.71
SGO_1580	*celB*	-22.44
SGO_2015	SGO_2015	-23.94
SGO_1581	SGO_1581	-33.28

S. gordonii Δ*argR* microarray data, regulation of genes in response to arginine depletion

(Note: locus tags have been changed)

Locus tag	Gene name	FC (abs)
SGO_0177	SGO_0177	78.53
SGO_RS00865	SGO_0175	63.27
SGO_RS00870	SGO_0176	53.09
SGO_RS03190	SGO_0648	9.75
SGO_RS03180	SGO_0646	8.01
SGO_RS03185	SGO_0647	7.77
SGO_RS03175	SGO_0645	7.40
SGO_RS06270	SGO_1277	6.10
SGO_RS09770	SGO_1997	4.69
SGO_RS04295	SGO_0874	4.51
SGO_RS06910	SGO_1409	4.18
SGO_RS05590	SGO_1137	4.12
SGO_RS00105	SGO_0021	4.10
SGO_RS05865	SGO_1194	3.86
SGO_RS06915	SGO_1410	3.84
SGO_RS02455	SGO_0496	3.68
SGO_RS03075	SGO_0624	3.59
xerS	SGO_1126	3.57
SGO_RS06905	SGO_1408	3.56
SGO_RS05870	SGO_1195	3.50
rnpA	SGO_0178	3.44
SGO_RS02135	SGO_0430	3.43
hisB	SGO_1407	3.40
SGO_RS00880	SGO_0179	3.32
SGO_RS00530	SGO_0106	3.30
SGO_RS08585	SGO_1751	3.27
SGO_RS06280	SGO_1280	3.27
SGO_RS04340	SGO_0883	3.26
SGO_RS03275	SGO_0665	3.26
SGO_RS02965	SGO_0602	3.25
SGO_RS01830	SGO_0368	3.25
SGO_RS05875	SGO_1196	3.23
SGO_RS08115	SGO_1656	3.22
SGO_RS06460	SGO_1317	3.19
SGO_RS02040	SGO_0409	3.16
SGO_RS02125	SGO_0427	3.16
SGO_RS00525	SGO_0105	3.13
SGO_RS00895	SGO_0182	3.07
SGO_RS01670	SGO_0339	3.06
SGO_RS02045	SGO_0410	3.06
SGO_RS05430	SGO_1105	3.06

SGO_RS04720	SGO_0964	3.05
SGO_RS09845	SGO_2012	3.01
SGO_RS08580	SGO_1750	3.00
SGO_RS02810	SGO_0571	3.00
SGO_RS00775	SGO_0157	2.94
SGO_RS02960	SGO_0601	2.92
SGO_RS06920	SGO_1411	2.89
hisH	SGO_1406	2.86
SGO_RS07810	SGO_1594	2.86
SGO_RS10205	SGO_2084	2.81
SGO_RS02775	SGO_0562	2.79
SGO_RS07660	SGO_1564	2.78
SGO_RS02780	SGO_0564	2.78
SGO_RS02050	SGO_0411	2.78
SGO_RS01645	SGO_0334	2.78
SGO_RS04250	SGO_0865	2.76
SGO_RS00890	SGO_0181	2.75
SGO_RS01980	SGO_0397	2.75
SGO_RS03390	SGO_0688	2.74
SGO_RS05580	SGO_1135	2.74
SGO_RS02585	SGO_0523	2.74
SGO_RS05450	SGO_1110	2.73
SGO_RS03360	SGO_0682	2.72
SGO_RS07655	SGO_1562	2.71
SGO_RS04640	SGO_0948	2.70
SGO_RS00885	SGO_0180	2.69
SGO_RS07890	SGO_1610	2.68
SGO_RS07350	SGO_1498	2.67
SGO_RS07650	SGO_1561	2.65
SGO_RS02970	SGO_0603	2.65
SGO_RS02815	SGO_0572	2.65
SGO_RS08590	SGO_1752	2.64
SGO_RS00450	SGO_0090	2.63
SGO_RS07685	SGO_1569	2.62
SGO_RS07825	SGO_1597	2.62
SGO_RS05255	SGO_1070	2.60
SGO_RS04685	SGO_0957	2.58
SGO_RS02770	SGO_0561	2.57
SGO_RS07815	SGO_1595	2.54
SGO_RS02080	SGO_0417	2.54
SGO_RS08110	SGO_1655	2.54
SGO_RS02850	SGO_0579	2.52
SGO_RS09775	SGO_1998	2.50
SGO_RS10200	SGO_2083	2.49
SGO_RS00475	SGO_0095	2.46
SGO_RS02510	SGO_0508	2.46

SGO_RS07285	SGO_1484	2.46
SGO_RS01370	SGO_0278	2.45
SGO_RS04725	SGO_0965	2.42
SGO_RS07355	SGO_1499	2.42
SGO_RS09970	SGO_2041	2.41
SGO_RS07280	SGO_1483	2.40
SGO_RS05575	SGO_1134	2.40
SGO_RS01320	SGO_0268	2.40
SGO_RS07645	SGO_1560	2.39
SGO_RS06285	SGO_1281	2.39
SGO_RS04145	SGO_0845	2.38
SGO_RS02590	SGO_0524	2.35
SGO_RS01975	SGO_0396	2.34
SGO_RS05595	SGO_1138	2.33
SGO_RS02085	SGO_0418	2.33
SGO_RS07820	SGO_1596	2.32
SGO_RS01380	SGO_0280	2.31
SGO_RS03580	SGO_0729	2.31
SGO_RS10075	SGO_2062	2.31
SGO_RS03280	SGO_0666	2.31
SGO_RS02935	SGO_0596	2.30
SGO_RS03490	SGO_0709	2.30
SGO_RS02595	SGO_0525	2.29
SGO_RS08150	SGO_1662	2.29
SGO_RS03605	SGO_0734	2.28
SGO_RS08985	SGO_1831	2.28
SGO_RS04825	SGO_0984	2.28
SGO_RS02075	SGO_0416	2.28
SGO_RS01985	SGO_0398	2.27
SGO_RS06890	SGO_1405	2.26
SGO_RS10285	SGO_2102	2.25
SGO_RS07675	SGO_1567	2.25
SGO_RS01990	SGO_0399	2.24
SGO_RS10280	SGO_2101	2.24
SGO_RS01815	SGO_0366	2.24
SGO_RS06560	SGO_1337	2.24
SGO_RS02515	SGO_0509	2.24
SGO_RS05585	SGO_1136	2.23
SGO_RS09975	SGO_2042	2.23
SGO_RS02845	SGO_0578	2.23
SGO_RS08800	SGO_1794	2.22
SGO_RS05810	SGO_1181	2.21
SGO_RS04495	SGO_0917	2.21
SGO_RS09965	SGO_2040	2.21
SGO_RS06795	SGO_1386	2.20
SGO_RS01820	SGO_0367	2.20

SGO_RS06790	SGO_1385	2.19
SGO_RS05940	SGO_1210	2.19
SGO_RS01970	SGO_0395	2.18
SGO_RS06785	SGO_1384	2.17
SGO_RS01435	SGO_0291	2.16
argD	SGO_1566	2.16
SGO_RS09465	SGO_1933	2.15
SGO_RS09765	SGO_1996	2.15
SGO_RS07680	SGO_1568	2.13
SGO_RS02940	SGO_0597	2.12
SGO_RS06885	SGO_1404	2.12
SGO_RS01325	SGO_0269	2.12
SGO_RS10290	SGO_2103	2.11
SGO_RS01425	SGO_0289	2.11
SGO_RS10530	SGO_2151	2.11
SGO_RS09470	SGO_1934	2.11
SGO_RS01430	SGO_0290	2.10
SGO_RS01330	SGO_0270	2.08
SGO_RS02730	SGO_0553	2.07
SGO_RS02525	SGO_0511	2.07
SGO_RS05805	SGO_1180	2.07
SGO_RS02520	SGO_0510	2.05
SGO_RS07490	SGO_1529	2.04
SGO_RS03285	SGO_0667	2.03
SGO_RS00250	SGO_0052	2.03
SGO_RS07390	SGO_1507	2.03
SGO_RS09005	SGO_1835	2.03
SGO_RS01220	SGO_0247	2.02
SGO_RS04130	SGO_0842	2.02
SGO_RS09780	SGO_1999	2.00
SGO_RS05455	SGO_1111	2.00
SGO_RS07385	SGO_1506	-2.01
SGO_RS07580	SGO_1547	-2.01
SGO_RS07640	SGO_1559	-2.01
SGO_RS06040	SGO_1230	-2.01
SGO_RS05200	SGO_1059	-2.01
SGO_RS04765	SGO_0972	-2.02
SGO_RS05715	SGO_1161	-2.02
SGO_RS07695	SGO_1571	-2.02
SGO_RS05325	SGO_1084	-2.04
SGO_RS08555	SGO_1745	-2.04
SGO_RS09195	SGO_1878	-2.05
SGO_RS07555	SGO_1542	-2.05
SGO_RS09095	SGO_1857	-2.06
SGO_RS01160	SGO_0235	-2.07
SGO_RS02600	SGO_0526	-2.08

SGO_RS01195	SGO_0242	-2.09
SGO_RS08185	SGO_1669	-2.09
ligA	SGO_1390	-2.10
SGO_RS09930	SGO_2030	-2.11
SGO_RS06735	SGO_1374	-2.11
rpsA	SGO_1234	-2.12
SGO_RS06045	SGO_1231	-2.12
SGO_RS01770	SGO_0358	-2.12
SGO_RS07575	SGO_1546	-2.12
scpB	SGO_1670	-2.13
SGO_RS09210	SGO_1881	-2.13
SGO_RS04930	SGO_1006	-2.14
SGO_RS01005	SGO_0205	-2.15
SGO_RS06930	SGO_1413	-2.16
SGO_RS04935	SGO_1007	-2.16
SGO_RS07530	SGO_1537	-2.16
SGO_RS00295	SGO_0061	-2.17
SGO_RS00300	SGO_0062	-2.17
SGO_RS09725	SGO_1986	-2.18
SGO_RS09385	SGO_1917	-2.18
SGO_RS03845	SGO_0783	-2.18
SGO_RS01205	SGO_0244	-2.20
SGO_RS09495	SGO_1939	-2.20
SGO_RS02140	SGO_0431	-2.22
SGO_RS00305	SGO_0063	-2.22
SGO_RS07980	SGO_1629	-2.23
SGO_RS00315	SGO_0065	-2.24
SGO_RS07095	SGO_1446	-2.26
SGO_RS09320	SGO_1903	-2.26
SGO_RS04770	SGO_0973	-2.27
SGO_RS00330	SGO_0068	-2.28
SGO_RS02155	SGO_0435	-2.28
SGO_RS02130	SGO_0429	-2.29
SGO_RS07190	SGO_1465	-2.29
SGO_RS01585	SGO_0322	-2.30
SGO_RS05275	SGO_1074	-2.30
SGO_RS09940	SGO_2033	-2.32
SGO_RS05920	SGO_1206	-2.32
SGO_RS04980	SGO_1015	-2.36
SGO_RS07550	SGO_1541	-2.36
SGO_RS01000	SGO_0204	-2.36
SGO_RS02165	SGO_0437	-2.37
SGO_RS01145	SGO_0232	-2.37
SGO_RS03530	SGO_0717	-2.38
SGO_RS01250	SGO_0253	-2.40
SGO_RS01775	SGO_0359	-2.41

SGO_RS03885	SGO_0792	-2.42
SGO_RS02615	SGO_0529	-2.42
SGO_RS05605	SGO_1140	-2.42
SGO_RS01200	SGO_0243	-2.43
SGO_RS07885	SGO_1609	-2.44
SGO_RS00085	SGO_0017	-2.44
SGO_RS02805	SGO_0570	-2.45
SGO_RS05490	SGO_1118	-2.45
SGO_RS02610	SGO_0528	-2.46
SGO_RS01355	SGO_0275	-2.46
SGO_RS06925	SGO_1412	-2.46
polC	SGO_1847	-2.49
SGO_RS03520	SGO_0715	-2.50
SGO_RS09485	SGO_1937	-2.52
SGO_RS03955	SGO_0806	-2.52
gatA	SGO_0436	-2.52
SGO_RS03890	SGO_0793	-2.56
SGO_RS06185	SGO_1259	-2.57
SGO_RS07525	SGO_1536	-2.58
SGO_RS07145	SGO_1456	-2.61
SGO_RS05035	SGO_1026	-2.63
SGO_RS02830	SGO_0575	-2.63
SGO_RS09925	SGO_2029	-2.64
SGO_RS08180	SGO_1668	-2.66
SGO_RS01360	SGO_0276	-2.66
SGO_RS09200	SGO_1879	-2.66
pheS	SGO_0859	-2.67
SGO_RS07515	SGO_1534	-2.68
engB	SGO_1139, *yihA, ysxC*	-2.68
SGO_RS03950	SGO_0805	-2.70
SGO_RS06805	SGO_1388	-2.72
SGO_RS02835	SGO_0576	-2.73
SGO_RS06295	SGO_1283	-2.74
SGO_RS09480	SGO_1936	-2.74
SGO_RS10095	SGO_2066	-2.76
era	SGO_0713, *bex, rbaA, sdgE, yqfH*	-2.79
SGO_RS09920	SGO_2028	-2.79
SGO_RS05270	SGO_1073	-2.80
SGO_RS03525	SGO_0716	-2.80
SGO_RS09615	SGO_1964	-2.83
SGO_RS07635	SGO_1558	-2.84
SGO_RS01590	SGO_0323	-2.84
SGO_RS08325	SGO_1698	-2.85
SGO_RS04220	SGO_0860	-2.86
SGO_RS10090	SGO_2065	-2.94
SGO_RS00710	SGO_0144	-2.98

SGO_RS07140	SGO_1455	-2.98
SGO_RS04225	SGO_0861	-3.02
SGO_RS02800	SGO_0569	-3.07
SGO_RS04940	SGO_1008	-3.08
SGO_RS07520	SGO_1535	-3.09
SGO_RS04985	SGO_1016	-3.17
SGO_RS04780	SGO_0975	-3.19
SGO_RS04795	SGO_0978	-3.21
SGO_RS08755	SGO_1785	-3.27
SGO_RS08295	SGO_1691	-3.28
SGO_RS08290	SGO_1690	-3.30
SGO_RS03960	SGO_0807	-3.32
SGO_RS08300	SGO_1692	-3.45
SGO_RS04990	SGO_1017	-3.47
SGO_RS08310	SGO_1694	-3.54
SGO_RS08320	SGO_1697	-3.54
SGO_RS08330	SGO_1699	-3.54
gyrB	SGO_1245	-3.57
SGO_RS03855	SGO_0785	-3.57
SGO_RS07630	SGO_1557	-3.58
SGO_RS04785	SGO_0976	-3.62
glyQ	SGO_0568	-3.63
SGO_RS06820	SGO_1391	-3.70
SGO_RS03815	SGO_0777	-3.75
SGO_RS08305	SGO_1693	-3.75
SGO_RS03860	SGO_0786	-3.76
SGO_RS04790	SGO_0977	-3.77
SGO_RS03850	SGO_0784	-3.77
SGO_RS09915	SGO_2027	-3.80
SGO_RS03965	SGO_0808	-3.96
SGO_RS05315	SGO_1082	-3.99
SGO_RS06110	SGO_1244	-4.02
SGO_RS08315	SGO_1695	-4.03
SGO_RS08315	SGO_1695	-4.20
SGO_RS03865	SGO_0787	-4.23
SGO_RS06180	SGO_1258	-4.25
SGO_RS06090	SGO_1240	-4.30
SGO_RS03820	SGO_0778	-4.34
SGO_RS08285	SGO_1689	-4.42
SGO_RS09900	SGO_2024	-4.43
SGO_RS06095	SGO_1241	-4.46
SGO_RS08280	SGO_1688	-4.51
SGO_RS08275	SGO_1687	-4.54
SGO_RS09890	SGO_2022	-4.55
SGO_RS02035	SGO_0408	-4.70
SGO_RS10265	SGO_2098	-4.72

SGO_RS06100	SGO_1242	-4.77
SGO_RS09875	SGO_2019	-4.77
SGO_RS09880	SGO_2020	-4.77
SGO_RS07800	SGO_1592	-4.86
SGO_RS09870	SGO_2018	-4.87
SGO_RS05040	SGO_1027	-4.88
SGO_RS09905	SGO_2025	-4.92
SGO_RS06105	SGO_RS06105	-4.98
ileS	SGO_0681	-5.04
SGO_RS09910	SGO_2026	-5.11
SGO_RS09895	SGO_2023	-5.13
SGO_RS06085	SGO_1239	-5.42
SGO_RS09885	SGO_2021	-5.43
SGO_RS09860	SGO_2016	-5.50
SGO_RS09865	SGO_2017	-5.80
SGO_RS02500	SGO_0506	-5.97
SGO_RS05855	SGO_1192	-6.02
SGO_RS07805	SGO_1593	-6.02
SGO_RS09855	SGO_2015	-6.05
SGO_0832	SGO_0832	-6.09
rplL	SGO_1191	-6.69
SGO_RS08270	SGO_1686	-8.49
SGO_RS08335	SGO_1700	-8.95
SGO_RS07715	SGO_1575	-13.07
SGO_RS07720	SGO_1576	-27.01
celC	SGO_1578	-27.16
SGO_RS07725	SGO_1577	-27.21
SGO_RS07735	SGO_1579	-29.56
SGO_RS07750	SGO_1582	-54.98
SGO_RS07745	SGO_1581	-81.68
celB	SGO_1580	-101.81

S. gordonii Δ*ahrC* microarray data, regulation of genes in response to arginine depletion

(Note: locus tags have been changed)

Locus tag	**Gene name**	**FC (abs)**
SGO_0177	SGO_0177	71.72
SGO_RS00865	SGO_0175	62.38
SGO_RS00870	SGO_0176	55.01
SGO_RS03190	SGO_0648	9.80
SGO_RS03175	SGO_0645	8.93
SGO_RS06270	SGO_1277	8.52
SGO_RS03185	SGO_0647	8.35
SGO_RS03180	SGO_0646	8.32
SGO_RS09770	SGO_1997	5.83
SGO_RS04295	SGO_0874	5.10
SGO_RS00105	SGO_0021	4.83
xerS	SGO_1126	4.56
SGO_RS02455	SGO_0496	4.29
SGO_RS05590	SGO_1137	4.10
SGO_RS06280	SGO_1280	3.57
SGO_RS04720	SGO_0964	3.56
SGO_RS02125	SGO_0427	3.55
SGO_RS05430	SGO_1105	3.53
SGO_RS01645	SGO_0334	3.48
SGO_RS08580	SGO_1750	3.47
SGO_RS09845	SGO_2012	3.44
SGO_RS02040	SGO_0409	3.38
SGO_RS05865	SGO_1194	3.37
SGO_RS00525	SGO_0105	3.35
SGO_RS06460	SGO_1317	3.34
SGO_RS08585	SGO_1751	3.32
SGO_RS04340	SGO_0883	3.31
SGO_RS07650	SGO_1561	3.31
SGO_RS08115	SGO_1656	3.29
SGO_RS05870	SGO_1195	3.27
rnpA	SGO_0178	3.22
SGO_RS00530	SGO_0106	3.21
SGO_RS08590	SGO_1752	3.21
SGO_RS10205	SGO_2084	3.19
SGO_RS07655	SGO_1562	3.17
SGO_RS01670	SGO_0339	3.16

SGO_RS02045	SGO_0410	3.15
SGO_RS07660	SGO_1564	3.13
SGO_RS02135	SGO_0430	3.11
SGO_RS00880	SGO_0179	3.10
SGO_RS01980	SGO_0397	3.06
SGO_RS02960	SGO_0601	3.03
SGO_RS06910	SGO_1409	3.02
SGO_RS02050	SGO_0411	3.01
SGO_RS03075	SGO_0624	3.01
SGO_RS05255	SGO_1070	3.01
SGO_RS02780	SGO_0564	2.99
SGO_RS02965	SGO_0602	2.99
SGO_RS04640	SGO_0948	2.97
SGO_RS00775	SGO_0157	2.97
SGO_RS04685	SGO_0957	2.97
SGO_RS03360	SGO_0682	2.96
SGO_RS10200	SGO_2083	2.95
SGO_RS06915	SGO_1410	2.89
SGO_RS03275	SGO_0665	2.87
SGO_RS02775	SGO_0562	2.86
SGO_RS07890	SGO_1610	2.85
SGO_RS02810	SGO_0571	2.84
SGO_RS00450	SGO_0090	2.81
SGO_RS02585	SGO_0523	2.79
SGO_RS01830	SGO_0368	2.78
SGO_RS02815	SGO_0572	2.78
SGO_RS07645	SGO_1560	2.77
SGO_RS02770	SGO_0561	2.73
SGO_RS00895	SGO_0182	2.70
SGO_RS09090	SGO_1856	2.69
SGO_RS05875	SGO_1196	2.69
SGO_RS05595	SGO_1138	2.66
SGO_RS06905	SGO_1408	2.66
SGO_RS05575	SGO_1134	2.64
SGO_RS01370	SGO_0278	2.63
SGO_RS01425	SGO_0289	2.61
SGO_RS02590	SGO_0524	2.58
SGO_RS06920	SGO_1411	2.58
SGO_RS04250	SGO_0865	2.58
SGO_RS05580	SGO_1135	2.58
SGO_RS08110	SGO_1655	2.57

SGO_RS05450	SGO_1110	2.57
hisB	SGO_1407	2.57
SGO_RS07685	SGO_1569	2.55
SGO_RS05810	SGO_1181	2.53
SGO_RS03390	SGO_0688	2.51
SGO_RS01320	SGO_0268	2.51
SGO_RS10280	SGO_2101	2.50
SGO_RS02510	SGO_0508	2.49
SGO_RS08495	SGO_1733	2.49
SGO_RS01435	SGO_0291	2.48
SGO_RS06285	SGO_1281	2.48
SGO_RS03605	SGO_0734	2.48
SGO_RS01975	SGO_0396	2.47
SGO_RS00885	SGO_0180	2.47
SGO_RS04725	SGO_0965	2.46
SGO_RS08800	SGO_1794	2.44
SGO_RS00890	SGO_0181	2.44
SGO_RS02970	SGO_0603	2.43
SGO_RS10285	SGO_2102	2.43
SGO_RS07825	SGO_1597	2.43
SGO_RS07350	SGO_1498	2.40
SGO_RS01380	SGO_0280	2.40
SGO_RS00475	SGO_0095	2.38
SGO_RS07810	SGO_1594	2.38
SGO_RS09970	SGO_2041	2.37
SGO_RS10075	SGO_2062	2.37
SGO_RS07815	SGO_1595	2.37
SGO_RS01430	SGO_0290	2.36
SGO_RS02935	SGO_0596	2.36
SGO_RS03280	SGO_0666	2.35
SGO_RS01325	SGO_0269	2.35
SGO_RS00430	SGO_0086	2.34
SGO_RS02515	SGO_0509	2.34
SGO_RS02080	SGO_0417	2.33
SGO_RS01330	SGO_0270	2.31
SGO_RS07355	SGO_1499	2.27
SGO_RS03580	SGO_0729	2.26
SGO_RS07820	SGO_1596	2.26
SGO_RS06560	SGO_1337	2.25
SGO_RS07285	SGO_1484	2.25
SGO_RS06790	SGO_1385	2.25

SGO_RS09465	SGO_1933	2.24
SGO_RS05425	SGO_1104	2.24
SGO_RS07280	SGO_1483	2.21
SGO_RS08985	SGO_1831	2.21
SGO_RS04145	SGO_0845	2.20
SGO_RS08160	SGO_1664	2.20
SGO_RS02850	SGO_0579	2.19
SGO_RS01990	SGO_0399	2.18
SGO_RS01985	SGO_0398	2.18
SGO_RS09765	SGO_1996	2.18
SGO_RS09775	SGO_1998	2.17
SGO_RS03570	SGO_0726	2.17
SGO_RS04475	SGO_0912	2.15
SGO_RS00250	SGO_0052	2.15
SGO_RS02075	SGO_0416	2.15
SGO_RS09470	SGO_1934	2.15
SGO_RS07680	SGO_1568	2.13
SGO_RS05940	SGO_1210	2.13
hisH	SGO_1406	2.13
SGO_RS07150	SGO_1457	2.13
SGO_RS02595	SGO_0525	2.12
SGO_RS06795	SGO_1386	2.11
SGO_RS02730	SGO_0553	2.11
SGO_RS08805	SGO_1795	2.11
SGO_RS00240	SGO_0050	2.10
SGO_RS01970	SGO_0395	2.10
SGO_RS08150	SGO_1662	2.10
SGO_RS01820	SGO_0367	2.10
SGO_RS05585	SGO_1136	2.10
SGO_RS06785	SGO_1384	2.10
SGO_RS05280	SGO_1075	2.09
SGO_RS02525	SGO_0511	2.07
SGO_RS02085	SGO_0418	2.07
SGO_RS02370	SGO_0479	2.07
SGO_RS07490	SGO_1529	2.05
SGO_RS05805	SGO_1180	2.04
SGO_RS04825	SGO_0984	2.04
SGO_RS08520	SGO_1738	2.04
SGO_RS02845	SGO_0578	2.03
SGO_RS05420	SGO_1103	2.03
SGO_RS00255	SGO_0053	2.03

SGO_RS04385	SGO_0894	2.03
SGO_RS03490	SGO_0709	2.03
SGO_RS04485	SGO_0915	2.02
SGO_RS09005	SGO_1835	2.02
SGO_RS03715	SGO_0756	2.01
SGO_RS09975	SGO_2042	2.01
SGO_RS01245	SGO_0252	2.01
SGO_RS03670	SGO_0747	2.01
SGO_RS09965	SGO_2040	2.01
SGO_RS04280	SGO_0870	2.00
SGO_RS09840	SGO_2011	2.00
SGO_RS00305	SGO_0063	-2.00
SGO_RS09095	SGO_1857	-2.00
SGO_RS05715	SGO_1161	-2.01
SGO_RS08555	SGO_1745	-2.02
SGO_RS00320	SGO_0066	-2.02
SGO_RS03555	SGO_0722	-2.02
SGO_RS09195	SGO_1878	-2.03
SGO_RS05325	SGO_1084	-2.04
SGO_RS04930	SGO_1006	-2.04
aspS	SGO_0434	-2.04
SGO_RS04935	SGO_1007	-2.04
SGO_RS07085	SGO_1444	-2.05
SGO_RS05320	SGO_1083	-2.05
SGO_RS01585	SGO_0322	-2.05
SGO_RS04770	SGO_0973	-2.06
SGO_RS01775	SGO_0359	-2.06
SGO_RS01160	SGO_0235	-2.07
SGO_RS02145	SGO_0432	-2.08
SGO_RS09320	SGO_1903	-2.08
SGO_RS01205	SGO_0244	-2.08
SGO_RS08060	SGO_1645	-2.08
SGO_RS07985	SGO_1630	-2.09
SGO_RS02250	SGO_0454	-2.09
SGO_RS00310	SGO_0064	-2.09
SGO_RS09385	SGO_1917	-2.10
SGO_RS06930	SGO_1413	-2.10
SGO_RS07535	SGO_1538	-2.12
SGO_RS06040	SGO_1230	-2.12
SGO_RS01005	SGO_0205	-2.12
SGO_RS07970	SGO_1627	-2.13

SGO_RS02805	SGO_0570	-2.14
SGO_RS07580	SGO_1547	-2.15
SGO_RS09495	SGO_1939	-2.16
SGO_RS08570	SGO_1748	-2.16
SGO_RS05205	SGO_1060	-2.16
SGO_RS07530	SGO_1537	-2.16
SGO_RS07095	SGO_1446	-2.17
SGO_RS07585	SGO_1548	-2.17
SGO_RS06735	SGO_1374	-2.19
SGO_RS03845	SGO_0783	-2.19
SGO_RS07550	SGO_1541	-2.20
SGO_RS01200	SGO_0243	-2.20
SGO_RS01355	SGO_0275	-2.20
cbiO	SGO_0016	-2.20
SGO_RS05200	SGO_1059	-2.21
SGO_RS09485	SGO_1937	-2.22
SGO_RS06045	SGO_1231	-2.25
scpA	SGO_1671	-2.25
SGO_RS00325	SGO_0067	-2.26
SGO_RS02155	SGO_0435	-2.28
SGO_RS06925	SGO_1412	-2.28
gatA	SGO_0436	-2.28
SGO_RS07575	SGO_1546	-2.29
SGO_RS07975	SGO_1628	-2.31
SGO_RS02600	SGO_0526	-2.32
SGO_RS08045	SGO_1642	-2.32
SGO_RS07555	SGO_1542	-2.33
SGO_RS03885	SGO_0792	-2.33
SGO_RS03890	SGO_0793	-2.34
SGO_RS06810	SGO_1389	-2.34
SGO_RS07525	SGO_1536	-2.34
SGO_RS05920	SGO_1206	-2.34
SGO_RS09210	SGO_1881	-2.35
SGO_RS00710	SGO_0144	-2.36
SGO_RS07640	SGO_1559	-2.36
SGO_RS02165	SGO_0437	-2.38
ligA	SGO_1390	-2.39
SGO_RS09480	SGO_1936	-2.39
SGO_RS02130	SGO_0429	-2.40
polC	SGO_1847	-2.41
SGO_RS02610	SGO_0528	-2.41

