Mutagenic Analysis of ArnD, an Enzyme

Essential for Polymyxin Resistance

Author:

Blake Marleau Wilcox-Snyder

Thesis Advisor:

Marcelo Sousa

Department of Chemistry and Biochemistry

University of Colorado, Boulder, CO

October 25th, 2012

Table of Contents

Abstract

Cystic Fibrosis (CF) is the most frequently inherited complication among Caucasians today, affecting 30,000 people in the U.S. alone. The major cause of mortality within CF patients is the development of chronic persistent infections by the opportunistic pathogen *Pseudomonas aeruginosa.* Typical antibacterial defense mechanisms used within healthy individuals such as Cationic Antimicrobial Peptides (CAMPs) are inhibited within CF patients. The inhibition is mediated by modification of the bacterial surface in the specific environment of the CF lung, namely: the covalent addition of $Ara4NH_4^+$ to Lipid A. This unique Lipid A modification utilized by *P. aeruginosa* within CF patients has been proven to establish a resistance against both CAMPs and last resort antibiotics such as Polymyxins. Abolishing this Lipid A modification would help control the pulmonary infections and thus holds significant promise as a means of extending the life expectancy and improving the quality of life of CF patients. The enzymes responsible for mediating a resistance against CAMPs are encoded by the seven-enzyme operon ArnBCADTEF. Here, I show that ArnD, an essential enzyme within the biosynthetic pathway, is capable of converting Undecaprenyl Phosphate-L-Ara4N-formyl to a novel species, which I propose is the deformylated intermediate Undecaprenyl Phosphate-L-Ara4N. I further show that the catalytic activity of ArnD is dependent on residues D43 and H233 as mutation of these residues dramatically reduces ArnD activity. This is consistent with ArnD catalyzing the deformylation reaction by a mechanism analogous to that of polysaccharide deacetylases, a family of metal dependent hydrolases. This would establish ArnD as the essential deformylase within the reaction pathway that leads to Lipid A $Ara4NH_4^+$ biosynthesis and bacterial resistance to CAMPs and clinical antibiotics of the same family such as Polymyxin B and Colistin.

Acknowledgements

Introduction:

Disease Background

CF is a genetic disease affecting approximately 30,000 people in the U.S. alone, and 70,000 worldwide [9]. CF is one of the most prevalent lethal hereditary disorders with a predicted median age of survival for a person with CF in the late 30s [9]. It is the most common inherited disorder amongst Caucasians; with an estimated 1 in 29 Caucasian Americans carrying the CF allele, and 1,000 new cases diagnosed annually [10].

In CF, homeostasis is upset due to a defect in the cystic fibrosis transmembrane-conductance regulator (CFTR) gene. This leads to a characteristic accumulation of thick mucus in the lungs, pancreatic ducts, and male reproductive tract. Within this environment, opportunistic pathogens such as *Pseudomonas aeruginosa* express genes that mediate resistance to the host immune system. This leads to bacterial colonization, bacterial overgrowth, and chronic, tissue-destructive inflammation [11]. These persistent infections are the main cause of mortality in CF.

This project is focused on understanding the resistance mechanisms that help bacteria colonize the CF lung.

Since CF is an autosomal recessive disease, an individual must have two copies of a mutated CFTR gene to express the disease phenotype. Thus, those affected by the disease are homozygous for mutations in the CFTR gene, while heterozygous individuals are relatively unaffected carriers. Approximately 70% of all CF patients carry the Δ508F mutation: a three base pair deletion in the CFTR gene (Figure 1)[4]. However, a small subset of patients carry the G551D mutation. In this case, the disease can be treated with ivacaftor ® (Vertex), a recently introduced CFTR inhibitor. This treatment method targets the CFTR gene mutation directly. On the other hand, there is no CFTR-targeted treatment currently available for the majority of the CF patient population. Therefore, it is important to develop treatments for the persistent infections that account for the majority of mortality associated with CF.
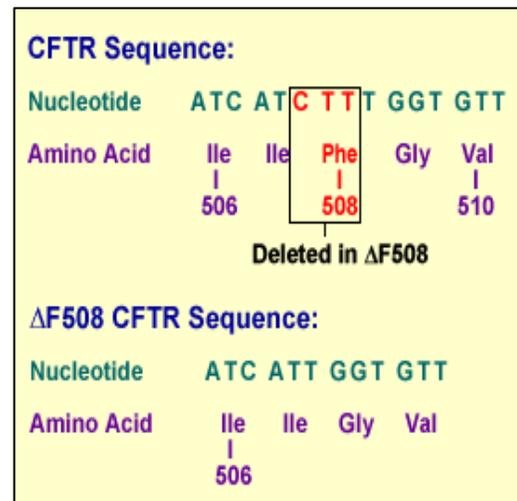
Figure 1: The ΔF508 deletion is the most common cause of cystic fibrosis. Ile507 remains unchanged because of degenerate code: ATC and ATT both code for isoleucine. Figure taken from [4]

Homozygous CFTR-mutants experience a defective transmembrane chloride ion channel resulting in an electrolyte imbalance. A normal sodium and chloride ion balance is needed to produce the thin mucus layer easily removed from the epithelial lining within wild-type (WT) individuals. The electrolyte imbalance in CF patients results in a

dehydrated mucus that is difficult to remove. This stationary sticky mucus layer is prone to colonization by Gram-negative bacteria, developing the chronic persistent infections that, as mentioned before, are the major cause of morbidity and mortality in CF [4].

Specifically, infections with *Pseudomonas aeruginosa* have been strongly correlated with CF mortality; approximately 80% of CF patients between the ages of 25 and 34 are infected with *P. aeruginosa* [8, 12]. These persistent infections in CF patients lead to inflammation, declining lung function, and death. Due to the fact that *P. aeruginosa* is the most frequently reported pathogen among CF patients, controlling these pulmonary infections holds significant promise as a means of extending the longevity and improving the quality of life of CF patients [9].

Numerous antimicrobial mechanisms act in unison to protect the healthy lung from infection within WT individuals, effectively forming a host defense system [13]. One way healthy individuals eliminate bacteria, including *P. aeruginosa,* from their bronchi is mucociliary clearance [14]. This mechanism encompasses the coordinated movement of cilia attached to the respiratory epithelium towards the pharynx. This synchronized movement carries microorganisms up towards the throat to be swallowed into the digestive environment of the stomach. Bacteria left behind by mucociliary clearance face a series of rapid-acting defense mechanisms provided by the innate immune response. Phagocytic cells are released from proximate lymph nodes such as the tonsils, and a number of Cationic Antimicrobial Peptides (CAMPs) are secreted, primarily from the airway epithelium.

The defensive mechanisms typically active within a healthy lung are inhibited within the CF lung. As a result of the mutated CFTR gene, mucociliary clearance is repressed due to the dehydrated viscous nature of the mucus layer covering the airway epithelium. This allows for bacteria to colonize the airway epithelium fluid layer. Once established, *P. aeruginosa* adopts a mucoid phenotype, overproducing alginate and forming a biofilm. This biofilm is highly sticky and makes it difficult for phagocytic cells and antibiotics to penetrate [15]. Furthermore, the effectiveness of CAMPs against *P. aeruginosa* appears to be reduced in CF patients. The loss of activity of CAMPs is explained by two specific alterations in the CF lung. First, there is an increased salt concentration in the airway surface fluid that reduces their effectiveness [13]. Second, and more specifically, a particular modification on the bacterial surface of *P. aeruginosa* mediates resistance to the CAMPs.

Cationic Antimicrobial Peptides (CAMPs)

CAMPs are found in all kingdoms of life. In mammals, they are produced in large quantities at sites of infection, such as the lungs, which are consistently bombarded by bacteria with nearly every breath. These peptides have broad-spectrum antibacterial, antifungal, antiviral and antiprotozoan properties [16]. Structural properties of CAMPs vary considerably, with a few common features. All CAMPs are encoded as larger precursors and subsequently modified [17], they are relatively small proteins, 12-50 amino acids, and are all amphipathic but positively charged. This positively charged feature of CAMPs mediates the electrostatic interaction the peptides have with the negatively charged microbial outer membrane, in the case of Gram-negative bacteria.

CAMPs target a fundamental difference between microbes and multicellular animals: the structure of their membranes. This distinguishing quality allows for CAMPs to target infectious microbes rather than host organism cells. This characterization of a foreign species can be more easily understood by looking more closely at bacterial targets, specifically Gram-negative bacteria.

Both Gram-positive and Gram-negative bacterial membranes have a peptidoglycan cell wall and a phospholipid bilayer with membrane-spanning proteins. Gram-negative bacteria, however, have an additional unique outer membrane (Figure 2). The Gram-negative outer membrane is an asymmetric bilayer that consists of phospholipids on the inner leaflet and Lipid A, the lipid anchor region of Lipo-Poly-Saccharide (LPS), on the outer leaflet [18]. Lipid A, or endotoxin, is recognized by the innate immune system. The outer leaflet faces the external environment and is rich in negatively charged phosphate headgroups [6]. This is in contrast to animal and plant membranes where most lipids facing the external environment are neutral, and negatively charged headgroups are

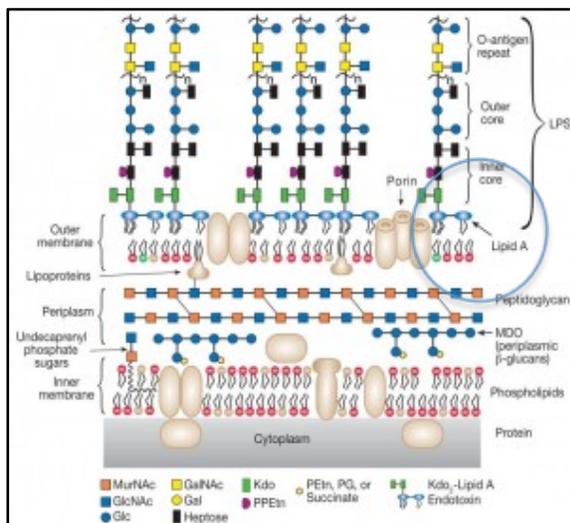Figure 2: Cartoon Model of Gram-negative bacterial membrane. Contains inner and outer membrane separated by peptidoglycan wall. LPS and Lipid A are anchored to outer leaflet of outer membrane. Image taken from [5].
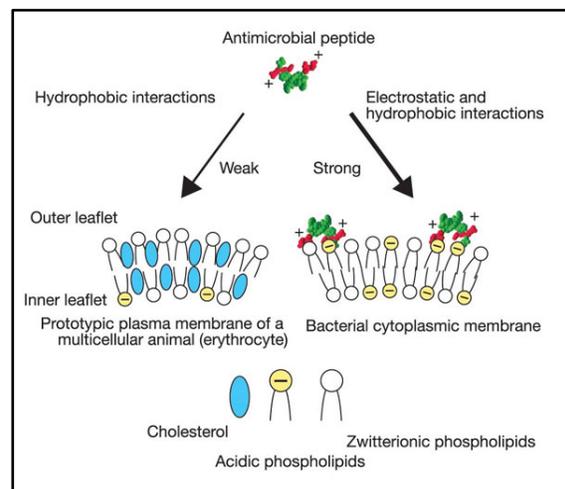
Figure 3: Cartoon Model depicting the how CAMPs differentiate between animal cell membranes and bacterial cell membranes. CAMPs bind negatively charged membranes with higher affinity. Image taken from [6].

localized to the inner leaflet. The different electrostatic properties of animal membranes allow for CAMPs to properly target microbes for destruction (Figure 3) [6]. Another feature, which protects animal cells, is cholesterol, which is present in animal membranes but is excluded from bacterial membranes. Cholesterol mitigates the activity of CAMPs either by stabilizing the non-microbial membrane or by interacting with the peptide directly [19].

CAMPs destroy microbes via permeabilization of their membranes. The Shai-Matsuzaki-Huang (SHM) model describes the general mechanism utilized by most CAMPs in their microbial identification and lysis (Figure 4). In short, the model proposes: first, an electrostatic interaction between the peptides and membrane, followed by lipid displacement, leading to an alteration in the membranes structure which further promotes apoptosis [20].
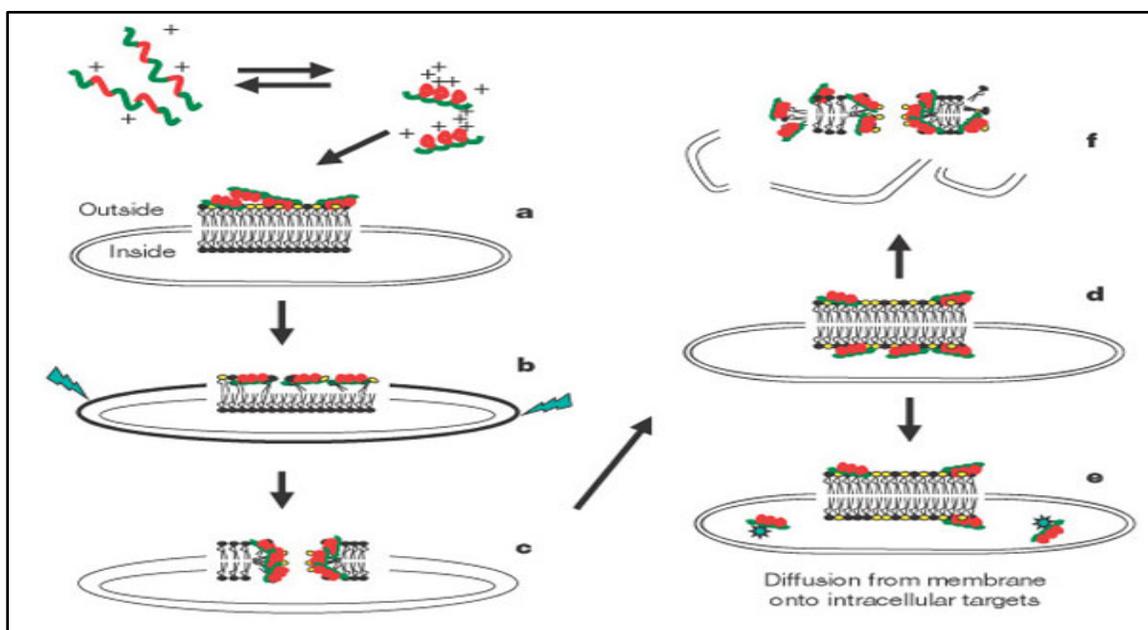
Figure 4:KEY: Phospholipid head groups: Yellow = (-) Charge / Black = No net charge
a) Electrostatic interaction between bacterial outer leaflet and peptides. b) Thinning of the outer leaflet due to peptide integration. Outer leaflet grows disproportionally to inner leaflet causing strain (Jagged blue arrows) c) Pore formation d) lipids and peptides transported into the inner leaflet. e) Peptides bind intracellular targets leading to cell death (in some cases). f) Destruction of cellular membrane and cell death (other cases). Image taken from [6].

The exact mechanism by which CAMPs kill microbes is still undetermined, although, many hypotheses have been presented. Fatal depolarization (of the normally energized membrane), formation of physical holes (causing outward diffusion of cellular contents), and improper distribution of membrane components are among of the most prevalent ideas published in the literature [19].

Despite their unclear killing mechanism, CAMPs remain attractive antimicrobial agents. They have intrinsic bactericidal activity and appear to promote the efficiency of inhaled antibiotic aerosols. This enhancement is believed to be a product of the above stated

disruption of the bacterial outer membrane, which facilitates the entry of additional molecules from the exterior, such as antibiotics [6, 21]. Once these molecules have reached the periplasmic space they then have a greater ability to integrate and transverse the cellular membrane to reach their intracellular cytoplasmic targets.

Due to their intrinsic bactericidal efficiency, a large commercial effort has been set forth to create CAMP analogs which represent a class of clinical antimicrobials. Aerosolized Polymyxins, specifically, have re-emerged as useful clinical antibiotics and have been utilized to treat infections with *P. aeruginosa* as well as other multidrug-resistant Gram-negative bacteria [22]. More specifically, Polymyxin E, otherwise known as colistin, is considered a last-resort treatment for chronic multidrug-resistant infections within current CF patients [22].

