# STUDIES ON THE BIOSYNTHESIS OF MENAQUINONE (VITAMIN K) AND ON THE CATABOLISM OF PSEUDOURIDINE

A Dissertation

by

NILKAMAL MAHANTA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

| | |
|---|---|
| Chair of Committee, | Tadhg P. Begley |
| Committee Members, | Frank M. Raushel |
| | David P. Barondeau |
| | Paul Straight |
| Head of Department, | Francois P. Gabbai |

May 2015

Major Subject: Chemistry

## ABSTRACT

The present work describes studies aimed at characterizing enzymes involved in bacterial metabolic pathways using a variety of biochemical methods, analytical techniques and structural studies. The first study explains the structural and biochemical characterization of a C-glycosidase involved in pseudouridine catabolic pathway. Our studies suggested that its mechanism is significantly different from the previously reported glycosidases. The second study describes the discovery and mechanistic characterization of a radical SAM enzyme involved in a new menaquinone biosynthetic pathway. This enzyme represents a unique reaction motif in radical SAM enzymology.

Pseudouridine (Ψ), the most abundant modification in RNA, is synthesized using Ψ synthase. Recently, a pathway for the degradation of Ψ was described in *Escherichia coli*. In this pathway, Ψ is first converted to Ψ 5′-monophosphate (ΨMP) by Ψ kinase and then ΨMP is degraded by ΨMP glycosidase to uracil and ribose 5- phosphate. The structural studies on the ΨMP glycosidase and its mutant (K166A) suggested that the reaction utilizes a Lys166-substrate adduct during catalysis. Biochemical studies on ΨMP glycosidase further confirmed the existence of a lysine adduct and allowed us to identify roles for specific active site residues. ΨMP glycosidase catalyzes the cleavage of the C−C glycosidic bond through a novel ribose ring-opening mechanism. This is the first mechanistically characterized C-glycosidase.

Menaquinone (MK, vitamin $K_2$) is a lipid soluble molecule that participates in the bacterial electron transport chain. In mammalian cells, MK functions as an essential vitamin for the activation of various proteins involved in blood clotting and bone metabolism. Recently, a new pathway for the biosynthesis of this cofactor was discovered in *Streptomyces coelicolor* A3(2) in which chorismate is converted to aminofutalosine in a reaction catalyzed by MqnA and an unidentified enzyme. Here, we reconstitute the biosynthesis of aminofutalosine and demonstrate that the missing enzyme (aminofutalosine synthase, MqnE) is a radical SAM enzyme that catalyzes the addition of the adenosyl radical to the double bond of 3-[(1-carboxyvinyl) oxy] benzoic acid. This is a new reaction type in the radical SAM superfamily. The substrate analogs based mechanistic investigation suggested that MqnE catalyzes a unique radical rearrangement reaction, unprecedented in biological chemistry.

Dedicated to my parents, my sister, my uncle and his family as well as all my near and dear ones in the family…

And to my wife and my son...

## ACKNOWLEDGEMENTS

First of all, I would like to thank my research advisor, Dr. Tadhg Begley for being an outstanding mentor and for giving me the opportunity to work in his lab on some very interesting projects. His constant guidance throughout my graduate study, scientific inputs in the projects and inspiration is highly acknowledged. I would also like to thank my committee members, Dr. Frank Raushel, Dr. David Barondeau and Dr. Paul Straight, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues in the Dr. Begley laboratory and the department faculty and staff for all their support and for making my time at Texas A&M University a great experience. In particular, I would like to thank Dmytro Fedoseyenko (Dr. Begley laboratory) for all his synthetic work in the menaquinone project, Siyu Huang and Katherine Hicks (Dr. Ealick laboratory, Cornell University) for their structural work, Benjamin Philmus (Dr. Begley laboratory), Fred Zinnel and Pei-Jing Pai (Dr. Russell laboratory) for some of the mass spectrometry work in the pseudouridine glycosidase project, Cynthia Kinsland (Dr. Ealick laboratory) and Dinesh Simkhada (Dr. Begley laboratory) for some of the cloning work and Lisa Cooper for being my mentor. I would also like to thank Dr. Howard Williams (Texas A&M University) for his assistance in the NMR analysis. I would also like to thank Dr. Steve Ealick (Cornell University, Ithaca, New York) and Dr. Tohru Dairi (Hokkaido University, Japan) for their collaboration in the research work described here. I also want to extend my

gratitude to the Welch Foundation and Department of Chemistry, Texas A&M University for their financial support during the course of this study.

Finally, thanks to parents, my sister, my uncle and his family as well as my near and dear ones in the family for their constant encouragement. And last not but the least, I would like to express my deep gratitude to my wife Maina, for her patience and love and to my little son Ayaan, for making my life wonderful during my graduate study.

## NOMENCLATURE

| | |
|---|---|
| Ψ | Pseudouridine |
| ΨMP | Pseudouridine-5'monophosphate |
| R5P | Ribose-5-phosphate |
| Ura | Uracil |
| ATP | Adenosyl 5'-triphosphate |
| MK | Menaquionone |
| OSB | O-succinyl benzoate |
| NADH | β-nicotinamide adenine dinucleotide, reduced form |
| NADPH | β-nicotinamide adenine dinucleotide phosphate, reduced form |
| DHNA | 1,4-dihydroxy-2-naphthoate |
| SAM | *S*-Adenosyl-L-Methionine |
| TPP | Thiamine pyrophosphate |
| Tris | tris-(hydroxymethyl) aminomethane |
| DHFL | dehypoxanthine futalosine |
| cDHFL | cyclic dehypoxanthine futalosine |
| PEP | phosphoenolpyruvate |
| NMR | Nuclear Magnetic Resonance |
| HPLC | High Performance Liquid Chromatography |
| LC-MS | Liquid Chromatography-Mass Spectrometry |

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# CHAPTER I

# INTRODUCTION: BIOSYNTHESIS AND CATABOLISM OF PSEUDOURIDINE

## Introduction

Pseudouridine (5-ribosyluracil, Ψ), **2** was the first modified nucleoside to be discovered and is the most abundant modification in RNA, existing in tRNAs, rRNAs, snRNAs, and snoRNAs.[1] Ψ is biosynthesized post-transcriptionally from RNA uridine moieties, **1** by a class of enzymes called pseudouridine synthase (Ψ synthase), which cleaves the glycosidic C−N bond and reconnects the uracil to the ribosyl moiety at the C5 position (Figure 1.1).[2] Pseudouridine reduces the conformational flexibility of RNA because the exposed, protonated N1 atom forms strong hydrogen bonds with structured water molecules, replacing weak interactions formed by C5 in uridine. Ψ in RNA provides the appropriate geometry and distance for coordination for a water molecule between its N1-H and the 5'-phosphates of both Ψ and the preceding residue forming a hydrogen bond. This restricts base conformation and mobility of backbone 5' to the site of pseudouridylation and also enhances base stacking.[1, 3]

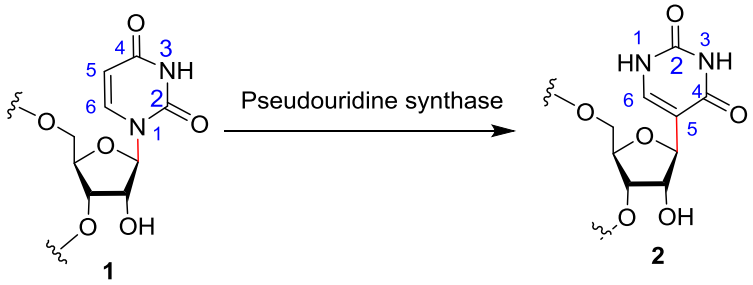



**Figure 1.1**: Biosynthesis of pseudouridine

Mammalian rRNA contains about 100 pseudouridines per ribosome and tRNAs contain an average of 3-4 of them.[1] Ψ is found in almost all tRNAs, at a notably universal position namely Ψ55, after which the T ΨC stem-loop is named (Figure 1.2). Other positions of tRNA at which it occurs in all three domains of life (archaebacterial, eubacteria and eukaryotes) as well as in organelles (such as mitochondria and chloroplasts) include the D stem and the anticodon stem and loop (ASL).[1] Ψ contributes to the stabilization of the specific structural motifs in which it occurs (Figure 2): for instance, the T ΨC loop (Ψ55), the D stem (Ψ13), the anticodon stem (features a strong closing base pair between Ψ39 and A31) and the anticodon loop (noncanonical or pseudo base-pairing between Ψ38 and residue 32).[4, 5]

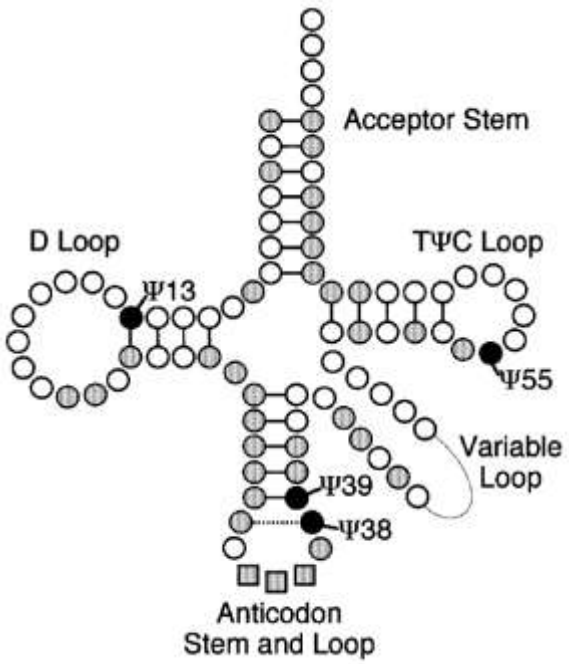



**Figure 1.2:** Distribution of Ψ residues in tRNA[4]

**Biological role of pseudouridine**

Through its unique ability to coordinate a structural water molecule via the free N1-H, Ψ exerts a subtle but significant rigidifying influence on the nearby sugar phosphate backbone in tRNA and also enhances base stacking. These effects of may underlie the biological role of most of the Ψ residues in RNA. During translation, Ψ is thought to modulate several interactions that tRNA molecules make with rRNAs and with mRNAs. The influence of modified nucleosides, like Ψ, on the local structure of the anticodon stem and loop (ASL) appears to be critical for the proper binding of tRNA to the ribosome (30S ribosomal subunit) which helps to foster correct codon-anticodon interactions.[6-10] This may increase the translational accuracy by decreasing the rate of peptide bond formation, thereby allowing more time for rejection of incorrect codon-anticodon pairs.[11] This improves the fidelity of protein biosynthesis. Ψ residues may also influence both rRNA folding and ribosome assembly.[12, 13] In the mature ribosome, Ψ probably contributes to the stabilization of local secondary or tertiary structure through RNA-RNA or RNA-protein interactions.[14] This fine tuning of the local higher order structure in rRNA could influence the speed and accuracy of decoding and proofreading during translation by fostering appropriate interactions with tRNA.[15]

It was observed that certain genetic mutants lacking specific Ψ residues in tRNA or rRNA exihibit difficulties in translation, display slow growth rates and fail to compete effectively with the wild-type strains in mixed culture.[1] In particular, normal growth is severely compromised in an *Escherichia coli* mutant deficient in a pseudouridine

synthase responsible for the formation of three closely spaced Ψ residues (Ψ1911, Ψ1915 and Ψ 1917) in the mRNA decoding region of the 23S rRNA.[16] The loss of function by homologues of a nuclear pseudouridine synthase in yeast (Cbf5p), *Drosophila* (mfl), rats (NAP57) and humans (dyskerin) leads to phenotypes ranging from severe physiological impairment to lethality.[17-21] For example, mutations in the human gene encoding dyskerin have been associated with the rare haematopoietic and malignant disorder, X-linked dyskeratosis congenita.[19, 22] Hence, these studies indicate that pseudouridylation of RNA[23] is an important event which confers a selective advantage in a biological context.