SGO_RS05490	SGO_1118	-2.42
SGO_RS10095	SGO_2066	-2.42
SGO_RS06295	SGO_1283	-2.44
rpsA	SGO_1234	-2.45
SGO_RS01250	SGO_0253	-2.46
SGO_RS00695	SGO_0140	-2.46
SGO_RS02140	SGO_0431	-2.47
SGO_RS05270	SGO_1073	-2.48
SGO_RS07190	SGO_1465	-2.48
SGO_RS04980	SGO_1015	-2.50
engB	SGO_1139, *yihA, ysxC*	-2.50
SGO_RS07145	SGO_1456	-2.51
SGO_RS01000	SGO_0204	-2.51
SGO_RS09930	SGO_2030	-2.51
scpB	SGO_1670	-2.52
SGO_RS10090	SGO_2065	-2.53
SGO_RS00330	SGO_0068	-2.53
SGO_RS00085	SGO_0017	-2.53
SGO_RS08185	SGO_1669	-2.55
SGO_RS07980	SGO_1629	-2.55
SGO_RS09935	SGO_2031	-2.55
SGO_RS09940	SGO_2033	-2.55
SGO_RS01590	SGO_0323	-2.56
SGO_RS01360	SGO_0276	-2.57
SGO_RS00315	SGO_0065	-2.58
SGO_RS05605	SGO_1140	-2.59
SGO_RS06185	SGO_1259	-2.63
SGO_RS03955	SGO_0806	-2.63
SGO_RS03520	SGO_0715	-2.65
SGO_RS02615	SGO_0529	-2.65
pheS	SGO_0859	-2.67
SGO_RS02800	SGO_0569	-2.70
SGO_RS09925	SGO_2029	-2.72
SGO_RS07885	SGO_1609	-2.73
SGO_RS03530	SGO_0717	-2.74
SGO_RS07140	SGO_1455	-2.76
SGO_RS08180	SGO_1668	-2.78
SGO_RS09200	SGO_1879	-2.80
SGO_RS03950	SGO_0805	-2.81
SGO_RS04220	SGO_0860	-2.86
SGO_RS02835	SGO_0576	-2.87

SGO_RS04940	SGO_1008	-2.88
SGO_RS04225	SGO_0861	-2.89
SGO_RS09615	SGO_1964	-2.93
SGO_RS02830	SGO_0575	-2.94
SGO_RS03525	SGO_0716	-2.98
SGO_RS09920	SGO_2028	-2.98
era	SGO_0713, *bex, rbaA, sdgE, yqfH*	-3.06
SGO_RS04780	SGO_0975	-3.07
SGO_RS05035	SGO_1026	-3.08
SGO_RS07515	SGO_1534	-3.12
glyQ	SGO_0568	-3.12
SGO_RS04985	SGO_1016	-3.13
SGO_RS08320	SGO_1697	-3.20
SGO_RS04795	SGO_0978	-3.26
SGO_RS07520	SGO_1535	-3.28
SGO_RS03960	SGO_0807	-3.33
SGO_RS06805	SGO_1388	-3.35
SGO_RS08755	SGO_1785	-3.35
SGO_RS04990	SGO_1017	-3.36
SGO_RS03855	SGO_0785	-3.40
SGO_RS08325	SGO_1698	-3.41
SGO_RS03860	SGO_0786	-3.43
SGO_RS04785	SGO_0976	-3.53
SGO_RS03865	SGO_0787	-3.54
SGO_RS03850	SGO_0784	-3.55
SGO_RS08295	SGO_1691	-3.56
SGO_RS06180	SGO_1258	-3.56
SGO_RS05315	SGO_1082	-3.61
SGO_RS03815	SGO_0777	-3.63
SGO_RS08290	SGO_1690	-3.65
SGO_RS06820	SGO_1391	-3.72
SGO_RS08315	SGO_1695	-3.74
gyrB	SGO_1245	-3.75
SGO_RS03965	SGO_0808	-3.83
SGO_RS04790	SGO_0977	-3.86
SGO_RS09915	SGO_2027	-3.87
SGO_RS10265	SGO_2098	-3.90
SGO_RS06110	SGO_1244	-3.91
SGO_RS07635	SGO_1558	-3.93
SGO_RS08300	SGO_1692	-3.96
SGO_RS08310	SGO_1694	-3.98

SGO_RS08330	SGO_1699	-3.99
SGO_RS05855	SGO_1192	-4.12
SGO_RS03820	SGO_0778	-4.14
SGO_RS06090	SGO_1240	-4.16
SGO_RS06095	SGO_1241	-4.28
SGO_RS02035	SGO_0408	-4.30
SGO_RS08315	SGO_1695	-4.40
SGO_RS06105	SGO_RS06105	-4.44
SGO_RS08305	SGO_1693	-4.44
ileS	SGO_0681	-4.51
SGO_RS07630	SGO_1557	-4.56
SGO_RS06100	SGO_1242	-4.57
rplL	SGO_1191	-4.87
SGO_RS08275	SGO_1687	-4.93
SGO_RS05040	SGO_1027	-5.00
SGO_RS06085	SGO_1239	-5.04
SGO_RS09905	SGO_2025	-5.11
SGO_RS02500	SGO_0506	-5.19
SGO_RS09900	SGO_2024	-5.24
SGO_RS07800	SGO_1592	-5.35
SGO_RS09910	SGO_2026	-5.36
SGO_RS08280	SGO_1688	-5.43
SGO_RS08285	SGO_1689	-5.44
SGO_RS09875	SGO_2019	-5.48
SGO_RS09870	SGO_2018	-5.52
SGO_RS09890	SGO_2022	-5.69
SGO_RS09860	SGO_2016	-5.90
SGO_RS09880	SGO_2020	-5.93
SGO_RS09865	SGO_2017	-6.29
SGO_RS09895	SGO_2023	-6.29
SGO_RS09885	SGO_2021	-6.33
SGO_RS07805	SGO_1593	-6.58
SGO_0832	SGO_0832	-6.75
SGO_RS09855	SGO_2015	-6.79
SGO_RS08270	SGO_1686	-8.97
SGO_RS07715	SGO_1575	-10.88
SGO_RS08335	SGO_1700	-10.95
SGO_RS07725	SGO_1577	-29.81
SGO_RS07720	SGO_1576	-30.83
celC	SGO_1578	-32.03
SGO_RS07735	SGO_1579	-34.01
SGO_RS07750	SGO_1582	-68.89
celB	SGO_1580	-113.64
SGO_RS07745	SGO_1581	-121.43

Appendix C: Relevant publication

Molecular Microbiology (2015) 97(2), 281–300 ■

doi:10.1111/mmi.13023
First published online 20 May 2015

Critical roles of arginine in growth and biofilm development by *Streptococcus gordonii*

Nicholas S. Jakubovics,[1]* Jill C. Robinson,[1] Derek S. Samarian,[2] Ethan Kolderman,[2] Sufian A. Yassin,[1] Deepti Bettampadi,[2] Matthew Bashton[3] and Alexander H. Rickard[2]
[1]*School of Dental Sciences, Newcastle University, Newcastle upon Tyne, UK.*
[2]*School of Public Health, Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA.*
[3]*Bioinformatics Support Unit, Newcastle University, Newcastle upon Tyne, UK.*

Summary

***Streptococcus gordonii* is an oral commensal and an early coloniser of dental plaque. *In vitro*, *S. gordonii* is conditionally auxotrophic for arginine in monoculture but biosynthesises arginine when coaggregated with *Actinomyces oris*. Here, we investigated the arginine-responsive regulatory network of *S. gordonii* and the basis for conditional arginine auxotrophy. ArcB, the catabolic ornithine carbamoyltransferase involved in arginine degradation, was also essential for arginine biosynthesis. However, *arcB* was poorly expressed following arginine depletion, indicating that *arcB* levels may limit *S. gordonii* arginine biosynthesis. Arginine metabolism gene expression was tightly co-ordinated by three ArgR/AhrC family regulators, encoded by *argR*, *ahrC* and *arcR* genes. Microarray analysis revealed that > 450 genes were regulated in response to rapid shifts in arginine concentration, including many genes involved in adhesion and biofilm formation. In a microfluidic salivary biofilm model, low concentrations of arginine promoted *S. gordonii* growth, whereas high concentrations (> 5 mM arginine) resulted in dramatic reductions in biofilm biomass and changes to biofilm architecture. Collectively, these data indicate that arginine metabolism is tightly regulated in *S. gordonii* and that arginine is critical for gene regulation, cellular growth and biofilm formation. Manipulating exogenous arginine concentrations may be an attractive approach for oral biofilm control.**

Accepted 3 April, 2015. *For correspondence. E-mail nick.jakubovics@ncl.ac.uk; Tel. (+44) 191 208 6796; Fax (+44) 191 208 8807.

Introduction

Oral streptococci including *Streptococcus sanguinis*, *S. mitis*, *S. oralis* and *S. gordonii* are pioneer colonisers of tooth surfaces and provide the foundations for the formation of mixed-species dental plaque biofilms (Kolenbrander *et al.*, 2010). These primary-colonising streptococci produce multiple cell surface protein adhesins that promote attachment to the salivary pellicle and aid the recruitment of other bacteria (Nobbs *et al.*, 2011). Shifts in the microbial population in dental plaque are responsible for the development of dental caries or periodontitis (Jakubovics and Kolenbrander, 2010). For the successful colonisation of tooth surfaces, streptococci must obtain key nutrients for growth and survival from the extracellular environment. Saliva provides the major source of nutrients for bacteria in nascent dental plaque. *In vitro*, however, many streptococci isolated from dental plaque grow poorly in human saliva (Kolenbrander, 2011). Growth may be enhanced by the presence of microbial consortia, which together provide a pool of extracellular enzymes that efficiently degrade DNA, proteins and complex salivary carbohydrates such as host mucins (Bradshaw *et al.*, 1994). Streptococci also benefit from the presence of lactate-utilising bacteria, which remove the waste products of metabolism (Johnson *et al.*, 2009; Ramsey *et al.*, 2011). Thus, in order to form biofilms, streptococci must be able to maximise the use of nutrients provided by saliva and partner species in the human oral cavity.

Following the development of chemically defined media, in the 1970s, it became clear that oral streptococci lack the biosynthetic machinery for several amino acids. For example, cysteine is broadly required by strains of *S. sanguinis* and the cariogenic species *S. mutans* (Cowman *et al.*, 1974; 1975; Terleckyj and Shockman, 1975). Most strains of *S. sanguinis* also required arginine, tyrosine and at least one branched chain amino acid for growth (Cowman *et al.*, 1975). In mutans streptococci, requirements for several amino acids are dependent upon the growth conditions employed. Indeed, *S. criceti* AHT, a member of the mutans group, required arginine when cultured aerobically, but grew without arginine under

strictly anaerobic conditions (Terleckyj and Shockman, 1975). More recently, we have found that *S. gordonii* DL1 (Challis) also requires arginine for aerobic, but not anaerobic, growth (Jakubovics *et al.*, 2008a). It is not clear why exogenous arginine is essential only under aerobic conditions. It seems unlikely that oxygen directly inhibits arginine biosynthesis as, to the best of our knowledge, oxygen does not directly inhibit any enzymes in the arginine biosynthesis pathway. Furthermore, *S. gordonii* biosynthesises arginine aerobically if a low concentration of arginine is provided initially (Jakubovics *et al.*, 2008a). It is possible that oxygen has an indirect effect, for example by increasing protein damage that in turn places an increased demand on arginine for cell growth. In line with this, cell–cell contact (coaggregation) with another pioneer coloniser of dental plaque, *Actinomyces oris* MG1, triggers the upregulation of *S. gordonii* arginine biosynthesis genes, protects *S. gordonii* from protein carbonylation and enables aerobic growth of *S. gordonii* in arginine-restricted conditions (Jakubovics *et al.*, 2008a,b). *A. oris* MG1 produces catalase and degrades hydrogen peroxide produced by *S. gordonii*. Addition of catalase alone enhances *S. gordonii* growth in low arginine but is not sufficient to allow aerobic growth following a rapid shift to medium lacking arginine (Jakubovics *et al.*, 2008b). Therefore, the observed growth arrest when *S. gordonii* is rapidly shifted to aerobic media without arginine may be due to a combination of (i) a lack of sufficient arginine to synthesise essential biosynthetic enzymes and initiate *de novo* arginine biosynthesis and (ii) additional requirements for arginine imposed by oxidative stress.

At present, the full pathway for arginine biosynthesis by *S. gordonii* is not entirely clear. The conventional arginine biosynthesis genes *argCJBD* and *argGH* for conversion of L-glutamate to L-ornithine and L-citrulline to L-arginine, respectively, are present. However, there is no clear *argF* gene encoding anabolic ornithine carbamoyltransferase (OTCase) to convert L-ornithine to L-citrulline (Fig. S1). Two genes, *pyrB* and *arcB*, encode proteins with significant homology to ArgF (Jakubovics *et al.*, 2008a). By analogy with *S. aureus* (Nuxoll *et al.*, 2012), it is likely that *arcB* fulfils the role of the anabolic OTCase in *S. gordonii*. It has been proposed that *arcB* may be co-transcribed with *arcA*, encoding arginine deiminase, and controlled by the P_{arcA} promoter which, in turn, is most active under high arginine and low oxygen (Dong *et al.*, 2002; Zeng *et al.*, 2006; Liu *et al.*, 2008). Therefore, it may be predicted that the expression of *arcB* is low under *in vitro* aerobic arginine-restricted conditions. This in turn may contribute to the conditional arginine auxotrophy phenotype of *S. gordonii*.

A recent analysis of the regulation of amino acid influx and efflux pathways in *Escherichia coli* identified just three different logical circuitries connecting transport, biosynthesis and utilisation (Cho *et al.*, 2012). The response to arginine sensing by the ArgR regulator involves repression of transport and biosynthesis genes and activation of the arginine utilisation pathway. This circuitry is indicative of a primary role for arginine as a signal or cue rather than as a key nutrient or substrate (Cho *et al.*, 2012). Genomes of bacteria of the order *Lactobacillales* typically encode two or more orthologues of *E. coli* ArgR or the related regulator AhrC of *Bacillus subtilis*. For example, *S. pneumoniae* has three ArgR/AhrC family regulators even though it is apparently auxotrophic for arginine (Kloosterman and Kuipers, 2011). The presence of multiple ArgR/AhrC family regulators potentially enables a wide range of responses to different arginine concentrations and is consistent with a role for arginine as an important chemical cue for gene regulation.

Here, we aimed to investigate the roles of arginine in gene regulation, growth and biofilm formation by *S. gordonii*. Specifically, we set out to (i) assess the role of *arcB* in arginine biosynthesis and conditional arginine auxotrophy, (ii) determine the functions of ArgR/AhrC family regulators in the expression of L-arginine uptake, biosynthesis and catabolism genes, (iii) identify global gene regulation responses to arginine in *S. gordonii* and (iv) investigate the impact of L-arginine on biofilm formation in an environmentally germane model system.

Results

ArcB is pivotal for arginine biosynthesis and catabolism in S. gordonii

In silico analysis had previously indicated that two genes in the *S. gordonii* genome have significant homology to the anabolic OTCase (ArgF) of *Lactococcus lactis* (Jakubovics *et al.*, 2008a). On the basis of homology and genome context, the products of these genes have been annotated PyrB (gene locus SGO_1109) and ArcB (SGO_1592). The product of the *arcB* gene has relatively strong homology to ArgF (66% identity), whereas the *pyrB* gene product is only 24% identical to ArgF. To determine whether either the *pyrB* or *arcB* gene plays a role in arginine biosynthesis, the *pyrB* or *arcB* genes of *S. gordonii* DL1 (Challis) were replaced with the non-polar *aphA3* kanamycin resistance determinant. Under anaerobic conditions, *S. gordonii* DL1, *S. gordonii arcB*::*aphA3* and *S. gordonii pyrB*::*aphA3* grew well in chemically defined medium (CDM): in each case cultures reached a final turbidity of > 200 Klett Units (KU) within 24 h after inoculation. In CDM lacking arginine, *S. gordonii* DL1 and *S. gordonii pyrB*::*aphA3* grew anaerobically to > 200 KU. However, no growth of *S. gordonii arcB*::*aphA3* was observed in this medium. Therefore, *arcB* appears to be essential for arginine biosynthesis, in addition to its previously identified role in arginine

catabolism (Dong *et al.*, 2002). *S. gordonii pyrB*::*aphA3* did not grow in CDM without uracil, indicating that *pyrB* likely encodes an aspartate carbamoyltransferase for pyrimidine biosynthesis. To ensure that the presence of the *aphA3* cassette did not affect growth of strains, mutants were also constructed in which *pyrB* or *arcB* was replaced with the non-polar *ermAM* erythromycin resistance determinant and similar patterns of growth were observed (data not shown).

The arcB gene is co-transcribed with arcA

The *S. gordonii arcB* gene is located within a six-gene cluster comprising *arcABCDTR*. The expression of *arcA* is induced in low pH and high arginine (Liu *et al.*, 2008). To assess whether *arcB* is co-transcribed with *arcA*, *S. gordonii* DL1 was cultured anaerobically in CDM supplemented with 5 mM arginine to mid-exponential phase (125–175 KU), and RNA was extracted. The presence of mRNA containing *arcA-arcB* was detected by reverse transcription polymerase chain reaction (RT-PCR) using primers 1446F/1447R, which span the *arcA* and *arcB* genes (Fig. 1). In the absence of reverse transcriptase, no products were detected with these primers. Therefore, it is evident that *arcB* is co-transcribed with *arcA*.

To determine whether *arcA* and *arcB* are subject to similar patterns of gene regulation, *S. gordonii* was cultured under conditions aimed to induce the expression of arginine biosynthesis gene expression (growth in CDM supplemented with 5 mM arginine, followed by a shift to no arginine), or conditions favouring induction of arginine catabolism genes (growth in CDM with 10 mM glucose, followed by addition of 50 mM arginine) (Zeng *et al.*, 2006). For the arginine catabolism conditions, a small amount of extra glucose (approximately 10% higher than unamended CDM) was included to maintain a low background level of *arcA* and *arcB* expression prior to addition of arginine, as it has been shown that adding 10 mM glucose to complex medium represses expression from the P_{arcA} promoter (Dong *et al.*, 2004; Zeng *et al.*, 2006). Expression of *arcA* and *arcB* decreased steadily for 5 min under both sets of conditions (Fig. 2). Following a shift from 5 mM arginine to no arginine, *arcA* and *arcB* expression continued to decrease for 45 min. Conversely, following arginine addition to CDM supplemented with 10 mM glucose, the expression of *arcA* and *arcB* increased dramatically after 5 min and continued to increase for 45 min. In all samples, the changes in expression of *arcA* and *arcB* were similar, and this is consistent with a shared regulatory mechanism for the two genes. However, it should be noted that gene regulation does not necessarily correlate with enzyme activity as ArcA and ArcB may be subject to post-transcriptional regulation (Liu *et al.*, 2008).

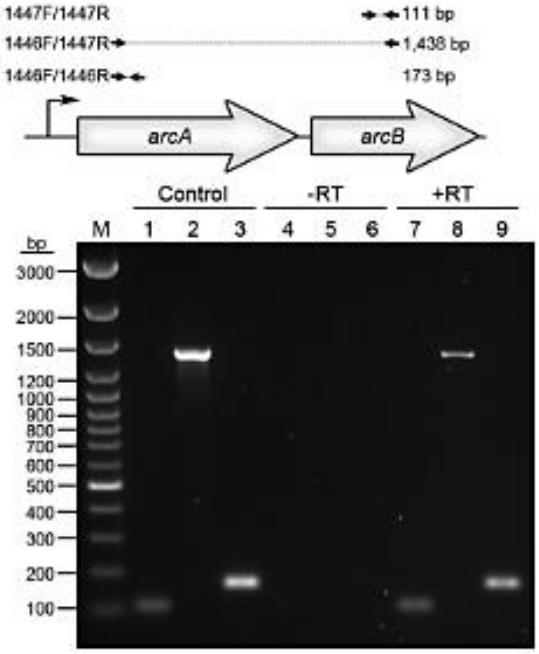


Fig. 1. Analysis of *arcA-arcB* gene transcript by RT-PCR. Combinations of 1446F, 1446R, 1447F and 1447R primers were used to amplify fragments of the *arcB* gene (111 bp; lanes 1, 4 and 7), *arcA* gene (173 bp; lanes 3, 6 and 9), or a region spanning *arcA-arcB* (1438 bp; lanes 2, 5 and 8). Positive control reactions (lanes 1–3) employed chromosomal DNA as a template for PCR. Alternatively, DNase I-treated RNA preparations were used as PCR templates either without reverse transcriptase (RT) (lanes 4–6) or with RT (lanes 7–9). The presence of a band at 1438 bp from cDNA template (lane 8) indicates that *arcA* and *arcB* are co-transcribed.

Increasing arcB *mRNA levels by genetic manipulation improves growth in low arginine*

We hypothesised that low levels of *arcB* expression during arginine-restrictive conditions may be responsible for the functional arginine auxotrophic phenotype of *S. gordonii* under aerobic laboratory conditions. To obtain increased *arcB* gene copies, plasmid pNJ-*arcB* was constructed in which the *arcB* gene was placed directly downstream of its native promoter P_{arcA} and was introduced into *S. gordonii arcB*::*aphA3*. The replication regions of pNJ-*arcB* originate from pTRKL2, which is maintained at 6–9 copies per cell (O'Sullivan and Klaenhammer, 1993). The complemented strain grew anaerobically in CDM without arginine and was able to grow aerobically at lower concentrations of arginine than *S. gordonii* DL1 (Table 1).

To further enhance the levels of *arcB* expression under arginine depletion, a copy of *arcB* was inserted into the *S. gordonii* chromosome downstream of *argD* and under control of the P_{argC} promoter (Fig. 3). Initially, attempts were made to introduce the *arcB* gene directly into the *S. gordo-*

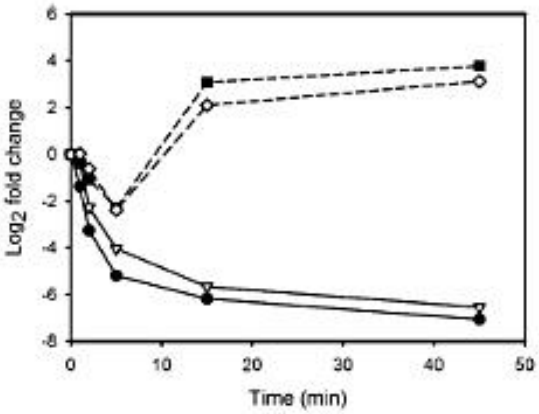


Fig. 2. Regulation of *arcA* and *arcB* genes in response to shifts in arginine concentrations. Cells were cultured in CDM supplemented with 10 mM glucose to mid-exponential phase, and arginine was added to a final concentration of 50 mM at time = 0 min (dashed lines). Alternatively, cells were cultured in CDM supplemented with 5 mM arginine to early exponential phase and, at time = 0 min, cells were harvested and resuspended in CDM lacking arginine. At intervals, aliquots were removed and expression of *arcA* (closed symbols) and *arcB* (open symbols) was determined by qRT-PCR. Total levels of RNA were normalised by comparison with 16S rRNA levels and relative levels compared with time = 0 min are shown.

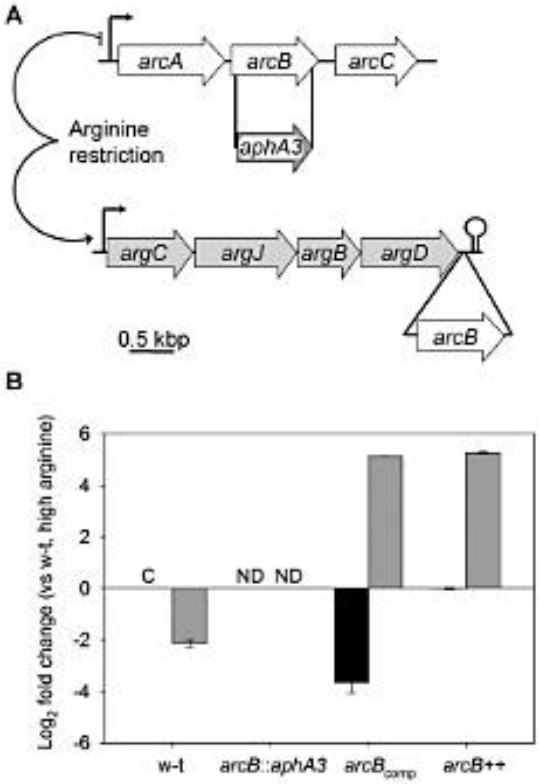


Fig. 3. Constructs for mutagenesis and complementation of *arcB*. A. The native position of *arcB* in the *arcABC* gene cluster is shown, along with the site of insertion of the *aphA3* cassette in the *arcB*::*aphA3* gene knockout construct. An additional strain was constructed in which *arcB* was replaced with the *ermAM* cassette at the same location (not shown). In the complementation strain, $arcB_{Comp}$, the *arcB* gene was inserted immediately downstream of *argD* and upstream of the predicted Rho-independent transcription terminator. An additional construct containing two copies of *arcB*, one downstream of *arcA* and one downstream of *argD* was also produced (*arcB*++). Predicted gene promoters are indicated by bent lines with arrows. Under arginine restriction, *arcA* promoter activity is reduced, whereas *argC* promoter activity is increased. Expression of *arcB* in the different strains after a shift from 5 mM arginine to either 5 mM arginine (black bars) or no arginine (grey bars) was determined by qRT-PCR (B). Values are means and SDs of log_2 fold change compared with *S. gordonii* DL1 in .5 mM arginine (marked as 'C' for comparator).

nii arcB::*aphA3* mutant with selection for transformants that were able to grow anaerobically on CDM agar without arginine. No transformants were obtained using this approach, even after several attempts. However, in control reactions, transformants were identified on CDM agar without arginine when the *arcB* complementation construct was introduced into *S. gordonii* DL1. The transformants contained a copy of *arcB* downstream of *argD* in addition to the native *arcB* gene within the *arcABC* operon, and this strain was labelled *S. gordonii arcB*++. To construct a strain with only one copy of *arcB*, located downstream of *argD*, the native copy of *arcB* in *S. gordonii arcB*++ was replaced with an *aphA3* kanamycin resistance cassette, generating *S. gordonii* $arcB_{Comp}$. The expression of *arcB* was assessed in each strain by quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) following anaerobic culture to mid-exponential phase in CDM

Table 1. Growth yield of *S. gordonii* strains in CDM amended to different arginine concentrations.

	Anaerobic (μM arginine)	Aerobic (μM arginine)							
Strain	0	0	8	16	32	64	128	256	512
w-t	+++[a]	–	–	–	++	+++	+++	+++	+++
arcB::*aphA3*	–	–	–	–	–	++	+++	+++	+++
pNJ-*arcB*	+++	–	+	++	+++	+++	+++	+++	+++
$arcB_{Comp}$	+++	–	–	++	+++	+++	+++	+++	+++
arcB ++	+++	–	+++	+++	+++	+++	+++	+++	+++

a. Values represent final growth yields after incubation in CDM amended to different concentrations of arginine for 96 h. Semi-quantitative assessment of growth from three independent experiments was as follows: < 51 KU (–), 51–150 KU (+), 151–250 KU (++) or > 250 KU (+++).

amended to 5 mM arginine, harvesting and re-suspension in either high (5 mM) arginine or no arginine (Fig. 3B). Expression of *arcB* in *S. gordonii* DL1 was fivefold lower in no arginine than in high arginine ($P < 0.001$). In 5 mM arginine, there were no significant differences in levels of *arcB* between *S. gordonii arcB*++ and the isogenic wild type. However, *arcB* levels were 12.5-fold lower in *S. gordonii* $arcB_{Comp}$ ($P < 0.001$). In no arginine, *arcB* was elevated in *S. gordonii arcB*++ and *S. gordonii* $arcB_{Comp}$ by 165-fold and 150-fold, respectively, in comparison with the wild type under arginine restriction ($P < 0.001$). Therefore, the relocation of *arcB* to a position immediately downstream of *argD* resulted in strong upregulation of *arcB* in response to arginine depletion, and a second copy of *arcB* under control of the P_{arcA} promoter in *S. gordonii arcB*++ prevented reduced *arcB* expression under high arginine.