Crucial to the utilization of this field of clinical antimicrobials is an understanding of the bacterial modification pathway that confers resistance to CAMPs. An understanding of this pathway would allow for development of resistance inhibitors promoting improved bactericidal activity and extended clinical life for the antibiotic.

CAMP Resistance Mechanism in *P. aeruginosa*
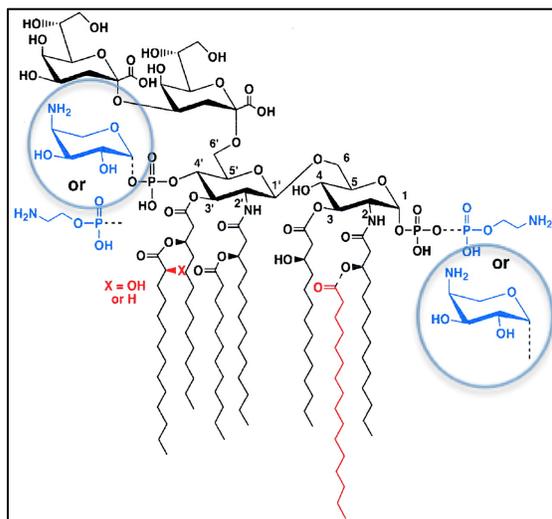
Figure 5: Molecular structure of Lipid A (Black) including some of the common bacterial modifications found in cells isolated from CF patients. the addition of L-Ara4N (blue), phosphoethanolamine (blue), and palmitoyl chain (red); the removal of hydroxymyrisotol chain (red); and the formation of 2-hydroxymyristate (red) [28].

Covalent modification of Lipid A is the main mechanism utilized by Gram-negative bacteria, including *P. aeruginosa,* to confer resistance to CAMPs and related antimicrobials. It has been demonstrated that bacteria make different forms of Lipid A in response to their environment [23]. The environment encountered by *P. aeruginosa* in the CF lung results in modification of its Lipid A with the positively charged sugar 4-amino-arabinose (Ara4NH$_4^+$)(Figure 5)[23, 24].

Covalent addition of Ara4NH$_4^+$ reduces the negative charge of outer bacterial membrane. Ara4NH$_4^+$ addition to Lipid A replaces a negatively charged phosphate with a positively charged primary amine. This modification confers resistance to CAMPs due to the decreased electrostatic interaction between the positively charged CAMPs and the negatively charged microbial outer leaflet. **This unique Lipid A modification utilized by *P. Aeruginosa* within CF patients has been proven to establish a resistance against both CAMPs and Polymyxins [23, 25].**

We hypothesize that inhibition of the biosynthetic pathway used to modify Lipid A with Ara4NH$_4^+$ in *P. aeruginosa* has utility in the treatment of CF patients. Inhibiting this pathway would effectively increase the susceptibility of the bacteria to host CAMPs and antibiotics that use a similar mechanism of action such as Polymyxins.

Biosynthetic Pathway for Lipid A- Ara4NH$_4^+$ Addition:

The enzymes responsible for mediating a resistance against CAMPs and Polymyxins via the biosynthetic addition of Ara4NH$_4^+$ to Lipid A are encoded by genes found at two different locations. Originally named Pmr (for Polymyxin Resistance), PmrE codes for UDP-glucose dehydrogenase, the first enzyme in the biosynthetic pathway. While, PmrHFIJKLM, at a different locus, is a seven-gene operon that encodes for the proteins that complete the biosynthetic pathway (Figure 6). This operon has been subsequently renamed ArnBCADTEF, (for Arabinose NH$_4^+$) following the order of reactions that
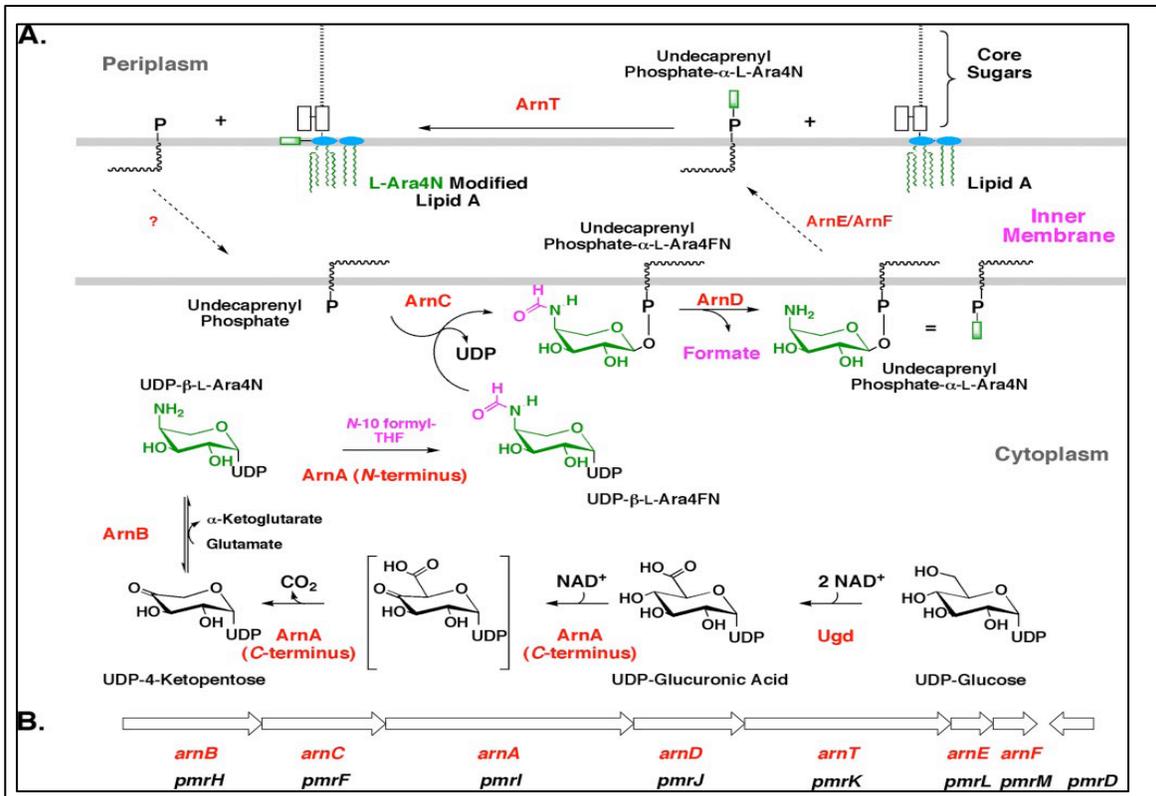
Figure 6: (A) The currently proposed pathway for the biosynthesis of Ara4N and its addition to Lipid A (taken from[1]). First, UDP-glucose is oxidized to UDP-glucuronic acid by Ugd. Next, the ArnA C-terminal domain converts UDP-glucuronic acid to novel UDP-4-keto-glucuronic acid intermediate which is subsequently decarboxylated to UDP-4-ketopentose. UDP-4-ketopentose is then transaminated by ArnB using glutamate as the amine donor to yield the novel sugar-nucleotide UDP-4-aminoarabinose (UDP-b-L-Ara4N). UDP-b-L-Ara4N is then formylated by the N-terminus of ArnA in order to drive the unfavorable reaction of ArnB forward. This transiently formylated product is used by ArnC to transfer Ara4N-formyl (AraFN) to undecaprenyl-phosphate with release of UDP. ArnD is proposed to deformylate undecaprenyl-phosphate-a-L-Ara4FN, which is then flipped to the periplasmic side of the inner membrane by the heterodimeric ArnE/ArnF. Last, ArnT transfers Ara4N from the undecaprenyl intermediate to Lipid A to yield the final product [Ara4N]-lipid A.

(B) Equivalence between the original Pmr and Current Arn nomenclatures.

result in covalent addition of Ara4NH$_4^+$ to Lipid A [26-28]. Genetic deletion studies within the ArnBCADTEF operon have been used to demonstrate that all of the enzymes in the biosynthetic pathway are essential for the modification [29]. Since all the enzymes in the pathway are essential it allows for multiple biological targets to be targeted for inhibitor design. This in turn results in a higher probability for successful inhibition and thus abolition of the Lipid A modification pathway.

The biosynthetic pathway, mediating the addition of Ara4NH4$^+$ to Lipid A, begins with the oxidation of a primary alcohol on UDP-glucose to a carboxylic acid in UDP-glucuronic acid. This reaction is catalyzed by the well-characterized dehydrogenase, Ugd (PmrE) [30]. Following this oxidation, the C-terminus of ArnA (PmrI) then catalyzes the oxidative decarboxylation of UDP-glucuronic acid to generate novel UDP-4-keto-pyranose [31, 32]. This intermediate is then transaminated using Glutamate as the amine donor by ArnB (PmrH) in an unfavorable reaction to produce UDP-ß-L-Ara4N with an equilibrium constant of 0.1 [33]. It has been shown that the N-terminus of ArnA then catalyzes the addition of a formyl group from N10-formyl-tetrahydrofolate to produce UDP-Ara4N. It has been suggested, since there have been no formyl groups detected on Lipid A, that this reaction occurs in order to drive the previous unfavorable reaction forward [33, 34]. It is important to notice that the formylation of UDP-Ara4N has been well characterized, however, there has been no detection of a formylated product attached to Lipid A. Thus, the pathway calls for a deformylase activity within one of the uncharacterized enzymes. The formylated product, an L-Ara4N-formyl moiety, is then transferred to undecaprenyl phosphate by ArnC (PmrF) forming undecaprenyl phosphate-αL-Ara4N-formyl. This crucial intermediate has been firmly established to accumulate within Polymyxin-resistant mutants of *Escherichia coli* and *Salmonella typhimurium* [35]. It has been hypothesized that this formylated intermediate is deformylated immediately by ArnD (PmrJ) followed by translocation of the Undecaprenyl Phosphate-L-Ara4N by the heterodimeric ArnE/ArnF protein complex. Finishing the pathway, ArnT (PmrK) then transfers the sugar moiety to the Lipid A unit effectively synthesizing [Ara4NH4$^+$]-Lipid-A.

ArnD, A hypothetical undecaprenyl phosphate-α-L-Ara4N-formyl deformylase

It has been determined through genetic mutation studies that the Protein ArnD is essential for [Ara4NH$_4^+$]-Lipid-A biosynthesis and thus CAMP resistance [29]. It has further been determined in our lab that ArnD is a membrane-associated enzyme (Lee, M. and Sousa M.S.; unpublished results). ArnD's currently proposed function within the pathway is to catalyze the deformylation reaction of undecaprenyl-phosphate-Ara4(N-formyl) to generate the final Ara4NH$_4^+$ structure that is attached downstream to Lipid A (Figure 7). Inhibition of ArnD would facilitate disrupting the modification pathway and effectively abolishing this type of resistance of Gram-negative bacteria against CAMPs and CAMP-like antibiotics such as Polymyxins.

This project focuses on the biochemical and structural characterization of ArnD. A detailed understanding of ArnD structure and mechanism is required in order to design specific inhibitors, which target the unique catalytic properties of ArnD. ArnD is a superb target for inhibitor design for two reasons. First, because ArnD is essential to the biosynthetic addition of Ara4NH$_4^+$ to Lipid A; and secondly, and even

more enticingly, because no known eukaryotic enzymes catalyzes the same reaction, giving its potential inhibitor a high specificity.
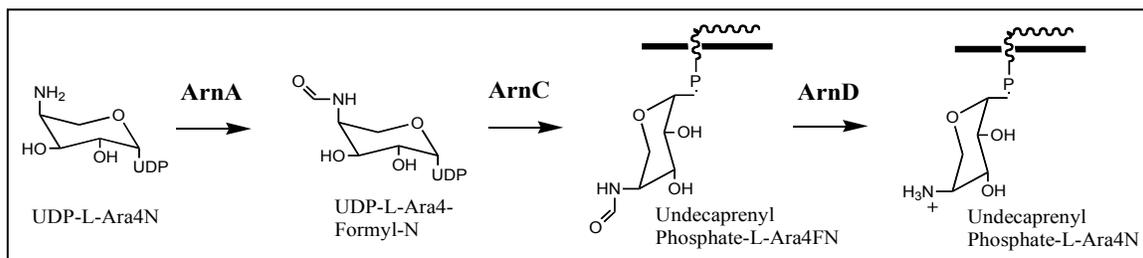
Figure 7: Sequence of reactions that lead to the synthesis of undecaprenyl-phosphate-Ara4N.

The complete genomic sequence of *Pseudomonas aeruginosa* was reported in Nature in 2000 by the Pathogenesis Corporation. This was the largest bacterial genome sequenced at the time, with almost 6.3 million base pairs, and provided crucial insight to the intrinsic drug resistance of *Pseudomonas aeruginosa* [36]. Among this insight, was the sequencing of the ArnBCADTEF operon, which is responsible for the biosynthetic addition of Ara4NH4$^+$ to Lipid A. This operon was found to be located between the 3.97 and 3.98 million base pair region of the *P. aeruginosa* genome (Figure 8) and provided the evidence that this resistance pathway was present in *P. aeruginosa.*
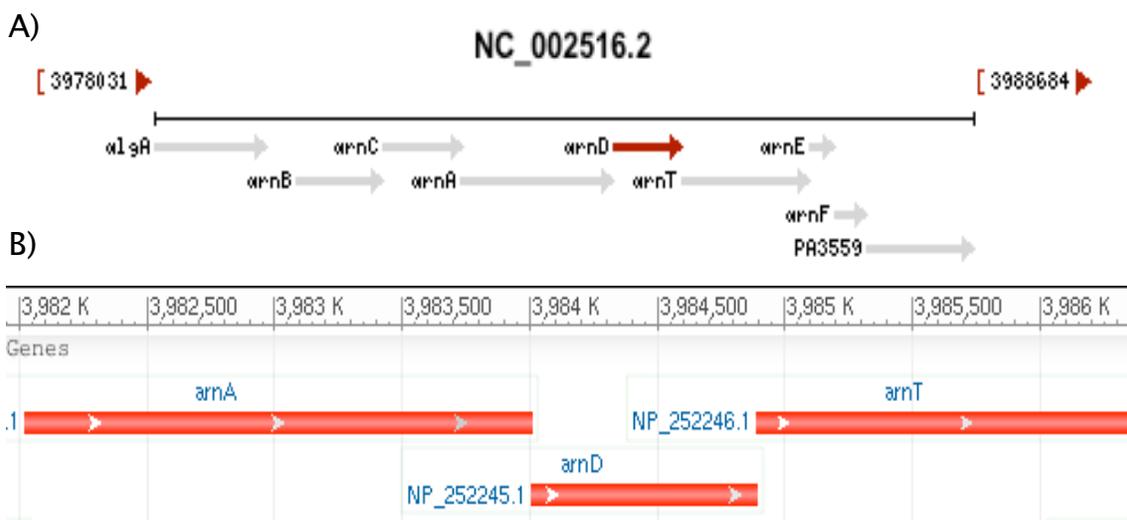
Figure 8: A) ArnD gene location within the *Pseudomonas aeruginosa* genome.
B) ArnD gene location within the ArnBCADTEF operon. Image from [36].

To increase the chances that the biochemical and structural characterization of these enzymes is successful, our lab has targeted the Arn enzymes from four model organisms: *Escherichia coli*, *Salmonella typhimurium*, *Yersinia pseudotuberculosis* and *Pseudomonas aeruginosa.* The amino acid sequence similarity amongst these orthologous proteins is ~80 %. This strongly suggests that the structural and biochemical

characterization of ArnD from any one of these organisms will provide a similar, if not identical, model for all the others.

