UCLA
UCLA Electronic Theses and Dissertations

Title
Identification and Characterization of a Novel Lipoprotein in Fusobacterium nucleatum

Permalink
https://escholarship.org/uc/item/3r90n33c

Author
Kaplan, Aida

Publication Date
2015

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Identification and Characterization of a Novel Lipoprotein in *Fusobacterium nucleatum*

A dissertation partially satisfying the requirements for the degree

Doctor of Philosophy

in Microbiology, Immunology and Molecular Genetics

by

Aida Kaplan

2015

ABSTRACT OF THE DISSERTATION

Identification and Characterization of a Novel Lipoprotein

in *Fusobacterium nucleatum*

by

Aida Kaplan

Doctor of Philosophy in Microbiology, Immunology and Molecular Genetics

University of California, Los Angeles, 2015

Professor Wenyuan Shi, Chair

Fusobacterium nucleatum is a Gram-negative bacterium that colonizes the oral cavity as well as other sites in the human body. It is known for its ability to adhere to a large number of different bacterial species and serving as a "bridging" organism. It plays an important role in the formation and maintenance of mature oral biofilms. At the same time it has been implicated in a number of important oral and systemic diseases and is considered an opportunistic pathogen. Very little is known about the outer membrane proteins of *F. nucleatum* and the molecular mechanisms of its adhesion. There is a large number of predicted membrane proteins that are expected to localize to either of the two membranes surrounding the cell. A number of

lipoproteins are predicted from sequence searches, yet none are currently characterized on a molecular level.

We have identified and characterized a novel hypothetical protein in *F. nucleatum* encoded by the ORF FN1253. The gene was shown to play a role in interspecies interactions with oral streptococci and was named *aid1* (Adhesion Inducing Determinant 1). Aid1 is important in modulation of adhesion to oral streptococci, as aberrant expression of the *aid1* gene leads to an uncharacteristic development of the dual species biofilm, which ultimately could lead to changes in the overall multispecies biofilm formation. The functional domains of the protein were also analyzed, and we showed that this small lipoprotein localizes to the inner leaflet of the outer membrane and requires the presence of its signal sequence as well as the anchoring cysteine residue for proper processing and maturation of the final protein product. Aberrant processing of the protein led to rapid turnover inside the cytosol, as the accumulation of the unprocessed protein appeared to be toxic to the cells. Expression of the *aid1* gene and the Aid1 protein production is tightly regulated, particularly in the presence of oral streptococci. We showed that the gene can be transcribed as two messenger RNA species, and only the longer polycistronic messenger is used for transcription.

To the best of our knowledge, this is the first molecular characterization of a fusobacterial lipoprotein, and the first hypothetical protein that is implicated in interspecies interactions within the oral cavity. This work also describes an interesting translational and transcriptional regulatory mechanism utilized by *F. nucleatum,* and is the first study to describe the regulation of a lipoprotein in this organism.

The dissertation of Aida Kaplan is approved

Robert P. Gunsalus

Asim Dasgupta

Ann M. Hirsch

Wenyuan Shi, Committee Chair

University of California, Los Angeles

2015

to my grandparents, Robert and Ashken

TABLE OF CONTENTS

ABSTRACT OF THE DISSERTATION ii

LIST OF FIGURES viii

LIST OF TABLES ix

ACKNOWLEDGEMENTS x

VITA xi

FOREWORD 1

CHAPTER 1: Characterization of *aid1*, a novel gene involved in *Fusobacterium nucleatum* interspecies interactions 5

Abstract 6

Introduction 7

Materials and Methods 9

Results 13

Discussion 17

Acknowledgements 22

CHAPTER 2: Intracellular localization and characterization of the functional domains of Aid1 32

Abstract 33

Introduction 34

Materials and Methods 37

Results 40

Discussion 42

CHAPTER 3: Characterization of transcriptional and translational regulation of *aid1* 50

Abstract 51

Introduction 52

Materials and Methods 55

Results 58

Discussion 61

CONCLUSION 74

REFERENCES 78

LIST OF FIGURES

Chapter 1

Figure 1. Differential expression of *aid1* in *F. nucleatum* single and dual species biofilms 23

Figure 2. *aid1* gene expression is time- and touch-dependent 24

Figure 3. Analysis and verification of *aid1* mutants 25

Figure 4. Quantitative coaggregation assay between *F. nucleatum* strains and *S. sanguinis* 27

Figure 5. Biofilm growth and morphology of *aid1* mutants with oral streptococci 28

Supplemental Figure 1. Construction of *aid1* mutant strains 29

Chapter 2

Figure 1. Aid1 is a predicted outer membrane lipoprotein 45

Figure 2. Aid1 localizes to the outer membrane of the cell 46

Figure 3. Aid1 is protected from degradation by enzymatic digest 48

Figure 4. Aid1 structural mutants are unstable in the cell 49

Chapter 3

Figure 1. *aid1* promoter constructs 67

Figure 2. *aid1* expression varies due to the length of the upstream sequence 68

Figure 3. *aid1* endogenous promoter sequence is necessary and sufficient for downregulation of the gene in the presence of *S. sanguinis* 69

Figure 4. Aid1 protein production requires co-transcription with FN1254 70

Figure 5. FN1254 expression is similar to that of *aid1* 72

Figure 6. *aid1* and FN1254 are regulated in a similar manner in the presence of *S. sanguinis* 73

LIST OF TABLES

Chapter 1

Table 1. Aggregation ability of *aid1* mutants 30

Table 2. Inhibition of cellular coaggregation 24

Chapter 3

Table 1. Primers used in this study 66

ACKNOWLEDGEMENTS

I thank my mentors, Dr. Wenyuan Shi for his contagious passion for science and endless creativity and Dr. Renate Lux for her patience and teaching me the fine details of scientific research.

I thank my committee members, Drs. Rob Gunsalus, Ann Hirsch and Asim Dasgupta for their help, support and invaluable insight throughout my challenging project.

I thank my lab mates, both past and present, for helping me stay sane and creating a great intellectual environment that has significantly contributed to my growth as a scientist. In particular, I'd like to thank Dr. Bruno Lima for being an amazing scientist that I will always look up to and becoming a great friend. Life in the lab wouldn't have been the same without him.

I would also like to thank my friends and family. My grandparents for filling our house with science textbooks and sparking my love for science at a young age, my mom for bringing me to the US and giving me the opportunity to get to where I am today and my brother for always being proud of me.

And last, but certainly not least, I save the biggest thank you to my best friend, mentor, collaborator and incredible husband, Dr. Chris Kaplan, without whose love and support I would have never survived the emotional ups and downs of graduate school.

VITA

2009	B.S. Microbiology, Immunology and Molecular Genetics University of California, Los Angeles
2011	Awarded UCLA MBI Whitcome Training Grant University of California, Los Angeles
2014	Awarded UCLA Graduate Division Dissertation Year Fellowship University of California, Los Angeles
2014	Sidney Rittenberg Award for Outstanding Graduate Students University of California, Los Angeles

PUBLICATIONS AND PRESENTATIONS

Kaplan, A., Kaplan, C. W., He, X., McHardy, I.H., Shi, W., Lux, R. (2014). Characterization of *aid1*, a novel gene involved in *Fusobacterium nucleatum* interspecies interactions. *Microbial ecology 68*(2), 379-387

Wang, R., **Kaplan, A.*,** Guo, L., Shi, W., Zhou, X., Lux, R., & He, X. (2012). The influence of iron availability on human salivary microbial community composition. *Microbial ecology*, *64*(1), 152-161. (*First co-author)

Kaplan, C. W., Sim, J. H., Shah, K. R., **Kolesnikova-Kaplan, A.**, Shi, W., & Eckert, R. (2011). Selective membrane disruption: mode of action of C16G2, a specifically targeted antimicrobial peptide. *Antimicrobial agents and chemotherapy*, *55*(7), 3446-3452.

Kaplan, A., Shi, W., Lux, R. Aid1 is a novel lipoprotein in *Fusobacterium nucleatum.* 93rd General Session & Exhibition of the International Association for Dental Research, March 2015, Boston, Massachusetts. Poster Presentation

Kaplan, A., Kaplan, C. W., He, X., McHardy, I. H., Shi, W., Lux, R. Characterization of a novel gene involved in *Fusobacterium nucleatum* interspecies interactions. 114th Annual Meeting of the American Society for Microbiology, May 2014, Boston, Massachusetts. Poster Presentation

Kaplan, A., Kaplan, C.W., Lux, R., Shi, W. Characterization of a novel gene involved in *Fusobacterium nucleatum* interspecies interactions. January 2014. UCLA Molecular Biology Institute Annual Retreat. Poster Presentation

Kaplan, A., Kaplan, C.W., Lux, R., Shi, W. *Aid1* modulates interactions of *Fusobacterium nucleatum* with oral streptococci. Annual Meeting of the International Association for Dental Research, March 2013, Seattle, Washington. Poster Presentation

Park, J., **Kaplan, A.,** Haake, S.K., Lux, R., Characterization of fusobacterial adhesins involved in coaggregation with *Porphyromonas gingivalis.* Annual Meeting of the International Association for Dental Research, March 2013, Seattle, Washington. Poster Presentation

Kaplan, A., Kaplan, C.W., Lux, R., Shi, W. Characterization of a novel protein involved in *Fusobacterium nucleatum* interactions with oral streptococci. October 2012. UCLA Molecular Biology Institute Annual Retreat. Poster Presentation

Kaplan, A., McHardy, I., Kaplan, C.W., Haake, S.K., Lux, R., Shi, W., Fn1253 plays a role in interspecies interactions of *F. nucleatum*. February 2012. UCLA School of Dentistry Poster Competition (**Second place award**). Poster Presentation

Park, J., **Kaplan, A.,** Lux, R., Haake, S.K., Identification and characterization of fusobacterial adhesins involved in coaggregation with *Porphyromonas gingivalis.* February 2012. UCLA School of Dentistry Poster Competition. Poster Presentation

Kaplan, A., McHardy, I.H., Kaplan, C.W., Lux, R., Shi, W. Identification and characterization of a novel protein involved in interspecies interactions of *Fusobacterium nucleatum*. October 2011, UCLA Molecular Biology Institute Annual Retreat. Poster Presentation

Foreword

The oral biofilm, also known as the dental plaque, is a great model system for understanding bacterial interspecies interactions. It is home to over 600 different species (Dewhirst et al. 2010) of bacteria that interact within the complex architecture of the biofilm, where each subset of species plays a special role in the maturation and maintenance of the mature dental plaque. It is a dynamic environment where different species require each other's presence for growth and nutrient acquisition. In the early days of discovery and characterization of oral biofilms the microorganisms comprising the dental plaque were characterized into early and late colonizing species. Many of these species adhere to one another and exhibit intraspecies aggregation and interspecies coaggregation (Kolenbrander & London 1992). Early colonizers are primarily Gram-positive species, which include streptococci and actinomyces (Palmer et al. 2003; Kolenbrander & London 1993). These species are capable of directly binding to the tooth surface by recognizing receptors in the salivary pellicle coating the tooth enamel (Nyvad & Kilian 1987). Late colonizing species are often Gram-negative species that do not have the ability to directly attach to oral surfaces or coaggregate with early colonizing microorganisms (Kolenbrander & London 1993). Early and late colonizing species interact with one another primarily through their interaction with a bridging organism *Fusobacterium nucleatum*. The ability of *F. nucleatum* to adhere to a large number of bacterial species, both early and late colonizers, makes it a key organism in the formation and maintenance of the dental plaque. While the importance of the adhesion properties of *F. nucleatum* has been recognized, only one fusobacterial adhesin for interspecies interaction has been characterized on a molecular basis (Kaplan et al. 2009) to date. The interaction between *F. nucleatum* and oral streptococci is arginine-inhibitable and mediated by the outer membrane protein RadD (Kaplan et al. 2009). While RadD plays a major role in the interactions between *F. nucleatum* and the early colonizing oral streptococci, the inactivation of the gene does not completely block coaggregation as observed in *in vitro* coaggregation assays (Kaplan et al. 2014). At the same time, addition of arginine to the coaggregation mixture completely prevents any binding of the *F. nucleatum* cells to the chains

of streptococci. This observation suggests the existence of additional factors involved in these interspecies interactions that have not yet been identified.

Along with the lack of molecular characterization of different adhesion mechanisms, there is very little known about the changes taking place in gene expression of different bacterial species upon mutual binding. A previous study has investigated transcriptional responses of *F. nucleatum* in the presence of oral streptococci as well as two representatives of the late colonizers, *T. denticola* and *T. forsythia* (McHardy 2011). In the presence of *S. sanguinis* and *S. gordonii* as representative early colonizing species, a number of genes was differentially regulated in *F. nucleatum.* Many of these genes encode metabolic proteins, transcriptional factors and stress related genes. A large number of the differentially regulated genes was uncharacterized and annotated as hypothetical proteins. While regulation of many of these genes is of great interest, one gene in particular stood out due to its differential regulation in all conditions that were tested. Annotated as FN1253 in the *F. nucleatum* ATCC 25586 strain, and the corresponding homologue in *F. nucleatum* ATCC 23726 strain being HMPREF0397_0433, the gene encodes a 67 amino acid long hypothetical protein. Through microarray analysis we discovered that the gene was significantly upregulated in single species *F. nucleatum* biofilms, but even more highly downregulated in the presence of oral streptococci, suggesting a role in interspecies interactions between *F. nucleatum* and *S. sanguinis* as a representative species of early colonizing oral streptococci (McHardy 2011). In both published genomes of strains ATCC 25586 and 23726, FN1253 is annotated as a hypothetical protein. Upon closer examination of the protein sequences we discovered a conserved lipobox motif and a predicted signal sequence, which suggested FN1253 to be a novel lipoprotein in *F. nucleatum.*

The first chapter of this dissertation illustrates the genetic characterization of FN1253 and its role in interspecies interactions between *F. nucleatum* and *S. sanguinis.* The chapter is a copy

of a manuscript published in the Journal of Microbial Ecology. It is reprinted with permission for academic and for no profit use via license number 3604980294318.

In the second chapter I sought to further characterize the protein encoded by ORF FN1253 and determine its cellular localization and trafficking.

The unique regulation of FN1253 is described in chapter three, where I characterize the necessary genetic elements for the gene expression, regulation in the presence of *S. sanguinis* and protein production. Chapters two and three are currently manuscripts in preparation.

This work has been supported by the NIH/NIDCR grant DE021108 and in part by the Whitcome Pre-doctoral Training grant and the UCLA Graduate Division Dissertation Year Fellowship.

Chapter 1: Characterization of *aid1*, a novel gene involved in *Fusobacterium nucleatum* interspecies interactions

Abstract

The oral opportunistic pathogen *Fusobacterium nucleatum* is known to interact with a large number of different bacterial species residing in the oral cavity. It adheres to a variety of Gram-positive bacteria, including oral streptococci via the arginine-inhibitable adhesin RadD. In this study, we describe a novel protein encoded by the predicted open reading frame FN1253 that appears to play a role in interspecies interactions of *F. nucleatum*, particularly with oral streptococci and related Gram-positive species. We designated FN1253 as *aid1* (Adherence Inducing Determinant 1). Expression analyses demonstrated that this gene was induced in *F. nucleatum* single species biofilms, while the presence of representative members of the oral microbiota known to adhere to *F. nucleatum* triggered its suppression. Inactivation as well as overexpression of *aid1* affected the ability of *F. nucleatum* to coaggregate with oral streptococci and the closely related *Enterococcus faecalis*, but not other Gram-positive oral species tested. Furthermore, overexpression of *aid1* led to a drastic change in the structure of dual species biofilms of *F. nucleatum* with oral streptococci. Aid1 function was abolished in the presence of arginine and found to be dependent on RadD. Interestingly, differential expression of *aid1* did not affect mRNA and protein levels of RadD. These findings indicate that RadD-mediated adhesion to oral streptococci involves more complex cellular processes than the simple interaction of adhesins on the surface of partner strains. Aid1 could potentially play an important role in facilitating RadD-mediated interaction with oral streptococci by increasing binding specificity of *F. nucleatum* to other microbial species.

Introduction

The oral cavity is a great model system for studying polymicrobial interactions since it is home to over 600 different recognized species of bacteria (Dewhirst et al. 2010) most of which are considered to be commensal (Aas et al. 2005). The microorganisms in the oral biofilm have been categorized into early and late colonizers (Kolenbrander & London 1993). Early-colonizing species are primarily Gram-positive, able to adhere directly to the tooth surface and form the basal layers of the oral biofilm. Late colonizers are comprised of mostly Gram-negative bacteria, including certain periodontal pathogens such as *Treponema denticola*, *Tannerella forsythia* and *Porphyromonas gingivalis* as well as others. Bacteria within the oral biofilm, also known as the dental plaque, form a complex network of direct and indirect interactions. The spatial distribution of different bacterial species is important in oral biofilm formation and architecture. Many of the known oral bacterial species do not directly adhere to one another; instead they interact indirectly via their mutual attachment to *Fusobacterium nucleatum* (Kolenbrander et al. 2002). *F. nucleatum* is a Gram-negative, anaerobic fusiform bacterium that has been associated with periodontal disease and a number of systemic diseases (Bastos et al. 2011; González et al. 2010; Hsiao et al. 2012; Lee et al. 2012; Okada et al. 2002; Sparks Stein et al. 2012; Han & Wang 2013). It is considered a "bridging organism" due to its ability to form a "colonization bridge" between species that do not directly interact, thus playing an integral role in the formation of a mature dental plaque. The physical attachment between interacting partner species is mediated by specific cellular adhesion proteins localized on their outer membranes. To characterize these important surface features in *F. nucleatum* on a molecular level, we employed a genetic system that was previously established in our laboratory and lead to discovery of the large outer membrane autotransporter protein RadD, which is required for effective binding to early colonizers (Kaplan et al. 2009). In order to investigate the transcriptional responses of *F. nucleatum* upon interactions with other species, we conducted

microarray analysis of *F. nucleatum* grown in the presence of representatives from both early and late colonizing species (McHardy 2011). These microarray data revealed that a small hypothetical protein encoded by FN1253 according to annotation of *F. nucleatum* ATCC 25586 (Kapatral et al. 2002) is induced in *F. nucleatum* single species biofilms but ubiquitously repressed in the presence of both early and late colonizers. Downregulation of this gene in dual species biofilms was more pronounced upon interaction with early colonizing streptococci.