The effects of relocating the *arcB* gene on growth of *S. gordonii* in low arginine were assessed by measuring the final growth yield of cells after culture in CDM amended to different arginine concentrations (Table 1). Anaerobically, all strains of *S. gordonii* except the *arcB* mutant grew strongly in the absence of arginine. In aerobic conditions, *S. gordonii* DL1 did not grow in CDM containing 16 μM arginine and grew moderately in 32 μM arginine. *S. gordonii arcB*::*aphA3* did not grow at any tested concentrations below 64 μM arginine. Moderate growth of *S. gordonii* $arcB_{Comp}$ was observed in CDM containing 16 μM arginine. Only *S. gordonii arcB*++ grew at very low arginine (8 μM), and none of the strains grew aerobically in medium without arginine. Therefore, the poor expression of *arcB* under low arginine conditions plays an important contribution to the lack of *S. gordonii* aerobic growth under low arginine.

Arginine biosynthesis and catabolism genes are co-ordinately regulated by three ArgR/AhrC family regulators

To identify the key regulators controlling arginine-dependent gene regulation in *S. gordonii*, the genome sequence of *S. gordonii* was BLAST-searched for genes encoding proteins with similarity to *Lactococcis lactis* ArgR or AhrC, *E. coli* ArgR or *B. subtilis* AhrC, and three sequences were identified (Fig. S2). ArcR (SGO_1588) has previously been characterised as an activator of arginine catabolism genes (Zeng *et al.*, 2006). The closest match to *E. coli* ArgR was encoded by SGO_2057 and is termed here *S. gordonii* ArgR. Searching with *B. subtilis* AhrC identified an *S. gordonii* AhrC orthologue, encoded by gene SGO_0697. Each of the predicted *S. gordonii* polypeptides include conserved amino acids that have been shown to be important for arginine-dependent transcriptional regulation (Fig. S2).

To investigate the function of *S. gordonii* ArgR/AhrC family regulators, each of the three genes (*arcR*, *argR* and *ahrC*) was disrupted by allelic exchange mutagenesis using a non-polar antibiotic insertion cassette, and double and triple mutants were produced as described in the *Experimental procedures*. To ensure that the observed effects of gene disruptions were not due to the introduction of antibiotic resistance cassettes, *argR*, *ahrC* and *arcR* were each disrupted with two different antibiotics, and patterns of regulation were shown to be similar in each mutant. Predicted Rho-independent terminators were identified downstream of *arcR* and *argR*. However, the *ahrC* gene is predicted to be in an operon with the DNA repair protein gene *recN* and SGO_0699 (Table S1). To ensure that the knockouts of *argR*, *ahrC* and *arcR* did not have polar effects on downstream genes, *S. gordonii* DL1 (wild type), *argR* and *ahrC* mutants were cultured in THB medium supplemented with 5 g l^{-1} yeast extract to mid-exponential phase (OD_{600} = 0.5–0.7). RNA was extracted and expression of *recN* and *mutS* (the gene downstream of *argR*) were assessed by qRT-PCR, normalised to levels of 16S rDNA expression. There were no significant differences in expression of *recN* or *mutS* in any of the strains (< 1.5 fold change between all strains). For *S. gordonii arcR*, the expression of the downstream gene *arcT* was assessed in high (0.5 mM) and no arginine as part of an ongoing microarray analysis, and in each case there was no difference in expression between the wild-type and mutant (data not shown). Therefore, disruptions of *argR*, *ahrC* and *arcR* did not have polar effects on downstream genes. In addition, the expression of *argR*, *ahrC* or *arcR* was not significantly altered in mutants lacking one or more of the ArgR/AhrC family regulators (data not shown).

The role of each regulator in controlling the expression of arginine metabolism genes in response to a shift in the arginine concentration was assessed by culturing strains anaerobically in CDM supplemented with 5 mM arginine to mid-exponential phase, harvesting and re-suspending in either CDM with 5 mM arginine or CDM without arginine. Expression of arginine metabolism genes was assessed by qRT-PCR (Fig. 4).

In *S. gordonii* DL1 (wild type), arginine biosynthesis genes *argC*, *argG* and $pyrA_b$ were strongly upregulated in response to a shift to no arginine (400-fold, 210-fold and 11-fold, respectively; $P < 0.001$ in each case) (Fig. 4A–C). Disruption of *argR* or *ahrC* resulted in strong expression of *argC* or $pyrA_b$ under 5 mM arginine or no arginine, indicating that both ArgR and AhrC are essential for down-regulation of *argC* and $pyrA_b$ in response to arginine. Disruption of *arcR* alone did not affect the expression of *argC* or $pyrA_b$ under 5 mM arginine and led to small but significant ($P < 0.01$) decreases in no arginine. Therefore, ArcR appears to play a minor role in promoting expression of *argC* or $pyrA_b$ in response to arginine depletion. By contrast, a clear role for ArcR was identified in regulation of *argG*. Disruption of *arcR* resulted in 11-fold increased

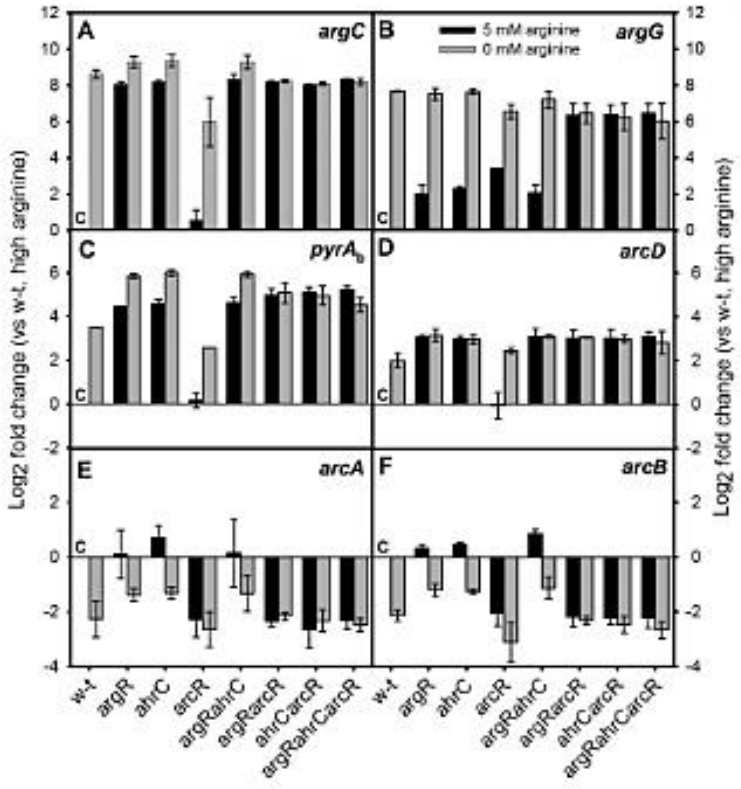


Fig. 4. Effects of disrupting ArgR/AhrC family regulators on the expression of arginine biosynthesis, transport and catabolism genes under high and no arginine. Anaerobically growing cells of *S. gordonii* DL1 (w-t) and isogenic *argR*, *ahrC* and *arcR* single mutants, *argR ahrC*, *argR arcR* and *ahrC arcR* double mutants and an *argR ahrC arcR* triple mutant were exposed to 5 mM arginine or no arginine for 30 min, and RNA was extracted. Expression of arginine biosynthesis genes (*argC*, $pyrA_b$ and *argG*), the gene encoding an arginine-ornithine antiporter (*arcD*) and arginine catabolism genes (*arcA* and *arcB*) was quantified by qRT-PCR. In each case, expression levels were compared with *S. gordonii* DL1 in 5 mM arginine (marked as 'C' for comparator). Bars represent means, and SDs from three independent experiments are shown. Note that different scales have been used for the *y*-axes.

expression of *argG* under high arginine compared with *S. gordonii* DL1 under the same conditions ($P < 0.001$). The effects of *arcR* disruption were independent of the presence or absence of ArgR or AhrC. Thus, the expression of *argG* in mutants disrupted in either *arcR* or *argR/ahrC* was partially reduced in 5 mM arginine compared with no arginine, whereas disruption of *arcR* in addition to *argR* and/or *ahrC* resulted in strong expression of *argG* independent of the arginine concentration.

On the *S. gordonii* genome, the *arcD* gene encoding an arginine-ornithine antiporter is immediately downstream and in the same direction as the *arcABC* genes, and it has been suggested that *arcD* may be part of the same operon (Dong *et al.*, 2002). However, in contrast to *arcA* or *arcB*, the expression of *arcD* was upregulated in low-arginine compared with 5 mM arginine (Fig. 4D; $P < 0.001$). In all mutants lacking *argR* or *ahrC*, expression of *arcD* was high, independent of arginine levels. Expression of *arcD* in the *arcR* single mutant was not significantly different from the wild type under 5 mM arginine or no arginine, indicating that ArcR does not regulate *arcD*. Using the promoter finding algorithm within Genome2D (http://genome2d.molgenrug.nl/), a putative promoter was identified immediately upstream of *arcD*. It is possible that *arcD* is also co-transcribed to some extent from the P_{arcA} promoter and that differences in mRNA stability across the transcript may also affect mRNA levels detected by qRT-PCR. Nevertheless, the above data strongly indicate that *arcD* expression is subject to different regulatory controls compared with *arcA* or *arcB*.

In *S. gordonii* DL1, expression of *arcA* and *arcB* was approximately four- to fivefold lower in CDM lacking arginine than in CDM containing 5 mM arginine (Fig. 4E and F; $P < 0.001$). Disruption of *argR* and/or *ahrC* did not significantly affect the expression of *arcA* and *arcB*. By contrast, arginine-dependent regulation of *arcA* and *arcB* was abrogated in all strains in which *arcR* was disrupted. In these mutants, *arcA* and *arcB* expression was low regardless of the arginine concentration, indicating the ArcR is required for optimal expression of *arcA* and *arcB* under high arginine. Together, the above data demonstrate that (i) ArgR and AhrC are both required for downregulation of arginine biosynthesis and transporter genes under high arginine, and (ii) ArcR acts independently of ArgR and AhrC to downregulate *argG* expression under high arginine. Furthermore, in agreement with previous observations (Dong *et al.*, 2002), ArcR is needed for upregulation of *arcA* and *arcB* genes under high arginine. This pattern of gene

regulation, in which both the biosynthesis and transporter genes are downregulated in high arginine, and catabolism genes are upregulated is similar to that identified in *E. coli* (Cho *et al.*, 2012).

Global gene regulation in response to arginine

The above data indicate that *S. gordonii* mounts a robust response to a shift in the arginine concentration involving the co-ordinated regulation of arginine biosynthesis, transport and catabolism genes. These experiments were performed using CDM supplemented to 5 mM arginine. The unamended CDM contains approximately 0.5 mM arginine, and we have previously observed that *S. gordonii* arginine biosynthesis genes are upregulated during batch growth in this medium, once arginine is depleted (Jakubovics *et al.*, 2008a). In preliminary experiments (not shown), we observed that arginine biosynthesis genes were strongly regulated in exponentially growing *S. gordonii* cells harvested and re-suspended in CDM without arginine compared with cells re-suspended in unamended CDM. We therefore chose to focus on comparing responses to 0.5 mM arginine with no arginine for studies on global arginine-mediated gene regulation.

A DNA microarray containing 2051 probes, covering > 95% of predicted *S. gordonii* genes, was designed and employed to assess global *S. gordonii* gene expression patterns in response to a 30 min exposure of anaerobically grown cells to high (0.5 mM arginine) or no arginine (see *Experimental procedures*). Initially, the microarray was validated by comparing microarray data with qRT-PCR for seven genes that had different levels of regulation in response to a shift from 5 mM arginine to no arginine and nine genes that were regulated by shifting from 0.5 mM arginine to no arginine (Fig. 5). All of the 16 genes analysed that were significantly regulated by microarray analysis were similarly regulated by qRT-PCR. By linear regression analysis there was a close correlation between data from microarrays and the combined data from qRT-PCR ($R^2 = 0.98$). The slope of the regression line was 0.94, indicating that the magnitude of gene regulation was similar independent of whether qRT-PCR or microarray was used and independent of whether cells were shifted to no arginine from 5 mM arginine or from 0.5 mM arginine.

In total, 464 genes were significantly regulated in response to arginine restriction, representing approximately 22.6% of all predicted *S. gordonii* genes. The complete list of regulated genes is presented in Table S1. Genes were assigned to clusters of orthologous groups based on predicted function (COGFun categories), and the number of genes in each group that were regulated in response to a shift in the arginine concentration are shown in Fig. S3. Overall, the COGFun group with the largest number of arginine-regulated genes was amino

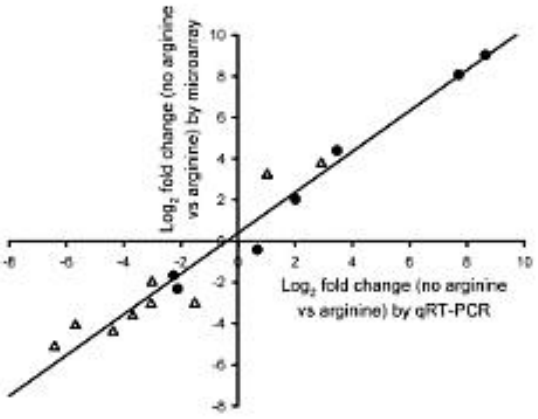


Fig. 5. Comparison between microarray data and qRT-PCR. Total RNA was extracted from *S. gordonii* following 30 min exposure to CDM either without or with arginine. The relative levels of gene expression in the absence of arginine compared with arginine-containing medium determined by microarray were plotted against levels assessed by qRT-PCR. The relative levels of expression of six arginine metabolism/transport genes (*argC*, *argG*, *pyrA*$_b$, *arcD*, *arcB* and *arcA*) and one control gene (*amyB*) in 0.5 mM arginine versus no arginine were assessed by microarray and compared with expression levels in 5 mM arginine versus no arginine, determined by qRT-PCR (closed circles). In addition, qRT-PCR was used to confirm the levels of expression of several genes (SGO_0846, *hsa*, *aspB*, *hisC*, *bfbC*, *bfbF*, SGO_1686, *wefE* and *wzg*) in the same RNA samples as those used for the microarray (open triangles). A linear regression line was drawn based on all the comparisons of qRT-PCR data with microarray data.

acid metabolism and transport (group E). In addition to arginine biosynthesis genes, a major group of genes encoding the histidine biosynthesis pathway was upregulated between 5- and 17-fold in no arginine. Genes encoding enzymes for biosynthesis of aromatic amino acids (*aroCBED*) and isoleucine/leucine/valine (*ilvH*, *ilvA*, *ilvB*, *ilvC*) were threefold to sixfold downregulated by arginine depletion. The oligopeptide transport system genes *hppH* and SGO_1716, the glutamine transport gene *glnQ* and an amino acid-binding permease gene (SGO_1727) were upregulated two- to sixfold in no arginine, whereas genes encoding the polyamine transporter (*potABCD*), putative amino acid permease (SGO_0985 and SGO_1482) and branched chain amino acid transport systems (*brnQ*, SGO_1626, SGO_1627, *braE*, *livH* and SGO_1630) were downregulated between 2- and 10-fold.

Several COGFun groups contained more members that were downregulated than upregulated when cells were exposed to CDM without arginine. In general, these pathways cover a diverse range of metabolic and biosynthetic pathways that are involved in cell maintenance and growth. Apart from genes with function unknown, only COGFun groups energy production and conversion (C), nucleotide metabolism and transport (F) and transcription

Table 2. Predicted operons containing genes that were strongly (> 10-fold) regulated by a shift from 0.5 mM arginine to no arginine in microarrays.

Locus	Description	Fold change (range)[a]
Upregulated (no arginine vs 0.5 mM arginine)		
SGO_1566–1569	ArgD/ArgB/ArgJ/ArgC, arginine biosynthesis	339.6 (260.9, 520.1)
SGO_0175–0177	ArgG/ArgH, arginine biosynthesis	267.6 (207.9, 342.4)
SGO_1656	Phosphoenolpyruvate carboxykinase	44.4
SGO_0645–0648	Hypothetical proteins	40.5 (32.4, 54.1)
SGO_0021	Hypothetical protein	26.7
SGO_1102–1106	PyrAa/PyrAb, arginine/pyrimidine biosynthesis	23.9 (18.7, 28.9)
SGO_0091–0094	Hypothetical proteins	12.1 (4.5, 36.6)
SGO_0874	Hypothetical protein	10.1
SGO_1401–1411	Histidine biosynthesis	9.4 (4.4, 17.0)
SGO_1831–1835	Hypothetical proteins	8.7 (5.5, 10.6)
Downregulated (no arginine vs 0.5 mM arginine)		
SGO_1575–1582	Bfb locus, biofilm formation and cellobiose PTS	–30.2 (–15.6, –81.9)
SGO_0831	Hypothetical protein	–13.7
SGO_0832	Hypothetical protein	–13.3
SGO_2098	RpsD, ribosomal protein S4	–12.5
SGO_1686–1700	Fab/acc locus, fatty acid biosynthesis	–11.4 (–6.4, –30.8)
SGO_0681	IleS, isoleucyl tRNA-synthetase	–11.1
SGO_2015–2028	Receptor polysaccharide biosynthesis	–8.6 (–4.0, –12.9)
SGO_0966–0978	Hsa, secondary secretion and glycosylation systems	–4.9 (–1.9, –10.8)

a. Fold increase (positive numbers) or decrease (negative numbers) in no arginine compared with 0.5 mM arginine. Where genes appear to be part of operons, the expression levels of the most strongly and most weakly expressed genes in the operon are shown (range).

(K) contained more members that were upregulated in no arginine than downregulated. Many of the genes involved in transcription encoded predicted transcription regulators, and it is possible that these were involved in co-ordinating the wider gene regulation response to arginine depletion. Overall, the effects of arginine depletion were consistent with an active reduction in cell growth.

Many of the genes that were most strongly regulated in response to a shift in arginine concentration were grouped in apparent operons. The structure of putative operons was predicted on the basis of gene location and orientation (Dehal *et al.*, 2010). Predicted operons with at least one gene that was regulated > 10-fold in response to arginine restriction are shown in Table 2. The most strongly regulated operons were those involved in arginine biosynthesis (*argCJBD* and *argGH*-SGO_0177) that were upregulated > 200-fold following a shift to no arginine. The arginine biosynthesis genes $pyrA_a$ and $pyrA_b$ were also strongly upregulated (– 24-fold) following arginine restriction. The histidine biosynthesis operon, SGO_1401–1411, was co-ordinately upregulated – 9-fold in low arginine. The SGO_1656 (*ppc*) gene was upregulated 44-fold in low arginine, and several single genes and putative multi-gene operons encoding hypothetical proteins were also strongly upregulated in response to arginine depletion. The most strongly downregulated operon in low arginine was the *bfb* gene locus encoding the cellobiose phosphotransferase system, which is also involved in biofilm formation and was downregulated – 30-fold. Other major multi-gene operons that were strongly downregulated in response to arginine depletion included the fatty acid biosynthesis operon (SGO_1686-SGO_1700), receptor polysaccharide biosynthesis (SGO_2015-SGO_2028) and the *hsa* gene locus encoding the Hsa adhesin and the secondary secretion apparatus (SGO_0966-SGO_0978). Single genes SGO_0831 and SGO_0832 encoding hypothetical proteins, *rpsD* encoding ribosomal protein S4 and *ileS* encoding isoleucyl tRNA synthetase were downregulated 11- to 14-fold in low arginine.

Gene regulation responses to arginine compared with other stimuli

To determine whether the observed regulatory responses were specific to arginine, or whether they were indicative of a more general stress response to amino acid depletion and growth arrest, cells were cultured in amino acid-replete CDM, and switched to CDM lacking L-arginine, L-histidine or branched chain amino acids (BCAA) L-leucine, L-isoleucine and L-valine. In each case amino acid depletion resulted in a rapid growth arrest (Fig. 6A), even though the *S. gordonii* genome encodes genes for biosynthesis of all these amino acids. After 30 min, the expression of 14 different genes in each medium was determined by qRT-PCR (Fig. 6B). The genes selected for this analysis included genes significantly upregulated by arginine depletion, genes downregulated and genes that were unchanged. The expression of two of the tested genes (*argC* and *asp5*) were significantly different between arginine depletion and depletion of either histidine or BCAA

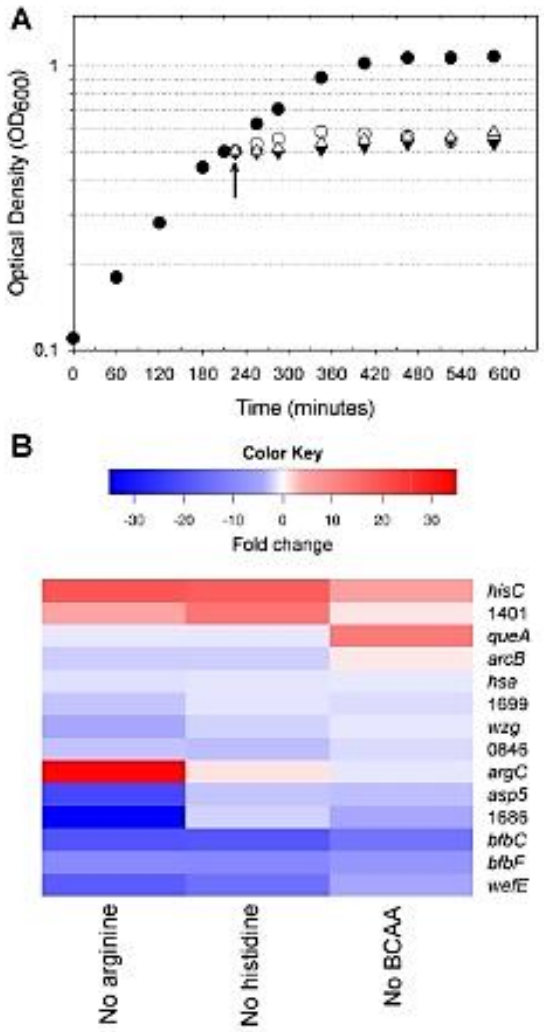


Fig. 6. Growth and gene expression in *S. gordonii* DL1 following depletion of arginine, histidine or BCAA.
A. Cells were cultured anaerobically in CDM to mid-exponential phase ($OD_{600} \sim 0.5$), harvested and resuspended in CDM (filled circles) or CDM lacking arginine (open circles), histidine (closed triangles) or BCAA (open triangles), indicated by an arrow, and growth was monitored until stationary phase.
B. 30 minutes after resuspension in different media, aliquots of cells were removed, and gene expression was monitored by qRT-PCR. Expression of 14 different genes is shown as a heatmap, and each colour represents the mean fold change compared with cells resuspended in CDM from four independent experiments.

($P < 0.005$). In addition, the expression of SGO_1686 was significantly different between CDM without arginine and CDM without histidine ($P < 0.05$). Several other genes appeared to be expressed at different levels following arginine depletion compared with depletion of BCAA, though the differences were not statistically significant. Therefore, the response to arginine depletion appears to involve a combination of stimulus-specific gene regulatory responses and a more general amino acid starvation stress response.

We have previously identified 23 genes in *S. gordonii* that were regulated in response to coaggregation with *A. oris*, including nine gene involved in arginine biosynthesis (Jakubovics *et al.*, 2008a). As arginine biosynthesis genes are regulated in response to changes in arginine concentration, we hypothesised that arginine may be a key signal for coaggregation sensing by *S. gordonii*. The effects of arginine depletion on the expression of the 23 coaggregation-regulated genes are shown in Fig. 7. In general, there was a strong correlation between the regulation of this set of genes under the two different conditions ($R^2 = 0.87$). The magnitude of the regulation was stronger in response to arginine restriction than to coaggregation (slope of the line = 0.44). Twenty-one of the 23 coaggregation-regulated genes were significantly changed under arginine restriction, as determined by the significance criteria outlined above. Only two genes (*spxB* and SGO_1308) that were regulated by coaggregation were not significantly regulated by arginine. Of these, *spxB* encodes pyruvate oxidase that is involved in the generation of hydrogen peroxide and may be important specifically in interbacterial interactions (Jakubovics *et al.*, 2008b). Overall, these data indicate that *S. gordonii* coaggregation-responsive genes are a subset of the genes regulated by arginine.

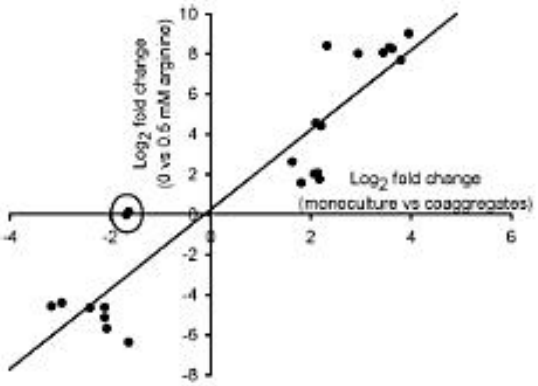


Fig. 7. Comparison between regulation of *S. gordonii* genes by coaggregation and by arginine depletion. The arginine-dependent expression of genes that had previously been identified as being regulated by coaggregation with *A. oris* was assessed using DNA microarrays. All genes that were significantly upregulated in monocultures compared with coaggregates were also upregulated in low arginine compared with high arginine. Most genes that were downregulated in monocultures were also downregulated in low arginine, with the exception of *spxB* (pyruvate oxidase) and SGO_1308 (hypothetical protein), which were not regulated by arginine (circled).

Effects of arginine on biofilm formation by S. gordonii

To assess the impact of L-arginine on biofilm formation by *S. gordonii*, cells were initially cultured for 24 h in a plastic 96 well microplate anaerobically in CDM adjusted to different concentrations of L-arginine. However, in this model, *S. gordonii* growth was reduced in low arginine concentrations, and the extent of biofilm formation, measured by staining with crystal violet, closely correlated with the amount of growth (linear regression, $R^2 = 0.94$; data not shown). This model was somewhat artificial as the mouth is an open system where nutrients are constantly replenished.

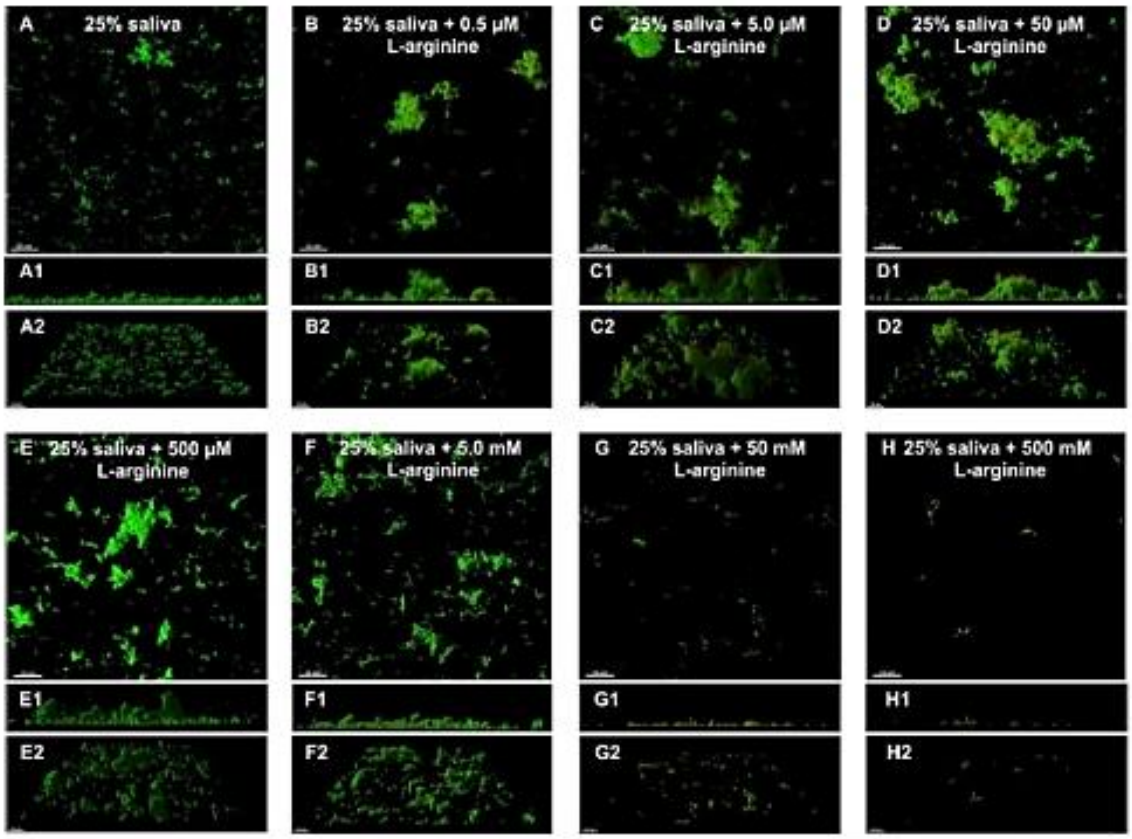
In order to assess the impact of L-arginine on *S. gordonii* biofilms grown under conditions representative of the oral cavity, a 24 channel Bioflux microfluidic system (Fluxion, San Francisco, CA) coupled to a Leica SPE CLSM (Leica, Exon, PA) was used. This system benefits from the requirement for only small volumes (< 1 ml) of saliva and carefully controlled flow and temperature conditions. Recently, oral care products have been developed that incorporate up to 8% (460 mM) L-arginine (Sullivan *et al.*, 2014), and it was of interest to determine whether either high or low L-arginine concentrations would affect *S. gordonii* biofilm formation. The high-throughput nature of the system made it possible to test a range (0.5 μM–500 mM) of arginine concentrations. Quantification of biofilm biomass and cell viability was enabled by Live/Dead stain and allowed arginine-dependent biofilm development to be characterised. An optimal range for enhanced biofilm development was observed when the 25% saliva was supplemented with between 0.5 μM and 500 μM arginine (Fig. 8).