**Biosynthesis of pseudouridine**

The pseudouridine synthases (Ψ synthases) isomerize uridine to pseudouridine in RNA and fall into six families.[24-26] Despite their insignificant global sequence similarity, the families share a common protein fold with a core β sheet and constellation of active site residues as evident from their crystal structures.[27] An active site Asp is the only amino acid conserved among all Ψ synthases and it is found to be essential for catalytic activity suggesting that this family of enzymes share a common mechanism.[28, 29] The mechanism for this enzyme is proposed based on several biochemical and structural studies using substrate analogs such as 5-fluorouridine and [$F^5$U]-RNA.[28, 29] However, a recent mechanistic proposal is put forward for pseudouridine synthase TruB from *E. coli*. (Figure 1.3).[30] The N-glycosyl bond in uridine **1** cleaves forming the oxocarbenium ion, **3** and uracil anion, **4**. The active site Asp residue traps the oxocarbenium ion forming an

enzyme covalent adduct, **5** at C1' while the uracil anion undergoes an $180^0$ rotation to juxtapose C5 of pyrimidine and C1' of the ribosyl moiety. This is followed by formation of the C1'-C5 glycosidic bond by the attack from the uracil anion on to the enzyme covalent adduct, forming **6**. A final tautomerization step probably assisted by the active site Asp lead to the formation of pseudouridine **2.**

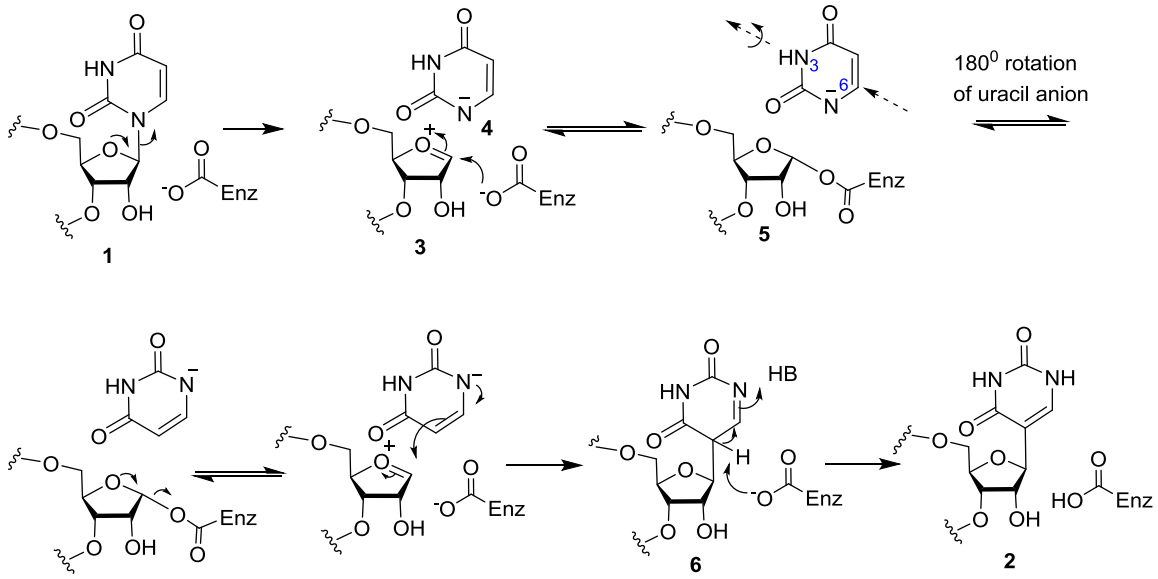
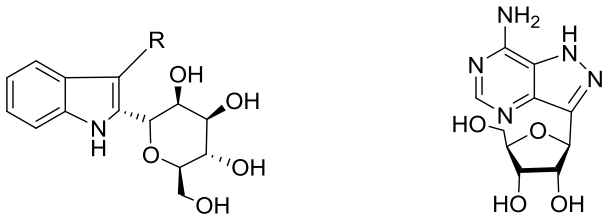



**Figure 1.3:** Mechanistic proposal for pseudouridine synthase.

**Catabolism of pseudouridine**

While Ψ biosynthesis has been studied extensively, little is known about Ψ catabolism. It is one of the few examples of molecules that have a glycosidic C-C bond. Other examples of such nucleosides include C-mannosyl residues bound to tryptophan[31] (mannosyltryptophan, **7,** Figure 1.4), the C-glucosides bound to catechol derivatives (C-

glucosylated form of microcin E492)[32] and some nucleoside analogues such as formycin A[33] (**8,** Figure 1.4). However, no enzyme that cleaves a C-C glycosidic bond is mechanistically characterized.




**7** **8**

**Figure 1.4**: Examples of other C-glycosides

It was reported that Ψ can be used by *E. coli* as a source of uracil for a strain deficient in pyrimidine synthesis (uridine auxotroph).[34] Ψ was shown to be phosphorylated in *E. coli* extracts by a kinase to pseudouridine 5'-phosphate (pseudouridine monophosphate, ΨMP) [35] which is further hydrolyzed to ribose-5-phosphate and uracil.[36] However, none of these enzymes were identified. Recently, the two enzymes responsible for Ψ degradation in *E. coli* were identified and their molecular characterization was achieved.[37] Ψ kinase (YeiC) first phosphorylates Ψ to pseudouridine 5′-monophosphate, **9** (ΨMP), with the help of adenosine triphosphate (ATP) and $Mg^{2+}$. ΨMP glycosidase (YeiN) then catalyzes the reversible cleavage of the C-C glycosidic bond to form ribose 5-phosphate, **10** (R5P) and uracil, **11** (Figure 1.5) and requires divalent metal ions such as $Mn^{2+}$ for its activity. While ΨMP glycosidase functions biologically in the cleavage direction, equilibrium strongly favors ΨMP synthesis with an equilibrium constant for

the degradation reaction of $2.3\times10^{-4}$.[37] In bacterial genomes Ψ kinase and ΨMP glycosidase are encoded by separate genes. In *E. coli* these are *yeiC* and *yeiN*, respectively, which are located in the same operon (Figure 1.6).[37] In some eukaryotes, the genes for Ψ kinase and ΨMP glycosidase are fused resulting in a bifunctional enzyme (Figure 1.6).[37] Enzymes metabolizing Ψ are not found in humans and most other higher organisms suggesting that Ψ is not metabolized in them and excreted in urine.[38]

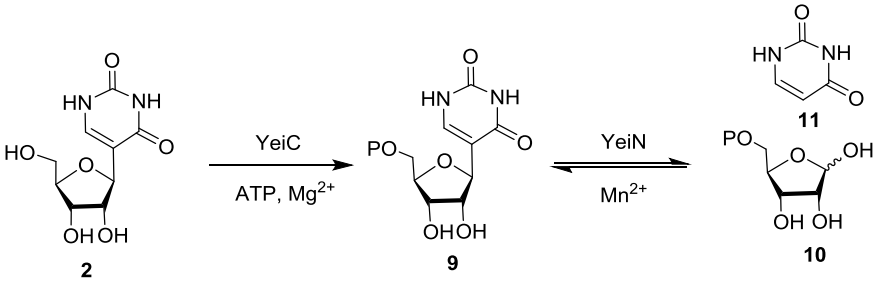



**Figure 1.5:** Catabolic pathway of pseudouridine

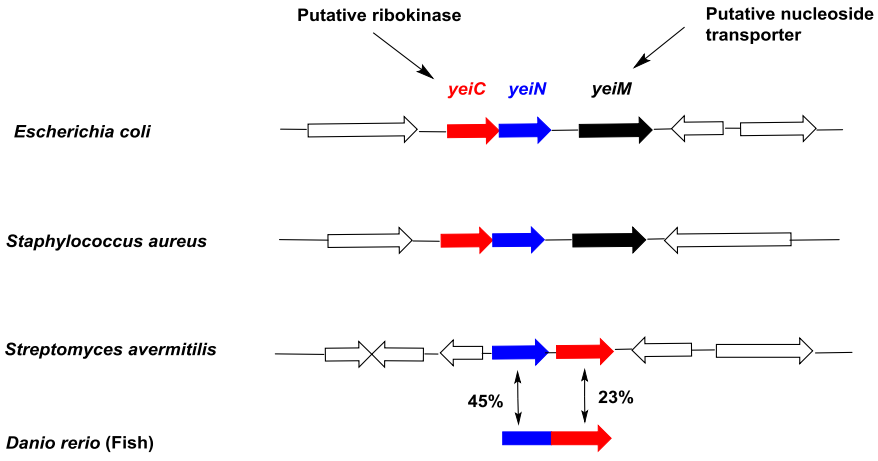



**Figure 1.6**: Organization of the genes involved in pseudouridine catabolism.

YeiN is the first molecularly identified enzyme that catalyzes a mechanistically intriguing hydrolytic cleavage of a glycosidic C-C bond.

**Glycosidase mechanisms**

While natural products containing O and N glycosidic bonds are widespread, natural products containing the C-glycosidic bond are much less prevalent.[39] The C-glycosidic bond is generally formed by an electrophilic aromatic substitution on an electron rich double bond. Enzymes that cleave C-N glycosidic bonds have been extensively studied. These include nucleoside and nucleotide hydrolases,[40-42] purine and pyrimidine nucleoside phosphorylases,[43] purine and pyrimidine phosphoribosyltransferases,[44, 45] and nucleoside deoxyribosyltransferases.[46] In addition to them, DNA repair enzymes such as endonuclease III also function as N- glycosylases to remove damaged nucleobases from DNA.[47] On the other hand, enzymes that cleave C-O glycosidic bonds are widespread in carbohydrate metabolism and have also been extensively studied mechanistically.[48, 49] Biochemical, biophysical and genetic studies suggest that these enzymes cleave C-N and C-O glycosidic bonds by a largely dissociative mechanism generating an oxocarbenium ion intermediate. General mechanisms for glycosidases cleaving C-O glycosidic bonds[49] are shown in Figure 1.7.

