Insights into the Mechanistic and Regulatory Properties of D-arabinose-5-phosphate Isomerases

By

David L. Cech

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Medicinal Chemistry)
in the University of Michigan
2017

Doctoral Committee:

Professor Ronald W. Woodard, Chair
Professor Harry L. T. Mobley
Professor Henry I. Mosberg
Professor John J. G. Tesmer

David L. Cech

cechd@umich.edu

ORCID iD: 0000-0002-8078-9720

TABLE OF CONTENTS

LIST OF TABLES v

LIST OF FIGURES vi

CHAPTER

I. Introduction 1

Background 1

D-Arabinose-5-Phosphate Isomerase 3

Dissertation Research Rationale 9

References 11

II. A Novel Glucose-6-phosphate Isomerase from *Listeria monocytogenes* 14

Summary 14

Introduction 15

Materials and Methods 16

Results 24

Discussion 33

Acknowledgments and Chapter Contributions 34

References 35

III. The Arabinose-5-phosphate Isomerase of *Bacteroides fragilis*: Insight Into Regulation of Single-domain Arabinose Phosphate Isomerases 38

Summary 38

Introduction 40

Materials and Methods 42
Results 52
Discussion 59
Acknowledgments and Chapter Contributions 62
References 63

IV. Identification of a D-Arabinose-5-Phosphate Isomerase in the Gram-positive *Clostridium tetani* 66
Summary 66
Introduction 67
Materials and Methods 68
Results 73
Discussion 81
Acknowledgments and Chapter Contributions 83
References 84

V. New Insights Into the Kdo Biosynthetic Pathway and Incorporation Into Lipid A-Like Molecules in *Arabidopsis thaliana* 86
Summary 86
Introduction 88
Materials and Methods 92
Results 103
Discussion 110
Acknowledgments and Chapter Contributions 112
References 114

VI. Structure of *E. coli* c3406 in Complex With A5P And Insight Into the Mechanism of D-Arabinose-5-Phosphate Isomerase 118
Summary 118
Introduction 119
Materials and Methods 120
Results and Discussion 124
Acknowledgments and Chapter Contributions 133
References 134

VII. Summary, Conclusions, and Future Directions 137

APPENDIX 141

Cystathionine β-Synthase Domain 141

References 143

LIST OF TABLES

Table

1.1: Kinetic constants for catalysis by the APIs of *E. coli*……..6

2.1: Strains, Plasmids, and Primers used in this study……..17

2.2: Kinetic constants for catalysis by Q723E8……..29

3.1: Strains, Plasmids, and Primers used in this study……..43

3.2: Kinetic constants for catalysis by various APIs……..55

4.1: Strains, Plasmids, and Primers used in this study……..69

4.2: Kinetic constants for the APIs of *E. coli* and CtAPI……..78

5.1: Strains, Plasmids, and Primers used in this study……..93

5.2: Kinetic constants for catalysis by various APIs……..105

6.1: Data collection and refinement statistics……..122

6.2: Strains, Plasmids, and Primers used in this study……..123

6.3: KdsD mutant characterization……..129

LIST OF FIGURES

Figure

1.1: Representative schematic of the Gram-negative membrane……..2

1.2: Constituents of LPS…..2

1.3: The Kdo biosynthetic pathway……..4

1.4: The SIS and CBS domains of APIs.……..5

2.1: Sequence alignment of Q723E8 and genomic context of *lmof2365_0531*..……..25

2.2: Clustal Omega alignment of the protein sequences of Q723E8 and the C-terminal end of various PGI……27

2.3: pH-rate profile for the conversion of G6P to F6P by Q723E8.……28

2.4: Effect of metal ions upon the glucose 6-phosphate isomerase activity of Q723E8.……29

2.5: The purification of Q723E8……30

2.6: Native molecular weight determination for Q723E8……31

2.7: Complementation of an A5P auxotroph on agar plates.……32

3.1: Gel filtration analysis of Q5LIW1……53

3.2: Effect of divalent metal ions on the activity of Q5LIW1……54

3.3: pH rate profile of Q5LIW1……55

3.4: Complementation of an A5P auxotroph on agar plates……56

3.5: Complementation of an A5P auxotroph on LB/agar plates…..57

3.6: Effect of CMP-Kdo on Q5LIW1 API activity……58

4.1: Alignment of sequences of *E. coli* APIs and CtAPI.....74

4.2: Genomic context of *gutQ* and *ctAPI.*……75

4.3: Standard curve from native molecular mass determination……76

4.4: pH rate profile of CtAPI……77

4.5: Complementation of an A5P auxotroph on an LB agar plate.....80

5.1: Kdo biosynthesis and incorporation in *E. coli*……88

5.2: A model of RG-II structure with side chains labeled.....89

5.3: Complementation of an A5P auxotroph on LB agar plates……105

5.4: Negative ion mode electrospray ionization FT-MS analysis of lipids.....107

5.5: LPS structures of observed masses of related peaks shown in Fig. 5.4……108

5.6: *Limulus amebocyte* lysate assay with total lipid extracts from 10-day old A. thaliana seedlings of wild type (col-0) and the atkdta-1 (SALK_035981) mutant.....109

6.1: Omit map of A5P in complex with c3406 API.……125

6.2: c3406 API active site.....126

6.3: Alignment of c3406 API APO structure and c3406 API in complex with A5P.……126

6.4: Phosphate binding pocket alignment.....127

6.5: Isomerization of select phosphosugars.....131

6.6: Proposed mechanism of c3406 API.....131

6.7: Active site comparison of rabbit PGI and c3406 API.....132

A.1: KpKpsF thermal melts……142

CHAPTER I

Introduction

Background

Historically, pathogenic bacteria have been a major human health problem. The rise of antimicrobial resistance has become an increasingly major concern. Multiple Gram-negative bacteria have become resistant to several classes of antibiotics and the rise of resistant bacteria has compromised the ability of healthcare providers to effectively treat infections (1,2). The CDC estimates that more than 2 million people in the United States are affected by antibiotic-resistant infections each year resulting in over 23,000 deaths (3). Many of the bacterial threats classified as urgent and serious by the CDC are Gram-negative. These threats call for the development of antibiotics with novel mechanisms of action. While many traditional antibiotics target the synthesis of cell wall peptidoglycan, the biosynthetic pathway of 2-keto-3-deoxy-D-*manno*-octulosonate (Kdo) has not been fully explored as a target for antibiotic development.

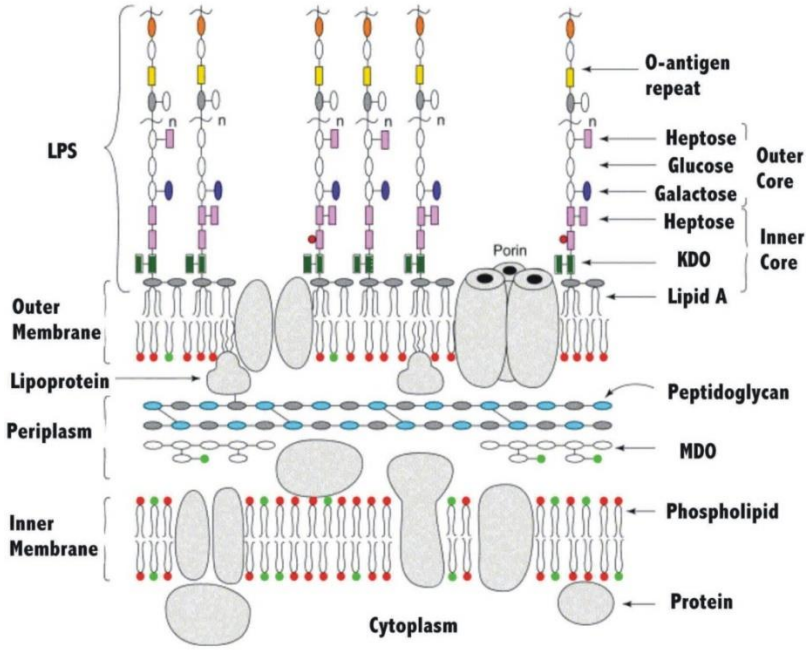

Fig. 1.1: Representative schematic of the Gram-negative membrane. Adapted from (1)

Kdo is an essential component of lipopolysaccharide (LPS), which is located on the cell surface of the outer membrane of Gram-negative bacteria (Fig. 1.1). The outer membrane provides for a higher level of intrinsic resistance to intracellular penetration of antimicrobials,

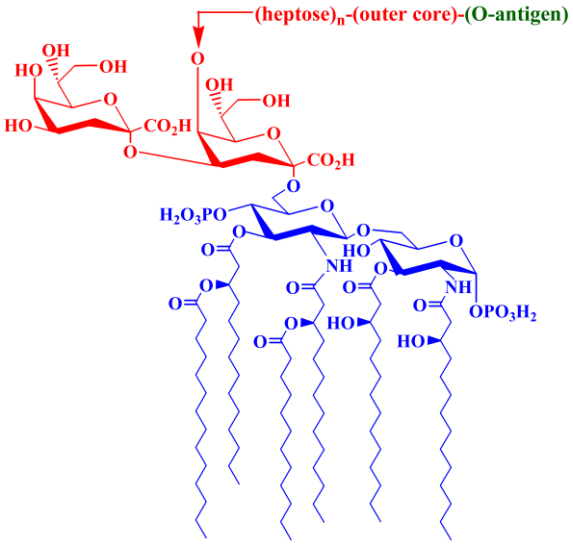

Fig. 1.2: Constituents of LPS. Blue: lipid A. Red: core oligosaccharide. Green: *O*-antigen. Adapted from (2).

which reduces the number of effective options for treatment when compared to Gram-positive bacteria (4). The outer membrane is an asymmetric lipid bilayer comprised of a variety of phospholipids in the inner leaflet and LPS on the outer leaflet. LPS, which accounts for 30% of the outer membrane gross weight (5), is composed of three constituents: lipid A, the core oligosaccharide, and the *O*-antigen (Fig. 1.2). The highly-conserved lipid A domain anchors LPS into the outer leaflet of the outer membrane. The core oligosaccharide, which usually contains at least one Kdo residue, is a conserved region which connects lipid A and the *O*-antigen. The core oligosaccharide in *E. coli* minimally requires the addition of Kdo to lipid IV_A (6). The O-antigen is the outermost region of LPS and is highly variable.

The presence of LPS (aka endotoxin), on the surface of the bacterium, allows for the host's immune system to recognize an infection through pathogen-associated molecular patterns (7). When an abnormal, imbalanced, response to an infection by the immune system occurs, several conditions can arise including fever, tachycardia, hypotension, multiple organ failure, and in some cases, death (8).

D-Arabinose-5-Phosphate Isomerase

The first step of the Kdo biosynthetic pathway (Fig. 1.3) is the reversible keto/aldol isomerization of the pentose phosphate pathway intermediate D-ribulose-5-phosphate (Ru5P) to D-arabinose-5-phosphate (A5P) catalyzed by arabinose-5-phosphate isomerase (API). A5P is condensed with phosphoenol pyruvate by Kdo8P synthase (KdsA) to form Kdo8P (9). Dephosphorylation of Kdo8P to Kdo occurs via Kdo8P phosphatase (KdsC) and is subsequently

activated with CTP to yield CMP-Kdo by cytidine-5'-monophosphatase KDO synthetase (KdsB) (10).

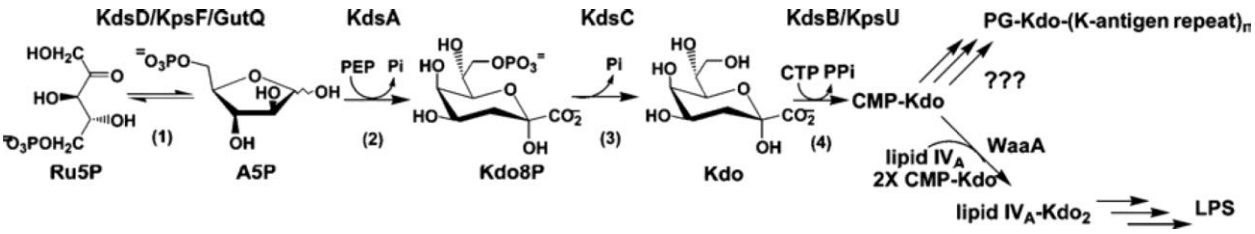

Fig. 1.3: The Kdo biosynthetic pathway. [1] D-arabinose-5-phosphate isomerase, [2] Kdo-8-phosphate synthase, [3] Kdo8P phosphatase, [4] cytidine-5'-monophosphate-Kdo synthetase. Adapted from (1,2).

APIs were originally identified by Volk while studying the metabolism of L-arabinose by *Propionibacterium pentosaceum* in the late 1950s (11). After initially naming the enzyme phosphoarabinoisomerase, he went on to partially purify the API from *P. pentosaceum* and show that is was specific for the interconversion of A5P and Ru5P (12).

While several organisms encode a single API, the genome of *E. coli* CFT073 contains four genes which encode APIs. Each of these APIs corresponds with a unique function. The first three of APIs KdsD, GutQ, and KpsF, are the traditional API, which contain both a sugar isomerase (SIS) domain and a tandem cystathionine β-synthase (CBS) domain (Fig. 1.4). KdsD is responsible for A5P necessary for LPS biosynthesis, GutQ is responsible for operon regulation, and KpsF is involved in capsule biogenesis (2).

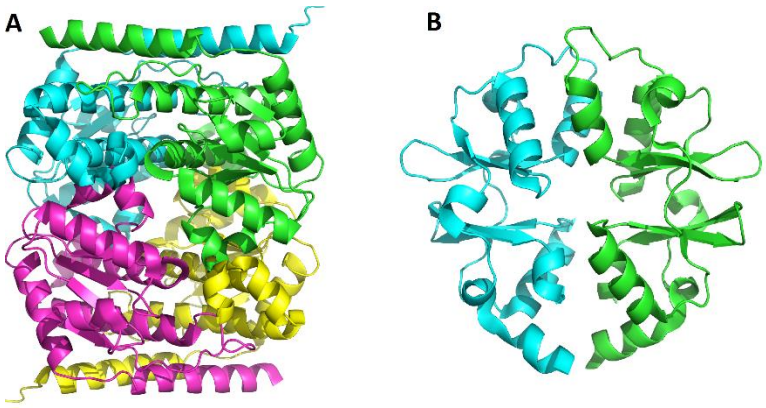

Fig. 1.4: The SIS and CBS domains of APIs. A) c3406 SIS domain. B) CBS domain of Klebsiella pneumoniae PDB ID: 3K2V (12).

KdsD

E. coli KdsD, the most studied API, was assayed and found to be a phosphosugar aldo-keto isomerase specific for the interconversion of Ru5P and A5P (13) and can serve as the sole A5P source of the cell (14). Upon further characterization of KdsD it was found to be a tetramer under native conditions (13), which has proven to be the norm for APIs. Full kinetic parameters for KdsD can be found in Table 1.1.

Table 1.1: Kinetic constants for catalysis by the APIs of *E. coli*.

Protein	k_{cat} (A5P to Ru5P, s^{-1})	K_m (A5P, mM)	k_{cat} / K_m (A5P, M^{-1} s^{-1})	k_{cat} (Ru5P to A5P, s^{-1})	K_m (Ru5P, mM)	k_{cat}/ K_m (Ru5P, M^{-1} s^{-1})	K_{eq}	Optimum pH	Subunit mass (Da)
KdsD[13]	157±4	0.61±0.06	$2.6x10^5$	255±16	0.35±0.08	$7.3x10^5$	0.50±0.06	8.4	35,084
GutQ[14]	218±4	1.2±0.1	$1.8x10^5$	242±11	0.64±0.08	$3.8x10^5$	0.47	8.25	33,909
KpsF[15]	15±1	0.57±0.04	$2.6x10^4$	19±2	0.30±0.03	$6.3x10^4$	0.48±0.02	7.75	35,447
c3406[16]	1.92±0.05	$8.8x10^3$	16.8±0.2	10.5±0.08	0.70±0.12	$1.5x10^4$	0.52	6.6	20,880

Various attempts have been made to obtain the crystal structure of *E. coli* KdsD, which consist of an N-terminal SIS domain (residues 1-200) followed by two CBS domains (residues 213-325), however, to date, all attempts have failed to obtain a full-length structure of the enzyme. However, the SIS domain of a catalytically inactive mutant K59A of *E. coli* KdsD (PDB ID: 2XHZ) (17) and the CBS domains of KdsD (PDB ID: 3FNA) (18) have been separately crystallized.

KpsF

The *kps* locus, which is necessary for the biosynthesis of polysialic acid capsule virulence factor in *E. coli* K1, contains at least two convergently transcribed operons. *kpsF* is the first gene of region 1 and is transcriptionally silent at lower temperatures (20°C). However, maximal transcription of *kpsF* can be seen at host body temperature (37°) (19). It has been shown that Group II K antigens exhibit temperature dependent regulation, with capsule expressions at 37°C but not 18°C and are under the transcriptional control of the same promotor as *kpsF* (19,20). The *kpsF* gene also encodes an intragenic Rho-dependent transcriptional terminator (RfaH) that

permits transcription regulation of region 1 in response to physiological stress (15,20). Maximal transcription of region 1 is regulated by the tyrosine-phosphorylated GTPase BipA, the global regulator heat-stable nucleoid-structuring protein (H-NS), integration host factor (IHF1), and promoters at 37°C (21). H-NS has been implicated in the thermoregulation of many genes, repressing transcription at low temperatures and activating gene expression at 37°C. The *kpsF* gene encodes an IHF1 site near the 5' end of the gene which suggests that IHF1 may play an indirect role in the activation of region 1 transcription (21). All of these regulators located within and around *kpsF* result in a coordinated repression of capsule synthesis at temperatures below 20°C and induction of capsule synthesis at 37°C (15).

KpsF was characterized from the group 2 capsule strain *E. coli* CFT073 in order to begin to better understand the role of Kdo in capsule biogenesis as *E. coli* K1 and K92 strains do not have Kdo substituted in the reducing end of their polysaccharides. KpsF was found to be tetrameric in its quaternary structure, and inhibited by Zn^{2+}, Cd^{2+}, and Hg^{2+}. The addition of EDTA increased activity by approximately 2 fold showing that there is not, however, a metal requirement for KpsF (15). It has been speculated that metal ion inhibition may be a post-transcriptional control mechanism *in vivo* (22), but it is unclear if this would be biologically relevant in the case of KpsF as the concentration of cytoplasmic unsequestered zinc is very low (15). Full kinetic parameters of KpsF can found in Table 1.1.

GutQ

The first G-API to be characterized, GutQ, was identified downstream from *gutR* of the glucitol operon in *E. coli* (23). The other seven genes in the glucitol operon encode for the EII^{Gut}

complex (formed by GutA, GutE, and GutB), a NADH-dependent dehydrogenase (GutD), a transcriptional repressor (GutR), and a transcriptional activator (GutM).

After the discovery of the function of KpsF and KdsD, it was determined that GutQ was indeed an arabinose 5-phosphate isomerase (Table 1.1). The role of GutQ in the *gut* operon remains unclear as a *gutQ* deletion in *E. coli* does not interfere with *gut* operon induction. Upregulation of *gutQ* occurs with the addition of D-glucitol. Further, the addition of varying concentrations of A5P appears to affect *gut* operon induction. This suggests that GutQ may function to synthesize A5P which then participates in the induction of the operon to modulate expression levels through an unknown mechanism (14).

SIS-domain APIs

The last type of API is the SIS-domain API. These API only contain an SIS domain and lack the CBS domains contained in traditional API. One of the most studied APIs in this category is c3406 API from *E. coli* CFT073, a very well-studied uropathogenic strain which upregulates a significant number of genes during UTI (16,24-26). Deletion of a cluster of genes, *c3405* to *c3410*, results in a significant attenuation of CFT073's ability to independently colonize mouse kidneys (27). Analysis of this cluster showed that one of the genes, *c3406*, encodes for a protein that has significant homology to the SIS domain of APIs. Other genes in the cluster encode for a putative 2-hydroxyacid dehydrogenase (*c3405*), a beta-cystathionase (*c3407*), a maltose and glucose-specific IIABC component (*c3408*), and an antiterminator (*c3409*) (16). *E. coli* CFT073 expresses group 2 K antigens and therefore already encodes for KdsD, GutQ, and KpsF. While these three enzymes contain two CBS domains, c3406 only consists of an SIS domain. The c3406 protein

shares 52%, 45%, and 40% identity with the SIS domains of KdsD, KpsF, and GutQ, respectively, from *E. coli* CFT073.

Analysis of the quaternary structure of the c3406 protein showed it likely exists as a tetramer. The enzymatic activity of c3406 was shown to be specific for the interconversion of Ru5P and A5P. Furthermore, it was shown that Zn^{2+}, Cd^{2+}, and Hg^{2+} all inhibit c3406's enzymatic activity, similar to KdsD. Perhaps the most striking differences between c3406 and the CBS-containing APIs is lower efficiency of c3406 as an API and the extreme pH-rate optimum, which is more than an entire pH unit lower than any characterized API (Table 1.1). Studies of an *E. coli* CFT073 Δ*c3406* strain showed that there was no significant difference in its ability to colonize rat bladder and kidneys compared to wild type (16). It is unclear what the biological function of the *c3405*-to-*c3410* operon is, but it appears as though c3406 API is not solely responsible for the phenotype that is observed when the entire operon is deleted (16).

Dissertation Research Rationale

APIs are the sole source of intracellular A5P utilized in the Kdo biosynthetic pathway. It is known that deletion of APIs leads to depletion of A5P pools and thus disrupts Kdo biosynthesis. It has long been thought that Kdo production was specific to Gram-negative bacteria. However, Kdo has been found in higher order plants. This means that A5P may be utilized for more than LPS biosynthesis.

To explore the hypothesis that A5P can be utilized for a variety of functions, an approach involving molecular biology, enzymology, and microbiology was applied. The first step was to

identify and characterize unique APIs from a variety of sources. In chapters II, III, IV, and V genes encoding potential APIs from Gram-positive bacteria, Gram-negative bacteria, and a plant were cloned, overexpressed, purified, and characterized. Utilizing the information gained form these studies, including the finding of a potential feedback inhibition pathway, a better understanding of the mechanism of APIs was sought in chapter VI, which could aid in the design of inhibitors of APIs.

References

1. Raetz, C. R. H., and Whitfield, C. (2002) Lipopolysaccharide Endotoxins. *Annual Review of Biochemistry* **71**, 635-700

2. Meredith, T. C. (2006) *D-Arabinose-5-Phosphate Isomerase from Escherichia coli.* PhD, University of Michigan

3. (2013) Antibiotic Resistance Threats in the United States. Centers for Disease Control and Prevention

4. Nikaido, H. (1989) Outer membrane barrier as a mechanism of antimicrobial resistance. *Antimicrob Agents Chemother* **33**, 1831-1836

5. Nikaido, H. (1996) Outer Membrane. in *Escherichia coli and Salmonella typhimurium : cellular and molecular biology* (Neidhardt, F. C. ed.), American Society for Microbiology, Washington, D.C. pp 29-47

6. Meredith, T. C., Aggarwal, P., Mamal, U., Lindner, B., and Woodard, R. W. (2006) Redefining the requisite lipopolysaccharide structure in *Escherichia coli*. *ACS chemical biology* **1**, 33-42

7. Silhavy, T. J., Kahne, D., and Walker, S. (2010) The bacterial cell envelope. *Cold Spring Harb Perspect Biol* **2**, a000414

8. Cohen, J. (2002) The immunopathogenesis of sepsis. *Nature* **420**, 885-891

9. Duewel, H. S., Sheflyan, G. Y., and Woodard, R. W. (1999) Functional and biochemical characterization of a recombinant 3-Deoxy-D-manno-octulosonic acid 8-phosphate synthase from the hyperthermophilic bacterium *Aquifex aeolicus*. *Biochemical and biophysical research communications* **263**, 346-351

10. Yi, L., Velasquez, M. S., Holler, T. P., and Woodard, R. W. (2011) A simple assay for 3-deoxy-d-manno-octulosonate cytidylyltransferase and its use as a pathway screen. *Anal. Biochem.* **416**, 152-158

11. Volk, W. A. (1959) The Enzymatic Formation of D-Arabinose 5-Phosphate from l-Arabinose and Adenosine Triphosphate by *Propionibacterium pentosaceum*. *J. Biol. Chem.* **234**, 1931-1936

12. Volk, W. A. (1960) Purification and Properties of Phosphoarabinoisomerase from *Propionibacterium pentosaceum*. *J. Biol. Chem.* **235**, 1550-1553

13. Meredith, T. C., and Woodard, R. W. (2003) *Escherichia coli* YrbH Is a D-Arabinose 5-Phosphate Isomerase. *J. Biol. Chem.* **278**, 32771-32777

14. Meredith, T. C., and Woodard, R. W. (2005) Identification of GutQ from *Escherichia coli* as a D-Arabinose 5-Phosphate Isomerase. *J. Bacteriol.* **187**, 6936-6942

15. Meredith, Timothy C., and Woodard, Ronald W. (2006) Characterization of *Escherichia coli* D-arabinose 5-phosphate isomerase encoded by kpsF: implications for group 2 capsule biosynthesis. *Biochemical Journal* **395**, 427-432

16. Mosberg, J., Yep, A., Meredith, T., Smith, S., Wang, P.-F., Holler, T., Mobley, H., and Woodard, R. (2011) A unique arabinose 5-phosphate isomerase found within a genomic island associated with the uropathogenicity of *Escherichia coli* CFT073. *J. Bacteriol.* **193**, 2981-2988

17. Gourlay, L. J., Sommaruga, S., Nardini, M., Sperandeo, P., Dehò, G., Polissi, A., and Bolognesi, M. (2010) Probing the active site of the sugar isomerase domain from *E. coli* arabinose-5-phosphate isomerase via X-ray crystallography. *Protein Science* **19**, 2430-2439

18. Cuff, M., Bigelow, L., Buck, K., and Joachimiak, A. The CBS Pair of possible D-arabinose 5-phosphate isomerase *yrbH* from *Escherichia coli* CFT073. *PDB ID: 3FNA*

19. Cieslewicz, M., and Vimr, E. (1996) Thermoregulation of *kpsF*, the first region 1 gene in the *kps* locus for polysialic acid biosynthesis in *Escherichia coli* K1. *J Bacteriol.* **178**, 3212-3220

20. Simpson, D. A., Hammarton, T. C., and Roberts, I. S. (1996) Transcriptional organization and regulation of expression of region 1 of the *Escherichia coli* K5 capsule gene cluster. *J. Bacteriol.* **178**, 6466-6474

21. Rowe, S., Hodson, N., Griffiths, G., and Roberts, I. S. (2000) Regulation of the *Escherichia coli* K5 Capsule Gene Cluster: Evidence for the Roles of H-NS, BipA, and Integration Host Factor in Regulation of Group 2 Capsule Gene Clusters in Pathogenic *E. coli*. *J. Bacteriol.* **182**, 2741-2745

22. Maret, W., Jacob, C., Vallee, B. L., and Fischer, E. H. (1999) Inhibitory sites in enzymes: zinc removal and reactivation by thionein. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 1936-1940

23. Yamada, M., and Saier, M. H., Jr. (1987) Glucitol-specific enzymes of the phosphotransferase system in *Escherichia coli*. Nucleotide sequence of the gut operon. *J Biol Chem* **262**, 5455-5463

24. Subashchandrabose, S., and Mobley, H. L. (2015) Virulence and Fitness Determinants of Uropathogenic *Escherichia coli*. *Microbiol Spectr* **3**

25. Garcia, E. C., Brumbaugh, A. R., and Mobley, H. L. (2011) Redundancy and specificity of *Escherichia coli* iron acquisition systems during urinary tract infection. *Infect Immun* **79**, 1225-1235

26. Snyder, J. A., Haugen, B. J., Buckles, E. L., Lockatell, C. V., Johnson, D. E., Donnenberg, M. S., Welch, R. A., and Mobley, H. L. (2004) Transcriptome of uropathogenic *Escherichia coli* during urinary tract infection. *Infect Immun* **72**, 6373-6381

27. Lloyd, A. L., Henderson, T. A., Vigil, P. D., and Mobley, H. L. T. (2009) Genomic Islands of Uropathogenic *Escherichia coli* Contribute to Virulence. *J. Bacteriol.* **191**, 3469-3481

CHAPTER II

A Novel Glucose-6-phosphate Isomerase from *Listeria monocytogenes*

The work described in this chapter has been published [Cech, D.L., Wang, P.F., Holt, M.C., Assimon, V.A., Schaub, J.M., Holler, T.P., Woodard, R.W., *A novel glucose 6-phosphate isomerase from Listeria monocytogenes.* Protein J, 2014. **33**(5): p. 447-56.]

Summary

D-Arabinose 5-phosphate isomerases (APIs) catalyze the interconversion of D-ribulose-5-phosphate and D-arabinose-5-phosphate (A5P). A5P is an intermediate in the biosynthesis of 3-deoxy-D-manno-octulosonate (Kdo), an essential component of lipopolysaccharide, which is found in the outer membrane of Gram-negative bacteria. The genome of the Gram-positive pathogen *Listeria monocytogenes* contains a gene encoding a putative sugar isomerase domain API, Q723E8, with significant similarity to c3406, the only one of four APIs from *Escherichia coli* CFT073 that lacks a CBS domain. However, *L. monocytogenes* lacks genes encoding any of the other enzymes of the Kdo biosynthesis pathway. Realizing that the discovery of an API in a Gram-positive bacterium could provide insight into an alternate physiological role of A5P in the cell, we prepared and purified recombinant Q723E8. We found that Q723E8 does not possess API activity, but instead is a novel GPI (D-glucose 6-phosphate isomerase). However, the GPI activity of Q723E8 is weak compared with previously described GPIs. *L. monocytogenes* contains an ortholog of the well-studied two-domain bacterial GPI. Based on this evidence, glucose utilization is likely not the primary physiological role of Q723E8.

Introduction

The Gram-positive bacterium *Listeria monocytogenes*, originally isolated from infected rabbits [1], is ubiquitous in the environment and causes disease in a variety of species, including *Homo sapiens* [2]. It is perhaps best known for causing food-borne illnesses, a trait enhanced by its ability to survive and grow under a wide range of environmental conditions and the long incubation time between exposure and the onset of illness [3]. *L. monocytogenes* is one of 8 species in the genus *Listeria* and the only human pathogen among them. Another species in the genus, *L. ivanovii*, is an animal pathogen [4]. *L. monocytogenes* is a facultative anaerobe that can grow in environments with temperatures ranging from 0 to 45 °C and pH from 4.1 to 9.6 [5]. It can survive on various food products for a very long time. Once ingested, *L. monocytogenes* is an intracellular pathogen capable of growth within cells of the gastrointestinal tract or in macrophages. Given the diversity of its habitat, one might expect *Listeria* to utilize a vast array of carbon substrates. However, under laboratory conditions, glucosamine, N-acetylglucosamine and glucose are the primary substrates for growth of *L. monocytogenes* [6].

We were intrigued by the appearance of the crystal structure of a 200 amino-acid protein (PDB ID: 3FXA) from *L. monocytogenes* strain 4b F2365, which was annotated as an API. This structure had substantial homology to c3406, one of the four arabinose-5-phosphate isomerases present in the Gram-negative pathogen *Escherichia coli* CFT073 [7]. Because the catalytic properties of most of the small proteins that consist of a single sugar isomerase domain [8] are underexplored, we cloned, purified and characterized this protein, which will be referred to by its UniProt ID: Q723E8. Contrary to the initial prediction of an API, this protein is capable of catalyzing the interconversion of glucose-6-phosphate (G6P) and fructose 6-phosphate (F6P).

Given its small size and lack of a metal cofactor, Q723E8 may represent the first member of a novel class of glucose-6-phosphate isomerases (GPIs).

Materials and Methods

Materials

Genomic DNA from *Listeria monocytogenes* type 4b strain F2365 was purchased from the American Type Culture Collection (catalogue # BAA-679D-5). PCR was performed using primers synthesized by Integrated DNA Technologies (Coralville, Iowa, USA) or Invitrogen (Carlsbad, CA, USA), the Failsafe™ PCR PreMix Selection Kit from Epicentre Biotechnologies (Madison, WI, USA), and an MJ Research PTC-200 Peltier Thermal Cycler. TA TOPO cloning was performed with the TOPO TA cloning kit from Invitrogen. Enzymes required for subcloning and site-directed mutagenesis were purchased from New England Biolabs (Ipswich, WI, USA) and Stratagene (La Jolla, CA, USA), respectively. Plasmid DNA purification was performed with the Promega Wizard Miniprep kit (Madison, WI, USA). DNA sequencing was performed by the University of Michigan Biomedical Resources Core Facility. Metal salts used for assays were purchased as high-purity solids from Alfa-Aesar and used without further purification.

Table 2.1 Strains, Plasmids, and Primers used in this study

Item	Description	Source
E. coli		
TOP10	F$^-$ *mcrA* Δ(*mrr*-*hsdRMS*-*mcrBC*) F80*lacZ*ΔM15 Δl*acX*74 *recA1 araD*139 Δ(ara-leu)7697 *galU galK rpsL* (StrR) endA1 *nupG*	Invitrogen
TCM15	BW30270(*ΔkdsDΔgutQ*)	Ref. 3
Rosetta 2 (DE3) pLysS	F$^-$ *ompT* hsdS$_B$($r_B^-m_B^-$) *gal dcm* (DE3) pLysSRARE2 (Cam^R)	EMD Millipore
Plasmids		
pCR2.1-TOPO	Library plasmid	Invitrogen
pT7-7	Expression vector	Ref. 10
pT7-7c3406	*E.coli* CFT073 c3406 inserted into NdeI/BamHI of pT7-7, Amp^R	Ref. 8
pT7-7LMOf2365_0531	*L. monocytogensis* 4b F22365 LMOf2365_0531 in NdeI/ BamHI of pT7-7, Amp^R	This study
Primers		
LMOF2365_0531.F	GGCATATGGATAAACAAGCTATTTTAGATAATATTCACCAAACAGG[a]	Invitrogen
LMOF2365_0531.R	CCGGATCCTTATTTATTTAATAATTTGTTTCCAACGGCACCACCAGG[b]	Invitrogen
LMOF2365_0531.S.SDM	GGTGCCGTTGGAAACAAATTATTGAACAAGTGACAGCACACTGGCGGCCG	Invitrogen
LMOF2365_0531.A.SDM	CGGCCGCCAGTGTGCTGTCACTTGTTCAATAATTTGTTTCCAACGGCACC	Invitrogen

Bacterial strains, plasmids, primers and growth media

The bacterial strains, plasmids, and primers used in this study are described in Table 2.1. *E. coli* TCM15 is a derivative of BW30270, *E. coli* K-12 MG1655; *rph*$^+$*fnr*$^+$, in which the *kdsD* and *gutQ* genes were disrupted using a phage λ Red recombinase system [10]. Strains were grown in LB medium [11]. TCM15 cultures were additionally supplemented with D-arabinose 5-phsophate (A5P; 15 μM) and G6P (10 μM).

Cloning, expression and purification of *L. monocytogenes* Q723E8

The lmof2365_0531 gene, which encodes Q723E8, was amplified from the genomic DNA of *L. monocytogenes* 4b strain F2365 using the primers LMOF2365_0531.F and LMOF2365_0531.R (Table 2.1). Agarose gel-purified PCR products were inserted into vector

pCR2.1-TOPO using the TA TOPO cloning kit. The gene was then subcloned by digestion of an insert-containing plasmid with *Nde*I and *Bam*HI, gel-purification of the insert, and then ligation into the similarly restricted expression vector pT7-7 [12]. Site-directed mutagenesis was performed using primers LMOF2365_0531.S.SDM and LMOF2365_0531.A.SDM (Table 2.1), PCR 2X premix buffer E, and the DNA polymerase Herculase II. The PCR mixture was digested with endonuclease *Dpn*I at 37 °C for 1 h to remove the template plasmid, and then used to transform chemically competent *E. coli* TOP 10 cells to ampicillin resistance. DNA sequencing of the resulting plasmid, pT7-lmof2365_0531, confirmed that the sequence of the lmof2365_0531 gene conformed to the sequence published in the Kyoto Encyclopedia of Genes and Genomes (www.genome.jp/dbget-bin/www_bget?lmf:LMOf2365_0531). The lmof2365_0531 gene was then transferred to the protein expression plasmid pET17b using the restriction endonucleases *Nde*I and *Bam*HI. The resulting expression plasmid was transformed into chemically competent *E. coli* Rosetta 2 (DE3) pLysS cells. A fresh transformant was grown in LB medium supplemented with 100 mg/L ampicillin and 30 mg/L chloramphenicol at 37 °C with shaking at 250 rpm until the $OD_{600\ nm}$ reached 0.8. The culture was cooled to 19 °C, and protein expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (0.4 mM). After 12 h of incubation at 19 °C, cells were harvested by centrifugation (6000 x g, 10 min, 4 °C). The cell pellet was resuspended in 20 mM Tris HCl buffer, pH 7.5, containing 0.1 M NaCl, and then sonicated on ice (3 cycles of 20 s, 2 min pauses between pulses) and clarified by centrifugation at 20,000 x g for 30 min. Ammonium sulfate was added slowly to the supernatant until 60% saturation at 4 ^{o}C. Precipitated proteins were removed by centrifugation (20,000 x g for 30 min) and the supernatant was dialyzed against 20 mM Tris HCl buffer, pH 7.5, overnight at 4 °C. After dialysis the protein was

loaded onto a Q-Sepharose anion exchange column equilibrated in 40 mM Tris HCl, pH 7.5, and eluted with a linear gradient of 0-600 mM NaCl over 15 column volumes at a flow rate of 0.5 mL/min. The fractions containing Q723E8 were concentrated using an Amicon ultra centrifugal filter (10,000 MWCO); exchanged into 20 mM 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES) buffer, pH 7.5, 100 mM NaCl, and 25% glycerol, divided into aliquots and stored at -80 °C.