FN1253 homologues are highly conserved across all fusobacterial species sequenced to date and, with no homologues found in other species for which genome sequences are available, it appears to be unique to fusobacteria. In this study, we investigated the role of FN1253 in microbial interactions involving *F. nucleatum.* We demonstrated that FN1253, which we denoted as *aid1* (Adherence Inducing Determinant gene 1), plays a role in interaction of *F. nucleatum* with oral streptococci. Aid1 function requires the presence of the previously identified adhesin RadD (Kaplan et al. 2009). To the best of our knowledge, this is the first hypothetical protein in the *F. nucleatum* genome that has been characterized thus far with a role in interspecies interactions.

Materials and Methods

Bacterial strains and culture conditions

F. nucleatum strains were grown on Columbia agar plates supplemented with 5% sheep blood or in Columbia broth (Difco, Detroit, MI) under anaerobic conditions (5% CO_2, 5% H_2, 90% N_2). Thiamphenicol (MP Biomedicals, Irvine, CA) at 5μg/ml was used for selection and maintenance of strains containing the *catP* determinant. Clindamycin (MP Biomedicals, Irvine, CA) at 0.4μg/ml was used for selection and maintenance of strains possessing the *ermB* determinant. *Streptococcus sanguinis* ATCC 10556 and *S. gordonii* ATCC 10558 were grown anaerobically in Todd-Hewitt (TH) broth (BD Difco, Detroit, MI) at 37°C. *Enterococcus faecalis* ATCC 19433 was grown aerobically at 37°C with shaking in Brain Heart Infusion (BHI) (BD Difco, Sparks, MD) broth. *Lactobacillus casei* ATCC 393 was grown aerobically in the presence of 5% CO_2 in Luria Berthani (LB) (BD Difco, Sparks, MD) broth supplemented with 1% yeast extract. *Staphylococcus epidermidis* ATCC 12228 was grown aerobically at 37°C with shaking at 250rpm in Tryptic Soy Broth (TSB) (BD Difco, Sparks, MD). *Veillonella atypica* PK1910 was grown anaerobically in TH broth (BD Difco, Detroit, MI) supplemented with 0.06% lactic acid at 37°C. *Porphyromonas gingivalis* W50 was grown in Columbia broth anaerobically at 37°C. *Treponema denticola* ATCC 35405 was grown in TYGVS anaerobically at 37°C as previously described (Ohta et al. 1986). *Tannerella forsythia* ATCC 43037 was grown in NAM medium (ATCC) anaerobically at 37°C.

Mutant construction

In this study we used *F. nucleatum* ATCC 23726, which in contrast to ATCC 25586 can be genetically modified. FN1253 is annotated as HMPREF0397_0433 (Genbank ID: ADVK00000000.1, NCBI BioProject Accession: PRJNA31471 ID: 31471) in *F. nucleatum* ATCC 23726. The Δ*aid1* inactivation strain was constructed by double homologous recombination

(Supplemental Figure 1A). A sequence including *aid1* and ~600-700bp flanking regions (that contain only truncated portions of the respective upstream and downstream genes) was amplified from wild type *F. nucleatum* ATCC23726 strain with *Taq* DNA polymerase (Invitrogen, Carlsbad, CA, USA) using forward (5'- TTTATTAAAACTTATGGGAGATAGATA-3') and reverse (5'- TCCAGAAGGAAAACAACCATCA-3') primers and subcloned into the pJET 2.1 cloning vector (Fermentas, Inc, Glen Burnie, MD, USA) to obtain construct pIP-*aid1* (Supplemental Figure 1A). The *catP* gene was amplified from plasmid pHS31 (Kaplan et al. 2005) using forward (5'-GTCACAGGATCCCAGTCGAAGTGGGCAAGT-3') and reverse (5'-GTCACCGGATCCCCGTATTTCTACGATGTTTTTGC-3') primers and subcloned into the pJET 2.1 vector. *catP* was excised with BamHI and ligated into a pIP-*aid1* digested with BamHI resulting in an insertion at nucleotide 38 within the *aid1* gene. The resulting plasmid pIP-*aid1::catP* was verified by restriction analysis and PCR. The plasmid was linearized with ScaI prior to transformation into *F. nucleatum.* For overexpression of *aid1* in *F. nucleatum* we constructed plasmid pEP-*aid1* as follows. The fragment carrying *aid1* including the upstream and downstream regions described above was excised from pIP-*aid1* using XbaI and XhoI and ligated into linearized pHS58 shuttle plasmid (Haake, personal communication) that carries an *ermB* erythromycin/clindamycin resistance cassette to generate the pEP-*aid1* expression vector. Plasmid pEP-*aid1* was transformed into wildtype *F. nucleatum* strain ATCC 23726 as well as its derivative lacking *radD* according to previously described protocols (Haake et al. 2000) to yield *Fn/*pEP-*aid1* and ∆*radD/*pEP-*aid1*, respectively.

Transcriptional Analysis

Genomic DNA was extracted from stationary-phase cells following standard protocols and used for generating reference standard curves. Total RNA was extracted using High-Pure RNA Isolation Kit (Roche, Palo Alto, CA, USA) according to manufacturer's instructions. 2μg of total RNA was used for cDNA synthesis using Transcriptor Universal cDNA Master (Roche, Palo

Alto, CA, USA) following the manufacturer's protocol. For qRT-PCR, SYBR green (Bio-Rad, Hercules, CA, USA) was used for fluorescence detection with the iCycler real-time PCR system (Bio-Rad), according to the manufacturer's instructions. *aid1* cDNA was amplified using 5'-TACAGGAGGTGCCGTAGCAG-3' forward and 5'-TTTTTGTTAATTCTCCAGCTCCA-3' reverse primers. Expression levels of 16S rRNA were determined using 5'-TTGGACAATGGACCGAGAGT-3' forward and 5'-GCCGTCACTTCTTCTGTTGG-3' reverse primers for normalization of the qRT-PCR data.

Coaggregation assay

Coaggregation assays were performed in coaggregation buffer (CAB) 150mM NaCl, 1mM Tris, 0.1mM $CaCl_2$, 0.1mM $MgCl_2$ (Cisar et al. 1979). Briefly, the cells were collected, washed and resuspended in phosphate-buffered saline (PBS). Equal numbers of bacterial cells were diluted in CAB or 25% pooled saliva in ddH_2O to a final concentration of 2×10^9 cell ml^{-1} in a 200µL clear PCR reaction tube. The tubes were vortexed for 5s and graded on a 0-4 scale after 10 min based on degree of coaggregation (Kolenbrander et al. 1990). Scores were assigned as follows: 0 - no visible coaggregation; 1 - small aggregates that stay suspended; 2 - larger aggregates that settle slowly and leave the supernatant turbid; 3 - large aggregates that settle quickly but leave the supernatant turbid; 4 - complete sedimentation with a clear supernatant. No autoaggregation of individual strains was observed in our experimental controls. For coaggregation inhibition assays, 50mM final concentration of either D-galactose, L-arginine, D-glucose or N-acetyl galactosamine were added to the reaction tube containing the *F. nucleatum* strain and mixed by vortexing before addition of the coaggregation partner strain.

Spectrophotometric coaggregation assay

Spectrophotometric co-aggregation experiments were performed according to our published procedures (Kaplan et al. 2009). Briefly, the cells were combined in CAB as described above

and after 10 min of incubation, the co-aggregation reactions were centrifuged at low speed (100 *g*) for 1min to pellet co-aggregating cells while leaving the non-aggregated bacteria in suspension. The supernatant was then removed without disturbing the pellet, and the optical density of the recovered supernatant was measured at 600 nm. Relative co-aggregation of species A and B was determined by dividing the difference between the total turbidity of each partner strain and the coaggregation supernatant turbidity by the total turbidity of each partner strain using the $\{[OD_{600}(A)+OD_{600}(B)]-OD_{600}(A+B)\} /[OD_{600}(A)+OD_{600}(B)]$ formula.

Biofilm growth

Biofilms were grown in eight-well chambered coverglasses (Nunc, Rochester, NY) pre-coated with 100µl of 50% saliva diluted in ddH_2O that was centrifuged for 5 min at 10,000 x *g* to remove debris. The chambers were UV sterilized for 1hr before inoculation. Overnight cultures of *F. nucleatum* (~10^7 cells), *S. sanguinis* and *S. gordonii* (~10^5 cells) were inoculated into the growth chamber wells containing 400µl of filter-sterilized BHI saliva broth (BHIS: 25% BHI and 25% saliva). Samples were incubated overnight under anaerobic conditions (5% H_2, 5% CO_2, 90% N_2) at 37°C. After 22-24hrs samples were fluorescently labeled with the nucleic acid staining dye SYTO9 (Invitrogen, Carlsbad, CA) and visualized using a PASCAL 5 confocal laser scanning microscope (CLSM) (Zeiss, Jena, Germany). The scanning module of the system was mounted onto an inverted microscope (Axiovert 200M) and samples were viewed through a 40x oil-immersion objective (Achroplan/N.A. 1.3). Excitation of 488nm with an argon laser in combination with a 505-530nm bandpass emission filter was used for SYTO9 fluorescence imaging.

Statistical analysis

Student's *t*-test was performed using Excel 2010 (Microsoft, Seattle, WA)

Results

Confirmation of the *aid1* gene expression profile

Previous microarray analysis in *F. nucleatum* strain ATCC 23726 revealed a unique expression pattern for an ORF that corresponds to FN1253 in the published genome of *F. nucleatum* strain ATCC 25586 (McHardy 2011). *F. nucleatum* ATCC 23726 was used for further characterization since it can be genetically manipulated (Kaplan et al. 2005). Homology searches with the recently published draft genome of *F. nucleatum* ATCC 23726 confirmed the presence of the corresponding ORF (HMPREF0397_0433) in this strain, albeit the original annotation indicated a truncated version of the protein. Analysis of flanking sequences revealed that HMPREF0397_0433 exhibits 100% homology with FN1253 indicating a mis-annotation in the ATCC 25586 genome, where the protein is missing the N-terminal portion.

Expression of FN1253 was induced in single species biofilms, and repressed in the presence of representative species of both early and late colonizers (McHardy 2011). Conditions tested included dual species biofilms of *F. nucleatum* ATCC 23726 with *S. sanguinis* ATCC 10556 and *S. gordonii* ATCC 10558 representing early colonizing streptococci, as well as *T. denticola* ATCC 35405 and *T. forsythia* ATCC 43037 as representatives of Gram-negative late colonizing species. The microarray data were validated via quantitative real-time PCR (qRT-PCR) with *aid1*-specific primers, which confirmed the differential regulation of *aid1* in single and dual species biofilms (Figure 1). Repression of *aid1* was most pronounced upon interaction with the streptococci.

Repression of *aid1* is time- and contact-dependent

Next, to determine whether physical interaction of *F. nucleatum* with streptococci plays a role in regulation of *aid1*, we analyzed the *aid1* mRNA levels in mixed culture. Cells were mixed

together with and without arginine, a known inhibitor of interactions of *F. nucleatum* with oral streptococci (Kaplan et al. 2009), and spun down into a pellet, since only few *F. nucleatum* cells start forming biofilms at early time points. qRT-PCR revealed a small decrease in *aid1* expression after 1hr incubation (Figure 2). At the 20hr time point the expression of *aid1* was significantly decreased. When 50mM arginine was added to the cell mixture there was no significant gene regulation present at both time points. Neither co-culture with *S. sanguinis* nor the addition of arginine affected the growth rate of *F. nucleatum.* Since fusobacterial interactions with streptococci are predominantly mediated by RadD, we performed similar analysis with a *F. nucleatum* mutant strain lacking this major adhesin. However, inactivation of *radD* did not have an effect on *aid1* levels, while addition of 50mM of arginine to the mixture blocked repression of *aid1* in a manner comparable to the wildtype parent.

Construction and basic characterization of *aid1* gene inactivation and overexpression derivatives

In order to characterize *aid1* on a molecular level, we constructed a gene inactivation mutant as well as a strain overexpressing the gene in wild type *F. nucleatum* ATCC 23726. Multiple attempts to obtain an inactivation mutant in which the gene was fully deleted were unsuccessful. As an alternative, we disrupted *aid1* by inserting a chloramphenicol/thiamphenicol resistance cassette (*catP*) 38 nucleotides downstream of the gene start site. The mutation was introduced into the *F. nucleatum* wild type ATCC 23726 strain via double homologous recombination of a construct carrying the *catP* gene and the corresponding chromosomal regions (*aid1*::*catP*) (Supplemental Figure 1A). This results in a non-functional truncated and out-of-frame final gene product. The overexpression mutant was constructed by introducing an expression vector carrying *aid1* under its own endogenous promoter (pEP-*aid1*) into *F. nucleatum* ATCC 23726 (Supplemental Figure 1B). The respective mutant derivatives were selected with appropriate antibiotics, thiamphenicol for the inactivation mutant and clindamycin for the overexpression

strain. Inactivation of *aid1* was confirmed by PCR (Figure 3A). Both mutants had growth rates comparable to the wild-type strain (data not shown). Expression levels of *aid1* were examined by qRT-PCR and the overexpression strain was found to produce approximately three times more *aid1* mRNA than the wild type parent strain (Figure 3B).

aid1 modulates the coaggregation ability of _F. nucleatum_ with oral streptococci and related species in a RadD-dependent manner

The *aid1* encoding gene is noticeably repressed in the presence of oral streptococci and Gram-negative late colonizers. We therefore investigated the involvement of *aid1* in the ability of *F. nucleatum* to coaggregate with a wide array of oral bacterial species. Since *F. nucleatum* does not autoaggregate using standard protocols, we first tested if lack or overexpression of *aid1* would alter this behavior and found that this is not the case (Table 1). We then investigated the ability of the *aid1* mutants to interact with other species using a standard coaggregation assay (Cisar et al. 1979). Equal numbers of cells were added into a tube with coaggregation buffer (CAB) and coaggregation was scored on a 0-4 scale after 10 mins as described in the Experimental Procedures section. The *aid1* inactivation strain displayed a reduced ability to aggregate with oral streptococci as well as closely related *Enterococcus faecalis* while its ability to coaggregate with other Gram-positive and Gram-negative species was not affected. The *F. nucleatum* strain overexpressing *aid1* exhibited an increased aggregation ability with oral streptococci, a feature that was not apparent during interaction with other Gram-positive and Gram-negative oral species (Table 1). Based on this data FN1253 was designated as *aid1* (Adherence Inducing Determinant gene 1). To further characterize the nature of this enhanced interaction with oral streptococci, we performed coaggregation assays with *S. sanguinis* as representative species in the presence of known inhibitors of fusobacterial interactions with other species as well as carbohydrates that generally do not interfere with coaggregation

(Kolenbrander & Andersen 1989). Enhanced coaggregation of the *aid1*-overexpressing mutant derivative with streptococci was only blocked by arginine (Table 2).

We reported previously that an arginine-inhibitable adhesion RadD plays an important role in mediating the coaggregation between *F. nucleatum* and a variety of Gram-positive oral bacteria, including oral streptococci (Kaplan et al. 2009). To investigate whether Aid1 requires RadD for its function, we introduced the pEP-*aid1* plasmid into a *radD* mutant derivative of ATCC 23726, which is unable to coaggregate and form biofilms with oral streptococci (Kaplan et al. 2009). Despite overexpression of *aid1*, lack of *radD* still abolished coaggregation of *F. nucleatum* with oral streptococci (Table 1).

All coaggregation data were quantified by performing a spectrophotometric coaggregation assay, which verified the differences in coaggregation under the various conditions tested in the visual coaggregation assay (Figure 4).

aid1 overexpression enhances the ability of _F. nucleatum_ to form biofilms with oral streptococci and alters biofilm morphology

Since *aid1* levels are significantly decreased after 20hrs of co-incubation with *S. sanguinis*, we investigated the ability of *aid1* mutant derivatives to form biofilms with *S. sanguinis*. Lack of *aid1* did not affect the ability of *F. nucleatum* to form dual species biofilms with *S. sanguinis.* The biofilms formed with the *aid1* inactivation mutant were indistinguishable from the ones formed by *S. sanguinis* and wild type *F. nucleatum*. In contrast, biofilms produced by the *aid1* overexpression mutant in the presence of *S. sanguinis* were consistently taller and morphologically different. Streptococcal cells were more evenly distributed throughout the height of the dual-species biofilm, as opposed to localizing closer to the bottom of the biofilm when grown with wildtype *F. nucleatum*. (Figure 5).

Discussion

In this study, we discovered that a small conserved hypothetical protein in *F. nucleatum* ATCC 23726, which we designated as Aid1 (Adhesion Inducing Determinant 1) plays a role in fusobacterial interspecies interactions and biofilm formation. According to the corresponding *F. nucleatum* ATCC 25586 annotation (Kapatral et al. 2002), Aid1 is encoded by a homolog of ORF FN1253. This ORF first caught our attention during microarray analysis, since it was significantly induced in *F. nucleatum* mono-species biofilms and repressed in the presence of other oral bacterial species (McHardy 2011). We confirmed the expression pattern of *aid1* by qRT-PCR and demonstrated a much greater reduction of gene expression in the presence of the Gram-positive oral streptococci compared to Gram-negative species (Figure 1). Neither inactivation nor overexpression of *aid1* had any discernable effect on general cell physiology, appearance, overall membrane composition or single species biofilm growth (data not shown), and interaction with Gram-negative species (Table 1). Aid1-dependent differences in coaggregation appear to be limited to oral streptococci and closely related species, since coaggregation of the *F. nucleatum aid1* mutant derivatives with more distantly related Gram-positive species such as *Lactobacillus casei* and *Staphylococcus epidermidis* was indistinguishable from the wildtype parent behavior. This apparent differential effect of Aid1 on binding to partner species was especially intriguing considering that *F. nucleatum* interacts with numerous oral bacterial species and serves as a colonization bridge between species that cannot attach to each other (Kolenbrander & London 1993; Whittaker et al. 1996). This unique characteristic of fusobacteria plays a major role in the formation and architecture of oral biofilms and fusobacterial interactions with other microorganisms have previously been characterized by specific inhibition with a variety of amino acids and carbohydrates (Kolenbrander & Andersen 1989; Kolenbrander et al. 1990). Futhermore, we recently identified the large outer membrane protein RadD as the major adhesin for interactions with a number of Gram-positive oral bacterial

species including oral streptococci, and demonstrated that lack of this large outer membrane protein disrupts dual-species biofilm architecture (Kaplan et al. 2009). In this study we found that Aid1-mediated effects are dependent on the presence of RadD (Table 1).