A pre-treatment of the glass surfaces in the Bioflux microfluidic device with L-arginine at concentrations up to 500 mM did not appear to affect initial adhesion of *S. gordonii* cells to the substratum (Fig. S4). Three-dimensional rendering showed that *S. gordonii* biofilms grown for 22 h in 25% saliva formed thin, patchy biofilms (Fig. 8A) with an average biovolume of 0.89 $\mu m^3/\mu m^2$. Supplementing 25% saliva with between 0.5 μM and 500 μM arginine resulted in significant ($P < 0.05$) increases in biofilm biovolume by up to threefold. This was coincident with an increase in average biofilm thickness, although only biofilms developed in 500 μM arginine were significantly thicker than biofilms grown without added arginine ($P < 0.05$). Biofilms were structured in heterogeneous stack-like microcolonies, and there was a great deal of variation in thickness within individual samples, resulting in high error bars for this parameter (Fig. 8B–E). Based upon biofilm biovolume, average biofilm thickness, and biofilm roughness, *S. gordonii* biofilms developed in 25% saliva supplemented with 5 mM arginine were not statistically different ($P > 0.05$) from those developed in non-supplemented 25% saliva. However, biofilms developed in 50–500 mM arginine were substantially altered in biofilm architecture, thickness and biomass. Architecturally, the biofilms were increasingly patchy as the arginine concentration increased, and the likelihood of detecting the presence of aggregated micro-colonies was reduced. When developed in saliva containing 500 mM arginine, the biomass was significantly ($P < 0.05$) reduced by 15-fold, as compared with no added arginine, and possessed significantly reduced average thickness (35-fold decrease, $P < 0.05$). Roughness, which is a description of the variation in biofilm thickness, was also significantly different ($P < 0.05$). High concentrations of L-arginine, up to 500 mM, did not affect the growth of *S. gordonii* in planktonic cultures in CDM (data not shown). In addition, viable counts of *S. gordonii* in unamended saliva or in saliva adjusted to 500 mM L-arginine remained stable over 24 h, indicating that high concentrations of L-arginine were not toxic to *S. gordonii* in saliva (data not shown). The addition of L-arginine (for all experiments, as HCl salt) had little effect on the pH in the growth medium. The pH of saliva without arginine or with different L-arginine concentrations varied between 7.1 and 7.9. In general, the pH of the effluent was slightly higher and ranged between 7.9 and 8.3. The above data indicate that arginine stimulates *S. gordonii* biofilm development at lower concentrations (0.5–500 μM) and retards biofilm development at higher concentrations (50–500 mM).

In addition to the architectural changes that were caused by the supplementation of L-arginine, subtle effects on biofilm viability were observed (Fig. 8). As inferred from pixel intensity analysis (red/green) of Live/Dead stained biofilms, low (0.5–5 μM) and high (500 mM) concentrations of arginine caused significantly more cell-death/damage, when compared with the unsupplemented saliva. Although significant ($P < 0.05$), these might be a little misleading as they might be caused in-part by the architectural changes of the biofilms (e.g. Fig. 8A versus 8B) or loss of the majority of the viable biofilm cells in the flowing saliva, due to dispersive or de-adhesive effects of arginine, leaving damaged/dead cells behind (e.g. Fig. 8A versus 8E). In order to further investigate the viability of *S. gordonii* downstream of the biofilm model, cells in the effluent were visualised (Fig. S5). Images clearly showed that there were abundant cell masses in both unsupplemented saliva and in saliva supplemented with 500 mM L-arginine, indicating that *S. gordonii* had grown in both media and that the vast majority of cells were viable.

Discussion

The work presented here demonstrates that arginine has a concentration-dependent effect on *S. gordonii* gene expression and can alter the ability of this oral bacterium to form biofilms. In other bacterial species such as *E. coli*, high levels of exogenous arginine lead to repression of



Biofilm Parameter	Control	0.5 μM	5 μM	50 μM	500 μM	5 mM	50 mM	500 mM
Viability (%)	84.00 (4.56)	69.99 (8.69)**	76.78 (5.37)**	82.34 (4.42)	81.24 (4.05)	86.30 (3.34)	81.65 (4.99)	75.37 (7.09)**
Biomass ($\mu m^3/\mu m^2$)	0.89 (0.89)	1.12 (1.11)*	2.53 (1.68)*	1.78 (0.64)*	2.30 (1.32)*	0.90 (1.11)	0.21 (0.22)	0.06 (0.02)*
Thickness (μm)	2.15 (2.31)	2.10 (2.20)	4.91 (3.48)	3.71 (1.39)	5.04 (2.67)*	1.77 (2.21)	0.31 (0.38)*	0.06 (0.04)*
Roughness	1.74 (0.13)	1.80 (0.06)	1.72 (0.07)	1.66 (0.05)	1.65 (0.06)	1.76 (0.16)	1.90 (0.08)**	1.98 (0.01)**

Data were derived from at least three separate microfluidics channels
*P<0.05 and **P<0.01: significant differences from the control

Fig. 8. CLSM micrographs showing *S. gordonii* biofilms developed in different concentrations of L-arginine in 25% human saliva. Biofilm images are rendered in the XY dimension (A–H), the XZ dimension (A1–H1) and XYZ dimension (A2–H2). Images are ordered by increasing arginine concentration: control/no added arginine, 0.5 μM arginine, 5 μM arginine, 50 μM arginine, 500 μM arginine, 5 mM arginine, 50 mM arginine, 500 mM arginine. Bar represents 20 μm. Associated table shows biofilm characteristics after development in different arginine concentrations. Values represent an average of at least nine images from three different microfluidic channels.

arginine biosynthesis and transport genes by the arginine-dependent regulator ArgR and to increased expression of the arginine catabolism operon *astCADBE* (Cho *et al.*, 2012). This regulatory circuitry is consistent with a proposed role for arginine in signalling, rather than simply functioning as an exogenous nutrient (Cho *et al.*, 2012).

Here, we have demonstrated that *S. gordonii* has a similar regulatory logic, as arginine biosynthesis genes (*argCJBD*, *pyrA*$_a$*pyrA*$_b$, *argGH*) and arginine transport (*arcD*) are repressed in high arginine, whereas arginine catabolism (*arcABC*) is upregulated. However, the regulatory circuitry is more complex in *S. gordonii* as (i) *arcB* appears to have

dual roles in biosynthesis and catabolism and (iii) arginine-dependent gene regulation in *S. gordonii* involves the concerted actions of three ArgR/AhrC family regulators.

It appears that ArcB, an ornithine carbamoyltransferase (OTCase), is essential for arginine biosynthesis in *S. gordonii* as strains disrupted in *arcB* were unable to grow anaerobically in the absence of arginine. This enzyme catalyses the carbamoylation of the δ-amino group of ornithine by carbamoylphosphate to produce citrulline and inorganic phosphate. The production of citrulline is thermodynamically favoured, and *in vitro* ArcB enzymes are always assayed in the anabolic direction (Sainz *et al.*, 1998). However, studies on the catabolic OTCase from *Pseudomonas aeruginosa* (ArcB) have shown that it is essentially unidirectional *in vivo* due to poor affinity for carbamoyl phosphate and high cooperativity for this substrate (Tricot *et al.*, 1993). It is not clear whether catabolic OTCases from other bacteria are also subject to allosteric regulation or whether they direct catalysis towards citrulline catabolism by coupling with carbamate kinase, the next enzyme in the catabolic pathway. The *P. aeruginosa* genome also contains an *argF* gene encoding an anabolic OTCase. Mutants lacking a functional *argF* grew on minimal medium without arginine only after prolonged incubation, indicating that ArcB was either unable to function in the anabolic direction, or that its anabolic OTCase activity was very weak (Haas *et al.*, 1977). Our data indicate that *S. gordonii* ArcB can function for arginine biosynthesis in *S. gordonii*, but biosynthesis is only sufficient to sustain rapid growth under certain conditions, such as the gradual depletion of arginine during exponential growth in CDM (Jakubovics *et al.*, 2008a). *S. gordonii* DL1 does not grow aerobically after a rapid shift to no arginine, possibly due to a lack of time to accumulate a pool of carbamoyl phosphate as a substrate for ArcB or due to increased demand for arginine in the presence of oxygen. Other strains of *S. gordonii* appear to have similar phenotypes as *S. gordonii* Blackburn, Channon, FSS2 FSS3, M5 and PK488 also failed to grow in CDM without arginine (data not shown). Growth of all strains except *S. gordonii* Channon was restored in CDM containing high (8 mM) arginine.

The ability of ArcB to function in an anabolic direction may also be limited by poor gene expression following arginine depletion. The *arcB* gene is part of a six gene *arcABCDTR* cluster, in which *arcR* is present in reverse orientation compared with the other genes (Dong *et al.*, 2002). The promoter upstream of *arcA* (P_{arcA}) has been mapped and shown to contain two CRE box consensus elements that are recognised by the carbon catabolite protein CcpA and a 27 bp element that is bound by ArcR (Dong *et al.*, 2002; Zeng *et al.*, 2006). The expression of *arcA* is repressed by glucose in the presence of CcpA and is induced under anaerobic conditions by the Fnr-like protein Flp and the two-component system VicRK, and in low pH by the two-component systems CiaRH and ComDE (Dong *et al.*, 2002; 2004; Liu *et al.*, 2008; Liu and Burne, 2009). Furthermore, in glucose-grown cells, arginine sensing by ArcR results in approximately fourfold induction of expression from P_{arcA} (Zeng *et al.*, 2006). We have now demonstrated that *arcB* is co-transcribed with *arcA* and that the expression of *arcB* is also decreased in low arginine conditions. Relocating the *arcB* gene to a location downstream of *argD* and upregulated following arginine restriction significantly improved growth in low arginine, suggesting that poor expression of *arcB* is a major restriction on arginine biosynthesis in *S. gordonii in vitro*. However, even though relocation of *arcB* increased the levels of *arcB* transcripts 150-fold in low arginine, it did not enable aerobic growth in the absence of arginine.

The presence of multiple ArgR/AhrC family regulators is common in the *Lactobacillales*, perhaps reflecting a critical role for arginine sensing in this group of organisms. For, example *Lactobacillus plantarum* and *Lactococcus lactis* each have two paralogues of ArgR and AhrC, *S. pneumoniae* has three, and the *Enterococcus faecalis* genome encodes four ArgR/AhrC family proteins (Paulsen *et al.*, 2003; Larsen *et al.*, 2004; Nicoloff *et al.*, 2004; Kloosterman and Kuipers, 2011). In *L. plantarum*, ArgR1 and ArgR2 are both required for repression of arginine biosynthesis genes under high arginine, and mutations in the DNA binding or oligomerisation domains of either *argR1* or *argR2* genes abolish arginine-dependent repression (Nicoloff *et al.*, 2004). Similarly, in *L. lactis*, ArgR and AhrC act interdependently to control arginine biosynthesis and catabolism gene expression, and it has been proposed that these may combine in the presence of arginine to form a heterohexameric complex that is an active repressor (Larsen *et al.*, 2004; 2008). However, the DNA-binding activities of AhrC and ArgR regulons are not completely equivalent, and promoter binding assays indicate that AhrC interferes with ArgR binding to the promoter upstream of the arginine catabolic operon (Larsen *et al.*, 2005). *S. pneumoniae* contains three ArgR/AhrC regulators, of which ArgR1 and AhrC have been shown to act cooperatively to repress the expression of at least five promoters in response to high arginine (Kloosterman and Kuipers, 2011). In contrast to *L. lactis*, the *S. pneumoniae* ArgR1 and AhrC proteins are not involved in the control of the arginine catabolism operon *arcABC*. The third *S. pneumoniae* ArgR paralogue has not been analysed to date.

To the best of our knowledge, our data represent the first holistic analysis of the roles of three ArgR/AhrC family regulators in any organism. As in other bacteria, *S. gordonii* ArgR and AhrC act cooperatively to repress the expression of arginine biosynthesis and transport genes in high arginine. ArcR has already been shown to induce arginine catabolism genes under high arginine (Zeng

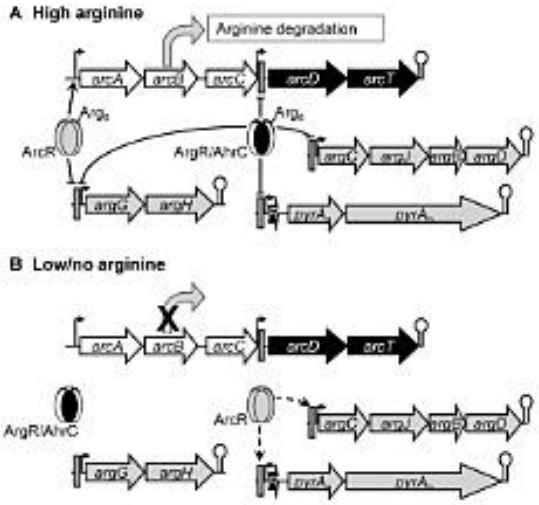


Fig. 9. A model of regulation of arginine metabolism genes by ArcR, ArgR and AhrC. ArgR and AhrC are dependent on each other for activity and here they are represented as a functional protein complex.
A. In the presence of arginine, ArcR, ArgR and AhrC are activated. This is shown as direct binding by six arginine residues (Arg_6). Activated ArgR/AhrC represses transcription of genes involved in arginine biosynthesis (shaded arrows) or accessory arginine-related functions (black arrows), indicated by lines with capped ends. In the presence of arginine, ArcR positively regulates expression of arginine catabolism genes (white arrows), shown by a line with an arrowhead, and negatively regulates *argGH* expression. Elements upstream of these genes that have consensus ARG box signatures are indicated by shaded boxes. Predicted promoters are indicated by thin right-facing arrows, and terminators are shown as loops and vertical lines.
B. In very low or no arginine, ArcR weakly upregulates (dashed lines) promoters upstream of *argC* and *pyrR*.

et al., 2006). Here, we have shown that ArcR also strongly represses *argGH* under high arginine. This presumably reduces the conversion of citrulline to arginine, and channels citrulline to the arginine catabolism pathway under high arginine (see Fig. S1). Under low arginine, ArcR had a minor stimulatory effect on the expression of *argCJBD* and $pyrA_aB_aA_b$. Therefore, all three ArgR/AhrC family regulators are required for the co-ordinated control of arginine metabolism gene expression in *S. gordonii*. A model for the functions of ArgR, AhrC and ArcR in the regulation of arginine metabolism genes is presented in Fig. 9. It is important to note that we have not investigated direct binding of ArgR/AhrC regulators to promoter regions, and it is possible that some regulatory effects may occur through other transcriptional regulators or by differential mRNA degradation.

Predictions of transcription factor binding sites at RegPrecise (http://regprecise.lbl.gov/RegPrecise/) or Genome2D (http://genome2d.molgenrug.nl/) databases identified putative ArgR/AhrC regulatory box elements upstream of a number of arginine-regulated genes, including *argC*, *argG*, *arcD*, *pyrR*, *serS*, *asd*, SGO_1716, SGO_1317, SGO_1656 and SGO_1716. However, these results must be interpreted with caution as searches also returned a number of 'false positives', where apparent regulatory elements were identified in genes that were not regulated in response to arginine depletion by microarray. It was not possible to search specifically for ArcR regulatory elements as the ArcR consensus element is not well established.

Global gene expression in response to arginine limitation or to disruption in ArgR/AhrC family regulators has been investigated in a number of bacteria. In *L. lactis*, disruption of *argR* and/or *ahrC* led to de-repression of arginine biosynthesis genes in high arginine (Larsen *et al.*, 2008). Disruption of *ahrC* also led to downregulation of arginine catabolism genes, whereas *argR* knockout resulted in slight increases in pyrimidine biosynthesis genes. In *S. pneumoniae*, growth in low arginine resulted in the upregulation of 13 genes including genes involved in amino acid or oligopeptide transport and arginine biosynthesis, and downregulation of five genes including *pyrD*, which is required for pyrimidine biosynthesis (Kloosterman and Kuipers, 2011). In *E. coli*, the ArgR regulon is extensive and includes 423 genes (Cho *et al.*, 2012). Many of these are controlled indirectly through the action of ArgR on other transcriptional regulators. Genes that are controlled directly by ArgR include those involved in arginine biosynthesis and transport, histidine biosynthesis and the biosynthesis of glutamate, aromatic amino acids and lysine. Our microarray analysis demonstrates that *S. gordonii* also mounts a major restructuring of gene expression in response to arginine restriction involving changes in expression of > 450 genes. As with other organisms, amino acid metabolism and transport are among the functions most strongly regulated by arginine. In addition, arginine modulates expression of genes in the pyrimidine metabolism pathway, which is closely linked to arginine metabolism. However, in *S. gordonii*, the overall impact of arginine restriction appears to be a reduction in processes associated with growth and metabolic activity such as protein synthesis, biosynthetic pathways and cell envelope biogenesis. In addition, among the most strongly regulated genes were those associated with adhesion and biofilm formation.

Selected genes were validated by qRT-PCR analysis, and we attempted to define the structure of operons based on *in silico* analyses combined with analysis of gene expression data. Several important operons were shown to be regulated by arginine. For example, Hsa, or its allelic variant GspB in certain strains of *S. gordonii*, is a critical adhesin for binding host glycoproteins and platelets (Takamatsu *et al.*, 2006; Jakubovics *et al.*, 2009; Pyburn *et al.*, 2011). The function of Hsa is dependent of secretion

by the SecA2-SecY2 system and five accessory secretory proteins (Asp proteins) encoded by genes present in the *hsa* locus (Yen *et al.*, 2013). Although the microarray data presented here indicated that all genes in the *hsa* locus were downregulated in response to arginine, there were differences in the level of regulation across the locus (Fig. S6). Generally, genes further downstream of *hsa* were more strongly regulated in response to arginine restriction than genes closer to *hsa*. It is likely that there are several promoters in the *hsa* gene locus and/or that there is selective degradation of mRNA from this region. Further studies will be required to identify the impact of fluctuations in extracellular arginine on Hsa function. Other major adhesion or biofilm formation loci were more consistently regulated in response to arginine. For example, all genes in the *bfb* (biofilm formation/cellobiose PTS) locus were downregulated > 15-fold following arginine restriction by microarray analysis (Table 2 and Fig. S6) and, in the case of *bfbC* and *bfbF*, strong downregulation was confirmed by qRT-PCR.

We have previously shown that *S. gordonii* responds to coaggregation with *Actinomyces oris* by upregulating genes involved in arginine biosynthesis and biofilm formation (Jakubovics *et al.*, 2008a). Here, we have demonstrated that arginine restriction influences the expression of 21 of the 23 genes that were shown to be responsive to coaggregation. The effects of coaggregation on gene expression were not seen in co-cultures of *S. gordonii* and *A. oris* in which cells were dispersed, indicating that physical contact between cells is required for gene regulation (Jakubovics *et al.*, 2008a). A number of previous studies have identified connections between arginine and biofilm formation in different bacteria. For example, at physiological concentrations found in cystic fibrosis sputum, arginine promotes *P. aeruginosa* biofilm formation and prevents swarming motility (Bernier *et al.*, 2011a). In model *P. aeruginosa* biofilms, arginine catabolism genes are upregulated compared with planktonic cells, leading to anaerobic metabolism and increased susceptibility to ciprofloxacin and tobramycin (Sauer *et al.*, 2002; Borriello *et al.*, 2004; Xu *et al.*, 2013). Similarly, arginine deiminase activity is upregulated in *S. aureus* and *S. pneumoniae* biofilms compared with planktonic cells (Zhu *et al.*, 2007; Allan *et al.*, 2014). In *E. faecalis*, the arginine-dependent regulators ArgR and AhrC are critical for biofilm formation *in vitro* (Kristich *et al.*, 2008), and AhrC is also required *in vivo* in a mouse model of catheter-associated urinary tract infection (Frank *et al.*, 2013). Interestingly, under static conditions in nutrient-rich media, we have observed that *S. gordonii* strains disrupted in the *arcR* gene form approximately 50% reduced biofilms compared with the isogenic wild-type progenitor (data not shown). This does not appear to be directly associated with the role of ArcR in regulating arginine biosynthesis, transport or catabolism genes as mutants in *argH*, *arcA*, *arcB* or *arcD* are not impaired in biofilm formation. Therefore it is possible that a different target of ArcR gene regulation plays a key role in biofilm formation, and we are currently investigating this hypothesis.

Using a microfluidic system that facilitates the growth of biofilms in flowing pooled human cell-fee saliva, we show that arginine has a major impact on biofilm formation in *S. gordonii*. Free arginine in whole saliva is generally very low, around 6 μM and increasing to approximately 8 μM following a protein-rich meal (Brand *et al.*, 1997). It has been suggested that microbial proteases may play a key role in releasing amino acids from salivary proteins (Syrjänen *et al.*, 1990), and therefore amino acid levels in the dental plaque microenvironment may depend upon the microbial species present. In the microfluidic model, arginine-supplemented saliva in the μM range enhanced biofilm development while the upper-mM range retarded biofilm development and altered biofilm architecture (Fig. 8). It is unclear why high concentrations of arginine cause reductions in the biomass of *S. gordonii* biofilms, though this observation is important as high concentrations of arginine are currently being incorporated into oral healthcare products (Sullivan *et al.*, 2014). The catabolism of arginine produces ammonia, which is alkaline, and trials are currently underway to assess the potential for arginine to be used as an anti-caries agent on the basis that the ammonia released by bacterial metabolism of arginine neutralises dental plaque acid (Nascimento *et al.*, 2009; 2013; 2014). It is possible that excessive alkaline production in biofilm cells of *S. gordonii* may trigger the release of cells from surfaces. Ammonia itself has been shown to be a signalling molecule that modulates biofilm formation and resistance to antibiotics (Nijland and Burgess, 2010; Bernier *et al.*, 2011b). However, in the microfluidics biofilm system used in this study, high concentrations of arginine did not result in significantly greater increases in pH of the effluent during biofilm growth than low arginine concentrations. Thus, arginine may inhibit cell–cell interactions directly. It is already known that arginine inhibits or retards coaggregation (Kamaguchi *et al.*, 1994; Levesque *et al.*, 2003) and autoaggregation (Merritt *et al.*, 2009). Changes in cell–cell interactions will likely be most apparent in biofilm models that incorporate fluid flow such as the microfluidic system employed in this study. Whether these effects extend to multi-species oral biofilms is currently being examined (manuscript in press).

In summary, our data strongly support the concept that arginine plays a major role in modulating key processes including growth and biofilm formation in *S. gordonii*. The regulatory response network for arginine is set up to allow arginine biosynthesis and growth when changes in external arginine are gradual, but to shut down cell growth in response to rapid depletion of arginine. We hypothesise that this regulatory architecture prevents *S. gordonii* from

over-committing resources to cell growth when arginine transiently reaches high concentrations, such as during a meal. High concentrations of arginine trigger the dispersal of *S. gordonii* from biofilms, which could potentially enable *S. gordonii* to relocate to distant sites in the mouth. We are now undertaking investigations into the mechanisms underlying this process. Early colonisers such as *S. gordonii* are critical for the initiation of dental plaque development and for recruitment of potentially pathogenic microorganisms. Ultimately, therefore, it may be possible to develop new strategies for oral biofilm control based on interfering with arginine sensing by oral bacteria.

Experimental procedures

Bacterial media and growth conditions

Streptococcus gordonii was routinely cultured in Todd Hewitt Broth (THB) medium (Difco, Detroit, MI) without shaking at 37°C or on THB solidified by the addition of 15 g l^{-1} Bacto-agar at 37°C in a candle jar. For some experiments, *S. gordonii* was cultured in CDM, prepared as described previously (Jakubovics *et al.*, 2008a) and incubated either aerobically or in an anaerobic environment under 90% N_2/5% H_2/5% CO_2. For gene regulation studies, L-arginine HCl (Sigma) was added to growth media as appropriate. Alternatively, when required, L-arginine was omitted from CDM (CDM-arg) or CDM was prepared without L-histidine or BCAA (L-leucine, L-isoleucine and L-valine). Alternatively, CDM was supplemented with L-arginine to a final concentration of 5 mM. Prior to growth in CDM, *S. gordonii* was cultured in TYEG medium containing (per l) 10 g Bacto tryptone, 5 g yeast extract, 3 g K_2HPO_4 and 2 g D-glucose, adjusted to pH 7.5 before autoclaving. For microfluidics biofilms, *S. gordonii* was initially cultured in Schaedler's medium (Difco) at 37°C without shaking. *E. coli* was cultured in Luria–Bertani (LB) medium or on LB medium solidified by the addition of 15 g l^{-1} Bacto-agar (Difco). For blue/white selection, 16 μl of 0.1 M isopropyl β-d-1-thiogalactopyranoside (IPTG) and 50 μl of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) were spread over solidified LB medium before adding cells. When required antibiotics were included in growth media at the following concentrations: erythromycin 2 μg ml^{-1} (for *S. gordonii*) or 100 μg ml^{-1} (for *E. coli*), ampicillin 50 μg ml^{-1}, spectinomycin 100 μg ml^{-1}, kanamycin 250 μg ml^{-1}.

Genetic manipulation of S. gordonii

Routine genetic manipulations were conducted as described by Sambrook and Russell (2001). All gene replacement constructs were generated using PCR overlap extension mutagenesis either with or without a cloning step in vector pGEM-T. Primers for mutagenesis are listed in Table S2. For disruption of *arcR*, *argR* or *ahrC* by insertion of an *ermAM* cassette, or disruption of *argR* by *aphA3* insertion, primers were designed to amplify approximately 500 bp regions upstream and downstream of the target genes from *S. gordonii* chromosomal DNA, with a central *Eco*RI restriction site. The 'F2' primer (forward primer for the region downstream of the gene of interest) contained a 5′ extension designed to overlap with the 'R1' primer (reverse primer for the upstream region). The upstream and downstream regions were PCR amplified, and the fragments were cleaned with the QIAquick PCR clean-up kit (Qiagen). Equimolar ratios of the products were combined and used as template for a second round of PCR. The product generated was cloned in pGEM-T vector (Promega) to generate pGEM-*arcR*, pGEM-*argR* or pGEM-*ahrC*, and used for transformation of *E. coli* JM109. To insert antibiotic resistance cassettes, *ermAM* or *aphA3* genes were PCR-amplified from plasmids pCM18 (Hansen *et al.*, 2001) or pSF151 (Tao *et al.*, 1992) with primers ermF1/R1 or aphA3F1/R1 that contained *Eco*RI restriction sites (Table S2). Fragments were cleaned, digested with *Eco*RI and ligated with pGEM-based plasmids to generate pGEM-*arcR*::*ermAM*, pGEM-*argR*::*ermAM*, pGEM-*argR*::*aphA3* or pGEM-*ahrC*::*ermAM* and used to transform *E. coli* JM109. Plasmids were screened for those that contained the antibiotic resistance cassette in the same orientation as the gene that it was replacing. Plasmid inserts were amplified with arcRF1/R2, argRF1/R2 or ahrCF1/R2 as appropriate, and products were used to transform *S. gordonii* DL1 (Challis) as previously described (Jakubovics *et al.*, 2005). All mutants were checked by PCR amplification and sequencing.