However, although the genes encoding several C-glycosynthases have been identified and overexpressed[50-57] and the structures of ligand free C-C bond forming glycosyltransferase, UrdGT2[58, 59] and ligand free C-glycosidase TM1464 from

*Thermotoga maritima* (an Indigoidine synthase, IndA or ΨMP glycosidase, YeiN like protein) have been determined[60], yet there is currently no example of a mechanistically well characterized C-glycosidase or glycosynthase.

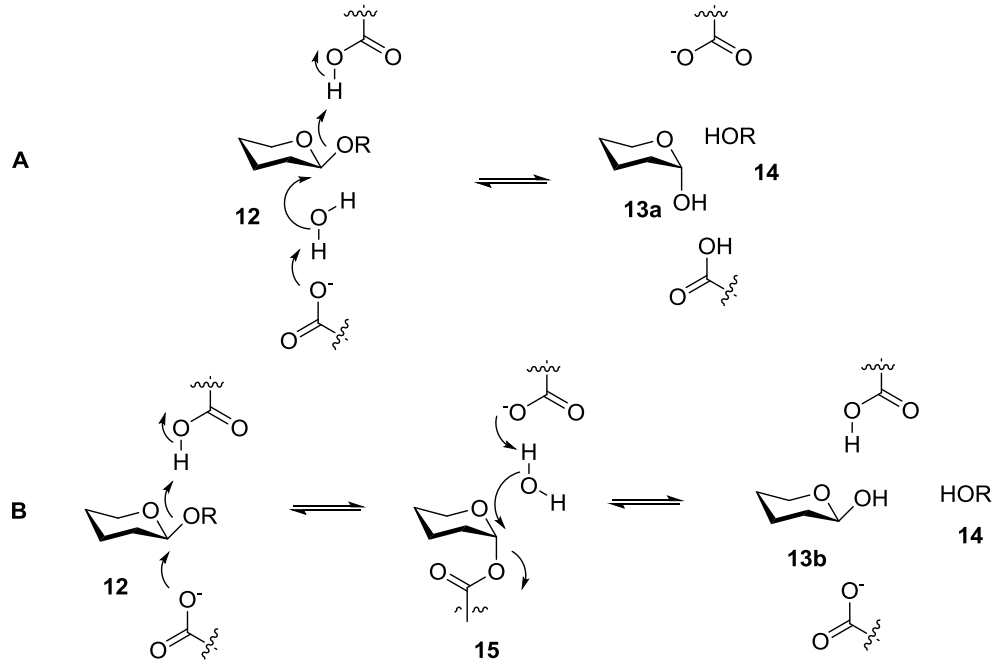



**Figure 1.7**: General mechanisms for inverting (A) and retaining (B) glycosidases

**Research opportunity**

From the above discussion, it is clear that, although there are reports of several naturally occurring C-glycosides, there is currently no example of a mechanistically well-characterized C-glycosidase. Recently identified ΨMP glycosidase (YeiN) represents one such case where biochemical and structural studies would lead to the understanding of the underlying mechanism of this kind of glycosidases. Moreover, it was observed that unlike the C-O and C-N bond cleaving glycosidases, YeiN shows metal ion cofactor

dependence. Hence, it will be interesting to know how different its mechanism would be from the well-established mechanisms of previously characterized glycosidases. The studies described herein are aimed at understanding the mechanism of the first ever identified C-C glycosidic bond cleaving enzyme (Figure 1.5).

## CHAPTER II

## PSEUDOURIDINE CATABOLISM: STUDIES ON THE PSEUDOURIDINE MONOPHOSPHATE (ΨMP) GLYCOSIDASE, YeiN [*#]

### Introduction

The pseudouridine (Ψ) catabolic pathway was discovered in *E. coli*.[37] In this pathway, pseudouridine kinase (YeiC) first phosphorylates the 5'-hydroxyl group of pseudouridine forming ΨMP. Then ΨMP glycosidase (YeiN) catalyzes the reversible conversion of ΨMP, **9** to ribose 5 phosphate, **10** and uracil, **11** (Figure 1.5). Based on the mechanism of Ψ synthase, we have a mechanistic proposal for YeiN (Figure 2.1).

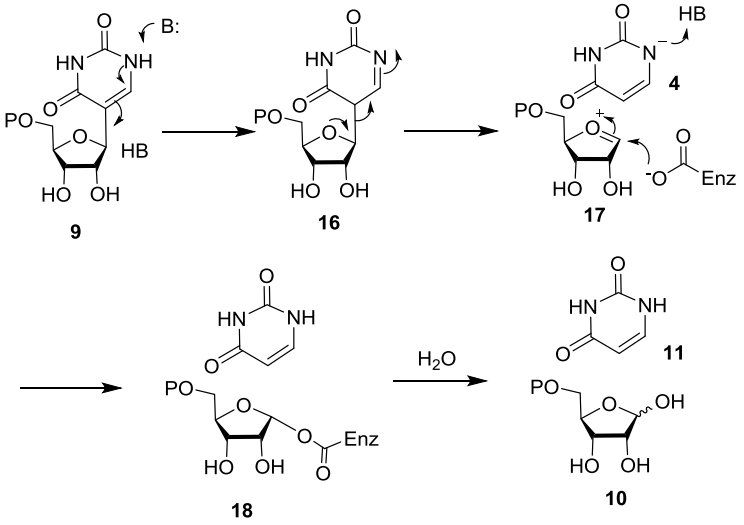



**Figure 2.1**: Mechanistic proposal for the ΨMP glycosidase, YeiN.

*Part of this chapter is reprinted with permission from "Pseudouridine Monophosphate Glycosidase: A New Glycosidase Mechanism" by Huang, S.; Mahanta, N.; Begley, T. P.; Ealick, S. E., 2012. *Biochemistry*, 51, 9245-9255, Copyright [2012] by American Chemical Society.

[#](Siyu Huang performed all the structural work shown here. Nilkamal Mahanta carried out all the biochemical studies. Cynthia Kinsland cloned YeiN and its mutants).

In this proposal, an acid-base residue assisted tautomerization would convert **9** to **16**, followed by the glycosidic bond cleavage in the latter, leading to the oxocarbenium ion intermediate **17** and uracil anion, **4**. Then **17** may be trapped by an enzyme active site residue (like Asp in Figure 1.3) forming a covalent enzyme-substrate adduct which may be further hydrolyzed to form ribose-5-phosphate, **10**. The uracil anion, after protonation would yield uracil, **11**. To investigate the mechanism of the glycosidase reaction and test the validity of our hypothesis, structural studies and biochemical experiments were carried out on YeiN. Interestingly, our results suggested a novel mechanism for ΨMP glycosidase, involving an covalent adduct between an active site lysine residue and C1' of the sugar moiety of the substrate during catalysis.[61]

**Results and discussion**

***HPLC analysis of the reaction mixture*:** Since, the YeiN catalyzed reaction is reversible; the reverse reaction was analyzed by high performance liquid chromatography in order to reconstitute the YeiN activity. Uracil (1mM) and ribose-5-phosphate, R5P (1mM) was incubated with YeiN (50 µM) in Tris-HCl buffer (pH 8.0) in presence of divalent metal ions ($Mn^{2+}$ and $Co^{2+}$, 1mM). Formation of a new compound at 2.1 min in the HPLC chromatogram was observed (Figure 2.2) and further confirmation by coupled YeiC and YeiN assay (Figures A1 and A2, Appendix A) and mass spectrometry and NMR spectroscopy (Figures A3 and A4, Appendix A) suggested that ΨMP is formed in the YeiN assay confirming the enzyme activity. This allowed us to carry on with further biochemical and structural studies on YeiN.

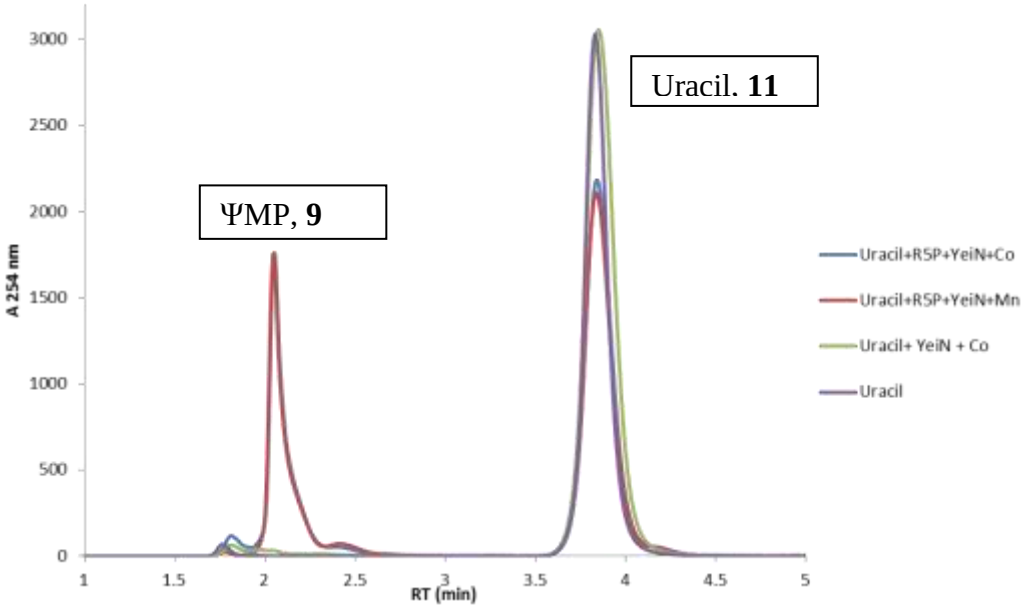



**Figure 2.2**: HPLC analysis of the YeiN reaction mixture.
[Blue trace: chromatogram of the full reaction mixture (YeiN+ R5P, **10** + Uracil, **11** + $Co^{2+}$) Red trace: chromatogram of the full reaction mixture (YeiN+ R5P, **10** + Uracil, **11** + $Mn^{2+}$). Both of them show the production of ΨMP, **9** at 2.1 min. Green trace: chromatogram of the control reaction mixture lacking substrate **10**. Purple trace: chromatogram of the authentic standard for uracil **11**.]