Investigation of Aid1, the novel protein that appears to modulate in particular fusobacterial interactions with oral streptococci and the closely related *E. faecalis*, led to the hypothesis that interspecies interactions of *F. nucleatum* are more complex than the simple binding of a fusobacterial surface adhesin with the corresponding adhesin on its partner species. Specifically, we propose that this important bridging organism distinguishes its binding partners by employing large outer membrane proteins for general attachment to larger groups of bacteria and fine-tuning these interaction by providing binding specificity via differential expression of small membrane-associated proteins such as Aid1. The idea that *F. nucleatum* can distinguish its binding partners is also supported by previous work on the inhibition of coaggregation, which demonstrated that *F. nucleatum* interacts with oral streptococci in a multimodal manner since different combinations of inhibitors were necessary to block these interactions (Takemoto et al. 1995).

Even though the presence of RadD is required for Aid1 function both proteins overlap in their involvement in the same process on inter-cellular interactions, their gene expression is independent of each other. Increased abundance of *aid1* appears to induce the ability of *F. nucleatum* to bind to oral streptococci without significantly increasing *radD* mRNA or protein levels (data not shown). Furthermore, expression of *aid1* is regulated in a contact- and time-dependent manner that is RadD-independent (Figure 2). Addition of arginine completely blocks all cellular interactions between *F. nucleatum* and *S. sanguinis* (Table 2, Figure 4) even when cells are coincubated in pellet form and abolishes the typical dramatic decrease of *aid1* mRNA in the presence of this partner species (Figure 2), which suggests that cellular contact is

required for activation of the signaling pathway resulting in regulation of the *aid1* gene. At the same time, inactivation of RadD, while still dramatic, does not completely eliminate coaggregation between *F. nucleatum* and *S. sanguinis*. This observation is more obvious in the more sensitive spectrophotometric coaggregation assay (Figure 4). Our data suggest that the reduced cell-cell contact in the absence of RadD is still sufficient to trigger the signaling cascade leading to repression of *aid1*, while addition of arginine completely blocks all interactions and thus the downstream transcriptional effects. These findings further support that RadD and Aid1 act independently of each other but are involved in the same intracellular interaction process. The *aid1* gene seems to be required for initial interaction with oral streptococci, since the Δ*aid1* mutant displays decreased coaggregation with *S. sanguinis* ability (Figure 4, Table 1). Downregulation of *aid1* occurs over time and at the time of initial contact with *S. sanguinis* the protein is still present in its unchanged amounts in the cell, therefore in the coaggregation assay lack of Aid1 leads to reduction in binding ability. Based on our data Aid1 appears to allow *F. nucleatum* cells to distinguish between different Gram-positive species. RadD, while being the primary adhesin, does not possess the necessary specificity to distinguish between different Gram-positive species including different streptococci and the interactions rely on additional proteins that facilitate cell-cell recognition and further attachment.

The interactions between different bacterial species are important in the formation and maintenance of oral biofilms. The oral biofilm is a structured bacterial community of cells growing attached directly to tooth and tissue surfaces (ten Cate 2006). Within the oral biofilm architecture each species occupies a certain niche, which is crucial in the formation and maintenance of the mature plaque. The importance of the spatial distribution of the different species has been described both *in vitro* (Kolenbrander 1993) and more recently *in vivo* (Zijnge et al. 2010). Gram-positive species, primarily streptococci, usually occupy the basal layers of the biofilm and are able to directly attach to the tooth surface, while *F. nucleatum* is distributed

throughout the biofilm, providing a scaffold by binding different species. To date only the fusobacterial adhesin RadD has been shown to play a role in the formation of a dual species biofilm between *F. nucleatum* and *S. sanguinis* (Kaplan et al. 2009). Our data indicate that *aid1* also appears to play a role in dual-species biofilm architecture of *F. nucleatum* and *S. sanguinis.* The observed spatial distribution of *S. sangunis* in a dual species biofilm with *F. nucleatum* strain overexpressing *aid1* (Figure 5) suggests that the downregulation of *aid1* is necessary for the normal formation of the biofilm involving *F. nucleatum* and oral streptococci and overexpressing the gene disrupts this natural architecture of the oral biofilm and alters the proper distribution of the streptococcal species. This can have potential detrimental effects on the downstream maturation of the biofilm by not allowing other species, primarily late colonizers, to adhere and integrate into the growing plaque. Inactivation of *aid1* had no significant effect on biofilm formation and structure, which is consistent with its role in interactions with oral streptococci and the repression of the gene in the presence of species belonging to this genus.

The *aid1* gene is predicted to encode a small hypothetical protein comprised of 67 amino acids. BLAST searches against available databases revealed that it is highly conserved across oral fusobacterial species, while no significant homology to any gene within other species sequenced to date was apparent. Protein analysis tools such as InterProScan (version 4.8), PredictProtein and SABLE protein structure prediction server predict that Aid1 is membrane-associated based on the presence of a predicted short N-terminal signal sequence, short helices that are indicative of interaction with the cellular membrane as well as a lipoprotein motif at the N-terminal end of the protein followed by a glycine zipper motif. Glycine zipper sequences are common motifs in membrane proteins and are usually found in transmembrane domains (Kim et al. 2005). These domains are also thought to be involved in protein oligomerization (Barwe et al. 2007; Plotkowski et al. 2007), that may allow Aid1 to form a larger protein complex on the bacterial membrane. While Aid1 is the first small hypothetical protein being characterized

in *F. nucleatum* interspecies interactions, another small fusobacterial membrane protein FadA has been previously identified and characterized for its ability to attach to eukaryotic cells (Han et al. 2005). FadA has been shown to oligomerize and thus form a large protein complex via leucine zipper motifs that allows it to act as an adhesin required in attachment and invasion of epithelial cells (Xu et al. 2007).

In summary, the nature of bacterial interactions involving *F. nucleatum* appears to be more complex than the contact between individual adhesins on the cell surfaces of the partner species. It appears that small proteins like Aid1 are able to define and fine-tune the specificity of large adhesins such as RadD. To best of our knowledge *aid1* is currently the first hypothetical fusobacterial protein identified to have a specific role in interspecies interactions of this important bridging organism with other species in the oral cavity. It appears to be required by *F. nucleatum* for recognition and discrimination for Gram-positive oral bacterial species. This study sheds light on the complexity of bacterial interactions in the oral cavity, suggesting that most of these interactions involve more than one adhesin. Bacteria in the oral cavity need mechanisms to identify and select their binding partners to create the large network of interactions that is present in the dental plaque. Other proteins may be involved in recognition of different streptococcal groups as well as specific species. Studies are currently underway to understand the deeper implications of *aid1* in the formation and maintenance of the oral biofilm as well as identification of other proteins involved in specific interactions of *F. nucleatum* with different oral bacterial species.

Acknowledgements

The pHS58 shuttle plasmid was kindly provided by Dr. Susan Kinder Haake. This work is supported by the Whitcome Pre-doctoral Training Grant to AK and NIH/NIDCR grant DE021108 to RL.

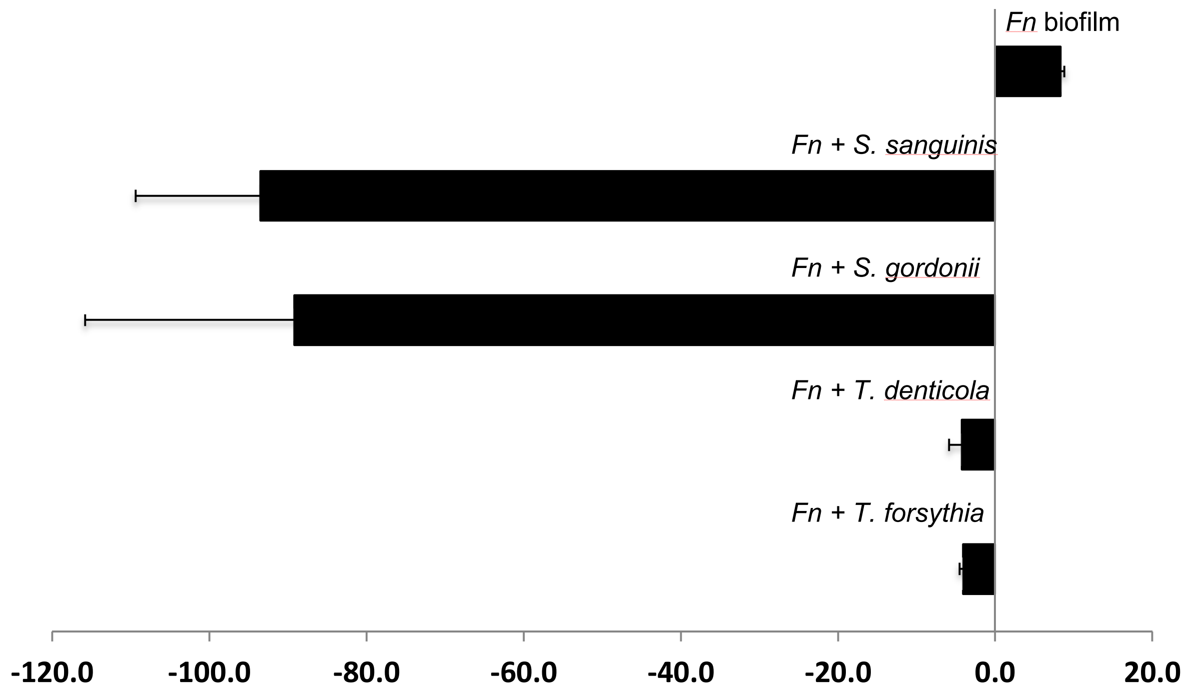


Figure 1. Differential expression of *aid1* in *F. nucleatum* single and dual species biofilms. *F. nucleatum* ATCC23726 cells were grown under biofilm conditions for 24hrs either alone or in the presence of known interacting partners. *aid1* gene expression was analyzed by quantitative real-time PCR (qRT-PCR). Three independent experiments were performed.

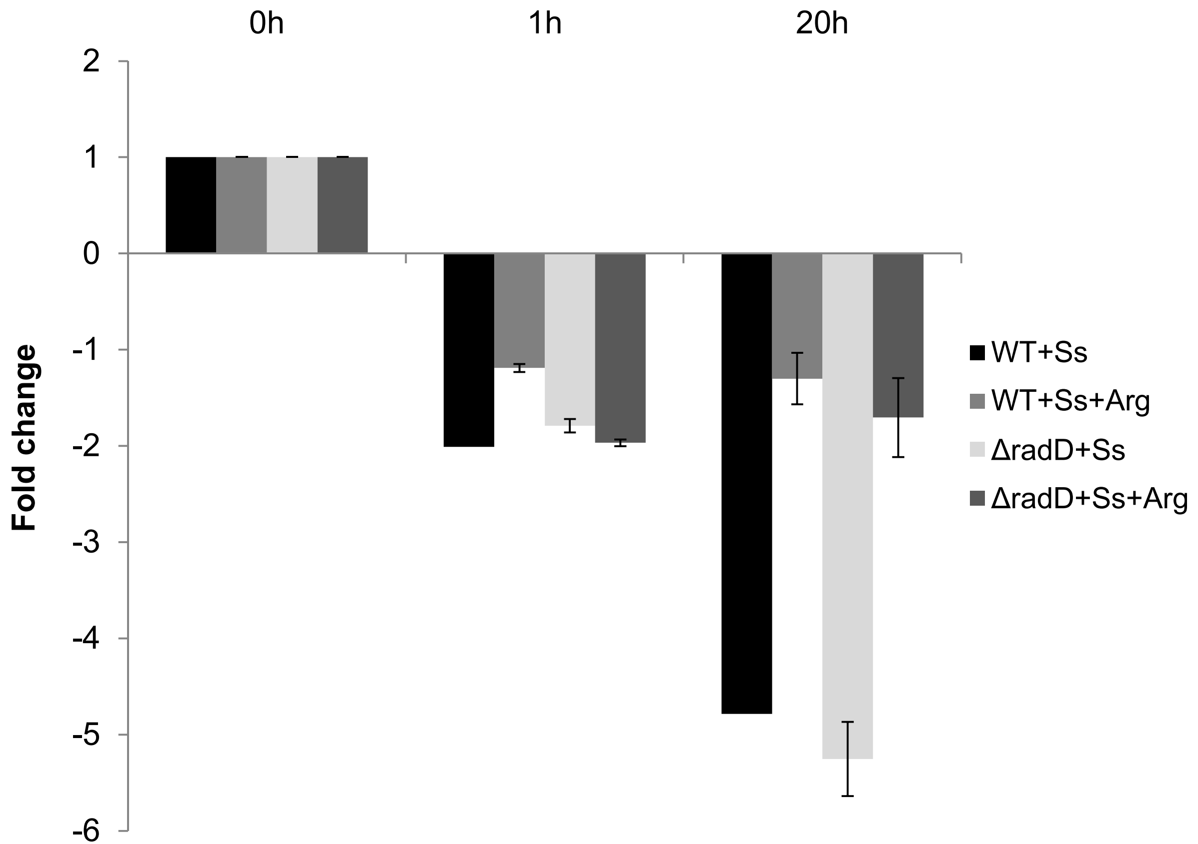


Figure 2. *aid1* gene expression is time- and touch-dependent.

aid1 gene expression was monitored by qRT-PCR at early (1hr) and late (20hr) time points when *F. nucleatum* was coincubated with *S. sanguinis* after being pelleted by centrifugation. Arginine was added as a negative control to block physical interactions between the cells. The data represent the average of three independent experiments.

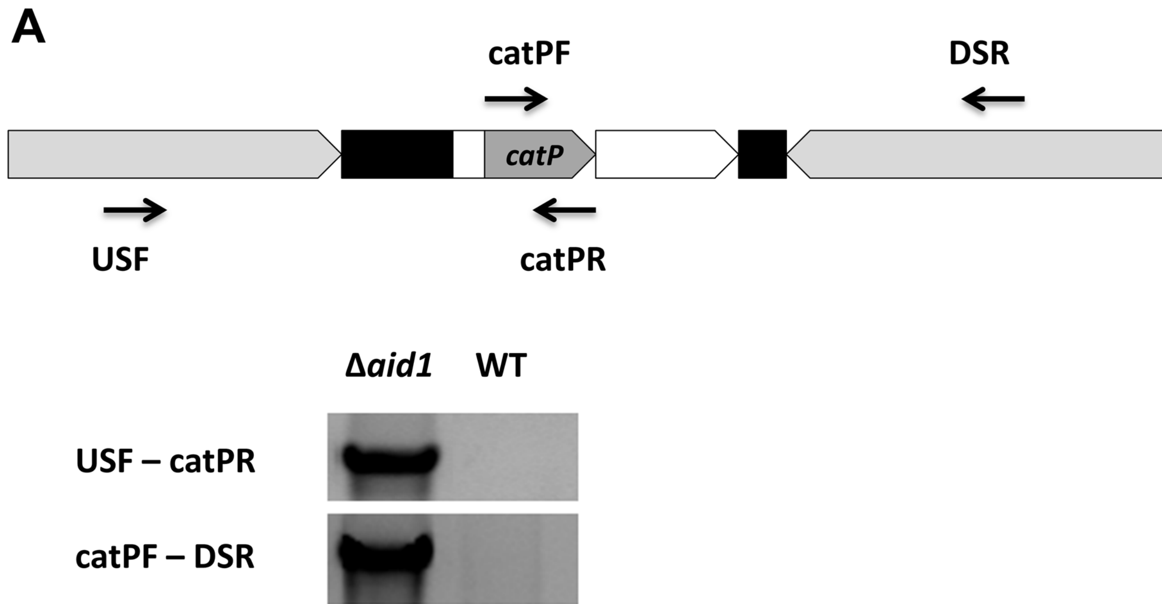
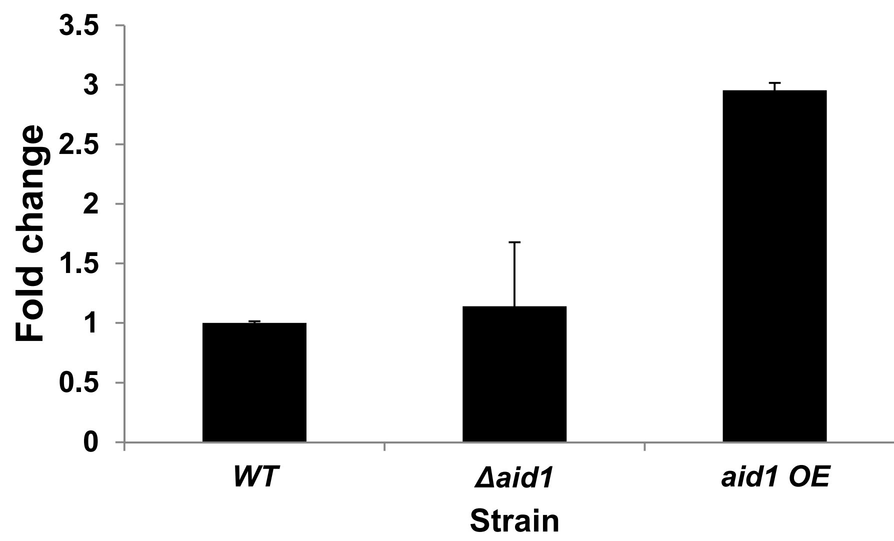


Figure 3. Analysis and verification of *aid1* mutants.