For disruption of *arcB* or *pyrB*, or replacement of *arcR* with the *aad9* spectinomycin resistance cassette amplified from plasmid pDL278 (LeBlanc *et al.*, 1992), mutagenesis was employed without a cloning step. Primers were designed to amplify approximately 500 bp upstream or downstream of the gene of interest. Extensions were added to the 5′ end of the 'R1' and 'F2' primers to overlap primers for amplification of the antibiotic resistance cassette. Following PCR amplification of the regions upstream and downstream of the gene of interest and the antibiotic resistance cassette, the three fragments were combined in equimolar quantities and used as template for a second round of PCR. For replacement of *arcB* with *aphA3*, the upstream and downstream regions of *arcB* were amplified with arcBF1/R1 and arcBF2/R2, respectively, and the *aphA3* cassette was amplified from plasmid pSF151 with primers aphA3F2/R2. To replace *arcB* with *ermAM*, *arcB* was amplified from *S. gordonii* chromosomal DNA using primers arcBF1/R3 and arcBF3/R2, in which 5′ overlap extensions were included in the 'R3' and 'F3' primers. The *ermAM* cassette was amplified from pCM18 with ermF1/

R1. Similarly, for *pyrB* mutagenesis, regions around the *pyrB* gene were amplified with pyrBF1/R1 and pyrBF2/R2 for disruption with *aphA3*, or with pyrBF1/R3 and pyrBF3/R2 for disruption with *ermAM*. Overall, these reactions generated products *arcB*::*ermAM*, *arcB*::*aphA3*, *pyrB*::*ermAM*, *pyrB*::*aphA3* or *arcB*::*aad9*. Fragments were cleaned and used for transformation of *S. gordonii* DL1 (Challis). A similar approach was employed to generate the *argD-arcB* complementation strain. Approximately 500 bp regions surrounding a predicted Rho-independent terminator (5′-AAAAGGATTCAGTTTGAGCTGGATTCTTTTT-3′) downstream of *argD* were amplified with primers argDF1/R1 and argDF2/R2 (Table S2). The *arcB* gene was amplified with primers arcBF4/R4. Following PCR amplification, the three products were mixed and used as a template for a second round of PCR with primers argDF1/R2. The *argD-arcB* fragment generated was cleaned and used for transformation of *S. gordonii* DL1. Transformants were selected on solidified CDM-arg medium. All transformants were checked by PCR amplification and DNA sequencing. For complementation of *arcB* mutants with *arcB* gene immediately downstream of the promoter P_{arcA}, the *arcB* gene region, P_{arcA} promoter and an approximately 5 kb region of plasmid pPE1010 (Egland *et al.*, 2004) were PCR amplified using primer pairs arcBF5/R5, ParcAF1/R1 and pPE_F1/R1, respectively, fused to generate plasmid pNJ-*arcB* and used for transformation of *E. coli* Stellar competent cells using the In-Fusion HD PCR ligation cloning kit (Clontech, Mountain View, CA). Plasmids were extracted, checked by DNA sequencing and used for transformation of *S. gordonii arcB*::*aphA3*.

Growth in chemically defined media

For assessing growth in CDM amended to different concentrations of L-arginine, cells were initially cultured on solidified TYEG medium for 96 h at 37°C and 5% CO_2. Individual colonies were subcultured to CDM and incubated for 24 h at 37°C, 5% CO_2. Cultures were diluted 1:100 in CDM and incubated for a further 24 h at 37°C, 5% CO_2. Cells were harvested by centrifugation, washed twice in CDM-arg and resuspended in CDM-arg. Cultures were used to inoculate CDM amended to various concentrations of arginine, to achieve an initial turbidity of between 25 and 30 Klett Units (KU), measured using a Klett-Summerson colorimeter with a 660 nm filter (Klett Manufacturing, New York). Cultures were incubated at 37°C anaerobically (90% N_2/5% H_2/5% CO_2) or aerobically (atmospheric CO_2) for 96–120 h, and the final growth yields were determined. Growth experiments were repeated three times independently and converted to a four point semi-quantitative scale [< 51 KU (–), 51–150 KU (+), 151–250 KU (++) or > 250 KU (+++)]. In most cases at least two of the three cultures had the same growth yield on the four-point scale, and this value was reported. Occasionally, three different values were obtained for the same strain at one arginine concentration, in which case the median value was given.

For experiments investigating transcriptional regulation in response to shifts in arginine, histidine or BCAA concentration, *S. gordonii* DL1 or isogenic ArgR-family regulator mutants were cultured in TYEG at 37°C for 24 h, with antibiotics as appropriate. Cells were subcultured to CDM and incubated at 37°C for 24 h. Cells were further subcultured and grown at 37°C in CDM to mid-exponential phase (140–160 KU). Cultures were split into two 4 ml portions, and each was harvested at 3800 *g*, 20°C for 10 min in a swing-out rotor. Cell pellets were resuspended in CDM-arg, CDM, CDM supplemented to 5 mM L-arginine or CDM without either L-histine or BCAA, and incubated at 37°C for up to a further 45 min.

RNA extraction and RT-PCR/qRT-PCR

Intracellular RNA was stabilised by the addition of 2 volumes of RNAProtect (Qiagen, Valencia, CA) and vortex mixing for 5 s, and RNA was extracted as previously described (Jakubovics *et al.*, 2008a). Briefly, cells were pelleted by centrifugation, the supernatant was removed and cells were stored at −70°C for up to 72 h. Cells were re-suspended in 1 ml Trizol reagent (Invitrogen, Carlsbad, CA), mixed with lysing matrix B (Qbiogene, Morgan, Irvine, CA) and disrupted in a FastPrep bead beater (Qbiogene). Subsequently, RNA was extracted using the Trizol manufacturer's protocol. Extracted RNA was treated for 1 h at 37°C with RQ1 DNase I (Promega, Madison, WI) and purified using RNeasy MinElute columns (Qiagen). A sample of RNA was analysed on a 0.8% (wt/vol) agarose gel containing 3% (vol/vol) formaldehyde to check for degradation. The concentration of RNA in each sample was estimated with a NanoDrop ND-1000 spectrophotometer (Labtech, Uckfield, East Sussex).

For RT-PCR and qRT-PCR analysis, samples were reverse transcribed with Superscript III reverse transcriptase (Invitrogen) and cleaned using MinElute columns (Qiagen). Primers 1446F/R and 1447F/R for RT-PCR analysis of *arcA*/*arcB* are described in Table S2. Reactions were carried out using REDTaq polymerase (Sigma-Aldrich, St Louis, MO) with the following thermocycle protocol: denaturation at 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, 56°C for 30 s and 72°C for 90 s, and a final elongation at 72°C for 5 min. Primers 1446F/R and 1447F/R were also used for qRT-PCR analysis of *arcA* and *arcB*. Other qRT-PCR primers are shown in Table S2, except the following that have previously been reported: 16SSgF1/R1 (16S rDNA), 0175F/R (*argG*), 1590F/R (*arcD*), 1569F/R (*argC*), 1104F/R ($pyrA_c$) and 1075F/R (*amyB*) (Jakubovics *et al.*, 2008a). Reactions (25 µl total volume) contained 0–10 ng cDNA template, 12.5 µl Power

SyBr Green PCR mix (Applied Biosystems, Foster City, CA) and forward/reverse primers each at 300 nM, with the exception of 16SSgF1/R1 reactions, which contained primers at 100 nM. An MX3005P thermocycler (Stratagene, La Jolla, CA) was employed for qRT-PCR using the protocol: 95°C for 10 min, 40 cycles of 95°C for 30 s, 56°C for 1 min and 72°C for 30 s, and a dissociation curve consisting of 95°C for 1 min, 56°C for 30 s, and incremental increases in temperature up to 95°C. Fluorescence readings were collected following the 56°C primer annealing step and throughout the dissociation curve. Specific amplification of the desired fragments was assessed by the presence of a single sharp fluorescence decrease during the dissociation phase and by analysis of representative samples on agarose gels. Reaction efficiencies were estimated by performing three independent reactions for each set of primers using dilutions of one *S. gordonii* cDNA sample as template over a 6-log range of concentrations. All primers gave reaction efficiencies > 80%. Relative quantities of transcripts were calculated from three independent experiments by normalising against the 16S rDNA gene as described previously (Jakubovics *et al.*, 2008a). The heatmap was drawn in R (R Core Team, 2014).

DNA microarray analysis

A microarray containing 2051 probes for *S. gordonii* genes was designed using the Agilent eArray platform (Agilent Technologies, Wokingham, Berkshire, UK). Custom settings were employed to design probes optimised for hybridisation at 65°C. Probe sequences and microarray data have been deposited in the Gene Expression Omnibus (GEO) database under accession numbers GSE51346 and GPL17786. An annotation file for the array was produced by aligning the probe sequences to the *S. gordonii* genome using Bowtie2 (Langmead and Salzberg, 2012). BEDTools (Quinlan and Hall, 2010) and custom Perl scripts were then used to produce the annotation for each probe using the GenBank file for the *S. gordonii* genome as a source of annotation. Samples of RNA from four independent experiments were sent to the Functional Genomics Unit, Birmingham University, UK, for reverse transcription, labelling and hybridisation. Data were analysed using GeneSpring software (Agilent). All data were normalised using the 75th percentile normalisation with baseline to median. Samples were taken from four independent experiments, and significant differences between expression levels in high or no arginine were assessed using *t*-tests with *P*-values corrected for multiple testing using Benjamini-Hochberg false discovery rate (FDR) correction (Reiner *et al.*, 2003) within Genespring GX 11 (Agilent) in conjunction with the custom annotation file. Genes were considered significantly regulated if they had FDR corrected *P*-value of ≤ 0.05 and the fold change was > 2. Functional Clusters of Orthologous Gene (COGFun) designations were taken from the MicrobesOnline database (http://meta.microbesonline.org/operons/gnc467705.html).

Saliva and inoculum preparation for biofilm experiments

Human saliva was collected from volunteers in accordance with the University of Michigan Institutional Review Board evaluated protocol (HUM00042954) described by Nance *et al.* (2013). Cell-free saliva (CFS) was used as the lone nutrient source. This was prepared using a similar protocol to that described by Rao *et al.* (2011). Briefly, saliva was gathered from five healthy adults who had not consumed anything but water for at least 2 h prior to collection. All donors had not taken any antibiotics for at least 3 months and did not smoke. The saliva from each donor was pooled, and 2.5 mM DTT was added before standing on ice for 10 min. The pooled saliva was then centrifuged at 20 000 *g* for 30 min. The supernatant was collected and diluted with distilled water to a final concentration of 25%. The 25% saliva was then filter sterilised using a 0.22 μm pore-size filter (Nalgene) to yield CFS. For long-term storage, CFS was separated into 30 ml aliquots and stored at −80°C. Prior to use, L-arginine HCl was added to CFS to final concentrations of 0.5 μM, 5 μM, 50 μM, 500 μM, 5 mM, 50 mM or 500 mM. A control with CFS containing no arginine was used for all experiments. Inocula were prepared by growing *S. gordonii* DL1 in 5 ml of Schaedler's broth that had been pre-reduced in an atmosphere containing 5% CO_2. The culture was grown under an atmosphere containing 5% CO_2 at 37°C until an OD_{600} of 0.4 was reached, whereupon it was used as an inoculum for microfluidics experiments.

Microfluidics system

Biofilms were developed using a Bioflux 200 microfluidics system (Fluxion, San Francisco, CA) with attached 48 well Bioflux microfluidics plates. Each Bioflux plate was conditioned with CFS prior to use. One hundred microlitres of each concentration of L-arginine-supplemented CFS was added in triplicate to each inlet well and flowed at 1.0 dyne cm^{-2} for 2–3 min at 20°C, followed by 20 min at 20°C with no flow in order to condition the channels for cell attachment. Exponentially growing *S. gordonii* DL1 cell suspensions in Schaedler's broth at an OD_{600} of 0.4 were flowed into the Bioflux system for 6 s from the outlet port at a speed of 1.0 dyne cm^{-2}, to facilitate inoculation of the viewing area and not further upstream (ie preventing contamination of the inlet reservoir). The plate was then incubated at 37°C for 45 min with no flow. The outlet wells containing the inoculum were aspirated, and 900 μl of each of the respective L-arginine concentrations of CFS

was added to each inlet well to bring the total inlet volume to approximately 1 ml. The CFS then flowed from the inlet to outlet wells for 20 h at 0.2 dyne cm^{-2} at 37°C.

Following incubation, all wells were aspirated, and the biofilms were washed with 100 µl of PBS for 20 min at 0.2 dyne cm^{-2} at 20°C. Biofilms were stained with BacLight LIVE/DEAD viability kit (Invitrogen, Grand Island, NY), prepared using 3 µl of each component (Styo-9 and propidium iodide) per 1 ml of PBS. After aspirating all PBS from the wells, 100 µl of the stain was added to each inlet well and flowed at 0.2 dyne cm^{-2} for 40 min at 20°C. All stain was then aspirated from the wells, and 100 µl of PBS was added to each inlet well and run at 0.2 dyne cm^{-2} for 20 min at room temperature to remove any excess stain.

Imaging and analysis of microfluidics biofilms

A Leica SPE confocal laser scanning microscope (CLSM) (Leica, Exon, PA) equipped with a 40 × 1.25 NA HCX PL APO infinity-corrected oil objective lens was used to image the biofilms. Images were obtained using a 488 nm laser set at 15% of maximum power, which allowed the excitation of both components of the BacLight LIVE/DEAD stain (Syto-9 and propidium iodide). The excitation capture range for Syto-9 was 510–540 nm, whereas the range for propidium iodide was 620–650 nm. A negative control of arginine-free CFS was used to calibrate the offset and gain for the microscope, which was then kept constant for image capturing of all biofilms within that experiment.

After image collection, IMARIS software (Bitplane, Zurich, Switzerland) was used to visualise the biofilms in 2D and 3D. Additionally, IMARIS was used to prepare the images for analysis using COMSTAT software (Heydorn *et al.*, 2000) and IMAGEJ (Schneider *et al.*, 2012). COMSTAT was used to quantify the biomass of the biofilm, as well as its average thickness and roughness. Thickness was measured from where one surface of the biofilm contacts the glass on the bottom of the plate to the top of the opposite surface in the channel in the field of view, whereas roughness was a measurement of thickness variability in a single field of view. For the LIVE/DEAD quantification, IMAGEJ was used. Green represented viable cells while red represented dead or damaged cells, and the pixel intensity of red and green was measured in throughout the image stacks using the *Histogram* function for each channel. The percentage of green and red, and therefore the percentage of viable and non-viable cells, was subsequently determined for each image stack. This was performed using Excel (Microsoft, Redmond, WA) to multiply the total pixels by the intensity (8 bit images, 0–255 in intensity levels). All software programs were run on a computer containing an Intel i5 processor (Intel, Santa Clara, CA) and a Radeon 5850 graphics card with 1 GB of RAM (AMD, Sunnyvale, CA).

Statistical tests

Testing for significant differences was performed by one-way analysis of variance using the Tukey's post-hoc test for pairwise comparisons.

Acknowledgements

The authors would like to thank Lesley Old, Claire Foley and Lydia Chapman, Newcastle University, for expert technical assistance and assistance with experiments and Lorraine Wallace, Functional Genomics, Proteomics and Metabolomics Facility, Birmingham University for very helpful advice with microarrays and for conducting the microarray hybridisations and scanning. We are also grateful to Dr Paul Kolenbrander for many helpful discussions and suggestions.

References

Allan, R.N., Skipp, P., Jefferies, J., Clarke, S.C., Faust, S.N., Hall-Stoodley, L., *et al.* (2014) Pronounced metabolic changes in adaptation to biofilm growth by *Streptococcus pneumoniae*. *PLoS ONE* **9:** e107015.

Bernier, S.P., Ha, D.G., Khan, W., Merritt, J.H., and O'Toole, G.A. (2011a) Modulation of *Pseudomonas aeruginosa* surface-associated group behaviors by individual amino acids through c-di-GMP signaling. *Res Microbiol* **162:** 680–688.

Bernier, S.P., Letoffe, S., Delepierre, M., and Ghigo, J.M. (2011b) Biogenic ammonia modifies antibiotic resistance at a distance in physically separated bacteria. *Mol Microbiol* **81:** 705–716.

Borriello, G., Werner, E., Roe, F., Kim, A.M., Ehrlich, G.D., and Stewart, P.S. (2004) Oxygen limitation contributes to antibiotic tolerance of *Pseudomonas aeruginosa* in biofilms. *Antimicrob Agents Chemother* **48:** 2659–2664.

Bradshaw, D.J., Homer, K.A., Marsh, P.D., and Beighton, D. (1994) Metabolic cooperation in oral microbial communities during growth on mucin. *Microbiology* **140:** 3407–3412.

Brand, H.S., Jorning, G.G., Chamuleau, R.A., and Abraham-Inpijn, L. (1997) Effect of a protein-rich meal on urinary and salivary free amino acid concentrations in human subjects. *Clin Chim Acta* **264:** 37–47.

Cho, B.K., Federowicz, S., Park, Y.S., Zengler, K., and Palsson, B.O. (2012) Deciphering the transcriptional regulatory logic of amino acid metabolism. *Nat Chem Biol* **8:** 65–71.

Cowman, R.A., Perrella, M.M., and Fitzgerald, R.J. (1974) Influence of incubation atmosphere on growth and amino acid requirements of *Streptococcus mutans*. *Appl Microbiol* **27:** 86–92.

Cowman, R.A., Perrella, M.M., Adams, B.O., and Fitzgerald, R.J. (1975) Amino acid requirements and proteolytic activity of *Streptococcus sanguis*. *Appl Microbiol* **30:** 374–380.

Dehal, P.S., Joachimiak, M.P., Price, M.N., Bates, J.T., Baumohl, J.K., Chivian, D., *et al.* (2010) MicrobesOnline: an integrated portal for comparative and functional genomics. *Nucleic Acids Res* **38:** D396–D400.

Dong, Y., Chen, Y.Y., Snyder, J.A., and Burne, R.A. (2002) Isolation and molecular analysis of the gene cluster for the

arginine deiminase system from *Streptococcus gordonii* DL1. *Appl Environ Microbiol* **68:** 5549–5553.
Dong, Y., Chen, Y.Y., and Burne, R.A. (2004) Control of expression of the arginine deiminase operon of *Streptococcus gordonii* by CcpA and Flp. *J Bacteriol* **186:** 2511–2514.
Egland, P.G., Palmer, R.J., Jr, and Kolenbrander, P.E. (2004) Interspecies communication in *Streptococcus gordonii–Veillonella atypica* biofilms: signaling in flow conditions requires juxtaposition. *Proc Natl Acad Sci USA* **101:** 16917–16922.
Frank, K.L., Guiton, P.S., Barnes, A.M., Manias, D.A., Chuang-Smith, O.N., Kohler, P.L., *et al.* (2013) AhrC and Eep are biofilm infection-associated virulence factors in *Enterococcus faecalis*. *Infect Immun* **81:** 1696–1708.
Haas, D., Holloway, B.W., Schambock, A., and Leisinger, T. (1977) The genetic organization of arginine biosynthesis in *Pseudomonas aeruginosa*. *Mol Gen Genet* **154:** 7–22.
Hansen, M.C., Palmer, R.J., Jr, Udsen, C., White, D.C., and Molin, S. (2001) Assessment of GFP fluorescence in cells of *Streptococcus gordonii* under conditions of low pH and low oxygen concentration. *Microbiology* **147:** 1383–1391.
Heydorn, A., Nielsen, A.T., Hentzer, M., Sternberg, C., Givskov, M., Ersboll, B.K., *et al.* (2000) Quantification of biofilm structures by the novel computer program COMSTAT. *Microbiology* **146:** 2395–2407.
Jakubovics, N.S., and Kolenbrander, P.E. (2010) The road to ruin: the formation of disease-associated oral biofilms. *Oral Dis* **16:** 729–739.
Jakubovics, N.S., Kerrigan, S.W., Nobbs, A.H., Stromberg, N., van Dolleweerd, C.J., Cox, D.M., *et al.* (2005) Functions of cell surface-anchored antigen I/II family and Hsa polypeptides in interactions of *Streptococcus gordonii* with host receptors. *Infect Immun* **73:** 6629–6638.
Jakubovics, N.S., Gill, S.R., Iobst, S.E., Vickerman, M.M., and Kolenbrander, P.E. (2008a) Regulation of gene expression in a mixed-genus community: stabilized arginine biosynthesis in *Streptococcus gordonii* by coaggregation with *Actinomyces naeslundii*. *J Bacteriol* **190:** 3646–3657.
Jakubovics, N.S., Gill, S.R., Vickerman, M.M., and Kolenbrander, P.E. (2008b) Role of hydrogen peroxide in competition and cooperation between *Streptococcus gordonii* and *Actinomyces naeslundii*. *FEMS Microbiol Ecol* **66:** 637–644.
Jakubovics, N.S., Brittan, J.L., Dutton, L.C., and Jenkinson, H.F. (2009) Multiple adhesin proteins on the cell surface of *Streptococcus gordonii* are involved in adhesion to human fibronectin. *Microbiology* **155:** 3572–3580.
Johnson, B.P., Jensen, B.J., Ransom, E.M., Heinemann, K.A., Vannatta, K.M., Egland, K.A., *et al.* (2009) Interspecies signalling between *Veillonella atypica* and *Streptococcus gordonii* requires the transcription factor CcpA. *J Bacteriol* **191:** 5563–5565.
Kamaguchi, A., Baba, H., Hoshi, M., and Inomata, K. (1994) Coaggregation between *Porphyromonas gingivalis* and mutans streptococci. *Microbiol Immunol* **38:** 457–460.
Kloosterman, T.G., and Kuipers, O.P. (2011) Regulation of arginine acquisition and virulence gene expression in the human pathogen *Streptococcus pneumoniae* by transcription regulators ArgR1 and AhrC. *J Biol Chem* **286:** 44594–44605.
Kolenbrander, P.E. (2011) Multispecies communities: interspecies interactions influence growth on saliva as sole nutritional source. *Int J Oral Sci* **3:** 49–54.
Kolenbrander, P.E., Palmer, R.J., Jr, Periasamy, S., and Jakubovics, N.S. (2010) Oral multispecies biofilm development and the key role of cell-cell distance. *Nat Rev Microbiol* **8:** 471–480.
Kristich, C.J., Nguyen, V.T., Le, T., Barnes, A.M., Grindle, S., and Dunny, G.M. (2008) 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* **74:** 3377–3386.
Langmead, B., and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9:** 357–359.
Larsen, R., Buist, G., Kuipers, O.P., and Kok, J. (2004) ArgR and AhrC are both required for regulation of arginine metabolism in *Lactococcus lactis*. *J Bacteriol* **186:** 1147–1157.
Larsen, R., Kok, J., and Kuipers, O.P. (2005) Interaction between ArgR and AhrC controls regulation of arginine metabolism in *Lactococcus lactis*. *J Biol Chem* **280:** 19319–19330.
Larsen, R., van Hijum, S.A., Martinussen, J., Kuipers, O.P., and Kok, J. (2008) Transcriptome analysis of the *Lactococcus lactis* ArgR and AhrC regulons. *Appl Environ Microbiol* **74:** 4768–4771.
LeBlanc, D.J., Lee, L.N., and Abu-Al-Jaibat, A. (1992) Molecular, genetic, and functional analysis of the basic replicon of pVA380-1, a plasmid of oral streptococcal origin. *Plasmid* **28:** 130–145.
Levesque, C., Lamothe, J., and Frenette, M. (2003) Coaggregation of *Streptococcus salivarius* with periodontopathogens: evidence for involvement of fimbriae in the interaction with *Prevotella intermedia*. *Oral Microbiol Immunol* **18:** 333–337.
Liu, Y., and Burne, R.A. (2009) Multiple two-component systems modulate alkali generation in *Streptococcus gordonii* in response to environmental stresses. *J Bacteriol* **191:** 7353–7362.
Liu, Y., Dong, Y., Chen, Y.Y., and Burne, R.A. (2008) Environmental and growth phase regulation of the *Streptococcus gordonii* arginine deiminase genes. *Appl Environ Microbiol* **74:** 5023–5030.
Merritt, J., Niu, G., Okinaga, T., and Qi, F. (2009) Autoaggregation response of *Fusobacterium nucleatum*. *Appl Environ Microbiol* **75:** 7725–7733.
Nance, W.C., Dowd, S.E., Samarian, D., Chludzinski, J., Delli, J., Battista, J., *et al.* (2013) A high-throughput microfluidic dental plaque biofilm system to visualize and quantify the effect of antimicrobials. *J Antimicrob Chemother* **68:** 2550–2560.
Nascimento, M.M., Gordan, V.V., Garvan, C.W., Browngardt, C.M., and Burne, R.A. (2009) Correlations of oral bacterial arginine and urea catabolism with caries experience. *Oral Microbiol Immunol* **24:** 89–95.
Nascimento, M.M., Liu, Y., Kalra, R., Perry, S., Adewumi, A., Xu, X., *et al.* (2013) Oral arginine metabolism may decrease the risk for dental caries in children. *J Dent Res* **92:** 604–608.
Nascimento, M.M., Browngardt, C., Xiaohui, X., Klepac-

Ceraj, V., Paster, B.J., and Burne, R.A. (2014) The effect of arginine on oral biofilm communities. *Mol Oral Microbiol* **29:** 45–54.
Nicoloff, H., Arsene-Ploetze, F., Malandain, C., Kleerebezem, M., and Bringel, F. (2004) Two arginine repressors regulate arginine biosynthesis in *Lactobacillus plantarum*. *J Bacteriol* **186:** 6059–6069.
Nijland, R., and Burgess, J.G. (2010) Bacterial olfaction. *Biotechnol J* **5:** 974–977.
Nobbs, A.H., Jenkinson, H.F., and Jakubovics, N.S. (2011) Stick to your gums: mechanisms of oral microbial adherence. *J Dent Res* **90:** 1271–1278.
Nuxoll, A.S., Halouska, S.M., Sadykov, M.R., Hanke, M.L., Bayles, K.W., Kielian, T., *et al.* (2012) CcpA regulates arginine biosynthesis in *Staphylococcus aureus* through repression of proline catabolism. *PLoS Pathog* **8:** e1003033.
O'Sullivan, D.J., and Klaenhammer, T.R. (1993) High- and low-copy-number *Lactococcus* shuttle cloning vectors with features for clone screening. *Gene* **137:** 227–231.
Paulsen, I.T., Banerjei, L., Myers, G.S., Nelson, K.E., Seshadri, R., Read, T.D., *et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299:** 2071–2074.
Pyburn, T.M., Bensing, B.A., Xiong, Y.Q., Melancon, B.J., Tomasiak, T.M., Ward, N.J., *et al.* (2011) A structural model for binding of the serine-rich repeat adhesin GspB to host carbohydrate receptors. *PLoS Pathog* **7:** e1002112.
Quinlan, A.R., and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26:** 841–842.
R Core Team. (2014) R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org.
Ramsey, M.M., Rumbaugh, K.P., and Whiteley, M. (2011) Metabolite cross-feeding enhances virulence in a model polymicrobial infection. *PLoS Pathog* **7:** e1002012.
Rao, D., Arvanitidou, E., Du-Thumm, L., and Rickard, A.H. (2011) Efficacy of an alcohol-free CPC-containing mouthwash against oral multispecies biofilms. *J Clin Dent* **22:** 187–194.
Reiner, A., Yekutieli, D., and Benjamini, Y. (2003) Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* **19:** 368–375.
Sainz, G., Tricot, C., Foray, M.F., Marion, D., Dideberg, O., and Stalon, V. (1998) Kinetic studies of allosteric catabolic ornithine carbamoyltransferase from *Pseudomonas aeruginosa*. *Eur J Biochem* **251:** 528–533.
Sambrook, J., and Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
Sauer, K., Camper, A.K., Ehrlich, G.D., Costerton, J.W., and Davies, D.G. (2002) *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. *J Bacteriol* **184:** 1140–1154.
Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9:** 671–675.
Sullivan, R., Rege, A., Corby, P., Klaczany, G., Allen, K., Hershkowitz, D., *et al.* (2014) Evaluation of a dentifrice containing 8% arginine, calcium carbonate, and sodium monofluorophosphate to prevent enamel loss after erosive challenges using an intra-oral erosion model. *J Clin Dent* **25:** A7–A13.
Syrjänen, S.M., Alakuijala, L., Alakuijala, P., Markkanen, S.O., and Markkanen, H. (1990) Free amino acid levels in oral fluids of normal subjects and patients with periodontal disease. *Arch Oral Biol* **35:** 189–193.
Takamatsu, D., Bensing, B.A., Prakobphol, A., Fisher, S.J., and Sullam, P.M. (2006) Binding of the streptococcal surface glycoproteins GspB and Hsa to human salivary proteins. *Infect Immun* **74:** 1933–1940.
Tao, L., LeBlanc, D.J., and Ferretti, J.J. (1992) Novel streptococcal-integration shuttle vectors for gene cloning and inactivation. *Gene* **120:** 105–110.
Terleckyj, B., and Shockman, G.D. (1975) Amino acid requirements of *Streptococcus mutans* and other oral streptococci. *Infect Immun* **11:** 656–664.
Tricot, C., Nguyen, V.T., and Stalon, V. (1993) Steady-state kinetics and analysis of pH dependence on wild-type and a modified allosteric *Pseudomonas aeruginosa* ornithine carbamoyltransferase containing the replacement of glutamate 105 by alanine. *Eur J Biochem* **215:** 833–839.
Xu, Z., Fang, X., Wood, T.K., and Huang, Z.J. (2013) A systems-level approach for investigating *Pseudomonas aeruginosa* biofilm formation. *PLoS ONE* **8:** e57050.
Yen, Y.T., Cameron, T.A., Bensing, B.A., Seepersaud, R., Zambryski, P.C., and Sullam, P.M. (2013) Differential localization of the streptococcal accessory Sec components and implications for substrate export. *J Bacteriol* **195:** 682–695.
Zeng, L., Dong, Y., and Burne, R.A. (2006) Characterization of cis-acting sites controlling arginine deiminase gene expression in *Streptococcus gordonii*. *J Bacteriol* **188:** 941–949.
Zhu, Y., Weiss, E.C., Otto, M., Fey, P.D., Smeltzer, M.S., and Somerville, G.A. (2007) *Staphylococcus aureus* biofilm metabolism and the influence of arginine on polysaccharide intercellular adhesin synthesis, biofilm formation, and pathogenesis. *Infect Immun* **75:** 4219–4226.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Appendix D: Conference and meeting attendance, and public engagement

North East Postgraduate Conference, Newcastle, 26th October 2012 (Poster)

Society for General Microbiology Autumn conference, Warwick University, 3rd-5th September 2012

Title: Characterisation of a biofilm-defective *arcR* mutant strain of *Streptococcus gordonii*

Abstract: *Streptococcus gordonii* is a bacterium commonly found in the healthy human mouth, and a primary coloniser of tooth surfaces. This gives it an important role in the establishment of dental plaque, which is a typical 'biofilm' - a multi-layered community of bacterial cells surrounded by a sticky protective matrix, found on the surface of teeth. Dental plaque is present in health but, if left unchecked, plays a causative role in the formation of both gum disease and tooth decay, the two most common oral diseases. A strain of *S. gordonii* has been generated that is defective in biofilm formation. This strain carries a deletion of the gene *arcR,* which encodes ArcR, a protein involved in the regulation of arginine catabolism. The role of ArcR in biofilm formation is currently unclear. The aim of this project was to investigate the *S. gordonii arcR* mutant with a view to characterising the contribution of ArcR to biofilm growth. A number of phenotypes were investigated, including cell surface hydrophobicity, cell surface protein and intracellular protein expression, appearance of biofilms (by atomic force microscopy imaging), and the relative fitness of the mutant in direct competition with the wild-type. It was found that the *arcR* deletion strain of *S. gordonii* exhibited structural differences from the wild-type - in addition to the unstable biofilm phenotype seen previously - including differences in cell surface protein expression. Future work aims to identify those proteins that are differentially expressed in order to understand the mechanisms of *S. gordonii* biofilm formation.