Molecular weight determination

The subunit mass of Q723E8 was determined utilizing electrospray ionization mass spectrometry on an LCT electrospray/time-of-flight spectrometer. The native MW (molecular weight) of Q723E8 was estimated by native gel electrophoresis using a Pharmacia PhastSystem. Samples of Q723E8 were run on gradient acrylamide gels using an 8-25% total acrylamide gradient (GE Healthcare PhastGel Gradient 8-25) under native conditions (GE Healthcare Phastgel Native Buffer Strips). Standards, run in triplicate, included BSA dimer (132.4 kDa), BSA monomer (66.2 kDa), and chicken egg white ovalbumin monomer (44.3 kDa). A relative mobility (R_F) value was determined for each protein standard, averaging the three replicates. A plot of R_F versus log(MW) was fit by linear regression using GraphPad Prism 5. The MW of Q723E8 was determined by experimentally measuring its R_F value, in duplicate, and calculating its log(MW) from the standard curve. Two experimental values of the MW of Q723E8 were calculated and these values were averaged to yield the MW.

Complementation experiments in the TCM-15 *E. coli* derivative

Electrocompetent cells were prepared using a previously described TCM15 variant [7] that carries a kanamycin resistance cassette. These electrocompetent TCM15 cells were transformed separately with pT7-lmof2365_0531, pT7-c3406, and the empty pT7-7 vector. Cells were grown on LB/agar plates supplemented with 100 mg/L ampicillin, 50 mg/L kanamycin, 15 µM A5P and 10 µM G6P. Single colonies were used to inoculate liquid LB media supplemented with the same concentrations of antibiotics, A5P, and G6P. After overnight growth at 37 °C, cells were washed twice with liquid LB medium to remove the A5P and G6P in the overnight culture. The washed cells were streaked on LB/agar plates with and without A5P/G6P, and incubated overnight at 37 °C.

D-Glucose 6-phosphate isomerase assays

The discontinuous cysteine-carbazole colorimetric assay, adapted to a 96-well microplate format [9], was used to measure the isomerization of G6P to F6P. Briefly, 25 µL aliquots of Q723E8 in 100 mM 3-(N-morpholino)propanesulfonic acid (MOPS) buffer, pH 7.0, containing 1 mM EDTA, were added to the wells of a 96-well assay microplate. The assay plate, which also contained aliquots of G6P in the same buffer in separate wells, was then incubated at 37 °C for 3 min using a PTC-200 Peltier Thermal Cycler (MJ Research). Reactions were initiated by adding of 25 µL G6P solution to the aliquots of Q723E8, and then incubating the plate at 37 °C for the specified time, typically 3 min. Reaction mixtures were quenched with 50 µL of 25 N H_2SO_4. Next, a 90 µL portion of each quenched reaction mixture was transferred to the appropriate well of a fresh 96-well flat bottom assay plate containing 10 µL of 0.12% ethanolic carbazole, 10 µL of 1.5%

aqueous cysteine-HCl, and 230 µL of 25 N H_2SO_4 per well. This microplate was incubated at room temperature (~21 °C) for 3 h before recording the absorbance of each well at 540 nm. All plates contained controls consisting of identical reactions that lacked either Q723E8 or G6P, and internal F6P standards in triplicate.

Substrate specificity

Recombinant Q723E8 was tested for its ability to convert aldoses to ketoses using the discontinuous cysteine-carbazole assay. For each potential substrate, a solution containing 250 nM enzyme, 100 mM Tris-HCl pH 7.5, and 0.8 mM substrate was incubated for 10 min at 37 °C, quenched with 25 N H_2SO_4, transferred to a separate microplate containing the cysteine-carbazole assay mixture, and incubated at room temperature (~21 °C) for 3 h. The absorbance at 540 nm was compared to that of an otherwise identical control reaction lacking enzyme. Determinations were performed in triplicate, and the test aldoses that produced an absorbance significantly greater than the control were considered potential substrates. The aldoses tested include A5P, G6P, D-ribose 5-phosphate, D-arabinose, D-mannose 6-phosphate, D-glucosamine 6-phosphate, and D-ribose.

Determining the pH-rate profile of Q723E8

To determine the optimal pH of the enzyme, we tested Q723E8's catalytic activity in a series of buffered solutions of varying pH using the cysteine-carbazole assay. Buffer solutions (200 mM buffer, 2 mM EDTA) were prepared at pHs from 5.25 to 9.5 in 0.25 unit increments at 37 °C. The buffer 2-(*N*-morpholino)ethanesulfonic acid (MES) was used from pH 5.25 to 6.5; MOPS buffer was used from pH 6.75 to 7.75; and 2,2'-(Propane-1,3-diyldiimino)bis[2-

(hydroxymethyl)propane-1,3-diol] (Bis-tris propane) buffer was used from pH 8.0 to 9.5. Final concentrations were 100 nM Q723E8, 5 mM G6P, 1 mM EDTA, and 100 mM buffer. Reactions were allowed to proceed for 0, 5, and 10 min before quenching with equal volumes of 25 N H_2SO_4. Quenched reactions were kept at room temperature (~21 ^oC) to develop for 3 h before recording the absorbance at 540 nm. The activity at each pH was determined in triplicate. Reaction rates were determined by fitting the 0, 5 and 10 min time points using linear regression.

Effect of divalent metal ions on the GPI activity of Q723E8

Samples of Q723E8 were diluted into buffer containing 200 mM MOPS pH 7.0, and either 20 μM EDTA or 20 μM divalent metal salt and incubated for 1 h at 4 °C. The remaining isomerase activity was assayed using the discontinuous cysteine-carbazole assay. Final concentrations in each assay were: 100 mM MOPS pH 7.0, 500 nM enzyme, 2 mM G6P, and 10 μM metal salt or EDTA. Reactions were incubated for 3 min at 37 °C and the color was allowed to develop for 3 h at room temperature. Metal salts included $BaCl_2{\cdot}2H_2O$, $MnCl_2{\cdot}4H_2O$, $ZnCl_2$, $NiCl_2{\cdot}6H_2O$, $CoSO_4{\cdot}7H_2O$, $CuSO_4$, $FeSO_4{\cdot}7H_2O$, $CdCl_2$, $MgCl_2$, $CaCl_2{\cdot}2H_2O$, and $HgCl_2$.

Enzyme kinetics

Kinetic parameters for the isomerization of G6P to F6P or A5P to D-ribulose 5-phsohpate (Ru5P) catalyzed by Q723E8 were determined using the discontinuous cysteine carbazole assay. Individual conditions were assayed in triplicate, with 500 nM Q723E8, 100 mM MOPS pH 7.0, 1 mM EDTA, and substrate concentrations ranging from 0.125 mM to 8 mM G6P or 0.156 mM to 10 mM A5P. Individual assays, pre-heated for 3 min at 37 ^oC, and appropriate controls (F6P standards, a reaction lacking enzyme, and a reaction lacking substrate), were allowed to react for

3 min at 37 °C, before being quenched with equal volumes of 25 N H_2SO_4. The reaction was incubated at room temperature (~21 ^oC) for 3 h to develop color. Control reactions demonstrated that less than 10% of the substrate was converted to product under all the conditions tested. Kinetic parameters were obtained by fitting the data to the Michaelis-Menten equation using nonlinear least-squares regression with GraphPad Prism 5 software.

Kinetic parameters for the isomerization of F6P to G6P were determined in triplicate using a coupled assay in which the G6P produced was oxidized to 6-phosphoglucono-δ-lactone by glucose 6-phosphate dehydrogenase and its cofactor, $NADP^+$. The progress of the reaction was monitored by following the absorbance, at 340 nm, of the NADPH produced [13]. Individual assays were performed in a quartz cuvette as follows: the coupling enzyme (*E. coli* D-glucose 6-phosphate dehydrogenase, gene name *zwf*) was diluted into 2x reaction buffer (200 mM MOPS, pH 7.0, containing 2 mM EDTA) to a final assay concentration of 1 μM of the coupling enzyme. A solution containing F6P (final concentration ranging from 0.01 mM to 3 mM) and $NADP^+$ (final concentration of 0.16 mM) was then added to the cuvette, which was subsequently allowed to incubate for 1 min at room temperature. The reaction was initiated by the addition of Q723E8 (final concentration 100 nM) and then the absorbance at 340 nm was measured continuously for 5 min, at 10 s intervals, on a Hewlett-Packard 8453 diode array spectrophotometer. Progress curves were analyzed by linear regression to determine initial reaction rates (nM/s). Kinetic parameters were obtained by fitting the rate versus substrate concentration data to the Michaelis-Menten equation using the nonlinear least-squares regression function of GraphPad Prism 5 software.

Equilibrium constant (K_{eq}) determination

Solutions containing 500 nM Q723E8, 100 mM MOPS pH 7.0, 1 mM EDTA, 10% D_2O, and 5 mM either F6P or G6P were allowed to reach equilibrium at room temperature (~21 ^oC). The solutions were then analyzed by ^{31}P NMR using a Varian 400 multinuclear NMR spectrometer with a 0.05 N phosphoric acid standard (sealed within a capillary tube) set to a shift of 0 ppm. Spectra were acquired using 64 scans with a ten-second delay between scans. Preliminary studies using longer relaxation times showed no change in the peak integrals, supporting the assumption that this delay time is greater than three times the T1 relaxation parameter for both G6P and F6P.

Results

Q723E8 orthologs are found in *L. monocytogenes, L innocua, L ivanovii,* and *L. seeligeri*

A basic local alignment search tool (BLAST) search of the Listeria sequences in GenBank using the amino acid sequence of Q723E8 as the query returned 293 entries from *L. monocytogenes* strains, 8 from *L. innocua* strains, 3 from *L. ivanovii* strains and 1 from a *L. seeligeri* strain, each sharing ≥ 92% identity with the query sequence. The remainder of the hits were full-length (~320 amino acids) arabinose 5-phosphate isomerases from a wide variety of bacterial species. The 293 *L. monocytogenes* strains included *L. monocytogenes* 4a L99, a strain representative of lineage I, and *L. monocytogenes* 4b CLIP80459, a strain representative of lineage III.

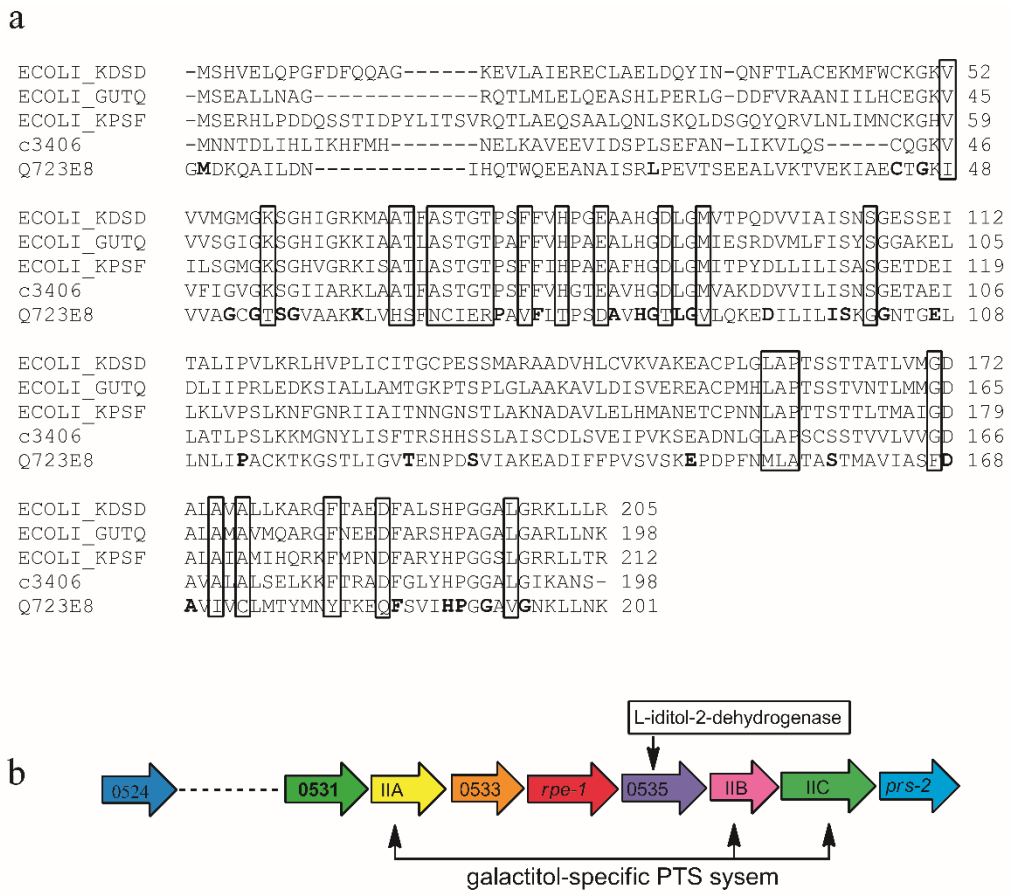

Fig. 2.1: Sequence alignment of Q723E8 and genomic context of lmof2365_0531. **a** Clustal Omega alignment of the protein sequences of Q723E8, *E. coli* CFT073 c3406, and the SIS domains of *E. coli* KdsD, GutQ and KpsF. Alignment performed using the program Clustal Omega on the EMBL-EBI server (www.ebi.ac.uk/tools/msa/clustalo). Bolded residues are conserved in all five structures. Residues in rectangular boxes are conserved in *E. coli* sequences, but different in Q723E8. **b** Genomic context of *L. monocytogenes* lmof2365_0531. Information obtained from KEGG (www.genome.jp/kegg-bin/show_genomemap?ORG=lmf&ACCESSION=LMOf2365_0531).

Q723E8 shares significant sequence identity with APIs and GPIs

A BLAST search of the RCSB Protein Data Bank using the amino acid sequence of c3406 as the query returned *L. monocytogenes* lmof2365_0531 [14] as the second-closest hit (with 33% identity), after the structure of the SIS domain from *E. coli* KdsD (51% identical) [15]. Assuming this gene could encode an API, we performed a Clustal Omega alignment of the protein given

sequence tag *L. monocytogenes* lmof2365_0531, to which we will refer using its UniProt ID, Q723E8, with c3406 and the SIS domains of *E. coli* KdsD, GutQ, and KpsF [CBS-domain containing APIs] (Fig. 2.1A) [16]. The alignment revealed that Q723E8 shares 31 conserved residues (Fig. 2.1A, bolded residues) including the putative catalytic residue His 193 (KdsD numbering) [15]. The sequence identity between Q723E8 and the SIS-domain of these four known D-arabinose 5-phosphate isomerases ranges between 30.8 and 33.3%. The genomic context of lmof2365_0531 is also very similar to that of *E. coli* gutQ.

The genomic context of lmof2365_0531 (Fig. 2.1B) [18] contains three open reading frames (0532, 0536 and 0537) that are annotated as a galactitol-specific phosphoenolpyruvate:sugar phosphotransferase system permease. Among these three phosphoenolpyruvate:sugar phosphotransferase system genes, there is an open reading frame (0535) that is annotated as an L-iditol dehydrogenase, and an open reading frame (rpe-1) encoding a D-ribulose 5-phosphate 3-epimerase, which interconverts D-ribulose 5-phosphate and D-xylulose 5-phosphate [18]. This is similar to the gut operon [17], which encodes a three-component sorbitol-specific permease (gutA, gutE, and gutB), which is part of a phosphoenolpyruvate:sugar phosphotransferase system that simultaneously imports D-sorbitol across the cell membranes and phosphorylates it [17]. The operon also contains a gene (gutD), which encodes an oxidoreductase that converts the D-sorbitol 6-phosphate into D-fructose 6-phosphate. The gutQ gene, which encodes a full-length API, is the last gene in the operon. However, it should be noted that Q723E8 also has 24 residues that are different than residues that are conserved among the SIS domains of c3406, KpsF, GutQ and KdsD (Fig 2.1A, boxes). This information led us to hypothesize that Q723E8 could possess GPI activity.

Fig. 2.2: Clustal Omega alignment of the protein sequences of Q723E8 and the C-terminal end of various PGI. PGI include rabbit PGI, PDB ID: 1G98; mouse PGI, PDB ID: 1U0F; *S. aureus* PGI, PDB ID: 3FF1; *V. cholerae* PGI, PDB ID: 3HJB. Alignment performed using the program Clustal Omega on the EMBL- EBI server (www.ebi.ac.uk/tools/msa/clustalo). Bolded residues are conserved in all five structures. Residues in rectangular boxes are conserved in *E. coli* sequences, but different in Q723E8

An alignment of Q723E8 with various SIS-domain GPIs within the PDB (Fig. 2.2) reveals that Q723E8 is a much smaller protein consisting of only 200 amino acids whereas the other four GPIs presented (rabbit GPI, PDB ID: 1G98; mouse GPI, PDB ID: 1U0F; *Staphyloccus aureus* GPI, PDB ID: 3FF1; *Vibrio cholerae* GPI, PDB ID: 3HJB) all contain over 500 amino acids and a typical SIS-domain containing GPI ranges between 450 and 600 amino acids. The alignment shows that these four GPIs and Q723E8 share only 5 completely conserved residues across all five proteins. When structurally comparing Q723E8 and mouse GPI (PDB ID: 1G98) the catalytic histidine H388 of 1G98 structure aligns to that of H83 in Q723E8. However, the catalytic E357 in the four known

GPIs aligns with T28 in Q723E8 in the alignment shown in Fig. 2.2, but structurally the nearest residue is C68.

Q723E8 possesses GPI activity

To determine if Q723E8 can catalyze the isomerization of sugars or phosphosugars other than A5P, we tested its ability to catalyze the isomerization of a series of aldoses, including D-ribose, D-ribose 5-phosphate, D-arabinose, D-glucose 6-phosphate, D-glucose 1-phosphate, D-glucosamine 6-phosphate, and D-mannose 6-phosphate, to ketoses. The only potential substrate that produced an assay signal that was significantly above the background was G6P (data not shown).

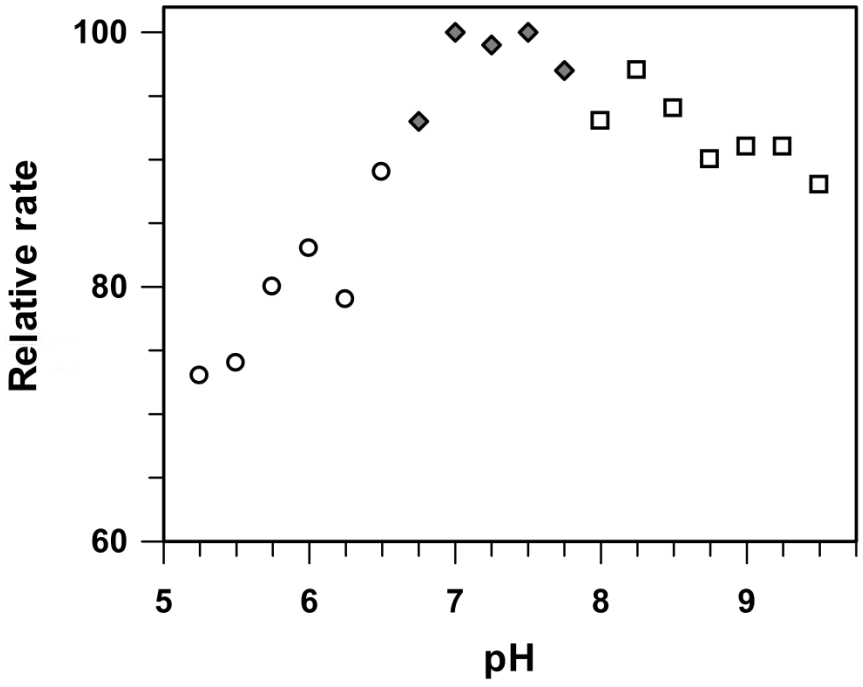

Fig. 2.3: pH-rate profile for the conversion of G6P to F6P by Q723E8. Open circle data points taken using MES buffer, filled diamond data points taken using MOPS buffer, open square data points taken using BTP buffer. Data are expressed as rate relative to the rate at pH 7.0, which was set at 100

The pH-rate profile measured using G6P as substrate (Fig. 2.3) revealed a broad pH optimum from approximately pH 7.0 to pH 9.0.

Table 2.2. Kinetic constants for catalysis by Q723E8.

F6P → G6P			G6P → F6P			
k_{cat} (s^{-1})	K_m (µM)	k_{cat}/K_m($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_m (µM)	k_{cat}/K_m($M^{-1}s^{-1}$)	K_{eq}
0.22 ± 0.014	69 ± 12	3.2 x 10^3	1.78 ± 0.07	0.95 ± 0.08	1.86 x 10^3	0.29

A detailed kinetic analysis of the GPI activity of Q723E8 at pH 7 was performed in the presence of EDTA. The results (Table 2.2) indicate that Q723E8 is a relatively poor GPI, when compared to the GPIs (EC 5.3.1.9) contained in the BRENDA database (data not shown) [19].

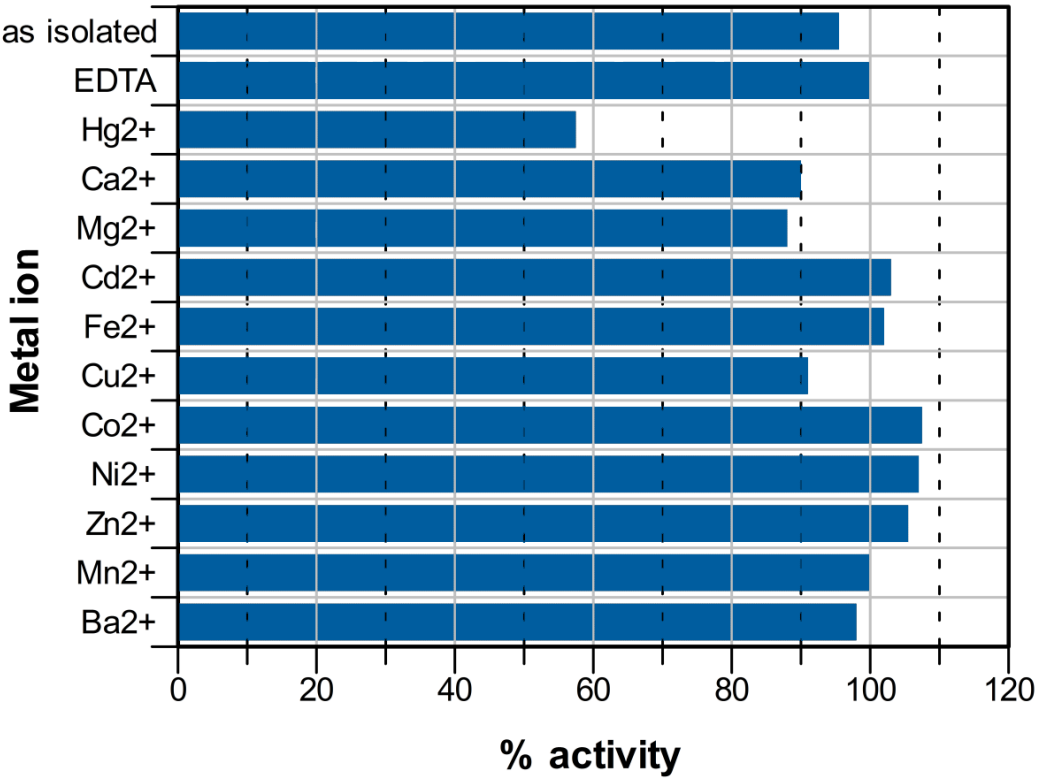

Fig. 2.4 Effect of metal ions upon the glucose 6-phosphate isomerase activity of Q723E8. The enzyme was pre-incubated with no additive (as isolated), EDTA, or a divalent metal ion, then assayed for glucose6-phosphate isomerase activity. The activity of the EDTA-containing sample was assigned a value of 100 %

Since APIs, such as KdsD, are known to co-purify with inhibitory divalent metal ions, and the cupin-type GPIs require metal ions for activity [22], we chose to assay the recombinant Q723E8 for the presence of co-purified metal ions, and test its activity in the presence of added divalent metal ions. High-resolution inductively coupled plasma mass spectroscopy of recombinant Q723E8, as isolated, revealed that the enzyme contained only traces of divalent metal ions. Addition of divalent metal ions to the assay mixture did not lead to significant enhancement or inhibition of catalysis, except in the case of divalent mercury ions (Fig. 2.4), which inhibited catalytic activity.

Q723E8 lacks API activity

To prepare protein samples for analysis, lmof2365_0531 was cloned from the genomic DNA of *L. monocytogenes* strain 4b f2365 using PCR. The PCR product was incorporated into plasmid pCR2.1 TOPO and then transferred to plasmid pT7-7. DNA sequencing of the resulting plasmid revealed undesired mutations at the 3-prime end of the lmof2365-0531 gene. After removing the undesired mutations via site directed mutagenesis the resulting protein expression

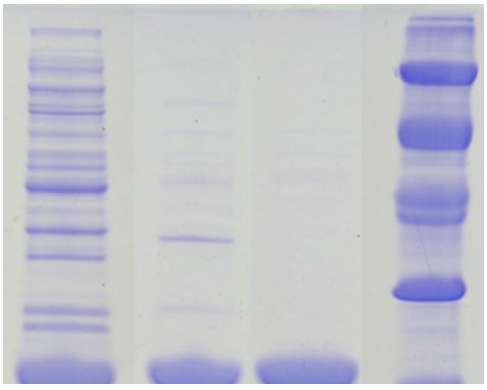

Fig. 2.5: The purification of Q723E8. Lane 1: crude lysate, lane 2: supernatant from 60 % ammonium sulfate precipitation, lane 3: final product from mono Q column, lane 4: low range SDS-PAGE standards

was observed to be low. The lmof2365_0531 gene was subsequently transferred to plasmid pET17b for recombinant expression in *E. coli* BL21(DE3). Selective ammonium sulfate precipitation of contaminating *E. coli* proteins, at 60% saturation, yielded significant purification. Further purification, post dialysis, using anion exchange chromatography yielded protein that was >95% pure by SDS-PAGE analysis (Fig. 2.5). LC-MS (liquid chromatography – mass spectrometry) analysis of the homogeneous protein showed a molecular weight of 21 315.9 Da, (calculated value 21 315.7 Da). Confident the desired protein had been expressed and purified;

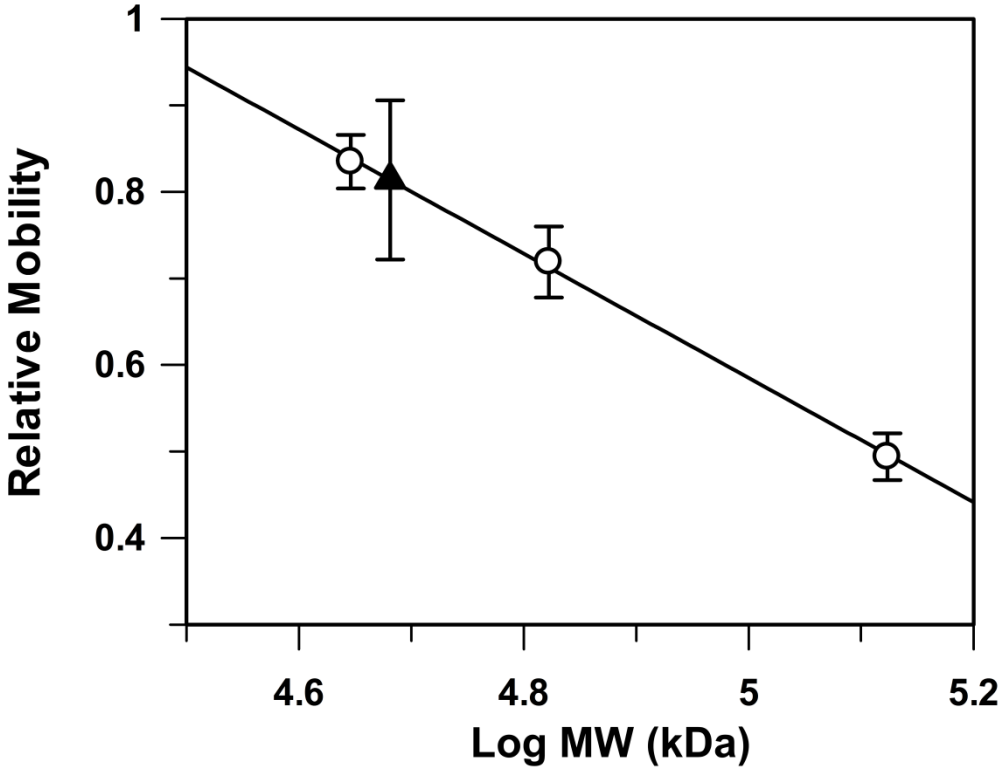

Fig. 2.6: Native molecular weight determination for Q723E8. Open symbols represent the protein standards. The closed triangle is located at the average relative mobility and molecular weight determined from two independent observations. Error bars represent the standard error of each measurement. Standard error was calculated from these two experimental MW values to yield the error associated with the measurement.

its quaternary structure was investigated using non-denaturing PAGE. Native Q723E8 migrated as a species with molecular weight 45 040 Da (2.1 times the monomer weight) suggesting that

Q723E8 exists as a dimer in solution (Fig. 2.6). To determine the potential API activity of recombinant Q723E8, the protein was assayed with concentrations of A5P up to 10 mM using the cysteine-carbazole method to detect the formation of Ru5P. No evidence of Ru5P production was observed, suggesting that Q723E8 has no API activity in vitro.

To test the possibility of an in vivo cytosolic cofactor, we tested the ability of Q723E8 to complement an API defect in *E. coli.* The deletion of both *kdsD* and *gutQ* produces a conditional ΔAPI construct, *E. coli* TCM15 [10]. This strain cannot survive without API activity unless the medium is supplemented with a combination of A5P, which is required for lipopolysaccharide production, and G6P, which is required to induce the transporter that allows A5P to enter the cell [10,20]. A5P is a high affinity, but non-inducing, substrate of the G6P-inducible hexose phosphate transport system (the Uhp system) [21]. This defect can be complemented by supplying an API

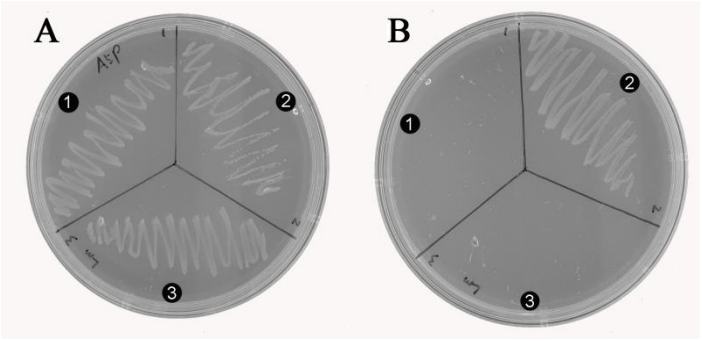

Fig. 2.7: Complementation of an A5P auxotroph on agar plates. **A** Agar plate containing LB medium supplemented with 15 lM A5P, 10 lM G6P, 0.1 mg/mL ampicillin, and 0.05 mg/mL kanamycin. **B** Agar plate containing LB supplemented with 0.1 mg/mL ampicillin and 0.05 mg/mL kanamycin. In both panels, the wedges were streaked with: TCM15 harboring pT7-7 (vector control, wedge 1), TCM15 harboring pT7-7 c3406 (c3406 control, wedge 2) and TCM15 harboring pT7-7Q723E8 (Q723E8, wedge 3)

gene in trans using the plasmid pT7-7, as shown previously with *c3406* [7]. Attempts to complement the API defect in *E. coli* strain TCM15 using the gene encoding Q723E8, sub-cloned

into pT7-7, failed to form colonies. This inability of *E. coli* TCM15 containing pT7-7Q723E8 to form colonies on agar media lacking A5P is identical to the phenotype displayed by the empty vector (Fig. 2.7), and distinctly different than the phenotype displayed by c3406. This suggests that Q723E8 lacks significant API activity in a bacterial cell. Therefore, we conclude that Q723E8 is not a functional API either in vivo or in vitro, despite its sequence similarity to *E. coli* c3406 and contextual similarity to *gutQ*.

Discussion

What began as a quest to find the first API in a Gram-positive organism, in an effort to further understand the physiological function of APIs not associated with LPS biosynthesis, has resulted in the discovery of a potentially new GPI. The quaternary structure of Q723E8, in solution, is also more similar to that of GPIs than that of APIs. We found Q723E8 to be a dimer in solution. The X-ray crystal structure of Q723E8 [14] suggests that the putative active site includes residues from two of the monomers within asymmetric unit. These results are similar to that of known GPIs, including both the typical bacterial GPI and cupin-type GPI [22,23].

The GPI activity of Q723E8 is extremely weak with a kcat two orders of magnitude slower than the GPI of *Bacillus stearothermophilus* [25]. and has a very broad pH-rate profile, causing us to question its physiological significance. The genome of *L. monocytogenes* strain 4b F2365 contains a gene (lmof2365-2338) that encodes a typical bacterial GPI [23] that shares >90% similarity with the GPI of the prototypical Gram-positive organism *Bacillus subtilis*. This GPI alone should be sufficient to support the aerobic and anaerobic fermentation of glucose that has been reported for *Listeria* species [24]. The GPI activity of Q723E8 seems, therefore, superfluous. Therefore, we speculate that the GPI activity of Q723E8 may not be this enzyme's main

physiological role within the cell. However, its physiologically relevant role may involve the isomerization of substrates that were not available for testing, but any additional role can only be speculated.

Acknowledgements and Chapter Contributions

This work was supported, in part, by National Institutes of Health grant AI-061531 (to R.W. Woodard). The authors thank Jim Windak and Paul Lennon of the University of Michigan Department of Chemistry Mass Spectroscopy facility for performing the LC-MS analysis of recombinant lmof2365_0531, and Ted Houston of the University of Michigan Department of Geology's W.M. Keck Elemental Geochemistry Laboratory for performing the metals analysis.

Cloning of the lmof2365_0531 gene was done by Assimon, V.A., the pH rate profile was completed by Schaub, J.M., native oligomeric state was done by Holt, M.C., protein purification was carried out by Wang, P.F., all other experiments and writing of the manuscript were completed by Cech, D.L., reviewing and editing of the manuscript was carried about by Holler, T.P., and Woodard, R.W.

References

1. Murray EGD, Webb RA, Swann MBR (1926) A disease of rabbits characterised by a large mononuclear leucocytosis, caused by a hitherto undescribed bacillus *Bacteruim Monocytogenes*. *J Pathol Microbiol* **29**:407-439.

2. Gray ML, Killenger AH (1966) *Listeria monocytogenes* and Listeric infections. *Bacteriol Rev* **30**:309-382.

3. Gandhi M, Chikindas ML (2007) *Listeria*: A foodborne pathogen that knows how to survive. *Int J Food Microbiol* **113**:1-15.

4. Hain T, Ghai R, Billion A, Kuenne CT, Steinweg C, Izar B, Mohamed W, Abu Mraheil M, Domann E, Schaffrath S, Kärst U, Goesmann A, Oehm S, Pühler A, Merkl R, Vorwerk S, Glase P, Garrido P, Rusinol C, Buchrieser C, Goebel W, Chakraborty T (2012) Comparitive genomics and transcriptomics of lineages I, II, and III strains of *Listeria monocytogenes*. *BMC Genomics* **13**:144.

5. Lingu B, Ricke SC, Johnson MG (2009) Growth, survival, proliferation and pathogenesis of *Listeria monocytogenes* under low oxygen or anaerobic conditions: A review. *Anaerobe* **15**:7-17.

6. Pine L, Malcolm GB, Brooks JB, Daneshvar MI (1988) Physiological studies on the growth and utilization of sugars by *Listeria* species. *Can J Microbiol* **35**:245-254.

7. Mosberg JA, Yep A, Meredith TC, Smith S, Wang P-F, Holler TP, Mobley HLT, Woodard RW (2011) A Unique Arabinose 5-Phosphate Isomerase Found within a Genomic Island Associated with the Uropathogenicity of *Escherichia coli* CFT073. *J Bacteriol* **193**: 2981–2988.