A. Analysis of Δ*aid1* mutant strain*.* Diagram of gene inactivation linear vector introduced into WT *Fn* ATCC 23726 to obtain the Δ*aid1* inactivation mutant. Confirmation of insertion into *aid1* gene by PCR analysis. Arrows indicate the location of the primers used for PCR amplification. Fragments were amplified from the mutant strain but not from the wild-type control. **B. Gene expression of *aid1* in mutant strains and verification of *aid1* overexpression strain.** *aid1* gene expression was verified using quantitative real-time PCR (qRT-PCR) using *aid1*-specific primers.

B



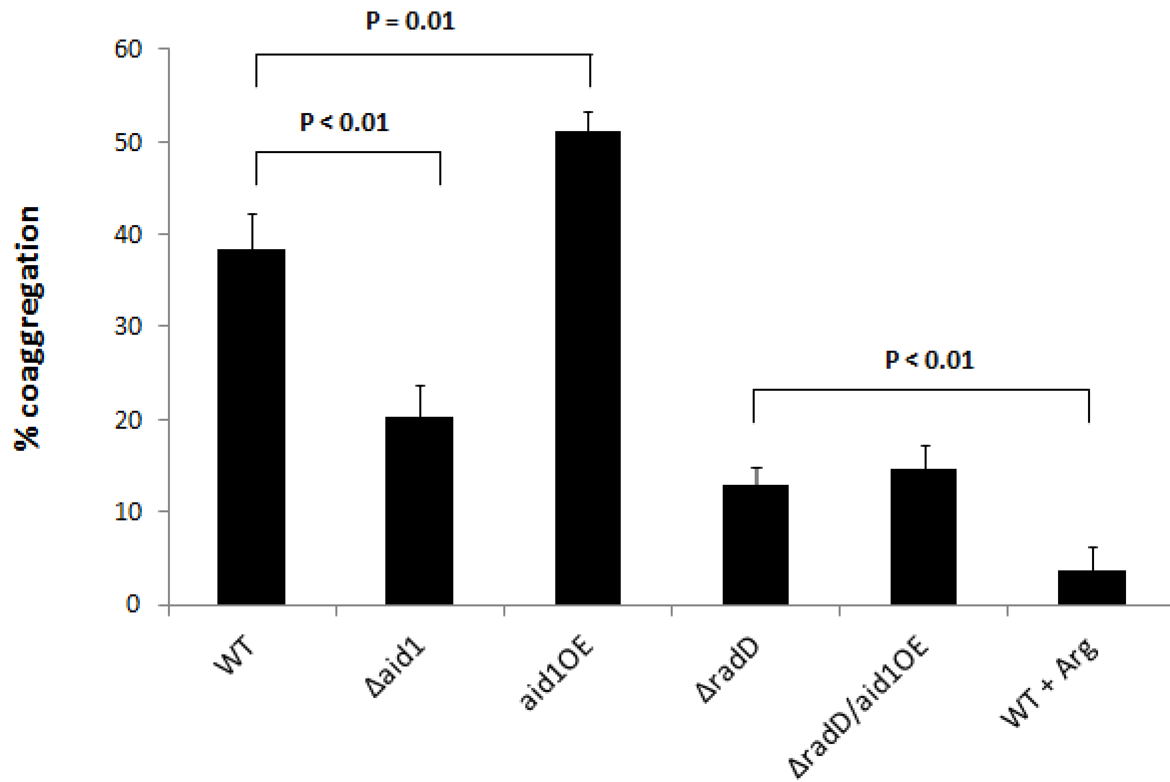


Figure 4. Quantitative coaggregation assay between *F. nucleatum* strains and *S. sanguinis*.

Wild type *F. nucleatum* and *aid1* mutant derivatives were mixed with equal number of *S. sanguinis* cells in coaggregation buffer (CAB) and allowed to aggregate for 10 mins. OD_{600} absorption was measured before and after cells were allowed to aggregate. The data represent the average of three independent experiments.

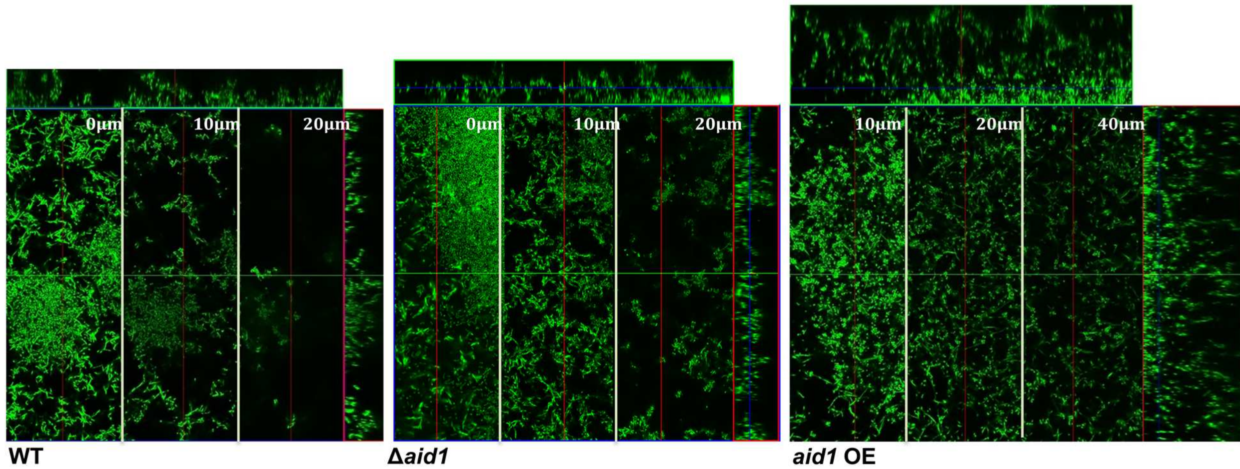
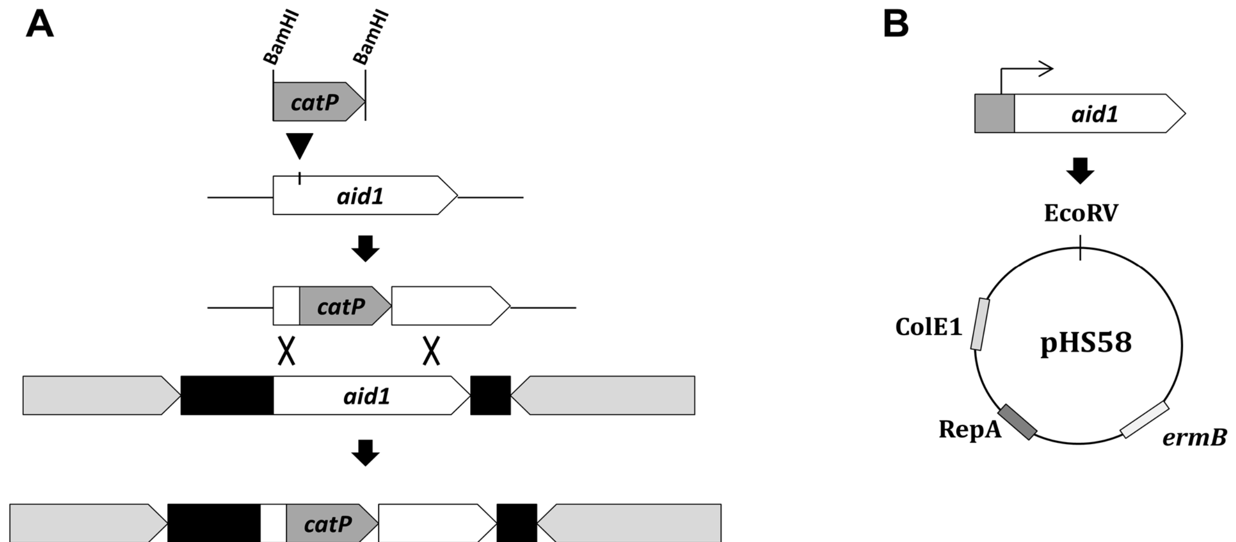


Figure 5. Biofilm growth and morphology of *aid1* mutants with oral streptococci.

Dual species biofilms were grown with wild type *F. nucleatum,* mutants lacking and overexpressing *aid1* together with *S. sangunis.* Biofilms were grown for 24 hrs under anaerobic conditions and stained with SYTO9. Sections from 20µm above the growth surface demonstrate the variation in biofilm morphology as well as the differences in the height of the biofilm. Representative images chosen from three independent experiments are shown.



Supplemental Figure 1. Construction of *aid1* mutant strains.

A. Construction of Δ*aid1* inactivation strain. *aid1* was inactivated via double homologous recombination by inserting a *catP* thiamphenicol/chloramphenicol resistance cassette into the *aid1* gene. **B. Construction of *Fn*/pEP-*aid1* and *radD*/pEP-*aid1* overexpression strains.** *aid1* was overexpressed in wild type *F. nucleatum* ATCC 23726 and in *radD* strains by introducing a shuttle plasmid carrying *aid1* gene with its own endogenous promoter into wild type *F. nucleatum* ATCC 23726 and Δ*radD* strains respectively.

	WT	Δ*aid1*	*aid1* OE	Δ*radD*	Δ*radD/aid1* OE
Auto-aggregation	**0**	**0**	**0**	**0**	**0**
Gram-positive oral					
S. sanguinis	**2**	**1**	**3**	**0**	**0**
S. gordonii	**2**	**1**	**3**	**0**	**0**
E. faecalis	**2**	**1**	**3**	**0**	**0**
L. casei	**2**	**2**	**2**	**0**	**0**
S. epidermidis	**4**	**4**	**4**	**2**	**2**
Gram-negative oral					
P. gingivalis	**3**	**3**	**3**	**3**	**3**
T. denticola	**0**	**0**	**0**	**0**	**0**
T. forsythia	**0**	**0**	**0**	**0**	**0**
V. atypica	**0**	**0**	**0**	**0**	**0**

Table 1. Aggregation ability of *aid1* mutants.

Wild type *F. nucleatum* and *aid1* mutant derivatives were mixed with other species in coaggregation buffer (CAB) and allowed to aggregate for 10 mins. The method of assigning coaggregation scores is described in the experimental procedures. Coaggregation experiments were performed at least five times.

	WT	Δ*aid1*	*aid1* OE
PBS	**2**	**1**	**4**
Arginine	**0**	**0**	**0**
Galactose	**2**	**1**	**3**
GlcNac	**2**	**1**	**4**
Glucose	**2**	**1**	**4**

Table 2. Inhibition of cellular coaggregation.

50mM final concentration of inhibitor was added to coaggregation mix in coaggregation buffer (CAB) and coaggregation was scored after 10 mins. Experiments were repeated independently at least five times.

Chapter 2: Intracellular localization and characterization of the functional domains of Aid1

Abstract

Fusobacterium nucleatum is an opportunistic oral pathogen known for its ability to interact with a large number of different bacterial species colonizing the oral cavity. It possesses a number of outer membrane proteins that are predicted to be involved in adhesion to host as well as other bacterial cells. At the same time, *F. nucleatum* genome reveals presence of a number of lipoproteins that are currently uncharacterized. We have recently characterized a small hypothetical protein Aid1 in *F. nucleatum* to be important for modulating the interaction with oral streptococci. Upon further investigation Aid1 was identified as a novel lipoprotein in *F. nucleatum* which possesses a conserved cysteine that follows a characteristic lipobox motif found in all bacterial lipoproteins. More detailed investigation confirmed localization of Aid1 to the inner leaflet of the outer membrane of the cell. Deletion of the N-terminal signal sequence resulted in no detectable levels of the protein. To the best of our knowledge, this is the first fusobacterial lipoprotein that has been biochemically characterized and described to function in modulation of the adhesion to oral streptococci

Introduction

Fusobacterium nucleatum is an important bridging organism in the oral cavity (Rickard et al. 2003) and has a number of large outer membrane proteins on its outer surface that are thought to play a role in adhesion to host cells and other bacterial species present in the oral biofilm. While many studies have shown the ability of *F. nucleatum* to interact with multiple representatives of the oral microbiome, only one fusobacterial adhesin has been characterized to date (Kaplan et al. 2009). Recent prediction in the sequenced *F. nucleatum* ATCC 25586 strain reveals presence of 27 putative lipoproteins, but others may be present as well (Babu et al. 2006). Strain ATCC 23726 that can be genetically modified and has been used for this study, was recently sequenced but not yet fully annotated, therefore it is not yet known how many predicted lipoproteins are present, but we expect the number to be similar to that of strain ATCC 25586, since these two strains are very closely related. At the same time, none of these lipoproteins are yet characterized. *F. nucleatum* genome analysis reveals the presence of all the enzymes necessary for lipoprotein biogenesis, as well as the LolCDE genes that constitute the lipoprotein transport system across Gram-negative membranes (Kapatral et al. 2002). Genes coding for LolA and LolB chaperones were not found, but it has been shown previously that there's much variation in the sequence of these genes and they are more conserved on the protein structure and functional levels, rather than the genetic level. This suggests that *F. nucleatum* might have genes encoding LolA and LolB that have not yet been identified.

Lipoproteins are an important class of bacterial membrane-associated proteins that are involved in multiple cellular functions, including signal transduction, adhesion, nutrient uptake, transport, antibiotic resistance and virulence (Kovacs-Simon et al. 2011). Two of the three enzymes required for biosynthesis of lipoproteins in Gram-negative bacteria are essential. During biogenesis, lipoproteins are first translated as preprolipoproteins. The preprolipoprotein contains

an N-terminal signal peptide that terminates with a conserved sequence, referred to as the lipobox, [LVI][ASTVI][GAS]C. The terminal cysteine residue is further modified through the attachment of a diacylglycerol by the lipoprotein diacylglyceryl transferase (Lgt), resulting in the prolipoprotein form. Next, lipoprotein signal peptidase (Lsp) cleaves the signal sequence of the lipidated prolipoprotein leaving the cysteine as the new amino-terminal residue. Finally, the cleaved prolipoprotein is modified by attachment of an amide-linked acyl group to the cysteine by lipoprotein N-acyl transferase (Lnt). Lipoproteins in Gram-negative bacteria are triacylated. Once the lipoproteins are modified, they are further localized to their final location in the cell membrane. In *E. coli* most lipoproteins are located at the periplasmic face of the outer membrane, with some attached to the peptidoglycal layer. Lipoproteins in Gram-negative bacteria are targeted to the periplasmic side of the membranes by the lipoprotein localization machinery (Lol). LolCDE complex along with the periplasmic chaperone, LolA are responsible for translocating mature lipoproteins to the outer membrane of a Gram-negative cell (Okuda & Tokuda 2011). The residue adjacent to the conserved cysteine determines whether the lipoprotein will be retained in the inner membrane or transported to the outer membrane. This is known as the "+2" rule and it requires an Asp residue to follow the cysteine in order for the protein to be retained in the inner membrane (Yamaguchi et al. 1988). The presence of any other amino acid results in translocation by the Lol system to the outer membrane. At the same time, it has also been shown that a His residue at +3 position also plays a role in targeting the protein to the outer membrane (Narita & Tokuda 2010). In *E. coli* all of the studied lipoproteins face the periplasmic space, but some Gram-negative bacteria have lipoproteins present on the outer surface of the outer membrane, although the translocation mechanism is not well understood.

The previous chapter identifies Aid1 as a novel hypothetical protein identified in *F. nucleatum* that plays a role in the modulation of interspecies interactions between *F. nucleatum* and oral

streptococci (Kaplan et al. 2014). Amino acid sequence analysis identified the first seventeen residues to constitute a signal sequence of the protein, with a terminal cysteine residue. The presence of a conserved lipobox characteristic of bacterial lipoproteins was also found. These *in silico* findings suggested Aid1 to be a novel fusobacterial lipoprotein. In this chapter we have utilized biochemical approaches to determine the subcellular localization of the Aid1 protein, as well as genetic approaches to investigate the key functional domains of Aid1, required for proper processing of the mature protein.

Materials and Methods

In silico analysis of the Aid1 sequence

Aid1 sequence comprised of 67 amino acids was analyzed for presence of a signal peptide using Singnal IP 4.0 (http://www.cbs.dtu.dk/services/SignalP/). The lipobox motif was verified using LipoP 1.0 server (http://www.cbs.dtu.dk/services/LipoP/) and confirmed using PSORTb v3.0 (http://www.psort.org/psortb/index.html)

Bacterial cell growth conditions

Wild type *F. nucleatum* ATCC 23726 cells were grown in Columbia broth (Difco), mutant derivatives carrying variations of the pHS58 plasmid were grown in Columbia broth supplemented with thiamphenicol (5μg/ml) for selection of mutant derivatives carrying the *catP* gene and clindamycin (0.4 μg/ml) for selection of mutant derivatives carrying the *ermB* gene.

Aid1 antibody generation and Western blot analysis

Anti-Aid1 polyclonal antibody was generated via Genomic Antigen Technology (SDIX) by immunizing rabbits with a proprietary vector carrying the Aid1 truncated sequence containing amino acids 18-67 that constitute the full length of the protein less the signal peptide.

Total cellular protein was isolated from pelleted cells. Briefly, the cells were centrifuged for 10 minutes at 4,600 x *g* and the pellet was resuspended in 2x loading buffer (100mM Tris-HCl pH 6.8, 4% SDS, 20% Glycerol, 5% BME). The solution was centrifuged at 13,000 x *g* to remove cellular debris and the supernatant was collected and boiled for 10 minutes. Samples were stored at -20°C until further use. 10-25μl of sample was loaded into precast Mini-PROTEAN Tris-Tricine gels (BioRad) and separated by SDS-PAGE at constant voltage of 150V for 50 minutes. Proteins were transferred to nitrocellulose membranes (Whatman) at constant

resistance of 400mA for 1 hour. 1μg/ml anti-Aid1 primary rabbit antibody was used to detect Aid1. 1:10,000 dilution of secondary goat-anti-rabbit HRP-conjugated antibody was used for chemiluminescent detection of the signal (Life Technologies). Images were taken using chemiluminescence setting on BioRad Gel Doc.