***Overall structure of the ΨMP glycosidase***: ΨMP glycosidase is a homotrimer[61] (Figure 2.3A) with one trimer per asymmetric unit in the crystal structure. Of the 312 possible residues, each protomer is complete except for the first two to five residues at the N-terminus and the C-terminal residue, which is missing in most of the protomers. ΨMP glycosidase shows a αβα fold. The central mixed β-sheet contains 11strands with 1↑4↑3↓2↓7↓8↓11↓9↓10↓5↑6↑ topology and is flanked by six α-helices on one side and seven α-helices on the other (Figure 2.3 B, C).[61] The trimer contains a buried surface area of 8650 Å2, accounting for 27.8% of the total surface area. The trimer interface is

formed primarily by helices α5 and α8 from one protomer and helices α9 and α10 from the adjacent protomer. The interface includes hydrogen bonds between Ile146 and Thr180* (an asterisk indicates residues from an adjacent protomer) and between Arg96 and Asn228*, and a salt bridge between Arg97 and Glu179*.

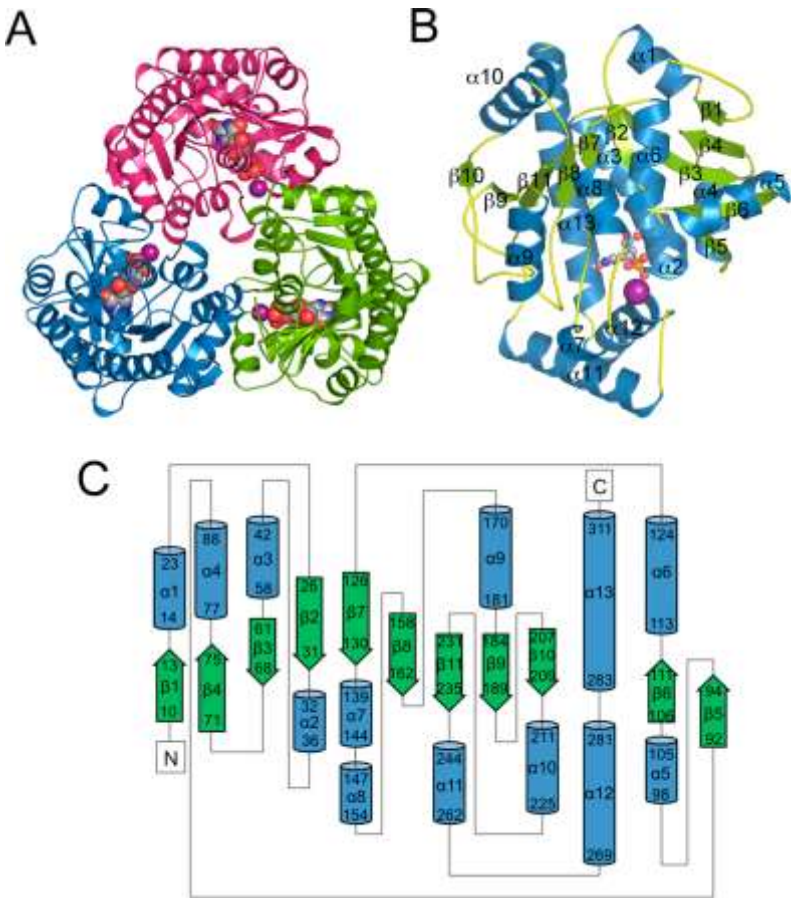



**Figure 2.3**: Structure of the ΨMP glycosidase.[61]
[(A) Ribbon diagram of the ΨMP glycosidase trimer color-coded by protomer and the active site is indicated by the space filling model. The Mn (II) ion is colored purple. (B) Ribbon diagram of a ΨMP glycosidase protomer labeled with color-coded α-helices (blue) and β-strands (green). (C) Topology diagram of ΨMP glycosidase. The green arrows represent β-strands and blue cylinders α-helices. The first residue number and the last residue number of each secondary structural element are indicated].

***Mn (II) binding site***: Mn (II) is required for ΨMP glycosidase activity[37] and for the formation of high-quality crystals. The Mn (II) coordination is octahedral with Mn (II)−oxygen distances ranging from 2.13 to 2.35 Å (Figure 2.4). Only one protein residue, Asp145, participates in coordination, and water molecules (W1−W5) occupy the five remaining positions. The coordinating oxygen atoms interact with protein atoms, substrate atoms (Figure 2.4), and other water molecules.W1 hydrogen bonds to Glu179* and a ΨMP phosphate oxygen atom.W2 hydrogen bonds to His137 and a ΨMP phosphate oxygen atom.W3 hydrogen bonds to Asp145 and the carbonyl oxygen atom of Gly266.W4 hydrogen bonds to Glu176* and Glu179*. W5 hydrogen bonds to a water molecule in the second sphere (W11) and a water molecule (W7) that bridges to the ΨMP phosphate. The water molecules directly involved in Mn (II) interactions are involved in additional hydrogen bonding interactions that extend into the remainder of the active site. In Figure 2.4, water molecules and protein atoms in the first and second coordination spheres are shown. The sulfate ion results from the crystallization conditions and occupies the phosphate binding site.

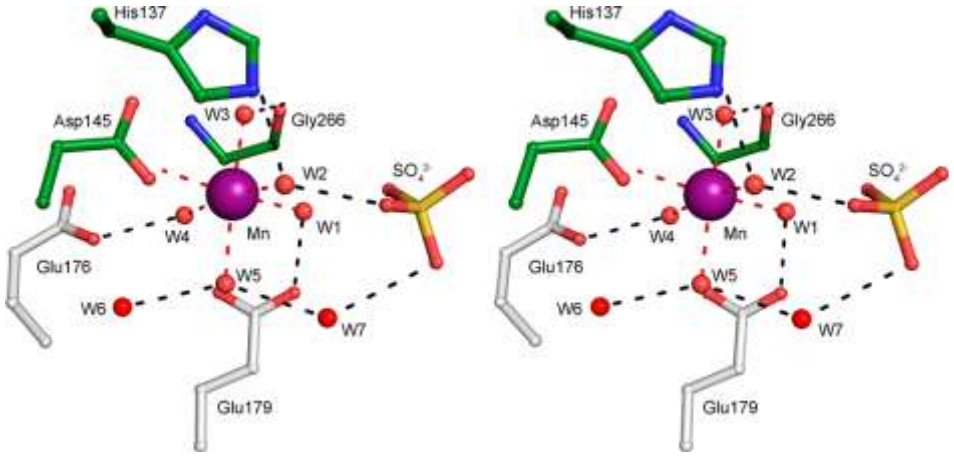



**Figure 2.4:** Stereoview of the Mn (II) binding site.[61]

***Structure of the ΨMP glycosidase−ring-opened ribose ΨMP adduct***: The structure of ΨMP glycosidase cocrystallized with R5P and uracil shows ring-opened ribose ΨMP covalently attached to Lys166.[61] Ring opening and closing are correlated with an approximately 90° rotation and a 2 Å shift of the uracil (Figure 2.5 A, B), thus occupying a binding site different from that of the K166A−ΨMP complex (Figure 2.7). In this binding site, O2 forms a hydrogen bond with Asn289. The ring-opened ribose retains hydrogen bonds between the 2′-hydroxyl group and Glu31 and between the 3′-hydroxyl group and Asp149; however, the hydrogen bond to O3′ requires a rotation of the Glu31 carboxylate about the Cγ−Cδ bond compared to the ΨMP complex (Figure 2.6). The phosphate group forms hydrogen bonds with Lys93, Ser147, W1, W2, and W7−W11. The water molecules near the uracil are not well defined, possibly because of the lower resolution (2.5 Å) of this structure.

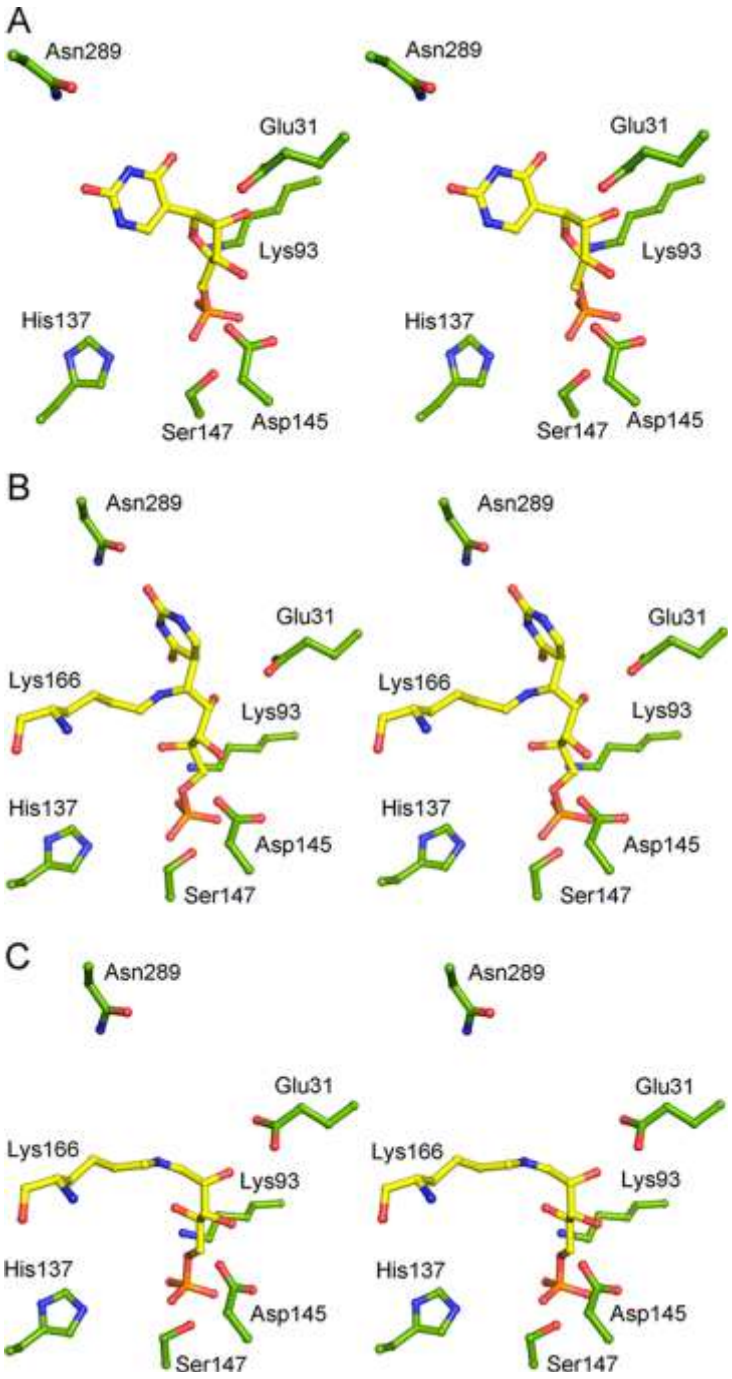



**Figure 2.5**: Comparison of the ΨMP glycosidase complexes.[61]
[(A) Stereoview of the K166A−ΨMP complex (B) Stereoview of the ΨMP glycosidase−ring-opened ribose ΨMP adduct (C) Stereoview of the ΨMP glycosidase−R5P adduct. Protein carbon atoms are colored green; ligand carbon atoms are colored yellow.]