8. Bateman A (1999) The SIS domain: a phosphosugar-binding domain. *Trends in Biochemical Sciences* **24**: 94–95.

9. Meredith TC, Woodard RW (2003) *Escherichia coli* YrbH Is a D-Arabinose 5-Phosphate Isomerase. J Biol Chem **278**: 32771–32777.

10. Meredith TC, Woodard, RW (2005) Identification of GutQ from *Escherichia coli* as a D-Arabinose 5-Phosphate Isomerase. *J Bacteriol* **187**:6936–6942.

11. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning : a laboratory manual, 1st edn. Cold Spring Harbor Laboratory Press: Cold Spring harbor, N.Y.

12. Tabor S (1985) A Bacteriophage T7 RNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes. *PNAS* **82**: 1074–1078.

13. Sandoval JM, Arenas FA, Vásquez CC (2011) Glucose-6-Phosphate Dehydrogenase Protects *Escherichia coli* from Tellurite-Mediated Oxidative Stress. *PLoS ONE* **6**: e25573.

14. Joint Center for Structural Genomics Crystal structure of a putative sugar isomerase (lmof2365_0531) from Listeria monocytogenes str.4b f2365 at 1.60 Å resolution.

15. Gourlay LJ, Sommaruga S, Nardini M, Sperandeo P, Dehò G, Polissi A, Bolognesi M (2010) Probing the active site of the sugar isomerase domain from *E. coli* arabinose-5-phosphate isomerase via X-ray crystallography. *Protein Science* **19**: 2430–2439.

16. Meredith TC, Woodard RW (2006) Characterization of *Escherichia coli* D-arabinose 5-phosphate isomerase encoded by kpsF: implications for group 2 capsule biosynthesis. *Biochemical Journal* **395**:427-432.

17. Yamada M, Yamada Y, Saier MH Jr. (1987) Physical and genetic characterization of the glucitol operon in *Escherichia coli*. *J Bacteriol* **169**: 2990–2994.

18. Nelson KE (2004) Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. *Nucleic Acids Research* **32**: 2386–2395.

19. Scheer M, Grote A, Chang A, Schomburg I, Munaretto C, Rother M, Sohngen C, Stelzer M, Thiele J, Schomburg D (2010) BRENDA, the enzyme information system in 2011. *Nucleic Acids Research* **39**: D670–D676.

20. Meredith TC, Aggarwal P, Mamat U, Lindner B, Woodard RW (2006) Redefining the Requisite Lipopolysaccharide Structure in *Escherichia coli. ACS Chem Biol* **1**: 33–42.

21. Eidels L, Rick PD, Stimler NP, Osborn MJ (1974) Transport of D-Arabinose-5-Phosphate and D-Sedoheptulose-7-Phosphate by the Hexose Phosphate Transport System of *Salmonella typhimurium*. *J Bacteriol* **119**: 138–143.

22. Hansen T, Schlichting B, Felgendreher M, Schonheit P (2005) Cupin-Type Phosphoglucose Isomerases (Cupin-PGIs) Constitute a Novel Metal-Dependent PGI Family Representing a Convergent Line of PGI Evolution. *J Bacteriol* **187**: 1621–1631.

23. Katz LA (1996) Transkingdom transfer of the phosphoglucose isomerase gene. *J Mol Evol* **43**: 453–459.

24. Romick TL, Fleming HP, McFeeters RF (1996) Aerobic and anaerobic metabolism of *Listeria monocytogenes* in defined glucose medium. *Appl Environ Microbiol* **62**: 304–307.

25. Meng, M, Chane, TL, Sun YJ, Hsiao, CD (1999) Probing the location and function of the conserved histidine residue of phosphoglucose isomerase by using an active site directed inhibitor N-bromoacetylethanolamine phosphate. *Protein Science* **8**: 2438-2443.

CHAPTER III

The Arabinose-5-phosphate Isomerase of *Bacteroides fragilis*: Insight Into Regulation of Single-domain Arabinose Phosphate Isomerases

The work described in this chapter has been published [Cech, D., Wang, P.F., Holler, T.P., Woodard, R.W., *Analysis of the arabinose-5-phosphate isomerase of Bacteroides fragilis provides insight into regulation of single-domain arabinose phosphate isomerases.* J Bacteriol, 2014. **196**(15): p. 2861-8.]

Summary

Arabinose-5-phosphate isomerases (APIs) catalyze the interconversion of D-ribulose-5-phosphate and D-arabinose-5-phosphate, the first step in the biosynthesis of 3-deoxy-D-manno-octulosonic acid (Kdo), an essential component of the lipopolysaccharide in Gram-negative bacteria. Classical APIs, such as *Escherichia coli* KdsD, contain a sugar isomerase domain and a tandem cystathionine beta-synthase domain. Despite substantial effort, little is known about structure-function relationships in these APIs. We recently reported an API containing only a sugar isomerase domain. This protein, c3406 from *E. coli* CFT073, has no known physiological function. In this study, we investigated a putative single-domain API from the anaerobic Gram-negative bacterium *Bacteroides fragilis*. This putative API (UniProt ID Q5LIW1) is the only protein encoded by the *B. fragilis* genome with significant identity to any known API, suggesting that it is responsible for lipopolysaccharide biosynthesis in *B. fragilis*. We tested this hypothesis by preparing recombinant Q5LIW1 protein (here referred to by the UniProt ID Q5LIW1), characterizing its API activity *in vitro*, and demonstrating that the gene encoding Q5LIW1

(GenBank ID YP_209877.1) was able to complement an API-deficient *E. coli* strain. We demonstrated that Q5LIW1 is inhibited by cytidine 5'-monophospho-3-deoxy- D-*manno*-2-octulosonic acid, the final product of the Kdo biosynthesis pathway, with a K_i of 1.91 µM. These results support the assertion that Q5LIW1 is the API that supports lipopolysaccharide biosynthesis in *B. fragilis* and is subject to feedback regulation by CMP-Kdo. The sugar isomerase domain of *E. coli* KdsD, lacking the two cystathionine beta-synthase domains, demonstrated API activity and was further characterized. These results suggest that Q5LIW1 may be a suitable system to study API structure-function relationships

Introduction

Arabinose-5-phosphate isomerases (APIs) catalyze the interconversion of D-ribulose-5-phosphate (Ru5P), the product of the oxidative phase of the pentose phosphate pathway (1), and D-arabinose-5-phosphate (A5P), the first intermediate in the biosynthesis of 3-deoxy-D-manno-octulosonic acid (Kdo). Kdo is an essential component of the cell envelope of Gram-negative bacteria (2), and is also found in certain algae and plants (3). Kdo biosynthesis and activation represents an attractive pathway for drug targeting because Kdo is not synthesized by humans (1).

The Woodard laboratory has identified four distinct API genes from various strains of the model Gram-negative organism *Escherichia coli*. The genome of *E. coli* K-12 encodes two distinct APIs, KdsD and GutQ. The *kdsD* gene, formerly known as *yrbH*, is located in the *yrb* gene cluster. Its product, KdsD, catalyzes the formation of A5P used in the biosynthesis of Kdo ((2)). The *gutQ* gene is found in the glucitol (sorbitol) operon, which encodes proteins involved in the utilization of glucitol as a sole carbon source. The physiological role of GutQ, which has KdsD-like levels of API activity and can produce sufficient A5P to support Kdo biosynthesis in the absence of KdsD ((3)), is somewhat of an enigma because A5P has not been directly implicated in the import or utilization of sorbitol. *E. coli* strains that produce Group II K-antigens (e.g. *E. coli* CFT073) contain a third API, KpsF. Since Group II K-antigens contain Kdo, KpsF is considered responsible for the biosynthesis of the Kdo required by the Group II K-antigen biosynthetic machinery (7). In support of this hypothesis, the *kpsF* gene is found in the *kps* cluster, which also encodes a homolog of the *kdsB* gene (*kpsU*). The *kdsB* gene encodes cytidine 5'-monophospho-3-deoxy-D-*manno*-2-octulosonic acid (CMP-Kdo) synthetase, another enzyme in the Kdo biosynthetic pathway. Finally,

E. coli CFT073 has a fourth API encoded by the gene given sequence tag c3406. This gene, though found on a genomic island implicated in virulence, is non-essential for the virulent phenotype ((4)). Like *gutQ*, c3406 is situated near a gene implicated in carbohydrate utilization.

Three of these APIs, KdsD, GutQ, and KpsF, consist of two domains – a sugar isomerase (SIS) domain and a tandem cystathionine beta synthase (CBS) domain. The purpose of the CBS domains in these proteins has not been established, but in other contexts CBS domains have been implicated in the modulation of enzyme activity ((5)). It is also not clear if and how single-domain APIs are regulated. The X-ray crystal structure of a full-length, CBS domain-containing API has remained elusive despite many attempts, including the crystallization of mutants ((6)). The lack of a full length API crystal structure along with very little knowledge about how APIs are regulated has impeded structure-function studies with this enzyme.

The API encoded by c3406, in contrast, consists of a single SIS domain and is therefore much smaller than KdsD, GutQ, and KpsF. A search for similar proteins in the RCSB Protein Data Bank revealed the structure (PDB ID: 3ETN) of an SIS-domain protein, YP_209877.1 (UniProt ID: Q5LIW1) from *Bacteroides fragilis* NCTC 9343, in complex with CMP-Kdo, the end-product of the Kdo biosynthesis pathway. *B. fragilis* is a Gram-negative anaerobic bacterium that incorporates a Kdo-containing lipopolysaccharide (LPS) in its cell envelope ((7)). A search of the *B. fragilis* NCTC 9343 genome found that it contains an open reading frame orthologous to all 4 enzymes in the well-established Kdo biosynthetic pathway of *E. coli* K-12 (*kdsA*-*kdsD*). The protein Q5LIW1, which shares 34% identity with c3406, is the only protein found in *B. fragilis* with substantial identity to any API. This information led us to hypothesize that Q5LIW1, like KdsD in *E. coli* , is the API that supports LPS biosynthesis in *B. fragilis*, and furthermore that its enzymatic activity may

be feedback regulated by CMP-Kdo. To test these hypotheses, Q5LIW1 was cloned, expressed and purified. Its activity as an API was subsequently characterized both in vitro and in vivo (*E. coli* TCM15 cells). Finally, the ability of CMP-Kdo to inhibit the API activity of Q5LIW1 was probed.

Previous reports from our lab have shown that c3406, which is the only API from *E. coli* that naturally lacks CBS domains, can complement an API defect in *E. coli* TCM15 (8). This information, along with the results of the studies of Q5LIW1 reported here, led us to question if a truncated *E. coli* KdsD containing only the SIS domain would retain sufficient API activity to complement the API deficiency in TCM15. The results of these experiments are also reported below.

Materials and Methods

Materials

Genomic DNA from *Bacteroides fragilis* NCTC 9343 was purchased from the American Type Culture Collection (catalogue # 25285D-5). Primers for PCR were synthesized by Integrated DNA Technologies (Coralville, Iowa, USA). Failsafe™ PCR PreMix Selection kit was purchased from Epicentre Biotechnologies (Madison, WI, USA). PCR was performed in a MJ Research PTC-200 Peltier Thermal Cycler. TA TOPO cloning was performed via the TA TOPO cloning kit purchased from Invitrogen (Grand Island, NY, USA). Enzymes for subcloning were purchased from New England Biolabs (Ipswich, WI, USA). The Promega Wizard Miniprep kit (Madison, WI, USA) was used for plasmid DNA purification. DNA sequencing was performed by the University of Michigan Biomedical Resources Core Facility. Metal salts used for assays were purchased as high-purity solids from Alfa-Aesar (Ward Hill, MA, USA) and used without further purification.

Bacterial strains, plasmids, primers, and growth media

The bacterial strains, primers, and plasmids used in this study are described in Table 3.1. *E. coli* TCM15 is a derivative of BW30270, *E. coli* K-12 MG1655 *rph*$^{+}$*fnr*$^{+}$, in which the *gutQ* and *kdsD* genes were disrupted via the phage λ Red recombinase system (6,12). All strains were grown in LB medium (13). TCM15 cultures were supplemented with A5P (15 μM) and D-glucose-6-phosphate (G6P; 10 μM). A5P is necessary to complement the API defect and G6P induces the transporter necessary for the uptake of A5P (8).

Table 3.1 Strains, Plasmids, and Primers used in this study.

Item	Description	Source
E. coli		
TOP10	F^{-} *mcrA* Δ(*mrr*-*hsdRMS*-*mcrBC*) F80*lacZ*ΔM15 Δ*lacX*74 *recA1 araD*139 Δ(ara-leu)7697 *galU galK rpsL* (StrR) endA1 *nupG*	Invitrogen
TCM15	BW30270(*ΔkdsDΔgutQ*)	Ref. 6
Rosetta 2 (DE3) pLysS	F^{-} *ompT* hsdS$_B$($r_B^-m_B^-$) *gal dcm* (DE3) pLysSRARE2 (Cam^R)	EMD Millipore
Plasmids		
pCR2.1-TOPO	Library plasmid	Invitrogen
pT7-7	Expression vector	Ref. 13
pT7-7-c3406	*E.coli* CFT073 c3406 inserted into NdeI/BamHI of pT7-7, Amp^R	Ref. 8
pT7-7-YP_209877.1	*B. fragilis* YCH46 BF0137 inserted into NdeI/BamHI of pT7-7, Amp^R	This study
pT7-7-yrbH	*E. coli* KdsD inserted into NdeI/BamHI of pT7-7, Amp^R	Ref. 5
pT7-7-EcKdsDΔ2CBS	Truncation of pT7-7-yrbH to the first 213 amino acids.	This study
Primers		
BF0137.F	CAATTAAGATCATATGATTGAATCTATTCAAGAACTC[a]	IDT
BF0137.R	CGAGTATAGTCTCTCCGGATCCGATTACTTTACGC[b]	IDT
EcKdsDΔ2CBS.F	CGATATTATGCATACGGGCTAAGAGATCCCGCATGTTAAGAAAACGG	IDT
EcKdsDΔ2CBS.R	CCGTTTTCTTAACATGCGGGATCTCTTAGCCCGTATGCATAATATCG	IDT

Cloning, expression, and purification of *B. fragilis* Q5LIW1

The YP_209877.1 gene, encoding Q5LIW1, was amplified from the genomic DNA of *Bacteroides fragilis* NCTC 9343 using the primers BF0137START and BF0137STOP (Table 3.1), which were designed to incorporate *Nde*I and *Bam*HI sites. PCR products were purified via extraction from a 1% (w/v) agarose gel using a Qiagen QIAquick Gel Extraction Kit, inserted into vector pCR2.1-TOPO using the TA TOPO cloning kit, and subcloned, after digestion of the insert

containing vector with *Nde*I and *Bam*HI, into similarly restricted, expression vector pT7-7 (14). The resulting plasmid, pT7-7-YP_209877.1, was transformed into *E. coli* TOP10 chemically competent cells. DNA sequencing of the resulting plasmid, pT7-7-YP_209877.1, confirmed that the YP_209877.1 gene was identical to the sequence published in the NCBI entry. This plasmid was subsequently transformed into *E. coli* Rosetta 2 (DE3) pLysS chemically competent cells. A fresh transformant was grown in LB medium supplemented with 100mg/l ampicillin and 30 mg/l chloramphenicol at 37°C while shaking at 250 rpm until the optical density at 600 nm (OD600) reached 0.6. The culture was cooled to 19°C, and expression of protein was induced with the addition of isopropyl β-D-1-thiogalactopyranoside to 0.42 mM. After 16 h of incubation at 19°C, cells were harvested by centrifugation (6000 x g, 10 min, 4°C). The pellet was suspended in 25 ml buffer B (40 mM Tris-HCl, pH 7.5), sonicated on ice (3 cycles of 20 s bursts, 2 min pauses between pulses), and clarified by centrifugation at 18,000 x g for 30 min. Solid ammonium sulfate was slowly added, while stirring, to the clarified lysate to reach 50% saturation at 4°C and allowed to stir for an additional 10 min. Precipitated proteins were removed by centrifugation at 15,000 x g for 25 min and solid ammonium sulfate was slowly added to the supernatant to reach 60% saturation (at 4°C) while stirring. The solution was allowed to stir for an additional 10 min after the final addition of ammonium sulfate. Q5LIW1 was pelleted, along with some contaminating proteins, by centrifugation at 15,000 x g for 25 min and resuspended in buffer A (40 mM Tris-HCl, pH 7.5, 1.4 M ammonium sulfate) prior to being passed through a Millipore PVDF 0.22 µm filter. The solubilized protein was loaded onto a Phenyl Superose column, which had been equilibrated with buffer A. Protein was eluted with an inverse linear gradient of 40-100% (v/v) buffer B in buffer A over 15 column volumes at 0.5 ml/min. Fractions containing, as determined by SDS-

PAGE gel, Q5LIW1 were pooled and concentrated using an Amicon ultra centrifugal filter (10,000 MWCO) and buffer exchanged into 40 mM Tris-HCl, pH 7.5 and 25% (v/v) glycerol and stored at -80°C.

Site-directed mutagenesis, expression, and purification of *E. coil* KdsDΔ2CBS

Plasmid template DNA (pT7-7-yrbH, approximately 0.5 pmole) was added to a PCR cocktail containing 1X *Pfu*Turbo buffer, 200 µM each dNTP, 3 µM EcKdsDΔ2CBS.F and EcKdsDΔ2CBS.R, and 2.5 U *Pfu*Tubro DNA polymerase. The reaction was incubated in a MJ research PTC-200 thermal cycler with standard site-directed mutagenesis parameters. The reaction mixture was digested with *DpnI* for 1 h at 37°C and purified via extraction from an agarose gel. The resulting plasmid pT7-7-EcKdsDΔ2CBS was transformed into *E. coli* TOP10 chemically competent cells. DNA sequencing of the plasmid pT7-7-EcKdsDΔ2CBS confirmed that the KdsD gene had been truncated to include a stop codon between the SIS domain and first CBS domain (resulting in a gene encoding for a 213 amino acid protein). The plasmid was subsequently transformed into *E. coli* BL21 (DE3) chemically competent cells. A fresh transformant was grown and cells were harvested as above. EcKdsDΔ2CBS was purified via a Hi-Load™ (16/10) Q-Sepharose fast flow column, followed by ammonium sulfate precipitation as previously described (5).

Molecular mass determination

The subunit mass of Q5LIW1 was determined via electrospray ionization mass spectrometry utilizing an LCT electrospray/time-of-flight spectrometer. The native molecular mass (MW) of Q5LIW1 was estimated by gel filtration chromatography on a HiPrep (26/60)

Sephacryl S-100 column. Standards, run in triplicate, included BSA dimer (132.4 kDa), BSA monomer (66.2 kDa), chicken egg white ovalbumin monomer (44.3 kDa), and cytochrome c (12.4 kDa). A Log (MW) vs. elution volume/void volume (V_e/V_o) plot was fit by linear regression, using Microsoft Excel. The MW of Q5LIW1 was determined experimentally by measuring its elution volume and calculating the Log (MW) from the standard curve.

Metal content analysis

Enzyme samples were prepared for metal content analysis by 24 h of dialysis at 4°C against 2 liters of metal-free 40 mM Tris-HCl pH 8.5 buffer and stored in metal-free glass vials. The divalent metal content, of each sample and the buffer used in dialysis as a control, was determined using high resolution inductively coupled plasma mass spectrometry on a Finnigan MAT Element instrument at the University of Michigan Department of Geology.

Substrate specificity

To determine recombinant Q5LIW1's ability to convert aldoses to ketoses, the enzyme was assayed with various aldoses including A5P, G6P, D-glucose-1-phosphate, D-ribose-5-phosphate, D-arabinose, D-mannose-6-phosphate, D-glucosamine-6-phosphate, and D-ribose using the discontinuous cysteine-carbazole assay (5). For each potential substrate, in triplicate, a 50 µl solution of 100 nM enzyme, 100 mM 2-Bis(2-hydroxyethyl)amino-2-(hydroxymethyl)-1,3-propanediol (Bis-tris propane), pH 8.5, and 10 mM substrate, was incubated for 10 min at 37°C before being quenched with an equal volume of 25 N H_2SO_4. A 90 µl aliquot was transferred to a separate microplate containing a freshly prepared cysteine-carbazole assay mixture (10 µl of a 0.12% ethanolic carbazole solution, 10 µl of 1.5% aqueous cysteine-HCl, and 230 µl of 25 N

H_2SO_4), and incubated at room temperature (~21°C) for 3 h to develop color. Enzyme substrates were identified by comparing the color of the reaction to the color of a no-enzyme control.

Complementation of a *kdsD*/*gutQ* defect in *E. coli*

Electrocompetent cells were prepared, from the *E. coli* TCM15 variant that carries a kanamycin resistance cassette, by growing the cells in LB medium until early log phase (OD600, ~0.5). The cells were then harvested via centrifugation and washed three times in ice-cold 10% glycerol (8). Empty pT7-7 vector, pT7-yrbH, pT7-7-EcKdsDΔ2CBS, pT7-7-c3406, and pT7-7-YP_209877.1 were separately transformed into electrocompetent *E. coli* TCM15 cells. Transformed cells were grown on LB/agar plates supplemented with 100mg/l ampicillin, 50mg/l kanamycin, 15 µM A5P, and 10 µM G6P. Liquid LB media supplemented with the same concentrations of antibiotics, A5P, and G6P, were inoculated with single colonies and grown overnight at 37°C. Cells were washed twice, by centrifugation at 6000 rpm and resuspension in liquid LB medium, to remove A5P and G6P in the overnight culture, streaked on LB/agar plates with and without A5P/G6P, and incubated overnight at 37°C. Genes that complemented the API defect allowed the transformed cells to grow both with and without added A5P/G6P.

Determining the pH-rate profile of Q5LIW1

The optimal pH for Q5LIW1 was determined by assaying Q5LIW1's enzymatic activity, using the discontinuous cysteine-carbazole assay, in a series of buffer solutions, of varying pH. Buffer solutions (200 mM buffer, 2 mM EDTA) were prepared in 0.25 pH increments from 5.0 to 9.5 at 37 ^{o}C. 2-(*N*-morpholino)ethanesulfonic acid (MES) buffer was used from pH 5.0 to 6.0 and Bis-tris propane buffer from pH 6.25 to 9.5. Final reaction concentrations were 100 mM buffer,

1 mM EDTA, 10 mM A5P, and 400 nM enzyme. Reactions were incubated at 37°C for 3 min before being quenched with equal volumes of 25 N H_2SO_4. Quenched reactions were developed, as above, and incubated at room temperature (~21°C) for 3 h before measuring the absorbance at 540 nm. The enzymatic activity at each pH was determined in triplicate and reaction rates were determined by fitting time points using linear regression.

Enzyme kinetics

Kinetic parameters for the isomerization of A5P to Ru5P or G6P to D-fructose-6-phosphate (F6P) catalyzed by Q5LIW1 and EcKdsDΔ2CBS were determined, in triplicate, using the discontinuous cysteine-carbazole assay with 100 nM Q5LIW1 or EcKdsDΔ2CBS, 100 mM Bis-tris propane, pH 8.5, 1 mM EDTA, and substrate concentrations ranging from 0.156 mM to 20 mM A5P or 0.195 mM to 25 mM G6P. Individual assays, which were pre-heated at 37°C for 3 min before the addition of enzyme to initiate the reaction, and appropriate controls, were allowed to react for 3 min at 37°C before being quenched with equal volumes of 25 N H_2SO_4. The quenched reactions were developed as described above. Control reactions, with Ru5P/F6P, showed that less than 10% of substrate was converted to product in all conditions tested. Kinetic parameters were obtained by fitting the data to the Michaelis-Menten equation using nonlinear least-squares regression with GraphPad Prism 5 software.

Assay of A5P isomerase activity

Kinetic parameters for the isomerization of Ru5P to A5P were determined, in triplicate, using a modified coupled Aminoff assay utilizing the 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase (Kdo8PS) from *Arabidopsis thaliana* (15,16). Reaction mixtures containing final

concentrations of phosphoenolpyruvate (PEP, 10 mM), Kdo8PS (0.15 mg/ml), Bis-tris propane pH 8.5 (100 mM), EDTA (1 mM), and Ru5P final concentrations ranging from 0 to 10 mM were heated, separately from mixtures of Q5LIW1 or EcKdsDΔ2CBS (100 nM final concentration) to 37°C for 3 min. The reaction was then initiated by the addition of Q5LIW1 or EcKdsDΔ2CBS and incubated for 2 min at 37°C. Reactions were quenched with an equal volume of 10% (w/v) trichloroacetic acid. To develop the color, 50 µl of each reaction mixture was transferred to a separate glass tube; after which 100 µl of a solution containing $NaIO_4$ (25 mM) and H_2SO_4 (0.125 N) was added followed by incubation for 10 min at room temperature (~21°C). The addition of 200 µl of a solution containing $NaAsO_2$ (2% w/v) and HCl (0.5 N) neutralized the excess $NaIO_4$. After the disappearance of the color, a solution containing 500 µl of 0.36% thiobarbituric acid, pH 9.0, was added and the reaction was incubated at 95°C for 10 min. These mixtures were transferred to a 96 well flat bottom plate and the absorbances were read at 549 nm. Data were fit using the nonlinear least-squares regression function of GraphPad Prism 5 software.

D-Glucose-6-phosphate isomerase assays

To determine the kinetic parameters for the isomerization of F6P to G6P catalyzed by Q5LIW1, a coupled assay was used, in which the G6P produced by Q5LIW1 was oxidized to 6-phosphoglucono-δ-lactone by glucose-6-phosphate dehydrogenase and the cofactor $NADP^+$. The formation of NADPH was measured at 340 nm, which allowed for reaction progress to be monitored (17). Individual reactions were performed in a quartz cuvette, in triplicate. The coupling enzyme (*E. coli* D-glucose-6-phosphate dehydrogenase, gene name *zwf*) was diluted into 2x reaction buffer (200 mM Bis-tris propane, pH 8.5, 2 mM EDTA), to a final coupling enzyme concentration of 1 µM. $NADP^+$ (0.16 mM final concentration) and F6P (0.01-3 mM final

concentration) were subsequently added to the reaction mixture, which was allowed to incubate for 1 min at room temperature (~21°C). The reaction was initiated by the addition of Q5LIW1 (final concentration 100 nM) and the absorbance, at 340 nm, was monitored, at 10 s intervals for 5 min, using a Hewlett-Packard 8453 diode array spectrophotometer. Initial reaction rates (nM/s) were determined by analysis of progress curves via linear regression. Kinetic parameters were obtained by fitting of the rate versus substrate concentration data, from assays performed in triplicate, to the Michaelis-Menten equation using nonlinear least-squares regression in GraphPad Prism 5.

Generation of CMP-Kdo

CMP-Kdo was generated using *E. coli* CMP-Kdo synthetase (KdsB) (18). In a typical reaction CTP (0.55 mM) and Kdo (1.1 mM), were added to 4x reaction buffer (400 mM Bis-tris propane pH 8.5, 5 mM $MgCl_2$), and the reaction was initiated by addition of KdsB to a final concentration of 13.2 µg/ml. The reaction was allowed to incubate for 10 min at room temperature before being quenched with EDTA (45 mM final concentration).

The Eikonogen assay (19,20), which allows for the direct correlation of the concentration of CMP-Kdo with the amount of inorganic pyrophosphate produced, was used to determine the concentration of inorganic pyrophosphate produced. In this assay, the CMP-Kdo solution (50 µl) treated with 2.5% ammonium molybdate (50 µl), of 0.5 M β-mercaptoethanol (50µl) and Eikonogen reagent (20 µl, 0.125 g sodium sulfite, 7.325 g sodium metabisulfite, and 0.125 g 1-amino-2-naphthol-4-sulfonic acid in 50 ml hot deionized water. The mixture was incubated at room temperature (~21°C) for 30 min and the absorbance was read at 540 nm. A malachite green assay, as previously described, was used to correct for inorganic phosphate concentrations (20).

Absorbances were compared to a standard curve prepared using a serial dilution of a standard aqueous solution of $Na_2P_2O_7$. Data was analyzed by linear regression in Microsoft Excel.

Inhibition kinetics with CMP-Kdo

To test if recombinant Q5LIW1 was inhibited by CMP-Kdo, a coupled assay in which the Ru5P produced by Q5LIW1-catalyzed isomerization of A5P was reduced to ribitol using an excess of NADPH and the enzyme Bcs1 (CDP-ribitol synthase) was performed. Reactions containing A5P (0.25 mM or 0.50 mM), Bis-tris propane buffer (100 mM), pH 8.5, EDTA (1 mM), NADPH (0.16 mM), and Bcs1 (96 µg/ml) (21) were incubated separately from mixtures containing Q5LIW1 (100 nM) and CMP-Kdo (final assay concentrations ranging from 0-68.51 µM) for 3 min at 37°C. The enzyme/inhibitor mixture was added to the reaction, mixed for 20 s, and incubated at 37°C for 10 min while monitoring the absorbance at 340 nm every 24 s using a SpectraMax M5 microplate reader (Molecular Devices). Absorbance values were plotted versus time in Microsoft Excel, and fit using linear regression. The slopes were converted to reaction rates in nM/s. Rates were plotted as a function of inhibitor concentration in GraphPad Prism 5 and fit using the nonlinear least squares technique, to a model of competitive inhibition based upon the Michaelis-Menten equation. Rates were converted to percent response of maximum and the Log(response/(1-response) was plotted vs. Log[CMP-Kdo] in Microsoft Excel to generate a Hill Plot (Fig. 3.7B).

Equilibrium constant (K_{eq}) determination

Solutions containing 100 mM Bis-tris propane buffer, pH 8.5, 1 mM EDTA, 10% D_2O, 500 nM Q5LIW1 or EcKdsDΔ2CBS, and a 5 mM concentration of F6P, G6P, A5P, or Ru5P were incubated at room temperature (~21°C) for 72 h, sufficient to reach equilibrium. These solutions

were analyzed by ^{31}P NMR using a Varian 400 multinuclear NMR spectrometer. A solution of 0.05 N phosphoric acid standard was sealed within a capillary tube and set to a value of 0 ppm. Spectra were acquired using 64 scans with a 10 s relaxation delay between scans. Preliminary studies with longer relaxation times did not show a change in peak integrations, which supports the assumption that the chosen delay time was greater than three times the T1 relaxation parameter for G6P, F6P, A5P, and Ru5P.

Effect of divalent metal ions on the API activity of Q5LIW1

To determine the effect of divalent metal ions on the API activity of Q5LIW1, samples of Q5LIW1 were diluted in buffer containing 108.2 mM Bis-tris propane, pH 8.5, and 10.82 µM EDTA or divalent metal salt, and incubated on ice for 30 min. The API activity was assayed using the Aminoff assay (15). Final concentrations in each reaction were: 100 nM Q5LIW1, 100 mM Bis-tris propane, pH 8.5, 0.15 mg/ml KdsA, 10 mM PEP, and 5 mM Ru5P. Reactions were incubated for 3 min at 37°C before being quenched with an equal volume of 10% (w/v) trichloroacetic acid, and the color developed as described above in the A5P isomerization assay. Metal salts tested included $BaCl_2 \cdot 2H_2O$, $MnCl_2 \cdot 4H_2O$, $ZnCl_2$, $NiCl_2 \cdot 6H_2O$, $CoSO_4 \cdot 7H_2O$, $CuSO_4$, $FeSO_4 \cdot 7H_2O$, $CdCl_2$, $MgCl_2$, $CaCl_2 \cdot 2H_2O$, and $HgCl_2$.

Results

Expression and characterization of Q5LIW1 and EcKdsDΔ2CBS.

(i) Q5LIW1 is one of a small group of single SIS-domain proteins.

A BLASTP search using YP_209877.1 as the query sequence revealed putative SIS-domain proteins in Gram-negative bacteria of the genera *Bacteroides*, *Parabacteroides*, *Dysgonomonas*, and *Porphyromonas*. In each case, the YP_209877.1 homolog was the only putative API

discovered within the genome; no CBS domain-containing APIs were identified. The *B. fragilis* NCTC 9343 genome contains orthologs of all the other enzymes in the *E. coli* K-12 Kdo biosynthesis pathway, including KdsA (locus YP_210629; E value, $7e^{-75}$), KdsB (locus YP_211897; E value, $2e^{-38}$), KdsC (locus YP_212734; E value, $4e^{-31}$), and WaaA (locus YP_213607; E value, $1e^{-42}$). Taken together, these observations suggest that Q5LIW1 and its homologs constitute a unique group of SIS-domain APIs responsible for supporting Kdo biosynthesis in these Gram-negative organisms.

(ii) Recombinant Q5LIW1 is a tetramer.

Locus YP_209877.1 was cloned by PCR from *B. fragilis* NCTC 9343 genomic DNA obtained from the ATCC and inserted into the plasmid pT7-7 (9). Overexpression in *E. coli* Rosetta (DE3) pLysS cells produced sufficient quantities of protein that were purified using ammonium sulfate fractionation and chromatography on a phenyl-Sepharose column. The purified protein migrated at approximately 21 kDa on an SDS-PAGE gel; its subunit molecular mass was determined to be 21,809.99 Da (calculated, 21,810.2) by electrospray mass spectroscopy. The quaternary

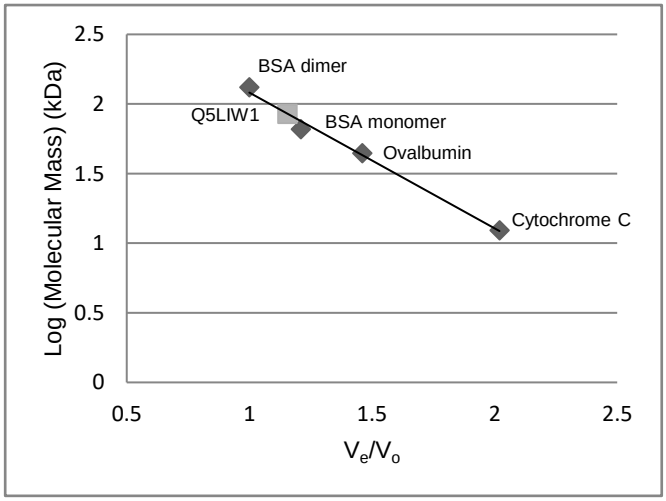

Fig. 3.1: Gel filtration analysis of Q5LIW1. Q5LIW1 shown as grey box. Standard curve from native molecular mass determination by gel filtration chromatography (V_e is the elution volume and V_o is the void volume).

structure of the protein was probed using gel filtration chromatography (Fig. 3.1). Native Q5LIW1 eluted with an apparent molecular mass of 85,871.3 Da, approximately 3.94 times the subunit mass. This result is almost identical to the results obtained with the APIs of *E. coli*, which behave as tetramers in solution (5, 6, 8). Metal content analysis showed that the purified protein, as isolated, did not contain any metals. Inhibition tests with added divalent metals showed that only Cu^{2+} had an inhibitory effect on the isomerization of A5P to Ru5P, amounting to a roughly 4-fold decrease in activity (Fig. 3.2).

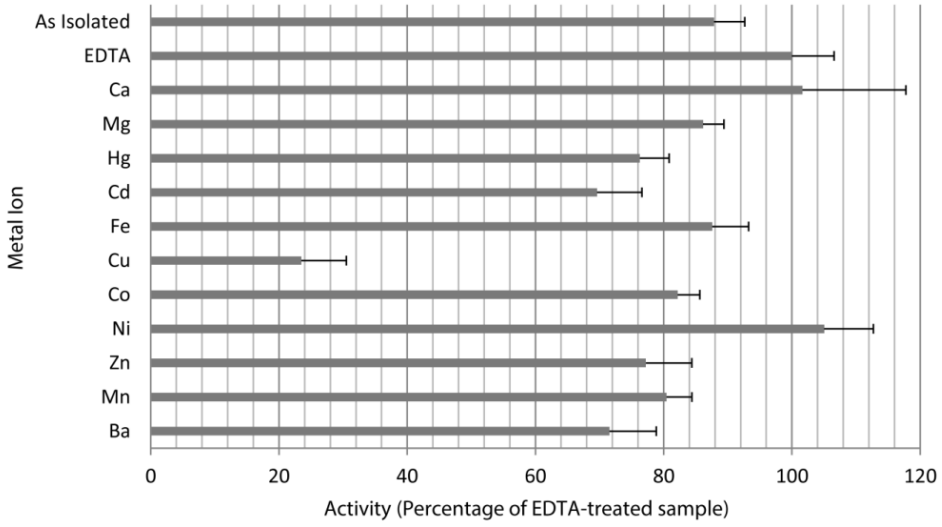

Fig. 3.2: Effect of divalent metal ions on the activity of Q5LIW1.

Based on a sequence comparison, *E. coli kdsD*, in the expression vector pT7-7, was truncated to produce a single SIS domain by site-directed mutagenesis to insert a stop codon after 213 amino acids (Asp214*). The resulting pT7-7-EcKdsDΔ2CBS was moved to *E. coli* BL21(DE3), expressed, and purified.