Mutant construction

Gene sequences carrying the appropriate mutations were synthesized by Integrated DNA Technologies, Inc with phosphorylated ends and cloned into previously cut and dephosphorylated pHS58 plasmid vector. Constructs were transformed into CopyCutter™ EPI400™ (Epicentre) chemically competent *E. coli* cells and selected with 300μg/ml erythromycin on LB plates. Plasmids recovered from the *E. coli* host were confirmed by PCR and sequencing and transformed into wildtype and ΔAid1 *F. nucleatum* strains. Transformants were selected on Columbia blood agar plates containing 0.4μg/ml clindamycin .

Bacterial membrane isolation

Wild type *F. nucleatum* cells were grown overnight in 100ml of Columbia broth and collected via centrifugation at ~8,000*g* at 4°C for 10 min. Cell pellet was resuspended in French press buffer (20mM Na_3PO_4, 150mM NaCl, pH=7.2) with protease inhibitors (Roche Dignostics, Indianapolis, IN) and passed through French pressure cell 3-4 times at 12,000 psi. Unbroken cells and debris were removed by centrifugation at 10,000*g* at 4°C for 10 min. Supernatant was collected and membranes were pelleted by ultracentrifugation at 150,000*g* at 4°C for 60 minutes. The pellet was resuspended in French press buffer and stored in -80°C until use.

Sucrose gradient centrifugation

Sucrose cushion was prepared by overlaying 500μl of sucrose solutions in decreasing concentration starting with the 60% sucrose solution on the bottom and ending with the 30%

solution on the top of the tube. Solubilized membranes were overlayed on top of the sucrose cushion and carefully transferred into the ultracentrifuge. Membranes were centrifuged at 190,000*g* at 4°C for 48 hours. After centrifugation 300µl aliquots were removed into separate tubes and the sucrose concentration of each sample was verified using a refractometer.

Protease sensitivity assay

Protease sensitivity assays were performed as described previously (Arambula et al. 2013) wish minor modifications. Briefly, wild-type *F. nucleatum* cells were cultured in Columbia broth to an OD_{600} of 1.0 and harvested by centrifugation. Resulting pellets were washed with PBS (pH 7.3) supplemented with 5mM MgCl. Cells were normalized by OD and treated with 0, 50, 100, 200 and 400 µg/mL of proteinase K (Sigma) for 1 hour at room temperature. Next, cells were harvested by centrifugation and washed with PBS + 5mM MgCl. Cell pellets were solubilized and total cell proteins were separated by SDS-PAGE and Aid1 was detected by Western blot using α-Aid1 polyclonal antibody.

Results

Aid1 is a predicted lipoprotein

To identify conserved motifs within Aid1 protein sequence we used a number of sequence prediction softwares to analyze the 67 amino acid sequence of the protein. Analysis using Signal IP 4.1 revealed the presence of a 17aa signal peptide with a terminal lipobox sequence [LXXC], which is highly conserved across prokaryotic lipoproteins. A terminal cysteine necessary for triacylation of lipoproteins is also present on the N-terminus of the predicted mature cleaved form of the protein. Presence of a histidine residue at +3 position is predictive of the protein being localized to the outer membrane along with the absence of a retention signal aspartic acid at position +2 and presence of a threonine residue instead. (Figure1).

Generation of the α-Aid1 genomic antibody

Due to highly hydrophobic nature of Aid1 (pK ~ 10) attempts to isolate or chemically synthesize the protein were unsuccessful. Therefore, we used the Genomic Antigen Technology for generating α-Aid1 in rabbits via DNA immunization. The antibody generation was performed by SDIX, Inc using proprietary protocols. The antibody titers achieved were 0.24-0.38 ng/mL. We tested the sera against total fusobacterial proteins and there was no cross reactivity in the small protein range relevant for Aid1 detection (data not shown). The α-Aid1 rabbit genomic antibody (rGAT) was used in all Western blot analyses and specifically detected the Aid1 protein (Figure 2A).

Aid1 localizes to the inner leaflet of the outer membrane of *F. nucleatum*

Next, we utilized biochemical approaches to investigate the cellular localization of Aid1 in *F. nucleatum*. Wild type cells were fractionated into cytoplasmic and membrane fractions using a French pressure cell. Proteins were separated using SDS-PAGE. Western blot analysis

revealed presence of Aid1 in the membrane fraction (Figure 2A). Cytoplasmic fraction also retained a small portion of the Aid1 protein. In the total cellular fraction total Aid1 was separated into two bands, reflecting the presence of two distinct species of protein in the cell (Figure 2A). We further performed sucrose gradient centrifugation to separate inner and outer membranes of wild type *F. nucleatum* cells. Western blot revealed the majority of the protein localized to the outer membrane fraction of the cell (Figure 2B), which was predicted from the protein sequence. Next, the cells were treated with increasing concentrations of pronase or proteinase K in order to degrade surface proteins. Even at the highest concentration of the enzymes, the Aid1 band was present on a Western blot in unchanged amount suggesting that the protein localizes to the inner leaflet of the outer membrane and is protected from enzymatic digest (Figure 3).

Dissection of key domains of the Aid1 protein

In order to further characterize Aid1 as a lipoprotein, we used a genetic approach to dissect the key domains of Aid1 that are required for localization and proper trafficking of the protein. In order to characterize the different functional domains of the protein, structural mutants were constructed. Amino acids 1-17 comprising the signal sequence of the lipoprotein were deleted and expressed on the pHS58 shuttle vector in the Δ*aid1* background strain. Similarly, a C18A mutation was introduced into the Aid1 sequence and also expressed in the Δ*aid1* background. Both *F. nucleatum* strains carrying the mutant versions of Aid1, did not produce any detectable levels of protein. At the same time, small amounts of protein were detected in the *E. coli* strains carrying the shuttle vectors (Figure 4). The C18A mutation caused a significant decrease in detectable protein levels, and the detected protein band appeared to be of slightly larger size as shown by Western blot. Deletion of the signal sequence, which should lead to the retention of the protein in the cytoplasm, resulted in no detectable protein levels in both *F. nucleatum* and *E. coli* strains even though there were detectable mRNA levels of *aid1* in all strains (data not shown).

Discussion

In this chapter we characterize Aid1 on a biochemical level, describing its subcellular localization as well as dissecting the functional domains necessary for proper processing and maturation of the protein. *In silico* analysis revealed a conserved lipobox motif ILVGC, that includes a terminal cysteine residue. The first 18 amino acids of the protein were predicted to contain the signal peptide by analysis using SignalP 4.1 (Figure 1). Furthermore, the cysteine residue was followed by a Thr and His residues at the +2 and +3 positions. The lack of an Asp at the +2 position and the presence of a His at the +3 position predict the protein to be translocated to the outer membrane of the cell based on previous studies done in *E. coli* as a model Gram-negative organism (Okuda & Tokuda 2011; Narita & Tokuda 2006).

With this information in mind, we investigated the cellular localization of Aid1 in *F. nucleatum.* A specific α-Aid1 antibody was developed through the rabbit Genomic Antigen Technology (rGAT) and it was highly specific and sensitive for Aid1 as shown by Western blot of cell lysates (Figure 2). In the total cell lysates Aid1 appeared as two distinct bands, suggesting that there are two different protein species present in the cell. We suspect that the two bands represent the preprolipoprotein and the prolipoprotein/lipoprotein forms of Aid1. At the same time, the cytoplasmic and membrane fractions only show the presence of a single band, which is expected since the preprolipoprotein form is only found in the cytoplasm and not stable in the membrane as it undergoes rapid modification. Interestingly, the band that is found in the membrane fraction and predicted to be the cleaved, mature lipoprotein form, runs higher than the species found in the cytoplasm which is predicted to be the longer, preprolipoprotein form. While this is unusual, it has been previously reported to also be the case for another small fusobacterial protein FadA that has both cleaved and uncleaved versions (Xu et al. 2007). This aberrant migration pattern can be due to the hydrophobic signal peptide present in the

preprolipoprotein form binding more SDS and resulting in a faster migration pattern. Also, the shape of the two molecules may differ, resulting in this unexpected migration pattern. Presence of different post-translational modifications can be another possible explanation, although unlikely from what is known about lipoprotein processing in the cell.

Aid1 localized to the outer membrane fraction in sucrose gradient fractionation experiments (Figure 2B) and was not degraded in the protease sensitivity assays (Figure 3B) suggesting that the mature protein faces the periplasmic space and is anchored to the inner leaflet of the outer membrane. Next, we analyzed the key functional domains of the protein by mutating the cysteine residue to an alanine, as well as constructing a truncated version of the protein that lacks the signal peptide. Without the signal peptide the protein is predicted to be retained in the cytoplasm, as it cannot be targeted to the cell membrane. A mutation in the cysteine residue is predicted to result in accumulation of the protein in the inner membrane since Lnt cannot recognize the lipoprotein for further processing. When the plasmids expressing the mutated versions of Aid1 were expressed in *F. nucleatum,* no detectable amounts of protein were observed on Western blot (data not shown). Since this is a membrane-associated protein, aberrant accumulation of the mutated protein in the compartments other than the outer membrane could lead to its rapid turnover and degradation. Western blots performed on the *E. coli* strains carrying the shuttle vectors shows a highly reduced amount of protein in the C18A mutant and no detectable amount of protein in the Δ1-17 truncated mutant. The protein band of the Aid1-C18A strain migrates slower than that of the wild-type, most likely due to it being the uncleaved form of the protein. The reduced amount of the protein is most likely due to degradation. The C18A mutation is predicted to trap the protein in the inner membrane and the cellular machinery likely keeps the protein levels to a minimum to prevent toxicity to the cell. Same may be true for the Δ1-17 truncated version of the protein. Accumulation of the mutant form of Aid1 in the cytosol may not be well- tolerated by the cell and thus trigger a rapid turnover

of the protein. We tested the mRNA levels of both constructs and the gene is being expressed at approximately 3-4 fold higher levels (data not shown) , yet there is no detectable amount of protein present in the cell.

Since Aid1 does not appear to be exposed on the outer surface of the cell, it is not likely to be an adhesin itself. Many lipoproteins play roles in signal transduction or modulation of transport systems (Kovacs-Simon et al. 2011). The co-aggregation phenotype described in chapter one is most likely due to an indirect role that Aid1 plays in modulation of the adhesion to oral streptococci. Genome sequence analysis reveals the presence of an operon upstream of Aid1 that encodes a putative C4 dicarboxylate transport system, which is a member of the tripartite ABC periplasmic transporter (TRAP) family (Fischer et al. 2010). In Gram-positive bacteria a number of transporters have been shown to use lipoproteins as their periplasmic binding proteins (Sutcliffe & Russell 1995). Also, recent studies have identified a novel lipoprotein that is involved in a TRAP transport system in *Treponema pallidum* (Deka et al. 2012). Very little is known about the TRAP family of transporters, since they were recently discovered and are not yet well-characterized (Fischer et al. 2010). According to its genome sequence *F. nucleatum* possesses many features of a Gram-positive bacterium, although morphologically it resembles a Gram-negative species (Kapatral et al. 2002). Several characterized lipoproteins in Gram-positive bacteria have been shown to serve as binding proteins for molecules that are being transported through ABC-type transporter systems (Cockayne et al. 1998; Dintilhac & Claverys 1997). Therefore it is possible that Aid1 is a lipoprotein that is part of the upstream TRAP transporter system and plays a role in binding the transported substrate. Some ABC transporters have also been implicated in adhesion mechanisms, so Aid1 can be indirectly involved in binding to oral streptococci. While the distinct function of Aid1 has not yet been elucidated and requires further investigation, we have been able to characterize its role in modulating the adhesion and the cellular localization of the protein.

Figure 1. Aid1 is a predicted outer membrane lipoprotein

Amino acid sequence of Aid1 showing the N-terminal 17 amino acid signal sequence (black), followed by a conserved cysteine residue (top arrow). The signal sequence also contains the LXXC lipobox motif directly preceding the Cys residue. Thr and His residues that follow the Cys residue constitute the outer membrane sorting signal (bottom arrow).

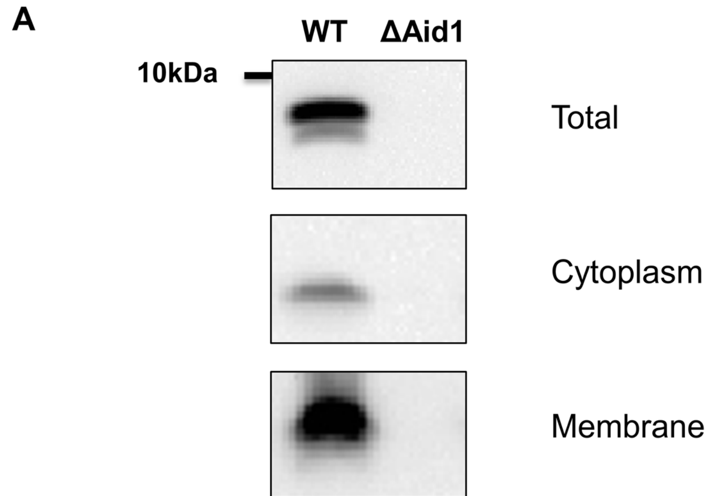
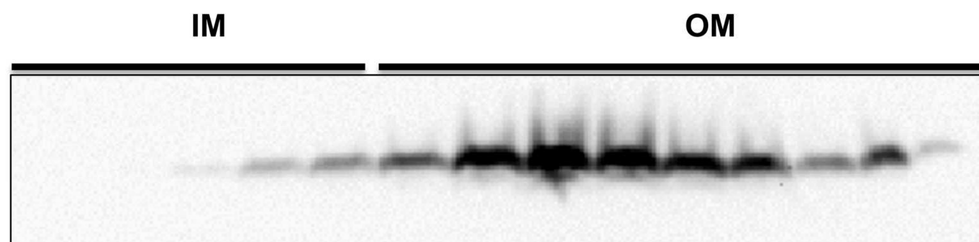


Figure 2. Aid1 localizes to the outer membrane of the cell

A. Aid1 is found in the membrane fraction of *F. nucleatum.* Wild-type and ΔAid1 *F. nucleatum* cells were collected and fractionated into cytoplasmic and membrane fractions using a French pressure cell. Proteins were separated by SDS-PAGE and subjected to Western blot analysis. Aid1 protein was detected using an α-Aid1 polyclonal antibody. **B. Aid1 localizes to the outer membrane of the cell.** Total cell membranes were fractionated using sucrose gradient centrifugation. Aid1 primarily localized to the outer membrane (OM) fractions as shown by Western blot analysis.

B



Fraction	2	4	6	8	10	12	14
% sucrose	32	38	43	47	51	54	56

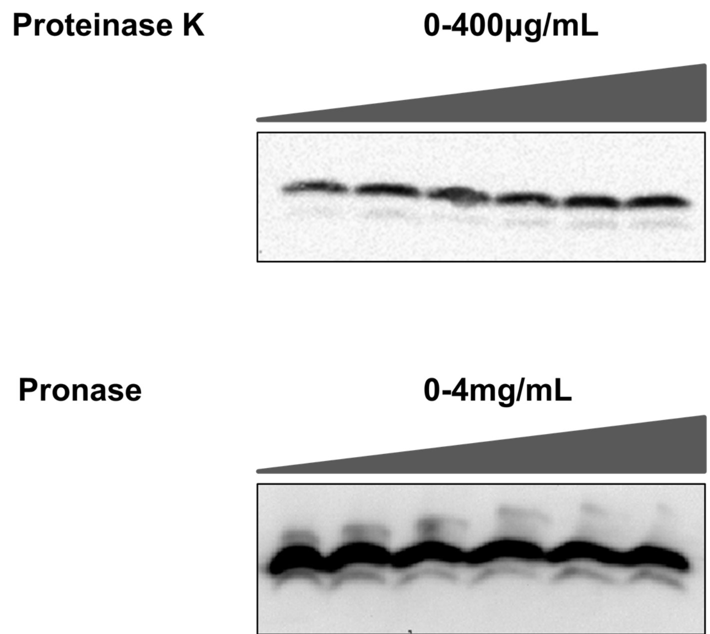


Figure 3. Aid1 is protected from degradation by enzymatic digest

Increasing concentrations of Proteinase K (0-400μg/mL) and Pronase (0-4mg/mL) were used to treat wild-type *F. nucleatum* cells. Total cell protein was analyzed by Western blot for the presence of Aid1.

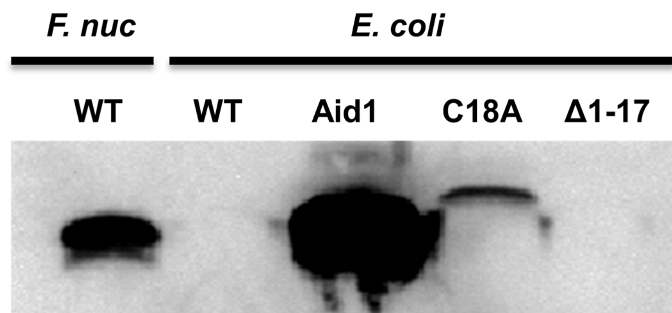


Figure 4. Aid1 structural mutants are unstable in the cell

Aid1 C18A and Δ1-17 mutants were expressed in *E. coli* cells and analyzed by Western blot. Wild-type *F. nucleatum* was used as a control.

Chapter 3: Characterization of transcriptional and translational regulation of *aid1*

Abstract

Aid1 is the first lipoprotein described in the opportunistic oral pathogen *Fusobacterium nucleatum* and it has been shown to play a role in modulation of adhesion to oral streptococci. The gene encoding this protein is downregulated upon *F. nucleatum* binding to oral streptococcal species as early shortly after initial contact, and the downregulation increases over time. At the same time, *aid1* is upregulated in *F. nucleatum* single species biofilms. Thus the ability of the cell to regulate the gene expression and protein production of *aid1* appears to be of great importance during growth and maturation of the oral biofilm. Using expression analysis tools we investigated the genetic elements surrounding *aid1* to determine their role in the regulation of *aid1* gene expression. Regulation of Aid1 protein production was also analyzed and found to require different genetic elements. We have mapped the minimum necessary *cis*-element required for downregulation of *aid1* in the presence of *Streptococcus sanguinis* as representative partner strain, and have described the genetic elements responsible for Aid1 protein production. This work illustrates an example of decoupled transcriptional and translational regulation of *aid1* and suggests the presence of additional mechanisms responsible for maintaining levels of the Aid1 protein according to environmental changes.