Preliminary protein expression experiments within this lab showed that *Salmonella* ArnD could be expressed in *E. coli* with good yields. Therefore the experiments described in this work were carried out with the *Salmonella* ortholog. *Salmonella* and *Pseudomonas* ArnD share high sequence conservation (64% identity, 75% similarity) (Figure 9). We are therefore confident that the results obtained with *Salmonella* ArnD are applicable to all members of the family, including *Pseudomonas*.

```
Pseudomonas_Aeruginosa    MIQAGLRIDVDTFRGTRDGVPRLLELLDEAGLKATFFFSVGPDNMGRHLW
Salmonella_               MTKVGLRIDVDTLRGTREGVPRLLATLHRHGVQASFFFSVGPDNMGRHLW
                          * :.********;****;****** *.. *;;*;**************

Pseudomonas_Aeruginosa    RLARPAFFWKMLRSRAASLYGWDILLAGTAWPGKPIGRELGPLMRRTQAA
Salmonella_               RLIKPRFLWKMLRSNAASLYGWDILLAGTAWPGKNIGNANAGIIRET-AT
                          ** ;* *;****** .************** **.  . ::*.* *;

Pseudomonas_Aeruginosa    GHEVGLHAWDHHAWQTHAGVWSVQQLGEQIRRGSDCLADILGQPVRCSAA
Salmonella_               YHETGLHAWDHHAWQTHSGHWSIRQLEEDIARGITALEAIIGKPVTCSAA
                           **.**************;* **;;** *;* ** .* *;*;** ****

Pseudomonas_Aeruginosa    AGWRADGRVVEAKQPFGFRYNSDCRGRGAFRPRLADGSPGIPQVPVNLPT
Salmonella_               AGWRADGRVVRAKEPFNLRYNSDCRGTTLFRPLLMPGQTGTPQIPVTLPT
                          **********.**;**.;******** *** * *..* **;**.***

Pseudomonas_Aeruginosa    FDEVVGPGLPREAYNDFILERFAAGRDN-VYTIHAEVEGLLLAPAFRELL
Salmonella_               WDEVIGPAVQAQSFNTWIISRMLQDKGTPVYTIHAEVEGIVHQPLFEDLL
                          ;***;**.:  :::* ;*;.*;  .;.. **********;:  * *.;**

Pseudomonas_Aeruginosa    RRAERRGIRFRPLGELLPDDPRSLPLAELVRGRLAGREGWLGVRQP----
Salmonella_               VRARDAGITFCPLGELLPASPESLPLGQIVRGHIPGREGWLGCQQAVSAS
                          **.  ** * ******* .*.****.;;***;;.******* ;*.
```

Figure 9: Key: * = Exact Match, : = Highly Conserved, . = Conserved, (Blank) = Not Conserved
Polypeptide sequence alignment comparison of ArnD within *Pseudomonas Aeruginosa* and *Salmonella Typhimuruim,* Illustrates orthologous nature and sequence conservation of ArnD within varying forms of bacterial species

Project Rationale and Hypothesis

A Blast search with ArnD showed weak but significant sequence similarity with members of the polysaccharide deacetylase superfamily (Figure 10). Polysaccharide deacetylases are metalloenymes, containing Zn^{2+} or Co^{2+} in their active site, that catalyze the hydrolysis of acetyl groups from polysaccharides. This suggests that ArnD may catalyze an analogous reaction: the hydrolysis of a formyl group from the sugar in undecaprenyl phosphate-αL-Ara4N-formyl
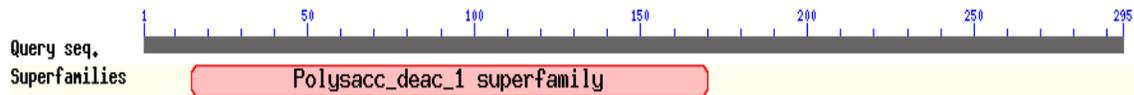
```
                 1        50          100          150          200          250      295
Query seq.    |_____|
Superfamilies        (  Polysacc_deac_1 superfamily  )
```

Figure 10: Blast search of ArnD sequence confirming relation to Polysaccharide Deacetylase Superfamily.

Characteristic of the polysaccharide deacetylase family are three amino acid residues coordinating the metal ion, while His and Asp residues are directly involved in the catalysis. Figure 11 shows the sequence alignment of ArnD with six polysaccharide deacetylases of known structure. Although there is low overall sequence similarity, the residues important for metal coordination and catalysis appear conserved (Figure 11 highlighted in blue and magenta, respectively).
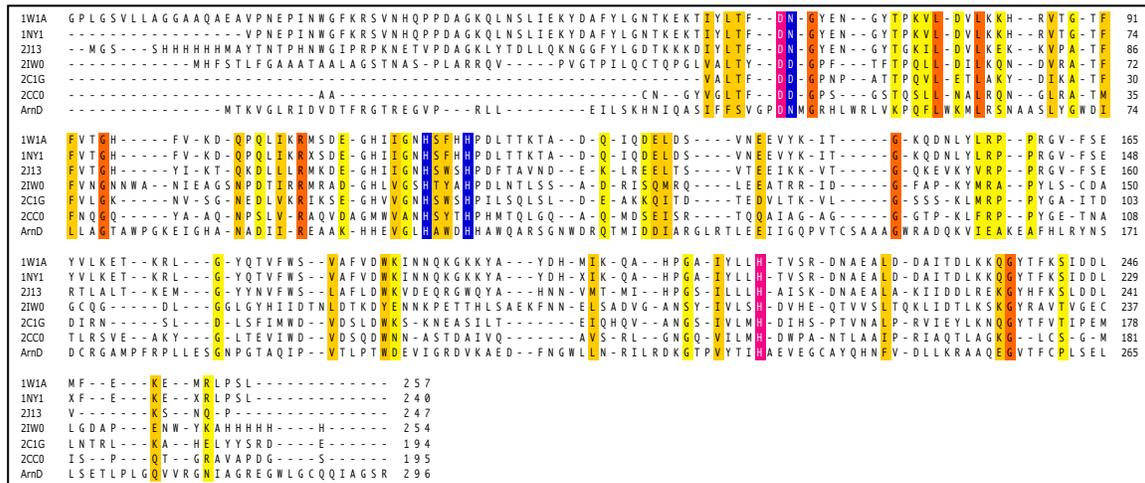
```
1W1A  GPLGSVLLAGGAAQAEAVPNEPINWGFKRSVNHQPPDAGKQLNSLIEKYDAFYLGNTKEKTIYLTF--DN-GYEN--GYTPKVL-DVLKKH--RVTG-TF   91
1NY1  ----------------VPNEPINWGFKRSVNHQPPDAGKQLNSLIEKYDAFYLGNTKEKTIYLTF--DN-GYEN--GYTPKVL-DVLKKH--RVTG-TF   74
2J13  --MGS---SHHHHHHMAYTNTPHNWGIPRPKNETVPDAGKLYTDLLQKNGGFYLGDTKKKDIYLTF--DN-GYEN--GYTGKIL-DVLKEK--KVPA-TF   86
2IW0  -----------MHFSTLFGAAATAALAGSTNAS-PLARRQV-----PVGTPILQCTQPGLVALTY--DD-GPF---TFTPQLL-DILKQN--DVRA-TF   72
2C1G  -------------------------------------------------VALTF--DD-GPNP--ATTPQVL-ETLAKY--DIKA-TF   30
2CC0  --------------------AA--------------------------CN--GYVGLTF--DD-GPS---GSTQSLL-NALRQN--GLRA-TM   35
ArnD  ------------MTKVGLRIDVDTFRGTREGVP---RLL---------EILSKHNIQASIFFSVGPDNMGRHLWRLVKPQFLWKMLRSNAASLYGWDI   74

1W1A  FVTGH-----FV-KD-QPQLIKRMSDE-GHIIGNHSFHHPDLTTKTA--D-Q-IQDELDS----VNEEVYK-IT-----G-KQDNLYLRP--PRGV-FSE  165
1NY1  FVTGH-----FV-KD-QPQLIKRXSDE-GHIIGNHSFHHPDLTTKTA--D-Q-IQDELDS----VNEEVYK-IT-----G-KQDNLYLRP--PRGV-FSE  148
2J13  FVTGH-----YI-KT-QKDLLLRMKDE-GHIIGNHSWSHPDFTAVND--E-K-LREELTS----VTEEIKK-VT-----G-QKEVKYVRP--PRGV-FSE  160
2IW0  FVNGNNWA--NIEAGSNPDTIRRMRAD-GHLVGSHTYAHPDLNTLSS--A-D-RISQMRQ----LEEATRR-ID-----G-FAP-KYMRA--PYLS-CDA  150
2C1G  FVLGK-----NV-SG-NEDLVKRIKSE-GHVVGNHSWSHPILSQLSL--D-E-AKKQITD----TEDVLTK-VL-----G-SSS-KLMRP--PYGA-ITD  103
2CC0  FNQGQ-----YA-AQ-NPSLV-RAQVDAGMWVANHSYTHPHMTQLGQ--A-Q-MDSEISR----TQQAIAG-AG-----G-GTP-KLFRP--PYGE-TNA  108
ArnD  LLAGTAWPGKEIGHA-NADII-REAAK-HHEVGLHAWDHHAWQARSGNWDRQTMIDDIARGLRTLEEIIGQPVTCSAAAGWRADQKVIEAKEAFHLRYNS  171

1W1A  YVLKET--KRL---G-YQTVFWS--VAFVDWKINNQKGKKYA---YDH-MIK-QA--HPGA-IYLLH-TVSR-DNAEALD-DAITDLKKQGYTFKSIDDL  246
1NY1  YVLKET--KRL---G-YQTVFWS--VAFVDWKINNQKGKKYA---YDH-XIK-QA--HPGA-IYLLH-TVSR-DNAEALD-DAITDLKKQGYTFKSIDDL  229
2J13  RTLALT--KEM---G-YYNVFWS--LAFLDWKVDEQRGWQYA---HNN-VMT-MI--HPGS-ILLLH-AISK-DNAEALA-KIIDDLREKGYHFKSLDDL  241
2IW0  GCQG------DL---GGLGYHIIDTNLDTKDYENNKPETTHLSAEKFNN-ELSADVG-ANSY-IVLSH-DVHE-QTVVSLTQKLIDTLKSKGYRAVTVGEC  237
2C1G  DIRN------SL---D-LSFIMWD--VDSLDWKS-KNEASILT------EIQHQV--ANGS-IVLMH-DIHS-PTVNALP-RVIEYLKNQGYTFVTIPEM  178
2CC0  TLRSVE--AKY---G-LTEVIWD--VDSQDWNN-ASTDAIVQ-------AVS-RL--GNGQ-VILMH-DWPA-NTLAAIP-RIAQTLAGKG--LCS-G-M  181
ArnD  DCRGAMPFRPLLESGNPGTAQIP--VTLPTWDEVIGRDVKAED--FNGWLLN-RILRDKGTPVYTIHAEVEGCAYQHNFV-DLLKRAAQEGVTFCPLSEL  265

1W1A  MF--E---KE--MRLPSL-------------  257
1NY1  XF--E---KE--XRLPSL-------------  240
2J13  V-------KS--NQ-P--------------  247
2IW0  LGDAP---ENW-YKAHHHHH----H-------  254
2C1G  LNTRL---KA--HELYYSRD----E-------  194
2CC0  IS--P---QT--GRAVAPDG----S-------  195
ArnD  LSETLPLGQVVRGNIAGREGWLGCQQIAGSR  296
```

Figure 11: Sequence alignment of ArnD with polysaccharide deacetylases of known structure. 1W1A: carbohydrate esterase *Bacillus subtilis*; 1NY1: polysaccharide deacetylase *Bacillus subtilis*; 2J13: carbohydrate esterase *Bacillus anthracis*; 2IW0: chitin deacetylase *Colletotrichum lindemuthianum*; 2C1G: peptidoglycan deacetylase *Streptococcus pneumoniae*; 2CC0: carbohydrate deacetylase *Streptomyces lividans*. Identical residues are shaded orange while conservative changes are shaded yellow. Residues involved in catalysis are highlighted in magenta while those involved in metal coordination are highlighted in blue.

14

The conservation of residues important for catalysis suggests that ArnD may utilize a hydrolysis mechanism similar to that of polysaccharide deacetylases. The catalytic mechanism of these metal dependent hydrolases is shown in Figure 12, using the Zn^{2+} dependent peptidoglycan deacetylase as an example [37-39].
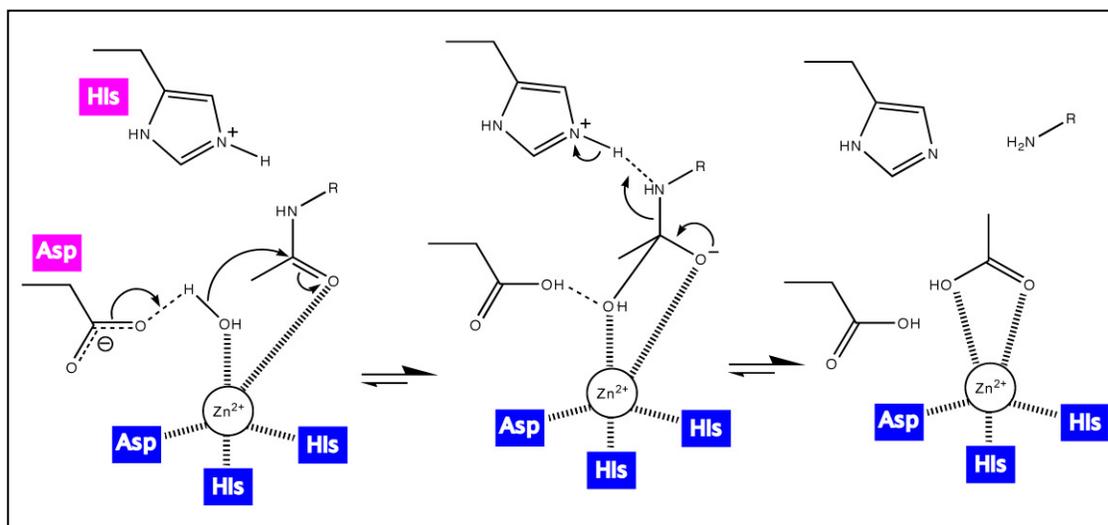

Figure 12: An example of the proposed catalytic mechanism for polysaccharide deacetylases and congruently for ArnD

In this mechanism a His and an Asp act as a general acid and base, respectively, to help hydrolyze the substrate. For this reaction a zinc-bound H_2O molecule is the nucleophile [39]. First, the zinc ion polarizes the side chain carbonyl oxygen on Asp. The aspartate residue then acts as a general base catalyst, activating the water molecule. This hydroxyde ion attacks the electrophilic carbonyl carbon of the substrate. The histidine in the catalytic pair then donates it's proton to the negatively charged intermediate, completing the hydrolysis. Within this catalytic model zinc acts to stabilize the tetrahedral oxyanion intermediate [39]. Within this example, zinc is used as the ionic species. However, it has been demonstrated that the coordinating metal ion is enzyme specific with Zn^{2+}[37, 39] and Co^{2+}[40] being the most frequent.

Given the necessity of a deformylase in the Lipid A-Ara4(NH_4^+) biosynthesis pathway and the conservation in ArnD of residues important for deacylation in related enzymes; it is my hypothesis that ArnD catalyzes the deformylation of undecaprenyl-phosphate-Ara4(N-formyl) utilizing a similar catalytic strategy as used by polysaccharide deacetylases.

Specific Goal:

I will test my hypothesis with point mutations of the putative catalytic residues: Asp_{43} and H_{233}; followed by an evaluation of the mutation's effect on ArnD activity. The catalytic residues were chosen first for mutation over the metal coordinating residues because only a single point mutation is likely necessary to affect activity. Mutating the metal coordinating residues may require mutation of two residues to confer a change in activity. Furthermore, a single point mutation is less likely to disturb the native fold of the enzyme. Ideally, the mutation would have no effect on the native fold, yet would not produce product do to the intrinsic loss of hydrolase activity. This circumstance would lend itself greatly towards crystallization and structure determination, which would further aid in specific inhibitor design.