Enzymatic properties of Q5LIW1 and EcKdsD2CBS.

(i) Recombinant Q5LIW1 is an API.

The substrate specificity of Q5LIW1 was investigated by testing its ability to catalyze the isomerization of a series of aldose phosphates. These experiments were performed at pH 6.5 based on the pH optimum observed for the c3406 protein (8). Q5LIW1 converted two of the substrates, A5P and G6P, to the corresponding ketoses. The pH-rate profile was investigated

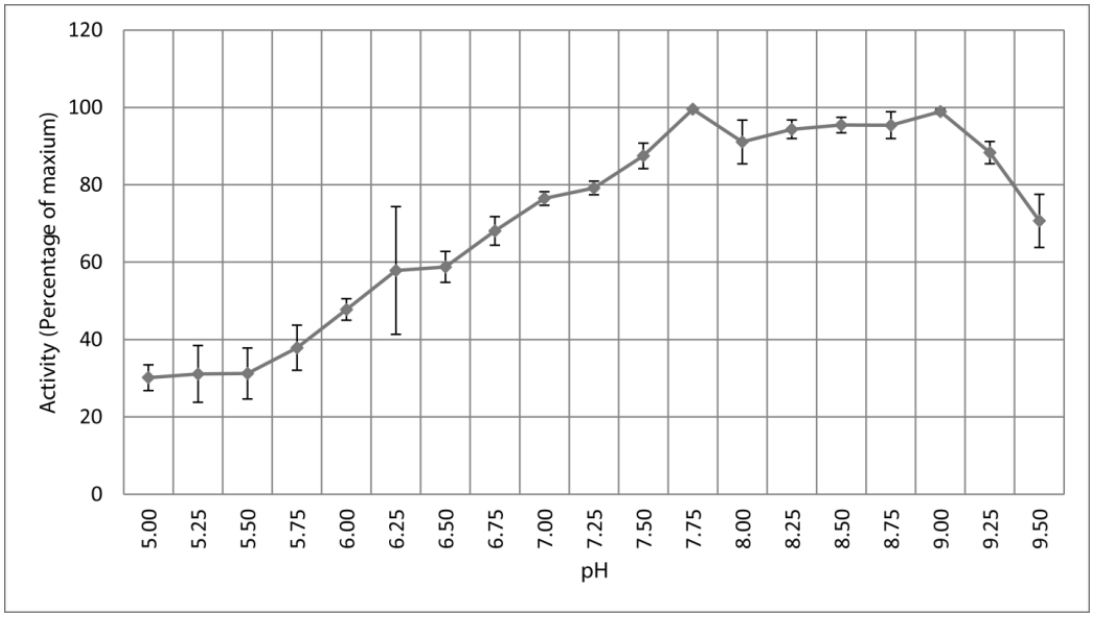

Fig. 3.3: pH rate profile of Q5LIW1.

using A5P as the substrate (Fig. 3.3). It was determined that Q5LIW1 has a very broad pH profile with an optimum ranging from 7.75 to 9.0.

Table 3.2 Kinetic constants for catalysis by various APIs.

Protein	K_m (A5P, mM)	k_{cat} / K_m (A5P, M^{-1} s^{-1})	k_{cat} (A5P to Ru5P, s^{-1})	k_{cat} (Ru5P to A5P, s^{-1})	K_m (Ru5P, mM)	k_{cat} / K_m (Ru5P, M^{-1} s^{-1})	K_{eq}	Optimum pH	Subunit mass (Da)
KdsD[Ref.5]	0.61±0.06	$2.6x10^5$	157±4	255±16	0.35±0.08	$7.3x10^5$	0.50±0.06	8.4	35,084
GutQ[Ref.6]	1.2±0.1	$1.8x10^5$	218±4	242±11	0.64±0.08	$3.8x10^5$	0.47	8.25	33,909
KpsF[Ref.7]	0.57±0.04	$2.6x10^4$	15±1	19±2	0.30±0.03	$6.3x10^4$	0.48±0.02	7.75	35,447
C3406[Ref.8]	1.92±0.05	$8.8x10^3$	16.8±0.2	10.5±0.08	0.70±0.12	$1.5x10^4$	0.52	6.6	20,880
Q5LIW1	34.54±12.4	$7.3 x10^3$	253±150	18.5±1.7	0.485±0.18	$3.8x10^4$	0.48	7.75-9	21,810
KdsD Δ2CBS	0.62±0.20	$4.7x10^4$	29±1	6.29±0.76	0.52±0.30	$1.2x10^4$	0.49	ND	22,585

(ii) Enzyme kinetics.

Kinetic parameters were determined for the interconversion of A5P and Ru5P at pH 8.5. A comparison of the kinetic constants obtained for Q5LIW1 and EcKdsDΔ2CBS with those previously obtained for the APIs of *E. coli* is shown in Table 3.2.

Kinetic constants for the conversion of F6P to G6P (phosphoglucose isomerase catalysis) were also determined for Q5LIW1: K_m of G6P, 0.92 ± 0.378 mM; k_{cat}/K_m of G6P, 8.14×10^3 M^{-1} s^{-1} ; k_{cat} of G6P to F6P, 7.486 ± 0.79 s^{-1} ; k_{cat} of F6P to G6P, 0.753 ±0.0386 s^{-1}; K_m of F6P, 0.192 ± 0.0395 mM; k_{cat}/K_m of F6P, 3.9×10^3 M^{-1} s^{-1}; K_{eq}, 0.28; optimum pH, 7.75 to 9; subunit mass, 21,810 Da.

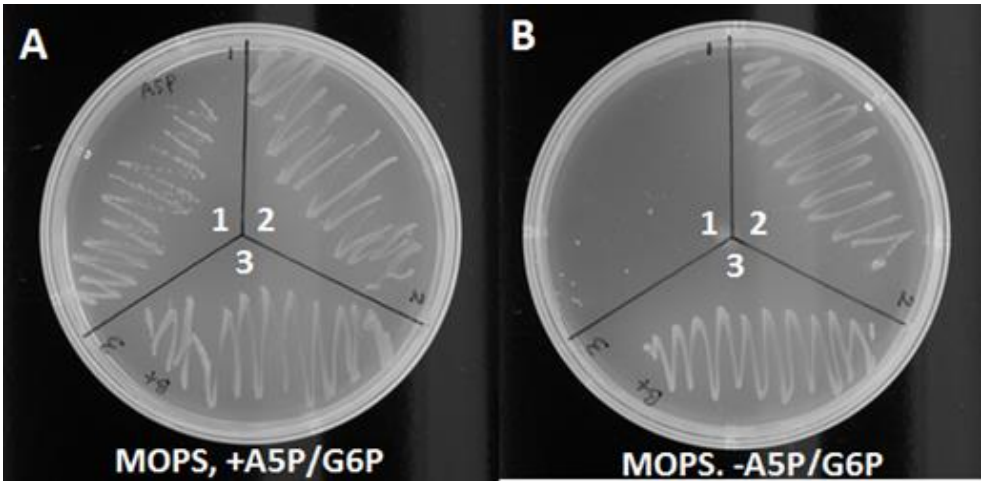

Fig. 3.4: Complementation of an A5P auxotroph on agar plates. (**A**) Agar plate containing MOPS medium supplemented with 15 µM A5P, 10 µM G6P, 0.1 mg/ml ampicillin, and 0.05 mg/ml kanamycin. (**B**) Agar plate containing MOPS medium supplemented with 0.1 mg/ml ampicillin, and 0.05 mg/ml kanamycin. In both panels, the wedges were streaked with: *E. coli* TCM15 harboring pT7-7 (vector control, wedge 1), *E. coli* TCM15 harboring pT7-7-c3406 (c3406 control, wedge 2), and *E. coli* TCM15 harboring pT7-7-Q5LIW1 (Q5LIW1, wedge 3).

Q5LIW1 and EcKdsDΔ2CBS are able to complement *E. coli* strain TCM15.

(i)Q5LIW1 complements the API defect in *E. coli* strain TCM15.

E. coli TCM15 (6) is a derivative of BW30270 in which both *kdsD* and *gutQ* have been deleted. To support the growth of TCM15, minimal medium must be supplemented with A5P (to support LPS biosynthesis) and G6P (to induce a transport system, uhp, that internalizes A5P) (22). This A5P/G6P auxotrophy can be circumvented by expression of an active API from a plasmid within the cell. As a test of the ability of Q5LIW1 to complement an API deficiency within bacterial cells, *E. coli* TCM15 was transformed with the plasmid pT7-7-YP_209877.1, pT7-7- c3406 (positive control), or pT7-7 (vector control). In this system, the plasmid inserts are expressed from a leaky T7 promoter. Equal numbers of cells were plated on morpholinepropanesulfonic acid (MOPS) minimal medium containing glycerol and either containing or lacking A5P/G6P (Fig. 3.4). Q5LIW1

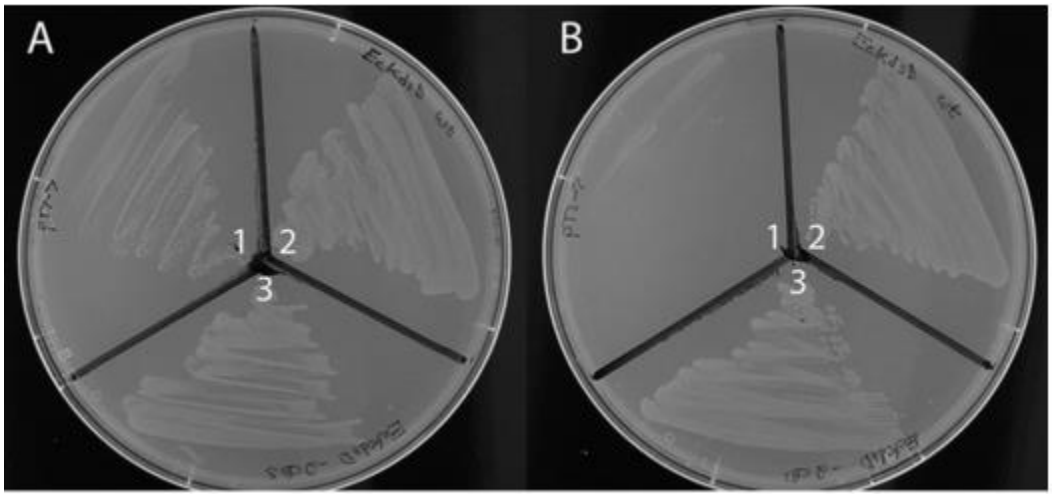

Fig. 3.5: Complementation of an A5P auxotroph on LB/agar plates. (**A**) Agar plate containing LB medium supplemented with 15 µM A5P, 10 µM G6P, 0.1 mg/ml ampicillin, and 0.05 mg/ml kanamycin. (**B**) Agar plate containing LB medium supplemented with 0.1 mg/ml ampicillin, and 0.05 mg/ml kanamycin. In both panels, the wedges were streaked with: *E. coli* TCM15 harboring pT7-7 (vector control, wedge 1), *E. coli* TCM15 harboring pT7-7-KdsD (KdsD wild type control, wedge 2), and *E. coli* TCM15 harboring pT7-7-KdsDΔ2CBS (KdsD truncated to contain only the SIS domain, wedge 3).

is able to complement the lack of A5P/G6P in the medium and is therefore an active API in *E. coli* TCM15.

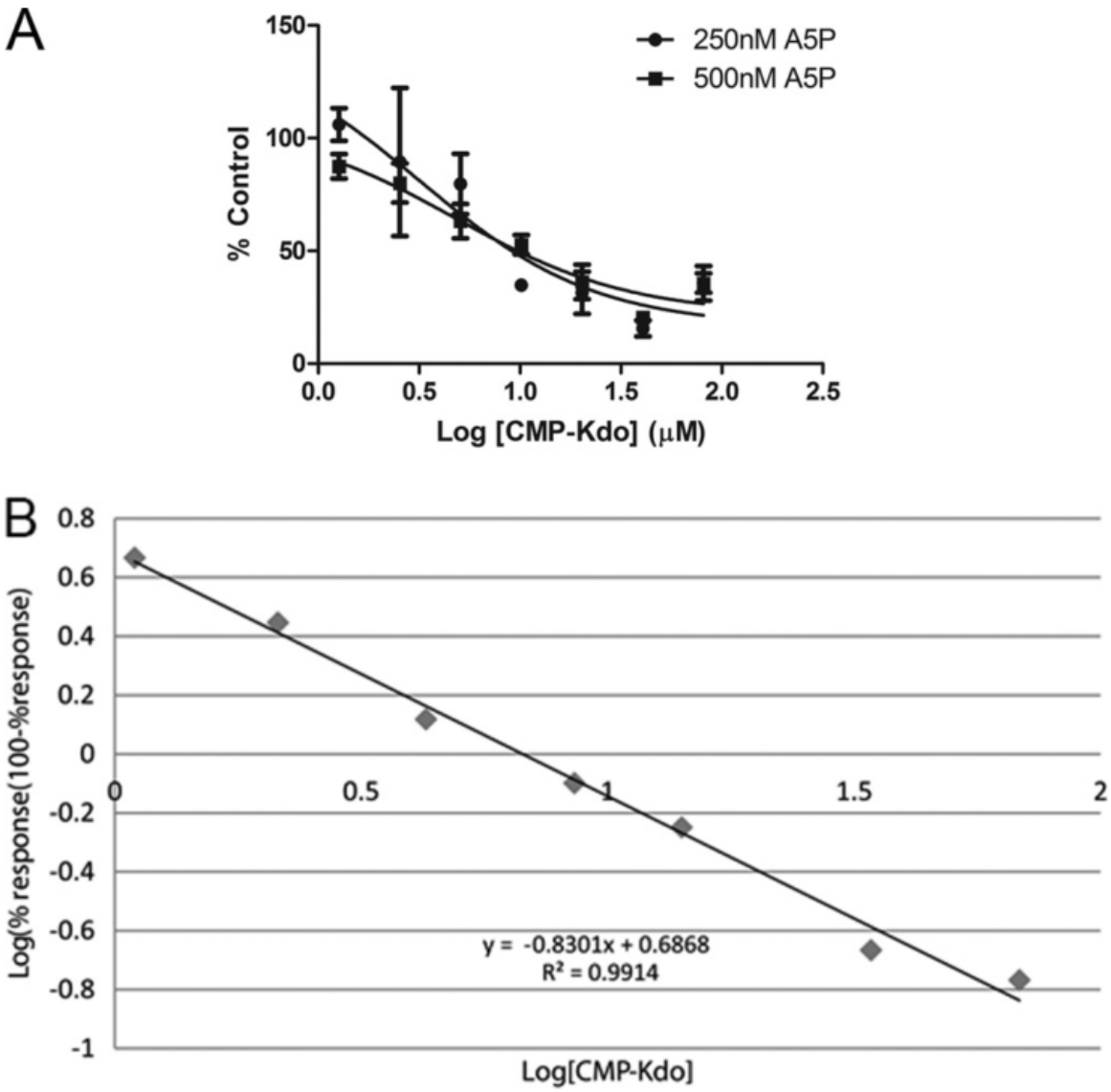

Fig. 3.6: Effect of CMP-Kdo on Q5LIW1 API activity. (A) Nonlinear regression fit of the effect of CMP-Kdo on Q5LIW1 API activity. Rates in nM/s were measured at CMP-Kdo concentrations ranging from 0 to 68.51 µM and various A5P concentrations (0.25mM and 0.5 mM). (B) Hill plot of CMP-Kdo binding to Q5LIW1, showing a 1:1 binding.

(ii) EcKdsDΔ2CBS complements the API defect in *E. coli* strain TCM15.

As a test of the ability of EcKdsDΔ2CBS to complement an API deficiency within bacterial cells, *E. coli* TCM15 was transformed with pT7-7 (vector control), pT7-yrbH (positive control, wild-type *kdsD*), or pT7-7-EcKdsDΔ2CBS. Equal numbers of cells were plated on LB medium either containing or lacking A5P/G6P (Fig. 3.5). EcKdsDΔ2CBS is able to complement the lack of A5P/G6P in the medium and therefore displays API activity in *E. coli* TCM15. CMP-Kdo inhibits the API activity of Q5LIW1. CMP-Kdo, the final product of the Kdo biosynthetic pathway, is used by the enzyme Kdo transferase (WaaA) to transfer Kdo to lipid IV_A. We tested the hypothesis that CMP-Kdo is a feedback inhibitor of Q5LIW1. CMP-Kdo, which is not commercially available and must be made in the lab, autohydrolyzes in aqueous solution (25°C, pH 7.5) with a half-life of approximately 34 min (23). This limitation was circumvented by performing API assays using CMP-Kdo that was generated in situ using CTP, Kdo, and CMP-Kdo synthetase (KdsB) from *E. coli*. This process generated high micromolar concentrations of CMP-Kdo, as determined using the Eikonogen assay (20). The addition of CMP-Kdo-containing mixtures to Q5LIW1 enzymatic assays caused a substantial decrease in the activity of Q5LIW1 (Ki of 1.91 ± 0.48 μM) (Fig. 3.6A). Addition of mixtures that lacked KdsB and were, therefore, devoid of CMP-Kdo did not inhibit the API activity of Q5LIW1. A Hill plot of these data shows a 1:1 binding ratio of Q5LIW1 to CMP-Kdo, suggesting that CMP-Kdo binds noncooperatively; however, we cannot rule out the possibility of negatively cooperative binding (Fig. 3.6B).

Discussion

The data presented in this report strongly support the hypothesis that Q5LIW1 is a physiologically relevant API that is responsible for the generation of A5P necessary to support

LPS biosynthesis in *B. fragilis*. A BLASTP search of the genomes of other Gram-negative obligate anaerobes, using Q5LIW1 as the query sequence, yielded several hits of species which contain putative SIS-domain proteins but do not contain a putative API with CBS domains. This suggests that single SIS-domain APIs may be utilized as the sole API in several Gram-negative obligate anaerobes. The slow growth of these obligate anaerobes should diminish the rate of Kdo biosynthesis required to sustain viability, compared with facultative anaerobes like *E. coli* , and thus diminish the rate at which A5P must be produced. Therefore, it is reasonable to conclude that an API such as Q5LIW1, which has a somewhat lower k_{cat}/K_m than that of *E. coli* KdsD, would be able to support cell viability. Experiments with *E. coli* TCM15 showed that Q5LIW1 could complement an API-deficient strain. This demonstrates that Q5LIW1 can serve as the sole source of A5P to support LPS biosynthesis in *E. coli* TCM15. The substrate specificity of Q5LIW1 was probed utilizing a series of aldose phosphates. Unlike previously characterized APIs, which are specific to the isomerization of Ru5P and A5P, Q5LIW1 was able to catalyze the isomerization of A5P to Ru5P as well as that of G6P to F6P. However, the phosphoglucose isomerase activity of Q5LIW1 is very weak and likely not physiologically relevant. The pH optimum for the isomerization of A5P to Ru5P by Q5LIW1 was found to include a broad range from 7.75 to 9.0. This pH optimum is similar to that of the two-domain APIs of *E. coli* (KdsD, GutQ, and KpsF) (5, 6, 7). Conversely, c3406, a SIS-domain API, has a sharp pH rate profile with an optimum pH of 6.6 (8). The kinetic parameters, for the isomerization of A5P as well as G6P to the corresponding ketoses, were determined. The kinetic profile of Q5LIW1 is very similar to that of c3406 (Table 3.2) for the isomerization of A5P, further suggesting that SIS-domain APIs may be able to serve as the sole A5P source to support LPS biosynthesis. These results led to the hypothesis that a

truncation of the two CBS domains of *E. coli* KdsD would result in a protein that would maintain API activity and potentially complement *E. coli* TCM15. EcKdsDΔ2CBS (a 213- amino-acid protein including only the SIS domain of *E. coli* KdsD) was cloned, overexpressed, and purified. Complementation studies with EcKdsDΔ2CBS showed that the truncated gene could complement the API defect in *E. coli* TCM15. The kinetic parameters of the truncated enzyme were also determined; EcKdsDΔ2CBS demonstrated a 5-fold decrease in catalytic efficiency in the A5P-to-Ru5P direction compared to wild-type KdsD (Table 3.2). In the Ru5P-to-A5P direction, the catalytic efficiency of EcKdsDΔ2CBS resembles the single-domain APIs of Q5LIW1 and c3406 more than it does the CBS domain-containing APIs of *E. coli*. The finding that single-domain APIs may serve as the sole API in several bacterial species led us to question how this unique group of APIs might be regulated. Because APIs catalyze the first step in Kdo biosynthesis, it is logical that the final product of the Kdo biosynthetic pathway, CMP-Kdo, may serve as a feedback inhibitor to regulate the pathway. In this paper, we demonstrated that CMP-Kdo inhibits Q5LIW1 with a K_i of 1.91 μM. Our inhibition experiments confirm the biological significance of CMP-Kdo within the active site of Q5LIW1 in the PDB (PDB ID 3ETN). These results give insight into a potential regulation mechanism of single-domain APIs. However, it is unclear if this regulation mechanism also applies *in vivo*. Based on the fact that Q5LIW1 may be regulated by CMP-Kdo feedback inhibition, we speculate that other single SIS-domain APIs, such as c3406, can also be regulated by CMP-Kdo. We further speculate that two-domain (CBS domain-containing) APIs may be regulated by CMP-Kdo, but it is unclear what role the CBS domain would play in this process. The discovery of this potential regulation mechanism of the single-domain API, Q5LIW1, has provided

a first glimpse of insight into understanding the regulation of APIs in the biosynthesis of Kdo and could serve as a model of structure-function studies of APIs.

Acknowledgements and Chapter Contributions

This work was supported by National Institutes of Health grant AI- 061531 (to R.W.W.). We thank Andrew Pratt for his helpful discussions and review of the manuscript, Kyle Heslip for the use of the Dotson lab plate reader, Victoria Assimon for the original cloning of the BF0137 gene, Ted Houston of the University of Michigan Department of Geology's W. M. Keck Elemental Geochemistry Laboratory for performing the metals analysis, and Jim Windak and Paul Lennon of the University of Michigan Department of Chemistry Mass Spectroscopy facility for performing the liquid chromatography-mass spectrometry analysis of recombinant Q5LIW1.

The purification scheme for Q5LIW1 was designed by Wang, P.F., all other experiments were completed by Cech, D.L., along with writing of the manuscript. Review and edits were completed with assistance from Holler, T.P., and Woodard, R.W.

References

1. Horecker, B. L. (2002) The Pentose Phosphate Pathway. *Journal of Biological Chemistry* **277**:47965–47971. doi:10.1074/jbc.X200007200

2. Meredith, T. C.; Aggarwal, P.; Mamat, U.; Lindner, B.; Woodard, R. W. (2006) Redefining the Requisite Lipopolysaccharide Structure in *Escherichia coli*. *ACS Chemical Biology* **1**:33–42. doi:10.1021/cb0500015

3. Royo, J.; Gímez, E.; Hueros, G. (2000) CMP-Kdo synthetase: a plant gene borrowed from gram-negative *eubacteria*. *Trends Genet.* **16**:432–433. doi:10.1016/S0168-9525(00)02102-8

4. Cipolla, L.; Polissi, A.; Airoldi, C.; Gabrielli, L.; Merlo, S.; Nicotra, F. (2011) New Targets for Antibacterial Design: Kdo Biosynthesis and LPS Machinery Transport to the Cell Surface. *Current Medicinal Chemistry* **18**:830–852. doi:10.2174/092986711794927676

5. Meredith, T. C., Woodard, R. W. (2003) *Escherichia coli* YrbH Is a D-Arabinose 5-Phosphate Isomerase. *Journal of Biological Chemistry* **278**:32771–32777. doi:10.1074/jbc.M303661200

6. Meredith, T. C.; Woodard, R. W. (2005) Identification of GutQ from *Escherichia coli* as a D-Arabinose 5-Phosphate Isomerase. *Journal of Bacteriology* **187**:6936–6942. doi:10.1128/JB.187.20.6936-6942.2005

7. Meredith, T. C.; Woodard, R. W. (2006) Characterization of *Escherichia coli* D-arabinose 5-phosphate isomerase encoded by kpsF: implications for group 2 capsule biosynthesis. *Biochemical Journal* **395**:427. doi:10.1042/BJ20051828

8. Mosberg, J. A.; Yep, A.; Meredith, T. C.; Smith, S.; Wang, P.-F.; Holler, T. P.; Mobley, H. L. T.; Woodard, R. W. (2011) A Unique Arabinose 5-Phosphate Isomerase Found within a Genomic Island Associated with the Uropathogenicity of *Escherichia coli* CFT073. *Journal of Bacteriology* **193**:2981–2988. doi:10.1128/JB.00033-11

9. Ignoul, S.; Eggermont, J. (2005) CBS domains: structure, function, and pathology in human proteins. *Am. J. Physiol., Cell Physiol.* **289**:C1369–1378. doi:10.1152/ajpcell.00282.2005

10. Gourlay, L. J.; Sommaruga, S.; Nardini, M.; Sperandeo, P.; Dehò, G.; Polissi, A.; Bolognesi, M. (2010) Probing the active site of the sugar isomerase domain from *E. coli* arabinose-5-phosphate isomerase via X-ray crystallography. *Protein Science* **19**:2430–2439. doi:10.1002/pro.525

11. Kumada, H.; Haishima, Y.; Kondo, S.; Umemoto, T.; Hisatsune, K. (1993) Occurrence of 2-keto-3-deoxyoctonate (KDO) and KDO phosphate in lipopolysaccharides of *Bacteriodes* species. *Current Microbiology* **26**:239–244. doi:10.1007/BF01577383

12. Datsenko, K. A.; Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences* **97**:6640–6645. doi:10.1073/pnas.120163297

13. Sambrook, J.; Fritsch, E. F.; Maniatis, T. (1989) Molecular cloning : a laboratory manual. vol. 1; Cold Spring Harbor Laboratory Press: Cold Spring harbor, N.Y.

14. Tabor, S. A Bacteriophage T7 RNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes. (1985) *Proceedings of the National Academy of Sciences* **82**:1074–1078. doi:10.1073/pnas.82.4.1074

15. Wu, J.; Patel, M. A.; Sundaram, A. K.; Woodard, R. W. (2004) Functional and biochemical characterization of a recombinant *Arabidopsis thaliana* 3-deoxy-D-manno-octulosonate 8-phosphate synthase. *Biochem. J.* **381**:185–193. doi:10.1042/BJ20040207

16. Sheflyan, G. Y.; Sundaram, A. K.; Taylor, W. P.; Woodard, R. W. (2000) Substrate Ambiguity of 3-Deoxy-D-manno-Octulosonate 8-Phosphate Synthase from *Neisseria gonorrhoeae* Revisited. *Journal of Bacteriology* **182**:5005–5008. doi:10.1128/JB.182.17.5005-5008.2000

17. Sandoval, J. M.; Arenas, F. A.; Vásquez, C. C. (2011) Glucose-6-Phosphate Dehydrogenase Protects *Escherichia coli* from Tellurite-Mediated Oxidative Stress. *PLoS ONE* **6**:e25573. doi:10.1371/journal.pone.0025573

18. Wu, J. Woodard, R. W. (2003) *Escherichia coli* YrbI Is 3-Deoxy-D-manno-octulosonate 8-Phosphate Phosphatase. *Journal of Biological Chemistry* **278**:18117–18123. doi:10.1074/jbc.M301983200

19. Flynn, R. M.; Jones, M. E.; Lipmann, F. (1954) A Colorimetric Determination of Inorganic Pyrophosphate. *J. Biol. Chem.* **211**:791–796.

20. Yi, L.; Velasquez, M. S.; Holler, T. P.; Woodard, R. W. (2011) A simple assay for 3-deoxy-d-manno-octulosonate cytidylyltransferase and its use as a pathway screen. *Analytical Biochemistry* **416**:152–158.

21. Pereira, M. P.; Brown, E. D. (2004) Bifunctional catalysis by CDP-ribitol synthase: convergent recruitment of reductase and cytidylyltransferase activities in *Haemophilus influenzae* and *Staphylococcus aureus*. *Biochemistry* **43**:11802–11812.

22. Eidels, L.; Rick, P. D.; Stimler, N. P.; Osborn, M. J. (1974) Transport of D-Arabinose-5-Phosphate and D-Sedoheptulose-7-Phosphate by the Hexose Phosphate Transport System of *Salmonella typhimurium*. *Journal of Bacteriology* **119**:138–143.

23. Lin, C. H.; Murray, B. W.; Ollmann, I. R.; Wong, C. H. (1997) Why is CMP-ketodeoxyoctonate highly unstable? *Biochemistry* **36**:780–785.

Chapter IV

Identification of a D-Arabinose-5-Phosphate Isomerase in the Gram-positive *Clostridium tetani*

Summary

D-Arabinose-5-phosphate isomerases (APIs) catalyze the interconversion of D-ribulose-5-phosphate and D-arabinose-5-phosphate. Various Gram-negative bacteria, such as the uropathogenic *E. coli* strain CFT073, contain multiple API paralogs (KdsD, GutQ, KpsF and c3406) that have been assigned various cellular functions. The D-arabinose-5-phosphate formed by these enzymes seems to play important roles in the biosynthesis of lipopolysaccharide and group 2 K-antigen capsules, as well as in the regulation of the cellular D-glucitol uptake and uropathogenic infectivity/virulence. The genome of a Gram-positive pathogenic bacterium, *Clostridium tetani*, contains a gene encoding a putative API, CtAPI, even though *C. tetani* lacks both LPS and capsid biosynthetic genes. To better understand the physiological role of A5P in this Gram-positive organism, recombinant CtAPI was purified and characterized. CtAPI displays biochemical characteristics similar to those of APIs from Gram-negative organisms and complements the API-deficiency of *E. coli* API knockout strain. Thus, CtAPI represents the first D-arabinose-5-phosphate isomerases to be identified and characterized from a Gram-positive bacterium.

Introduction

In Gram-negative microorganisms, D-arabinose-5-phosphate isomerase (API) catalyzes the interconversion of D-ribulose-5-phosphate (Ru5P), a product of the pentose phosphate pathway, and D-arabinose-5-phosphate. All APIs contain a catalytic sugar isomerase (SIS) domain; most also contain a C-terminal tandem cystathionine β-synthetase (CBS) domain, which is thought to play a regulatory role.

The cellular functions of the API paralogues have been assigned based upon their activity and the genomic context of the genes encoding them. The two-domain APIs of *E. coli* have been classified as L-API (involved in lipopolysaccharide [LPS] biosynthesis), G-API (part of the glucitol phosphotransferase operon), and K-API (involved in K-antigen biosynthesis). L-APIs, such as KdsD from *E. coli,* are the first enzyme in the pathway responsible for the biosynthesis of 3-deoxy-D-*manno*-octulosonic acid (Kdo), an essential component of both the lipopolysaccharide (LPS) of Gram-negative bacteria and the rhamnogalacturonan-II of certain plants (1-3). While still not fully understood, G-APIs, such as GutQ, appear to function in the biosynthesis of A5P to modulate expression levels of the *gut* operon (4). The gene encoding GutQ (*gut*Q) is the last open reading frame of the glucitol (sorbitol) operon, which encodes the proteins that transport D-glucitol across the inner bacterial membrane as D-glucitol-6-phosphate. This material is eventually metabolized to D-fructose-6-phosphate. Finally, the K-APIs, such as KpsF from *E. coli* CFT073, are considered responsible for the biosynthesis of the Kdo required for capsule biogenesis. This type of API is found in bacteria with K-antigens, in which a single Kdo molecule serves as a link between the K-antigen polysaccharide and its lipid α-glycerophosphate membrane anchor or multiple Kdo molecules serve as a repeating component of the capsular polysaccharide (5). Single-domain APIs

have been identified and characterized from pathogenic bacteria such as *Bacteroides fragilis* and *E. coli* CFT073, a uropathogenic strain of *E. coli*. Because the single-domain API Q5LIW1 is apparently the only API encoded by the *B. fragilis* genome, it is thought to produce the Kdo required for LPS biosynthesis. The gene encoding the single-domain API present in *E. coli* CFT073, c3406, is located on a genomic island that has been shown to be important for virulence in mice. However, the cellular role of c3406 remains unclear.

In this study, we identified and characterized a putative API from the Gram-positive organism, *Clostridium tetani*, the etiologic agent of tetanus.

Materials and Methods

Materials

An *E. coli* codon-optimized synthetic gene encoding the API from *C. tetani*, CTC00908, was purchased from GenScript (Piscataway, NJ, USA). Enzymes for subcloning were purchased from New England Biolabs (Ipswich, WI, USA). Plasmid DNA was purified using the Promega Wizard Miniprep kit (Madison, WI, USA). DNA sequencing was performed by the University of Michigan Biomedical Resources Core Facility.

Table 4.1 Strains, plasmids and primers used in this study.

Item	Description	Source
E. coli		
TOP10	F$^-$ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) F80*lac*ZΔM15 Δ*lacX*74 *recA1 araD*139 Δ(ara-leu)7697 *galU galK rpsL* (StrR) endA1 *nupG*	Invitrogen
TCM15	MG1655 *rph*$^+$ (*ΔkdsDΔgutQ*)	Ref. (4)
BL21 (DE3)	F$^-$ *ompT gal dcm lon hsdS*$_B$($r_B^- m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S)	Novagen
Plasmids		
pCR2.1-TOPO	Library plasmid	Invitrogen
pT7-7	Expression vector	Ref. (6)
pT7-7-c3406	*E. coli* CFT073 c3406 inserted into NdeI/BamHI of pT7-7, Amp^R	Ref. (7)
pT7-7-CtAPI	Codon optimized *C. tetani* WP_023437760.1 inserted into NdeI/BamHI of pT7-7, Amp^R	This study
pT7-7-yrbH	*E. coli kdsD* inserted into NdeI/BamHI of pT7-7, Amp^R	Ref. (8)
pET19b	T7 expression vector for N-terminal His-tag fusions, Amp^R	Novagen
pET19b-CtAPI	Codon optimized *C. tetani* CTC00908 inserted into NdeI/BamHI of pET19b, Amp^R	This Study

Bacterial strains, plasmids, primers, and growth media

The bacterial strains and plasmids used in this study are listed in Table 4.1. All strains were grown in LB medium (9). TCM15 is a derivative of BW30270 (*E. coli* K-12 MG1655 *rph*$^+$), in which the *kdsD* and *gutQ* genes have been disrupted using the phage λ Red recombinase system (4,10). TCM15 cultures were supplemented with G6P (10 μM) and A5P (15 μM). G6P is required for induction of the uhp transporter, which allows the uptake of A5P required for growth (11).

Cloning, expression, and purification of *C. tetani API*

The synthetic, *E. coli* codon-optimized *C. tetani* API gene was received with N-terminal *NdeI* and C-terminal *BamHI* restriction sites in pUC57. The *NdeI* to *BamHI* fragment containing the *C. tetani* API gene was sub-cloned into pET19b (to produce pET19b-CtetAPI for protein production) and pT7-7 (to produce pT7-ctetAPI for use in complementation experiments). The plasmid pET19b-CtetAPI was inserted into chemically competent *E. coli* BL21(DE3) cells. A fresh

transformant was grown at 37°C in LB medium supplemented with 100 mg/L ampicillin and 30 mg/L chloramphenicol while shaking at 250 rpm until the optical density at 600 nm reached 0.6. Recombinant protein expression was induced at 19°C with the addition of isopropyl-β-D-1-thiogalactopyranoside to a final concentration of 0.5 mM. After 16 h of incubation at 19°C and 250 rpm, the cells were harvested by centrifugation at 6,000 *x g* and 4°C for 10 min. The resulting cell pellet was resuspended in 50 mL of buffer A (20 mM Tris, pH 8.0, 300 mM NaCl) and lysed by sonication on ice using multiple 20 s bursts with 2 min pauses between bursts. The lysate was clarified by centrifugation (20,000 *x g*, 4°C, 45 min) and the supernatant was filtered through a Millipore 0.22 µm PVDF filter. The filtered solution was loaded onto a 5 mL GE HisTrap™ HP nickel column pre-equilibrated with buffer A, which was then washed with 20 column volumes of buffer A supplemented with 100 mM imidazole. The protein was eluted with buffer A supplemented with 500 mM imidazole. The eluted protein solution was buffer-exchanged into 20 mM Tris pH 8.0, 100 mM NaCl prior to concentration in an Amicon ultracentrifugal filter (10,000 molecular-weight cut off) and stored at 4°C.

Native oligomeric state determination

The native molecular mass of CtAPI was estimated by gel filtration chromatography on a HiPrep (26/60) Sephacryl S-100 column. Standards included Bovine serum albumin (BSA) dimer (132.4 kDa), BSA monomer (66.2 kDa), chicken egg white ovalbumin (44.3 kDa), and trypsin inhibitor (28 kDa). A plot of log (molecular mass) versus elution volume/void volume (V_e/V_o) was fitted by linear regression in Microsoft Excel. The molecular mass of CtAPI was determined experimentally by measuring its elution volume on the same column, in triplicate, and calculating the log(molecular mass) from the standard curve.