Introduction

The importance of differential gene regulation in oral bacteria, particularly in the context of multispecies communities, has been gaining increasing appreciation in recent years. Bacteria within the biofilm are constantly communicating with representatives of their own as well as different species. The presence of different microbial species, together with environmental cues, affects gene expression of individual cells as well as the overall bacterial community. The physical interaction of bacteria has been studied *in vitro* through coaggregation assays, where equal numbers of cells are mixed together and allowed to form physical aggregates. Dual species biofilms also provide good representation of interactions between two species as they are allowed to grow together and form interactions that resemble those in the more complex dental plaque. Previous studies have mostly focused on analyzing contact-dependent gene regulation in some of the interacting partners of *F. nucleatum,* such as *S. gordonii* (Jakubovics et al. 2008), *P. gingivalis* (Simionato et al. 2006), and *Actinomyces oris* (Frias et al. 2001). Transcriptional responses have also been described in *F. nucleatum* cells during autoaggregation (Merritt et al. 2009). Our efforts have focused on understanding the changes in gene expression in *F. nucleatum* upon binding to other species, and their underlying molecular mechanisms. In a previous study we performed microarray analysis on *F. nucleatum* cells grown together in dual species biofilms with representatives of both early and late colonizers (McHardy 2011) and described the changes in gene expression affected by the presence of these partner species. While, the microarray and transcriptomic analyses serves as an initial step in understanding the differential gene regulation, they do not provide insights into molecular mechanisms of these changes and further characterization is necessary.

Bacterial gene expression and regulation has been extensively studied in many species, with the majority of the work being performed in *E. coli* and *B. subtilis.* Bacteria utilize a number of regulatory mechanisms to control their gene expression in response to environmental cues (Ptashne & Gann 2002). Transcription initiation as well as posttranscriptional control mechanisms have been identified and described in a number of bacterial species. Prokaryotes can regulate transcription elongation and termination, translation initiation and termination as well as mRNA stability (Babitzke 2004; Condon 2003; Copeland 2003; Storz et al. 2004; Winkler et al. 2004). Although much is currently known about the mechanisms of prokaryotic transcription and translation, many factors affecting mRNA stability and degradation are still poorly understood. In the recent years a number of *cis*- and *trans*-acting regulatory RNAs have also been described (Gottesman 2004; Storz et al. 2004). Numerous recent reports have shown presence of multiple non-coding RNAs, called riboswitches, that can receive intracellular metabolic signals and regulate gene expression (Winkler & Breaker 2005). These signals trigger structural changes of the RNA, which consequently control regulation of associated genes. Bacterial genes are often organized in polycistronic operons, where multiple genes can be transcribed and translated at the same time. The organization into operons allow the cells to coordinate simultaneous expression of genes. But often, different gene products are needed in various stoichiometric amounts and additional steps are required for modulating the amounts of gene product produced by these genes from an operon. This step can be regulated at the level of translation but also mRNA stability and is referred to as modulation.

At the same time, very little is known about the gene regulatory mechanisms in the oral opportunistic pathogen *F. nucleatum.* The lack of genetic tools and challenges transforming the cells with plasmid DNA constructs have significantly hindered progress of characterization of many essential functions in *F. nucleatum* on a molecular level. A number of studies have

attempted at characterizing the regulation of certain genes in *F. nucleatum.* Serganov and colleagues have recently described crystal structures of an *F. nucleatum* riboswitch bound to Flavin-mononucleotide, riboflavin and antibiotic roseoflavin and have discussed the molecular mechanisms of FMN-based control of gene expression in *F. nucleatum* as well as other bacterial strains (Serganov et al. 2009). Other regulatory RNAs that regulate riboflavin gene expression (RFN elements) in *F. nucleatum* were found by Vitreschak and colleagues (Vitreschak et al. 2002). These studies provide great insight into some of the molecular mechanisms of gene regulation in this bacterium, yet do not provide a great overview of different mechanisms of transcriptional and translational regulation in *F. nucleatum.* This study delves into characterization of the gene regulatory mechanisms of a novel fusobacterial gene and adds to our existing knowledge of fusobacterial gene regulation.

Aid1 is the first fusobacterial lipoprotein that is being characterized on a molecular level, and this chapter focuses on a detailed molecular characterization of the genetic elements necessary for the differential regulation of the gene in the presence of *S. sanguinis.* A *cis*-element necessary for the downregulation of *aid1* in the presence of *S. sanguinis* is identified and described. Based on our results and observations, we propose that *aid1* gene expression and Aid1 protein production are differentially regulated and decoupled from one another.

Materials and Methods

Bacterial strains and culture conditions

F. nucleatum strains were grown on Columbia agar plates supplemented with 5% sheep blood or in Columbia broth (Difco, Detroit, MI, USA) under anaerobic conditions (5% CO_2, 5% H_2, 90% N_2). Thiamphenicol (MP Biomedicals, Irvine, CA, USA) at 5μg/ml was used for selection and maintenance of strains containing the *catP* determinant. Clindamycin (MP Biomedicals, Irvine, CA, USA) at 0.4μg/ml was used for selection and maintenance of strains possessing the *ermB* determinant. *Streptococcus sanguinis* ATCC 10556 was grown anaerobically in Todd-Hewitt (TH) broth (BD Difco, Detroit, MI, USA) at 37°C.

Construction of *aid1* promoter strains

Strains expressing *fipA*pro-*aid1* and *aid1*(S) were constructing by ligating synthesized gene sequences into pHS58 shuttle vector. Gene sequences were synthesized by Integrated DNA Technologies, Inc (IDT). The *fipA*pro-*aid1* sequence is a direct fusion of the *fipA* promoter sequence from *F. nucleatum* and full sequence of the *aid1* gene. The *aid1*(S) sequence included the full *aid1* sequence in addition to 291bp of the gene's upstream region. The total sequence length for both constructs was 498bp. Fragments were synthesized by IDT DNA with phosphorylated ends and ligated into digested and dephosphorylated pHS58 shuttle vector (Kaplan et al. 2014) and transformed into CopyCutter™ EPI400™ *E. coli* strain (Epicentre, Madison, WI, USA) and selected on 300mg/ml erythromycin LB plates. The *aid1*(L) expressing strain was constructed by ligation of a PCR-amplified fragment into pHS58 shuttle vector and selection on 300mg/ml erythromycin LB plates. Primers used for PCR amplification of the fragment are listed in Table 1. Plasmids were isolated using Wizard® *Plus* SV Miniprep DNA

Purification Kit (Promega, Madison, WI, USA) and used for electroporation into the Δ*aid1 F. nucleatum* strain using standard electroporation protocols as described previously (Kaplan et al. 2014). Strains were recovered overnight and plated on Columbia agar (Difco, Detroit, MI, USA) supplemented with 0.4μg/ml clindamycin (MP Biomedicals, Irvine, CA, USA) and incubated for 3-5 days under anaerobic conditions (5% CO_2, 5% H_2, 90% N_2).

Bacterial co-cultures

Equal numbers of *F. nucleatum* ATCC 23726 and *S. sanguinis* cells from overnight cultures were mixed together in a tube and pelleted for 10 minutes at 4600 x *g.* Pellets were incubated anaerobically at 37°C for 4 or 24 hours unless noted otherwise. After incubation cells were pelleted again for 5 minutes at 4600 x *g,* medium was removed and the pellets were frozen at -80°C prior to RNA isolation.

Transcriptional Analysis

Genomic DNA was extracted from stationary-phase cells following standard protocols and used for generating reference standard curves. Total RNA was extracted using the High-Pure RNA Isolation Kit (Roche, Palo Alto, CA, USA) according to manufacturer's instructions. One μg of total RNA served as template for cDNA synthesis with the PrimeScript 1st strand cDNA Syntesis Kit (Takara, Otsu, Shiga Japan) following the manufacturer's protocol. For qRT-PCR, SYBR green (Bio-Rad, Hercules, CA, USA) was used for fluorescence detection with the MyiQ real-time PCR system (Bio-Rad), according to the manufacturer's instructions. Expression levels of 16S rRNA were determined for normalization of the qRT-PCR data. Primers used for qRT-PCR analysis are listed in Table 1.

Operon analysis

Primers amplifying the intergenic region between the FN1254 and *aid1* genes were used for PCR amplification. Primer sequences are listed in Table 1. *F. nucleatum* genomic DNA (gDNA) was isolated using standard isopropanol precipitation technique. RNA isolation and cDNA generation was performed as described above. PCR amplification was carried out for 30 cycles using the Phusion® High Fidelity Master Mix (New England BioLabs, Ipswich, MA, USA) at 55°C primer melting temperature.

Western blot

Total cellular protein was isolated from pelleted cells as described in Chapter 2. Briefly, the cells were centrifuged for 10 minutes at 4,600*g* and the pellet was resuspended in 2x loading buffer (100mM Tris-HCl pH 6.8, 4% SDS, 20% Glycerol, 5% BME). The suspension was centrifuged at 13,000*g* to remove cellular debris and the supernatant was collected and boiled for 10 minutes. Samples were stored at -20°C until further use. 10-25µl of sample was loaded into precast Mini-PROTEAN Tris-Tricine gels (BioRad, Hercules, California, USA) and ran at a constant 150V voltage for 50 minutes. Proteins were transferred to nitrocellulose membranes (Whatman, Dassel, Germany) at a constant resistance of 400mA for 1 hour. One µg/ml anti-Aid1 primary rabbit antibody was used to detect Aid1. A 1:10,000 dilution of the secondary goat-anti-rabbit HRP-conjugated antibody was used for chemiluminescent detection of the signal (Life Technologies, Grand Island, NY, USA). Images were taken using chemiluminescence setting on the BioRad Gel Doc.

Results

The endogenous promoter of *aid1* is necessary and sufficient for downregulation of the gene in the presence of *S. sanguinis*

Our previous study has shown that *aid1* expression is downregulated when *F. nucleatum* is co-incubated with *S. sanguinis* (Kaplan et al. 2014). To characterize the molecular mechanisms of this regulation we investigated the *aid1* promoter region in order to identify the minimum necessary genetic components for regulation of the *aid1* gene. Plasmids carrying fragments of different lengths containing *aid1* and upstream sequences were constructed by ligating these fragments into the pHS58 shuttle vector (Figure 1), followed by transformation into *Δaid1 F. nucleatum*. The resulting mutant strains were further used for characterization of *aid1* regulation. Relative gene expression of *aid1* in planktonic cell cultures was measured using qRT-PCR using *aid1*-specific primers (Table 1). The *aid1*(S) and *aid1*(M) constructs (Kaplan et al. 2014) showed approximately 3-fold increase in mRNA production in planktonic cells, while the *aid1* (L) containing strain exhibited a much higher overexpression of the gene with a ~20-fold increase in mRNA levels in planktonic cells (Figure 2). When the respective mutant *F. nucleatum* strains were co-incubated with *S. sanguinis*, we observed downregulation of *aid1* in all strains that carried the endogenous *aid1* promoter, but not in the *fipA*pro-*aid1* strain that expresses *aid1* under the exogenous constitutively active *fipA* promoter (Figure 3).

Aid1 protein production requires transcription from the FN1254 promoter

After observing significant differences in the mRNA production level of the different *aid1* strains, we investigated the Aid1 protein production in these cells. Aid1 protein levels were analyzed via Western blotting using a polyclonal α-Aid1 antibody. Strains carrying the Aid1(S) and (M)

constructs did not produce any significant levels of Aid1 protein. Even though residual protein was detected, the levels were drastically lower compared to wild-type. Only the strain transformed with Aid1(L) produced a significant amount of protein that appeared to be even higher than in wild-type *F. nucleatum* (Figure 4A).

Since the *aid1*(L) carrying strain is the only one that contains the full FN1254 sequence on the shuttle vector, we performed an operon analysis on the cDNA isolated from all the *aid1* promoter strains by amplifying the intergenic region between FN1254 and *aid1*. Primers used for the analysis are listed in Table 1. Genomic DNA (gDNA) and cDNA samples from planktonic cells were used for the operon analysis. The gDNA sample, along with the wild-type cDNA and *aid1* (L) samples resulted in a distinct band of the expected size of ~350bp (Figure 4B). All other strains that did not contain the FN1254 sequence in its entirety did not show presence of a band from PCR analysis. Operon analysis was also performed on cDNA collected from *F. nucleatum* strains that were co-incubated with *S. sanguinis* and similar results were observed (data not shown).

FN1254 gene expression is regulated in a similar manner to *aid1*

After observing the co-transcription of FN1254 and *aid1* in *F. nucleatum*, we investigated the expression levels and regulation of FN1254. *aid1* mRNA levels are slightly increased in the strain containing *aid1*(S) (Figure 2) and significantly increased in the strain carrying *aid1*(L). We analyzed the gene expression levels of FN1254 in the same strains using qRT-PCR. FN1254 expression levels were similar to gene expression levels of *aid1.* In the strain expressing *aid1* from the plasmid *aid1*(S), mRNA levels from the chromosomal copy of FN1254 were approximately 3.5 fold higher than in the wild-type. The strain carrying *aid1*(L), in which FN1254

is encoded on the plasmid as well as the chromosome, produced ~13-fold higher FN1254 mRNA levels which is similar to the high expression level of *aid1* in the same strain (Figure 5).

Furthermore, we sought to investigate whether FN1254 is regulated in a similar manner to *aid1. F. nucleatum* strains carrying the *aid1* promoter constructs were co-incubated with *S. sanguinis* for 4 hours and gene expression levels of *aid1* and FN1254 were analyzed using qRT-PCR. FN1254 was downregulated in both the *aid1*(S) and *aid1*(L) strains very similar to *aid1.* The FN1254 gene expression was reduced by approximately 2.5 fold in the wild-type as well as both mutant strains, while *aid1* was consistently downregulated ~3-3.5 fold in all strains (Figure 6).

Discussion

The ORF encoding Aid1 first caught our attention as the only gene that was differentially regulated in *F. nucleatum* cells grown planktonically or as a biofilm and compared to dual species biofilms in the presence of representatives of both early and late colonizers. Its downregulation was ubiquitous in the presence of both oral streptococci and the late colonizing *T. denticola* and *T. forsythia*, although more pronounced in the presence of *S. sanguinis* and *S. gordonii.* This tight regulation of the *aid1* expression in the presence of other species suggests its importance in interspecies interactions. The DNA regions surrounding *aid1* contain a predicted nitroreductase FN1254, and a predicted TRAP transporter operon transcribed in the same direction as *aid1* upstream of the gene, and a putative membrane antigen downstream that is transcribed in the opposite direction. Sequence analysis reveled the presence of a Shine-Dalgarno (SD) sequence -7 nucleotides upstream of *aid1* and a predicted promoter sequence (data not shown). At the same time, *aid1* can be expressed from plasmid constructs carrying the gene with its endogenous promoter, which suggested that it can be transcribed as a monocistronic fragment (Figure 2). We expressed *aid1* in the Δ*aid1* deletion background on a shuttle vector carrying fragments that contained the gene with different lengths of upstream sequences (Figure1). The *aid1* gene was expressed at approximately 3-fold higher than wild-type levels from the *aid1*(S) and *aid1*(M) constructs but it was highly overexpressed (~20-fold) in the strain carrying *aid1* (L), that contains a fragment including the full length FN1254 gene with the upstream FN1254 promoter sequence (Figure 2). Interestingly, this difference in mRNA levels was also reflective of the Aid1 protein production (Figure 4A). Protein levels in the *aid1*(L) construct were not only easily detectable but also appeared higher than those of the wild-type cells, while the *aid1*(S) and *aid1*(M) strains produced barely detectable levels of Aid1 protein (Figure 4A). This data suggests that while the *aid1* endogenous promoter is sufficient for the

transcription of the gene, and the levels of transcription are approximately ~3-fold higher than that of the wild-type, this promoter does not appear to be as strong as the FN1254 promoter which is able to drive transcription to ~20-fold higher levels than wild-type. This could potentially be due to the FN1254 having the preferred promoter to *aid1* or FN1254pro being a more efficient promoter. It is possible that a yet unidentified *trans*-acting element stabilizes the FN1254 promoter and allows a much more rigorous transcription. At the same time, the short mRNA fragment that is generated does not contain the necessary elements for protein production. Even though *aid1* has a SD sequence, it may not be a strong ribosomal binding site. Often the physical distance of the SD sequence can determine the efficiency of translation (Kozak 2005). It is possible that the *aid1*(L) fragment contains a certain *cis*-element necessary for efficient translation of *aid1* mRNA. The data suggests that the regulatory element is upstream of the FN1254 start site since the *aid1*(M) strain also does not produce any significant amount of Aid1 protein, although it contains almost the entire FN1254 gene sequence (511/515 nucleotides). FN1254 also has a predicted promoter sequence upstream of the gene and a SD sequence at -6 nucleotides from the start site. It is likely that Aid1 protein translation is stabilized by an element that is present in the polycistronic mRNA fragment that is transcribed from the FN1254 promoter which enhances translation. We confirmed the presence of a longer polycistronic mRNA fragment by performing operon analysis PCR using primers that amplify the junction between FN1254 and *aid1*. We observed an amplicon in the wild-type *F. nucleatum* genomic DNA (gDNA), as well as wild-type cDNA and the *aid1*(L) containing strain, which suggests that *aid1* can also be transcribed as a part of a two gene polycistronic messenger (Figure 4B).