Characterisation of a biofilm-defective *arcR* mutant strain of *Streptococcus gordonii*

Jill C. Robinson (080838465), Dr. Nicholas Jakubovics[1], Dr. Matthew German[2], Prof. Waldemar Vollmer[3]
1. Oral Biology Lab, School of Dental Sciences, Newcastle University,
2. Dental Materials, School of Dental Sciences, Newcastle University,
3. Centre for Bacterial Cell Biology, Institute of Cell and Molecular Biosciences, Newcastle University.

Centre for Oral Health Research

Newcastle University

Introduction

Streptococcus gordonii is a primary coloniser of tooth surfaces and plays a key role in the establishment of dental biofilms[1]. A deletion strain of *S. gordonii*, lacking the gene *arcR*, is defective in biofilm formation. The *arcR* gene encodes the ArcR regulatory protein of the arginine deiminase system (ADS), which is involved in maintaining biofilm pH homeostasis and protecting against lethal acidification in the oral environment by converting arginine to ammonia and ATP[2]. This project aimed to further characterise the *arcR* mutant strain in order to understand the role of ArcR in biofilm formation.

Methods

- Crystal violet assay: Biofilms were grown in 6-well plates, and stained with crystal violet. The absorbance was read at A570 to give a measure of biofilm growth.
- Microbial adhesion to hexadecane (MATH) assay: Cell surface hydrophobicity was measured by adding n-hexadecane to an aqueous suspension of cells. The proportion of cells partitioning into the hexadecane phase ("% hydrophobicity") was determined by the reduction in A450 of the aqueous phase.
- SDS-PAGE: Cellular and surface proteins were extracted and separated on an SDS-PAGE gel to compare major protein expression in the wild-type and mutant strains.
- Atomic force microscopy (AFM) air imaging and force spectroscopy: AFM was used to test the stiffness of the surface of the biofilms by tapping the cell surfaces using the silicon probe. Air imaging was carried out on air-dried dehydrated biofilms.

Results

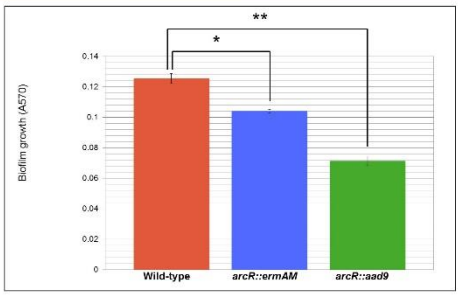


Figure 1. Quantitative analysis of the role of *arcR* in biofilm formation. Biofilms formed by *Streptococcus gordonii* DL1 (wild-type) and two mutants, PK3351 (*arcR::ermAM*) and PK3354 (*arcR::aad9*), were assessed by crystal violet staining. (T-test: * p <0.05, ** p <0.001).

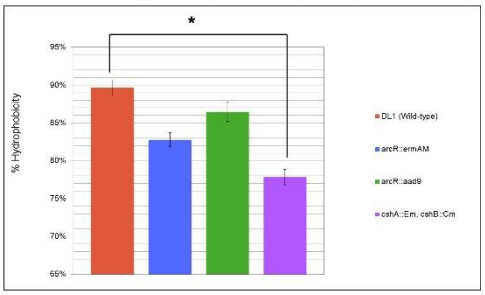


Figure 2. Cell surface hydrophobicity determined by MATH assay. A *cshA/cshB* deletion strain, known to have a lower cell surface hydrophobicity than wild-type, was used as control alongside the wild-type and two mutant strains. (T-test: * p <0.05).

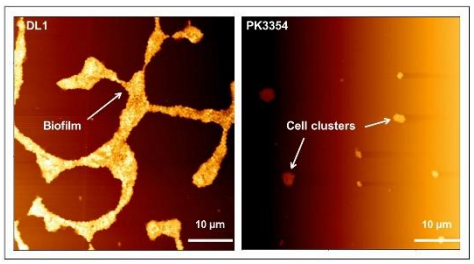


Figure 3. AFM air imaging of DL1 wild-type growth and PK3354 mutant growth. The mutant strain has been largely removed from the plate by the washing stage of the sample preparation, before imaging occurred.

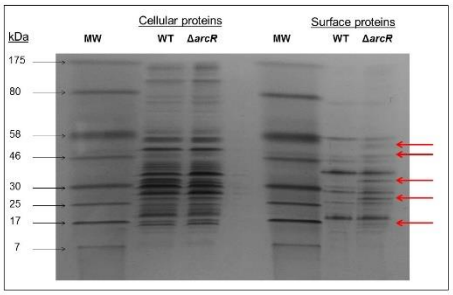


Figure 4. The protein profiles of wild-type and mutant strains. Lanes; WT) DL1 wild-type cell and surface proteins; Δ*arcR*) PK3354 mutant cell and surface proteins; MW) Molecular weight ladder. The red arrows on the right of the image mark out surface proteins that appear to have differential expression in the mutant to the wild-type strain.

Discussion and conclusions

1. The *arcR* mutants were impaired in their ability to form stable biofilms.
2. There were no significant differences between the cell surface hydrophobicity of the wild-type and *arcR* mutants.
3. AFM air imaging showed the wild-type strain formed biofilms, whereas any biofilm formed by the *arcR* mutant was removed during sample preparation, leaving only cell clusters remaining.
4. One-dimensional SDS-PAGE gel indicated there are differences in cell surface protein expression between the wild-type and *arcR* mutant.

Future work

- To determine whether extracellular factors are important in ArcR-dependent biofilm formation, growth and biofilm formation of the *arcR* mutant will be investigated in competition assays, in which the wild-type and mutant are grown in co-culture.
- Further SDS-PAGE analysis will be carried out to confirm the apparent differences in cell surface protein expression.
- AFM force spectroscopy will be used on biofilm samples to test the stiffness levels of the wild-type and *arcR* mutant biofilms. AFM imaging and force spectroscopy under liquid conditions will also be performed to test differences in growth and cell stiffness in hydrated conditions.

References

1. Kolenbrander, P. E., Palmer, R. J., Periasamy, S. and Jakubovics, N. S. (2010) 'Oral multispecies biofilm development and the key role of cell–cell distance', *Nat Rev Micro*, 8, (7), pp. 471-480.
2. Liu, Y. and Burne, R.A. (2009) 'Multiple Two-Component Systems Modulate Alkali Generation in Streptococcus gordonii in Response to Environmental Stresses', *Journal of Bacteriology*, 191(23), pp. 7353-7362.

Centre for Oral Health Research (COHR) research afternoon, Newcastle, 8th May 2013 (Poster)

Title: Mechanisms of biofilm formation by *Streptococcus gordonii*

Mechanisms of biofilm formation by *Streptococcus gordonii*

J.C. Robinson[1], N. Jakubovics[1], M. German[2], W. Vollmer[3]

1. Oral Biology, School of Dental Sciences, Newcastle University
2. Dental Materials, School of Dental Sciences, Newcastle University
3. Centre for Bacterial Cell Biology, School of Biomedical Sciences, Newcastle University

Introduction

Streptococcus gordonii is a primary coloniser of tooth surfaces and plays a key role in the establishment of dental biofilms[1]. A deletion strain of *S. gordonii*, lacking the gene *arcR*, was previously shown to be defective in biofilm formation. The *arcR* gene encodes the ArcR regulatory protein of the arginine deiminase system (ADS), which is involved in maintaining biofilm pH homeostasis, and protecting against lethal acidification in the oral environment by the conversion of arginine to ammonia and ATP[2].

By microarray analysis, ArcR was found to control expression of the gene SGO_0846, encoding a putative cell surface protein[3] of unknown function. The aim of this project was to investigate the effect of *arcR* on SGO_0846 expression and biofilm formation.

Methods

- Biofilm competition assays: Equal concentrations of the wild-type (WT) and *arcR* deletion strains were grown together in co-culture biofilms, to assess the relative fitness of each strain.
- qPCR: Using cDNA that had been extracted previously, looking to quantify SGO_0846 expression levels. Samples consisted of cDNA extracted from both WT and *arcR* deletion strains, in the presence and absence of arginine (to allow for ArcR effects). Primers specific for 16S ribosomal cDNA and the SGO_0846 gene were used.
- SGO_0846 mutagenesis: Insertion of an *ermAM* erythromycin-resistance gene into the middle of the SGO_0846 gene, in order to knock out the gene and observe the effects of deletion, specifically upon the biofilm phenotype.

Results

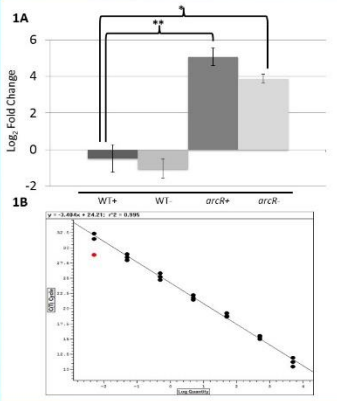


Figure 1. A) Graph of quantitative PCR (qPCR) data, showing the relative fold changes of SGO_0846 expression in the wild-type (WT) strains in the presence and absence of arginine, in comparison to the *arcR* deletion strain in the same conditions. Error bars show the standard error of mean (SEM). * Paired T-test p = <0.05, ** Paired T-test p = <0.01. B) Graph showing standard curve for qPCR, with standard curve consisting of 10-fold dilutions of chromosomal DNA. Red point has been discounted from the standard curve.

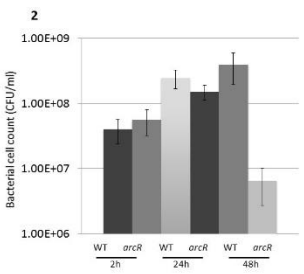


Figure 2. Graph to show the results of the biofilm competition assay, showing WT and *arcR* deletion strain co-culture biofilm cell counts over three different timepoints – 2h, 24h and 48h. Results are from four replicates, error bars show the SEM.

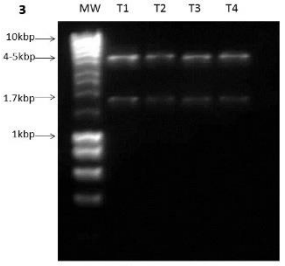


Figure 3. 1% agarose gel showing *EcoR1* restriction digest fragments of plasmids from four transformant colonies (T1-4) of *E. coli*. Recombinant plasmid contains SGO_0846 gene, which forms the fragment at 1.7kpb; remainder of plasmid is at 4-5kpb. MW = 0.2-10kpb.

Discussion

The qPCR data showed that WT expression levels of SGO_0846 were significantly different to both *arcR* in the presence (p=0.007) and absence (p=0.012) of arginine, corroborating the results of the gene microarray carried out previously (Figure 1).

The biofilm competition assay data appears to show that the *arcR* deletion strain growth within a co-culture biofilm drops off after 48h (Figure 2), a result which is also seen in the planktonic counts for the biofilms (data not shown here). This suggests that the *arcR* deletion strain may be impaired in survival after 24h, which may indicate that the *arcR* deletion strain is being out-competed by the WT strain. This would show that the WT strain is not compensating the *arcR* deletion strain in any way which would allow equal growth.

Significant progress has also been made towards the deletion of SGO_0846, with the production of a recombinant plasmid containing the SGO_0846 gene (Figure 3), which has been successfully transformed into an *E. coli* strain before the addition of the erythromycin-resistance cassette.

Future work

- qPCR analysis of other genes of interest from the microarray experiment, which also showed differential expression within the *arcR* deletion strain in comparison to the WT strain.
- Tests on the growth of the *arcR* deletion strain over a 48h time period, due to the apparent inability of the strain to survive within a biofilm culture – this will demonstrate whether this observation is only found in the biofilm, or whether it is an inherent decreased survival rate in the *arcR* deletion strain.
- Insertion of the *ermAM* erythromycin resistance cassette into the SGO_0846 gene, and subsequent transformation of the gene knockout into *Streptococcus gordonii* through homologous recombination onto the bacterial chromosome, in both the WT and the *arcR* deletion strains.
- Observation of the effects of SGO_0846 deletion on the biofilm phenotype in the WT and *arcR* deletion strains, to investigate whether deletion of SGO_0846 allows restoration of normal biofilm phenotype within the *arcR* deletion strain; which would suggest a role for ArcR in the regulation of SGO_0846 expression, and a role for SGO_0846 in *S. gordonii* biofilm formation.

References

1. Kolenbrander, P.E., Palmer, R.J., Periasamy, S. and Jakubovics, N.S. (2010). 'Oral multispecies biofilm development and the key role of cell–cell distance', *Nat Rev Micro*, 8(7), pp. 471-480.
2. Liu, Y. and Burne, R.A. (2009). 'Multiple Two-Component Systems Modulate Alkali Generation in Streptococcus gordonii in Response to Environmental Stresses', *Journal of Bacteriology*, 191(23), pp. 7353-7362.
3. NCBI (2011-2012). *Gene*. Available at: http://www.ncbi.nlm.nih.gov/gene (Accessed: 18th June).

Eurobiofilms 2013, Ghent University - Belgium, 9th-12th September 2013 (Poster)

Title: Role of the regulatory protein ArcR in *Streptococcus gordonii* biofilm formation

Abstract: Streptococcus gordonii is a primary coloniser of the oral cavity and is key in the establishment of dental plaque. Previous studies in our laboratory have shown that the arcR gene, encoding a transcriptional regulator (ArcR), is essential for biofilm formation by S. gordonii. Further, by microarray analysis, ArcR was found to control the expression of a putative cell surface protein of unknown function, termed SGO_0846. The aim of this study was to characterise the effect of arcR on SGO_0846 expression and biofilm formation.

This was achieved through competition assays, to measure the relative fitness of an arcR deletion strain in comparison to the wild-type DL1 strain; quantitative PCR, used to confirm and quantify levels of expression for the SGO_0846 gene in this strain and in DL1; and mutagenesis of the SGO_0846 gene, to observe the effects of deletion of this gene on the biofilm phenotype.

Competition assays showed a difference in relative fitness between the wild-type and arcR deletion strains, with the arcR mutant being out-competed by DL1 over a prolonged period of time. Quantitative PCR confirmed the effect of ArcR on SGO_0846 expression levels, with SGO_0846 expression being higher in the arcR deletion strain. Significant progress has also been made towards mutagenesis of the SGO_0846 gene. In conclusion, these results demonstrate that ArcR strongly regulates the expression of the SGO_0846 gene, and furthermore the biofilm formation of Streptococcus gordonii.

The role of the regulatory protein ArcR in *Streptococcus gordonii* biofilm formation

Jill C. Robinson[1], Dr. Nicholas Jakubovics[1],
Dr. Matthew German[2], Prof. Waldemar Vollmer[3]
1. Oral Biology, Centre for Oral Health Research, Newcastle University
2. Dental Materials, Centre for Oral Health Research, Newcastle University,
3. Centre for Bacterial Cell Biology, Institute of Cell and Molecular Biosciences, Newcastle University

Introduction

Streptococcus gordonii, an oral commensal organism and primary coloniser of the surfaces within the oral cavity, is known to play an important role in the formation of dental plaque[1]. Preliminary studies have shown that the *arcR* gene, encoding the arginine-dependent transcriptional regulator ArcR, is required for stable biofilm formation in *S. gordonii.* Deletion of *arcR* results in a strain that forms biofilms that appear superficially similar to wild-type biofilms, but are extremely unstable and easily break apart.

Aims

- To quantify differences in biofilm formation by the *arcR* mutant
- Measure cell surface hydrophobicity in the *arcR* mutant
- Investigate whether *arcR* mutant can be complemented by co-culture with wild-type

Methods

- Atomic force microscopy (AFM) – air imaging and force spectroscopy: air imaging was carried out on air-dried dehydrated biofilms. Force spectroscopy was used to measure the stiffness of the biofilms by tapping and dragging a silicon probe across the cell surfaces.
- Crystal violet assay: biofilms were grown in 6-well plates, stained in crystal violet, and the absorbance measured at A_{570} in order to quantify biofilm growth.
- Microbial adhesion to hexadecane (MATH) assay: the addition of *n*-hexadecane to an aqueous suspension of cells allowed the measurement of cell surface hydrophobicity. The proportion of cells partitioning into the hexadecane phase ("% hydrophobicity") was determined by the reduction in A_{450} of the aqueous phase.
- Competition assays – planktonic and sessile: Equal concentrations of the wild-type *S. gordonii* and *arcR* mutant strains were grown together in planktonic and biofilm co-cultures, to assess the relative fitness of each strain.

Results

1

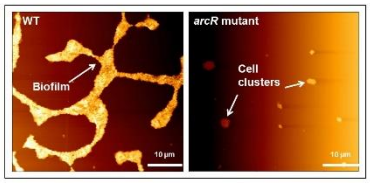


Figure 1. AFM air imaging of *S. gordonii* wild-type (WT) growth and *arcR* mutant strain growth. Following washing, there is little of the *arcR* mutant strain left on the plate.

2

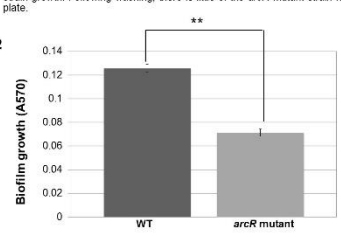


Figure 2. Quantitative analysis of the role of *arcR* in biofilm formation. Biofilms formed by *S. gordonii* wild-type (WT) and *arcR* mutant strains were assessed by crystal violet staining. (T-test: ** p <0.001).

3

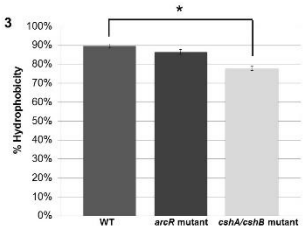
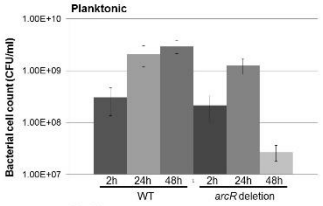


Figure 3. Cell surface hydrophobicity determined by MATH assay. A *cshA/cshB* deletion strain, known to have a lower cell surface hydrophobicity than wild-type, was used as control alongside the wild-type and *arcR* mutant strains. (T-test: * p <0.05).

4A



4B

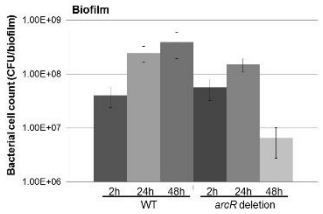


Figure 4. Relative fitness of the wild-type (WT) and *arcR* mutant strains as assessed by competition assay. Bacterial cell counts are shown for each strain over three timepoints – 2h, 24h and 48h. 4A) shows the planktonic cell counts taken from the media above the biofilm; 4B) shows the bacterial counts for the biofilms themselves.

Discussion and conclusions

- The *arcR* mutant strain is significantly impaired in biofilm formation.
- Both the crystal violet assays and the AFM imaging showed reduced biofilm formation in the *arcR* mutant strain, in comparison to the wild-type.
- The MATH assay showed no significant differences between the cell surface hydrophobicity of the *arcR* mutant strain and the wild-type.
- Competition assays – both planktonic and biofilm – showed an impairment in long-term survival (>48h) in the *arcR* mutant strain, in addition to the previously observed impairment in biofilm formation. This could be due to hydrogen peroxide build-up during growth – it may be that the *arcR* strain is more susceptible to hydrogen peroxide exposure than the wild-type.

Future work

- AFM imaging and force spectroscopy of hydrated biofilms under liquid conditions, as dehydrated biofilms were artificial in appearance.
- Tests to investigate mechanism by which *arcR* affects planktonic survival.
- Measure hydrogen peroxide levels and effects on *arcR* mutant survival.
- Gene expression array carried out on the wild-type and *arcR* mutant strains, under high and no arginine conditions, to observe which genes are affected in expression by the *arcR* gene or differing arginine conditions.

References

1. Kolenbrander, P.E., Palmer, R.J., Periasamy, S. and Jakubovics, N.S. (2010). 'Oral multispecies biofilm development and the key role of cell–cell distance', *Nat Rev Micro*, 8(7), pp. 471-480.

International Association for Dental Research General Session, Boston – Massachusetts, USA, 11th-14th March 2015 (Oral presentation)

Title: Impact of Amino Acid Depletion on Streptococcus gordonii Gene Expression

Abstract:

Objectives: Streptococcus gordonii is an oral commensal bacterium, and one of the early colonizers of tooth surfaces during the initial stages of dental plaque formation. Recent work within our group has shown that S. gordonii gene expression is dramatically altered in response to a rapid depletion of arginine, and that many genes associated with biofilm formation are regulated. However, it is not known whether this extensive response is specific to arginine, or whether it is part of a more generalized stress response. We hypothesize that arginine plays a unique role in the regulation of biofilm-associated genes, and this study aimed to assess S. gordonii gene regulation in response to depletion of different amino acids.

Methods: S. gordonii was cultured in chemically-defined medium (CDM). Cells were harvested, resuspended in CDM or CDM lacking arginine, histidine, or branched-chain amino acids (BCAA), and RNA was extracted after 30 min. Gene expression was analyzed by qRT-PCR.

Results: Shifting S. gordonii from amino acid-replete CDM to CDM lacking arginine or BCAA resulted in bacterial growth arrest, whereas exponential growth was maintained in CDM lacking histidine. There was some overlap observed between the gene regulation responses to depletion of specific amino acids, however, some of the genetic responses were unique to arginine. In particular, genes involved in the Hsa cell surface receptor synthesis and the cellobiose phosphotransferase system were specifically regulated in response to arginine depletion.

Conclusions: In conclusion, arginine has a quantifiable effect on the regulation and expression of a number of genes within the S. gordonii genome, some of which are directly involved in the process of biofilm formation. Future work aims to determine the precise roles of these genes in the formation and maintenance of Streptococcus gordonii biofilms.

Public Engagement

British Science Festival 2013, Newcastle University – What's Hiding in Your Mouth?

Helped to organise and run a free event at the British Science Festival, in September 2013. This event was aimed at children, to educate them about oral hygiene and the microbiology of the mouth.

Soapbox Science, Newcastle, June 2015

Volunteered as a helper at the Soapbox Science event, held in Newcastle city centre and aimed at promoting awareness of women in science, and current scientific research.

References

Aas, J. A., Paster, B. J., Stokes, L. N., Olsen, I. and Dewhirst, F. E. (2005) 'Defining the normal bacterial flora of the oral cavity', *Journal of Clinical Microbiology*, 43(11), pp. 5721-32.

Ahn, J., Yang, L., Paster, B. J., Ganly, I., Morris, L., Pei, Z. and Hayes, R. B. (2011) 'Oral Microbiome Profiles: 16S rRNA Pyrosequencing and Microarray Assay Comparison', *PLoS ONE*, 6(7), p. e22788.

Allison, D. G. (2000) *Community structure and co-operation in biofilms*. Cambridge University Press.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) 'Gapped BLAST and PSI-BLAST: a new generation of protein database search programs', *Nucleic Acids Research*, 25(17), pp. 3389-402.

Altschul, S. F., Wootton, J. C., Gertz, E. M., Agarwala, R., Morgulis, A., Schaffer, A. A. and Yu, Y. K. (2005) 'Protein database searches using compositionally adjusted substitution matrices', *FEBS Journal*, 272(20), pp. 5101-9.

Anderl, J. N., Franklin, M. J. and Stewart, P. S. (2000) 'Role of Antibiotic Penetration Limitation in *Klebsiella pneumoniae* Biofilm Resistance to Ampicillin and Ciprofloxacin', *Antimicrobial Agents and Chemotherapy*, 44(7), pp. 1818-1824.

Angst, E. (1923) 'The fouling of ships bottoms by bacteria', *Report, Bureau Construction and Repair. United States Navy Department, Washington, DC*.

Arakawa, T., Ejima, D., Tsumoto, K., Obeyama, N., Tanaka, Y., Kita, Y. and Timasheff, S. N. (2007) 'Suppression of protein interactions by arginine: A proposed mechanism of the arginine effects', *Biophysical Chemistry*, 127(1–2), pp. 1-8.

Atlas, R. M. and Bartha, R. (1997) *Microbial Ecology*. 4th edition edn. Menlo Park, CA, USA: Benjamin/Cummings Science Publishing.

Babu, J. P. and Dabbous, M. K. (1986) 'Interaction of Salivary Fibronectin with Oral Streptococci', *Journal of Dental Research*, 65(8), pp. 1094-1100.

Back, C. R., Douglas, S. K., Emerson, J. E., Nobbs, A. H. and Jenkinson, H. F. (2015) '*Streptococcus gordonii* DL1 adhesin SspB V-region mediates coaggregation via receptor polysaccharide of *Actinomyces oris* T14V', *Molecular Oral Microbiology*, 30(5), pp. 411-424.

Barcelona-Andrés, B., Marina, A. and Rubio, V. (2002) 'Gene Structure, Organization, Expression, and Potential Regulatory Mechanisms of Arginine Catabolism in *Enterococcus faecalis*', *Journal of Bacteriology*, 184(22), pp. 6289-6300.

Barnard, J. P. and Stinson, M. W. (1999) 'Influence of Environmental Conditions on Hydrogen Peroxide Formation by *Streptococcus gordonii*', *Infection and Immunity*, 67(12), pp. 6558-6564.

Bayles, K. W. (2007) 'The biological role of death and lysis in biofilm development', *Nature Reviews Microbiology*, 5(9), pp. 721-726.

Beilharz, K., van Raaphorst, R., Kjos, M. and Veening, J.-W. (2015) 'Red fluorescent proteins for gene expression and protein localization studies in *Streptococcus pneumoniae* and efficient transformation with Gibson assembled DNA', *Applied and Environmental Microbiology*, 81(20), pp. 7244-52.

Bensing, B. A., López, J. A. and Sullam, P. M. (2004) 'The *Streptococcus gordonii* Surface Proteins GspB and Hsa Mediate Binding to Sialylated Carbohydrate Epitopes on the Platelet Membrane Glycoprotein Ibα', *Infection and Immunity*, 72(11), pp. 6528-6537.

Berger, U. (1985) 'Prevalence of *Gemella haemolysans* on the pharyngeal mucosa of man', *Medical Microbiology and Immunology*, 174(5), pp. 267-74.

Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A. and Duncan, K. (2002) 'Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling', *Molecular Microbiology*, 43(3), pp. 717-731.

Bjarnsholt, T. (2011) 'Introduction to Biofilms', in Bjarnsholt, T., Jensen, P. Ø., Moser, C. and Høiby, N. (eds.) *Biofilm Infections*. Springer New York, pp. 1-9.

Booth, I. R. (1985) 'Regulation of cytoplasmic pH in bacteria', *Microbiological Reviews*, 49(4), pp. 359-378.

Brand, H. S., Jörning, G. G. A., Chamuleau, R. A. F. M. and Abraham-Inpijn, L. (1997) 'Effect of a protein-rich meal on urinary and salivary free amino acid concentrations in human subjects', *Clinica Chimica Acta*, 264(1), pp. 37-47.