Substrate specificity

To assess the substrate specificity of CtAPI, a panel of naturally abundant carbohydrates was tested using the colorimetric cysteine-carbazole assay. The following potential substrates were tested: D-arabinose-5-phosphate, D-arabinose, D-glucose-6-phosphate, D-mannose-6-phosphate, D-glucoseamine-6-phosphate, D-glucose-1-phosphate, and D-ribose-5-phosphate. Final reaction conditions included 100 mM 2,2'-(Propane-1,3-diyldiimino)bis[2-(hydroxymethyl)propane-1,3-diol] (Bis-tris propane) pH 8.0, 1 mM EDTA, 100 mM CtAPI, and 1 mM potential substrate. After 10 min at 37°C, reaction mixtures (each substrate was assayed in triplicate) were quenched with an equal volume of 25 N H_2SO_4. Aliquots (90 µL) of the quenched reaction mixtures were mixed with a freshly prepared cysteine-carbazole assay mixture (10 µL of 0.12% ethanolic carbazole solution, 10 µL of 1.5% aqueous cysteine-HCl, and 230 µL of 25 N H_2SO_4) and incubated at 21°C for 3 h to develop the color (8,12). Enzyme substrates were identified by comparing the absorbance of reactions to the absorbance of enzyme-omitted controls at 540 nm.

Complementation of a *kdsD*/*gutQ* defect in *E. coli* TCM15

Electrocompetent TCM15 cells were prepared as previously described (3) and separately transformed with pT7-ctetAPI, pT7-c3406 (7), and empty pT7-7 vector. Cells were grown on LB/agar plates supplemented with 100 µg/mL ampicillin, 15 µM A5P, and 10 µM G6P. Aliquots of liquid LB medium supplemented with the same concentrations of A5P, G6P, and ampicillin were inoculated with single colonies of the three transformants and grown overnight at 37°C. Cells were washed twice with unsupplemented LB medium (to remove the A5P and G6P) and then an

equal number of cells were streaked onto LB/agar/ampicillin plates supplemented with or lacking A5P and G6P before incubation overnight at 37°C.

Determining the pH-rate profile of CtAPI

To determine the optimal pH of its enzymatic activity, CtAPI's ability to isomerize A5P to Ru5P was tested in a series of buffered solutions of varying pH using the cysteine-carbazole assay. Buffered solutions (200 mM buffer, 2 mM EDTA) were prepared at pHs ranging from 5.00 to 8.00 in 0.25 unit increments at 37°C. The buffer 2-(*N*-morpholino)ethanesulfonic acid (MES) was used from pH 5.00 to 6.00, while bis-tris propane buffer was used from pH 6.25 to 8.00. Final reaction concentrations were 100 nM CtAPI, 5 mM A5P, 1 mM EDTA, and 100 mM buffer. Reactions, tested in triplicate, were incubated at 37°C for 3 min before being quenching with equal volumes of 25 N H_2SO_4. Quenched reactions were processed using the cysteine-carbazole assay mixture as described above, and then the absorbance at 540 nm was recorded after 3 h of color development. Relative reaction rates were determined using linear regression.

Enzyme kinetics

Kinetic parameters for the isomerization of Ru5P to A5P were determined, in triplicate, using a modified coupled Aminoff assay in which the A5P was converted to Kdo-8-phosphate using phosphoenolpyruvate (PEP) and the 3-deoxy-D-*manno*-octulosonate-8-phosphate synthase (Kdo8PS) from *Arabidopsis thaliana* (13,14). Reactions were initiated by adding the Ru5P substrate to mixtures containing (final concentrations) 100 mM CtAPI, 100 mM bis-Tris propane (pH 6.75), 10 mM PEP, 0.10 mg/mL Kdo8PS, and 1 mM EDTA. Samples of Ru5P ranging from 0 to 10 mM (final concentration) were incubated separately from the reaction mixtures at 37°C for 3

min before being used to initiate the reactions. The initiated reactions were incubated for 3 min at 37°C and then quenched with an equal volume of 10% (wt/vol) trichloroacetic acid. Color was developed as previously described (3,13). Absorbance was read at 549 nm and the data was fit using the nonlinear least-squares regression function of GraphPad Prism 5 software.

Kinetic parameters for the CtAPI-catalyzed isomerization of A5P to Ru5P were also determined, in triplicate, using the discontinuous cysteine-carbazole assay described above. Final reaction mixtures contained 100 mM CtAPI, 100 mM Bis-Tris propane (pH 6.75), 1 mM EDTA, and A5P concentrations ranging from 0 to 8 mM.

Equilibrium constant (K_{eq}) determination

Mixtures containing 100 mM Bis-Tris propane buffer, pH 6.75, 1 mM EDTA, 10% D_2O, 500 nM CtAPI, and either 5 mM A5P or 5 mM Ru5P were incubated at room temperature (21°C) for 48 h, which was sufficient time to reach equilibrium. The equilibrium constant was determined using ^{31}P nuclear magnetic resonance (NMR) as previously described (3,8). NMR experiments were performed using a Varian 400 multinuclear NMR spectrometer.

Results

Expression and characterization of *C. tetani* API

The primary sequence and genomic context of CtAPI

A BLASTP search of the non-redundant protein database was conducted using the sequence of *E. coli* CFT037 c3406 protein as the query. The search returned several putative SIS-domain proteins in Gram-positive bacteria of the genera *Clostridium, Lactobacillus*, and others. One of these results was the sequence WP_023437760.1 (referred to as *ctAPI*), which encodes a putative API within the genome of *C. tetani,* which we have designated CtAPI. A subsequent

BLASTP search using the sequence of CtAPI as the query revealed putative SIS-domain proteins in Gram-positive bacteria of several genera, including *Bacillus, Staphylococcus* and *Lactobacillus.* These results suggest the presence of APIs in multiple Gram-positive organisms. A multiple sequence alignment of the SIS domain sequences of CtAPI and the known APIs of E. coli (KdsD, GutQ, KpsF and c3406) is shown in Fig 4.1.

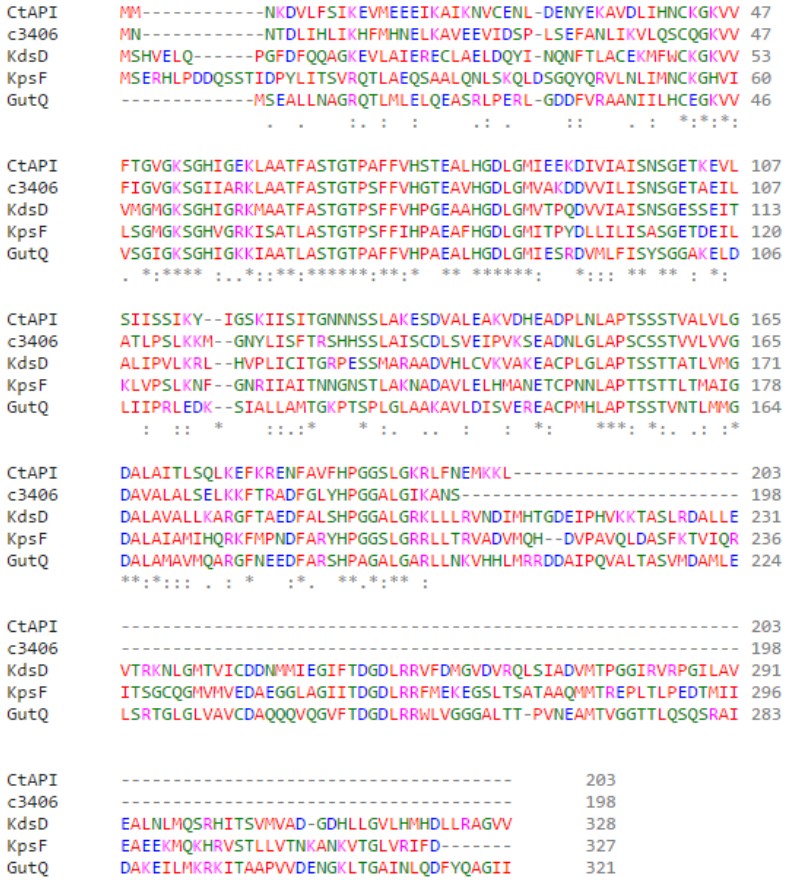

Fig. 4.1: Alignment of sequences of *E. coli* APIs and CtAPI. Sequences were aligned using Clustal W. * denotes residues is absolutely conserved, : denotes conservative, and . denotes semi-conservative.

To obtain evidence regarding the role played by CtAPI and its homologs in Gram-positive bacteria, the genome of C. tetani was searched for evidence of Kdo or K-antigen biosynthesis pathways. BLASTP searches were conducted using the sequences of other enzymes within the *E.*

coli Kdo biosynthetic pathway (KdsC, KdsA, KdsB and WaaA), as well as BLAST searches using the genes *kdsA*, *kdsC*, *kdsB, waaA.* These searches failed to provide evidence for a Kdo biosynthetic pathway in *C. tetani*. Similar searches were conducted using the sequences of the other enzymes in the K-antigen biosynthetic pathway (KpsFEDUCS) and the genes that encode them (*kpsFEDUCS*). These searches failed to provide evidence for the presence of a K-antigen biosynthetic pathway in *C. tetani*. Finally, the genomic context of *ctAPI* was investigated. The gene encoding CtAPI is positioned within an operon harboring several genes encoding proteins involved in carbohydrate metabolism and transport, (Fig. 4.2 B). This operon (CTC00903-CTC00911) includes genes assigned or annotated as a *rbsD/fucU* mutarotase, a LacI family transcriptional regulator, a D-ribose transporter ATP-binding protein, a ribose ABC transporter permease, a ribose ABC transporter substrate-binding protein, CtAPI, a ribokinase, a D-glucose:D-fructose oxidoreductase, and a hypothetical protein, respectively.

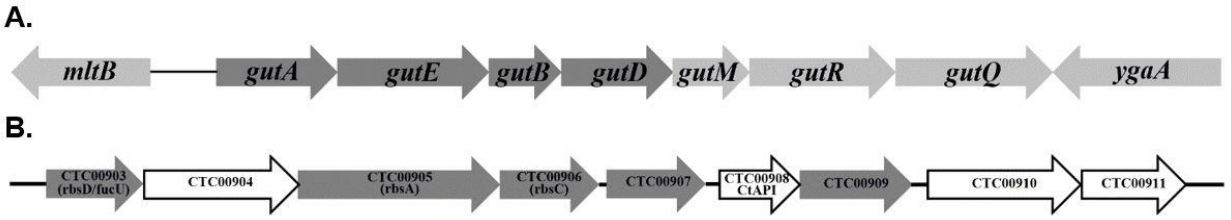

Fig. 4.2: Genomic context of *gutQ* and *ctAPI*. **A.** Genomic context of *gutQ* in *E. coli* K-12 MG1655. Genes involved in sorbitol metabolism and transport are shaded in dark gray. **B.** Genomic context of *ctAPI* in *C. tetani*. Genes involved in ribose transport and metabolism are shaded in gray.

<u>*Recombinant CtAPI is a tetramer*</u>.

Recombinant CtAPI protein carrying an N-terminal His-tag was overexpressed in *E. coli* BL21(DE3) from a synthetic, *E. coli* codon-optimized gene. After purification with a GE HisTrap™ HP column, the recombinant CtAPI was determined to be essentially homogenous (>95% pure)

using SDS-PAGE. The purified, His-tagged CtAPI migrated at approximately 26 kDa on an SDS-PAGE gel, consistent with the calculated subunit molecular mass of 25,740.1 Da. The oligomeric state of native CtAPI was determined using gel-filtration chromatography. Native CtAPI eluted with an apparently molecular mass of 109.2 kDa, which is 4.26 times the calculated subunit mass (Fig. 4.3).

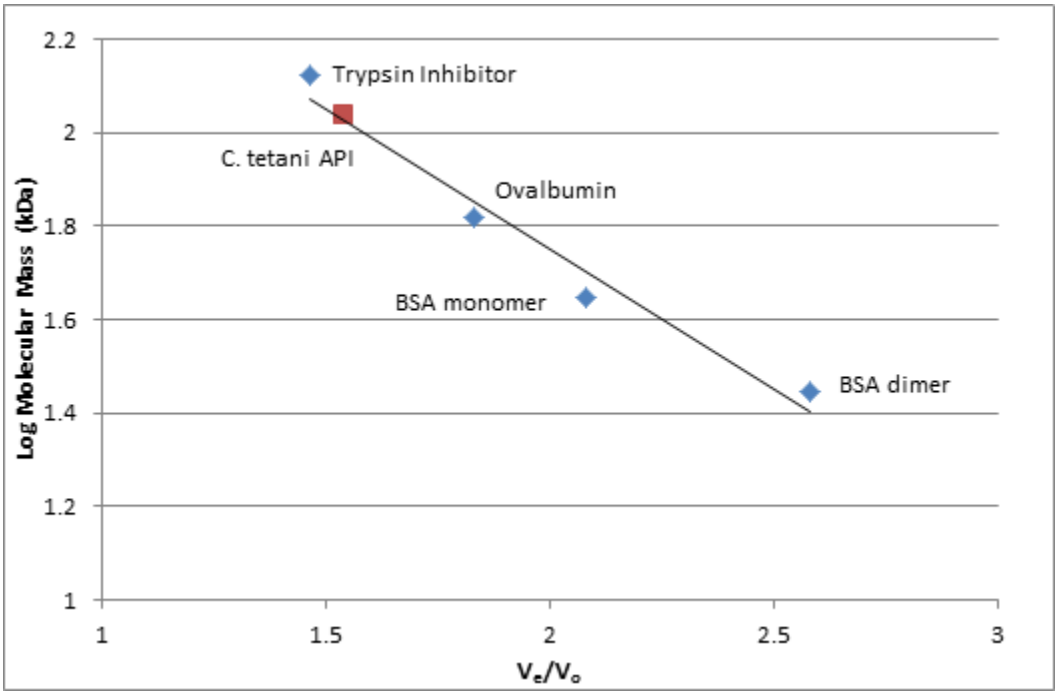

Fig. 4.3: Standard curve from native molecular mass determination. Standard curve from native molecular mass determination by gel filtration chromatography (V_e: elution volume, V_o: void volume, Red square: CtAPI, Blue diamond: Protein Standards).

Enzymatic properties of CtAPI

Recombinant CtAPI is an API.

To determine the substrate specificity of CtAPI, its ability to catalyze the isomerization of a series of aldoses and aldose phosphates, including D-arabinose, D-arabinose-5-phosphate, D-glucose-6-phosphate, D-mannose-6-phosphate, D-glucosamine-6-phosphate, D-glucose-1-phosphate, and D-ribose-5-phosphate, was tested. Of the various substrates tested, CtAPI only converted A5P to the corresponding ketose, Ru5P. A pH-rate profile for the isomerization of A5P to Ru5P was determined; optimal isomerase activity occurred at pH 6.5 (Fig. 4.4).

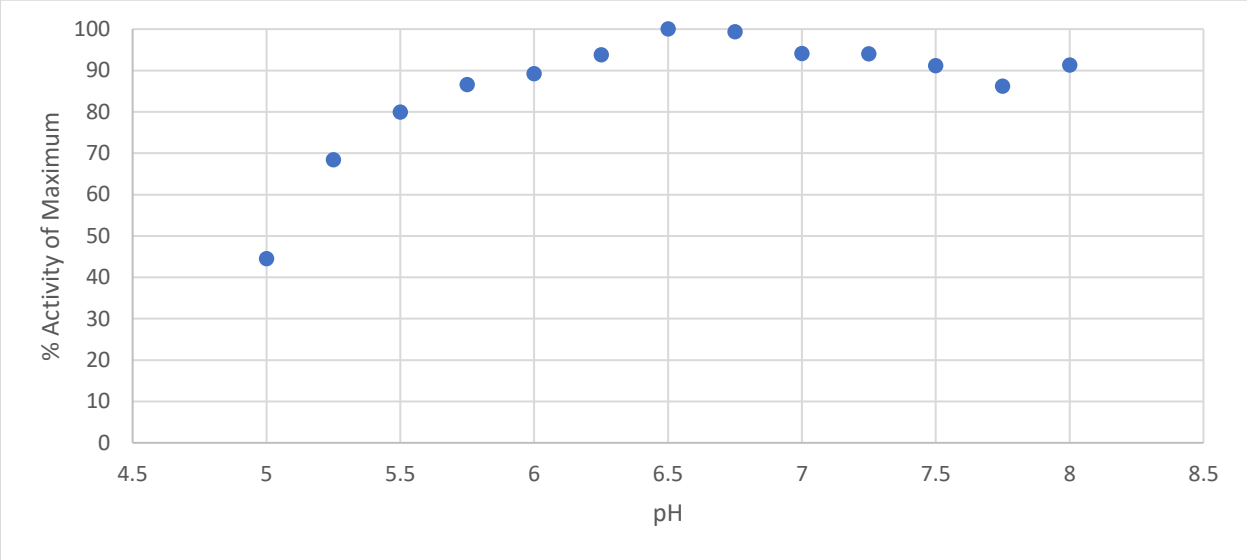

Fig. 4.4: pH rate profile of CtAPI. The rate of isomerization from A5P to Ru5P was measured at 37ºC in a series of pH environments ranging from 5.0 to 8.0.

Enzyme kinetics.

Kinetic parameters for the interconversion of Ru5P and A5P were determined at pH 6.5. In the conversion of Ru5P to A5P, CtAPI was found to have a k_{cat}/K_m of 1.3 X 10^4 $M^{-1}{\cdot}s^{-1}$. In the conversion of A5P to Ru5P, the k_{cat}/K_m was determined to be 5.4 X 10^4 $M^{-1}{\cdot}s^{-1}$. These data, along with a comparison of the kinetic constants determined with those previously obtained for *E. coli* APIs are presented in Table 4.2.

Table 4.2 Kinetic constants for the APIs of *E. coli* and CtAPI.

Protein	k_{cat} (A5P to Ru5P, s^{-1})	K_m (A5P, mM)	k_{cat} / K_m (A5P, M^{-1} s^{-1})	k_{cat} (Ru5P to A5P, s^{-1})	K_m (Ru5P, mM)	k_{cat}/ K_m (Ru5P, M^{-1} s^{-1})	K_{eq}	Optimum pH	Subunit mass (Da)
KdsD[Ref(8)]	157 ±4	0.61 ±0.06	$2.6x10^5$	255±16	0.35±0.08	$7.3x10^5$	0.50±0 .06	8.4	35,084
GutQ[Ref(4)]	218 ±4	1.2 ±0.1	$1.8x10^5$	242±11	0.64±0.08	$3.8x10^5$	0.47	8.25	33,909
KpsF[Ref(5)]	15 ±1	0.57 ±0.04	$2.6x10^4$	19±2	0.30±0.03	$6.3x10^4$	0.48±0 .02	7.75	35,447
C3406[Ref(7)]	16.8 ±0.2	1.92 ±0.05	$8.8x10^3$	10.5±0.08	0.70±0.12	$1.5x10^4$	0.52	6.6	20,880
CtAPI	102 ±5	1.89 ±0.24	$5.4x10^4$	86.7±3.79	6.65±0.15	$1.3x10^4$	0.45	6.5	21,764

CtAPI is able to complement the API deficiency of *E. coli* strain TCM15

TCM15 is an *E. coli* strain prepared by deleting the only two API genes (*kdsD* and *gutQ*) present in the genome of BW32070. Thus, TCM15 is unable to synthesize A5P and is inviable in the absence of a source of A5P because it is unable to synthesize LPS without A5P. In order to support the growth of TCM15, the culture medium must be supplemented with A5P, to support LPS biosynthesis, and G6P, to induce the uhp transport system that internalizes A5P (11). Expression of an active API from a plasmid within the cell will circumvent this A5P/G6P auxotrophy. As a test

of the ability of CtAPI to complement the API deficiency of TCM15, the plasmids pT7-7-CtAPI, pT7-7-c3406 (API positive control), and pT7-7 (vector control) were transformed into TCM15 cells. The API-encoding inserts in these plasmids are expressed from a leaky T7 promotor (15). Equal numbers of cells were plated on LB medium either containing or lacking A5P/G6P (Fig. 4.5). TCM15 cells harboring either pT7-7-CtAPI or c3406 grew on both of the experimental media, while cells harboring the vector control grew only on the medium containing A5P/G6P. These results indicate that CtAPI, like c3406, is an active API *in vivo*.

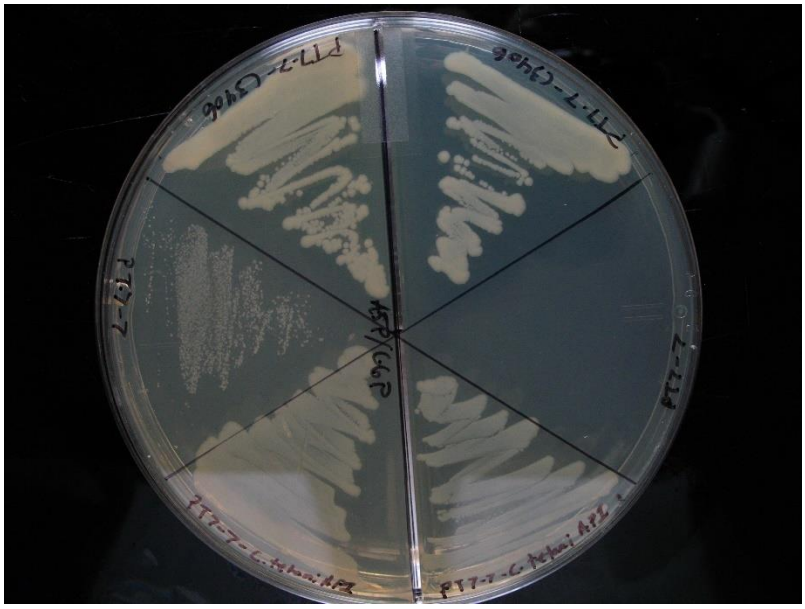

Fig. 4.5: Complementation of an A5P auxotroph on an LB agar plate. (Left) Agar plate containing LB medium supplemented with 15 µM A5P, 10 µM G6P, and 0.1 mg/mL ampicillin. (Right) Agar plate containing LB medium supplemented with 0.1 mg/mL ampicillin. In both halves, the wedges were streaked with: *E. coli* TCM15 harboring pT7-7-c3406 (positive control, wedge 1), *E. coli* TCM15 harboring pT7-7 (vector control, wedge 2), *E. coli* TCM15 harboring pT7-7-CtAPI (CtAPI, wedge 3).

Discussion

A BLASTP search using the protein sequence of c3406 led to the discovery of a putative API, which we have designated CtAPI, in the Gram-positive bacterium *C. tetani*. Both CtaPI and c3406 have a predicted SIS-domain and lack the tandem CBS domain found in the full-length APIs KdsD, GutQ and KpsF. Like all APIs characterized to date, CtAPI was found to be a tetramer in solution. A multiple sequence alignment of CtAPI with c3406, GutQ, KpsF, and KdsD (Fig 4.1) shows that the SIS domain of these proteins is fairly conserved. CtAPI shares primary sequence identities of 55%, 45%, and 52%, with KdsD, GutQ, and c3405, respectively. The five residues previously suggested to be important for catalysis (16,17), namely K52, E78, H81, E104, and H186 (CtAPI numbering) are conserved among all APIs known to date. These results demonstrate that CtAPI possesses the primary and quaternary structural characteristics typical of APIs, suggesting that CtAPI possesses API activity.

To test this hypothesis, the enzymatic activity of CtAPI was investigated. CtAPI was found to specifically interconvert D-ribulose-5-phosphate and D-arabinose-5-phosphate, confirming that CtAPI has API activity *in vitro*. The optimal pH for the isomerization was determined to be 6.5, which is essentially identical to the pH optimum (6.6) of c3406. The kinetic parameters of CtAPI are also comparable, particularly in terms of k_{cat}/K_m, to those of c3406 and *E. coli* KpsF (Table 4.2). This enzymatic activity was sufficient to complement the API-deficiency of bacterial strain *E. coli* TCM15 (Fig. 4.5), indicating that CtAPI can serve as a functional API *in vivo*. This evidence provides strong support for the assertion that CtAPI is the first API characterized from a Gram-positive bacterium.

Upon finding that CtAPI is indeed an API, we questioned what role an API would play in a Gram-positive bacterium. BLASTP searches of the proteome of *C. tetani* using the sequences of the other enzymes in the Kdo biosynthesis pathway failed to find proteins with sufficient sequence similarity to indicate similar function. This indicates that the A5P produced by CtAPI is unlikely to be used for Kdo biosynthesis and should not be considered a K-API. Similar searches conducted using the enzymes of the K-antigen biosynthesis pathway failed to find evidence of a K-antigen biosynthesis pathway in *C. tetani*, indicating that the A5P produced by CtAPI is also unlikely to be used for the synthesis of K antigen and that CtAPI should not be considered a K-API.

We did, however, find evidence suggesting a G-API-like role. The genomic context of the gene encoding CtAPI is remarkably similar to that of *gutQ* in *E. coli.* The *gutQ* operon (Fig. 4.2 A) contains seven genes: *gut(srl)AEBDMRQ* (4,18-21). GutA, GutE, and GutB form the complex that transports D-glucitol-6-phosphate across the inner membrane. GutR and GutM are a transcriptional repressor and a transcriptional activator, respectively (4,18-21). Thus, both CtAPI and GutQ are situated within an island associated with transport and metabolism of a sugar, sorbitol in the case of GutQ and perhaps ribose in the case of CtAPI. GutQ is thought to modulate the expression level of the *gut* operon through the synthesis of A5P. However, it has been challenging to study the effect of A5P synthesis on the *gut* operon due to the requirement for A5P in LPS biosynthesis within Gram-negative bacteria.

Synthesis of the data accumulated during this study leads us to speculate that CtAPI is serving a role in *C. tetani* that is similar to the role played by GutQ in *E. coli*. In this model, A5P is used as a regulatory molecule to induce/repress, through an unknown mechanism, the operon

containing the gene encoding the API. However, we cannot exclude the possibility that A5P functions in the regulation of additional operons likely involving ribose metabolism. We also cannot exclude the possibility that *C. tetani* is using CtAPI for the modulation of other carbohydrate transport and metabolism pathways. The identification of CtAPI within the Gram-positive *C. tetani,* which does not require the biosynthesis of LPS, provides a unique system with which to study the regulation of operons via A5P synthesis.

Acknowledgements and Chapter Contributions

Size exclusion chromatography was carried out by Markin, K., All other experiments were completed by Cech, D.L., writing of the manuscript was completed by Cech, D.L., and Woodard, R.W. The authors would like to thank Dr. Tod P. Holler for editing the manuscript.

References

1. Horecker, B. L. (2002) The Pentose Phosphate Pathway. *J. Biol. Chem.* **277**, 47965-47971

2. Royo, J., Gímez, E., and Hueros, G. (2000) CMP–KDO synthetase: a plant gene borrowed from Gram-negative eubacteria. *Trends in Genetics* **16**, 432-433

3. Cech, D., Wang, P. F., Holler, T. P., and Woodard, R. W. (2014) Analysis of the arabinose-5-phosphate isomerase of *Bacteroides fragilis* provides insight into regulation of single-domain arabinose phosphate isomerases. *J Bacteriol* **196**, 2861-2868

4. Meredith, T. C., and Woodard, R. W. (2005) Identification of GutQ from *Escherichia coli* as a D-Arabinose 5-Phosphate Isomerase. *J. Bacteriol.* **187**, 6936-6942

5. Meredith, Timothy C., and Woodard, Ronald W. (2006) Characterization of *Escherichia coli* D-arabinose 5-phosphate isomerase encoded by kpsF: implications for group 2 capsule biosynthesis. *Biochemical Journal* **395**, 427-432

6. Tabor, S. (1985) A Bacteriophage T7 RNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes. *PNAS* **82**, 1074-1078

7. Mosberg, J., Yep, A., Meredith, T., Smith, S., Wang, P.-F., Holler, T., Mobley, H., and Woodard, R. (2011) A unique arabinose 5-phosphate isomerase found within a genomic island associated with the uropathogenicity of *Escherichia coli* CFT073. *J. Bacteriol.* **193**, 2981-2988

8. Meredith, T. C., and Woodard, R. W. (2003) *Escherichia coli* YrbH Is a D-Arabinose 5-Phosphate Isomerase. *J. Biol. Chem.* **278**, 32771-32777

9. Miller, J. H. (1972) *Experiments in molecular genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

10. Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *PNAS* **97**, 6640-6645

11. Eidels, L., Rick, P. D., Stimler, N. P., and Osborn, M. J. (1974) Transport of D-arabinose-5-phosphate and D-sedoheptulose-7-phosphate by the hexose phosphate transport system of *Salmonella typhimurium*. *J Bacteriol* **119**, 138-143

12. Dische, Z., Borenfreund, E. (1951) A New Spectrophotometric method for the Detection and Determination of Keto Sugars and Trioses. *J Biol Chem* **192**, 583-587

13. Wu, J., Patel, M. A., Sundaram, A. K., and Woodard, R. W. (2004) Functional and biochemical characterization of a recombinant *Arabidopsis thaliana* 3-deoxy-D-manno-octulosonate 8-phosphate synthase. *The Biochemical journal* **381**, 185-193

14. Sheflyan, G. Y., Sundaram, A. K., Taylor, W. P., and Woodard, R. W. (2000) Substrate Ambiguity of 3-Deoxy-D-manno-Octulosonate 8-Phosphate Synthase from *Neisseria gonorrhoeae* Revisited. *J. Bacteriol.* **182**, 5005-5008

15. Mertens, N., Remaut, E., and Fiers, W. (1995) Tight transcriptional control mechanism ensures stable high-level expression from T7 promoter-based expression plasmids. *Biotechnology (N Y)* **13**, 175-179

16. Sommaruga, S., Gioia, L. D., Tortora, P., and Polissi, A. (2009) Structure prediction and functional analysis of KdsD, an enzyme involved in lipopolysaccharide biosynthesis. *Biochemical and Biophysical Research Communications* **388**, 222-227

17. Gourlay, L. J., Sommaruga, S., Nardini, M., Sperandeo, P., Dehò, G., Polissi, A., and Bolognesi, M. (2010) Probing the active site of the sugar isomerase domain from *E. coli* arabinose-5-phosphate isomerase via X-ray crystallography. *Protein Science* **19**, 2430-2439

18. Lengeler, J. (1975) Mutations affecting transport of the hexitols D-mannitol, D-glucitol, and galactitol in *Escherichia coli* K-12: isolation and mapping. *J Bacteriol* **124**, 26-38

19. Lengeler, J. (1975) Nature and properties of hexitol transport systems in *Escherichia coli*. *J Bacteriol* **124**, 39-47

20. Yamada, M., and Saier, M. H., Jr. (1987) Glucitol-specific enzymes of the phosphotransferase system in *Escherichia coli*. Nucleotide sequence of the gut operon. *J Biol Chem* **262**, 5455-5463

21. Yamada, M., and Saier, M. H., Jr. (1987) Physical and genetic characterization of the glucitol operon in *Escherichia coli*. *J. Bacteriol.* **169**, 2990-2994

Chapter V

New Insights Into the Kdo Biosynthetic Pathway and Incorporation Into Lipid A-Like Molecules in *Arabidopsis thaliana*

Summary

The sugar 3-deoxy-D-*manno*-octulosonic acid (Kdo) is found in the rhamnogalacturonan II of higher plants and the lipopolysaccharide of Gram-negative bacteria. The Kdo biosynthetic pathway, which is conserved among plants and Gram-negative bacteria, begins with the isomerization of D-ribulose-5-phosphate to D-arabinose-5-phosphate (A5P) catalyzed by D-arabinose-5-phosphate isomerase (API). In Gram-negative bacteria, Kdo incorporation into lipopolysaccharide is catalyzed by Kdo transferase. Although many bacterial APIs and Kdo transferases have been characterized, knowledge of their orthologs in plants is limited. Herein, we report the characterization of a putative API (AtAPI) and a putative Kdo transferase (AtWaaA) from the plant *Arabidopsis thaliana*. Our results demonstrate that AtAPI, the only protein encoded by the *A. thaliana* genome that shares significant sequence identity with any known API, has enzymatic properties similar to those of bacterial APIs. AtAPI can also complement an API-deficient *Escherichia coli* strain, supporting the hypothesis that this enzyme initiates the Kdo biosynthetic pathway in *A. thaliana*. Our results also demonstrate that AtWaaA, which was previously proposed to catalyze the incorporation of Kdo into rhamnogalacturonan II, can transfer Kdo to lipid A in a Kdo transferase-deficient *E. coli* strain. These results suggest that Kdo

may be incorporated into lipid A-like molecules in *A. thaliana,* a role not yet described in plants, in addition to its role in rhamnogalacturonan II synthesis.

Introduction

The eight-carbon acidic sugar 3-deoxy-D-*manno*-octulosonic acid (Kdo), normally found in the lipopolysaccharide (LPS) of Gram-negative bacteria, has also been shown to be present in

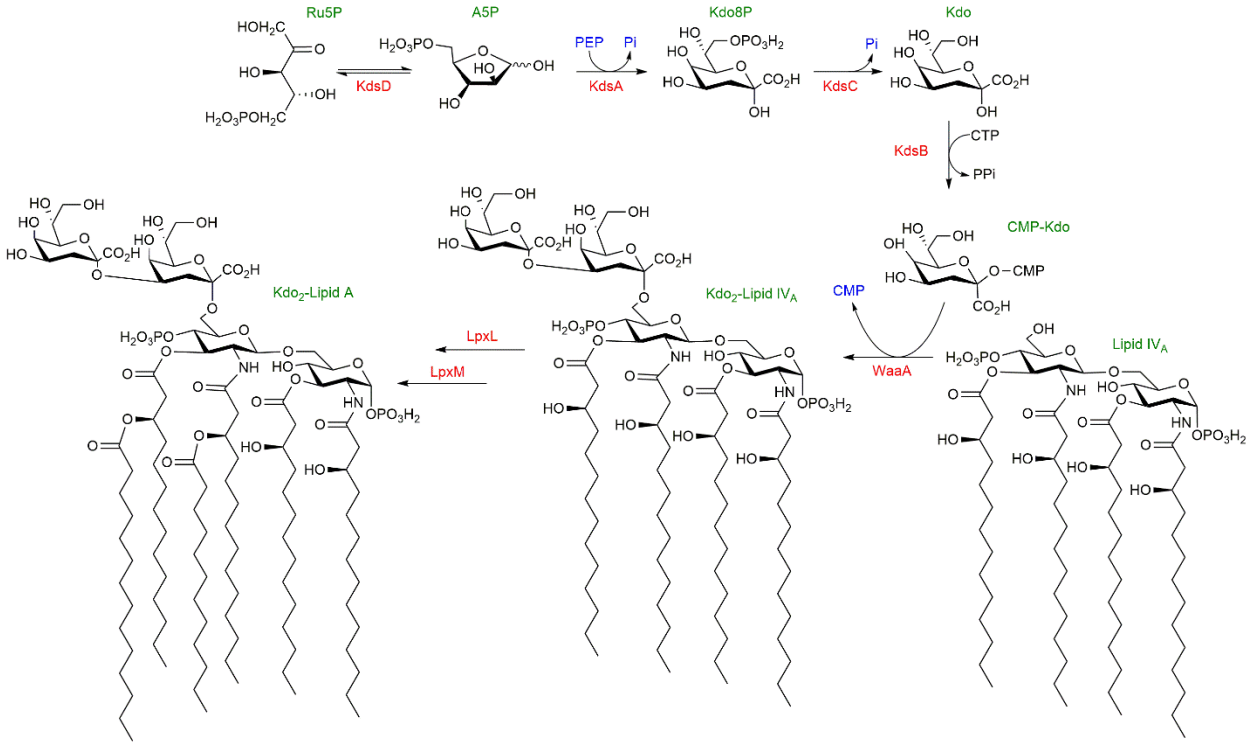

Fig. 5.1: Kdo biosynthesis and incorporation in *E. coli*. The enzymes (in red) involved are D-arabinose-5-phosphate isomerase, Kdo8P synthase, Kdo8P phosphatase, and CMP-Kdo synthetase. Two CMP-activated Kdo molecules are transferred to lipid IV_A via Kdo transferase before the addition of secondary acyl chains by LpxL and LpxM.

some green algae and higher plants (1,2). The biosynthesis and enzymatic activation of Kdo (Fig. 5.1) generally involves four sequential enzymes: D-arabinose-5-phosphate (A5P) isomerase (KdsD), Kdo-8-phosphate synthase (KdsA), Kdo-8-phosphate phosphatase (KdsC), and cytidine-5'-monophosphate-Kdo synthetase (KdsB) (3). These enzymes catalyze the synthesis of Kdo through the isomerization of the pentose pathway intermediate D-ribulose-5-phosphate (Ru5P) to A5P, condensation of A5P with phosphoenol pyruvate to yield Kdo-8-phosphate and inorganic

phosphate, dephosphorylation of Kdo-8-phosphate to Kdo and inorganic phosphate, and the activation of Kdo with CTP to yield CMP-Kdo and pyrophosphate, respectively. CMP-Kdo is transferred to lipid IV_A by the Kdo transferase WaaA, previously referred to as KdtA by Raetz (3,4).