Choice of mutations:

The catalytic residues, Asp_{43} and His_{233}, must be mutated into the most conservative mutations in order to maintain the native fold of the protein. For a protein folding in solution: the fold is primarily due to the hydrophobic effect, which is an entropically driven process.

The mutations that are most conservative were determined by Jonson et. al. and can be viewed in Figure 13 [2]. This figure describes the correlation between each individual amino acid residue and their distribution within known protein structures. Green regions represent a positive correlation, indicating residues that inhabit a similar protein environment, while red areas indicate negative correlation and white regions demonstrate a low degree of correlation either positive or negative. As can be seen in the figure, aspartate, D, shows a high degree of correlation with only two residues: serine and asparagine or S and N, respectively. On the other hand, histidine shows very little positive or negative correlations with any amino acid residues. However, some positive correlation can be seen in tyrosine, tryptophan and arginine, (Y, W and R, respectively). Therefore, 5 specific point mutations will be made to WT ArnD: aspartate 43 to asparagine and serine (D43N and D43S), and histidine 233 to tyrosine, tryptophan and arganine (H233W, H233Y, and H233R). Due to the conservative nature of these mutations, we expect them to not disrupt the structure of the enzyme.
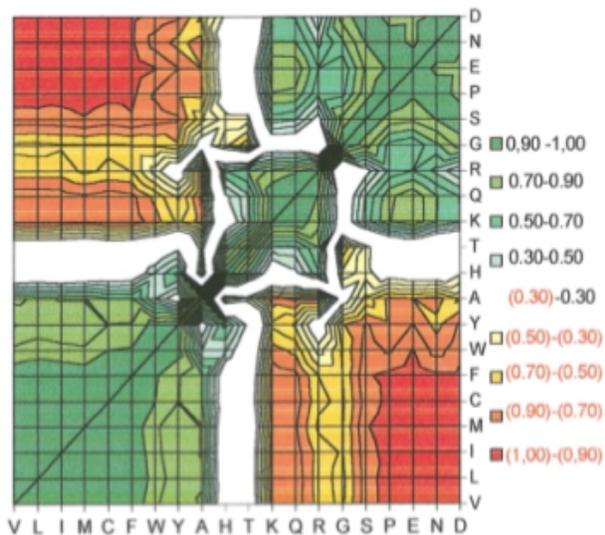

Figure 13: Demonstrates various correlations between different amino acids and their relative distribution within known enzymes. Green Areas represent a positive correlation. Red Areas represent a negative correlation. White areas have a low degree of correlation. Figure take from [2, 3].

Following mutagenesis of an expression plasmid, ArnD and its mutants will be extracted from membranes and purified optimizing the yield of non-aggregated protein. The enzymes will then be assayed for their ability to catalyze the deformylation of undecaprenyl-phosphate-Ara4(N-formyl) to yield undecaprenyl-phosphate-Ara4N. The substrate for ArnD is not commercially available and will be produced by sequential enzymatic transformations of UDP-Glucuronic acid. TLC separations will be used to monitor product formation for each enzymatic step as well as evaluation of ArnD activity.

My project tests whether ArnD does indeed catalyze the deformylation reaction of undecaprenyl-phosphate-Ara4(N-formyl) to yield undecaprenyl-phosphate-Ara4N. Further my experiments test whether the mechanism of this deformylation is similar to the mechanism utilized by the polysaccharide deacetylase family. The expectation is that these experiments will validate the role of ArnD in Lipid A-Ara4NH$_4^+$ biosynthesis and provide valuable information for development and evaluation of inhibitors. These may prove valuable leads for development of clinical inhibitors to treat *Pseudomonas* infections in CF patients and extend the clinical life of antibiotics such as colistin.

Results

Mutated Plasmid Design and Amplification

The plasmid used for gene introduction, cloning, and expression was a modified parent vector pET28a(+) containing a Kanamycin resistance cassette (Figure 14). This vector was modified in our lab to introduce additional restrictions sites and tags for purification resulting in plasmid pMS174. The gene coding for Salmonella ArnD had been previously cloned into this vector (M.S. Lee unpublished data) resulting in the full-length protein with a 6 X Histidine tag at its C-terminus (Figure 15).
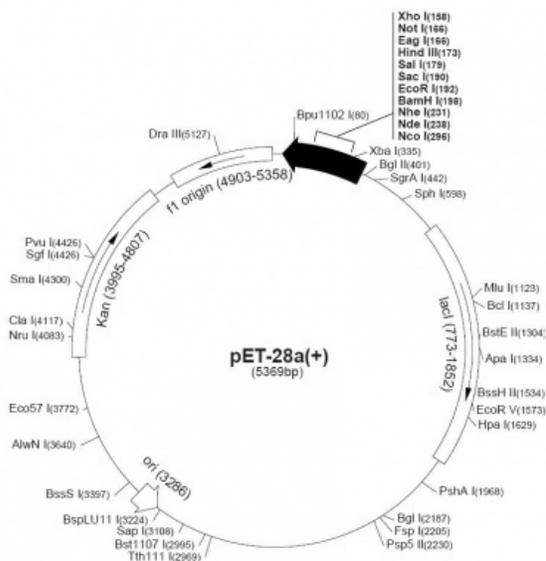

Figure 14: pET-28a(+) vector map. Important features are: MCS region, TEV-site, lac1, 6x His TAG at 5' end of MCS, and Kanamycin Resistant cassette. Figure taken from [3].
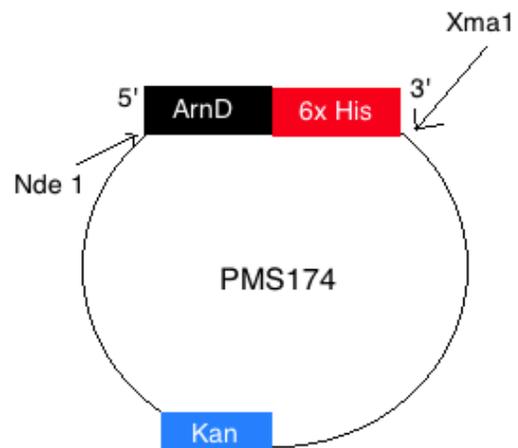
Figure 15: pMS174 vector map. Important features are: ArnD gene, 6x His TAG at 3' end of ArnD, and Kanamycin Resistant cassette.

As described in detail in materials and methods, the Quikchange method (Stratagene) was used to introduce point mutations in the ArnD expression plasmid. Oligonucleotide primers were designed to specifically target the codons within the ArnD gene, responsible for the addition of Asp_{43} and H_{233} into the polypeptide chain. Upper and lower primers were synthesized for each mutation: D43N, D43S, H233W, H233Y, H233R. These primers were ~40 bases in length, with a minimum GC content of 40%, and ~15 bases on either side of the desired mutation codon. These parameters ensure proper annealing to the parent vector during subsequent PCR thermocycling and amplification. Following PCR amplification of the plasmid, digestion with *Dpn1* endonuclease served to degrade the parent (wild type) vector so that only the new, mutated plasmid remained. Figure 16 shows the analysis of the mutagenized PCR products performed on 0.8% agarose gel electrophoresis, depicting clean amplified

products. The intensity and location of the bands confirms successful generation of the desired mutant plasmids.

The original vector was ~6Kb and the addition of the ArnD gene was roughly ~1.1 Kb in length. The mutated plasmids were used to transform XL-10 competent *E. coli* cells which were plated on LB medium supplemented with Kanamycin. Individual colonies were used to prepare plasmid DNA. Then, individual clones were selected and submitted for sequencing using the T7 promoter and terminator primers, confirming the presence of desired mutations in all five mutants.
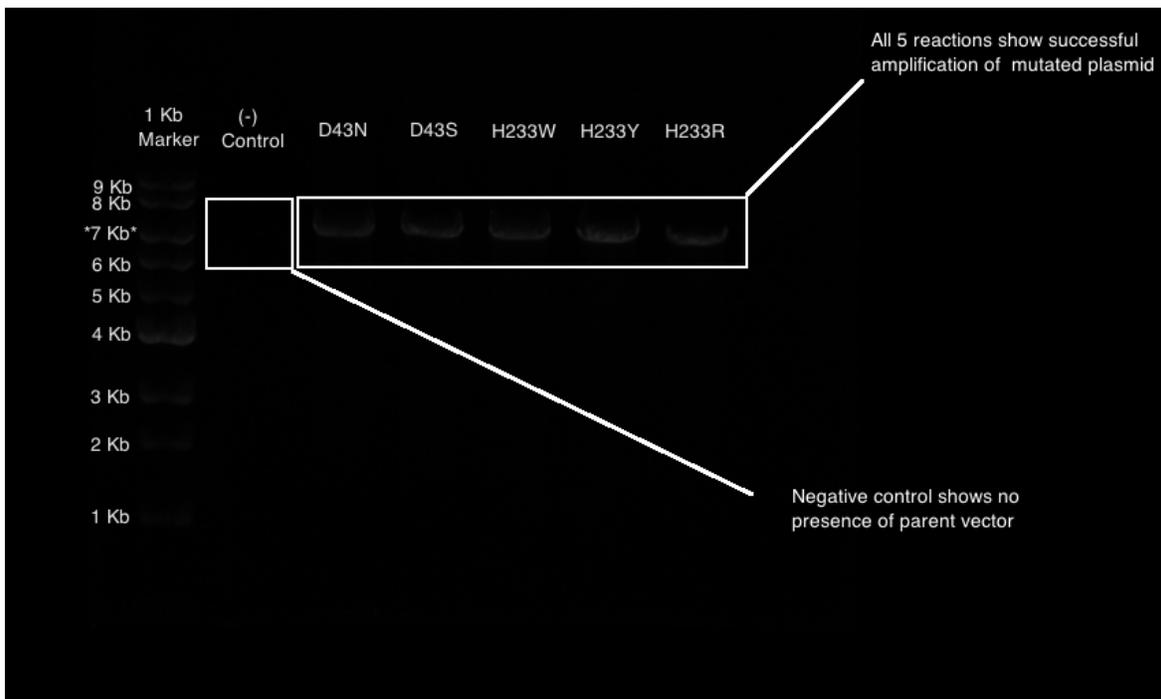

Figure 16: 0.8% agarose gel in 0.5% TAE Buffer stained by EtBr depicting PCR amplification of pMS174 vector containing ArnD gene using cycling parameters used under Stratagene protocol (see material and methods). The insert band was detected just above 7 kb on 0.8% agarose gel, which suggests proper insertion and amplification of the mutated ArnD gene vector ~7.1 kb in size. Negative control reaction represents an identical reaction run under the same conditions without the addition of Pfu turbo polymerase. 5 μL of sample was added to each well, ran at 135 V for ~30 min.

Protein Expression and Purification

Protein expression was carried out using *E. coli* Bl-21 competent cells as a host. Prior to large-scale protein preparations, small-scale expression tests were conducted to evaluate expression of the desired protein, ArnD. Several colonies were selected for each individual mutation and compared using SDS PAGE before and after expression induction (Figure 17).
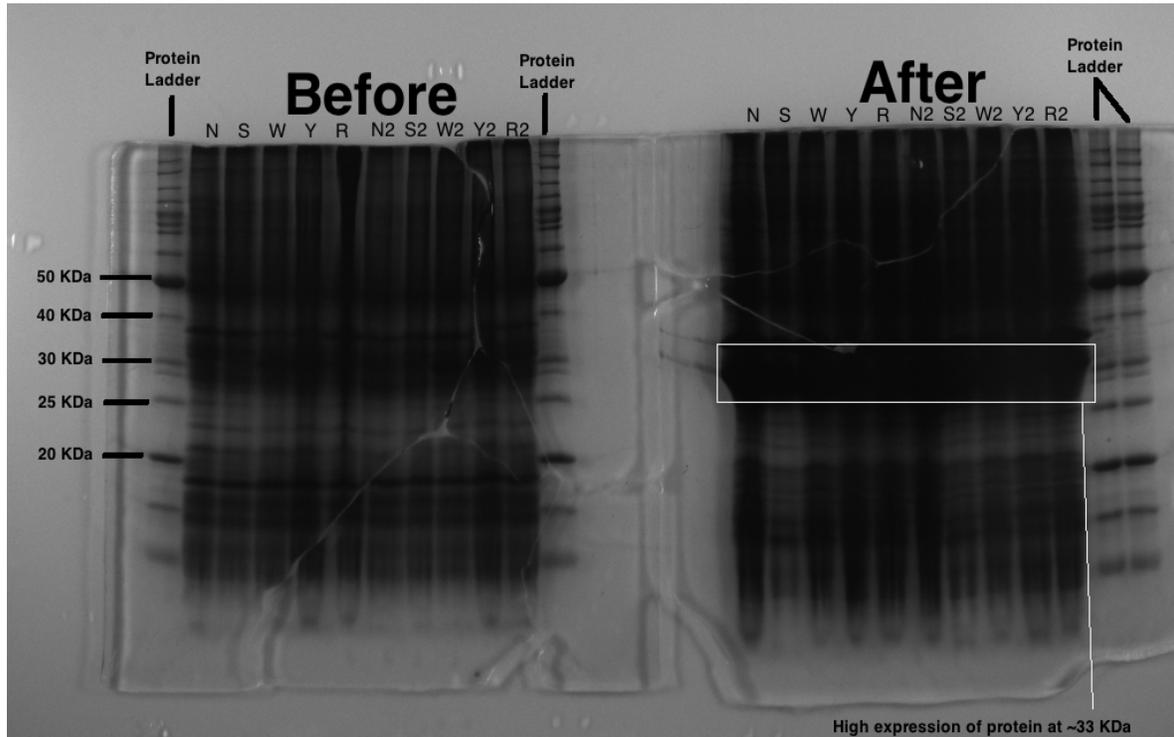

Figure 17: SDS-PAGE gel depicting expression optimization conditions of ArnD. Before Gel: Lane 1 is a standard protein marker, Lanes 2 through 11 contain before induction (-IPTG) samples of 2 separate colonies of the 5 different mutations. After Gel: Lanes 1- 10 depict the same 2 separate colonies of the 5 different mutation used in the before gel after induction (+IPTG) overnight at 20°C. All mutant constructs express a high level of protein at the desired location, however, only the colonies which expressed best were carried forward for large-scale expression and purification.

The basic protocol for protein purification includes: (1) cell lysis, (2) protein extraction from the membrane, (3) affinity purification using Ni-NTA agarose and (4) size exclusion chromatography (SEC) to evaluate the presence of aggregates in the purified sample.

Small-scale purification allowed for the refinement of expression conditions. Expression induction was carried out at various temperatures to maximize the yield of unaggregated protein as judged by SEC. The best two experimental conditions are shown in Figure 18. The results demonstrate the smallest proportion of aggregated protein (First peak around Fraction 28) as compared to unaggregated protein (second peak around Fraction 44) when the expression is carried out at 20 C, or approximately room temperature.
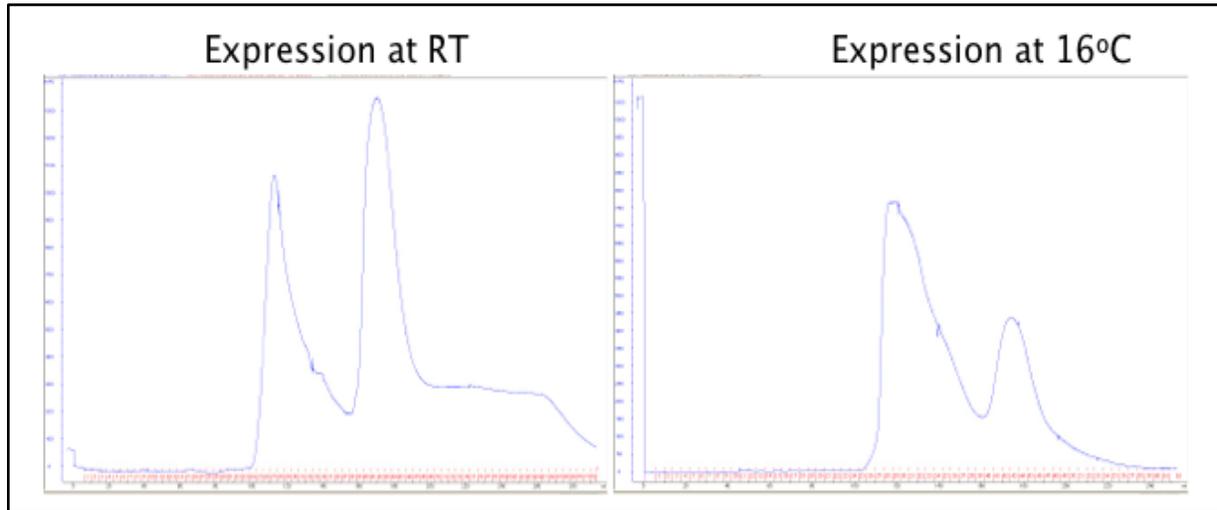

Figure 18: Y-Axis: Absorption at 280 nm; X-Axis: mL eluted/fraction #
Elution of WT ArnD expressed in two separate conditions after Ni-NTA Affinity Chromatography & HiLoad 26/60 Superdex 200 Size Exclusion Chromatography. In Both Graphs: there is an initial peak ~ Fraction 28 representing aggregate protein followed by a second peak of the desired detergent soluablized membrane protein at ~Fraction 44. At RT, there is an increase in absorbance of the second, solubalized ArnD, peak both relative to the first peak and in general. At RT, Aggregate: Soluabalized 1100 mAU: 1350 mAU; at 16° C, Aggregate: Soluabalized 800 mAU: 400 mAU.