***Structure of the ΨMP glycosidase−ribose-5-phosphate (R5P) adduct*:** Crystals of ΨMP treated with R5P show a structure in which the R5P is in the ring-opened form and is covalently attached to Lys166 via an imine.[61] The R5P superimposes closely with the ring-opened ribose in the ΨMP adduct and forms hydrogen bonds between the 2-hydroxyl group and Glu31 and between the 3-hydroxyl group and Asp149 (Figure 2.5 B, C). The phosphate binding site is essentially the same as for the previous complex. Water molecules in the uracil binding site are less ordered and more variable from protomer to protomer compared to those of the K166A−ΨMP complex (Figure 2.6).

***ΨMP glycosidase active site***: The active site, as defined by the structure of the K166A−ΨMP complex, is located in a cleft formed by helices α12 and α13 and the loops following strands β2, β7, and β8 (Figure 2.6 A, B).[61] The ΨMP ribose is in a C3′-endo conformation; the glycosidic torsion angle is anti, and the C4′−C5′ bond is in a gauche, trans conformation. Including the Mn (II) coordination sphere, the active site contains 19 well-ordered water molecules that are present in all three protomers. ΨMP is surrounded by 15 of the water molecules, and these mediate many of the active site contacts with the protein. One of the ΨMP phosphate oxygen atoms forms hydrogen bonds with W1 from the Mn (II) coordination sphere, Lys93, and water molecule W11 (Figure 2.6).

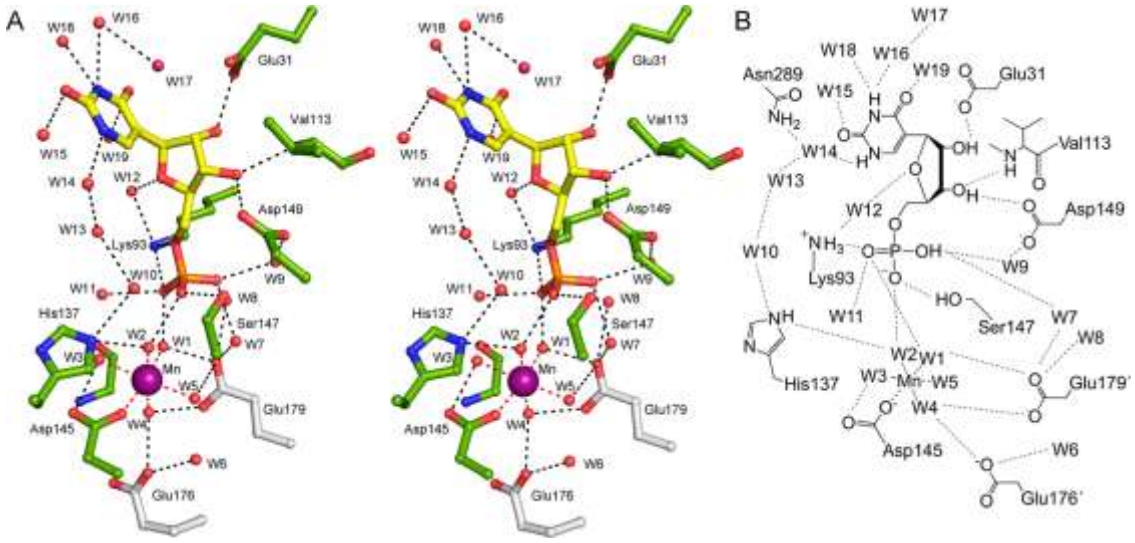



**Figure 2.6:** ΨMP glycosidase active site.[61]
[(A) Stereoview of the K166A−ΨMP active site: ΨMP was observed in the active site upon cocrystallization of K166A with either ΨMP or R5P and uracil. ΨMP and Mn (II) form extensive hydrogen bond connections with the active site residues and water molecules. Interactions are shown by a network of dashed lines (B) Schematic diagram of the K166A−ΨMP complex.]

A second phosphate oxygen atom forms hydrogen bonds with W2 from the Mn(II) coordination sphere, Ser147, and water molecule W10. The third phosphate oxygen atom forms hydrogen bonds with three water molecules (W7−W9), which interact through multiple contacts with the protein (Ser95, Thr112, Ala148 NH, Asp149, and Glu179*). The ribose O2′-hydroxyl group hydrogen bonds to Glu31; the O3′-hydroxyl group hydrogen bonds to Asp149 and the Val113 amide group, while O4′ hydrogen bonds with water molecule W12, which in turn hydrogen bonds to Lys93 and W19. The ΨMP uracil forms no direct hydrogen bonds with protein residues; however, N1, O2, N3, and O4 form hydrogen bonds with five water molecules. One water molecule (W14)

hydrogen bonds to N1 and W15 and forms a bridge to the ΨMP phosphate through two additional water molecules (W10 and W13).The second water molecule (W16) hydrogen bonds to O2, the side chain of Asn289, W17, and W18, which is positioned over the center of the uracil ring and hydrogen bonds to the amide nitrogen and carbonyl oxygen atoms of Gly132. The third water molecule (W15) also hydrogen bonds to O2, the amide nitrogen atom of Ala166, and W14. The fourth water molecule (W18) hydrogen bonds to N3, the carbonyl oxygen atom of Gly38, and the side chain of Asn289. The fifth water molecule (W19) hydrogen bonds to O4, the carbonyl oxygen atom of His37, and an additional water molecule (W12), which hydrogen bonds to O4′ of ΨMP and Lys93.[61]

***Detection of a Lys166- R5P covalent adduct by mass spectrometry***: To confirm the findings from the structural study, biochemical experiments were carried out on YeiN. The protein-substrate covalent adduct formation is investigated by intact protein mass spectrometry. The mass of wild-type ΨMP glycosidase was determined to be 35453.4 Da by ICR-MS analysis (Figure 2.8A).[61]The mass of the reduced R5P complex was 35668.4 Da, corresponding to a mass increase of 215 Da (Figure 2.8A). This is consistent with a doubly protonated reduced R5P-derived imine of mass 214 Da due to the adduct formation with lysine 166 (Figure 2.7). Adduct formation was not observed when only uracil was incubated with ΨMP glycosidase (Figure 2.8A).

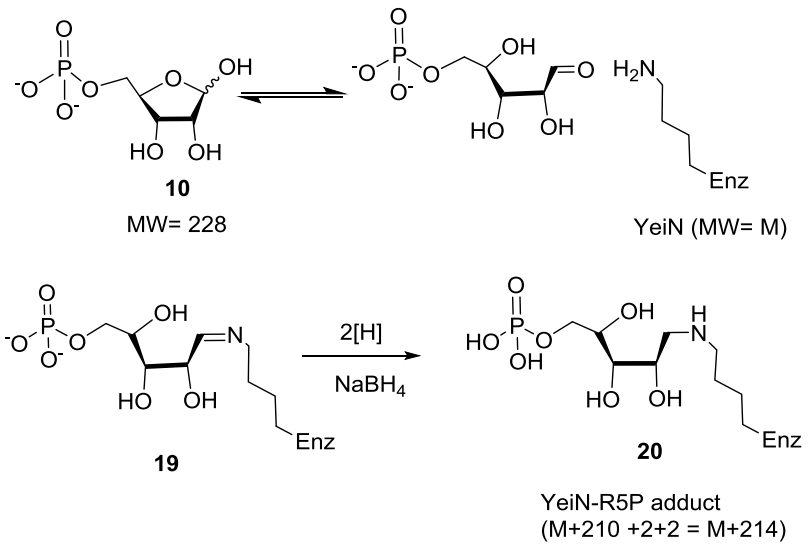



**Figure 2.7**: Predicted mass of the YeiN-R5P adduct

The site of imine formation was identified by trypsin digestion of the reduced ΨMP glycosidase R5P adduct (Figure 2.8B−E).[61]Trypsin is a protease that cleaves after lysine and arginine residues, and lysine modification blocks the cleavage reaction as trypsin cannot recognize the modified site. Trypsin digestion of the unmodified enzyme yielded the GAEHTFDISADLQELANTNVTVVCAGAK peptide (mass of 2930.4079 Da, amino acids 139−166). For ΨMP glycosidase incubated with R5P, trypsin digestion didn't cut at the usual K166 residue but rather cut at the next lysine, K191 yielding the GAEHTFDISADLQELANTNVTVVCAGAKSILDLGLTTEYLETFGVPLIGYQTK peptide (mass of 5896.8759 Da, amino acids 139−191). This demonstrates that Lys166 is the R5P-modified residue. This was confirmed by similar ICR-MS analysis of the K166A mutant for which no mass increase was observed for the R5P-treated sample (Figure 2.8A).