LPS, also known as endotoxin, is an essential part of the outer membrane of Gram-negative bacteria and is associated with endotoxic shock, septicemia, and other immunological responses (5). Bacteria with compromised LPS are generally more susceptible to antibiotics and less pathogenic (6,7). LPS contains three main components: lipid A, which is embedded in the outer membrane; an oligosaccharyl core; and O-antigen (5). In LPS, Kdo residues are attached to lipid IV_A and link the extracellular carbohydrate domain of LPS to the membrane-embedded lipid IV_A (8).

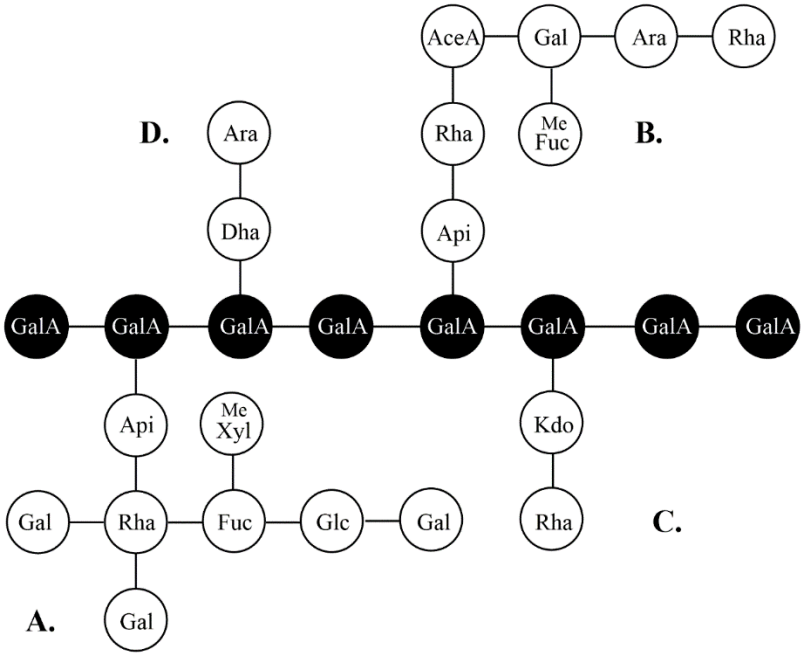

Fig. 5.2: A model of RG-II structure with side chains labeled. Abbreviations are AceA: Aceric acid, Api: Apiose, Ara: Arabinose, Dha: 3-Deoxy-lyxo-heptulosaric acid, Fuc: Fucose, Gal: Galactose, GalA: Galacturonic Acid, Glc: Glucose, Kdo: 3-Deoxy-D-*manno*-octulosonic acid, Me Fuc: 2-*O*-methyl-L-fucose, Me Xyl: 2-*O*-methyl-D-xylose, Rha: Rhamnose

In vascular plants, Kdo is part of the Rha-α(1-5)Kdo side chain C of the highly conserved cell wall component rhamnogalacturonan II (RG-II; Fig. 5.2) (9,10). RG-II is found mainly as a dimer crosslinked by a borate diester bond between the residues of side chain A. This dimer is important for cell wall integrity (11). Studies have shown that a normal RG-II structure is necessary for the crosslinking of RG-II by borate and plant growth (10). Kdo-deficient mutants of *Arabidopsis thaliana* are impaired in pollen tube elongation, which is important in carrying sperm cells to the ovule (9,12-14).

The biosynthesis and enzymatic activation of Kdo in *A. thaliana* is analogous to that in Gram-negative bacteria. Homologs of the bacterial *kdsD*, *kdsA*, and *kdsB* genes have been annotated in the *A. thaliana* genome (15). In *A. thaliana,* At3g54690 encodes a putative A5P isomerase (AtAPI), At1g79500 and At1g16340 encode two potentially functional Kdo-8-phosphate synthases, At1g53000 encodes a functional CMP-Kdo synthetase, and At5g03770 encodes a putative Kdo transferase (AtWaaA) (12,16,17). There are no significant homologs of bacterial Kdo-8-phosphate phosphatases in *A. thaliana*, but knockouts of this gene in Gram-negative bacteria are not lethal (16). All of these *A. thaliana* enzymes have been characterized except AtAPI and AtWaaA, which have received little attention other than the disruptional mutation of the two genes (12,16-18).

All the APIs studied to date have come from Gram-negative bacteria. Full-length APIs, such as *E. coli* KdsD, KpsF, and GutQ, contain a sugar isomerase (SIS) domain and a tandem cystathionine beta-synthase (CBS) domain, which is thought to be regulatory. The SIS-domain APIs, such as c3406 from *E. coli* CFT073 and Q5LIW1 of *Bacteroides fragilis*, are another form of API that lacks the tandem CBS domain of the full-length APIs (19-23).

Null mutants of the Kdo biosynthetic genes in *A. thaliana* appear to be nonviable, likely due to the gametophyte-lethality of these mutations (9,12,18). However, null mutants of AtWaaA are phenotypically indistinguishable from the wild-type. In these null mutants, the sugar composition of RG-II remains identical to that of wild-type *A. thaliana*, including the presence of wild-type amounts of the Rha-α(1-5)Kdo side chain (15,17). This has led to the speculation that AtWaaA does not incorporate Kdo into RG-II, as a deficiency in RG-II would impair proper pollen tube elongation (11). Furthermore, AtWaaA-GFP fusions have been shown to localize to ring-like structures surrounding the mitochondria (17). This is also inconsistent with the suggestion that AtWaaA is involved in the synthesis of RG-II, because RG-II synthesis is thought to occur in the Golgi apparatus. Rather, it is consistent with the hypothesis that AtWaaA is involved in the transfer of Kdo to an unidentified acceptor molecule in the mitochondrion.

A. thaliana contains homologs of all the genes necessary to produce the lipid A precursor lipid IV_A (Fig. 5.1) except *lpxH*, which is also absent in some Gram-negative bacteria (24,25). Most of the *lpx* genes, which are involved in the biosynthesis of lipid IV_A, have either been shown or predicted to be targeted to the *A. thaliana* mitochondrion (16,17,25). It has been reported that *A. thaliana* produces the lipid IV_A precursor known as lipid X (25). Lipid X levels were highest in the mitochondria and elevated in chloroplasts, compared with whole-cell homogenates. Null mutants of *AtlpxA*, *AtlpxD2*, *AtlpxB*, *AtlpxK*, and *AtwaaA*, along with *AtlpxC*-RNA*i* knock-down plants, were phenotypically indistinguishable from wild-type plants (25). Various lipid A precursors were detected in *A. thaliana* null mutants and knock-downs. However, lipid IV_A was detected only in *atkdta-1*, an At3g54690 mutant strain, and Kdo-lipid IV_A, Kdo_2-lipid IV_A, and the putative penta- or hexa-acylated derivatives were not detected (25).

In this study, designed to broaden knowledge of the Kdo biosynthetic pathway in *A. thaliana,* we characterized AtAPI and explored the possibility that Kdo is added to a lipid A-like molecule by AtWaaA.

Materials and Methods

Materials

The synthetic, *E. coli* codon-optimized *A. thaliana* API gene (Locus Name: At3g54690) was purchased from Genewiz (South Plainfield, NJ, USA). The synthetic, *E. coli* codon-optimized *A. thaliana waaA* (NP_195997.2) gene was purchased from GenScript (Piscataway, NJ, USA). Enzymes for subcloning were purchased from New England Biolabs (Ipswich, WI, USA). Plasmid DNA was purified using the Promega Wizard Miniprep kit (Madison, WI, USA). DNA sequencing was performed by the University of Michigan Biomedical Resources Core Facility.

Table 5.1: Stains, plasmids, and primers used in this study.

Item	Description	Source
E. coli		
TOP10	F- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) F80*lacZ*ΔM15 Δ*lacX*74 *recA1 araD*139 Δ(ara-leu)7697 *galU galK rpsL* (StrR) endA1 *nupG*	Invitrogen
TCM15	BW30270(*ΔkdsDΔgutQ*)	Ref. (21)
BL21 (DE3)	*F⁻ ompT gal dcm lon hsdS$_B$(r$_B$⁻m$_B$⁻) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB⁺]$_{K-12}$(λ^S)*	NEB
KPM22	TCM15 *yhjD400*	Ref. (5)
KPM50	KPM22 *kdsD*$^+$	Ref. (26)
KPM53	KPM50 Δ*waaC*	Ref. (26)
KPM54	KPM50 Δ*waaA*	This Study
KPM56	KPM53 Δ*waaA*	Ref. (26)
Plasmids		
pT7-7	T7 Expression vector, Amp^R	Ref. (27)
pET-16b	T7 expression vector for N-terminal His-tag fusions, Amp^R	Novagen
pET-19b	T7 expression vector for N-terminal His-tag fusions, Amp^R	Novagen
pT7-7-c3406	*E. coli* CFT073 c3406 inserted into NdeI/BamHI of pT7-7, Amp^R	Ref. (22)
pET-19b-AtAPI	*E. coli* codon optimized A.*thaliana* At3g54690 inserted into NdeI/BamHI of pET-19b, Amp^R	This study
pET-16b-AtWaaA	*A. thaliana* At5g03770 (from U17171) inserted into NdeI/BamHI of pET-16b, Amp^R	This study
pET-16b-EcWaaA	*E. coli waaA* inserted into NdeI/BamHI of pET-16b, Amp^R	Ref. (26)
pUM212	pET-16b carrying the *waaA* gene of *E. coli* K-12, Amp^R	Ref. (26)
pUC57-Kan-*AtAPI*	PUC57-Kan carrying the *E. coli* codon optimized *A. thaliana API* gene, Kan^R	Genewiz
U17171	pENTR/SD/D-TOPO_AT5G03770.1	ABRC/Ref. (28)
Primers		
AtWaaA.F	TGTCCATATGAAGCTCGGAGTGTTTGTATAC	IDT
AtWaaA.R	CACTGGATCCTCACTCGAGTTTGCATTCAATGTGATTTCTTG	IDT
ECOwaaAH1	ACAGCTAAATACATAGAATCCCCAGCACATCCATAAGTCAGCTATTTACTGTGTAGGCTGGAGCTGCTTC	MWG
ECOwaaAH2	TAATGGGATCGAAAGTACCCGGATAAATCGCCCGTTTTTGCATAACAACCCATATGAATATCCTCCTTAG	MWG

Bacterial strains, plasmids, primers, and growth media

The bacterial strains, plasmids, and primers used in this study are listed in Table 5.1. All strains were grown in LB medium. The auxotrophic strain TCM15, dependent on exogenous A5P for growth, is a derivative of BW30270 in which the *kdsD* and *gutQ* genes have been disrupted using the phage λ Red recombinase system (21,29). TCM15 cultures were supplemented with 10 µM G6P and 15 µM A5P. G6P is necessary for induction of the uhp transporter, which allows for the uptake of A5P (30).

Cloning, expression, and purification of *A. thaliana* API

The *E. coli* codon-optimized synthetic gene encoding AtAPI was cloned into the *Nde*I and *Bam*HI sites of the plasmid pUC57-kan (31). The insert was then subcloned into pET-19b and inserted into chemically competent *E. coli* TOP10 cells. After verification of the resulting plasmid using DNA sequencing, the pET-19b-AtAPI plasmid was inserted into chemically competent *E. coli* BL21 (DE3) cells for protein expression. A fresh transformant was grown, with shaking (250 rpm), at 37 °C in LB medium supplemented with 100 µg/mL ampicillin until the culture reached an optical density at 600 nm of 0.6. After cooling the culture to 19 °C, expression of protein was induced by the addition of isopropyl-β-D-1-thiogalactopyranoside to 0.42 mM. The culture was incubated for 16 h at 19 °C before cells were harvested by centrifugation (6,000 *x g*, 10 min, 4 °C). Cell pellets were suspended in 50 mL of buffer A (20 mM Tris-HCl, pH 8.0, 300 mM NaCl) prior to being sonicated on ice (5 cycles of 20-s bursts, 2-min pauses between bursts). Lysed cells were clarified by centrifugation at 20,000 *x g* for 45 min at 4 °C. The resulting supernatant was filtered through a 0.45 µm Millipore polyvinylidene difluoride (PVDF) filter and loaded onto a HisTrap HP 5 mL column (GE Healthcare) that had been pre-equilibrated, with 10 mL buffer A. The column

was washed with 100 mL of buffer A supplemented with 100 mM imidazole, and protein was eluted with 20 mL buffer A supplemented with 500 mM imidazole. Fractions containing AtAPI, as determined using SDS-PAGE, were pooled. The pooled fractions were dialyzed against buffer containing 20 mM Tris-HCl, pH 8.0, and 100 mM NaCl, and then concentrated using an Amicon ultracentrifugal filter (10,000 molecular weight cutoff). Purified protein was stored at 4 °C prior to use.

Synthesis of Na_2Ru5P

The synthesis of Na_2Ru5P followed the procedure of Pontremoli and Mangiarotti (32), with modifications. A mixture containing 1.54 mmol of D-gluconate-6-phosphate, 0.015 mmol of $NADP^+$, and 2.92 mmol of sodium pyruvate was dissolved in 50 mL of glycylglycine buffer (37.5 µM, pH 7.6) at room temperature. Next, 6-phosphogluconate dehydrogenase and lactate dehydrogenase were added to final concentrations of 500 nM and 588 nM, respectively, and the solution was incubated at 37 °C for 5 h. Similar to Pontremoli and Mangiarotti (31), reaction progress was monitored using the disappearance of D-gluconate-6-phosphate via the carbazole-cysteine assay (33). When D-gluconate-6-phosphate was no longer detected, the reaction mixture was chilled to 0 °C and 650 mg of activated charcoal was stirred into the reaction to remove nucleotides. This mixture was filtered and the filtrate was loaded onto a DEAE Sepharose column that had been pre-equilibrated with water. The column was washed with water and then eluted with a gradient of 0 to 120 mM NaCl. The fractions containing Ru5P, identified using the carbazole-cysteine assay, were concentrated to a semi-solid by rotary evaporation under vacuum. The residue was dissolved in 20 mL methanol, and then concentrated via rotary evaporation under vacuum to dryness. To remove insoluble by-products, the product was

dissolved in 10 mL methanol and centrifuged at 9,000 *x g* for 10 min. The pellet was discarded and the supernatant was concentrated by rotary evaporation under vacuum to dryness. Finally, the product was heated under vacuum to remove any residual methanol and water. The final yield of Na_2Ru5P was 53%.

Native oligomeric state

The native molecular mass of AtAPI was approximated using gel filtration chromatography on a HiPrep (26/60) Sephacryl S-100 column. Bovine serum albumin (BSA) dimer, (132.4 kDa), BSA monomer (66.2 kDa), chicken egg white ovalbumin monomer (44.3 kDa), and cytochrome *c* (12.4 kDa) were used as standards. Samples were run in triplicate using 20 mM Tris-HCl, pH 8.0, 50 mM NaCl as the eluent. A plot of log molecular mass vs. elution volume/void volume was fit using linear regression with Microsoft Excel. The oligomeric state of AtAPI was determined using the elution volume of the native protein and the standard curve.

Substrate specificity

To determine the substrate specificity of AtAPI, reaction mixtures containing 200 nM AtAPI, 100 mM Bis-Tris Propane, pH 8.0, and 1 mM EDTA were incubated for 3 min at 37 °C prior to the addition of 10 mM of D-arabinose, D-ribose-5-phosphate, G6P, D-glucose-1-phosphate, D-glucosamine-6-phosphate, D-mannose-6-phosphate, or D-arabinose-5-phosphate, also held at 37 °C, to initiate the reaction. After incubation for 3 min at 37 °C, reactions were quenched with an equal volume of 12.5 M H_2SO_4. Formation of the corresponding ketoses was visualized using the cysteine-carbazole assay (23,33). All plates contained no-enzyme controls and no-substrate controls in separate reactions. The assays were performed in triplicate.

Complementation of *E. coli* TCM15

TCM15 cells were grown at 37 °C to the mid-exponential growth phase (OD_{600} ~0.5) in LB medium supplemented with 15 µM A5P and 10 µM G6P. These cells were pelleted by centrifugation at 3300 *x g* and 4 °C for 10 min, washed by resuspension in ice cold water, re-pelleted, and then washed again. The washed cell pellet was resuspended in 50 µL of water for electroporation. Empty pT7-7 vector, pET-19b-AtAPI, and pT7-7-c3406 were separately inserted into the electrocompetent TCM15 cells. Transformed cells were grown on LB agar plates supplemented with 100 µg/mL ampicillin, 15 µM A5P, and 10 µM G6P. Single colonies taken from each plate were grown overnight in liquid LB medium supplemented as above. The resulting cells were washed twice with LB medium (by suspending the cells in LB medium, then pelleting them at 3300 *x g* and 4 °C) to remove the A5P and G6P present in the overnight cultures. Cells were then suspended in LB medium and streaked onto sectors of an LB agar plate with or without A5P and G6P supplementation (as above). Plates were incubated at 37 °C for 16 h before obtaining the photograph shown in Fig. 5.3.

pH-rate profile of AtAPI

The optimal pH for AtAPI activity was determined by performing the discontinuous cysteine-carbazole assay (see below) with a series of buffer solutions at varying pHs. Buffers ranged in pH from 5.25 to 8.75 in increments of 0.25 pH units at 37 °C. Buffers with pH values from pH 5.25 to 6.00 contained final concentrations of 100 mM MES, 1 mM EDTA, while buffers ranging from pH 6.25 to 8.75 contained 100 mM bis-tris propane (BTP), and 1 mM EDTA. Buffer solutions containing final concentrations of 400 nM enzyme were incubated separately from solutions of 10 mM A5P at 37 °C for 3 min. Equal volumes of A5P were added to the reaction mixtures and allowed to react for 3 min before being quenched with an equal volume of 12.5 M

H_2SO_4. Color was developed using the cysteine-carbazole solution described above. Plates contained controls in which enzyme was omitted and all reactions were carried out in triplicate.

Enzyme kinetics

Aminoff assay

Kinetic parameters for the isomerization of Ru5P to A5P were determined, in triplicate, using a coupled assay employing the Kdo-8-phosphate synthase from *A. thaliana* (16). Reaction mixtures containing final concentrations of 100 mM BTP, pH 7.75, 1 mM EDTA, 10 mM phosphoenolpyruvate, 3.2 µM Kdo-8-phosphate synthase, and 100 nM AtAPI were incubated separately from mixtures of Ru5P at 37 °C for 3 min. Reactions were initiated with the addition of Ru5P, and then allowed to react for 3 min at 37 °C before quenching with an equal volume of 10% (w/v) trichloroacetic acid. Kdo-8-phosphate formed in this reaction was quantitated using the Aminoff assay formatted for a 96-well plate (23). Kdo concentrations were determined from a standard curve of Kdo on the same plate. Data was fitted using the nonlinear least-squares regression function of GraphPad Prism 5 software.

Discontinuous cysteine-carbazole assay

Kinetic parameters for the isomerization of A5P to Ru5P were determined, in triplicate, using the discontinuous cysteine-carbazole assay (23). Reaction mixtures, which contained final concentrations of 100 nM AtAPI, 1 mM EDTA, 100 mM BTP pH 7.75, and A5P substrate solutions were preheated separately at 37 °C. Reactions were initiated by mixing the substrate and reaction mixtures to obtain final A5P concentrations ranging from 0.156 to 10 mM. Reaction mixtures and controls omitting enzyme were incubated for 3 min at 37 °C before quenching with equal volumes of 12.5 M H_2SO_4. Ru5P concentrations were determined using a standard curve that was present

on the same plate. The data were fit using the nonlinear least-squares regression function of GraphPad Prism 5 software.

Equilibrium constant (K_{eq})

Solutions containing 10% D_2O, 100 mM BTP pH 7.75, 1 mM EDTA, 500 nM AtAPI, and 5 mM of either A5P or Ru5P were incubated at room temperature (~21 °C) for 48 hours to allow the enzymatic reaction to reach equilibrium. Equilibrium constant was determined by ^{31}P NMR as previously described (23).

Kdo_2-lipid IV_A production by AtWaaA vs EcWaaA

The AT5G03770.1 gene, which is located within the AT5G03770 locus, was amplified from the plasmid U17171 (Arabidopsis Biological Resource Center) using the primers AtWaaA.F and AtWaaA.R (Table 5.1). The primers were designed to incorporate *Nde*I and *Bam*HI sites. PCR products were purified via electrophoresis through a 1% (w/v) agarose gel, excising the appropriately sized band and isolating the DNA using a Qiagen QIAquick gel extraction kit. After digestion with *Nde*I and *Bam*HI, the restricted fragments were purified via agarose gel as described above. Purified, digested PCR products were inserted into similarly digested and purified pET-16b vector (34). The resulting plasmid, pET-16b-AtWaaA, was inserted into chemically competent *E. coli* TOP10 cells.

Separate aliquots of electrocompetent KPM56 cells were transformed with pET-16b vector, pET-16b-AtWaaA, and pET-16b-EcWaaA (26). Transformed cells were grown on LB agar plates supplemented with 100 mg/L ampicillin. Single transformant colonies were used to inoculate 2 -mL aliquots of liquid LB medium supplemented with ampicillin, and then grown overnight at 37 °C. The overnight culture was diluted 1:100 in fresh LB medium and incubated at

37 °C at 200 rpm overnight. Cells were pelleted by centrifugation at 4388 *x g* and 4 °C for 30 min (Eppendorf Centrifuge 5804R). The cell pellet was resuspended in 30 mL of phosphate buffered saline and centrifuged again. Next, the cells were resuspended in 20 mL of ethanol and stirred at 4 °C overnight. Cells were centrifuged before being resuspended in 20 mL of acetone. The suspension was stirred at 4 °C for 2 h before centrifugation. The cell pellet was suspended in 20 mL of acetone and stirred at 4 °C for 2 h before a final centrifugation. Biomass was recovered by suspending the cells in a small volume of diethyl ether and filtration via indirect vacuum. The biomass was dried overnight and then pulverized to fine particulates.

LPS purification and analysis

LPS samples were purified by a modified version of the phenol-chloroform-petroleum ether procedure (35). A 3-mL aliquot of PCPI solution composed of 2 parts 90% phenol: 5 parts chloroform: 8 parts petroleum ether was added to each 300-mg sample of dried biomass. The suspension was allowed to stir at room temperature for 60 min. These samples were centrifuged for 10 min at 4000 *x g* in glass centrifuge tubes, then the supernatant was added to a round-bottomed flask. This procedure was repeated a second time with the supernatants being combined. Petroleum ether was removed from these samples via rotary evaporation under vacuum. The remainder of the LPS samples was dialyzed against 10 L of water with stirring. This dialysis was repeated twice with fresh solution. The LPS-containing solution, now in water, was ultrasonicated in a water bath for 10 min to homogenize the sample and then centrifuged for 10 min at 6000 *x g* at room temperature to remove debris. The supernatant containing LPS was frozen and lyophilized overnight leaving LPS as a white solid. Samples were resuspended in LPS-free water (Braun water) to a concentration of 1 mg/mL for mass spectrometric analysis.

LPS samples (1 mg/mL) were diluted 15-fold in a solution containing 50% 2-propanol supplemented with 4 mM triethylamine and 0.35 mM acetic acid. The resulting suspension had a pH of approximately 8.5. All samples were analyzed in the negative ion mode using a Q Exactive Plus instrument (Thermo, Bremen, Germany) with electrospray ionization. Samples were delivered with a syringe pump at a flow rate of 5 µL/min using a sheath gas flow of 5 L/min, a transfer capillary temperature of 250 °C, and an S-lens level of 100. Spectra were recorded at highest resolution (280,000 FWHM defined at m/z 200) with either no further collisional activation, or all-ion fragmentation activated with NCE set to 25 V.

Generation of *E. coli* KPM54

The *E. coli* Δ*waaA* strain KPM54 was constructed from KPM50, a KPM22 *kdsD*$^{+}$ derivative of BW30270 carrying the *yhjD400* suppressor allele that enables the strain to tolerate null mutations in essential Kdo pathway genes (26). To construct the chromosomal *waaA* deletion, the phage λ Red recombinase procedure was used essentially as described (29). The insert cassette targeting the *waaA* gene was constructed using plasmid pKD4 as the template and the primer pair ECOwaaAH1/ECOwaaAH2. Transformation of KPM50/pKD46 with the *waaA::kan* targeting cassette allowed the deletion to occur. Finally, the kanamycin resistance marker was excised in the presence of plasmid pCP20 to yield KPM54. The helper plasmid pKD46 was cured at 42 °C as recommended (29), whereas pCP20 was removed from KPM54/pCP20 at 37 °C to accommodate the temperature-sensitive phenotype of lipid IV_A-expressing strains (5).

Rescue of *E. coli* KPM54

Aliquots of electrocompetent KPM54 cells, prepared as described above, were separately transformed with pET-19b, pET-19b-AtWaaA, and pET-19b-EcWaaA. After initial growth on

LB/Amp plates, cells were grown with shaking (200 rpm) in liquid LB medium at 37 °C, streaked onto MacConkey agar, and incubated overnight at 37 °C (36).

A. thaliana seedlings and growth conditions

A. thaliana wild-type Col-0 (CS70000) and the *atkdta-1* (SALK_035981) mutation were obtained from the *Arabidopsis* Biological Resource Center (37). Seeds were surface sterilized by treatment with 50% bleach and 0.2% Triton X-100 for 10 min then washed six times with sterile distilled water. To break dormancy seeds were suspended in sterile water, and kept in the dark for 3 days at 4 °C. Seedlings were grown under sterile conditions in Murashige-Skoog medium supplemented with 0.1% Plant Preservative Mixture in long day cycles (16 h light/8 h dark) for 10 days (25).

Lipid extraction from _A. thaliana_ seedlings

Seedlings were collected by sterile filtering and ground in a mortar and pestle with 4 mL of ice cold buffer (0.3 M sorbitol, 5 mM EGTA, 5 mM EDTA, 20 mM Hepes pH 8.0, 8 mM cysteine, 10 mM Tricine) per g fresh plant material. The homogenate was passed through 2 layers of Miracloth®, protein concentration was determined via Bradford assay, and samples were aliquoted to contain 1 mg of protein. Lipids were extracted, using the Bligh-Dyer method (38), from aliquots containing equal amounts of protein by diluting to 1.6 mL in phosphate-buffered saline (39). Each 1.6 mL sample was extracted for 1 h at room temperature in a single phase Bligh-Dyer system containing a mixture of chloroform/methanol/water (1:2:0.8, vol/vol). The addition of chloroform and aqueous HCl, resulted in a biphasic system consisting of chloroform/methanol/0.1 M HCl (2:2:1.8), in which the lower phases were collected and the solvent was removed by rotary evaporation under vacuum. Lipids were stored at -80 °C.

Limulus Amebocyte Lysate test

Solid lipid samples were resuspended to 50 mg/mL in endotoxin free water. Samples were tested for endotoxin utilizing the microplate method for QCL-1000 quantitative chromogenic LAL test kit (BioWhittaker, Inc.). Briefly, 50 µL of sample/standard were dispensed into a pyrogen free microplate pre-equilibrated to 37 °C. Next, 50 µL of LAL was added and the plate was incubated at 37 °C for 10 min before the addition of 100 µL of substrate solution, prewarmed to 37 °C, to each reaction well. After an additional 6 min of incubation, 50 µL of 10% sodium dodecyl sulfate was added to stop the reaction. Absorbance was measured at 405 nm and data was fit using linear regression in Microsoft excel.

Results

Expression and quaternary structure of AtAPI

BLASTP searches of the *A. thaliana* (taxid: 3702) proteome were conducted using *E. coli* KdsD and *E. coli* CFT073 c3406 as query sequences. Both searches returned the same homologous sequence annotated as At3g54690 or NCBI accession number NP_191029.1. The protein sequence, which we will refer to as AtAPI, shares 31% and 35% amino acid identity with KdsD and c3406, respectively. A protein expression vector, pET-19b-AtAPI, was generated and used to overexpress AtAPI in *E. coli* BL21 (DE3). The recombinant, His-tagged enzyme was purified with a HisTrap HP column and determined to be homogenous using SDS-PAGE.

Purified, His-tagged AtAPI migrated at approximately 42 kDa on an SDS-PAGE gel. The subunit molecular mass of the protein was calculated to be 41,724.9 Da. The quaternary structure of the native protein was determined using this calculated mass and gel filtration

chromatography. Native AtAPI eluted with an apparent molecular mass of 165.2 kDa, which is 3.96 times the calculated subunit mass, indicating that AtAPI is a tetramer.

Enzymatic properties of AtAPI *in vitro*

The substrate specificity of AtAPI was determined by probing its ability to catalyze the isomerization of a series of aldoses and aldose phosphates, including D-arabinose, D-ribose-5-phosphate, D-glucose-6-phosphate (G6P), D-glucose-1-phosphate, D-glucosamine-6-phosphate, D-mannose-6-phosphate, and D-arabinose-5-phosphate. These experiments were performed at pH 8.0 based on the pH optima of the *E. coli* APIs. Of the substrates tested, AtAPI converted only A5P to the corresponding ketose. A pH-rate profile for the isomerization of A5P to Ru5P demonstrated that AtAPI has optimal isomerase activity at pH 7.75.

Kinetic parameters were determined for the interconversion of Ru5P and A5P at pH 7.75. AtAPI, in the conversion of Ru5P to A5P, had a k_{cat}/K_m of 1.2 X 10^3 $M^{-1}{\cdot}s^{-1}$. In the A5P to Ru5P direction, the k_{cat}/K_m was found to be 2.2 X 10^3 $M^{-1}{\cdot}s^{-1}$. A comparison of the kinetic constants determined for AtAPI with those previously obtained for the two-domain APIs of *E. coli* is presented in Table 5.2.

Table 5.2 Kinetic constants for catalysis by various APIs.

Protein	k_{cat} (A5P to Ru5P, s^{-1})	K_m (A5P, mM)	k_{cat} / K_m (A5P, M^{-1} s^{-1})	k_{cat} (Ru5P to A5P, s^{-1})	K_m (Ru5P, mM)	k_{cat}/ K_m (Ru5P, M^{-1} s^{-1})	K_{eq}	Optimum pH	Subunit mass (Da)
KdsD[19]	157 ±4	0.61 ±0.06	$2.6x10^5$	255±16	0.35±0.08	$7.3x10^5$	0.50±0.06	8.4	35,084
GutQ[21]	218 ±4	1.2 ±0.1	$1.8x10^5$	242±11	0.64±0.08	$3.8x10^5$	0.47	8.25	33,909
KpsF[20]	15 ±1	0.57 ±0.04	$2.6x10^4$	19±2	0.30±0.03	$6.3x10^4$	0.48±0.02	7.75	35,447
AtAPI	4.69 ±0.50	2.15 ±0.63	$2.2x10^3$	1.28±0.30	1.07±0.09	$1.2x10^3$	0.43	7.75	37,749

AtAPI complements the API defect in *E. coli* strain TCM15

E. coli TCM15 is a derivative of BW30270 in which both *kdsD* and *gutQ* have been deleted (21). To support the growth of TCM15, the medium must be supplemented with both A5P (to support LPS biosynthesis) and G6P (to induce the transport system, *uhp*, which internalizes A5P). Alternatively, the *kdsD/gutQ* deficiency can be complemented by *in vivo* expression of an active

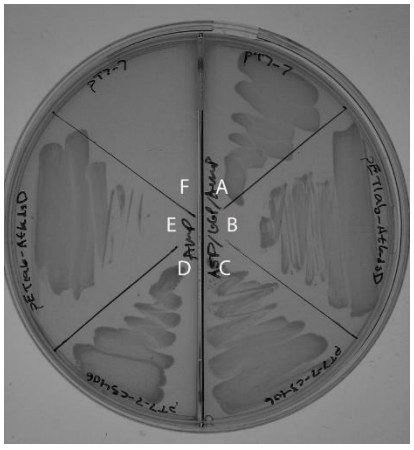

Fig. 5.3: Complementation of an A5P auxotroph on LB agar plates. (Right half) Agar plate containing LB medium supplemented with 15 µM A5P, 10 µM G6P, 0.1 mg/mL ampicillin. Wedges were streaked with: TCM15 harboring (A) pT7-7, (B) pET19b-AtAPI, and (C) pT7-7-c3406. (Left half) Agar plate containing LB medium supplemented with 0.1 mg/mL ampicillin. Wedges were streaked with: TCM15 harboring: (D) pT7-7-c3406 (API positive control), (E) pET19b-AtAPI, and (F) pT7-7.

API from a plasmid (21,30). To test the ability of AtAPI to complement an API deficiency in *E. coli*, the plasmids pET-19b-AtAPI, pT7-7 (vector control), and pT7-7-c3406 (positive control) were inserted into TCM15 cells. These plasmid vectors express their encoded proteins from a leaky T7 promoter. Equal numbers of cells were plated on LB medium containing or lacking A5P/G6P (Fig. 5.3). The presence of AtAPI complemented the API deficiency of TCM15, indicating it is an active API in these cells.

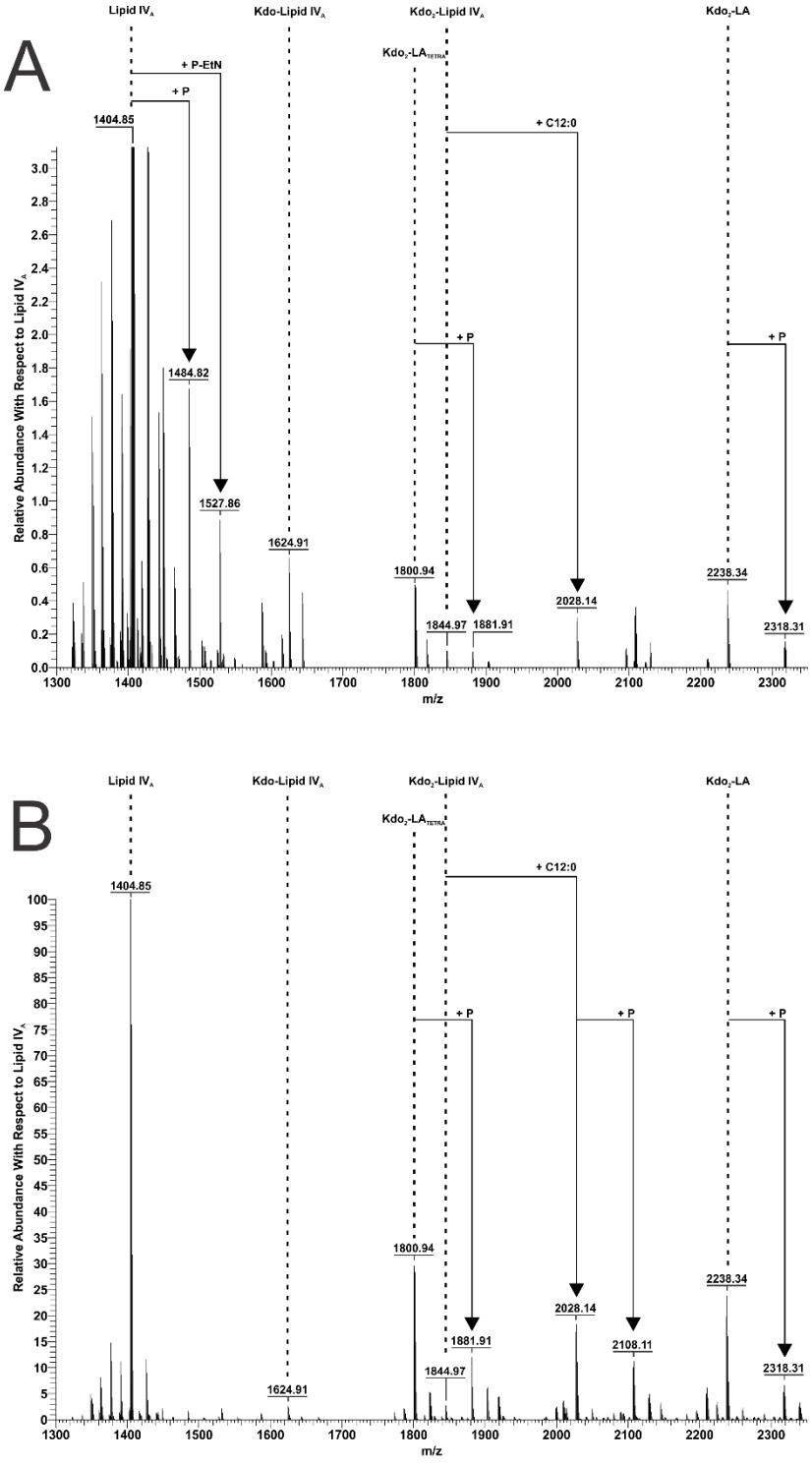

Fig. 5.4: Negative ion mode electrospray ionization FT-MS analysis of lipids. (A) from KPM56 harboring pET-16b-EcWaaA. (B) Corresponding analysis of lipids from KPM56 harboring pET-16b-AtWaaA. Scale changes between (A) and (B) for more clear visualization.