Breeuwer, P., Drocourt, J., Rombouts, F. M. and Abee, T. (1996) 'A Novel Method for Continuous Determination of the Intracellular pH in Bacteria with the Internally Conjugated Fluorescent Probe 5 (and 6-)-Carboxyfluorescein Succinimidyl Ester', *Applied and Environmental Microbiology*, 62(1), pp. 178-83.

Breitling, R., Armengaud, P., Amtmann, A. and Herzyk, P. (2004) 'Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments', *FEBS Letters*, 573(1–3), pp. 83-92.

Broman, K., Stalon, V. and Wiame, J. M. (1975) 'The duplication of arginine catabolism and the meaning of the two ornithine carbamoyltransferases in *Bacillus licheniformis*', *Biochemical and Biophysical Research Communications*, 66(2), pp. 821-827.

Brooun, A., Liu, S. and Lewis, K. (2000) 'A Dose-Response Study of Antibiotic Resistance in *Pseudomonas aeruginosa* Biofilms', *Antimicrobial Agents and Chemotherapy*, 44(3), pp. 640-646.

Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K. and Jones, D. T. (2013) 'Scalable web services for the PSIPRED Protein Analysis Workbench', *Nucleic Acids Research*, 41(Web Server issue), pp. W349-57.

Burne, R. A. and Marquis, R. E. (2000) 'Alkali production by oral bacteria and protection against dental caries', *FEMS Microbiology Letters*, 193(1), pp. 1-6.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J. and Wittwer, C. T. (2009) 'The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments', *Clinical Chemistry*, 55(4), pp. 611-622.

Chávez de Paz, L. E., Hamilton, I. R. and Svensäter, G. (2008) 'Oral bacteria in biofilms exhibit slow reactivation from nutrient deprivation', *Microbiology*, 154(7), pp. 1927-1938.

Chiang, W.-C., Nilsson, M., Jensen, P. Ø., Høiby, N., Nielsen, T. E., Givskov, M. and Tolker-Nielsen, T. (2013) 'Extracellular DNA Shields against Aminoglycosides in *Pseudomonas aeruginosa* Biofilms', *Antimicrobial Agents and Chemotherapy*, 57(5), pp. 2352-2361.

Chikwamba, R. K., Scott, M. P., Mejía, L. B., Mason, H. S. and Wang, K. (2003) 'Localization of a bacterial protein in starch granules of transgenic maize kernels', *Proceedings of the National Academy of Sciences*, 100(19), pp. 11127-11132.

Christensen, G. D., Simpson, W. A., Younger, J. J., Baddour, L. M., Barrett, F. F., Melton, D. M. and Beachey, E. H. (1985) 'Adherence of coagulase-negative staphylococci to plastic tissue culture plates: a quantitative model for the adherence of staphylococci to medical devices', *Journal of Clinical Microbiology*, 22(6), pp. 996-1006.

Christopher, A. B., Arndt, A., Cugini, C. and Davey, M. E. (2010) 'A streptococcal effector protein that inhibits *Porphyromonas gingivalis* biofilm development', *Microbiology*, 156(11), pp. 3469-3477.

Ciampi, M. S. and Roth, J. R. (1988) 'Polarity effects in the *hisG* gene of *Salmonella* require a site within the coding sequence', *Genetics*, 118(2), pp. 193-202.

Cisar, J. O., Sandberg, A. L., Abeygunawardana, C., Reddy, G. P. and Bush, C. A. (1995) 'Lectin recognition of host-like saccharide motifs in streptococcal cell wall polysaccharides', *Glycobiology*, 5(7), pp. 655-662.

Cole, J. N., Ramirez, R. D., Currie, B. J., Cordwell, S. J., Djordjevic, S. P. and Walker, M. J. (2005) 'Surface Analyses and Immune Reactivities of Major Cell Wall-Associated Proteins of Group A *Streptococcus*', *Infection and Immunity*, 73(5), pp. 3137-3146.

Costerton, J. W., Geesey, G. G. and Cheng, K. J. (1978) 'How bacteria stick', *Scientific American*, 238(1), pp. 86-95.

Costerton, J. W., Stewart, P. S. and Greenberg, E. P. (1999) 'Bacterial Biofilms: A Common Cause of Persistent Infections', *Science*, 284(5418), pp. 1318-1322.

Cowman, R. A., Perrella, M. M., Adams, B. O. and Fitzgerald, R. J. (1975) 'Amino acid requirements and proteolytic activity of *Streptococcus sanguis*', *Applied Microbiology*, 30(3), pp. 374-80.

Cramton, S. E., Gerke, C., Schnell, N. F., Nichols, W. W. and Götz, F. (1999) 'The Intercellular Adhesion (*ica*) Locus Is Present in *Staphylococcus aureus* and Is Required for Biofilm Formation', *Infection and Immunity*, 67(10), pp. 5427-5433.

Crooks, G. E., Hon, G., Chandonia, J. M. and Brenner, S. E. (2004) 'WebLogo: a sequence logo generator', *Genome Research*, 14(6), pp. 1188-90.

Cunin, R., Glansdorff, N., Piérard, A. and Stalon, V. (1986) 'Biosynthesis and metabolism of arginine in bacteria', *Microbiological Reviews*, 50(3), pp. 314-352.

Cusumano, Z. T., Watson, M. E. and Caparon, M. G. (2014) '*Streptococcus pyogenes* Arginine and Citrulline Catabolism Promotes Infection and Modulates Innate Immunity', *Infection and Immunity*, 82(1), pp. 233-242.

Cvitkovitch, D. G., Li, Y.-H. and Ellen, R. P. (2003) 'Quorum sensing and biofilm formation in Streptococcal infections', *Journal of Clinical Investigation*, 112(11), pp. 1626-1632.

D'Amico, S., Collins, T., Marx, J.-C., Feller, G. and Gerday, C. (2006) 'Psychrophilic microorganisms: challenges for life', *EMBO Reports*, 7(4), pp. 385-389.

Daep, C. A., Novak, E. A., Lamont, R. J. and Demuth, D. R. (2011) 'Structural Dissection and In Vivo Effectiveness of a Peptide Inhibitor of *Porphyromonas gingivalis* Adherence to *Streptococcus gordonii*', *Infection and Immunity*, 79(1), pp. 67-74.

Dalton, T., Dowd, S. E., Wolcott, R. D., Sun, Y., Watters, C., Griswold, J. A. and Rumbaugh, K. P. (2011) 'An *In Vivo* Polymicrobial Biofilm Wound Infection Model to Study Interspecies Interactions', *PLoS ONE*, 6(11), p. e27317.

Dam, P., Olman, V., Harris, K., Su, Z. and Xu, Y. (2007) 'Operon prediction using both genome-specific and general genomic information', *Nucleic Acids Research*, 35(1), pp. 288-98.

Das, T., Sharma, P. K., Busscher, H. J., van der Mei, H. C. and Krom, B. P. (2010) 'Role of Extracellular DNA in Initial Bacterial Adhesion and Surface Aggregation', *Applied and Environmental Microbiology*, 76(10), pp. 3405-3408.

de Jong, A., Pietersma, H., Cordes, M., Kuipers, O. and Kok, J. (2012) 'PePPER: a webserver for prediction of prokaryote promoter elements and regulons', *BMC Genomics*, 13(1), pp. 1-10.

Degnan, B. A., Fontaine, M. C., Doebereiner, A. H., Lee, J. J., Mastroeni, P., Dougan, G., Goodacre, J. A. and Kehoe, M. A. (2000) 'Characterization of an Isogenic Mutant of *Streptococcus pyogenes* Manfredo Lacking the Ability To Make Streptococcal Acid Glycoprotein', *Infection and Immunity*, 68(5), pp. 2441-2448.

Degnan, B. A., Palmer, J. M., Robson, T., Jones, C. E. D., Fischer, M., Glanville, M., Mellor, G. D., Diamond, A. G., Kehoe, M. A. and Goodacre, J. A. (1998) 'Inhibition of Human

Peripheral Blood Mononuclear Cell Proliferation by *Streptococcus pyogenes* Cell Extract Is Associated with Arginine Deiminase Activity', *Infection and Immunity*, 66(7), pp. 3050-3058.

Devroede, N., Thia-Toong, T. L., Gigot, D., Maes, D. and Charlier, D. (2004) 'Purine and pyrimidine-specific repression of the *Escherichia coli carAB* operon are functionally and structurally coupled', *Journal of Molecular Biology*, 336(1), pp. 25-42.

Dewhirst, F. E., Chen, T., Izard, J., Paster, B. J., Tanner, A. C. R., Yu, W.-H., Lakshmanan, A. and Wade, W. G. (2010) 'The Human Oral Microbiome', *Journal of Bacteriology*, 192(19), pp. 5002-5017.

Diep, B. A., Gill, S. R., Chang, R. F., Phan, T. H., Chen, J. H., Davidson, M. G., Lin, F., Lin, J., Carleton, H. A., Mongodin, E. F., Sensabaugh, G. F. and Perdreau-Remington, F. (2006) 'Complete genome sequence of USA300, an epidemic clone of community-acquired methicillin-resistant *Staphylococcus aureus*', *Lancet*, 367(9512), pp. 731-9.

Dineshkumar, T. K., Thanedar, S., Subbulakshmi, C. and Varshney, U. (2002) 'An unexpected absence of queuosine modification in the tRNAs of an *Escherichia coli* B strain', *Microbiology*, 148(12), pp. 3779-3787.

Doern, C. D. and Burnham, C.-A. D. (2010) 'It's Not Easy Being Green: the Viridans Group Streptococci, with a Focus on Pediatric Clinical Manifestations', *Journal of Clinical Microbiology*, 48(11), pp. 3829-3835.

Dogan, J., Gianni, S. and Jemth, P. (2014) 'The binding mechanisms of intrinsically disordered proteins', *Physical Chemistry Chemical Physics*, 16(14), pp. 6323-6331.

Dong, Y., Chen, Y.-Y. M. and Burne, R. A. (2004) 'Control of Expression of the Arginine Deiminase Operon of *Streptococcus gordonii* by CcpA and Flp', *Journal of Bacteriology*, 186(8), pp. 2511-2514.

Dong, Y., Chen, Y. Y., Snyder, J. A. and Burne, R. A. (2002) 'Isolation and molecular analysis of the gene cluster for the arginine deiminase system from *Streptococcus gordonii* DL1', *Applied Environmental Microbiology*, 68(11), pp. 5549-53.

Donlan, R. M. (2001) 'Biofilms and device-associated infections', *Emerging Infectious Diseases*, 7(2), pp. 277-281.

Donlan, R. M. (2002) 'Biofilms: Microbial Life on Surfaces', *Emerging Infectious Diseases*, 8(9), pp. 881-890.

Dueholm, M. S., Petersen, S. V., Sonderkaer, M., Larsen, P., Christiansen, G., Hein, K. L., Enghild, J. J., Nielsen, J. L., Nielsen, K. L., Nielsen, P. H. and Otzen, D. E. (2010) 'Functional amyloid in *Pseudomonas*', *Molecular Microbiology*, 77(4), pp. 1009-20.

Durand, J. M., Okada, N., Tobe, T., Watarai, M., Fukuda, I., Suzuki, T., Nakata, N., Komatsu, K., Yoshikawa, M. and Sasakawa, C. (1994) '*vacC*, a virulence-associated chromosomal locus of *Shigella flexneri,* is homologous to *tgt*, a gene encoding tRNA-guanine transglycosylase (Tgt) of *Escherichia coli* K-12', *Journal of Bacteriology*, 176(15), pp. 4627-4634.

Edwards, A. M., Grossman, T. J. and Rudney, J. D. (2007) 'Association of a high-molecular weight arginine-binding protein of *Fusobacterium nucleatum* ATCC 10953 with adhesion to secretory immunoglobulin A and coaggregation with *Streptococcus cristatus*', *Oral Microbiology and Immunology*, 22(4), pp. 217-24.

Efron, D. T., Kirk, S. J., Regan, M. C., Wasserkrug, H. L. and Barbul, A. (1991) 'Nitric oxide generation from L-arginine is required for optimal human peripheral blood lymphocyte DNA synthesis', *Surgery*, 110(2), pp. 327-34.

Egland, P. G., Palmer, R. J. and Kolenbrander, P. E. (2004) 'Interspecies communication in *Streptococcus gordonii–Veillonella atypica* biofilms: Signaling in flow conditions requires juxtaposition', *Proceedings of the National Academy of Sciences of the United States of America*, 101(48), pp. 16917-16922.

Eren, A. M., Borisy, G. G., Huse, S. M. and Mark Welch, J. L. (2014) 'Oligotyping analysis of the human oral microbiome', *Proceedings of the National Academy of Sciences*, 111(28), pp. E2875-E2884.

Etienne, W., Meyer, M. H., Peppers, J. and Meyer, R. A., Jr. (2004) 'Comparison of mRNA gene expression by RT-PCR and DNA microarray', *BioTechniques*, 36(4), pp. 618-20, 622, 624-6.

Facklam, R. (2002) 'What happened to the streptococci: overview of taxonomic and nomenclature changes', *Clinical Microbiology Reviews*, 15(4), pp. 613-30.

Farr, S. B. and Kogoma, T. (1991) 'Oxidative stress responses in *Escherichia coli* and *Salmonella typhimurium*', *Microbiological Reviews*, 55(4), pp. 561-585.

Featherstone, J. D. B. (2000) 'The science and practice of caries prevention', *Journal of the American Dental Association*, 131(7), pp. 887-99.

Featherstone, J. D. B. (2008) 'Dental caries: a dynamic disease process', *Australian Dental Journal*, 53(3), pp. 286-291.

Ferro, K., Bender, G. and Marquis, R. (1983) 'Coordinately repressible arginine deiminase system in *Streptococcus sanguis*', *Current Microbiology*, 9(3), pp. 145-149.

Fey, P. D. and Olson, M. E. (2010) 'Current concepts in biofilm formation of *Staphylococcus epidermidis*', *Future Microbiology*, 5(6), pp. 917-933.

Finkel, S. E. and Kolter, R. (2001) 'DNA as a Nutrient: Novel Role for Bacterial Competence Gene Homologs', *Journal of Bacteriology*, 183(21), pp. 6288-6293.

Flemming, H. C. and Wingender, J. (2010) 'The Biofilm Matrix', *Nature Reviews Microbiology*, 8(9), pp. 623-633.

Frank, K. L., Guiton, P. S., Barnes, A. M. T., Manias, D. A., Chuang-Smith, O. N., Kohler, P. L., Spaulding, A. R., Hultgren, S. J., Schlievert, P. M. and Dunny, G. M. (2013) 'AhrC and Eep Are Biofilm Infection-Associated Virulence Factors in *Enterococcus faecalis*', *Infection and Immunity*, 81(5), pp. 1696-1708.

Franklin, N. C. and Luria, S. E. (1961) 'Transduction by bacteriophage P-1 and the properties of the *lac* genetic region in *E. coli* and *S. dysenteriae*', *Virology*, 15, pp. 299-311.

Frillingos, S., Sahin-Toth, M., Wu, J. and Kaback, H. R. (1998) 'Cys-scanning mutagenesis: a novel approach to structure function relationships in polytopic membrane proteins', *FASEB Journal*, 12(13), pp. 1281-99.

Gardan, R., Rapoport, G. and Débarbouillé, M. (1997) 'Role of the transcriptional activator RocR in the arginine-degradation pathway of *Bacillus subtilis*', *Molecular Microbiology*, 24(4), pp. 825-837.

Garrett, T. R., Bhakoo, M. and Zhang, Z. (2008) 'Bacterial adhesion and biofilms on surfaces', *Progress in Natural Science*, 18(9), pp. 1049-1056.

Geesey, G. G., Richardson, W. T., Yeomans, H. G., Irvin, R. T. and Costerton, J. W. (1977) 'Microscopic examination of natural sessile bacterial populations from an alpine stream', *Canadian Journal of Microbiology*, 23(12), pp. 1733-1736.

Girard, F., Vachon, V., Préfontaine, G., Marceau, L., Su, Y., Larouche, G., Vincent, C., Schwartz, J.-L., Masson, L. and Laprade, R. (2008) 'Cysteine Scanning Mutagenesis of α4, a Putative Pore-Lining Helix of the *Bacillus thuringiensis* Insecticidal Toxin Cry1Aa', *Applied and Environmental Microbiology*, 74(9), pp. 2565-2572.

GSL Biotech (2015) *SnapGene® software* [Computer program]. Available at: http://www.snapgene.com/.

Hajishengallis, G. and Lamont, R. J. (2012) 'Beyond the red complex and into more complexity: the polymicrobial synergy and dysbiosis (PSD) model of periodontal disease etiology', *Molecular Oral Microbiology*, 27(6), pp. 409-419.

Hajishengallis, G., Liang, S., Payne, Mark A., Hashim, A., Jotwani, R., Eskan, Mehmet A., McIntosh, Megan L., Alsam, A., Kirkwood, Keith L., Lambris, John D., Darveau, Richard P. and Curtis, Michael A. (2011) 'Low-Abundance Biofilm Species Orchestrates Inflammatory Periodontal Disease through the Commensal Microbiota and Complement', *Cell Host & Microbe*, 10(5), pp. 497-506.

Hall-Stoodley, L. and Stoodley, P. (2009) 'Evolving concepts in biofilm infections', *Cellular Microbiology*, 11(7), pp. 1034-1043.

Handley, P. S., Correia, F. F., Russell, K., Rosan, B. and DiRienzo, J. M. (2005) 'Association of a novel high molecular weight, serine-rich protein (SrpA) with fibril-mediated adhesion of the oral biofilm bacterium *Streptococcus cristatus*', *Oral Microbiology and Immunology*, 20(3), pp. 131-140.

Hansen, M. C., Palmer, R. J., Jr., Udsen, C., White, D. C. and Molin, S. (2001) 'Assessment of GFP fluorescence in cells of *Streptococcus gordonii* under conditions of low pH and low oxygen concentration', *Microbiology*, 147(5), pp. 1383-91.

Härtel, T., Eylert, E., Schulz, C., Petruschka, L., Gierok, P., Grubmüller, S., Lalk, M., Eisenreich, W. and Hammerschmidt, S. (2012) 'Characterization of Central Carbon

Metabolism of *Streptococcus pneumoniae* by Isotopologue Profiling', *Journal of Biological Chemistry*, 287(6), pp. 4260-4274.

He, X., Hu, W., Kaplan, C., Guo, L., Shi, W. and Lux, R. (2012) 'Adherence to Streptococci Facilitates *Fusobacterium nucleatum* Integration into an Oral Microbial Community', *Microbial Ecology*, 63(3), pp. 532-542.

Hermansson, M. (1999) 'The DLVO theory in microbial adhesion', *Colloids and Surfaces B: Biointerfaces*, 14(1–4), pp. 105-119.

Hillman, J. D., Socransky, S. S. and Shivers, M. (1985) 'The relationships between streptococcal species and periodontopathic bacteria in human dental plaque', *Archives of Oral Biology*, 30(11-12), pp. 791-5.

Hinsa, S. M., Espinosa-Urgel, M., Ramos, J. L. and O'Toole, G. A. (2003) 'Transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas fluorescens* WCS365 requires an ABC transporter and a large secreted protein', *Molecular Microbiology*, 49(4), pp. 905-918.

Hoskins, J., Alborn, W. E., Arnold, J., Blaszczak, L. C., Burgett, S., DeHoff, B. S., Estrem, S. T., Fritz, L., Fu, D.-J., Fuller, W., Geringer, C., Gilmour, R., Glass, J. S., Khoja, H., Kraft, A. R., Lagace, R. E., LeBlanc, D. J., Lee, L. N., Lefkowitz, E. J., Lu, J., Matsushima, P., McAhren, S. M., McHenney, M., McLeaster, K., Mundy, C. W., Nicas, T. I., Norris, F. H., O'Gara, M., Peery, R. B., Robertson, G. T., Rockey, P., Sun, P.-M., Winkler, M. E., Yang, Y., Young-Bellido, M., Zhao, G., Zook, C. A., Baltz, R. H., Jaskunas, S. R., Rosteck, P. R., Skatrud, P. L. and Glass, J. I. (2001) 'Genome of the Bacterium *Streptococcus pneumoniae* Strain R6', *Journal of Bacteriology*, 183(19), pp. 5709-5717.

Hunt, S. M., Werner, E. M., Huang, B., Hamilton, M. A. and Stewart, P. S. (2004) 'Hypothesis for the Role of Nutrient Starvation in Biofilm Detachment', *Applied and Environmental Microbiology*, 70(12), pp. 7418-7425.

Irwin, C., Farmer, A., Willer, D. and Evans, D. (2012) 'In-Fusion® Cloning with Vaccinia Virus DNA Polymerase', in Isaacs, S. N. (ed.) *Vaccinia Virus and Poxvirology*. Humana Press, pp. 23-35.

Ishibashi, M., Tsumoto, K., Tokunaga, M., Ejima, D., Kita, Y. and Arakawa, T. (2005) 'Is arginine a protein-denaturant?', *Protein Expression and Purification*, 42(1), pp. 1-6.

Jakubovics, N. (2013) 'Extracellular Proteins and DNA in the Matrix of Oral Biofilms', in Jakubovics, N. and Palmer, R. J. (eds.) *Oral Microbial Ecology: Current Research and New Perspectives*. Caister Academic Press.

Jakubovics, N., Brittan, J. L., Dutton, L. C. and Jenkinson, H. F. (2009) 'Multiple adhesin proteins on the cell surface of *Streptococcus gordonii* are involved in adhesion to human fibronectin', *Microbiology*, 155(11), pp. 3572-3580.

Jakubovics, N., Gill, S. R., Iobst, S. E., Vickerman, M. M. and Kolenbrander, P. E. (2008) 'Regulation of Gene Expression in a Mixed-Genus Community: Stabilized Arginine Biosynthesis in *Streptococcus gordonii* by Coaggregation with *Actinomyces naeslundii*', *Journal of Bacteriology*, 190(10), pp. 3646-3657.

Jakubovics, N., Robinson, J. C., Samarian, D. S., Kolderman, E., Yassin, S. A., Bettampadi, D., Bashton, M. and Rickard, A. H. (2015) 'Critical roles of arginine in growth and biofilm development by *Streptococcus gordonii*', *Molecular Microbiology*, 97(2), pp. 281-300.

Jakubovics, N., Strömberg, N., Van Dolleweerd, C. J., Kelly, C. G. and Jenkinson, H. F. (2005) 'Differential binding specificities of oral streptococcal antigen I/II family adhesins for human or bacterial ligands', *Molecular Microbiology*, 55(5), pp. 1591-1605.

Jones, D. T. (2007) 'Improving the accuracy of transmembrane protein topology prediction using evolutionary information', *Bioinformatics*, 23(5), pp. 538-44.

Jones, D. T., Taylor, W. R. and Thornton, J. M. (1994) 'A model recognition approach to the prediction of all-helical membrane protein structure and topology', *Biochemistry*, 33(10), pp. 3038-49.

Kanehisa, M. and Goto, S. (2000) 'KEGG: Kyoto Encyclopedia of Genes and Genomes', *Nucleic Acids Research*, 28(1), pp. 27-30.

Kaplan, C., Lux, R., Haake, S. K. and Shi, W. (2009) 'The *Fusobacterium nucleatum* outer membrane protein RadD is an arginine-inhibitable adhesin required for inter-species adherence and the structured architecture of multispecies biofilm', *Molecular Microbiology*, 71(1), pp. 35-47.

Kaplan, C., Ma, X., Paranjpe, A., Jewett, A., Lux, R., Kinder-Haake, S. and Shi, W. (2010) '*Fusobacterium nucleatum* outer membrane proteins Fap2 and RadD induce cell death in human lymphocytes', *Infection and Immunity*, 78(11), pp. 4773-8.

Kaplan, J., Ragunath, C., Ramasubbu, N. and Fine, D. H. (2003) 'Detachment of *Actinobacillus actinomycetemcomitans* Biofilm Cells by an Endogenous β-Hexosaminidase Activity', *Journal of Bacteriology*, 185(16), pp. 4693-4698.

Keijser, B. J. F., Zaura, E., Huse, S. M., van der Vossen, J. M. B. M., Schuren, F. H. J., Montijn, R. C., ten Cate, J. M. and Crielaard, W. (2008) 'Pyrosequencing analysis of the Oral Microflora of healthy adults', *Journal of Dental Research*, 87(11), pp. 1016-1020.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. and Sternberg, M. J. E. (2015) 'The Phyre2 web portal for protein modeling, prediction and analysis', *Nature Protocols*, 10(6), pp. 845-858.

Kilian, M., MIKKELSEN, L. and HENRICHSEN, J. (1989) 'Taxonomic study of viridans streptococci: description of *Streptococcus gordonii* sp. nov. and emended descriptions of *Streptococcus sanguis* (White and Niven 1946), *Streptococcus oralis* (Bridge and Sneath 1982), and *Streptococcus mitis* (Andrewes and Horder 1906)', *International Journal of Systematic Bacteriology*, 39(4), pp. 471-484.

Kiliç, A. O., Tao, L., Zhang, Y., Lei, Y., Khammanivong, A. and Herzberg, M. C. (2004) 'Involvement of *Streptococcus gordonii* Beta-Glucoside Metabolism Systems in Adhesion, Biofilm Formation, and *In Vivo* Gene Expression', *Journal of Bacteriology*, 186(13), pp. 4246-4253.

Kinder, S. A. and Holt, S. C. (1993) 'Localization of the *Fusobacterium nucleatum* T18 adhesin activity mediating coaggregation with *Porphyromonas gingivalis* T22', *Journal of Bacteriology*, 175(3), pp. 840-50.

Kingsford, C. L., Ayanbule, K. and Salzberg, S. L. (2007) 'Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake', *Genome Biology*, 8(2), p. R22.

Kloosterman, T. G. and Kuipers, O. P. (2011) 'Regulation of Arginine Acquisition and Virulence Gene Expression in the Human Pathogen *Streptococcus pneumoniae* by Transcription Regulators ArgR1 and AhrC', *Journal of Biological Chemistry*, 286(52), pp. 44594-44605.

Kluyver, A. J. and van Niel, C. B. (1957) *The Microbe's Contribution to Biology* Cambridge, Massachusetts: Harvard University Press.

Kolderman, E., Bettampadi, D., Samarian, D., Dowd, S. E., Foxman, B., Jakubovics, N. S. and Rickard, A. H. (2015) 'L-Arginine Destabilizes Oral Multi-Species Biofilm Communities Developed in Human Saliva', *PLoS ONE*, 10(5), p. e0121835.

Kolenbrander, P. E., Andersen, R. N. and Moore, L. V. (1989) 'Coaggregation of *Fusobacterium nucleatum, Selenomonas flueggei, Selenomonas infelix, Selenomonas noxia*, and *Selenomonas sputigena* with strains from 11 genera of oral bacteria', *Infection and Immunity*, 57(10), pp. 3194-203.

Kolenbrander, P. E. and London, J. (1993) 'Adhere today, here tomorrow: oral bacterial adherence', *Journal of Bacteriology*, 175(11), pp. 3247-3252.

Kolenbrander, P. E., Palmer, R. J., Periasamy, S. and Jakubovics, N. S. (2010) 'Oral multispecies biofilm development and the key role of cell–cell distance', *Nature Reviews Microbiology*, 8(7), pp. 471-480.

Kreth, J., Merritt, J. and Qi, F. (2009a) 'Bacterial and Host Interactions of Oral Streptococci', *DNA and Cell Biology*, 28(8), pp. 397-403.

Kreth, J., Vu, H., Zhang, Y. and Herzberg, M. C. (2009b) 'Characterization of Hydrogen Peroxide-Induced DNA Release by *Streptococcus sanguinis* and *Streptococcus gordonii*', *Journal of Bacteriology*, 191(20), pp. 6281-6291.

Kristich, C. J., Nguyen, V. T., Le, T., Barnes, A. M., Grindle, S. and Dunny, G. M. (2008) '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', *Applied Environmental Microbiology*, 74(11), pp. 3377-86.

Lamont, R. J., El-Sabaeny, A., Park, Y., Cook, G. S., Costerton, J. W. and Demuth, D. R. (2002) 'Role of the *Streptococcus gordonii* SspB protein in the development of *Porphyromonas gingivalis* biofilms on streptococcal substrates', *Microbiology*, 148(6), pp. 1627-1636.

Lang, N. P., Wilson, T. G. and Corbet, E. F. (2000) 'Biological complications with dental implants: their prevention, diagnosis and treatment ', *Clinical Oral Implants Research*, 11, pp. 146-155.

Langmead, B. and Salzberg, S. L. (2012) 'Fast gapped-read alignment with Bowtie 2', *Nature Methods*, 9(4), pp. 357-359.

Larsen, R., Buist, G., Kuipers, O. P. and Kok, J. (2004) 'ArgR and AhrC Are Both Required for Regulation of Arginine Metabolism in *Lactococcus lactis*', *Journal of Bacteriology*, 186(4), pp. 1147-1157.