Additional small-scale purification tests were performed to define the best detergent for membrane extraction (Figure 19). The effect of detergent in the protein yield was evaluated by SEC (using a Superdex 200 column) and SDS PAGE analysis of the fractions. As shown in Figure 19, Triton-X-100 solubilizes the membrane protein ArnD better than the detergents: *n*-Dodecyl β-D-maltoside (DDM), N,N-Dimethyldodecylamine N-oxide (DDAO), Deoxycholate, 3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS) , and n-Octyl-β-D-Glucopyranoside (OG) at the concentrations used. This conclusion was made based on the thick band indicating high protein concentration both after membrane solubilization, Ni-NTA elution and similarly after SEC. Furthermore, this high concentration of protein is also confirmed by spectral analysis of the eluted SEC column at 280 nm.
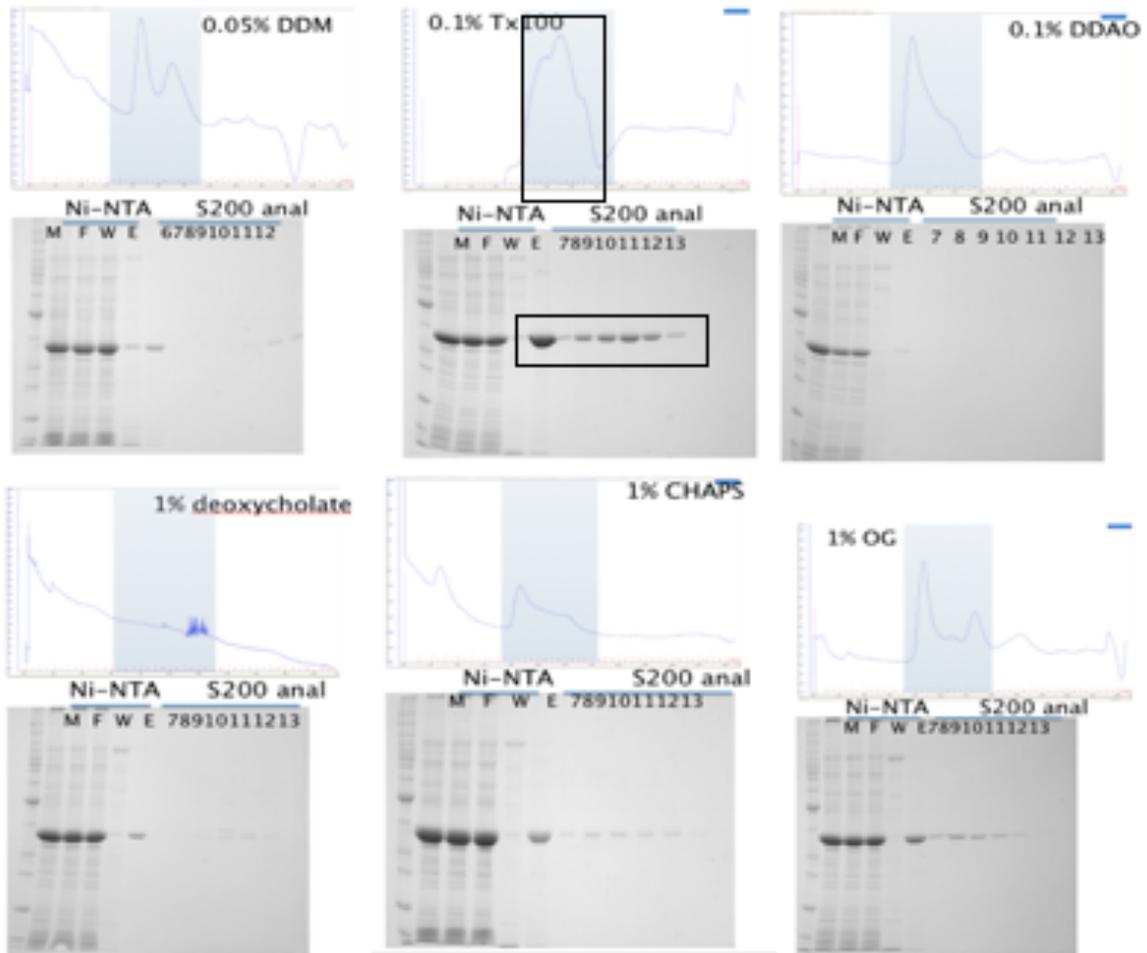
Figure 19: For each graph for each individual detergent: Y-Axis: Absorption at 280 nm;   X-Axis: mL eluted/fraction number

Elution of WT ArnD solubilized in six separate detergents after Ni-NTA Affinity Chromatography & HiLoad 26/60 Superdex 200 Size Exclusion Chromatography. Triton X-100 (top middle) was found to be the best detergent for the solubilization of the membrane associated enzyme ArnD. This can be seen in the various graphs where Triton X-100 has the most defined single absorbance peak at ~ Fraction 10 ~1400 mAU. This is further supported by SDS PAGE analysis. By loading the same volume it can be seen, by the thick band in lane E, that there is the highest protein concentration in Triton- X-100 detergent as opposed to the others tested. This is also further exemplified by the lanes 7-13, indicating a single peak of protein at ~33KDa, the size of WT ArnD.

Following the refinement of expression and purification of experimental parameters, each mutant, as well as WT ArnD, was expressed and purified under identical conditions. Approximately 6 L of *E. Coli* cells transformed with the appropriate plasmid were grown, induced and harvested. Wild-type and mutant ArnD were purified using Ni-NTA affinity chromatography. This method of purification is highly selective for the desired protein, ArnD, due to the high affinity of binding of the resin to the attached C-terminus 6 x His TAG on the ArnD constructs. Elution of ArnD, and ArnD mutants, is achieved by competitive binding of imidazole to the Ni-NTA resin. Along the expression and purification sequence, various aliquots were taken for SDS PAGE analysis to confirm proper expression and purification. Figure 20 shows a representative SDS-PAGE evaluating the large-scale purification of ArnD H233R. These purification procedures and corresponding gels were reproducible and essentially identical between each construct.
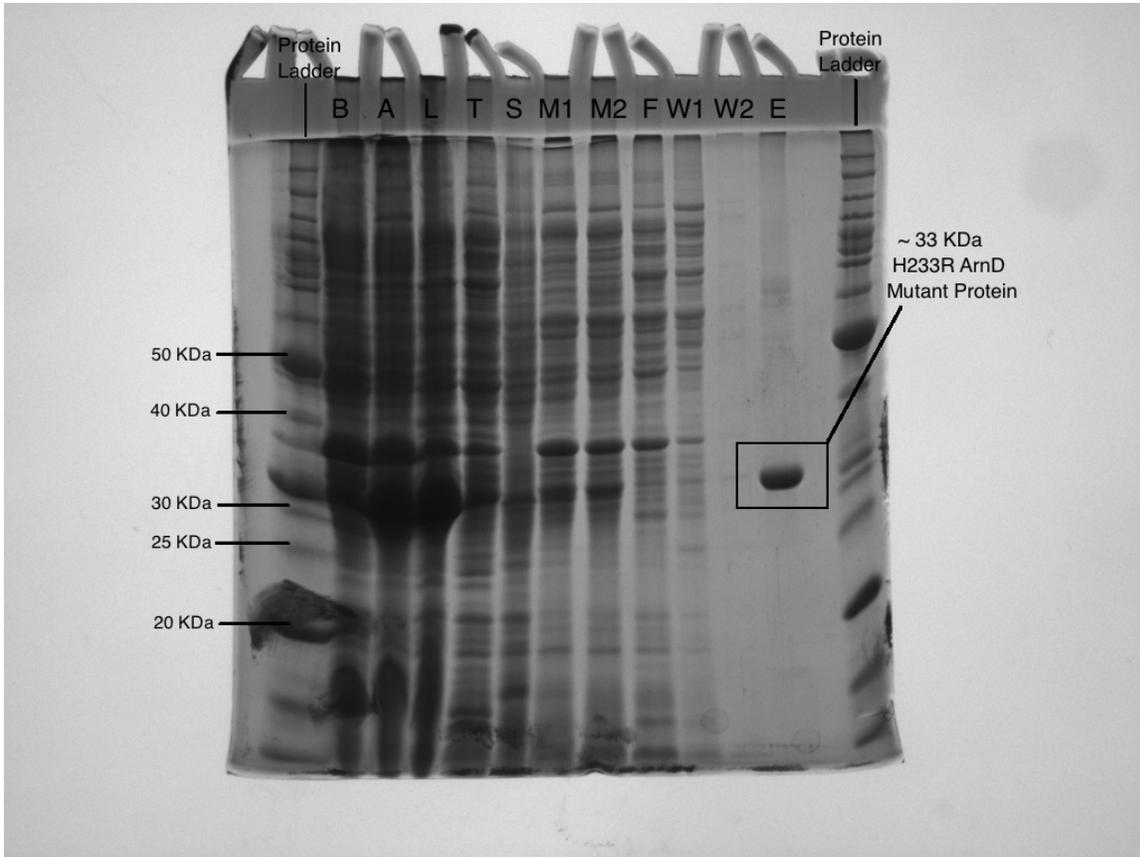
Figure 20: Similar gels were produced for WT ArnD as well as all 5 mutants:
12% SDS-PAGE gel depicting expression and purification of H233R mutant ArnD. (A detailed explanation of the various lanes can be found within materials and methods) Lane 1 is a standard protein marker, (B) is before induction (-IPTG). (A) is after induction (+IPTG), (L) is lysate, (T) is total soluble and membrane protein, (S) is soluble protein, (M1) is total membrane protein, (M2) is actual membrane protein, (F) is Flow-Through Ni-NTA, (W1) is wash with 0.1% Triton X-100, (W2) is wash with x5 CMC DDM, (E) is eluted protein from Ni-NTA. overnight at 20°C. All mutants as well as WT ArnD expressed a large amount of protein, followed by successful purification. This can be seen by looking at the Elution lane (E). There is a significant amount of desired protein just above the 30 KDa marker. This is representative of the ~33KDa H233R mutant ArnD construct.

As can be seen in Figure 20, there is a protein band located slightly above the 30 KDa marker. This band corresponds with the ~33KDa enzyme ArnD. After induction with ITPG and expression overnight there is a dramatic increase in the concentration of ArnD present. Cell lysis was performed with a French Press. Sonication should be avoided because membrane proteins tend to aggregate upon sonication. A low speed centrifugation step removed cell debris and a large amount of ArnD. This removal is most likely due to misfolding and association into inclusion bodies. The remaining fraction undergoes two centrifugation steps resulting in a membrane pellet. This pellet is solubilized using the detergent Triton X-100, ran through a Ni-NTA, and eluted with imidazole (lane E Figure 20). This column represents the elution of ArnD, at ~33KDa in the presence of very little contaminant protein. This sample was concentrated and subjected to Size Exclusion Chromatography.

Figure 21 depicts the final SEC step for the large-scale purification of ArnD. The SDS PAGE analysis of the fractions show that ArnD is highly purified. As before, these gels were highly reproducible and relatively identical between each construct; all mutant and WT constructs demonstrated similar protein yield, purity and location during elution. The purified proteins were frozen at -70 C for later activity analysis.
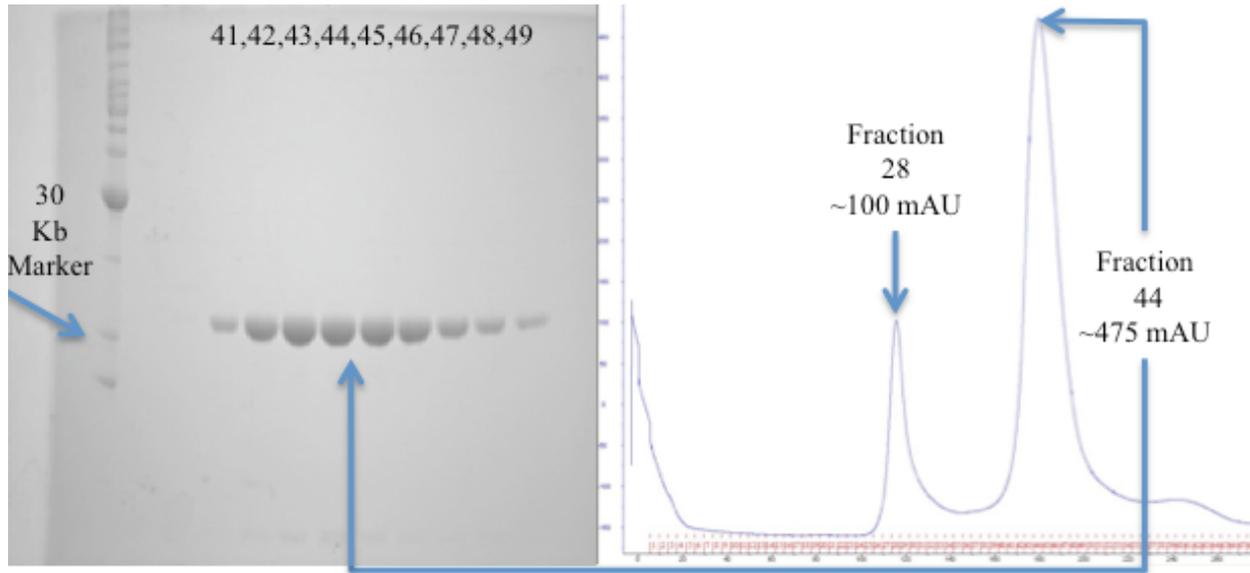

Figure 21: Similar gels and elutions were produced for WT ArnD as well as the 5 mutants:
Left:12% SDS-PAGE gel depicting purification of WT ArnD. (A detailed explanation of the various lanes can be found within materials and methods) There are 9 Fractions: 41-49 all containing some concentration of desired protein just above the 30 KDa marker. This is representative of the ~33KDa WT ArnD construct. The top of the peak, where there was maximum absorbance at 280 nm, was at the end of Fraction 44 This can be visualized by the thick bands in lanes 44 and 45.
Right: UV-Vis Spectral analysis at 280 nm of S200 SEC representative of all constructs both WT and Mutant. As can be seen by the figure there is a minimal first peak located around Fraction 28 which represents aggregated protein and other contaminants. The next sharp high absorption monomeric peak demonstrates a high concentration of purified ArnD mutant protein for activity analysis.