Since the length of the sequence preceding *aid1* appears to play a role in mRNA and protein production, we investigated how the different sequences preceding *aid1* affected the regulation of the gene in the presence of *S. sanguinis.* Interestingly, all three constructs that included the

aid1 endogenous promoter followed the same downregulation pattern observed in the wild-type *F. nucleatum* cells. Relative gene expression was reduced by approximately 3-4 fold after 4 hour co-incubation with *S. sanguinis* when compared to baseline mRNA levels for each respective strain. Meanwhile, when the promoter sequence was changed to the foreign constitutive *fipApro* promoter, *aid1* expression was no longer regulated and no downregulation was observed (Figure 3). This data suggests that the *aid1* mRNA levels are regulated irrespective of the size of the messenger in the presence of *S. sanguinis* partner species and that the sequence including the endogenous promoter is necessary and sufficient for downregulation of the gene when *F. nucleatum* is co-incubated with *S. sanguinis.*

Production and downregulation of *aid1* mRNA does not appear to require the presence of FN1254 on the messenger, yet the polycistronic mRNA species containing both FN1254 and *aid1* is necessary for Aid1 protein production. To investigate the role of FN1254 in the regulation of *aid1* translation and transcription, we next looked at the regulation of FN1254 itself. First, we analyzed the expression of FN1254 in planktonic cells and observed a very similar change in expression level of the gene compared to *aid1.* In the *aid1(*L) strain FN1254 expression was approximately 13-fold increased over wild-type levels. This is likely due to the additional expression of the gene from the plasmid, since it carries a construct that includes both *aid1* and full size FN1254 gene. The expression of both genes is highly increased and follows a similar pattern. More surprisingly, in the *aid1*(S) strain we observed a similar pattern of FN1254 expression, where it was approximately 3-fold increased over wild-type levels. This is more unexpected, since the plasmid doesn't carry the FN1254 sequence, therefore the increased expression must come from the genomic copy. While the overexpression of both genes isn't very dramatic, the increase in their expression is comparable at ~3 fold. This suggests that the cells require a strict stoichiometric amount of both genes to be expressed. Also, when the mutant *F. nucleatum* strains were co-incubated with *S. sanguinis* we observed a downregulation

in the expression of the FN1254 gene. While the absolute values were lower than those of *aid1,* ~2 fold vs. 3-4 fold reduction, the downregulation was detectable and consistent. Expression data was obtained by comparing the 0hr time point to the 4hr time point of each strain with respect to itself. This data corroborated with the previous finding of the importance of the stoichiometry between the *aid1* and FN1254 expression levels. The *aid1*(L) strain carries the FN1254 on the plasmid, but the *aid1*(S) strain does not. Yet, FN1254 was downregulated in both upon co-incubation with *S. sanguinis.* It appears that the chromosomal copy of the FN1254 gets downregulated as the levels of *aid1* are reduced in the cell. This data suggests that the protein product of FN1254 is not likely to be the regulatory element necessary for the regulation of *aid1* expression. The gene function of FN1254 is not yet known, but based on sequence predictions it may have enzymatic activity. The regulatory elements required for protein production appear to be upstream of the FN1254 gene and are necessary for Aid1 production. Since the two genes can be transcribed as a polycistronic message, it is not surprising that the expression and downregulation is similar. Yet, the data from the strain that only has a chromosomal copy of FN1254 following the same expression and downregulation pattern of the gene, suggests that the number of mRNA copies or protein molecules of Aid1 and FN1254 affect each other's production levels.

In summary, the endogenous promoter of *aid1* is necessary and sufficient for downregulation of the gene in the presence of *S. sanguinis*, while Aid1 protein production is subject to a more complicated regulatory mechanism that involves translation from a longer polysistronic mRNA as well as a specific stoichiometric balance between Aid1 and FN1254. Attempts to inactivate the FN1254 gene in order to demonstrate the role of FN1254 protein in the regulation of *aid1* have been unsuccessful, but as better genetic tools become available, it may be possible and important for furthering our understanding how the *F. nucleatum* cells regulate expression of *aid1.* The role of the short *aid1* messenger in protein production, gene regulation or the cell's

phenotype is yet to be elucidated. To the best of our knowledge, this is the first report of lipoprotein gene regulation in *F. nucleatum* as well as the first study that shows differential transcription and translation mechanisms in this organism.

Name	Primer sequence	Source
aid1 qPCR Fwd	5'-TACAGGAGGTGCCGTAGCAG-3'	Kaplan et al., 2014
aid1 qPCR Rev	5'-TTTTTGTTAATTCTCCAGCTCCA-3'	Kaplan et al., 2014
16S qPCR Fwd	5'-TTGGACAATGGACCGAGAGT-3'	Kaplan et al., 2014
16S qPCR Rev	5'-GCCGTCACTTCTTCTGTTGG-3'	Kaplan et al., 2014
FN1254 Fwd	5'-CTGGGGCAATGATAGCTTTT-3'	This study
FN1253 Rev	5'-TCCAGAAGGAAAACAACCATCA-3'	This study
Linker Fwd	5'-GGCAGAAGAACTTGGATTAGG-3'	This study
Linker Rev	5'- GGCACCTCCTGTAGCAGTTT-3'	This study

Table 1. Primers used in this study

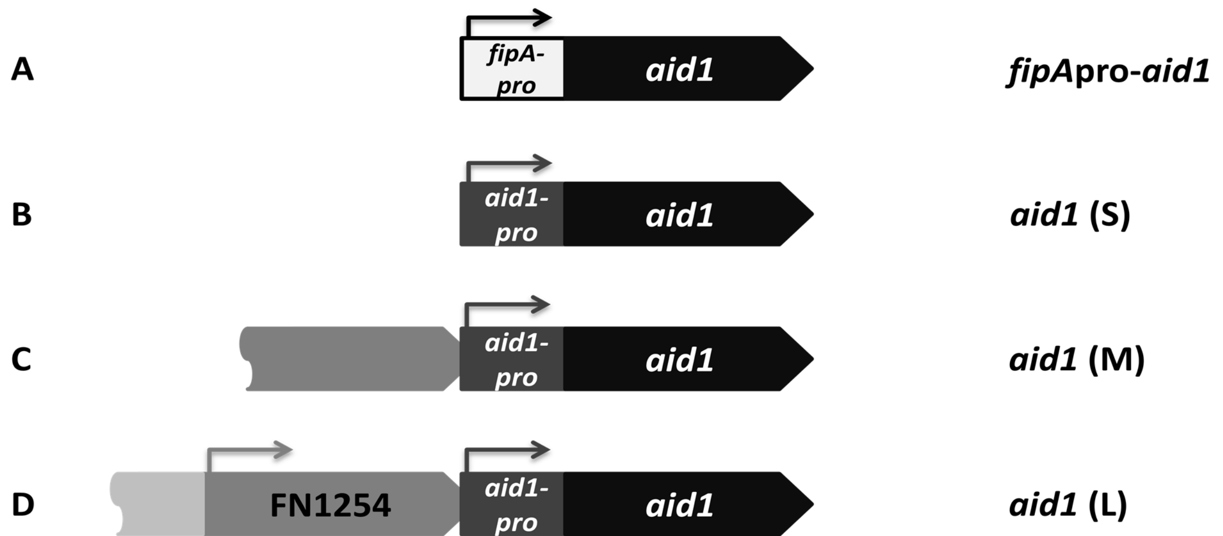


Figure 1. *aid1* promoter constructs

A. *aid1* expressed under fusobacterial FipA protein promoter. **B.** *aid1* expressed under its endogenous promoter. **C.** *aid1* expressed under its own promoter with additional upstream sequence. **D.** *aid1* and ORF FN1254 with their endogenous promoters.

All constructs were introduced into *F. nucleatum* Δ*aid1* strain on a pHS58 shuttle vector.

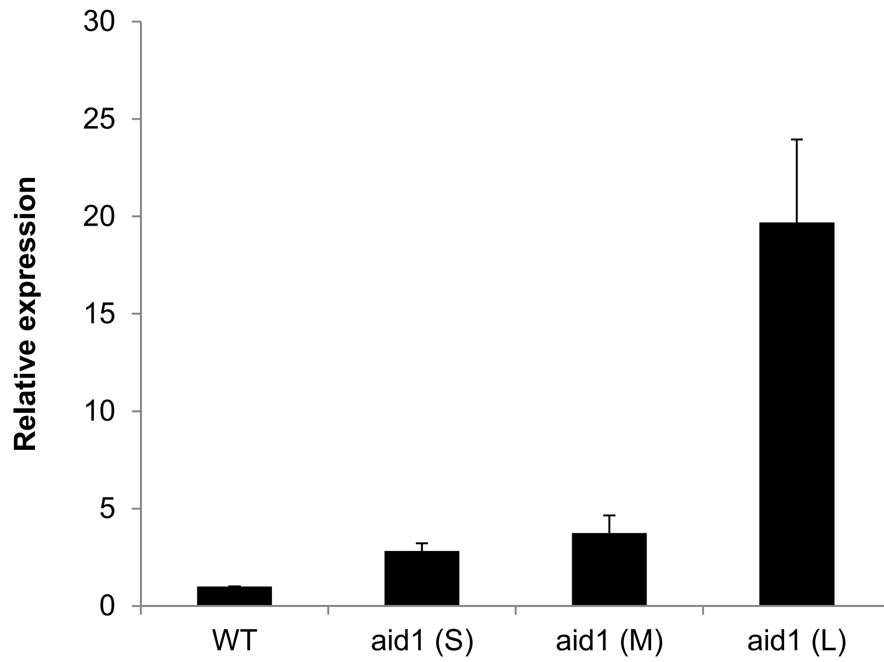


Figure 2. *aid1* expression varies due to the length of the upstream sequence

aid1 gene expression in planktonic cultures was determined by qRT-PCR in strains expressing short (S), medium (M) and long (L) constructs compared to wild-type levels.

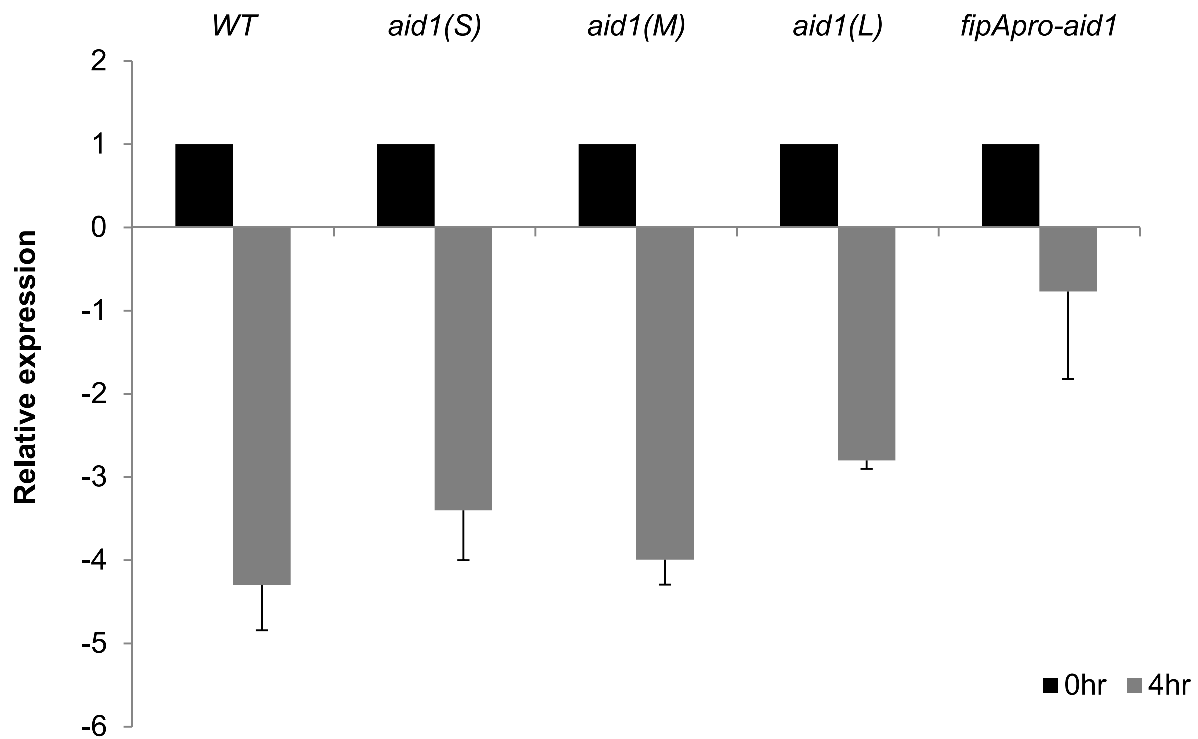


Figure 3. *aid1* endogenous promoter sequence is necessary and sufficient for downregulation of the gene in the presence of *S. sanguinis.*

Equal numbers of *F. nucleatum* and *S. sanguinis* cells were pelleted together and incubated anaerobically for 4 hours. Expression of *aid1* was analyzed by qRT-PCR and normalized to 16S rRNA levels.

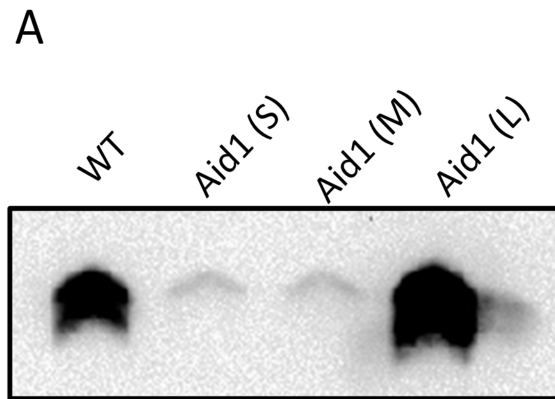
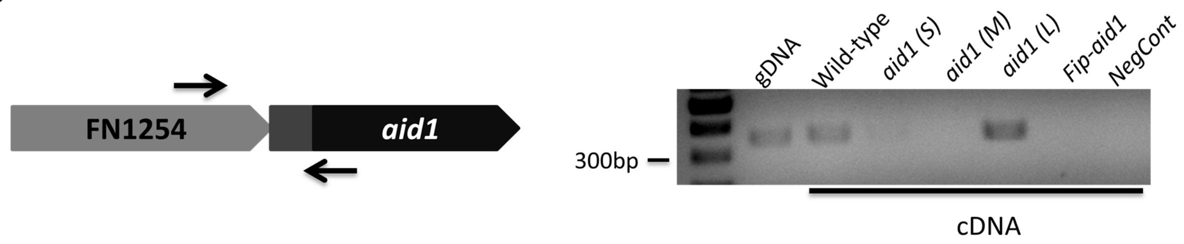


Figure 4. Aid1 protein production requires co-transcription with FN1254

A. Aid1 protein levels in strains carrying S, M and L constructs were determined by Western blot analysis.

B



B. FN1254-*aid1* junction was amplified using PCR. Genomic DNA (gDNA) from wild-type *F. nucleatum* cells was compared to cDNA samples from wild-type, Aid1(S), (M) and (L) strains, as well as to the FipApro-*aid1* strain. Negative control sample had no input nucleic acid.

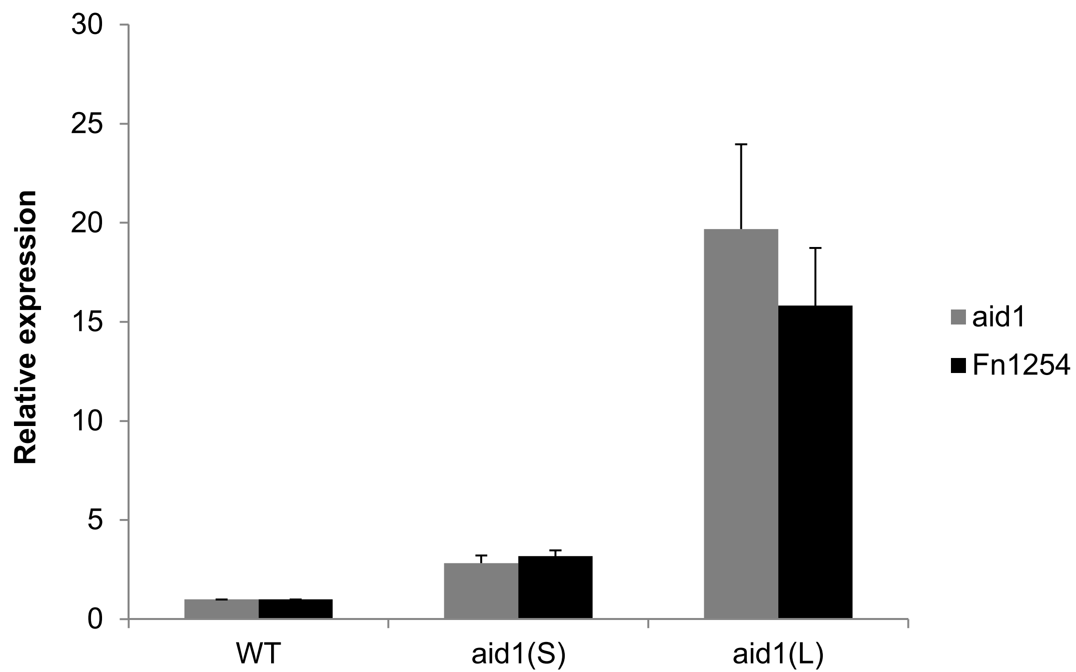


Figure 5. FN1254 expression is similar to that of *aid1*

Gene expression of *aid1* and FN1254 was analyzed by qRT-PCR and standardized to expression of 16S ribosomal RNA expression.