Larsen, R., Kok, J. and Kuipers, O. P. (2005) 'Interaction between ArgR and AhrC Controls Regulation of Arginine Metabolism in *Lactococcus lactis*', *Journal of Biological Chemistry*, 280(19), pp. 19319-19330.

Larsen, R., van Hijum, S. A. F. T., Martinussen, J., Kuipers, O. P. and Kok, J. (2008) 'Transcriptome Analysis of the *Lactococcus lactis* ArgR and AhrC Regulons', *Applied and Environmental Microbiology*, 74(15), pp. 4768-4771.

Latasa, C., Roux, A., Toledo-Arana, A., Ghigo, J.-M., Gamazo, C., Penadés, J. R. and Lasa, I. (2005) 'BapA, a large secreted protein required for biofilm formation and host colonization of *Salmonella enterica* serovar Enteritidis', *Molecular Microbiology*, 58(5), pp. 1322-1339.

Lauro, F. M. and Bartlett, D. H. (2008) 'Prokaryotic lifestyles in deep sea habitats', *Extremophiles*, 12(1), pp. 15-25.

Layton, J. C. and Foster, P. L. (2003) 'Error-prone DNA polymerase IV is controlled by the stress-response sigma factor, RpoS, in *Escherichia coli*', *Molecular Microbiology*, 50(2), pp. 549-61.

Lazarevic, V., Whiteson, K., Huse, S., Hernandez, D., Farinelli, L., Østerås, M., Schrenzel, J. and François, P. (2009) 'Metagenomic study of the oral microbiota by Illumina high-throughput sequencing', *Journal of Microbiological Methods*, 79(3), pp. 266-271.

Lee, S. F., Li, Y. H. and Bowden, G. H. (1996) 'Detachment of *Streptococcus mutans* biofilm cells by an endogenous enzymatic activity', *Infection and Immunity*, 64(3), pp. 1035-8.

Lei, B., Mackie, S., Lukomski, S. and Musser, J. M. (2000) 'Identification and Immunogenicity of Group A *Streptococcus* Culture Supernatant Proteins', *Infection and Immunity*, 68(12), pp. 6807-6818.

Lemos, J. A., Lin, V. K., Nascimento, M. M., Abranches, J. and Burne, R. A. (2007) 'Three gene products govern (p)ppGpp production by *Streptococcus mutans*', *Molecular Microbiology*, 65(6), pp. 1568-1581.

Ligtenberg, A. J., Walgreen-Weterings, E., Veerman, E. C., de Soet, J. J., de Graaff, J. and Amerongen, A. V. (1992) 'Influence of saliva on aggregation and adherence of *Streptococcus gordonii* HG 222', *Infection and Immunity*, 60(9), pp. 3878-3884.

Lin, T., Donald J, L. and Joseph J, F. (1992) 'Novel streptococcal-integration shuttle vectors for gene cloning and inactivation', *Gene*, 120(1), pp. 105-110.

Lin, X., Lamont, R. J., Wu, J. and Xie, H. (2008) 'Role of Differential Expression of Streptococcal Arginine Deiminase in Inhibition of *fimA* Expression in *Porphyromonas gingivalis*', *Journal of Bacteriology*, 190(12), pp. 4367-4371.

Litou, Z. I., Bagos, P. G., Tsirigos, K. D., Liakopoulos, T. D. and Hamodrakas, S. J. (2008) 'Prediction of cell wall sorting signals in gram-positive bacteria with a hidden markov model: application to complete genomes', *Journal of Bioinformatics and Computational Biology*, 6(2), pp. 387-401.

Liu, P., Shi, W., Zhu, W., Smith, J. W., Hsieh, S.-L., Gallo, R. L. and Huang, C.-M. (2010) 'Vaccination targeting surface FomA of *Fusobacterium nucleatum* against bacterial co-aggregation: Implication for treatment of periodontal infection and halitosis', *Vaccine*, 28(19), pp. 3496-3505.

Liu, Y.-L., Nascimento, M. and Burne, R. A. (2012) 'Progress toward understanding the contribution of alkali generation in dental biofilms to inhibition of dental caries', *International Journal of Oral Science*, 4(3), pp. 135-140.

Liu, Y. and Burne, R. A. (2009) 'Multiple two-component systems modulate alkali generation in *Streptococcus gordonii* in response to environmental stresses', *Journal of Bacteriology*, 191(23), pp. 7353-62.

Liu, Y., Dong, Y., Chen, Y.-Y. M. and Burne, R. A. (2008) 'Environmental and Growth Phase Regulation of the *Streptococcus gordonii* Arginine Deiminase Genes', *Applied and Environmental Microbiology*, 74(16), pp. 5023-5030.

Loo, C. Y., Corliss, D. A. and Ganeshkumar, N. (2000) '*Streptococcus gordonii* Biofilm Formation: Identification of Genes that Code for Biofilm Phenotypes', *Journal of Bacteriology*, 182(5), pp. 1374-1382.

Macrina, F. L., Tobian, J. A., Jones, K. R., Evans, R. P. and Clewell, D. B. (1982) 'A cloning vector able to replicate in *Escherichia coli* and *Streptococcus sanguis*', *Gene*, 19(3), pp. 345-353.

Maghnouj, A., de Sousa Cabral, T. F., Stalon, V. and Vander Wauven, C. (1998) 'The *arcABDC* Gene Cluster, Encoding the Arginine Deiminase Pathway of *Bacillus licheniformis,* and Its Activation by the Arginine Repressor ArgR', *Journal of Bacteriology*, 180(24), pp. 6468-6475.

Makarova, K. S., Mironov, A. A. and Gelfand, M. S. (2001) 'Conservation of the binding site for the arginine repressor in all bacterial lineages', *Genome Biology*, 2(4), pp. 1-8.

Mao, F., Dam, P., Chou, J., Olman, V. and Xu, Y. (2009) 'DOOR: a database for prokaryotic operons', *Nucleic Acids Research*, 37, pp. D459-63.

Marouni, M. J., Ziomek, E. and Sela, S. (2003) 'Influence of group A streptococcal acid glycoprotein on expression of major virulence factors and internalization by epithelial cells', *Microbial Pathogenesis*, 35(2), pp. 63-72.

Marsh, P. D. (1994) 'Microbial ecology of dental plaque and its significance in health and disease', *Advances in Dental Research*, 8(2), pp. 263-71.

McNab, R., Jenkinson, H. F., Loach, D. M. and Tannock, G. W. (1994) 'Cell-surface-associated polypeptides CshA and CshB of high molecular mass are colonization determinants in the oral bacterium *Streptococcus gordonii*', *Molecular Microbiology*, 14(4), pp. 743-754.

Miethke, M., Westers, H., Blom, E.-J., Kuipers, O. P. and Marahiel, M. A. (2006) 'Iron Starvation Triggers the Stringent Response and Induces Amino Acid Biosynthesis for Bacillibactin Production in *Bacillus subtilis*', *Journal of Bacteriology*, 188(24), pp. 8655-8657.

Mishra, A., Wu, C., Yang, J., Cisar, J. O., Das, A. and Ton-That, H. (2010) 'The *Actinomyces oris* Type 2 Fimbrial Shaft FimA Mediates Coaggregation with Oral Streptococci, Adherence to RBC and Biofilm Development', *Molecular Microbiology*, 77(4), pp. 841-854.

Moat, A. and Foster, J. (1995) 'Amino acids, purines and pyrimidines', *Microbial Physiology*, pp. 462-514.

Monroe, D. (2007) 'Looking for Chinks in the Armor of Bacterial Biofilms', *PLoS Biology*, 5(11), p. e307.

Moorthy, S. and Watnick, P. I. (2004) 'Genetic evidence that the *Vibrio cholerae* monolayer is a distinct stage in biofilm development', *Molecular Microbiology*, 52(2), pp. 573-587.

Mulcahy, H., Charron-Mazenod, L. and Lewenza, S. (2008) 'Extracellular DNA Chelates Cations and Induces Antibiotic Resistance in *Pseudomonas aeruginosa* Biofilms', *PLoS Pathogens*, 4(11), p. e1000213.

Murray, P. A., Prakobphol, A., Lee, T., Hoover, C. I. and Fisher, S. J. (1992) 'Adherence of oral streptococci to salivary glycoproteins', *Infection and Immunity*, 60(1), pp. 31-38.

Myers, B. M., Tietz, P. S., Tarara, J. E. and Larusso, N. F. (1995) 'Dynamic measurements of the acute and chronic effects of lysosomotropic agents on hepatocyte lysosomal pH using flow cytometry', *Hepatology*, 22(5), pp. 1519-1526.

NCBI (2011-2012) *Gene*. Available at: http://www.ncbi.nlm.nih.gov/gene (Accessed: 21st September 2015).

NCBI (2014) *Protein*. Available at: http://www.ncbi.nlm.nih.gov/protein (Accessed: 23rd October 2015).

NCBI (2015) *SGO_RS04150 (SGO_0846) cell wall protein*. Available at: http://www.ncbi.nlm.nih.gov/gene/?term=sgo_0846 (Accessed: 4th October 2015).

Nobbs, A. H., Jenkinson, H. F. and Jakubovics, N. S. (2011) 'Stick to Your Gums', *Journal of Dental Research*, 90(11), pp. 1271-1278.

Nobbs, A. H., Lamont, R. J. and Jenkinson, H. F. (2009) '*Streptococcus* Adherence and Colonization', *Microbiology and Molecular Biology Reviews*, 73(3), pp. 407-450.

Noguchi, S., Nishimura, Y., Hirota, Y. and Nishimura, S. (1982) 'Isolation and characterization of an *Escherichia coli* mutant lacking tRNA-guanine transglycosylase. Function and biosynthesis of queuosine in tRNA', *Journal of Biological Chemistry*, 257(11), pp. 6544-50.

North, A. K., Smith, M. C. M. and Baumberg, S. (1989) 'Nucleotide sequence of a *Bacillus subtilis* arginine regulatory gene and homology of its product to the *Escherichia coli* arginine repressor', *Gene*, 80(1), pp. 29-38.

Nugent, T. and Jones, D. T. (2009) 'Transmembrane protein topology prediction using support vector machines', *BMC Bioinformatics*, 10(159).

Nuxoll, A. S., Halouska, S. M., Sadykov, M. R., Hanke, M. L., Bayles, K. W., Kielian, T., Powers, R. and Fey, P. D. (2012) 'CcpA Regulates Arginine Biosynthesis in *Staphylococcus aureus* through Repression of Proline Catabolism', *PLoS Pathogens*, 8(11), p. e1003033.

Ohmori, H., Friedberg, E. C., Fuchs, R. P. P., Goodman, M. F., Hanaoka, F., Hinkle, D., Kunkel, T. A., Lawrence, C. W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G. C., Wang, Z. and Woodgate, R. (2001) 'The Y-Family of DNA Polymerases', *Molecular Cell*, 8(1), pp. 7-8.

Olsen, K. N., Budde, B. B., Siegumfeldt, H., Rechinger, K. B., Jakobsen, M. and Ingmer, H. (2002) 'Noninvasive Measurement of Bacterial Intracellular pH on a Single-Cell Level with Green Fluorescent Protein and Fluorescence Ratio Imaging Microscopy', *Applied and Environmental Microbiology*, 68(8), pp. 4145-4147.

Palmer, R. J., Kazmerzak, K., Hansen, M. C. and Kolenbrander, P. E. (2001) 'Mutualism versus Independence: Strategies of Mixed-Species Oral Biofilms *In Vitro* Using Saliva as the Sole Nutrient Source', *Infection and Immunity*, 69(9), pp. 5794-5804.

Parsek, M. and Singh, P. (2003) 'Bacterial Biofilms: An Emerging Link to Disease Pathogenesis', *Annual Review of Microbiology*, 57(1), pp. 677-701.

Peano, C., Wolf, J., Demol, J., Rossi, E., Petiti, L., De Bellis, G., Geiselmann, J., Egli, T., Lacour, S. and Landini, P. (2015) 'Characterization of the *Escherichia coli* σS core regulon by Chromatin Immunoprecipitation-sequencing (ChIP-seq) analysis', *Scientific Reports*, 5(10469).

Percival, S., Malic, S., Cruz, H. and Williams, D. (2011) 'Introduction to Biofilms', in Percival, S., Knottenbelt, D. and Cochrane, C. (eds.) *Biofilms and Veterinary Medicine*. Springer Berlin Heidelberg, pp. 41-68.

Periasamy, S., Joo, H.-S., Duong, A. C., Bach, T.-H. L., Tan, V. Y., Chatterjee, S. S., Cheung, G. Y. C. and Otto, M. (2012) 'How *Staphylococcus aureus* biofilms develop their characteristic structure', *Proceedings of the National Academy of Sciences*, 109(4), pp. 1281-1286.

Petersen, H., Keane, C., Jenkinson, H. F., Vickerman, M. M., Jesionowski, A., Waterhouse, J. C., Cox, D. and Kerrigan, S. W. (2010) 'Human platelets recognize a novel surface protein, PadA, on *Streptococcus gordonii* through a unique interaction involving fibrinogen receptor GPIIbIIIa', *Infection and Immunity*, 78(1), pp. 413-22.

Petersen, T., Brunak, S., von Heijne, G. and Nielsen, H. (2011) 'SignalP 4.0: discriminating signal peptides from transmembrane regions', *Nature Methods*, 8(10), pp. 785-786.

Pihlstrom, B. L., Michalowicz, B. S. and Johnson, N. W. (2005) 'Periodontal diseases', *Lancet*, 366(9499), pp. 1809-20.

Potera, C. (1999) 'Forging a Link Between Biofilms and Disease', *Science*, 283(5409), pp. 1837-1839.

Rahnama, M., Czupkałło, Ł., Kozicka-Czupkałło, M. and Łobacz, M. (2014) 'Gingival Crevicular Fluid – Composition and Clinical Importance in Gingivitis and Periodontitis', *Polish Journal of Public Health*, 124(2), pp. 96-98.

Ramasubbu, N., Thomas, L. M., Ragunath, C. and Kaplan, J. B. (2005) 'Structural analysis of dispersin B, a biofilm-releasing glycoside hydrolase from the periodontopathogen *Actinobacillus actinomycetemcomitans*', *Journal of Molecular Biology*, 349(3), pp. 475-86.

Ramboarina, S., Garnett, J. A., Zhou, M., Li, Y., Peng, Z., Taylor, J. D., Lee, W.-c., Bodey, A., Murray, J. W., Alguel, Y., Bergeron, J., Bardiaux, B., Sawyer, E., Isaacson, R., Tagliaferri, C., Cota, E., Nilges, M., Simpson, P., Ruiz, T., Wu, H. and Matthews, S. (2010) 'Structural Insights into Serine-rich Fimbriae from Gram-positive Bacteria', *Journal of Biological Chemistry*, 285(42), pp. 32446-32457.

Rangannan, V. and Bansal, M. (2011) 'PromBase: a web resource for various genomic features and predicted promoters in prokaryotic genomes', *BMC Research Notes*, 4(257).

Raushel, F. M., Thoden, J. B. and Holden, H. M. (1999) 'The amidotransferase family of enzymes: molecular machines for the production and delivery of ammonia', *Biochemistry*, 38(25), pp. 7891-9.

RegPrecise (2009-2015) *Regulon of ArgR in Streptococcus gordonii str. Challis substr. CH1*. Available at: http://regprecise.lbl.gov/RegPrecise/regulon.jsp?regulon_id=34822 (Accessed: 26th September 2015).

Reiner, A., Yekutieli, D. and Benjamini, Y. (2003) 'Identifying differentially expressed genes using false discovery rate controlling procedures', *Bioinformatics*, 19(3), pp. 368-375.

Rickard, A. H., Gilbert, P., High, N. J., Kolenbrander, P. E. and Handley, P. S. (2003) 'Bacterial coaggregation: an integral process in the development of multi-species biofilms', *Trends in Microbiology*, 11(2), pp. 94-100.

Rosan, B. and Lamont, R. J. (2000) 'Dental plaque formation', *Microbes and Infection*, 2(13), pp. 1599-1607.

Rosen, G., Genzler, T. and Sela, M. N. (2008) 'Coaggregation of *Treponema denticola* with *Porphyromonas gingivalis* and *Fusobacterium nucleatum* is mediated by the major outer sheath protein of *Treponema denticola*', *FEMS Microbiology Letters*, 289(1), pp. 59-66.

Rozen, S. and Skaletsky, H. J. (2000) 'Primer3 on the WWW for general users and for biologist programmers', in Krawetz, S. and Misener, S. (eds.) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Totowa, NJ: Humana Press, pp. 365-386.

Sakanaka, A., Kuboniwa, M., Takeuchi, H., Hashino, E. and Amano, A. (2015) 'Arginine-Ornithine Antiporter ArcD Controls Arginine Metabolism and Interspecies Biofilm Development of *Streptococcus gordonii*', *The Journal of Biological Chemistry*, 290(35), pp. 21185-21198.

Schneider, T. D. and Stephens, R. M. (1990) 'Sequence logos: a new way to display consensus sequences', *Nucleic Acids Research*, 18(20), pp. 6097-6100.

Schulz, C., Gierok, P., Petruschka, L., Lalk, M., Mäder, U. and Hammerschmidt, S. (2014) 'Regulation of the Arginine Deiminase System by ArgR2 Interferes with Arginine Metabolism and Fitness of *Streptococcus pneumoniae*', *mBio*, 5(6).

Seneviratne, C. J., Zhang, C. F. and Samaranayake, L. P. (2011) 'Dental plaque biofilm in oral health and disease', *Chinese Journal of Dental Research*, 14(2), pp. 87-94.

Shibli, J. A., Melo, L., Ferrari, D. S., Figueiredo, L. C., Faveri, M. and Feres, M. (2008) 'Composition of supra- and subgingival biofilm of subjects with healthy and diseased implants', *Clinical Oral Implants Research*, 19(10), pp. 975-982.

Shimada, T., Yamazaki, Y., Tanaka, K. and Ishihama, A. (2014) 'The Whole Set of Constitutive Promoters Recognized by RNA Polymerase RpoD Holoenzyme of *Escherichia coli*', *PLoS ONE*, 9(3), p. e90447.

Shrivastava, I. H., Capener, C. E., Forrest, L. R. and Sansom, M. S. P. (2000) 'Structure and Dynamics of K Channel Pore-Lining Helices: A Comparative Simulation Study', *Biophysical Journal*, 78(1), pp. 79-92.

Slany, R. K. and Kersten, H. (1994) 'Genes, enzymes and coenzymes of queuosine biosynthesis in procaryotes', *Biochimie*, 76(12), pp. 1178-1182.

Slotboom, D. J., Konings, W. N. and Lolkema, J. S. (2001) 'Cysteine-scanning Mutagenesis Reveals a Highly Amphipathic, Pore-lining Membrane-spanning Helix in the Glutamate Transporter GltT', *Journal of Biological Chemistry*, 276(14), pp. 10775-10781.

Socransky, S., Haffajee, A., Cugini, M., Smith, C. and Kent, R. (1998) 'Microbial complexes in subgingival plaque', *Journal of Clinical Periodontology*, 25(2), pp. 134-144.

Starlinger, P. and Saedler, H. (1972) 'Insertion mutations in microorganisms', *Biochimie*, 54(2), pp. 177-185.

Stepanović, S., Vuković, D., Dakić, I., Savić, B. and Švabić-Vlahović, M. (2000) 'A modified microtiter-plate test for quantification of staphylococcal biofilm formation', *Journal of Microbiological Methods*, 40(2), pp. 175-179.

Stewart, P. S. (2003) 'Diffusion in Biofilms', *Journal of Bacteriology*, 185(5), pp. 1485-1491.

Stewart, P. S. and Costerton, J. W. (2001) 'Antibiotic resistance of bacteria in biofilms', *The Lancet*, 358(9276), pp. 135-138.

Stuart, M. R., Chou, L. S. and Weimer, B. C. (1999) 'Influence of Carbohydrate Starvation and Arginine on Culturability and Amino Acid Utilization of *Lactococcus lactis* subsp. *lactis*', *Applied and Environmental Microbiology*, 65(2), pp. 665-673.

Sullivan, R., Rege, A., Corby, P., Klaczany, G., Allen, K., Hershkowitz, D., Goldder, B. and Wolff, M. (2014) 'Evaluation of a dentifrice containing 8% arginine, calcium carbonate, and sodium monofluorophosphate to repair acid-softened enamel using an intra-oral remineralization model', *The Journal of Clinical Dentistry*, 25(1 Spec No A), pp. A14-9.

Sutherland, I. (1999) 'Polysaccharases in Biofilms — Sources — Action — Consequences!', in Wingender, J., Neu, T. and Flemming, H.-C. (eds.) *Microbial Extracellular Polymeric Substances*. Springer Berlin Heidelberg, pp. 201-216.

Svensäter, G., Larsson, U. B., Greif, E. C. G., Cvitkovitch, D. G. and Hamilton, I. R. (1997) 'Acid tolerance response and survival by oral bacteria', *Oral Microbiology and Immunology*, 12(5), pp. 266-273.

Takahashi, Y., Takashima, E., Shimazu, K., Yagishita, H., Aoba, T. and Konishi, K. (2006) 'Contribution of Sialic Acid-Binding Adhesin to Pathogenesis of Experimental Endocarditis caused by *Streptococcus gordonii* DL1', *Infection and Immunity*, 74(1), pp. 740-743.

Takamatsu, D., Bensing, B. A., Prakobphol, A., Fisher, S. J. and Sullam, P. M. (2006) 'Binding of the Streptococcal Surface Glycoproteins GspB and Hsa to Human Salivary Proteins', *Infection and Immunity*, 74(3), pp. 1933-1940.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) 'MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0', *Molecular Biology and Evolution*, 30(12), pp. 2725-2729.

Terleckyj, B. and Shockman, G. D. (1975) 'Amino acid requirements of *Streptococcus mutans* and other oral streptococci', *Infection and Immunity*, 11(4), pp. 656-664.

Traxler, M. F., Summers, S. M., Nguyen, H.-T., Zacharia, V. M., Hightower, G. A., Smith, J. T. and Conway, T. (2008) 'The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*', *Molecular Microbiology*, 68(5), pp. 1128-1148.

Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J. S. and Arakawa, T. (2004) 'Role of Arginine in Protein Refolding, Solubilization, and Purification', *Biotechnology Progress*, 20(5), pp. 1301-1308.

UniProtKB (2007) *UniProtKB - A8AWI5 (A8AWI5_STRGC)*. Available at: http://www.uniprot.org/uniprot/A8AWI5 (Accessed: 9th October 2015).

Vacca-Smith, A. M., Jones, C. A., Levine, M. J. and Stinson, M. W. (1994) 'Glucosyltransferase mediates adhesion of *Streptococcus gordonii* to human endothelial cells *in vitro*', *Infection and Immunity*, 62(6), pp. 2187-2194.

Vu, B., Chen, M., Crawford, R. and Ivanova, E. (2009) 'Bacterial Extracellular Polysaccharides Involved in Biofilm Formation', *Molecules*, 14(7), pp. 2535-2554.

Wagner, J., Gruz, P., Kim, S.-R., Yamada, M., Matsui, K., Fuchs, R. P. P. and Nohmi, T. 'The *dinB* Gene Encodes a Novel *E. coli* DNA Polymerase, DNA Pol IV, Involved in Mutagenesis', *Molecular Cell*, 4(2), pp. 281-286.

Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F. and Jones, D. T. (2004a) 'The DISOPRED server for the prediction of protein disorder', *Bioinformatics*, 20(13), pp. 2138-2139.

Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F. and Jones, D. T. (2004b) 'Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life', *Journal of Molecular Biology*, 337(3), pp. 635-645.

Watnick, P. and Kolter, R. (2000) 'Biofilm, City of Microbes', *Journal of Bacteriology*, 182(10), pp. 2675-2679.

Whitman, W. B., Coleman, D. C. and Wiebe, W. J. (1998) 'Prokaryotes: The unseen majority', *Proceedings of the National Academy of Sciences of the United States of America*, 95(12), pp. 6578-6583.

Wilks, J. C. and Slonczewski, J. L. (2007) 'pH of the Cytoplasm and Periplasm of *Escherichia coli*: Rapid Measurement by Green Fluorescent Protein Fluorimetry', *Journal of Bacteriology*, 189(15), pp. 5601-5607.

Wingender, J., Jaeger, K.-E. and Flemming, H.-C. (1999) 'Interaction Between Extracellular Polysaccharides and Enzymes', in Wingender, J., Neu, T. and Flemming, H.-C. (eds.) *Microbial Extracellular Polymeric Substances*. Springer Berlin Heidelberg, pp. 231-251.

Wood, S. R., Kirkham, J., Shore, R. C., Brookes, S. J. and Robinson, C. (2002) 'Changes in the structure and density of oral plaque biofilms with increasing plaque age', *FEMS Microbiology Ecology*, 39(3), pp. 239-244.

World Health Organization (2012) *Oral Health*. Available at: http://www.who.int/mediacentre/factsheets/fs318/en/ (Accessed: 30th October 2015).

Wu, J. and Xie, H. (2010) 'Role of Arginine Deiminase of *Streptococcus cristatus* in *Porphyromonas gingivalis* Colonization', *Antimicrobial Agents and Chemotherapy*, 54(11), pp. 4694-4698.

Wu, T., Cen, L., Kaplan, C., Zhou, X., Lux, R., Shi, W. and He, X. (2015) 'Cellular Components Mediating Coadherence of *Candida albicans* and *Fusobacterium nucleatum*', *Journal of Dental Research*, 94(10), pp. 1432-8.

Wurmbach, E., Yuen, T. and Sealfon, S. C. (2003) 'Focused microarray analysis', *Methods*, 31(4), pp. 306-316.

Xiong, Y. Q., Bensing, B. A., Bayer, A. S., Chambers, H. F. and Sullam, P. M. (2008) 'Role of the serine-rich surface glycoprotein GspB of *Streptococcus gordonii* in the pathogenesis of infective endocarditis', *Microbial Pathogenesis*, 45(4), pp. 297-301.

Xu, X., He, J., Xue, J., Wang, Y., Li, K., Zhang, K., Guo, Q., Liu, X., Zhou, Y., Cheng, L., Li, M., Li, Y., Li, Y., Shi, W. and Zhou, X. (2015) 'Oral cavity contains distinct niches with dynamic microbial communities', *Environmental Microbiology*, 17(3), pp. 699-710.

Yao, Y., Berg, E. A., Costello, C. E., Troxler, R. F. and Oppenheim, F. G. (2003) 'Identification of Protein Components in Human Acquired Enamel Pellicle and Whole Saliva Using Novel Proteomics Approaches', *Journal of Biological Chemistry*, 278(7), pp. 5300-5308.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S. and Madden, T. L. (2012) 'Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction', *BMC Bioinformatics*, 13(134).

Zeng, L., Dong, Y. and Burne, R. A. (2006) 'Characterization of *cis*-Acting Sites Controlling Arginine Deiminase Gene Expression in *Streptococcus gordonii*', *Journal of Bacteriology*, 188(3), pp. 941-949.

Zhou, M., Boekhorst, J., Francke, C. and Siezen, R. J. (2008) 'LocateP: genome-scale subcellular-location predictor for bacterial proteins', *BMC Bioinformatics*, 9(173).

Zhu, Y., Weiss, E. C., Otto, M., Fey, P. D., Smeltzer, M. S. and Somerville, G. A. (2007) '*Staphylococcus aureus* Biofilm Metabolism and the Influence of Arginine on Polysaccharide Intercellular Adhesin Synthesis, Biofilm Formation, and Pathogenesis', *Infection and Immunity*, 75(9), pp. 4219-4226.

Zhuo, T., Rou, W., Song, X., Guo, J., Fan, X., Kamau, G. G. and Zou, H. (2015) 'Molecular study on the *carAB* operon reveals that *carB* gene is required for swimming and biofilm formation in *Xanthomonas citri* subsp. citri', *BMC Microbiology*, 15(1), p. 225.

Zijnge, V., van Leeuwen, M. B. M., Degener, J. E., Abbas, F., Thurnheer, T., Gmür, R. and M. Harmsen, H. J. (2010) 'Oral Biofilm Architecture on Natural Teeth', *PLoS ONE*, 5(2), p. e9321.

Zogaj, X., Nimtz, M., Rohde, M., Bokranz, W. and Romling, U. (2001) 'The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix', *Molecular Microbiology*, 39(6), pp. 1452-63.

Zúñiga, M., Champomier-Verges, M., Zagorec, M. and Pérez-Martínez, G. (1998) 'Structural and Functional Analysis of the Gene Cluster Encoding the Enzymes of the Arginine Deiminase Pathway of *Lactobacillus sake*', *Journal of Bacteriology*, 180(16), pp. 4154-4159.

Zúñiga, M., Miralles, M. d. C. and Pérez-Martínez, G. (2002) 'The Product of *arcR*, the Sixth Gene of the *arc* Operon of *Lactobacillus sakei*, Is Essential for Expression of the Arginine Deiminase Pathway', *Applied and Environmental Microbiology*, 68(12), pp. 6051-6058.