ArnD Substrate Formation

Due to the lack of a commercially available substrate for ArnD, I had to produce it. To this end, I took advantage of radioactively labeled UDP-glucuronic acid and purified enzymes of the Arn pathway to enzymatically produce undecaprenyl phosphate-α-L-Ara4N-formyl. The progress of the reaction was evaluated using Thin Layer Chromatography (TLC) and exposure to a PhosphoImager screen (Figure 22).
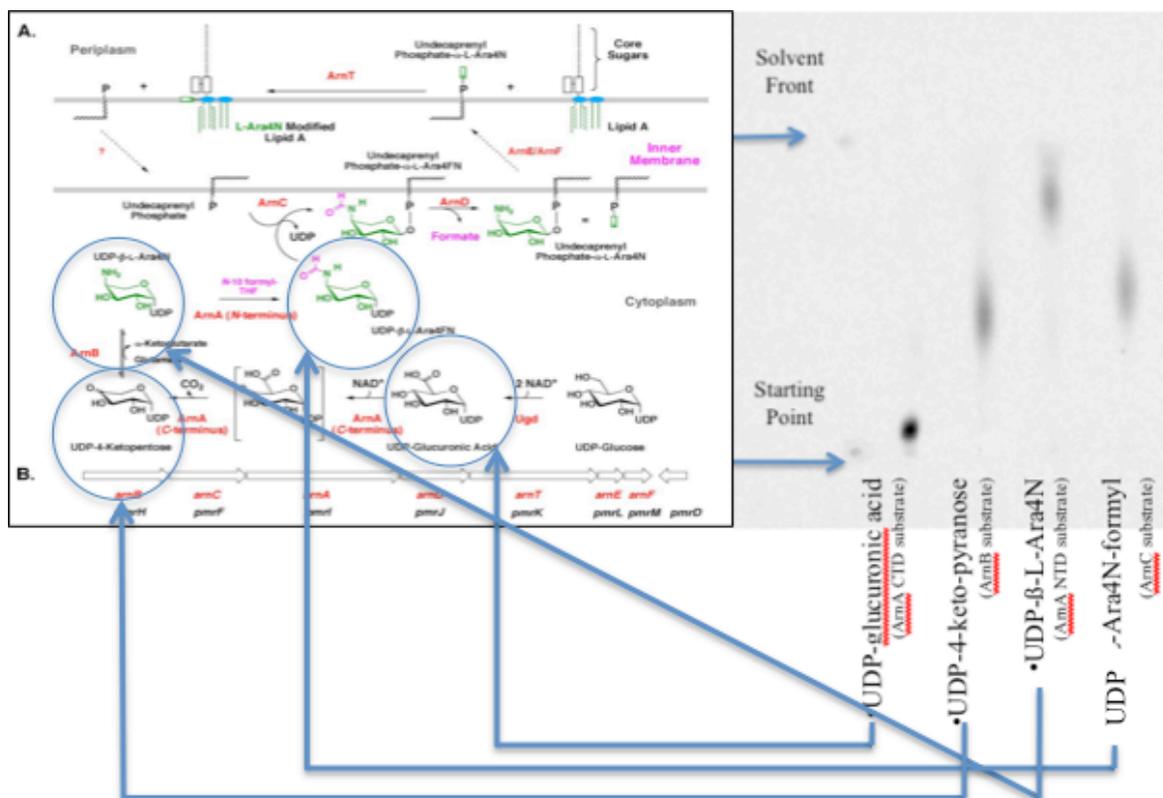

Figure 22: Left: The Arn Reaction pathway, in order to give reference to the TLC plate
Right: PEI Cellulose TLC plate ran in .4 M LiCl and .25 M Acetic Acid. The lanes contain: the initial substrate (UDP-[^{14}C]glucuronic acid, first from the left) and the products after sequential incubation of the initial substrate with ArnA-CTD (dehydrogenase), ArnB, and ArnA-NTD (transformylase) (second, third and fourth lane, respectively). The figure illustrates the relatively complete conversion of substrate to product in each step.

It can be seen by comparing the first and second lanes in Figure 22 that there is complete conversion to product catalyzed by the C-terminal (dehydrogenase) domain of ArnA in the presence of exogenous NAD$^+$. The product, UDP-4-keto-pyranose, is the substrate for ArnB. The transamination reaction catalyzed by ArnB in the presence of glutamate as the amine-donor, has an equilibrium constant of approximately 0.1 [33]. Thus, in order to generate a complete conversion of reactant to product a relatively high concentration of Glu was used to drive the reaction forward. The ArnB product, UDP-ß-L-Ara4N, can be visualized in the third lane with the farthest migration. This molecule has a protonated primary amine which holds a positive charge within the solvent system used and does not interact strongly with the PEI stationary phase. The last naturally soluble product in the Arn pathway is UDP-L-Ara4N-formyl, represented by the final lane of the TLC plate.

Similar to the previous lanes, there appears to be a complete conversion of reactant to the formylated product after the addition of full length ArnA to the reaction mixture along with exogenous N-10-methenyltetrahydrofolate. The Rf values for UDP-Glucuronic Acid, UDP- 4-keto-pyranose, UDP-ß-L-Ara4N, and UDP-L-Ara4N-formyl are 0.09, 0.45, 0.85 and .54 respectively. These results were obtained several times by myself and are consistent with the Rf values available in the literature where the identity of the products was verified by NMR and mass spectrometry [31, 33, 41].

The next step in obtaining the substrate for ArnD was the transfer of the N-formylated sugar to the undecaprenyl phosphate carrier. This was accomplished by mixing the previous reaction products with undecaprenyl phosphate solubilized in Triton X-100 and purified ArnC [41]. The conversion to the lipid-anchored sugar was monitored by TLC in Silica plates followed by PhosphoImager analysis (Figure 23). As before, the conversion to Undecaprenyl-Phosphate-L-Ara-4N-formyl appears fairly complete, and the Rf= 0.37 of the product is consistent with the literature [41].
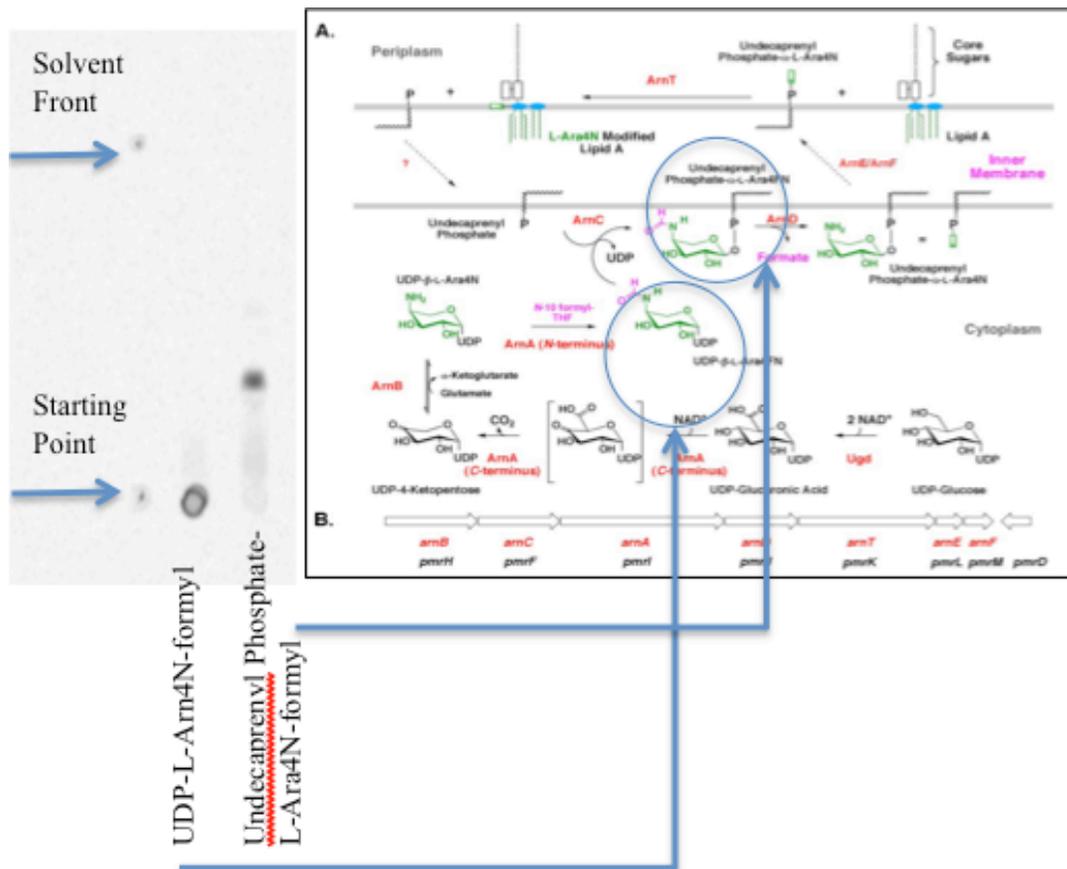
Figure 23: Left: The Arn Reaction pathway, in order to give reference to the TLC plate
Right: TLC Silica gel 60 F_{254} plate ran in (64/25/3.6/0.4) ($CHCl_3$/MeOH/H_2O/NH_4OH). The two lanes represent the migration of product along the Arn pathway. Starting with ArnA-NTD Product, UDP-L-Arn4N-formyl, first from left and the product after incubation with ArnC and Undecaprenyl Phosphate resuspended in Triton X-100. The figure illustrates the relatively complete conversion of substrate to product.

Evaluation of ArnD activity

With radiolabeled Undecaprenyl-Phosphate-L-Ara-4N-formyl in hand, I was now in a position to test if ArnD can convert this substrate to the deformylated product. In complementary experiments in our lab, Dr. MyeongSeon Lee was able to crystallize ArnD for structure determination by X-ray crystallography. During his initial characterization of the crystals, he observed X-ray absorption and fluorescence characteristic of Co^{2+} (Lee M.S. and Sousa M.C. unpublished results). These results are consistent with the hypothesis that ArnD is a metal dependent hydrolase and strongly suggests that the native metal was indeed Co^{2+}. Therefore, I assayed the activity of ArnD supplementing the reaction mixture with 0.5 mM $CoCl_2$ and different concentrations of purified, WT ArnD (Figure 24).
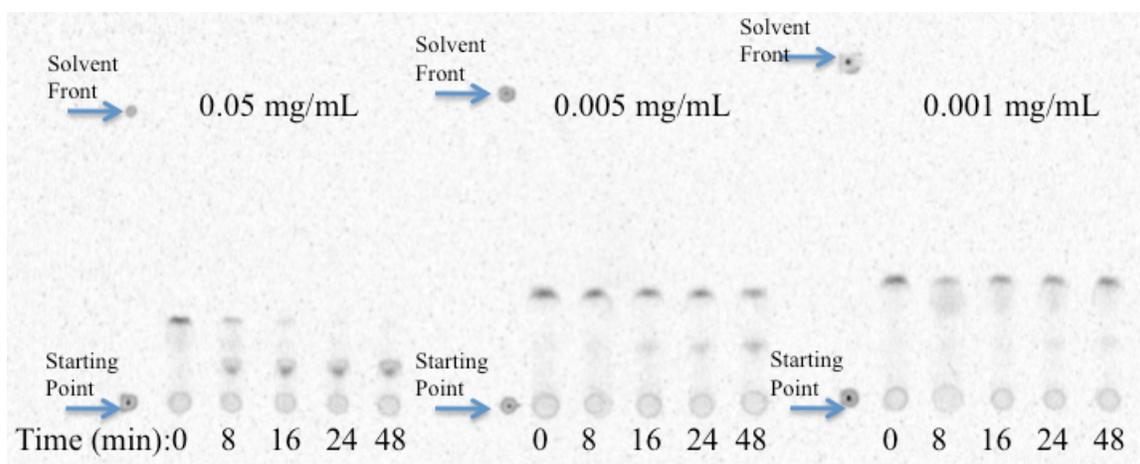

Figure 24: TLC Analysis of varying amounts of ArnD incubated with 0.5 mM $CoCl_2$ and approximately 0.8μM Undecaprenylphosphate-L-[^{14}C]Ara4N-formyl (in a mixture containing all the preceding Arn pathway enzymes and reactants).

As shown in Figure 24, there is clear formation of a species with migration slower from that of the initial substrate (time t=0). The substrate is fully converted to this slower-migrating species in 48 minutes in the presence of 0.05 mg/ml ArnD. The amount of slower-migrating species formed decreases drastically as the concentration of WT ArnD is reduced. This suggests that the slower-migrating product is due to a reaction catalyzed by ArnD. This reaction is proposed to be the deformylation of Undecprenyl-Phosphate-L-Ara4N-formyl to Undecaprenyl-Phosphate-L-Ara4N.

I utilized a similar reaction mixture and strategy to evaluate the activity of the five ArnD mutants. As seen in Figure 25, the reaction for WT ArnD was reproducible and yielded essentially complete conversion into product within 48 minutes. Conversely, all five mutants showed dramatic reduction of activity. However, D43N appears to have some residual activity as a small amount of product can be seen in the later time points (Figure 25, panel D43N).

These results are consistent with my hypothesis that ArnD is an enzyme which catalyzes a reaction with the substrate undecaprenyl-phosphate-Ara4(N-formyl) to a slower-migrating species, we expect this reaction to be a deformylation producing undecaprenyl-phosphate-Ara4N. These results are also consistent with the notion that ArnD uses a catalytic strategy similar to that of polysaccharide deacetylases. This is supported by the presence of Co^{2+} in natively folded ArnD as well as the explicit loss of activity in all mutants of the putative catalytic duo aspartate 44 and histidine 233 characteristic of these enzymes.
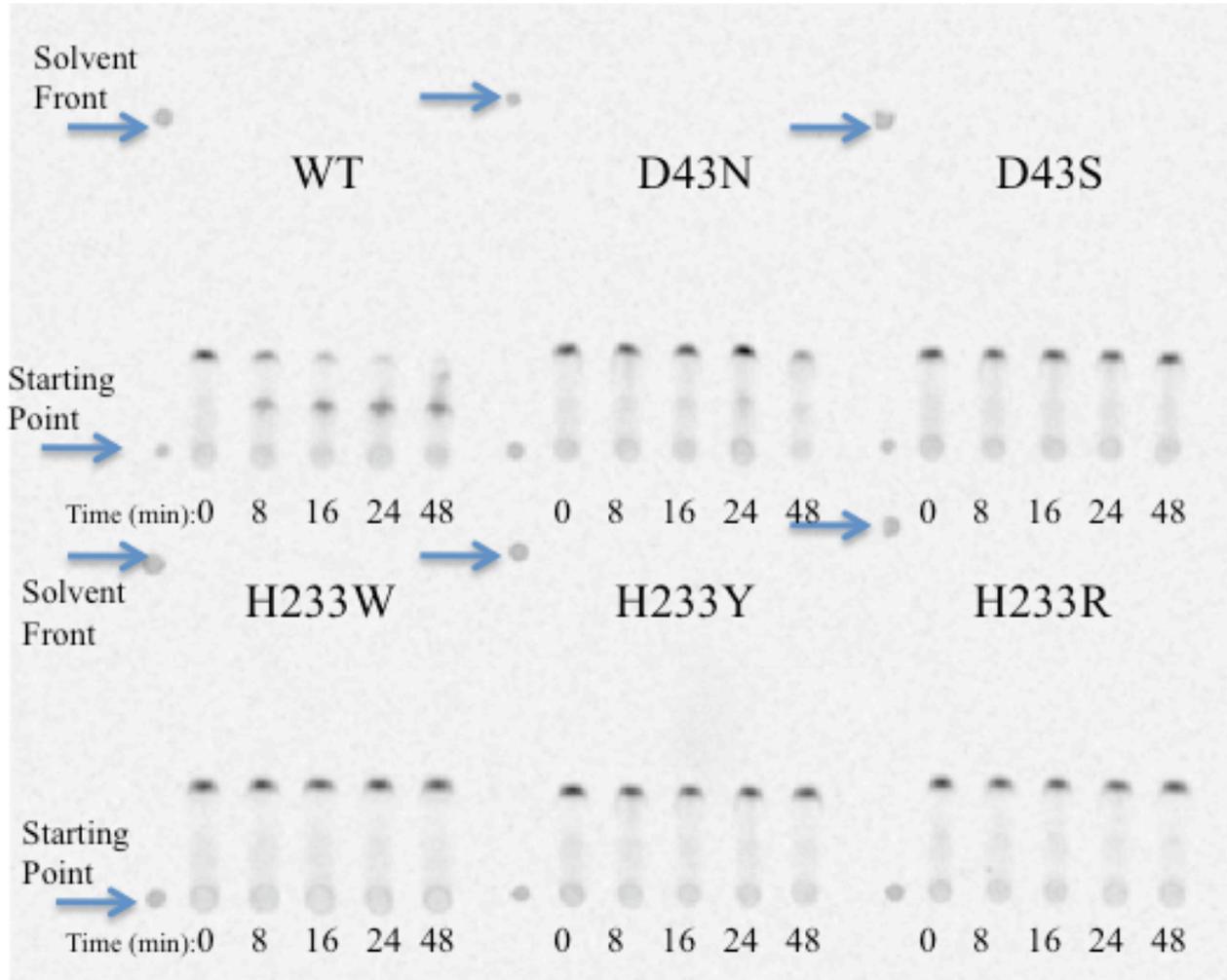

Figure 26: Identical assay as preformed in Figure 25, with duplicate results of WT ArnD and results of the 5 mutant ArnD constructs:
TLC Analysis of ArnD activity, 0.05 mg/mL WT ArnD and 5 mutants incubated individually with 0.5 mM $CoCl_2$ and approximately 0.8μM Undecaprenylphosphate-L-[^{14}C]Ara4N-formyl (in a mixture containing all the preceding Arn pathway enzymes and reactants).