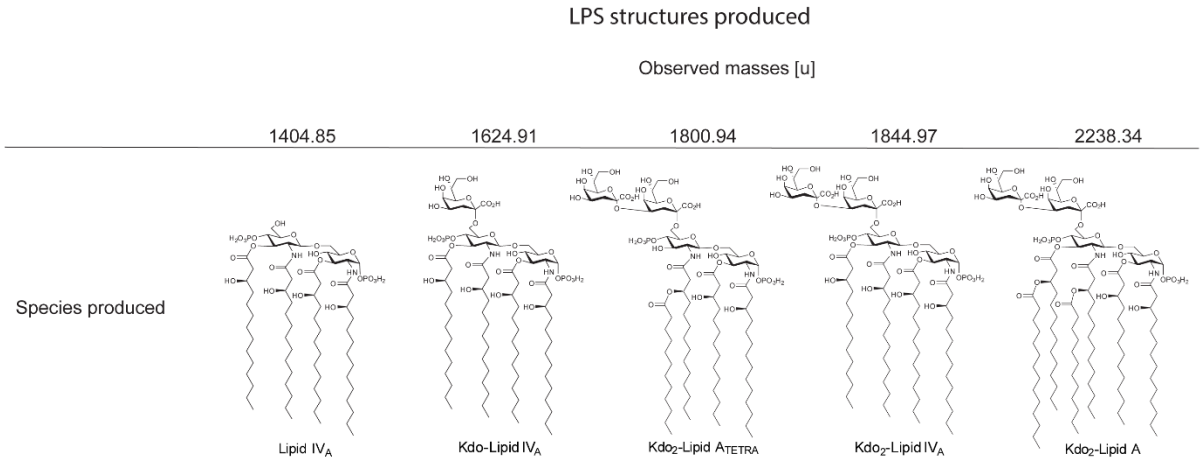

Fig.5.5: LPS structures of observed masses of related peaks shown in Fig. 5.4

AtWaaA transfer of Kdo to lipid IV_A

Electrospray ionization FT-MS analysis of LPS samples obtained from KPM56 cells expressing AtWaaA or *E. coli* WaaA (EcWaaA) showed that both samples produce species with a defined mass of 2238.34 u, which corresponds to Kdo_2-Lipid A (calculated mass 2238.34 u). The most common component in both LPS samples was lipid IV_A (exact and calculated mass 1404.85 u). Both samples also contained a deacylated species, lipid A_{TETRA} (40), with a mass of 1800.94 u (calculated mass of 1800.94 u). For the EcWaaA sample, lipid A_{TETRA} and Kdo_2-lipid A are present with abundances of ~32% and ~25% respectively compared to lipid IV_A at 100%, while the AtWaaA sample contains <1% abundance of each species (Fig. 5.4, 5.5).

To test if AtWaaA can complement a Kdo transferase deficiency in *E. coli*, the plasmids pET-16b-AtWaaAopt, pET-16b-EcWaaA, and pET-16b were inserted separately into KPM54, which lacks a Kdo transferase; grown in LB medium; and plated on MacConkey agar. Although it is known that *E. coli* strains with LPS core defects are hypersensitive to bile salts, to our knowledge the minimal LPS structure required for growth on MacConkey agar is unknown. Studies (not shown) in our laboratory have determined that *E. coli* K12 cells must minimally be able to produce Hep-$(Kdo)_2$-lipid A in order to grow on MacConkey agar. AtWaaA was not able to complement the Kdo transferase deficiency.

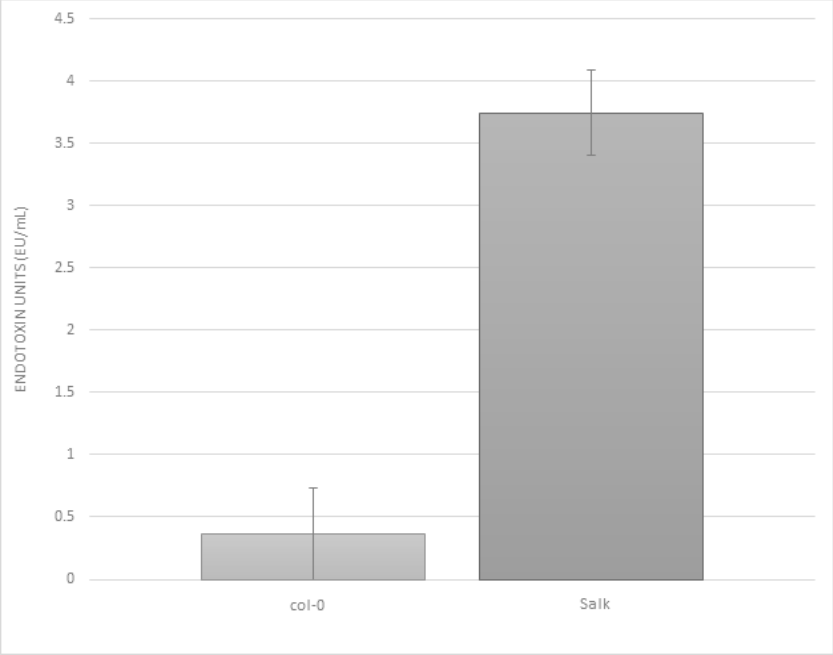

Fig. 5.6: *Limulus* amebocyte lysate assay with total lipid extracts from 10-day old *A. thaliana* seedlings of wild type (col-0) and the *atkdta-1* (SALK_035981) mutant.

Limulus Amebocyte Lysate (LAL) test of total lipids from Col-0 and atkdta-1

To determine the number of endotoxin units present in lipid samples extracted from Col-0 and *atkdta-1* an LAL assay was used. It is important to note that samples of standards of synthetic Lipid IV_A, Kdo_2-Lipid A (Avanti lipids), and *E. coli* 0111:B4 endotoxin give positive results in the LAL assay. Tests of total lipid extractions from Col-0 and *atkdta-1* (Fig. 5.6) showed lipids extracted from Col-0 to have 0.009 EU/mL, while *atkdta-1* had a calculated value of 3.661 EU/mL.

Discussion

Kdo is an integral component of both the LPS of Gram-negative bacteria and the RG-II of plant cell walls. *E. coli* strains lacking Kdo become susceptible to various antibiotics and detergents, while *A. thaliana* mutants deficient in Kdo are impaired in pollen tube elongation (5,11). In the first step of Kdo biosynthesis, Ru5P is converted to A5P in a reaction catalyzed by API.

APIs from Gram-negative bacteria have been extensively investigated (19-23); however, little is known about APIs outside Gram-negative bacteria. API null mutants in *A. thaliana* are nonviable, as RG-II appears to play an important role in primary cell wall formation and Kdo is required for proper RG-II formation (17,18). In this study, we showed that the putative API of *A. thaliana*, encoded by At3g54690, is a physiologically relevant API. BLASTP searches of the *A. thaliana* proteome with the sequences of c3406, a SIS-domain API, and *E. coli* KdsD, a full-length API, identified AtAPI as a probable API. Like most APIs characterized to date, AtAPI appears to be specific for the interconversion of Ru5P and A5P. Its pH optimum for this isomerization, 7.75, is similar to that of the bacterial APIs. Like the APIs KdsD, KpsF, and GutQ of *E. coli*, AtAPI contains a sugar isomerase domain and a tandem CBS domain and is a tetramer in solution. The kinetic

profile of AtAPI shows that it has a lower *kcat*/*Km* than that of the *E. coli* APIs. However, the lower enzymatic efficiency of AtAPI is still sufficient to complement the API-deficient bacterial strain *E. coli* TCM15 (Fig. 5.3). These findings suggest that AtAPI can serve as an API, and therefore as a source of A5P, in *A. thaliana.*

The putative Kdo transferase from *Arabidopsis thaliana,* AtWaaA*,* is predicted to have the same GT-B type fold as that found in the WaaAs of Gram-negative bacteria. The sequence of AtWaaA shares sequence identities of 36% and 27% (E value $1x10^{-63}$ and $6x10^{-29}$) with the Kdo transferases from *E. coli* and *Aquifex aeolicus*, respectively. In addition to a high degree of sequence identity, AtWaaA retains the highly conserved CMP-binding residues V277, P278, R279, H280 (V210, P211, R212, H213 in *A. aeolicus*) and the highly conserved putative Kdo-binding residues G345, H346, N347, and E350 (G271, H272, N273, E276 in *A. aeolicus* (26)) found in the C-terminal domain, which is known to bind the CMP-Kdo activated sugar nucleotide.

Previous reports have suggested that AtWaaA is unlikely to be involved in the transfer of Kdo to RG-II (11,17). In this study, we explored the possibility that AtWaaA may transfer activated Kdo to a lipid A-like molecule. Samples of KPM56 expressing AtWaaA were shown to contain Kdo_2-lipid A, lipid A_{TETRA}, and lipid IV_A (Fig. 5.4, 5.5). However, KPM56 expressing AtWaaA showed a relative abundance of less than 1% Kdo_2-lipid A, compared with 100% lipid IV_A, while KPM56 expressing EcWaaA showed a relative abundance of ~25% Kdo_2-lipid A compared with 100% lipid IV_A. We cannot, however, rule out the possibility that the low relative abundance of Kdo-glycosylated lipid IV_A stems from factors like a low membrane localization of AtWaaA in *E. coli*, which could be due to differences in the N-terminal membrane spanning region of AtWaaA compared with that of EcWaaA.

Having determined that AtWaaA, like *E. coli* WaaA. can transfer two Kdo residues to lipid IV_A in KPM56, we sought evidence that AtWaaA could rescue the Kdo transferase deficiency in KPM54. KPM54 differs from KPM56 in that KPM54 contains the heptosyl transferase (WaaC) necessary to elaborate LPS beyond Kdo_2-lipid A (26). However, when KPM54 containing pET-16b, pET-16b-AtWaaAopt, or pET-16b-EcWaaA, was plated on MacConkey agar, only the cells harboring the pET-16b-EcWaaA vector were able to survive. Therefore, it appears that AtWaaA does not have sufficient Kdo transferase activity to support growth on MacConkey agar and complement the Kdo transferase deficiency of KPM54.

The results presented here demonstrate that AtAPI is capable of serving as a physiologically relevant API for the production of Kdo in *A. thaliana*. The observation that AtWaaA can add two Kdo residues to lipid IV_A when grown in a Kdo transferase-deficient *E. coli* strain demonstrates that the enzyme exhibits bifunctional transferase activity. This result led us to speculate that the Kdo produced within *A. thaliana* is incorporated into both RG-II, as previously known, and a yet to be identified lipid A-like molecule.

Acknowledgements and Chapter Contributions

The authors would like to thank Brigitte Kunz at Research Center Borstel for her technical assistance with mass spectrometry of LPS samples, Tod Holler for help in editing the manuscript, and the *Arabidopsis* Biological Resource Center for the *A. thaliana* seeds and genes used in this study.

Cech, D.L. conducted most of the experiments and analyzed the results in addition to writing the majority of the manuscript. Pratt, A.C. purified LPS samples of KPM56, analyzed the experiments shown in figure 4 and 5, and wrote the experimental for those sections. Madak, J.T.

performed the synthesis of Ru5P and wrote the experimental for that section. Schwudke, D. provided technical assistance for generation of data for figures 4 and 5. Mamat, U. generated KPM54 and wrote the experimental for that section of the manuscript along with collaborating for figures 4 and 5. Woodard, R.W. conceived and coordinated the study along with contributing to the editing of the manuscript. All authors analyzed the results and approved the final version of the manuscript.

References

1. Becker, B., Lommerse, J. P. M., Melkonian, M., Kamerling, J. P., and Vliegenthart, J. F. G. (1995) The Structure of an Acidic Trisaccharide Component from a Cell-Wall Polysaccharide Preparation of the Green-Alga *Tetraselmis-Striata Butcher*. *Carbohydrate Research* **267**, 313-321

2. York, W. S., Darvill, A. G., McNeil, M., and Albersheim, P. (1985) 3-Deoxy-D-Manno-2-Octulosonic Acid (KDO) is a Component of Rhamnogalacturonan II, A Pectic Polysaccaride in the Primary Cell Walls of Plants. *Carbohydr. Res.* **138**, 109-126

3. Raetz, C. R. H. (1990) Biochemistry of Endotoxins. *Annual Review of Biochemistry* **59**, 129-170

4. Rick, P. D. (1987) Lipopolysaccharide biosynthesis, in *Escherichia coli* and *Salmonella typhimurium*, *ASM Publications*, American Society of Microbiology, Washington, D.C.

5. Meredith, T. C., Aggarwal, P., Mamal, U., Lindner, B., and Woodard, R. W. (2006) Redefining the requisite lipopolysaccharide structure in *Escherichia coli*. *ACS chemical biology* **1**, 33-42

6. Kropinski, A. M. B., Chan, L., and Milazzo, F. H. (1978) Susceptibility of Lipopolysaccharide-Defective Mutants of *Pseudomonas-Aeruginosa Strain Pao* to Dyes, Detergents, and Antibiotics. *Antimicrob Agents Chemother* **13**, 494-499

7. Roantree, R. J., Kuo, T. T., and MacPhee, D. G. (1977) The effect of defined lipopolysaccharide core defects upon antibiotic resistances of *Salmonella typhimurium*. *J Gen Microbiol* **103**, 223-234

8. Raetz, C. R. H., and Whitfield, C. (2002) Lipopolysaccharide Endotoxins. *Annual Review of Biochemistry* **71**, 635-700

9. Delmas, F., Seveno, M., Northey, J. G., Hernould, M., Lerouge, P., McCourt, P., and Chevalier, C. (2008) The synthesis of the rhamnogalacturonan II component 3-deoxy-D-manno-2-octulosonic acid (Kdo) is required for pollen tube growth and elongation. *Journal of experimental botany* **59**, 2639-2647

10. O'Neill, M. A., Ishii, T., Albersheim, P., and Darvill, A. G. (2004) Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annual review of plant biology* **55**, 109-139

11. Dumont, M., Lehner, A., Bouton, S., Kiefer-Meyer, M. C., Voxeur, A., Pelloux, J., Lerouge, P., and Mollet, J. C. (2014) The cell wall pectic polymer rhamnogalacturonan-II is

required for proper pollen tube elongation: implications of a putative sialyltransferase-like protein. *Annals of botany* **114**, 1177-1188

12. Kobayashi, M., Kouzu, N., Inami, A., Toyooka, K., Konishi, Y., Matsuoka, K., and Matoh, T. (2011) Characterization of *Arabidopsis* CTP:3-Deoxy-D-manno-2-Octulosonate Cytidylyltransferase (CMP-KDO synthetase), the Enzyme that Activates KDO During Rhamnogalacturonan II Biosynthesis. *Plant and Cell Physiology* **52**, 1832-1843

13. Johnson, M. A., and Lord, E. (2006) Extracellular Guidance Cues and Intracellular Signaling Pathways that Direct Pollen Tube Growth. **3**, 223-242

14. Palanivelu, R., and Tsukamoto, T. (2012) Pathfinding in angiosperm reproduction: pollen tube guidance by pistils ensures successful double fertilization. *Wiley Interdiscip Rev Dev Biol* **1**, 96-113

15. Smyth, K. M., and Marchant, A. (2013) Conservation of the 2-keto-3-deoxymanno-octulosonic acid (Kdo) biosynthesis pathway between plants and bacteria. *Carbohydrate research* **380C**, 70-75

16. Wu, J., Patel, M. A., Sundaram, A. K., and Woodard, R. W. (2004) Functional and biochemical characterization of a recombinant *Arabidopsis thaliana* 3-deoxy-D-manno-octulosonate 8-phosphate synthase. *The Biochemical journal* **381**, 185-193

17. Séveno, M., Séveno-Carpentier, E., Voxeur, A., Menu-Bouaouiche, L., Rihouey, C., Delmas, F., Chevalier, C., Driouich, A., and Lerouge, P. (2010) Characterization of a putative 3-deoxy-D-manno-2-octulosonic acid (Kdo) transferase gene from *Arabidopsis thaliana*. *Glycobiology* **20**, 617-628

18. Lalanne, E., Michaelidis, C., Moore, J. M., Gagliano, W., Johnson, A., Patel, R., Howden, R., Vielle-Calzada, J. P., Grossniklaus, U., and Twell, D. (2004) Analysis of transposon insertion mutants highlights the diversity of mechanisms underlying male progamic development in *Arabidopsis*. *Genetics* **167**, 1975-1986

19. Meredith, T. C., and Woodard, R. W. (2003) *Escherichia coli* YrbH Is a D-Arabinose 5-Phosphate Isomerase. *J. Biol. Chem.* **278**, 32771-32777

20. Meredith, Timothy C., and Woodard, Ronald W. (2006) Characterization of *Escherichia coli* D-arabinose 5-phosphate isomerase encoded by kpsF: implications for group 2 capsule biosynthesis. *Biochemical Journal* **395**, 427-432

21. Meredith, T. C., and Woodard, R. W. (2005) Identification of GutQ from *Escherichia coli* as a D-Arabinose 5-Phosphate Isomerase. *J. Bacteriol.* **187**, 6936-6942

22. Mosberg, J., Yep, A., Meredith, T., Smith, S., Wang, P.-F., Holler, T., Mobley, H., and Woodard, R. (2011) A unique arabinose 5-phosphate isomerase found within a genomic island associated with the uropathogenicity of *Escherichia coli* CFT073. *J. Bacteriol.* **193**, 2981-2988

23. Cech, D., Wang, P. F., Holler, T. P., and Woodard, R. W. (2014) Analysis of the arabinose-5-phosphate isomerase of *Bacteroides fragilis* provides insight into regulation of single-domain arabinose phosphate isomerases. *J Bacteriol* **196**, 2861-2868

24. Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J., and Swanson, R. V. (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**, 353-358

25. Li, C., Guan, Z., Liu, D., and Raetz, C. R. H. (2011) Pathway for lipid A biosynthesis in *Arabidopsis thaliana* resembling that of *Escherichia coli*. *PNAS* **108**, 11387-11392

26. Mamat, U., Schmidt, H., Munoz, E., Lindner, B., Fukase, K., Hanuszkiewicz, A., Wu, J., Meredith, T., Woodard, R., Hilgenfeld, R., Mesters, J., and Holst, O. (2009) WaaA of the hyperthermophilic bacterium *Aquifex aeolicus* is a monofunctional 3-deoxy-D-manno-oct-2-ulosonic acid transferase involved in lipopolysaccharide biosynthesis. *The Journal of biological chemistry* **284**, 22248-22262

27. Tabor, S. (1985) A Bacteriophage T7 RNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes. *PNAS* **82**, 1074-1078

28. Yamada, K., Lim, J., Dale, J. M., Chen, H., Shinn, P., Palm, C. J., Southwick, A. M., Wu, H. C., Kim, C., Nguyen, M., Pham, P., Cheuk, R., Karlin-Newmann, G., Liu, S. X., Lam, B., Sakano, H., Wu, T., Yu, G., Miranda, M., Quach, H. L., Tripp, M., Chang, C. H., Lee, J. M., Toriumi, M., Chan, M. M., Tang, C. C., Onodera, C. S., Deng, J. M., Akiyama, K., Ansari, Y., Arakawa, T., Banh, J., Banno, F., Bowser, L., Brooks, S., Carninci, P., Chao, Q., Choy, N., Enju, A., Goldsmith, A. D., Gurjal, M., Hansen, N. F., Hayashizaki, Y., Johnson-Hopson, C., Hsuan, V. W., Iida, K., Karnes, M., Khan, S., Koesema, E., Ishida, J., Jiang, P. X., Jones, T., Kawai, J., Kamiya, A., Meyers, C., Nakajima, M., Narusaka, M., Seki, M., Sakurai, T., Satou, M., Tamse, R., Vaysberg, M., Wallender, E. K., Wong, C., Yamamura, Y., Yuan, S., Shinozaki, K., Davis, R. W., Theologis, A., and Ecker, J. R. (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* **302**, 842-846

29. Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *PNAS* **97**, 6640-6645

30. Eidels, L., Rick, P. D., Stimler, N. P., and Osborn, M. J. (1974) Transport of D-arabinose-5-phosphate and D-sedoheptulose-7-phosphate by the hexose phosphate transport system of *Salmonella typhimurium*. *J Bacteriol* **119**, 138-143

31. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119

32. Pontremoli, S. R., and Mangiarotti, G. (1962) A Simple Method for the Preparation of D-Ribulose 5-Phsophate. *J. Biol. Chem.* **237**, 643-645

33. Dische, Z., Borenfreund, E. (1951) A New Spectrophotometric method for the Detection and Determination of Keto Sugars and Trioses. *J Biol Chem* **192**, 583-587

34. Studier, F. W., and Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology* **189**, 113-130

35. Galanos, C., Luderitz, O., and Westphal, O. (1969) A new method for the extraction of R lipopolysaccharides. *Eur J Biochem* **9**, 245-249

36. MacConkey, A. T. (1908) Bile Salt Media and their advantages in some Bacteriological Examinations. *Journal of Hygiene* **8**, 322

37. Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., Stevenson, D. K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C. C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D. E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W. L., Berry, C. C., and Ecker, J. R. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**, 653-657

38. Bligh, E. G., and Dyer, W. J. (1959) A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* **37**, 911-917

39. Dulbecco, R. (1954) Plaque Formation and Isolation of Pure Lines with Poliomyelitis Viruses. *Journal of Experimental Medicine* **99**, 167-182

40. Zahringer, U., Salvetzki, R., Wagner, F., Lindner, B., and Ulmer, A. J. (2001) Structural and biological characterisation of a novel tetra-acyl lipid A from *Escherichia coli* F515 lipopolysaccharide acting as endotoxin antagonist in human monocytes. *Journal of Endotoxin Research* **7**, 133-146

Chapter VI

Structure of *E. coli* c3406 in Complex With A5P and Insight Into the Mechanism of D-Arabinose-5-Phosphate Isomerase.

Summary

The lipopolysaccharide biosynthetic pathway is an underutilized target for novel antimicrobials to combat multidrug resistant Gram-negative bacteria. A key enzyme in this pathway is D-arabinose-5-phosphate isomerase (API), which catalyzes the isomerization of D-ribulose-5-phosphate and D-arabinose-5-phosphate. To better understand APIs, the sugar isomerase domain protein c3406 API was crystallized at 1.80 Å and in complex with A5P at 1.62 Å resolution. Crystals of c3406 API were soaked with D-arabinose-5-phosphate, which appears bound in the active site interacting with active site residues K53, H82, and H187. Mutation of these residues in both c3406 API and a traditional two domain API, *E. coli* KdsD, resulted in mutants comparable in active site functionality between the two enzymes, indicating that results from SIS domain APIs are relevant to traditional APIs. Additionally, a mechanism for isomerization performed by c3406 API via a *cis*-enediol is proposed.

Introduction

Lipopolysaccharide (LPS) is important to the integrity and permeability of the outer membrane of Gram-negative bacteria making its presence essential to bacterial viability during infection (1). One conserved component of the LPS is the unique sugar 3-deoxy-D-*manno*-octulosonate (Kdo), which connects the core oligosaccharide to lipid A. Kdo is currently only known to exist in Gram-negative bacteria and in some higher order plants (1-3). Although loss of the Kdo biosynthesis machinery can be overcome in the laboratory by suppressor mutations or unique growth conditions, it is essential when bacteria colonize a host and therefore Kdo biosynthesis remains a target for the design of novel antibacterials (1,4).

The biosynthesis of activated Kdo, CMP-Kdo, is accomplished in four steps starting with D-ribulose-5-phosphate (Ru5P), an end-product of the pentose phosphate pathway (5). The first step is the isomerization of Ru5P to D-arabinose-5-phosphate (A5P) catalyzed by arabinose-5-phosphate isomerase (API, E.C. 5.3.1.13). The loss of API genes is lethal and must be circumvented by addition of A5P in the media or a *yhjD* suppressor mutation which results in a hypersensitivity to antibiotics (1,6). A better understanding of the architecture of the API active site, including the identity and location of key catalytic residues, would accelerate attempts to design novel inhibitors of this enzyme.

Several APIs have been identified which fall into one of two categories, traditional APIs and SIS-domain APIs (6-10). Traditional APIs contain two domains: one sugar isomerase (SIS) domain and a tandem cystathionine beta-synthase (CBS) domain. Despite prolonged interest in APIs as targets for antibacterial therapy, attempts to acquire a structure of full-length *E. coli* APIs (KdsD, GutQ, or KpsF) have been unsuccessful. However, there have been crystal structures of

API fragments reported. Crystal structures of the CBS domains of KdsD from *E. coli* CFT073 (PDB ID: 3FNA) and from *Klebsiella pneumoniae* (PDB ID: 3K2V) are available. Two X-ray crystal structures of SIS domain API exist; a K59A inactive mutant of *E. coli* KdsD (PDB ID: 2XHZ) (11) and an API SIS-domain homolog from *Bacteroides fragilis* (BfAPI, PDB ID: 3ETN) (12). The structure of BfAPI contains the inhibitor CMP-Kdo, the end product of the Kdo biosynthetic pathway, bound in the active site (12). However, BfAPI has relaxed substrate specificity compared to all currently characterized APIs and may not be the best model for active site studies (7).

To probe the active site of APIs and identify catalytic active site residues, it would be helpful to have a structure of an API with substrate bound in the active site. We report herein X-ray crystal structures of the SIS-domain API c3406 with and without ligand. The ligand complexed stucture contains A5P bound in the active site. The uncomplexed structure of c3406 API and was determined at 1.80 Å resolution. These structures provide insight into active site residues, key active site interactions, and will provide valuable information for the design of inhibitors of APIs.

Materials and Methods

Crystallization of recombinant *c3406 API*

Recombinant c3406 API was overexpressed and purified as previously described (10). Crystals were grown by hanging drop vapor diffusion method at 21 °C. The hanging drops contained a mixture of equal parts protein solution (12 mg/mL c3406, 20 mM Tris-HCl pH 8.0, 1 mM DTT), and reservoir solution (0.1 M sodium citrate tribasic pH 5.6, 0.5 M NaCl, 2% v/v ethylene imine polymer). The hanging drops were allowed to equilibrate with 0.5 mL of reservoir solution. The crystals grew as colorless rods.

The crystals to be bound with ligand were soaked overnight in fresh hanging drop solution containing 0.1 M sodium citrate tribasic pH 5.6, 0.5 M NaCl, 2% v/v ethylene imine polymer, 5 mM A5P over 0.5 mL of reservoir solution. Crystals (soaked and unsoaked) were flash frozen at -180 °C.

Data collection and structure determination

A diffraction data set from a single crystal of c3406 API (approximately 0.1 x 0.2 x 0.5 mm in size) complexed with and without A5P were collected at the APS beamline 21 ID-D (LS-CAT) at Argonne National Laboratory. The crystals belong to space group I222 (Table 6.1). The data sets were processed with HKL2000 (13).

The structures of c3406 API were determined by molecular replacement using the previously reported structure of an SIS-domain sugar-phosphate isomerase (PBD ID: 3FXA), which shares 33% identity with c3406, as a search model. The structures were first refined in BUSTER before being iteratively refined and rebuilt utilizing REFMAC and COOT, respectively (14,15). The omit map of A5P was created by deletion of the ligand from the complexed structure followed by refinement in REFMAC (14). The omit map was calculated with the coefficients F_o-F_c and contoured at 4 σ to model the complexed A5P molecule. The crystal structure coordinates and structure factors were deposited in the Protein Data Bank with the PDB accession number 5UQI and 5VHU.

Table 6.1: Data collection and refinement statistics.

	c3406 (5UQI)	c3406 (5VHU)
Data collection		
Space group	I222	I222
Cell dimensions		
***a*, *b*, *c* (Å)**	70.89, 71.83, 76.85	71.83, 71.49, 77.23
***α*, *β*, *γ* (°)**	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution (Å)	17.96-1.62 (1.66-1.63)	52.60-1.80 (1.85-1.80)
Rsym	0.042 (0.184)	0.072 (0.191)
Completeness (%)	99.3 (96.6)	99.85
Redundancy	6.7 (6.5)	14.7 (14.6)
***I*/σ(*I*)**	6.58 (at 1.63Å)	10.7 (at 1.80Å)
Refinement		
Resolution (Å)	17.96-1.62	52.60-1.80
No. reflections	24950	17785
R_{work} / R_{free}	0.179 / 0.201	0.175 / 0.208
No. atoms		
Protein	1440	1405
Ligand (A5P/SO_4)	14	5
Water	106	85
Wilson *B* factor	19.5	13.9
Average *B*, all atoms	23.0	14.96
Average *B*, ligand (A5P/SO_4)	33.01	28.48
RMS deviations		
Bond lengths (Å)	0.018	0.019
Bond angles (°)	1.997	1.827

Site directed mutagenesis, expression, and purification of c3406 mutants

Mutation primers were designed utilizing the central overlapping primer method (16). Plasmid DNA template (pT7-7-c3406) was added to a PCR cocktail containing 1 × *Pfu*Turbo buffer, 200 µM of each deoxynucleoside triphosphate (dNTP), 3 µM of both forward and reverse mutation primers (Table 6.2), and 2.5 U *Pfu*Turbo DNA polymerase. The reaction mixtures were incubated with standard site-directed mutagenesis parameters. Following digestion with DpnI for 1 h at 37°C each reaction mixture was purified via extraction from an agarose gel. The resulting plasmids (Table 6.2) were transformed into *E. coli* TOP10 chemically competent cells. DNA sequencing of the resulting plasmids confirmed that each mutation had occurred. Subsequent site-directed mutagenesis to obtain additional mutants (double and triple) was carried out with the resulting plasmids in a similar manner. The resulting plasmids were

transformed into competent *E. coli* BL21 (DE3) cells. Protein was expressed and purified as described above.

Table 6.2: Stains, plasmids, and primers used in this study.

Item	Description	Source
E. coli		
TOP10	F⁻ *mcrA* Δ(*mrr*-*hsdRMS*-*mcrBC*) F80*lacZ*ΔM15 Δ*lacX*74 *recA1 araD*139 Δ(ara-leu)7697 *galU galK rpsL* (StrR) endA1 *nupG*	Invitrogen
BL21 (DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS*	NEB
Plasmids		
pT7-7	T7 Expression vector, Amp^R	Ref. (17)
pT7-7-c3406	*E. coli* CFT073 c3406 inserted into NdeI/BamHI of pT7-7, Amp^R	Ref. (10)
pT7-7-c3406-K53A	*E. coli* CFT073 c3406 K53A inserted into NdeI/BamHI of pT7-7, Amp^R	This study
pT7-7-c3406-H82A	*E. coli* CFT073 c3406 H82A inserted into NdeI/BamHI of pT7-7, Amp^R	This study
pT7-7-c3406-H187A	*E. coli* CFT073 c3406 H187A inserted into NdeI/BamHI of pT7-7, Amp^R	This study
pT7-7-c3406-K53A-H82A	*E. coli* CFT073 c3406 K53A H82A inserted into NdeI/BamHI of pT7-7, Amp^R	This study
pT7-7-c3406-K53A-H187A	*E. coli* CFT073 c3406 K53A H187A inserted into NdeI/BamHI of pT7-7, Amp^R	This study
pT7-7-c3406-H82A-H187A	*E. coli* CFT073 c3406 H82A H187A inserted into NdeI/BamHI of pT7-7, Amp^R	This study
pT7-7-c3406-K53A-H82A-H187A	*E. coli* CFT073 c3406 K53A H82A H187A inserted into NdeI/BamHI of pT7-7, Amp^R	This study
Primers		
c3406-K53A.F	GAAAAGTTGTTTTTATTGGTGTTGGTGCGTCCGGTATTATTGCCAGAAAACTC	IDT
c3406-K53A.R	CAATAATACCGGACGCACCAACACCA	IDT
c3406-H82A.F	GTTCACGGTACGGAAGCGGTAGCCGGCGACCTTGGAATGGTGGC	IDT
c3406-H82A.R	CAAGGTCGCCGGCTACCGCTTCC	IDT
c3406-H187A.F	CTCGTGCTGATTTCGGCTTATATGCTCCAGGAGGTGCACTCGGC	IDT
c3406-H187A.R	CACCTCCTGGAGCATATAAGCCGA	IDT

Equilibrium constant (K_{eq}) determinations

Mixtures containing 100 mM Bis-Tris propane buffer pH 6.5, 1 mM EDTA, 10% D_2O, 500 nM protein, and 10 mM A5P were incubated at 37°C for 48 hours which is sufficient for the enzymatic reaction to reach equilibrium (7). Reaction mixtures at equilibrium were analyzed by ^{31}P nuclear magnetic resonance (NMR) using a Varian 500 multinuclear NMR spectrometer. A sealed capillary tube with a solution of 16.67 mM phosphoric acid standard was set to a value of 0 ppm as an internal standard. Spectra were acquired using 32 scans with a 10 s relaxation time, as previously described to be greater than three times the T1 relaxation parameter for A5P and Ru5P (7).

Results and Discussion

The crystal structures of c3406 API with and without bound A5P, were solved at 1.80 Å and 1.62 Å, respectively. The asymmetric unit contains one monomer, which is part of a homotetramer. The protomer consists of residues 1-188 and adopts an α/β flavodoxin-like fold.

A omit maps calculated with the coefficients F_o-F_c and contoured at 4 σ contains electron density for a bound A5P molecule and sulfate ion in the active site of the protein (Fig 6.1).

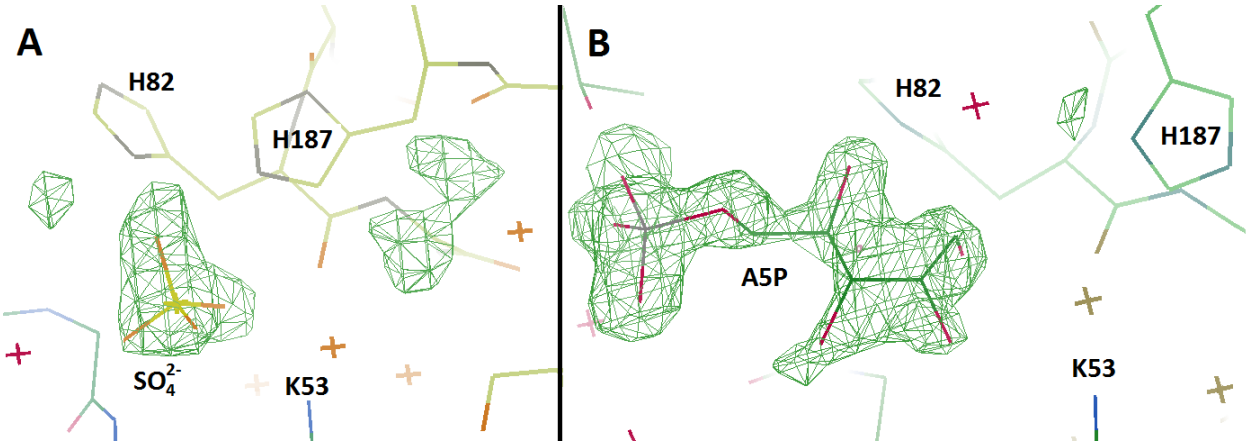

Fig. 6.1: Omit map of c3406 API. A) c3406 API with a sulfate ion bound. B) c3406 API with A5P complexed. Fobs-Fcalc density is contoured at 4 σ.

The structure of c3406 API structure is very similar to that of the *E. coli* KdsD (PDB ID: 2XHZ) (18) with a root-mean-square deviation (RMSD) of 0.902 Å when c3406 API and *E. coli* KdsD are aligned by all atoms. Both structures contain two crevices with internal twofold symmetry, resulting in two adjacent active sites in each. Each active site contains residues from three protomers (Fig. 6.2) including K53, H82, and H187 (c3406 numbering). These three predicted catalytic residues, originating from separate protomers, indicate that tetramerization of the protein is essential for catalysis. Other predicted active site residues also included E105 and E146 based on the crystal structure of the a mutant *E. coli* KdsD (18) but these residues do not appear to directly interact with A5P.