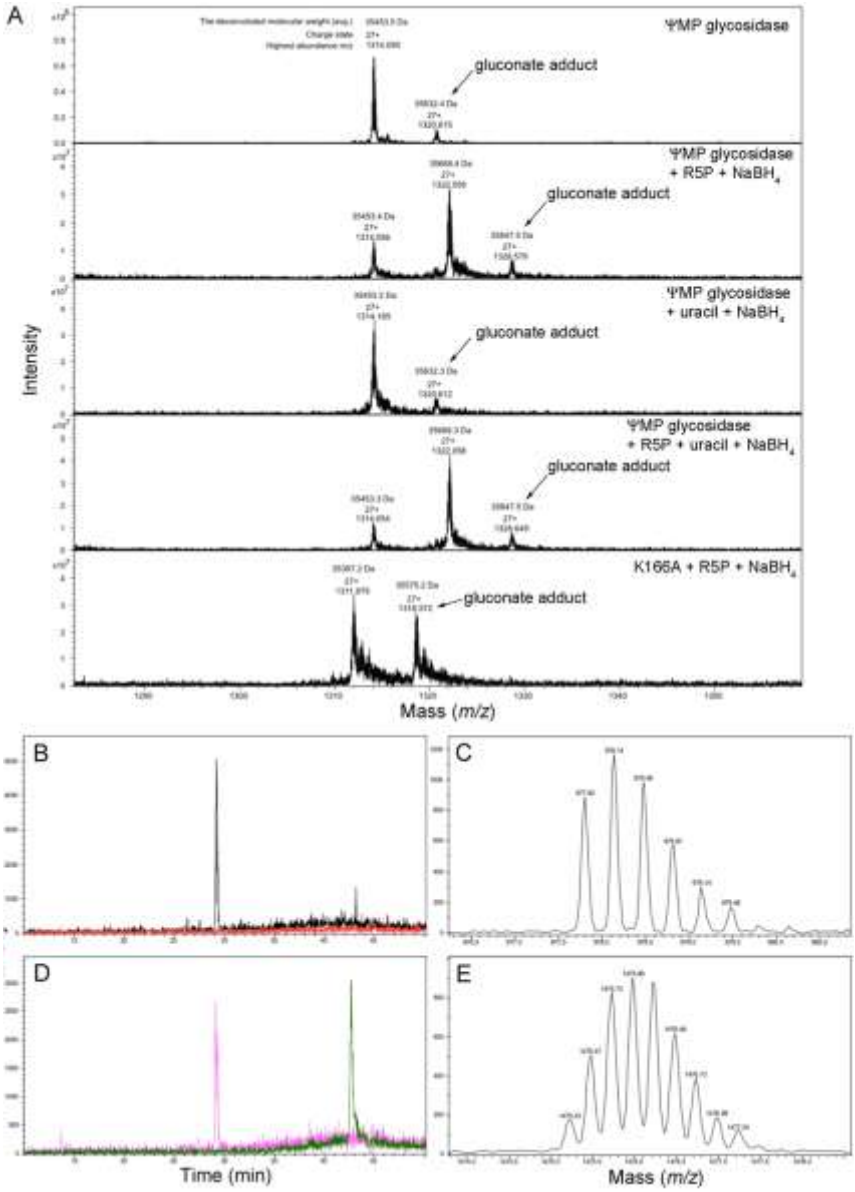



**Figure 2.8**: Analysis of the YeiN-R5P adduct by mass spectrometry.[61]
[(A) ICR-MS data for the identification of the borohydride-reduced ΨMP glycosidase−R5P adduct: The observed 179 Da adduct corresponds to enzyme gluconoylation. (B) MS chromatograms of the control sample. The black trace corresponds to data for the unmodified peptide, while the red trace corresponds to data for the R5P-bound peptide, which is not seen in the ΨMP glycosidase control. (C) Triply charged state of the unmodified peptide in the control ΨMP glycosidase sample. (D) MS chromatograms of the reduced ΨMP glycosidase−/R5P adduct. The pink trace corresponds to data for the unmodified peptide, while the green trace corresponds to data for the R5P-bound modified peptide. (E) Quadruply charged state of the R5P-bound modified peptide in the reduced ΨMP glycosidase−R5P adduct.]

***Steady state kinetics parameters of ΨMP glycosidase and its mutants*:** Steady state kinetic parameters were determined by measuring the production of ΨMP during the reverse reaction while varying the uracil concentration at a saturating R5P concentration (Figure 1.5). Michaelis−Menten analysis (Figure 2.9) of the native enzyme indicated a $K_m$ value of 170 μM for uracil and a $k_{cat}$ of 3.74 $s^{-1}$.[61]This corresponds to a catalytic efficiency, $k_{cat}/K_m$, of 22 × 103 $M^{-1}$ $s^{-1}$. Steady state kinetic parameters were also determined for the active site mutants E31A, K93A, K166A, and N289A (Table 1 and Figure A5 of Appendix A). The activity of D149A was below the detection limit even at an enzyme concentration as high as 100 μM and no activity was detected after several attempts (written as ND in Table 2.1). K93A has a higher *K*m value than the wild-type enzyme. E31A, K93A, K166A, and N289A have significantly lower *k*cat values, with E31A and K166A showing the largest effects.

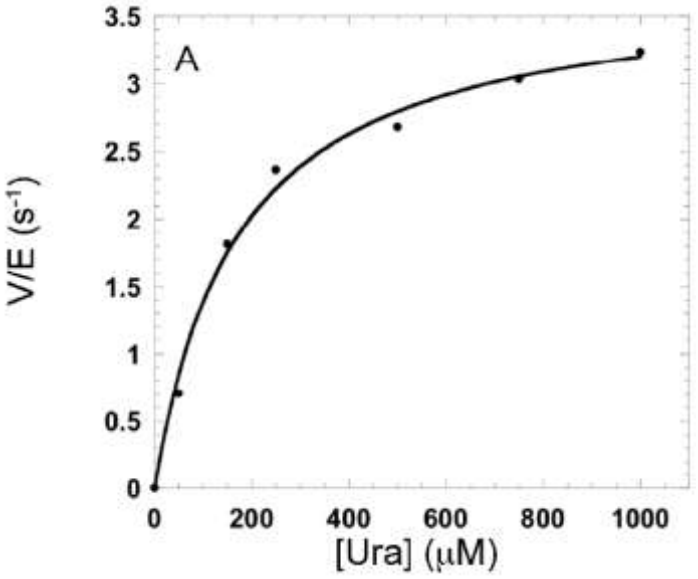



**Figure 2.9:** Michaelis-Menten plot for the ΨMP glycosidase.[61]

**Table 2.1**: Steady state kinetic parameters for the ΨMP glycosidase and mutants.[61]

| Enzyme | $k_{cat}$ ($s^{-1}$) | $K_m$ (μM) | $k_{cat}/K_m$ ($M^{-1}$ $s^{-1}$) |
|---|---|---|---|
| YeiN | 3.74 ± 0.14 | 170 ± 22 | (22 ± 4) x $10^3$ |
| K166A | 0.0013 ± 0.0001 | 163 ± 23 | 8.0 ± 1.5 |
| E31A | 0.0005 ± 0.0001 | 191 ± 39 | 2.6 ± 0.8 |
| K93A | 0.23 ± 0.01 | 214 ± 18 | (1.1 ± 0.1) x $10^3$ |
| N289A | 0.22 ± 0.01 | 300 ± 22 | (7.3 ± 0.8) x $10^2$ |
| D149A | ND | ND | ND |

***Pre-steady state kinetics analysis of ΨMP glycosidase reaction*:** In a rapid quench experiment, burst phase was observed for the ΨMP glycosidase catalyzed reaction indicating that the formation of product is much faster than product release step. The plot of turn over number vs. time(s) is shown in Figure 2.10. The data was fitted into the equation shown in the graph and from this equation; the burst amplitude ($m_2$), burst rate ($m_3$) and linear rate ($m_1$) were determined to be 0.7±0.03, 124±14 $s^{-1}$ and 1.01±0.11 $s^{-1}$ respectively (Table 2.2).

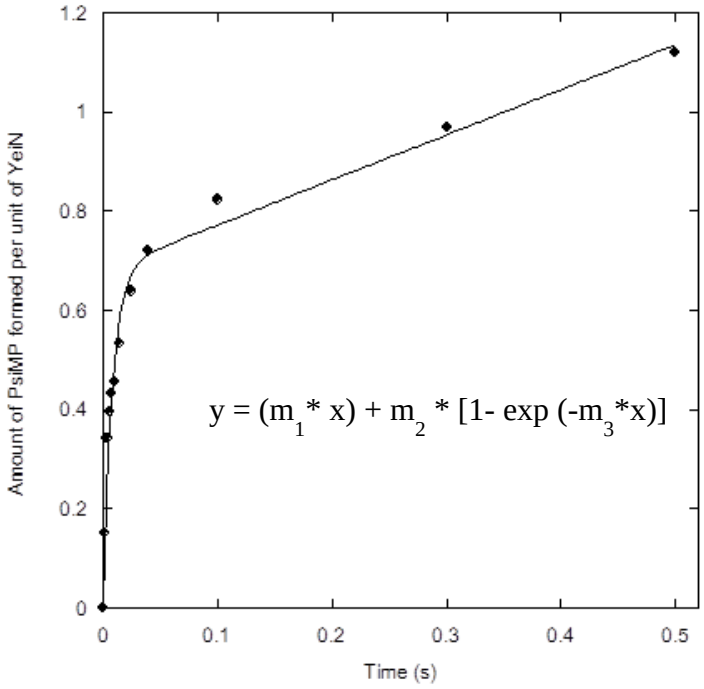



**Figure 2.10:** Pre-steady state kinetics for the ΨMP glycosidase.

**Table 2.2**: Pre steady state kinetics parameters for the ΨMP glycosidase

| | |
|---|---|
| $m_1$ | 1.01 ± 0.11 |
| $m_2$ | 0.70 ± 0.03 |
| $m_3$ | 124 ± 14 |

***Mechanistic implications*:** The mechanism of pseudouridine synthase has been extensively studied and is outlined in Figure 1.3.[30]The microscopic reverse of this reaction would seem to be a reasonable starting hypothesis for the reaction catalyzed by

ΨMP glycosidase (Figure 2.1). Our structural studies identified four snapshots of the *E. coli* ΨMP glycosidase reaction coordinate.[61] Wild-type *E. coli* ΨMP glycosidase showed the unliganded state, although a sulfate ion from the crystallization solutions occupied the phosphate binding site. Incubation of the K166A mutant with R5P and uracil resulted in a structure with the substrate ΨMP bound in the active site. Native ΨMP glycosidase incubated with R5P and uracil showed an intermediate in which ring-opened ribose ΨMP is covalently attached to Lys166. Finally, ΨMP glycosidase incubated with R5P alone resulted in a structure with ring opened R5P attached to Lys166. The structural studies suggested that the reaction involves an intermediate imine with Lys166. This was confirmed by mass spectrometry of the wild type enzyme and trypsin digested fragments. Kinetic studies of active site mutants suggested roles for individual amino acid active site residues. Comparison of the ΨMP glycosidase structures [60, 61] suggested a role for conformational changes in ΨMP cleavage and product release. This series of experiments demonstrated that the ΨMP glycosidase-catalyzed reaction occurs by a mechanism that is significantly different from the microscopic reverse of the pseudouridine synthase mechanism.[30]

An alternative mechanistic proposal is outlined in Figure 2.11. Ring opening of the ΨMP ribose gives **21**. This reaction requires acid−base catalysis. As no amino acids are apparent in the structure for this role (Figure 2.1), the ring-opening reaction probably occurs by a water-mediated protonation−deprotonation reaction. The conformation of the C glycosidic bond is reasonable for a concerted process (the C6−C5−C1−O4a

dihedral angle is 50°, while the optimal angle would be 90°). Related ribose ring-opening reactions are catalyzed by GTP cyclohydrolase.[62] Following ribose ring opening, Lys166 undergoes a conjugate addition to the C−C double bond of **21** to give **22** with the proton coming from Glu31. Subsequent cleavage of the C−C glycosidic bond by a retroaldol-type reaction releases the uracil anion, which is stabilized by hydrogen bonding of N1 to the hydroxyl of Thr130 and the amide NH of Gly131 and Gly132. Surprisingly, there are no stabilizing hydrogen bonds to the C2 or C4 carbonyl oxygen atoms of uracil. Hydrolysis of imine **19** gives **23**, which will then cyclize to **10**. In addition to the structures, the covalent linkage between Lys166 and the substrate is supported by the characterization of the R5P Lys166 imine by mass spectrometry.