Discussion

In the experiments described above there was successful introduction of point mutations in the ArnD expression plasmid. The intensity and location of the vector bands after PCR amplification confirmed successful generation of the desired mutant plasmids and was further confirmed by sequencing. One of the potential problems of making point mutations is improper annealing to the WT gene. Since there is not proper Watson-Crick base pairing at the point mutation site it is possible that the oligonucleotide primer will not bind. In addition, it is also possible that during PCR amplification random undesired mutations are introduced along with the desired ones. This fact was taken in to consideration and thus two separate reactions for each mutation were submitted for sequencing. Sequencing results confirmed the absence of undesired mutations, and the presence of the designed ones.

Intrinsic with the introduction of point mutations to WT proteins is the possibility of misfolding. The exact structural and electrostatic properties of each individual residue are unique and thus changing even a single residue has the potential to disrupt the native fold of the enzyme. However, this has not prevented the use of point mutation studies in regard to determining putative catalytic residues. Mutagenic analysis has led to the determination of catalytic residues for many enzymes both within our lab as well as labs across the globe. One way that point mutations are typically introduced is via the replacement of any given residue with alanine. By replacing the side chain on any given amino acid with a simple methyl group, as long as the side chain is active in catalysis, this typically inhibits the enzyme activity. Despite this, in regard to the two residues mutated within this experiment, aspartate and histidine, there is very little structural or electrostatic similarity to alanine. Due to these varying features, it is unlikely that His or Asp would inhabit a similar enzymatic environment as Ala would. Thus, in both cases, aspartate and histidine, were not mutated to Ala because of the potential to disrupt the enzyme folding. Instead, the mutations were designed to minimize the disturbance of the proteins native fold. In theory, these mutations have a greater probability mimicking a relatively similar environment of the active site both spatially and electrostatically from mutant ArnD as compared to WT ArnD. Such an environment would further promote proper substrate/enzyme binding and allow a better evaluation of the specific role of the side chains in catalysis.

By using a series of mutations for a single residue, one increases the likelihood of obtaining a loss of activity not due to protein misfolding, but rather due to the loss of a crucial catalytic residue. The mutation D43S is a mutation determined by Jonson et. Al to have a 0.9-1.0 degree of correlation (Figure 13), the highest range possible. This would then suggest that a mutation from Asp to Ser would have minimal disturbance as compared to a mutation of Asp to another residue. Despite this, it seems possible that the different structural properties of the side chains of aspartate and serine could perturb the putative active site. Therefore, I also mutated D43 to asparagine, which is nearly identical to aspartate, with only the relatively mild alteration of a single functional group from carboxylic acid to an amide. Thus, a D43N point mutation is one of the most conservative mutations possible.

Nevertheless, protein misfolding is always a possibility when point mutations are being introduced. The only way one can confirm that the native fold of the enzyme is conserved and not perturbed by the mutations is through crystallization studies followed by structure determination. Despite the fact that we cannot say positively at this moment in time that the mutants folded correctly, there are distinctive clues which lead us to this notion. First, after lysing, the cellular lysate was centrifuged in order to remove any of the unlysed cells, various cellular debris, and inclusion bodies containing a majority of the misfolded enzymes within the cell. It can be seen in Figure 20 that total protein concentration, from the cellular lysate to the after centrifuged product, is decreased dramatically due to the removal of all misfolded enzymes during the expression period. Second, the S200 size exclusion analysis also suggests a properly folded enzyme. Unfolded proteins that are were not removed in inclusions bodies remain as aggregated proteins. A misfolded enzyme often has the potential of exposing normally hidden hydrophobic residues, which results in aggregation and elution early in size exclusion chromatography. This aggregation can be visualized in Figure 21 in the UV trace of the S200 elution. The first, smaller absorption peak, is representative of the earlier, larger, aggregated elution at Fraction 28. The second higher absorption peak is representative of the solubilized membrane protein ArnD in WT and all 5 mutant forms. Finally when looking at the purified WT ArnD alongside purified mutant ArnD the solutions look identical. There is no cloudiness in the either solution, which is a common visual indicator of protein aggregation.

Assuming that there was a not intrinsic loss of enzymatic activity due to the misfolding of the 5 mutants, it must still then be determined how the mutants produce a loss of activity. It was stated that the specific point mutations made were the most conservative available. Is it possible then that these replaced residues can also support the same function in catalysis?

The D43N mutation replaces the active base in catalysis with a side chain pKa ~4 to an amide. The amide group replacing the carboxylic acid is a very weak base, however, it is possible, that Asn positions the water molecule correctly in the active site and still supports, with much lower efficiency, the hydrolysis of the formate group of Undecaprenyl Phosphate-L-Ara-4N-formyl. This could explain the residual activity observed for D43N in Figure 25.

For D43S the loss of catalysis seems more obvious, it is unlikely that the serine side chain will act as a general base deprotonating the water molecule to become a positively charged species. Although, in some circumstances the pKa of the side chain alcohol can be reduced within the enzyme allowing it's alkoxide form to act as a base, it can be seen in Figure 25 that the D43S mutation does not support conversion of the ArnD substrate to product. Similarly, mutation of H233 to W, Y or R confers a dramatic loss of activity. The histidine which has a pKa ~7 depending on the relative environment is an ideal molecule for proton donation to the developing product. W, Y, R side chain available protons have the relative pKas of approximately 21, 17 and 12, respectively. Therefore, these residues would not be good general acids, consistent with the lack of activity seen in the H233 mutations.

The TLC analysis by PhosphoImaging demonstrated a clear, full, conversion of ArnC product and ArnD substrate to a new product; confirmation of the identity of this product will require MS and NMR. However, we believe that ArnD did indeed deformylate Undecaprenyl Phosphate Ara4N-formyl. The direct increase in product formation with all other variables held constant suggests that ArnD is indeed catalyzing a reaction and it is not some contaminant. In addition, on a Silica gel plate, polar molecules bind to the stationary phase while less charged molecule migrate further up the solvent front. This trend is consistent in our proposed identity of the migrating species.

Due to the novel nature of these results from any published data, a number of future studies can be devised. Crucial to confirming the above stated results is a scale-up of the reactions with non-radioactively labeled substrates, purification by HPLC and submission to Mass Spectrometry and NMR analysis. Furthermore these mutants should be crystallized under the same conditions which WT ArnD had been successfully crystallized for structure determination by X-Ray diffraction. These studies will serve to undoubtedly confirm the identity of ArnD as a deformylase with His_{233} and Asp_{43} active in catalysis.

My results are consistent with the reaction mechanism proposed for polysaccharide deacytlases shown in Figure 13. Therefore, analogs mimicking the tetrahedral transition-state may prove to be good inhibitors of the reaction mechanism.

In order to give insight into potential specific inhibitor design, the class of zinc proteases can be utilized. These enzymes, similarly to ArnD, are metalloenzymes which catalyze a hydrolysis reaction. For this reason, specific inhibitors, designed for zinc proteases, which target the divalent metal as a crucial member of catalysis, could be extrapolated to be useful within ArnD. In addition to this, transition-state analogs could be created that mimic the tetrahedral intermediate undergone in the reaction mechanism after nucleophilic attack of the activated water molecule on the carbonyl carbon of the deformylated product.

Materials & Methods

Site-Directed Mutagenesis: Using the JSH-C vector generously received from MyeongSeon Lee, five individual point mutations were introduced at two separate locations. These point mutations were made using mutagenic oligonucleotide primers amplified by PCR according to the Quickchange method (Stratagene protocol). First, the two primers, upper and lower, for each point mutation were designed and ordered from (Integrated DNA Technologies [IDT], Inc.) (Table 1). Next, the ordered HPLC-purified oligonucleotide primers were used to set up mutagenic PCRs containing:
5 µL of 10X Pfu Buffer, ~20 ng of JSH-C DNA vector, 125 ng of oligonucleotide primer (both upper and lower), 1 µL 10 µM dNTP, and ~40 µL of ddH2O to obtain a final volume of 49 µL. Lastly, 1 µL of Pfu turbo polymerase (Stratagene)was added followed by subsequent PCR thermocycling. Thermocycling parameters used were:
(1) 95ºC for 3 minutes, (2) 95ºC for 1 minute, (3) 55ºC for 1 minute, (4) 72ºC for 13 minutes (2 min/kb plasmid), repeat steps 2-4 16X, (5) 72ºC for 10 minutes, (6) 4ºC until sample is used for analysis and further experimentation. Further, the template plasmid DNA was digested by adding 1 µL *Dpn I* restriction endonuclease (New England BioLabs, Inc.) and incubating for 1 h at 37°C to hydrolyze then adding 1 additional µL Dpn I and incubating overnight. DNA Agarose Gels: In order to measure for sufficient and appropriate amplification of our mutated vector: 10µl of all the crude PCR products were run in Tris-Acetate-EDTA (TAE) buffer (0.04 M Tris, 0.001 M EDTA, 0.02 M acetic acid) at 160 V on a 0.8 % agarose gel supplemented with intercalater EtBr.

Mutation	Primer Sequences (mutant codons are bold and underlined, coding strands are presented above non-coding strands, all strands written 5'-3')
D43N	5'-GCT TTT TCT TCA GCG TCG GGC CG**A AC**A ATA TTG GAC GTC ATC TTT GG-3' 5'-CCA AAG ATG ACG TCC CAT ATT **GTT** CGG CCC GAC GCT GAA GAA AAA GC-3'
D43S	5'-GCT TTT TCT TCA GCG TCG GGC CG**A GC**A ATA TTG GAC GTC ATC TTT GG-3' 5'-CCA AAG ATG ACG TCC CAT ATT **GCT** CGG CCC GAC GCT GAA GAA AAA GC-3'
H233W	5'-GCA CGC CGG TAT ATA CCA TC**T GG**G CGG AAG TCG AAG GTA TTG TCC-3' 5'-GGA CAA TAC CTT CGA CTT CCG C**CC A**GA TGG TAT ATA CCG GCG TGC-3'
H233Y	5'-GCA CGC CGG TAT ATA CCA TC**T AT**G CGG AAG TCG AAG GTA TTG TCC-3' 5'-GGA CAA TAC CTT CGA CTT CCG C**AT A**GA TGG TAT ATA CCG GCG TGC-3'
H233R	5'-GCA CGC CGG TAT ATA CCA TC**C GT**G CGG AAG TCG AAG GTA TTG TCC-3' 5'-GGA CAA TAC CTT CGA CTT CCG C**AC G**GA TGG TAT ATA CCG GCG TGC-3'

Table 1: Primer Sequences for Mutagenic PCR Amplification

Small-Scale Expression of ArnD Mutants: Similar to the transformation procedure in XL-10 cells heat shock was used in order to uptake the mutated plasmid. A mixture of 20 µL BL-21 Rosetta cells (Novagen) and 1 µL of purified plasmid DNA was incubated on ice for 30 min, then heated for 30 sec at 42°C, and last incubated 2 min on ice. After 180 uL of LB broth containing 50 µg/ml kanamycin was added, the cells were incubated for 1 hour in a 37°C shaker and the 200 µL was plated on an LB selective growth plate supplemented with 50 µg/mL kanamycin. The plate was incubated overnight at 37°C. After colonial growth, a single colony was inoculated in ~4 mL LB broth supplemented with 50 µg/mL kanamycin, grown in a 37°C shaker to an OD_{600} of ~0.6, and cooled on ice for 10 min. Prior to expression a 20 µL aliquot was taken for SDS-PAGE analyzation. Expression was induced by the addition of 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG, Gold Bio Technology Inc.), and cultures were allowed to grow overnight in a 20°C shaker. After induction, the remaining 180 µL aliquot was centrifuged at 13,000 rpm for 10 min, the supernatant was removed, and the pellet was resuspended in 20 µL of MilliQ ddH_2O.

Large Scale Expression & Purification of ArnD Mutants: The following procedure was preformed independently for all 5 mutants:
A mixture of 20 µL BL-21 Rosetta cells (Novagen) and 1 µL of purified mutated plasmid DNA was incubated on ice for 30 min, then heated for 30 sec at 42°C, and last incubated 2 min on ice. After, 180 uL of LB broth containing 100µg/ml kanamycin was added, the cells were incubated for 1 hour in a 37°C shaker and the 200 µL was plated on an LB selective growth plate supplemented with 50 µg/mL