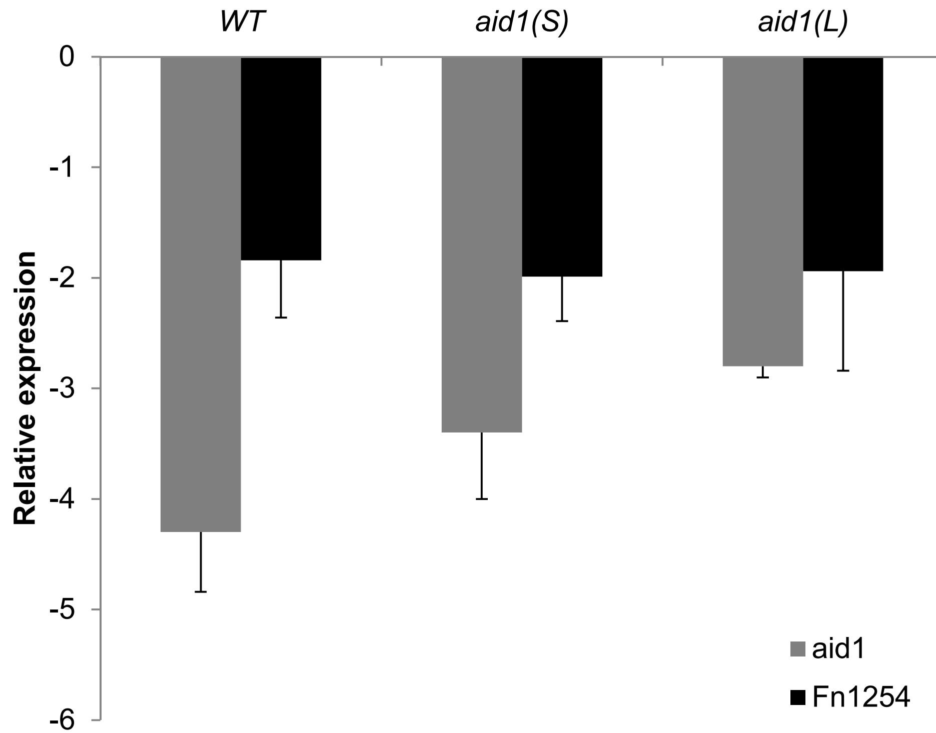


Figure 6. ***aid1*** **and FN1254 are regulated in a similar manner in the presence of** ***S. sanguinis***

The *F. nucleatum* strains were co-incubated with equal cell numbers of *S. sanguinis* for 4 hours. Total cell RNA was isolated and used for the analysis of gene expression levels of *aid1* and FN1254 using qRT-PCR.

Conclusion

F. nucleatum has long been a challenging organism to work with due to the lack of genetic tools and unique features not shared with other known organisms. In this work we describe a novel uncharacterized lipoprotein in *F. nucleatum* for its role in interspecies interactions with oral streptococci. To date, only one fusobacterial adhesin, RadD, has been characterized on a molecular level (Kaplan et al. 2009). Several other outer membrane proteins are being characterized for their roles in pathogenesis and adhesion to human tissues (Xu et al. 2007). FadI is the only other fusobacterial protein that has been proposed to be a lipoprotein (Gupta et al. 2010), yet its molecular characteristics, including cellular localization and functional domains, as well as its regulation have not yet been described.

In this work we sought to characterized a novel hypothetical protein in *Fusobacterium nucleatum.* The gene was originally discovered through a microarray analysis performed on *F. nucleatum* that was grown in different conditions, including single species planktonic and biofilm cultures, as well as dual species biofilms with early and late colonizing partner species. Encoded by the ORF FN1253, this gene was ubiquitously regulated in all conditions tested. In the presence of partner species, particularly the oral streptococci, FN1253 was significantly downregulated. We constructed inactivation and overexpression mutants of FN1253 and observed a change in coaggregation ability of these *F. nucleatum* mutants with *S. sanguinis,* as representative early colonizing partner strain. The FN1253 overexpression strain also had an aberrant dual species biofilm phenotype when grown with *S. sanguinis*. Increased expression of the gene resulted in an abnormal architecture of the dual species biofilms, where the streptococci were able to bind throughout the height of the biofilm, instead of primarily localizing to the bottom layers. Because of these observations we renamed FN1253 to *aid1* (Adhesion Inducing Determinant 1), and sought to further characterize the gene on a molecular level in order to better understand its biological function in the cell (Kaplan et al. 2014).

In order to proceed with the biochemical characterization of the Aid1 protein, a specific α-Aid1 antibody was generated using Genomic Antigen Technology (GAT) that utilizes DNA immunization. Protein sequence analysis predicted Aid1 to be a lipoprotein and localize to the outer membrane of the cell. We performed cell fractionations and separated the membrane fractions from the cytosol, as well as inner and outer membrane fractions and showed that Aid1 localizes to the outer membrane of the cell. We also demonstrated that Aid1 is localized in the inner leaflet of the outer membrane. Deletion of the signal peptide and mutation in the conserved cysteine residue both resulted in rapid turnover of the protein in the *F. nucleatum* cells and no detectable protein amounts, suggesting that the accumulation of the unprocessed protein molecules is toxic to the cell and is quickly degraded. Aid1 is the first fusobacterial lipoprotein that has been shown to play a role in interspecies interactions and is being characterized on a molecular level. Previous studies have identified potential lipoproteins in *F. nucleatum* but have not dissected their processing and cellular localization, making Aid1 unique in that respect.

aid1 is differentially regulated in the presence of *S. sanguinis* and requires physical contact of the two species for a rapid response and downregulation of expression. We sought to investigate the genetic elements responsible for the regulation of *aid1* expression. Our data revealed that *aid1* can be transcribed from its own endogenous promoter as well as a part of a polycistronic mRNA transcribed from the promoter of the upstream FN1254 gene. The endogenous promoter was necessary and sufficient for the downregulation of *aid1* in the presence of *S. sanguinis,* as we observed no change in expression when the promoter sequence was replaced with a foreign promoter of the *fipA* gene. On the other hand, the Aid1 protein production requires transcription of the full polycistronic messenger RNA that includes the upstream FN1254 gene. Mutant strains harboring plasmids that only contained *aid1* with the

endogenous promoter had very small, almost undetectable levels of protein. Meanwhile, the mutant strain carrying the long construct produced protein levels higher than that of the wild-type strain. The FN1254 gene was also expressed and regulated similarly to *aid1* corroborating the dual transcription data. Interestingly, the expression of the two genes followed the same pattern even in the mutant strains that harbored *aid1* on a plasmid and FN1254 on the chromosome, suggesting that the two genes share a common master regulator irrespective of whether or not they are transcribed on the same mRNA.

In conclusion, we have identified and characterized a novel hypothetical protein in *F. nucleatum aid1* which is the first fusobacterial lipoprotein characterized on a deeper molecular level. We also have described some of the genetic elements necessary for the regulation of the *aid1* gene. Very little is known about gene regulation machinery in *F. nucleatum* and this work adds to our knowledge of how genes are regulated in this organism. At the same time, there are still unanswered questions on the molecular mechanisms of the function and regulation of *aid1.* Further characterization of the role of Aid1 in modulation of adhesion is necessary. Also, the identification of the master regulator responsible for downregulation of *aid1* will add to our understanding of its function within the cell.

References

Aas, J.A. et al., 2005. Defining the normal bacterial flora of the oral cavity. *Journal of Clinical Microbiology*, 43(11), pp.5721–5732.

Arambula, D. et al., 2013. Surface display of a massively variable lipoprotein by a Legionella diversity-generating retroelement. *Proceedings of the National Academy of Sciences of the United States of America*, 110(20), pp.8212–7. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3657778&tool=pmcentrez&rendertype=abstract.

Babitzke, P., 2004. Regulation of transcription attenuation and translation initiation by allosteric control of an RNA-binding protein: The Bacillus subtilis TRAP protein. *Current Opinion in Microbiology*, 7(2), pp.132–139.

Babu, M.M. et al., 2006. A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *Journal of Bacteriology*, 188(8), pp.2761–2773.

Barwe, S.P. et al., 2007. Janus model of the Na,K-ATPase beta-subunit transmembrane domain: distinct faces mediate alpha/beta assembly and beta-beta homo-oligomerization. *Journal of molecular biology*, 365(3), pp.706–714.

Bastos, J.A. et al., 2011. Identification of periodontal pathogens and severity of periodontitis in patients with and without chronic kidney disease. *Archives of Oral Biology*, 56(8), pp.804–811.

Ten Cate, J.M., 2006. Biofilms, a new approach to the microbiology of dental plaque. *Odontology*, 94(1), pp.1–9.

Cisar, J.O., Kolenbrander, P.E. & McIntire, F.C., 1979. Specificity of coaggregation reactions between human oral streptococci and strains of Actinomyces viscosus or Actinomyces naeslundii. *Infection and Immunity*, 24(3), pp.742–752.

Cockayne, A. et al., 1998. Molecular cloning of a 32-kilodalton lipoprotein component of a novel iron-regulated Staphylococcus epidermidis ABC transporter. *Infection and Immunity*, 66(8), pp.3767–3774.

Condon, C., 2003. RNA processing and degradation in Bacillus subtilis. *Microbiology and molecular biology reviews : MMBR*, 67(2), pp.157–174, table of contents.

Copeland, P.R., 2003. Regulation of gene expression by stop codon recoding: Selenocysteine. *Gene*, 312(1-2), pp.17–25.

Deka, R.K. et al., 2012. Structural, bioinformatic, and in vivo analyses of two Treponema pallidum lipoproteins reveal a unique TRAP transporter. *Journal of molecular biology*, 416(5), pp.678–96. Available at: http://www.sciencedirect.com/science/article/pii/S0022283612000691 [Accessed April 29, 2015].

Dewhirst, F.E. et al., 2010. The human oral microbiome. *Journal of Bacteriology*, 192(19), pp.5002–5017.

Dintilhac, A. & Claverys, J.P., 1997. The adc locus, which affects competence for genetic transformation in Streptococcus pneumoniae, encodes an ABC transporter with a putative lipoprotein homologous to a family of streptococcal adhesins. *Research in Microbiology*, 148(2), pp.119–131.

Fischer, M. et al., 2010. Caught in a TRAP: substrate-binding proteins in secondary transport. *Trends in microbiology*, 18(10), pp.471–8. Available at: http://www.sciencedirect.com/science/article/pii/S0966842X10001113 [Accessed April 29, 2015].

Frias, J., Olle, E. & Alsina, M., 2001. Periodontal pathogens produce quorum sensing signal molecules. *Infection and Immunity*, 69(5), pp.3431–3434.

González, O.A. et al., 2010. HIV-1 reactivation induced by the periodontal pathogens Fusobacterium nucleatum and Porphyromonas gingivalis involves toll-like receptor 4 and 9 activation in monocytes/macrophages. *Clinical and Vaccine Immunology*, 17(9), pp.1417–1427.

Gottesman, S., 2004. The small RNA regulators of Escherichia coli: roles and mechanisms*. *Annual review of microbiology*, 58, pp.303–328.

Gupta, S. et al., 2010. Fusobacterium nucleatum-associated beta-defensin inducer (FAD-I): identification, isolation, and functional evaluation. *The Journal of biological chemistry*, 285(47), pp.36523–36531.

Haake, S.K. et al., 2000. Native plasmids of Fusobacterium nucleatum: characterization and use in development of genetic systems. *Journal of bacteriology*, 182(4), pp.1176–1180.

Han, Y.W. et al., 2005. Identification and characterization of a novel adhesin unique to oral fusobacteria. *Journal of Bacteriology*, 187(15), pp.5330–5340.

Han, Y.W. & Wang, X., 2013. Mobile microbiome: oral bacteria in extra-oral infections and inflammation. *Journal of dental research*, 92(6), pp.485–91. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3654760&tool=pmcentrez&rendertype=abstract.

Hsiao, W.W. et al., 2012. Microbial transformation from normal oral microbiota to acute endodontic infections. *BMC Genomics*, 13(1), p.345.

Jakubovics, N.S. et al., 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.

Kapatral, V. et al., 2002. Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586. *Journal of Bacteriology*, 184(7), pp.2005–2018.

Kaplan, A. et al., 2014. Characterization of aid1, a Novel Gene Involved in Fusobacterium nucleatum Interspecies Interactions. *Microbial Ecology*, 68(2), pp.379–387.

Kaplan, C.W. et al., 2005. Fusobacterium nucleatum apoptosis-inducing outer membrane protein. *Journal of dental research*, 84(8), pp.700–704.

Kaplan, C.W. et al., 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.

Kim, S. et al., 2005. Transmembrane glycine zippers: physiological and pathological roles in membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 102(40), pp.14278–14283.

Kolenbrander, P. & London, J., 1992. Ecological Significance of Coaggregation among Oral Bacteria. In K. C. Marshall, ed. *Advances in Microbial Ecology SE - 4*. Advances in Microbial Ecology. Springer US, pp. 183–217. Available at: http://dx.doi.org/10.1007/978-1-4684-7609-5_4.

Kolenbrander, P.E., 1993. Coaggregation of human oral bacteria: potential role in the accretion of dental plaque. *Journal of Applied Bacteriology*, 74, p.79S–86S. Available at: http://dx.doi.org/10.1111/j.1365-2672.1993.tb04344.x.

Kolenbrander, P.E. et al., 2002. Communication among oral bacteria. *Microbiology and molecular biology reviews : MMBR*, 66(3), pp.486–505, table of contents.

Kolenbrander, P.E. & Andersen, R.N., 1989. Inhibition of coaggregation between Fusobacterium nucleatum and Porphyromonas (Bacteroides) gingivalis by lactose and related sugars. *Infection and Immunity*, 57(10), pp.3204–3209.

Kolenbrander, P.E., Andersen, R.N. & Moore, L. V, 1990. Intrageneric coaggregation among strains of human oral bacteria: potential role in primary colonization of the tooth surface. *Applied and environmental microbiology*, 56(12), pp.3890–3894.

Kolenbrander, P.E. & London, J., 1993. Adhere today, here tomorrow: Oral bacterial adherence. *Journal of Bacteriology*, 175(11), pp.3247–3252.

Kovacs-Simon, A., Titball, R.W. & Michell, S.L., 2011. Lipoproteins of bacterial pathogens. *Infection and Immunity*, 79(2), pp.548–561.

Kozak, M., 2005. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene*, 361(1-2), pp.13–37.

Lee, H.-R. et al., 2012. Fusobacterium nucleatum GroEL induces risk factors of atherosclerosis in human microvascular endothelial cells and ApoE(-/-) mice. *Molecular oral microbiology*, 27(2), pp.109–23. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22394469.

McHardy, I.H., 2011. *Analysis of Interspecies Interactions of Periodontal Bacteria Using Microarrays*. Ann Arbor: University of California, Los Angeles. Available at: http://search.proquest.com/docview/1022181168?accountid=14512.

Merritt, J. et al., 2009. Autoaggregation response of Fusobacterium nucleatum. *Applied and Environmental Microbiology*, 75(24), pp.7725–7733.

Narita, S. ichiro & Tokuda, H., 2010. Sorting of bacterial lipoproteins to the outer membrane by the Lol system. *Methods in molecular biology (Clifton, N.J.)*, 619, pp.117–129.

Narita, S.I. & Tokuda, H., 2006. An ABC transporter mediating the membrane detachment of bacterial lipoproteins depending on their sorting signals. *FEBS Letters*, 580(4), pp.1164–1170.

Nyvad, B. & Kilian, M., 1987. Microbiology of the early colonization of human enamel and root surfaces in vivo. *Scandinavian journal of dental research*, 95(5), pp.369–380.

Ohta, K., Makinen, K.K. & Loesche, W.J., 1986. Purification and characterization of an enzyme produced by Treponema denticola capable of hydrolyzing synthetic trypsin substrates. *Infection and Immunity*, 53(1), pp.213–220.

Okada, M. et al., 2002. *Microbiological, immunological and genetic factors in family members with periodontitis as a manifestation of systemic disease, associated with hematological disorders.*,

Okuda, S. & Tokuda, H., 2011. Lipoprotein Sorting in Bacteria. *Annual Review of Microbiology*, 65(1), pp.239–259.

Palmer, R.J. et al., 2003. Coaggregation-mediated interactions of streptococci and actinomyces detected in initial human dental plaque. *Journal of Bacteriology*, 185(11), pp.3400–3409.

Plotkowski, M.L. et al., 2007. Transmembrane domain of myelin protein zero can form dimers: Possible implications for myelin construction. *Biochemistry*, 46(43), pp.12164–12173.

Ptashne, M. & Gann, A., 2002. *Genes and Signals.*, Available at: http://www.amazon.com/dp/0879696338.

Rickard, A.H. et al., 2003. Bacterial coaggregation: an integral process in the development of multi-species biofilms. *Trends in Microbiology*, 11(2), pp.94–100. Available at: http://www.sciencedirect.com/science/article/pii/S0966842X02000343 [Accessed April 29, 2015].

Serganov, A., Huang, L. & Patel, D.J., 2009. Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature*, 458(7235), pp.233–237.

Simionato, M.R. et al., 2006. Porphyromonas gingivalis genes involved in community development with Streptococcus gordonii. *Infect Immun*, 74(11), pp.6419–6428. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16923784.

Sparks Stein, P. et al., 2012. Serum antibodies to periodontal pathogens are a risk factor for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, 8(3), pp.196–203. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3712346&tool=pmcentrez&rendertype=abstract.

Storz, G., Opdyke, J.A. & Zhang, A., 2004. Controlling mRNA stability and translation with small, noncoding RNAs. *Current Opinion in Microbiology*, 7(2), pp.140–144.

Sutcliffe, I.C. & Russell, R.R., 1995. Lipoproteins of gram-positive bacteria. *Journal of Bacteriology*, 177(5), pp.1123–1128. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC176714/.

Takemoto, T. et al., 1995. Characteristics of multimodal co-aggregation between Fusobacterium nucleatum and streptococci. *Journal of periodontal research*, 30(4), pp.252–257.

Vitreschak, A.G. et al., 2002. Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic acids research*, 30(14), pp.3141–3151.

Whittaker, C.J., Klier, C.M. & Kolenbrander, P.E., 1996. Mechanisms of adhesion by oral bacteria. *Annual review of microbiology*, 50, pp.513–552.

Winkler, W.C. et al., 2004. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*, 428(6980), pp.281–286.

Winkler, W.C. & Breaker, R.R., 2005. Regulation of bacterial gene expression by riboswitches. *Annual review of microbiology*, 59, pp.487–517.

Xu, M. et al., 2007. FadA from Fusobacterium nucleatum utilizes both secreted and nonsecreted forms for functional oligomerization for attachment and invasion of host cells. *Journal of Biological Chemistry*, 282(34), pp.25000–25009.

Yamaguchi, K., Yu, F. & Inouye, M., 1988. A single amino acid determinant of the membrane localization of lipoproteins in E. coli. *Cell*, 53(3), pp.423–432.

Zijnge, V. et al., 2010. Oral biofilm architecture on natural teeth. *PLoS ONE*, 5(2).