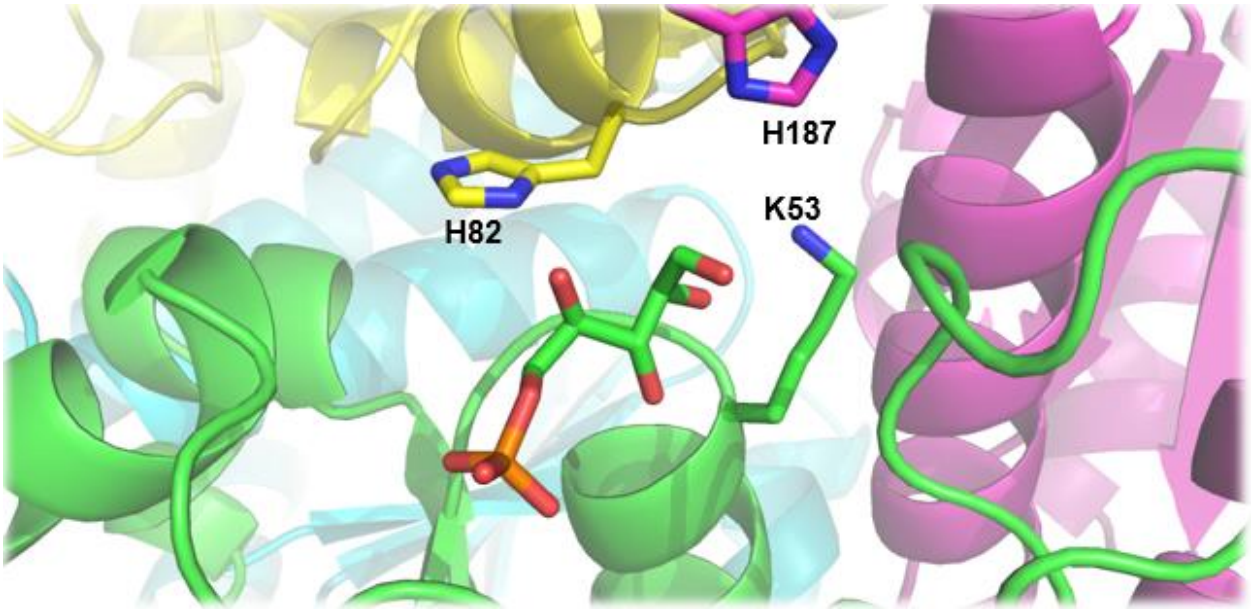

Fig. 6.2: c3406 API active site. The active site of c3406, interacting with A5P, containing residues (K53, H82, H187) which come from three protomers (shown in green, yellow, and purple).

The unbound and complexed A5P c3406 API structures align with an RMSD of 0.125 Å

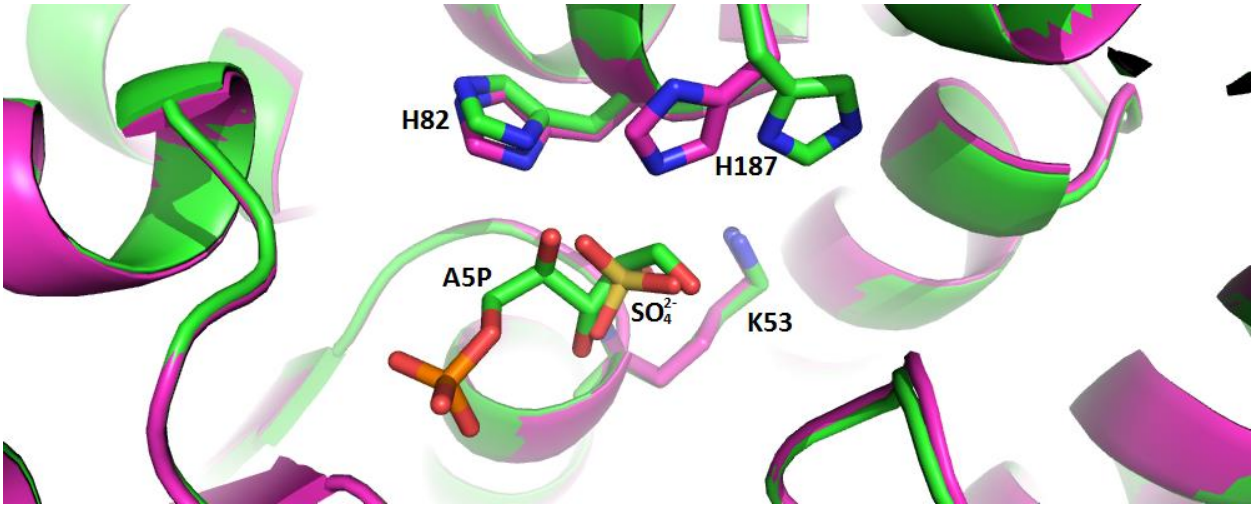

Fig. 6.3: Alignment of the c3406 API unbound structure (sulfate shown in yellow) and c3406 API in complex with A5P. c3406 API in complex shown in green, c3406 API unbound shown in purple..

when aligning all atoms (Fig. 6.3). The major difference between the two structures is the movement of H187 further away from where ligands bind in the A5P complexed structure. The

shift of H187 could be to accommodate the binding and catalysis of the A5P ligand and confirms the expectation of flexibility in the active site. Additionally, the APO c3406 API structure contains a sulfate ion in the active site, which was likely carried over from protein purification. The presence of the sulfate ion in the APO c3406 API structure could indicate that a negative charge is preferred for protein oligomerization. However, the structure of mutant *E. coli* KdsD contains no such charge in the reported structure.

It was previously proposed that the serine/threonine phosphate binding pocket in KdsD

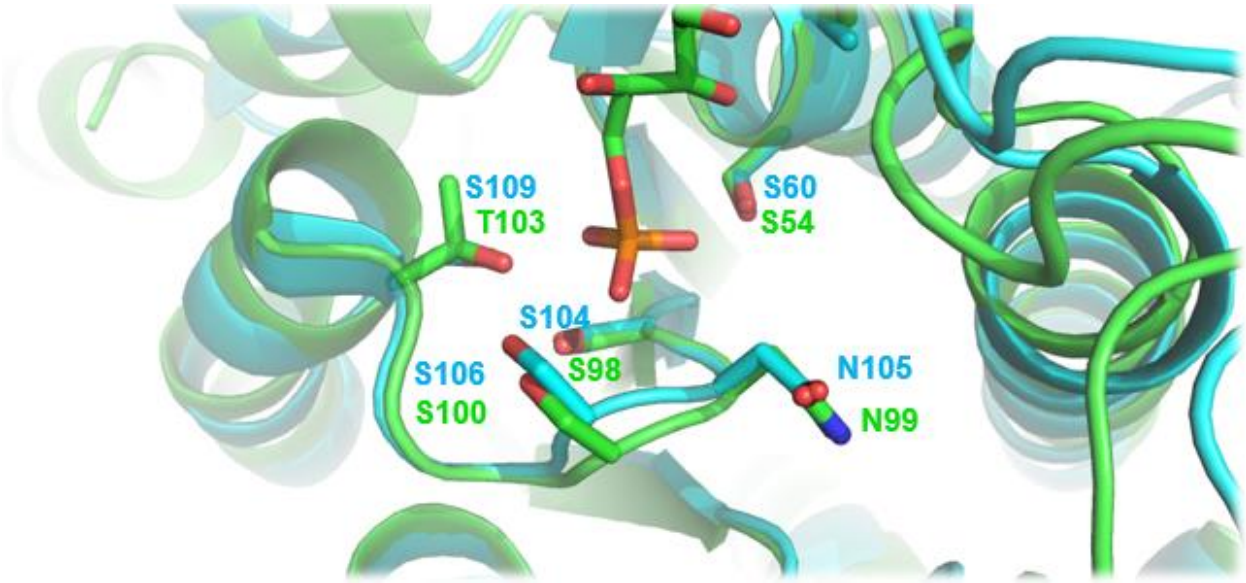

Fig. 6.4: Phosphate binding pocket alignment. Alignment of KdsD (cyan) and c3406 API (green) showing residues involved in the binding of the phosphate of A5P.

was comprised of S54, T64, N99, S103, T155, S157, T158, and T161 (c3406 numbering). An alignment (Fig 6.4) of KdsD (2XHZ) (18) with c3406 API (in complex with A5P) shows that S54, S98, S100, S109, and the backbone amide of N99 appear to be involved in binding the phosphate of the A5P. In contrast to the catalytic residues, these phosphate binding residues reside within a single protomer of the protein.

The similarity between KdsD and c3406 API structures indicates that active site mutants between c3406 and KdsD should be comparable. Mutants of KdsD (Table 6.3) show that indeed K53, H82, and H187 (c3406 numbering) are important to the catalytic activity. These data were corroborated by analogous K53A, H82A, and H193A mutants in c3406, which show comparable K_{eq} by ^{31}P NMR to *E. coli* KdsD. In addition, the pH rate profile of c3406 shows a sharp peak at an optimal pH of approximately 6.6 which suggests two histidine residues may participate in the catalytic mechanism (10).

Table 6.3: KdsD mutant characterization.

	A5P → Ru5P		Ru5P → A5P		$K_{eq\ [Ru5P]/[A5P]}$		
	k_{cat} (s^{-1})	K_M (mM)	k_{cat} (s^{-1})	K_M (mM)	Haldane Eq.	^{31}P NMR A5P → Ru5P	^{31}P NMR Ru5P →A5P
WT* (KdsD Numbering)	146 ± 8	0.6 ± 0.1	256 ± 16	0.4 ± 0.08	0.38	0.34	0.34
E85A	ND[a]	ND[a]	6.0 ± 1.1	1.3 ± 0.5	__	0.29	0.37
E85C	29 ± 1	0.9 ± 0.1	13 ± 0.9	0.6 ± 0.1	__	0.29	0.24
E85D	17 ± 1	0.6 ± 0.1	7.3 ± 0.3	0.24 ± 0.04	1.49	0.26	0.30
E85Q	ND[a]	ND[a]	13.9 ± 0.9	0.6 ± 0.1	0.93	0.29	0.30
E111A	57 ± 2	0.7 ± 0.1	72 ± 3	0.6 ± 0.1	0.68	0.29	0.36
E152A	3.4 ± 0.2	0.1 ± 0.05	ND[a]	ND[a]	__	0.54	0.40
E152C	5.7 ± 0.2	0.7 ± 0.1	ND[a]	ND[a]	__	0.35	0.39
E152D	25 ± 1	0.5 ± 0.1	58 ± 3	0.7 ± 0.2	0.60	0.42	0.67
E152Q	40 ± 1	1.1 ± 0.1	63 ± 3	0.5 ± 0.1	0.28	0.24	0.38
K59A	1.1 ± 0.1	5.3 ± 0.9	54 ± 14	12 ± 4	0.046	0.23	0.35
K66A	ND[a]	ND[a]	8.7 ± 0.9	2.1 ± 0.4	__	0.09	0.65
H62A	141 ± 4	2.1 ± 0.5	181 ± 12	1.6 ± 0.3	0.59	0.33	0.33
H82A	103 ± 6	1.7 ± 0.7	ND[a]	ND[a]	__	0.24	0.36
H88A	13.4 ± 0.8	0.9 ± 0.2	140 ± 10	2.5 ± 0.3	0.26	0.24	0.57
H193A	ND[a]	ND[a]	ND[a]	ND[a]	__	SP[c]	0.58
H193K	160 ± 1	1.3 ± 0.1	43 ± 8	0.8 ± 0.2	2.53	0.26	0.28
H193Q	ND[a]	ND[a]	5.6 ± 1.2	2.7 ± 0.8	__	SP[c]	0.76
H193N	ND[a]	ND[a]	ND[a]	ND[a]	__	SP[c]	0.53

n≥3, [a]Not Detectable *KdsD numbering

The reversible interconversion of Ru5P and A5P requires a proton to move between the first and second carbons of the carbohydrate. There are at least two potential mechanisms by which this isomerization step is thought to occur. Enzymes like xylose isomerase transfer this hydrogen via a hydride shift, which requires the presence of a metal co-factor (19,20). Studies of the APIs from *E. coli* show that they can be inhibited by divalent metals and do not require a metal co-factor for catalysis (6,8-10). The lack of a metal co-factor leads to speculation that the isomerization proceeds by the second potential mechanism; base-catalyzed movement of a proton between C-2 and C-1, with a *cis*-enediol intermediate (21).

This *cis*-enediol based mechanism, in ribose-5-phosphate isomerase (Rpi), involves deprotonation of C2 by the carboxylate of glutamic acid which forms the *cis*-enediolate intermediate (22). The negatively charged O1 is stabilized in an oxyanion hole formed by a chloride anion interaction with five backbone amide nitrogens (22), which is not present in the structure of c3406. A lysine residue has been speculated to be important for orienting the glutamic acid which is speculated to act in the proton transfer from C2 to C1. The active site lysine is speculated to play a role in polarizing the C2 hydroxyl group of R5P for proton transfer via a hydrogen bond between O2 and the positively charged side chain of lysine (22). Due to both Rpi and API sharing a common intermediate (Fig 6.5), we hypothesized that API would utilize a similar mechanism.

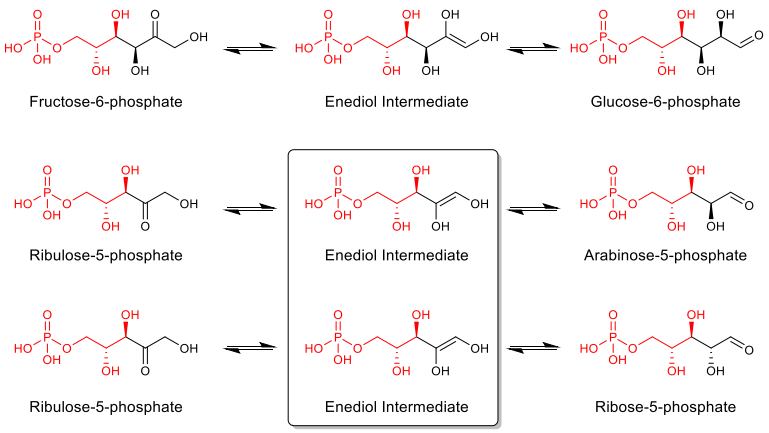

Fig. 6.5: Isomerization of select phosphosugars. F6P, Ru5P, G6P, A5P, and R5P share common stereochemistry (shown in red). F6P, G6P, A5P, and R5P are shown in linear form for clarity (these linear forms exist only as a small percentage in solution).

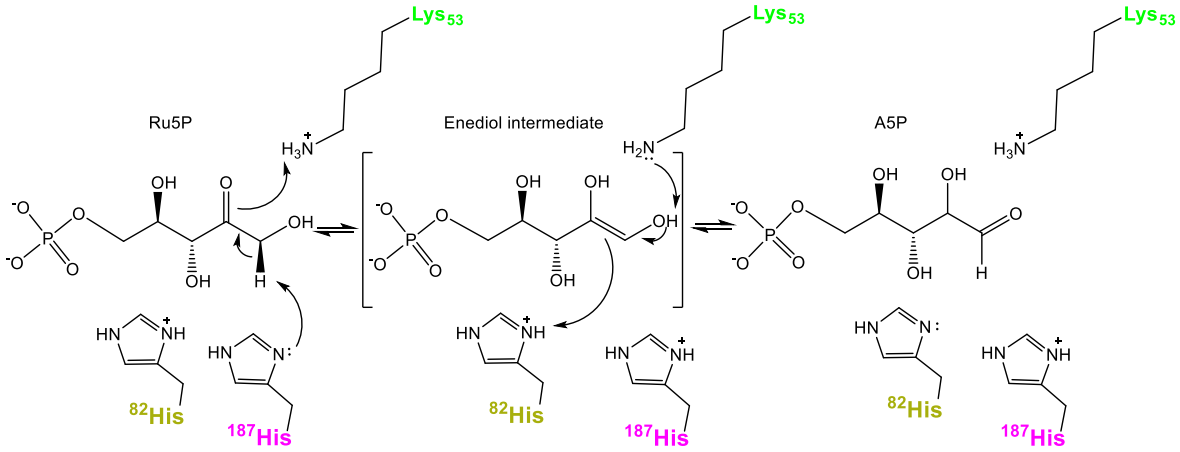

Fig. 6.6: Proposed mechanism of c3406 API.

The most likely mechanism for the isomerization of Ru5P to A5P by c3406 based on the crystal structures and active site mutations is through a *cis*-enediol intermediate involving residues K53, H82, and H187. In this putative mechanism (Fig. 6.6) deprotonation of the pro-(*S*) C1 hydrogen of Ru5P occurs via H187, followed by the protonation of the C2 carbonyl oxygen by Lys53 to form the enediol intermediate. In order for the reaction to proceed in the forward direction, Lys53 must deprotonate the hydroxy group of C1, followed by the protonation of pro-

(*R*) C2 by His82 to form A5P. The active site is likely regenerated by bulk solvent due to its solvent exposure.

This predicted mechanism is similar to the suggested mechanism of phosphoglucose isomerase (PGI), contrary to our hypothesis of Rpi similarity. An alignment of c3406 API in complex with A5P with the rabbit PGI (PDB ID: 1HOX) results in an RMSD of 3.533. The active site of rabbit PGI (Fig 6.7A) is structurally similar to that of c3406 API in complex with A5P (Fig 6.7B). It is also interesting to note that G6P, although 1 carbon longer than A5P, shares the same stereochemical configuration as A5P. This is in contrast to R5P, which differs in configuration with A5P at the C2 position (Fig. 6.5). It is conceivable that this stereochemical difference plays an important role in the catalytic mechanism.

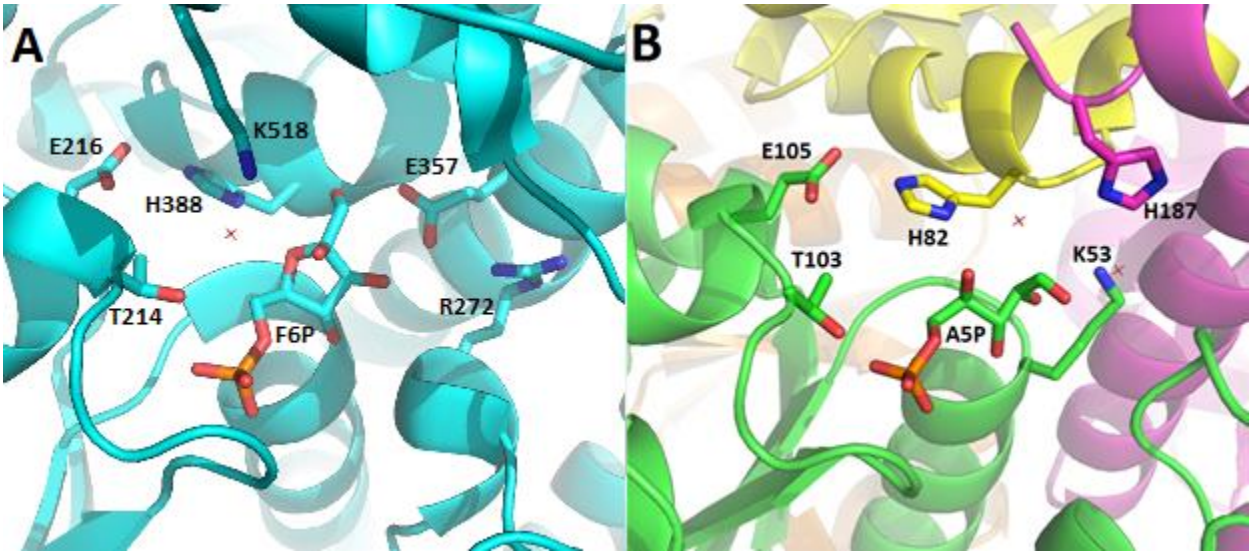

Fig. 6.7: Active site comparison of rabbit PGI and c3406 API. A) Rabbit PGI (PDB ID: 1HOX) in complex with fructose-6-phosphate. B) c3406 API in complex with A5P.

The isomerization of fructose-6-phosphate and glucose-6-phosphate in rabbit PGI involves the opening/closing of the cyclic form of F6P and G6P utilizing T214, E216, H388, K518, and a coordinated water (23). The isomerization step occurs as E357 deprotonates C1 and a

proton is donated by R272 to the C2 hydroxyl group, forming the *cis*-enediol intermediate. The reaction proceeds by transfer of the proton from E357 to C2 before the molecule undergoes ring closure (23).

In conclusion, the structure of c3406 API with A5P bound presented here is the first structure of an API SIS-domain with substrate or product bound to the enzyme. The A5P is bound in its linear form and is oriented so that its phosphate group interacts with S54, S98, S100, S109, and the backbone amide of N99. The c3406 API structures have confirmed key active site residues, namely K53, H82, and H187, and the c3406 A5P complex, along with analysis of active site mutants in *E. coli* KdsD, lead us to propose a *cis*-enediol mechanism for the isomerization of Ru5P and A5P. These studies have significantly improved our understanding of the API active site, and give valuable insight for the future development of inhibitors of the Kdo biosynthetic pathway.

Acknowledgements and Chapter Contributions

The authors would like to thank Dr. John Tesmer for his helpful suggestions and guidance in ligand and structure determination.

The manuscript was prepared by Cech, D.L. with editing assistance from Pratt, A.C., Holler, T.P., and Woodard, R.W. Cheng, M.C., Barraza, S., and Yep, A. performed *E. coli* KdsD mutant production and characterization. Markin, K. assisted with site-directed mutagenesis of c3406. Cech, D.L. performed all other experiments.

References

1. Meredith, T. C., Aggarwal, P., Mamal, U., Lindner, B., and Woodard, R. W. (2006) Redefining the requisite lipopolysaccharide structure in *Escherichia coli*. *ACS chemical biology* **1**, 33-42

2. Becker, B., Lommerse, J. P. M., Melkonian, M., Kamerling, J. P., and Vliegenthart, J. F. G. (1995) The Structure of an Acidic Trisaccharide Component from a Cell-Wall Polysaccharide Preparation of the Green-Alga *Tetraselmis-Striata Butcher*. *Carbohydrate Research* **267**, 313-321

3. York, W. S., Darvill, A. G., McNeil, M., and Albersheim, P. (1985) 3-Deoxy-D-Manno-2-Octulosonic Acid (KDO) is a Component of Rhamnogalacturonan II, A Pectic Polysaccaride in the Primary Cell Walls of Plants. *Carbohydr. Res.* **138**, 109-126

4. Mamat, U., Meredith, T. C., Aggarwal, P., Kuhl, A., Kirchhoff, P., Lindner, B., Hanuszkiewicz, A., Sun, J., Holst, O., and Woodard, R. W. (2008) Single amino acid substitutions in either YhjD or MsbA confer viability to 3-deoxy-D-manno-oct-2-ulosonic acid-depleted *Escherichia coli. Molecular Microbiology* **67**, 633-648

5. Raetz, C. R. H. (1990) Biochemistry of Endotoxins. *Annual Review of Biochemistry* **59**, 129-170

6. Meredith, T. C., and Woodard, R. W. (2005) Identification of GutQ from *Escherichia coli* as a D-Arabinose 5-Phosphate Isomerase. *J. Bacteriol.* **187**, 6936-6942

7. Cech, D., Wang, P. F., Holler, T. P., and Woodard, R. W. (2014) Analysis of the arabinose-5-phosphate isomerase of *Bacteroides fragilis* provides insight into regulation of single-domain arabinose phosphate isomerases. *J Bacteriol* **196**, 2861-2868

8. Meredith, T. C., and Woodard, R. W. (2003) *Escherichia coli* YrbH Is a D-Arabinose 5-Phosphate Isomerase. *J. Biol. Chem.* **278**, 32771-32777

9. Meredith, Timothy C., and Woodard, Ronald W. (2006) Characterization of *Escherichia coli* D-arabinose 5-phosphate isomerase encoded by kpsF: implications for group 2 capsule biosynthesis. *Biochemical Journal* **395**, 427-432

10. Mosberg, J., Yep, A., Meredith, T., Smith, S., Wang, P.-F., Holler, T., Mobley, H., and Woodard, R. (2011) A unique arabinose 5-phosphate isomerase found within a genomic island associated with the uropathogenicity of *Escherichia coli* CFT073. *J. Bacteriol.* **193**, 2981-2988

11. Gourlay, L. J., Sommaruga, S., Nardini, M., Sperandeo, P., Deho, G., Polissi, A., and Bolognesi, M. (2010) Probing the active site of the sugar isomerase domain from *E. coli* arabinose-5-phosphate isomerase via X-ray crystallography. *Protein Sci* **19**, 2430-2439

12. Chiu, H. J., Grant, J. C., Farr, C. L., Jaroszewski, L., Knuth, M. W., Miller, M. D., Elsliger, M. A., Deacon, A. M., Godzik, A., Lesley, S. A., and Wilson, I. A. (2014) Structural analysis of arabinose-5-phosphate isomerase from *Bacteroides fragilis* and functional implications. *Acta Crystallogr D Biol Crystallogr* **70**, 2640-2651

13. Otwinowski, Z., and Minor, W. (1997) [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307-326

14. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D: Biological Crystallography* **53**, 240-255

15. Emsley, P., Cowtan, K., Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography* **D60**, 2126-2132

16. Wang, H., Zhou, N., Ding, F., Li, Z., Chen, R., Han, A., and Liu, R. (2011) An efficient approach for site-directed mutagenesis using central overlapping primers. *Anal. Biochem.* **418**, 304-306

17. Tabor, S. (1985) A Bacteriophage T7 RNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes. *PNAS* **82**, 1074-1078

18. Gourlay, L. J., Sommaruga, S., Nardini, M., Sperandeo, P., Dehò, G., Polissi, A., and Bolognesi, M. (2010) Probing the active site of the sugar isomerase domain from *E. coli* arabinose-5-phosphate isomerase via X-ray crystallography. *Protein Science* **19**, 2430-2439

19. Allen, K. N., Lavie, A., Farber, G. K., Glasfeld, A., Petsko, G. A., and Ringe, D. (1994) Isotopic exchange plus substrate and inhibition kinetics of D-xylose isomerase do not support a proton-transfer mechanism. *Biochemistry* **33**, 1481-1487

20. Zheng, Y. J., Merz, K. M., Jr., and Farber, G. K. (1993) Theoretical examination of the mechanism of aldose-ketose isomerization. *Protein Eng* **6**, 479-484

21. Airoldi, C., Sommaruga, S., Merlo, S., Sperandeo, P., Cipolla, L., Polissi, A., and Nicotra, F. (2011) Targeting Bacterial Membranes: Identification of *Pseudomonas aeruginosa* D-Arabinose-5P Isomerase and NMR Characterisation of its Substrate Recognition and Binding Properties. *Chembiochem* **12**, 719-727

22. Hamada, K., Ago, H., Sugahara, M., Nodake, Y., Kuramitsu, S., and Miyano, M. (2003) Oxyanion hole-stabilized stereospecific isomerization in ribose-5-phosphate isomerase (Rpi). *J Biol Chem* **278**, 49183-49190

23. Lee, J. H., Chang, K. Z., Patel, V., and Jeffery, C. J. (2001) Crystal Structure of Rabbit Phosphoglucose Isomerase Complexed with Its Substrated-Fructose 6-Phosphate. *Biochemistry* **40**, 7799-7805

Chapter VII

Summary, Conclusions, and Future Directions

This dissertation focused on studying D-arabinose-5-phosphate (A5P) sources and potential uses, by the cell, for the resulting A5P. The D-arabinose-5-phosphate isomerases (APIs) studied herein came from a variety of sources including Gram-negative bacteria, Gram-positive bacteria, and plants. Insight into the role of APIs, their mechanism, and the role of A5P they produce has been gained through kinetic characterization, inhibition studies, mutagenesis studies, and the X-ray crystal structure of an API.

In chapter II, a putative API from the Gram-positive organism, *Listeria monocytogenes,* was characterized. The putative API, Q723E8, showed no API activity but possessed D-glucose-6-phosphate isomerase (GPI) activity. Q723E8 is a unique GPI as it is an SIS domain protein with roughly 200 amino acids while typical PGIs, such as those in *Staphyloccus aureus* and *Vibrio cholera*e, contain over 500 amino acids. Q723E8 could therefore represent a novel class of GPIs with an architecture similar to APIs.

In chapter III, the API from *Bacteroides fragilis* was characterized. The crystal structure of BfAPI, an SIS domain protein, in complex with CMP-Kdo was previously deposited in the PDB (PDB ID: 3ETN). CMP-Kdo is the end product of the Kdo biosynthetic pathway and was bound in the active site of the enzyme. Inhibition studies of the BfAPI showed that CMP-Kdo can inhibit the enzyme with a K_i of 1.91 μM *in vitro* and we speculated that CMP-Kdo could also serve as a feedback inhibitor *in vivo*. BfAPI, being the sole API in *B. fragilis*, represents the only SIS domain

API characterized to date that is responsible for being the sole source of A5P used in lipopolysaccharide production. Furthermore, BfAPI possesses an additional function, in the form of GPI activity. All other APIs characterized to date are specific to the isomerization of A5P and Ru5P.

In chapter IV, a putative API from *Clostridium tetani* was characterized. CtAPI is the first API characterized from a Gram-positive organism. The genomic context of the CtAPI gene appears to be similar to that of *gutQ* in *E. coli* K-12. Both genes are situated within an operon associated with the transport and metabolism of a sugar other than A5P. This raises the possibility of A5P being utilized to regulate the ribose metabolism operon, in which the gene encoding CtAPI is located in a similar manner to sorbitol and *gutQ* in *E. coli*. The finding of an API in a Gram-positive organism should allow for the confirmation and further study of the A5P regulation of the operon in an environment devoid of interfering A5P utilizing pathways such as LPS biosynthesis.

In chapter V, the API of *Arabidopsis thaliana* was characterized. This is the first API characterized in a plant and supports the biosynthesis of Kdo for use in rhamnogalacturonan-II (RG-II). Additionally, the possibility of Kdo being transferred to a lipid IV_A-like molecule was explored as it had previously been shown that *A. thaliana* produces lipid IV_A and Kdo. Our results showed that AtWaaA exhibits bifunctional Kdo transferase activity in an *E. coli* KPM56 system by transferring two Kdo molecules to lipid IV_A of *E. coli*. These results led us to speculate on the possibility that Kdo is not only being used in *A. thaliana* for RG-II biosynthesis, but also in an unidentified lipid A-like molecule.

In chapter VI, the sugar isomerase domain API, c3406, was probed through X-ray crystallography and mutagenesis to better understand the mechanism of APIs. The first crystal

structure of an active API with a substrate bound was presented. The active site of c3406 contains three residues (K53, H82, and H187) which are within proximity to interact with C1 and C2 of A5P. Mutagenesis studies of these residues show select mutants to be inactive, while others exhibit a 10 to 100 fold loss of activity in both c3406 and *E. coli* KdsD. A mechanism is proposed based on the crystal structure, mutagenesis studies, and pH rate profiles in which these three residues are utilized for the isomerization of Ru5P and A5P. This proposed mechanism is similar to that of GPIs and could be indicative of why BfAPI, in chapter III, showed relaxed substrate specificity.

In conclusion, the identification and characterization of APIs from a variety of sources including a Gram-positive bacterium, Gram-negative bacteria, and a plant indicate that APIs and the A5P they produce are not unique to Gram-negative bacteria, but are in fact much more widespread than previously thought. In studying these APIs, I have explored the enzymes that produce A5P for both lipopolysaccharide biosynthesis and RG-II biosynthesis. Additionally, I have speculated upon the possibility of A5P as a regulatory molecule and its use in the biosynthesis of Kdo for addition to a lipid A-like molecule in *A. thaliana.* However, the role of A5P in Gram-positive organisms and the potential biosynthesis of a lipid A-like molecule in *A. thaliana* are both poorly understood and warrant further investigation.

The determination of the X-ray crystal structure, identification of CMP-Kdo as a potential feedback API inhibitor *in vivo*, mutagenesis studies, and a proposed mechanism pave the way for a more thorough understanding of the active site of API and the development of new antimicrobials targeting the Kdo biosynthetic pathway. The validation of the proposed mechanism will aid in the design of inhibitors of APIs. The exploration of these SIS domain and full-length APIs in this thesis has been crucial in expanding our understanding of APIs and the role

of the SIS domain in this protein. This information could be complemented by further research on the interaction of CBS domains (see appendix) in full length APIs, the role they play, and the possibility of feedback inhibition by CMP-Kdo *in vivo*.

APPENDIX

Cystathionine β-Synthase Domain

Cystathionine β-synthase (CBS) domains have been found in proteins throughout all kingdoms of life (1). CBS domains are generally about 60 residues in length and usually occur in pairs. CBS domains fold into an α/β structure with internal symmetry (2). The pair of domains contains an interdomain cavity that represents a ligand-binding site (1). Many CBS domains function as regulatory units within their respective protein providing an allosteric binding site, as is speculated to be the case in APIs. The most common regulatory ligand for CBS domains are adenine nucleotides (1). The binding of these regulatory ligands can lead to either activation or inhibition of the enzyme (3,4).

Studies have been conducted in our laboratory on the CBS domains of APIs which warrant further investigation. Preliminary thermal shift assays in our laboratory have shown that CMP

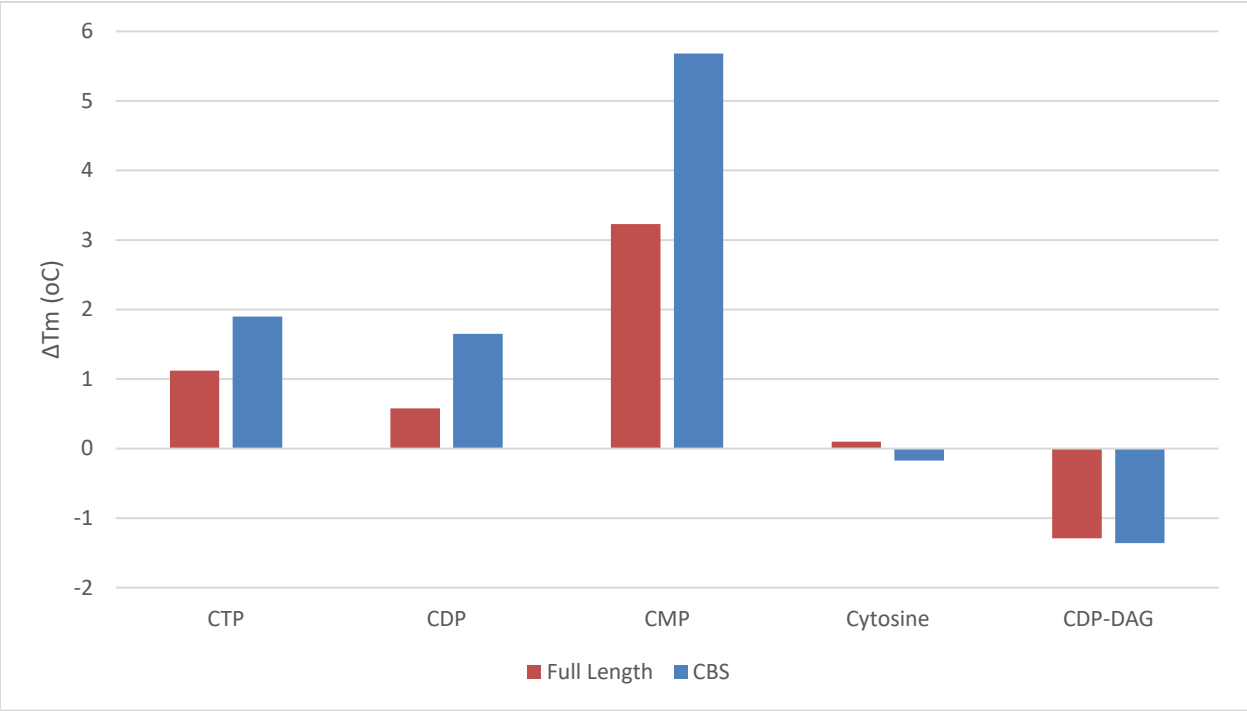

Fig. A.1: *Klebsiella* pneumoniae KpsF thermal melts. Thermal shift assay of KpKpsF comparing full-length protein and the CBS domains. Values shown are average melting temperature shifts with n=3.

binds both protein constructs containing only the CBS domain of APIs and full-length APIs separately (Figure A.1). This information, along with the *in vitro* inhibition studies of CMP-Kdo with BfAPI, lead us to speculate that CMP-Kdo could interact with both domains of full length APIs. Additional studies with longer lived CMP-Kdo analogs could help test this hypothesis.

References

1. Baykov, A. A., Tuominen, H. K., and Lahti, R. (2011) The CBS Domain: A Protein Module with an Emerging Prominent Role in Regulation. *ACS chemical biology* **6**, 1156-1163

2. Zhang, R., Evans, G., Rotella, F. J., Westbrook, E. M., Beno, D., Huberman, E., Joachimiak, A., and Collart, F. R. (1999) Characteristics and crystal structure of bacterial inosine-5'-monophosphate dehydrogenase. *Biochemistry* **38**, 4691-4700

3. Lucas, M., Encinar, J. A., Arribas, E. A., Oyenarte, I., García, I. G., Kortazar, D., Fernández, J. A., Mato, J. M., Martínez-Chantar, M. L., and Martínez-Cruz, L. A. (2010) Binding of S-Methyl-5'-Thioadenosine and S-Adenosyl-l-Methionine to Protein MJ0100 Triggers an Open-to-Closed Conformational Change in Its CBS Motif Pair. *Journal of Molecular Biology* **396**, 800-820

4. Tuominen, H., Salminen, A., Oksanen, E., Jämsen, J., Heikkilä, O., Lehtiö, L., Magretova, N. N., Goldman, A., Baykov, A. A., and Lahti, R. (2010) Crystal Structures of the CBS and DRTGG Domains of the Regulatory Region of *Clostridium perfringens* Pyrophosphatase Complexed with the Inhibitor, AMP, and Activator, Diadenosine Tetraphosphate. *Journal of Molecular Biology* **398**, 400-413