Kanamycin. The plate was incubated overnight at 37°C. After colonial growth, a single colony was inoculated in ~4 mL LB broth supplemented with 50 µg/mL Kanamycin, grown in a 37°C shaker to an OD_{600} of ~0.6, and cooled on ice for 10 min. This mixture was subsequently mixed in equal parts with 100% glycerol to obtain a 50% glycerol stock and stored at -70° C for further experimentation. Subsequently, each mutant was grown independently in 65 mL LB broth supplemented with 50 µg/mL Kanamycin by addition of a small amount (~ >1 µL) of 50% glyerol cell culture. This culture was incubated overnight at 37°C, then 10 mL of the overnight culture was used to inoculate 6 x 1 L LB broth flasks supplemented with 50 µg/mL Kanamycin. These cultures were incubated at 37°C to an OD_{600} of ~0.6, and cooled on ice for 30 min. A 1 mL aliquot was removed and labeled Before Expression (B). Expression was induced by the addition of 0.5 mM IPTG to each of the 6 x 1 L cultures. Expression proceeded overnight at Room Temperature (RT) ~20° C. A 0.5 mL aliquot was removed and labeled After Expression (A).Cells were then harvested by means of centrifugation at 5,000 RPM for 30 min. at 4° C. The cell pellet was resuspended in buffer containing: 25 mM Tris-HCl at pH 8.0, 300 mM NaCl , 5 mM BME, 10% Glycerol, and complete EDTA-free protease inhibitor cocktail (Roche) used at 1 tablet per 100 mL of buffer. Cells were lysed by use of French Press at a 1000 PSI. A 20 µL aliquot was removed and labeled Lysis (L). Excess cell debris was removed from the cell lysis by use of a low-speed centrifugation at 15,000xg for 30 min at 4° C. After centrifugation, a 20 µL aliquot of supernatant was removed and labeled total protein, containing both the soluble and membrane proteins (T). The supernatant was then removed for ultra-centrifugation at 40,000 RPM for 3 Hrs at 4° C. After ultra-centrifugation, a 20 µL aliquot of the supernatant was removed and labeled soluble protein (S). The pellet was then removed and dissolved in resuspension buffer containing: 25 mM Tris-HCl at pH 8.0, 300 mM NaCl , 5 mM BME, 10% Glycerol, 1% Triton X-100 (Sigma). A 20 µL aliquot of the dissolved pellet was removed and labeled total membrane fraction (M1). A high-speed centrifugation was then preformed at 40,000 x g for 30 min at 4° C to remove all undissolved particles. A 20 µL aliquot of the supernatant was removed and labeled actual membrane fraction (M2). The supernatant was further purified by use of a ~10 mL Ni-NTA agarose column (Qiagen) pre-equilibrated with: 25 mM Tris-HCl at pH 8.0, 300 mM NaCl, 5 mM BME, 10% Glycerol, 1% Triton X-100. ~4 column volumes of M2, membrane protein supernatant, was added to the column (F). The column was then washed to dissociate non-specific binding of undesired proteins first by running ~4.5 column volumes of: 25 mM Tris-HCl at pH 8.0, 300 mM NaCl, 5 mM BME, 10% Glycerol, 0.1% Triton X-100, and 15 mM imidazole (W1). Next, a second wash was performed to exchange the detergent surrounding ArnD by running ~4.5 column volumes of: 25 mM Tris-HCl at pH 8.0, 150 mM NaCl, 5 mM BME, 10% Glycerol , 0.0435% or x5 CMC DDM, and 15 mM imidazole (W2). ArnD was eluted by running ~3.5 column volumes of: 25 mM Tris-HCl at pH 8.0 , 300 mM NaCl, 5 mM BME, 10% Glycerol, 0.0435% or x5 CMC DDM, and 300 mM imidazole (E). This eluted sample of ~35 mL was concentrated to ~5mL using a 50 KDa MWCO concentrator, and then the concentrated protein was further purified by use of a S200 size exclusion chromatography. The sample was first filtered through 0.20 µm nylon syringe filter, and next loaded on HiLoad 26/60 Superdex 200 Prep Grade Column (Amersham Pharmacia Biotech) pre-equilibrated with: 25 mM Tris-HCl at pH 8.0, 300 mM NaCl, 5 mM BME, 10% Glycerol, 0.1% Triton X-100, and 15 mM imidazole. In an identical buffer the enzyme was eluted into 4 mL fractions measuring absorbance by use of a spectrophotometer at 280 nm. These fractions as well as aliquoted portions along the purification process were analyzed via SDS PAGE. The fractions containing purified protein were pooled and concentrated using a 50 KDa MWCO concentrator, the total protein concentration was determined by a Lowry assay. The purified protein was stored at -70° C until further experimentation.

Sodium-Dodecyl- Sulfate Polyacrilamide Gel Electrophoresis (SDS PAGE): The following gel was made for the analysis of all protein samples: ~7.5 mL running gel: 12% Acrylamide mix, Tris-HCl, pH 8.8, 0.1% SDS, 0.1% Ammonium Persulfate (APS), 0.003 mL TEMED, 2.475 mL ddH20; ~1mL Stacking Gel: 5% Acrylamide mix, Tris-HCl, pH 6.5, 0.1% SDS, 0.1% Ammonium Persulfate (APS), 0.001 mL TEMED, 0.68 mL ddH20. 5 uL of protein ladder was added to one lane while 10 uL of sample was added to the other various wells. The gel was ran for ~1 Hr at a constant voltage of 150V. The gel was then stained for ~1Hr in 50% MeOH 10% Acetic Acid and 0.25% Brilliant Blue R-250 and then destained in ddH20.

Lowry Assay: First, x20, x50, x200 dilutions were made of the purified WT ArnD as well as the various 5 mutant constructs. Then using Bovine Serum Albumin (BSA) 400 mg/mL, 300 mg/mL, 200 mg/mL, 100 mg/mL, 50 mg/mL and 0 mg/mL concentrations were made in order to form a standard curve. Pipette the

various samples into ependorfs and add 1mL ddH20. Then add 150 uL 1% Deoxycholate in .1 M NaOH, mix and incubate at RT for 5 min. Add 200 uL 50% TCA, mix and incubate on ice for 15 min, spin down the precipitate in micro fuge 14,000 RPM for 10 min. Pipette off supernatant and add 1.15 mL of dilutent mixture containing: 19.6 mL of 2% Na_2CO_3 in water, 200 μL of 1% $CiSo_4 \cdot 5H_2O$ in water, and 200 μL 1%KNa Tartate•45H_2O in water. Vortex until resuspended, incubate at RT for 10 min. Add 200 μL of dilute 1:1 Folin-Ciocalteau reagent with ddH_2O and incubate in the dark for ~1 hour. Measure the Absorbance at 750nm. Use the BSA to form a line of best fit then by using the slope the concentration of unknown enzyme can be obtained.

ArnD Substrate Formation: The following reaction sequence was incubated in order to produce the product of ArnD, enzymatically, with radiolabeled UDP-Glucuronic acid as the substrate. First: 50 mM HEPES, 0.5 mg/mL ArnA-CTD, 2 mM DTT, 3mM NAD, 0.8 μM *UDP-Glucuronic Acid, was incubated overnight. Following the first incubation 1mM L-Glutamate and 0.5 mg/mL ArnB was added to the reaction mixture and incubated overnight. Next, 0.5 mg/mL Full length ArnA was added along with 1 mM N-10-metyltetrahydrofolate produced as explained in [31]. 11 μL of this entire mixture of product and reactants along with the accompanying enzymes was then separated and pipetted into a reaction tube containing 2 μL 1% Triton-X-100, 2 μL 500 mM HEPES pH7.5, 2 mM Mn^{2+}, and 0.5 mg/mL purified ArnC, in previously evaporated Undecaprenyl Phosphate (Polish Academy of Sciences) with ddH_2O to make a final volume of 20 μL. Formation of a product was confirmed by TLC, where a signal was shown that subsequently migrated after the addition of 0.05 mg/mL purified ArnD.

WT ArnD Activity Analysis: In order to measure the activity of ArnD various concentrations were used of ArnD were added to the reactant mixture of ArnC product with the simultaneous addition of 2 μL of 5 mM $CoCl_2$. The three concentrations of ArnD shown are 0.05, 0.005 and 0.001 mg/mL, respectively. These reactions were ran for 48 min taking and plating 2 μL aliquots at the time pints 8 min, 16 min, 24 min, 48 min.

Mutant ArnD Activity Analysis: In order to measure the activity of mutant ArnD constructs, identical assays were performed as above in WT Activity Analysis. 0.05 mg/mL concentrations of ArnD mutants were added to the reactant mixture of ArnC product with the simultaneous addition of 2 μL of 5 mM $CoCl_2$. . These reactions were ran for 48 min taking and plating 2 μL aliquots at the time pints 8 min, 16 min, 24 min, 48 min.

TLC Analysis: Two separate assays were performed:
- A PEI Cellulose Plate (EMD Chemicals) was first soaked in methanol and dried. Next, the plate was spotted by: UDP-glucuronic acid, UDP-4-keto-glucuronic acid, UDP-4-ketopentose UDP-4-aminoarabinose (UDP-β-L-Ara4N), and Ara4N-formyl (UDP-β-L-Ara4FN). The solvent used was: 0.4 M LiCl (Sigma), 0.25 M Acetic Acid (Mallinckrodt Chemicals).
- A TLC Silica Gel 60 F_{254} plate (EMD Chemicals) was spotted by: UDP-β-L-Ara4FN, undecaprenyl-phosphate-L-Ara4FN, undecaprenyl-phosphate-L-Ara4N. The solvent used was (64/24/3.6/.4) (CHCl3/MeOH/H2O/NH4SO4 (Mallinckrodt Chemicals, Fischer Scientific).

References

1. Yan, A., Z. Guan, and C.R. Raetz, *An undecaprenyl phosphate-aminoarabinose flippase required for polymyxin resistance in Escherichia coli.* J Biol Chem, 2007. 282(49): p. 36077-89.

2. Jonson, P.H. and S.B. Petersen, *A critical view on conservative mutations.* Protein Eng, 2001. 14(6): p. 397-402.

3. Novagen, *pET-28a(+) vector Map.*

4. Energy, U.S.D.o., *Cystic Fibrosis Gene.* 2003.

5. *Cell Wall.*

6. Zasloff, M., *Antimicrobial peptides of multicellular organisms.* Nature, 2002. 415(6870): p. 389-95.

7. Tomasz, A., *Multiple-Antibiotic-Resistant Pathogenic Bacteria - a Report on the Rockefeller-University Workshop.* New England Journal of Medicine, 1994. 330(17): p. 1247-1251.

8. Corporation, N.P. *What is Pseudomonas aeruginosa?* 2012 [cited 2012; Available from: http://www.tobitime.com/info/cf-pseudomonas-aeruginosa/What-is-Pseudomonas-aeruginosa.jsp?usertrack.filter_applied=true&NovaId=2935376911791733696.

9. Foundation, C.F., *About CF: Causes, Signs & Symptoms of Cystic Fibrosis.* 2012, Cystic Fibrosis Foundation.

10. Encyclopedia, A.D.A.M.M., *Cystic Fibrosis - PubMed Health.* 2012.

11. Govan, J.R. and V. Deretic, *Microbial pathogenesis in cystic fibrosis: mucoid Pseudomonas aeruginosa and Burkholderia cepacia.* Microbiol Rev, 1996. 60(3): p. 539-74.

12. FitzSimmons, S.C., *The changing epidemiology of cystic fibrosis.* J Pediatr, 1993. 122(1): p. 1-9.

13. Cole, A.M. and T. Ganz, *Antimicrobial peptides and proteins in the CF airway.* Methods Mol Med, 2002. 70: p. 447-64.

14. Sleigh, M.A., J.R. Blake, and N. Liron, *The Propulsion of Mucus by Cilia.* American Review of Respiratory Disease, 1988. 137(3): p. 726-741.

15. Lim, J., et al., *Studying the Effect of Alginate Overproduction on Pseudomonas aeruginosa Biofilm by Atomic Force Microscopy.* Journal of Nanoscience and Nanotechnology, 2011. 11(7): p. 5676-5681.

16. Okuyama-Nishida, Y., et al., *Prevention of Death in Bacterium-Infected Mice by a Synthetic Antimicrobial Peptide, L5, through Activation of Host Immunity.* Antimicrobial Agents and Chemotherapy, 2009. 53(6): p. 2510-2516.

17. Wilcox, S., *The New Antimicrobials: Cationic Peptides.* BioTeach Journal, 2004. 2: p. 88-91.

18. Miller, S.I., R.K. Ernst, and M.W. Bader, *LPS, TLR4 and infectious disease diversity.* Nature Reviews Microbiology, 2005. 3(1): p. 36-46.

19. Matsuzaki, K., *Why and how are peptide-lipid interactions utilized for self-defense? Magainins and tachyplesins as archetypes.* Biochimica Et Biophysica Acta-Biomembranes, 1999. 1462(1-2): p. 1-10.

20.   Matsuzaki, K., *Why and how are peptide-lipid interactions utilized for self-defense? Magainins and tachyplesins as archetypes.* Biochim Biophys Acta, 1999. 1462(1-2): p. 1-10.

21.   Zasloff, M., *Antibiotic peptides as mediators of innate immunity.* Curr Opin Immunol, 1992. 4(1): p. 3-7.

22.   Zavascki, A.P., et al., *Polymyxin B for the treatment of multidrug-resistant pathogens: a critical review.* Journal of Antimicrobial Chemotherapy, 2007. 60(6): p. 1206-1215.

23.   Ernst, R.K., et al., *Specific lipopolysaccharide found in cystic fibrosis airway Pseudomonas aeruginosa.* Science, 1999. 286(5444): p. 1561-1565.

24.   Ernst, R.K., et al., *Unique lipid A modifications in Pseudomonas aeruginosa isolated from the airways of patients with cystic fibrosis.* Journal of Infectious Diseases, 2007. 196(7): p. 1088-1092.

25.   Guo, L., et al., *Lipid A acylation and bacterial resistance against vertebrate antimicrobial peptides.* Cell, 1998. 95(2): p. 189-98.

26.   Gunn, J.S., et al., *Genetic and functional analysis of a PmrA-PmrB-regulated locus necessary for lipopolysaccharide modification, antimicrobial peptide resistance, and oral virulence of Salmonella enterica serovar typhimurium.* Infect Immun, 2000. 68(11): p. 6139-46.

27.   Zhou, Z., et al., *Lipid A modifications characteristic of Salmonella typhimurium are induced by NH4VO3 in Escherichia coli K12. Detection of 4-amino-4-deoxy-L-arabinose, phosphoethanolamine and palmitate.* J Biol Chem, 1999. 274(26): p. 18503-14.

28.   Zhou, Z., et al., *Lipid A modifications in polymyxin-resistant Salmonella typhimurium: PMRA-dependent 4-amino-4-deoxy-L-arabinose, and phosphoethanolamine incorporation.* J Biol Chem, 2001. 276(46): p. 43111-21.

29.   Gunn, J.S., et al., *Constitutive mutations of the Salmonella enterica serovar Typhimurium transcriptional virulence regulator phoP.* Infect Immun, 2000. 68(6): p. 3758-62.

30.   Moriarity, J.L., et al., *UDP-glucuronate decarboxylase, a key enzyme in proteoglycan synthesis: cloning, characterization, and localization.* J Biol Chem, 2002. 277(19): p. 16968-75.

31.   Breazeale, S.D., A.A. Ribeiro, and C.R.H. Raetz, *Oxidative decarboxylation of UDP-glucuronic acid in extracts of polymyxin-resistant Escherichia coli. Origin of lipid a species modified with 4-amino-4-deoxy-L-arabinose.* J Biol Chem, 2002. 277(4): p. 2886-96.

32.   Gatzeva-Topalova, P.Z., P. Andrew, and M.C. Sousa, *Crystal structure and mechanism of the Escherichia coli ArnA (PmrI) transformylase domain. An enzyme for lipid A modification with 4-amino-4-deoxy-L-arabinose and polymyxin resistance.* Biochemistry, 2005. 44(14): p. 5328-5338.

33.   Breazeale, S.D., A.A. Ribeiro, and C.R.H. Raetz, *Origin of lipid A species modified with 4-amino-4-deoxy-L-arabinose in polymyxin-resistant mutants of Escherichia coli. An aminotransferase (ArnB) that generates UDP-4-deoxy-L-arabinose.* J Biol Chem, 2003. 278(27): p. 24731-9.

34.   Raetz, C.R.H. and C. Whitfield, *Lipopolysaccharide endotoxins.* Annual Review of Biochemistry, 2002. 71: p. 635-700.

35.   Trent, M.S., et al., *Accumulation of a polyisoprene-linked amino sugar in polymyxin-resistant Salmonella typhimurium and Escherichia coli: structural characterization and transfer to lipid A in the periplasm.* J Biol Chem, 2001. 276(46): p. 43132-44.

36.   Stover, C.K., et al., *Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen.* Nature, 2000. 406(6799): p. 959-964.

37.   Blair, D.E., et al., *Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor.* Proc Natl Acad Sci U S A, 2005. 102(43): p. 15429-34.

38.   Whittington, D.A., et al., *Crystal structure of LpxC, a zinc-dependent deacetylase essential for endotoxin biosynthesis.* Proc Natl Acad Sci U S A, 2003. 100(14): p. 8146-50.

39.   Hernick, M. and C.A. Fierke, *Zinc hydrolases: the mechanisms of zinc-dependent deacetylases.* Arch Biochem Biophys, 2005. 433(1): p. 71-84.

40.   Martinou, A., D. Koutsioulis, and V. Bouriotis, *Expression, purification, and characterization of a cobalt-activated chitin deacetylase (Cda2p) from Saccharomyces cerevisiae.* Protein Expr Purif, 2002. 24(1): p. 111-6.

41.   Breazeale, S.D., et al., *A formyltransferase required for polymyxin resistance in Escherichia coli and the modification of lipid A with 4-Amino-4-deoxy-L-arabinose. Identification and function oF UDP-4-deoxy-4-formamido-L-arabinose.* J Biol Chem, 2005. 280(14): p. 14154-67.