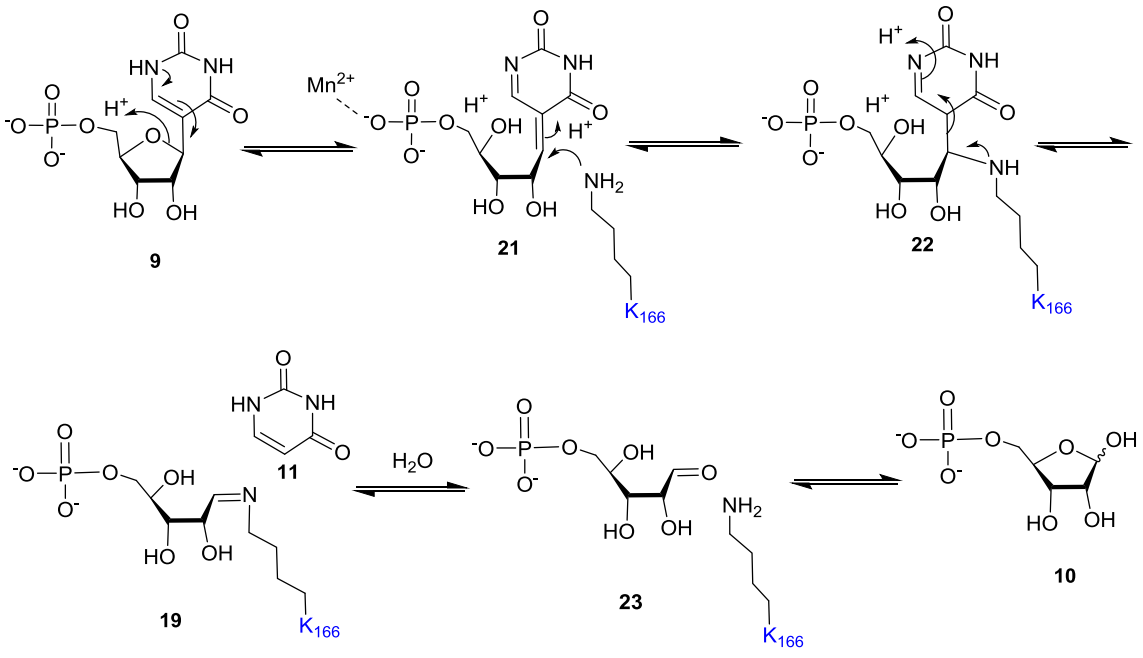



**Figure 2.11**: Mechanistic proposal for the ΨMP glycosidase.[61]

An alternative mechanism in which Lys166 displaces the pyrimidine from the ribose, as found, for example, in 8-oxoguanine DNA glycosylase[63] is unlikely. Lysine 166, modeled into the structure of the enzyme ΨMP complex, is not suitably positioned for direct base displacement, and the structure of the ring-opened ribose ΨMP complex is not consistent with this mechanism. The structure of UrdGT2, with the substrates modeled into the active site, suggests that this C−C glycosyl transferase does not use an imine intermediate.[58] In this regard; our analysis suggests that ΨMP glycosidase may use a new glycosidase mechanism.[61]

***Implications from the steady state kinetics parameters*:** This mechanistic proposal allows us to suggest testable functions for the active site residues.[61] Lys166 plays a key role in facilitating the C glycosyl bond cleavage reaction, and as expected, the $k_{cat}$ value of the K166A mutant is 2900-fold lower than that of the wild type, with almost no change in $K_m$. It is interesting that low levels of catalytic activity are observed for this mutant, suggesting that water can replace lysine in the mutant. Glu31 is likely to be the proton source for the lysine conjugate addition reaction (**9** to **21** in Figure 19). Consistent with this role, the $k_{cat}$ value of the E31A mutant is 7500-fold lower than that of the wild type, with little change in $K_m$. Lys93 contributes one of several interactions involved in phosphate binding. The $k_{cat}$ value of the K93A mutant is 17-fold lower than that of the wild type, with only a modest increase in $K_m$. Asn289 hydrogen bonds to two water molecules that are in turn hydrogen bonded to the C2 carbonyl and to N3. These interactions stabilize the negative charge on the uracil in the conversion of **22** to **19** and

**11**. Consequently, the $k_{cat}$ value of the N289A mutant is 17-fold lower than that of the wild type, with a small increase in $K_m$. Asp149 forms a hydrogen bond to the substrate C3′ alcohol. We would therefore expect that the activity of the D149A mutant would be comparable to that of the wild type, which is inconsistent with our finding that it shows no detectable activity. One possibility is that the D149A mutant has misfolded; however, CD spectra of the wild type and D149A are nearly identical (unpublished data). Another possibility is that Asp149 serves indirectly as an acid and/or base through a charged network. We could not get any activity for this mutant under standard assay conditions.[61]

***NMR analysis of the ΨMP glycosidase reaction mixture***: Besides these biochemical studies, attempts have also been made to trap and identify the proposed intermediates using NMR spectroscopy (data not shown). Reactions were performed using both wild type ΨMP glycosidase (YeiN) and K166A mutant and $^{13}C$ labeled substrates ($^{13}C_1$ ribose-5 phosphate and $^{13}C_{4,\ 5}$ -uracil) and $Mg^{2+}$ (as the reaction is slower than with $Mn^{2+}$) and were quenched at various time points by heating at 100°C for few minutes and analyzed by NMR spectroscopy. Reactions were also monitored in real time scale by using NMR (added enzyme in the NMR tube carrying all other components and monitored after every 5 min). However, all our attempts to trap intermediates biochemically (compounds **21** and **22** as proposed in the mechanism, Figure 2.11) were unsuccessful as reaction always resulted in the formation of the final product pseudouridine monophosphate, **9.** This suggested that probably these intermediates are

not released from the enzyme active site during the catalytic cycle or may be present in very low concentration in the solution to be detected by NMR spectroscopy.

***Role of the hydrated metal binding site***: From the structural studies[61], it was clear that ΨMP glycosidase contain heavily hydrated, octahedrally coordinated Mn (II) ions in which the ligands are an aspartate side chain and five water molecules. In addition, the water molecules in the second sphere (W6−W9) are also present in both structures, and all side chains in the second sphere are conserved, including Glu176 and Glu179 from the neighboring protomer. This heavily hydrated metal binding site is unusual and likely plays a role in anchoring the ΨMP phosphate. Two phosphate oxygen atoms form hydrogen bonds to W1 and W2, which are in the Mn (II) coordination sphere. The phosphate group also hydrogen bonds to Ser147 and Lys93, which are conserved in both structures. While the ribose and uracil both undergo movement during catalysis, the phosphate remains in the same position, with no hydrogen bonding changes. ΨMP glycosidase shows the highest activities with Fe(II), Co(II), and Mn(II), while Zn(II) and, to a lesser extent, Ni(II) were inhibitory.[37] These observations suggest that the ionic radius of the metal ion may play a role in properly positioning the substrate.

***Role of conformational changes in catalysis***: Both the protein and substrate undergo conformational changes during catalysis. The uracil undergoes a 90° rotation and a 2 Å shift after ring opening and Lys166 adduct formation.[61] In the glycosidase ΨMP complex, uracil forms no hydrogen bonds with protein atoms; however, after shifting

position, uracil hydrogen bonds through O2 to Asn289, which is located at the opposite end of the active site relative to the phosphate. Repositioning of the uracil also places N1 near the amide nitrogen atoms of Gly131 and Gly132, which are absolutely conserved. Heavy reliance in the binding site on water mediated interactions is consistent with the need for major repositioning of an intermediate during catalysis.

**Conclusion**

The ΨMP glycosidase complexes described here give four snapshots of the reaction coordinate and provide key structural insights into the enzymatic hydrolysis of C-glycosides. The mass spectrometric studies and the steady state kinetics analysis of the ΨMP glycosidase and the active site mutants also provide substantial evidence in support of the structural data. Most of the characterized glycosidases utilize a dissociative mechanism to reversibly cleave C−N and C−O glycosidic bonds. During such a mechanism, the opening of the ribose ring was not observed and there is no requirement for metal ion cofactors. In contrast, the structural and biochemical studies reported here demonstrate that ΨMP glycosidase requires divalent metal ions (such as $Mn^{2+}$ or $Co^{2+}$) and utilizes a new mechanism involving ribose ring opening and subsequent covalent linkage between C1′ and an active site lysine, thus setting the stage for a facile C glycosyl bond fragmentation by a novel retro aldol type reaction mechanism. To the best of our knowledge, this is the first mechanistically characterized example of a C- glycosidase.

**Experimental procedures**

***Chemical reagents***: R5P and uracil were purchased from Sigma-Aldrich. Tris(2-carboxyethyl)phosphine (TCEP) was purchased from Hampton Research. Kanamycin and IPTG were purchased from Fischer Scintific. Ni-NTA resin was obtained from Qiagen (Valencia, CA).

***Cloning, overexpression and purification of ΨMP glycosidase:*** Standard methods were used for DNA restriction endonuclease digestion, ligation, and transformation of DNA.[62-64]. Plasmid DNA was purified with a GeneJet miniprep kit (Fermentas, Glen Burnei, MD). DNA fragments were separated by agarose gel electrophoresis, excised, and purified with the Zymoclean gel DNA recovery kit (Zymo Research, Orange, CA). *E. coli* strain MachI (Invitrogen, Madison, WI) was used as the recipient for transformations during plasmid construction and for plasmid propagation and storage. An Eppendorf Mastercycler and Phusion DNA polymerase (New England Biolabs, Ipswich, MA) were used for polymerase chain reaction (PCR). All restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs. *E. coli* strain BL21 (DE3) and the pET overexpression systems were purchased from Novagen (Madison, WI). The *yeiN* gene was amplified via PCR from *E. coli* K12 genomic DNA using the following primers: upstream primer 5′-GGG TAG CAT ATG TCT GAA TTA AAA ATT TCC CCT G-3′ (inserts an *Nde*I site at the start codon of the yeiN open reading frame) and downstream primer 5′-CCC TAC TCG AGT TAA CCC GCG AGA CGC TGA TAT TC-3′ (inserts an *Xho*I site after the end of the *yeiN* open reading

frame). The purified PCR product was digested with *Nde*I and *Xho*I, purified, and ligated into similarly digested pTHT, a pET-28-derived vector, which allows attachment of a modified six histidine tag followed by a tobacco etch virus protease (TEV-protease) cleavage site onto the N-terminus of the expressed protein. Colonies were screened for the presence of the insert, and a representative plasmid was designated pEcYeiN.THT. The PCR-derived DNA was sequenced and shown to contain no errors.

The yeiN gene was further transformed into *E. coli* strain BL21 (DE3) (Novagen). The cells were grown overnight in a 10 mL starter culture in Luria-Bertani (LB) medium[65] containing 30 μg/mL kanamycin, then transferred to cultures containing 1.5 L of LB medium, and incubated at 37 °C while being shaken until an OD600 of 0.8 was achieved. The culture was then induced with 1 mM isopropyl 1-β-D-galactopyranoside (IPTG) and incubated overnight at 15 °C. Cells were harvested by centrifugation at 10000g for 30 min at 4 °C, resuspended in lysis buffer containing 50 mM tris(hydroxymethyl)-aminomethane (Tris) (pH 8.0) 300 mM NaCl, and 10 mM imidazole, and lysed by sonication. The lysate was centrifuged at 40000g for 30 min, and the supernatant was loaded onto a column containing 5 mL of Ni-NTA resin (Qiagen) preequilibrated with lysis buffer. The column was then washed with 3 times the volume of wash buffer containing 50 mM Tris (pH 8.0), 300 mM NaCl, and 30 mM imidazole for 1.5 h. The protein was eluted with 50 mM Tris (pH 8.0), 300 mM NaCl, and 250 mM imidazole. The eluted protein was then subjected to size exclusion chromatography using an ACTA Explorer fast performance liquid chromatography

system with a HiLoad 26/60 Superdex 200 prep grade column (GE Healthcare). The resulting protein was more than 95% pure as judged by sodium dodecyl sulfate−polyacrylamide gel electrophoresis analysis (unpublished experiments). The protein was then concentrated to 15 mg/mL using an Amicon concentrator (30 kDa molecular mass cutoff filter, Millipore), flash-frozen, and stored at −80 °C.

***ΨMP glycosidase mutagenesis*:** Site-directed mutagenesis was performed by a standard PCR protocol using PfuUltraII DNA polymerase per the manufacturer's instructions (Agilent) and *Dpn*I to digest the methylated parental DNA prior to transformation. In addition to the forward and reverse primers required to introduce the mutation, a third primer was designed to screen for the presence of the mutation by colony PCR (Table 1 Appendix A). For screening, the primers designated "sF" were paired with the T7T primer (5′-GCTAGTTATTGCTCAGCGG-3′) and the primers designated "sR" were screened with the T7Plac primer (5′-TATAGGGGAATTGTGAGCGG-3′). All the mutants were verified by sequencing.

***Enzymatic synthesis of ΨMP*:** ΨMP glycosidase (2.5 mg), uracil (100 mM), and R5P were incubated in a buffer containing 0.5 mM MnCl2 and 25 mM HEPES (pH 7.1) at 25°C for 1 h. The enzyme was removed using an Amicon 30 kDa molecular mass cutoff (Millipore) filter. The reaction product was identified as ΨMP by co-migration with an authentic sample of ΨMP during high-performance liquid chromatography (HPLC) analysis as well as MS and NMR analysis.

***Crystallization of the ΨMP glycosidase*[61]**: ΨMP glycosidase was crystallized using the hanging drop vapor diffusion method at 22 °C. The initial crystallization condition was determined using sparse matrix screens Crystal Screen 1 and 2 (Hampton Research). The optimized reservoir conditions included 20% polyethylene glycol 4000, 0.2 M sodium acetate, and 0.1 M Tris (pH 7.0). ΨMP glycosidase was incubated with 4 mM $MnSO_4$ before crystallization. The drops contained 1.5 μL of a protein solution and 1.5 μL of a reservoir solution. Prismatic crystals grew within 2 days to a size of 200 μm × 50 μm × 50 μm. The ΨMP glycosidase−R5P complex was cocrystallized under similar conditions except that the protein was incubated with 4 mM $MnCl_2$ and 2 mM R5P for 30 min prior to crystallization. The ΨMP glycosidase−ring-opened ribose ΨMP complex was obtained by cocrystallizing ΨMP glycosidase with 4 mM $MnCl_2$, 2 mM R5P, and saturated uracil. The K166A−ΨMP complex was prepared by cocrystallizing the mutant enzyme with R5P and uracil (by Siyu Huang, Ealick lab, Cornell University).

***Data collection and processing***: Prior to data collection, crystals were soaked in a cryoprotectant solution of 5% glycerol in mother liquor to avoid damage during vitrification. Data sets for ΨMP glycosidase and the complexes were collected at Northeast Collaborative Access Team (NE-CAT) beamline 24- ID-C at the Advanced Photon Source (APS) using an ADSC Quantum 315 detector (Area Detector Systems Corp.) at a wavelength of 0.9795 Å with 1 s exposure times and 1° oscillation angles. The data set for the K166A−ΨMP complex was collected at Cornell High Energy Synchrotron Source (CHESS) using an ADSC Quantum 270 detector at a wavelength of

0.9180 Å with 1 s exposure times and 1° oscillation angles. The detector distances were 275, 350, 325, and 200 cm for the ΨMP glycosidase, the R5P complex, the ring-opened ribose ΨMP complex, and the K166A−ΨMP complex, respectively. Data were indexed, integrated, and scaled using the HKL2000 program suite.[66] Data collection statistics are reported[61] (by Siyu Huang, Ealick lab, Cornell University).

***Structure determination and refinement***: The structure of *E.coli* ΨMP glycosidase was determined by molecular replacement using MOLREP.[67] The search model was a monomer of Protein Data Bank (PDB) entry 1VKM, which has a sequence 39% identical to that of ΨMP glycosidase, after modification by CHAINSAW.[68] 1VKM was originally reported to be an indigoidine (IndA)-like protein from *Thermotoga maritima*[60] but later was shown to be a ΨMP glycosidase.[37] The model was refined through successive rounds of manual model building using COOT[69] and restrained refinement with REFMAC5.[70] Water molecules were then included after the model converged, followed by two additional rounds of refinement. The structure of the ΨMP glycosidase−ringopened R5P adduct was determined by using the structure of unliganded ΨMP glycosidase as the model and refining with PHENIX.[71] The structure of the ΨMP glycosidase−ringopened ribose ΨMP adduct, which crystallized in a different space group, was determined by molecular replacement using MOLREP[67] using a monomer from the refined ΨMP glycosidase as the search model. The structure of the K166A−ΨMP complex was determined by using the structure of the ΨMP glycosidase−ring-opened ribose ΨMP adduct as the model and refining with PHENIX.[71]

The ligand contents and alternative side chain conformations were determined by Fobs − Fcal maps and composite omit maps from PHENIX.[71] All ligands were clearly observed in the composite omit map but were not placed until the last stage of the refinement. The final refinement statistics are reported[61] (by Siyu Huang, Cornell University).

***Comparison with other protein structures***: A structural similarity search for ΨMP glycosidase was performed using DALI.[72] Not surprisingly, PDB entry 1VKM, which was used as the search model during molecular replacement, showed the highest degree of similarity with a Z score of 42.2 (39% identical sequence)[61] (by Siyu Huang, Ealick lab, Cornell University).

***HLPC analysis of the ΨMP glycosidase reaction mixture*:** HPLC analysis following a linear gradient, at a flow rate of 1 mL/min, was used with absorbance detection at 254 nm. Solvent A was water; solvent B was 100 mM $K_2HPO_4$ (pH 6.6), and solvent C was methanol: 100% B at 0 min, 10% A and 90% B at 5 min, 25% A, 60% B, and 15% C at 7 min, 25% A, 60% B, and 15% C at 17min, 30% A, 40% B, and 30% C at 19 min, 100% B at 21 min, and 100% B at 30 min. The column used was a Supelcosil LC-18-T HPLC column (15 cm × 4.6 mm, 3 μm particle size).

***Determination of kinetics parameters for the ΨMP glycosidase reaction:*** Uracil concentrations were varied while the R5P concentration was kept at saturation (1 mM) and the Mn (II) concentration at 250 μM. ΨMP formation was monitored by HPLC

analysis. Each reaction mixture was divided into 90 μL aliquots, and 10 μL of 10 μM ΨMP glycosidase was added to achieve the final concentration of 1μM. For each set of reactions, the reaction was quenched at 20s, 40 s, 1 min, 2 min, 3 min, 5 min, and 10 min time points when the mixture was heated at 100 °C for 2 min. In the case of K166A, the uracil concentrations were varied while the R5P concentration was kept at 1 mM and the Mn(II) concentration at 250 μM for five different sets of reactions. Each reaction mixture was divided into 85 μL aliquots, and 15 μL of 250 μM K166A was added to start the reaction (final concentration of 37.5 μM). For each set of reactions, the reaction was quenched at 20 s, 40 s, 1 min, 2 min, 3 min, 5 min, and 10 min time points when the mixture was heated at 100 °C for 2 min. All samples were analyzed by HPLC, and the amount of ΨMP formed was plotted against time for each set of substrate concentrations. The initial slopes from each set of reactions were then plotted against the substrate concentration and fit to the Michaelis−Menten equation using KaleidaGraph (Synergy Software). The mutant E31A, K93A, H137A, D149A, and N289A activities were measured using a similar procedure (Figure A5, Appendix A).

***Biochemical characterization of the ΨMP glycosidase-R5P adduct*:** A reaction mixture containing 2 mM R5P and 1 mM Mn(II) in 50 mM Tris-HCl, 100 mM NaCl, and 2 mM TCEP (pH 8.0) was incubated at room temperature for 1 h with ΨMP glycosidase (50 μM), followed by addition of sodium borohydride (2 mg) and further incubation at room temperature for 45 min. The sample was then buffer-exchanged into 10 mM ammonium acetate (pH 8.0) using a Bio-Rad desalting column and analyzed by ICR-MS. Control

samples, one containing 2 mM uracil and the other lacking R5P and Mn(II), were similarly prepared and analyzed. A similar experiment was also performed on the K166A mutant of the enzyme. To locate the site of adduct formation; trypsin digestion analysis was performed on each of these samples as follows. Ten microliters of guanidine hydrochloride (6 M) was added to each sample (50 μL containing 100 μg of ΨMP glycosidase), followed by 1 μL of DTT (200 mM). After 1 h incubation at room temperature, 10 μL of iodoacetamide (200 mM) was added, and the reaction mixture was further incubated in the dark for 1 h. Twenty-nine microliters of ammonium bicarbonate buffer was further added to reduce the final concentration of guanidine hydrochloride to 0.6 M (final volume of 100 μL). One microliter of trypsin in 50 mM ammonium bicarbonate buffer (pH 8.0, 1 μg/ μL) was added, and the samples were incubated at 37 °C for 20 h. The samples were then analyzed by LC-MS.[61,73]

***Rapid quench experiment for the ΨMP glycosidase reaction***: The reaction mixture containing YeiN (500 μM) preincubated with ribose 5-phosphate (1.5mM) and $Mn^{2+}$ (1.5 mM) in syringe A was reacted with Uracil (500μM) present in syringe B in the KinTeK rapid quench apparatus. The reaction was quenched using 5M HCl at various time points and quenched reaction mixtures were collected. The reaction samples were neutralized by 5M potassium phosphate buffer (pH ~10) to pH 6.5 and were analyzed by HPLC. The amount of pseudouridine monophosphate formed per unit enzyme concentration was plotted against time to determine the pre steady state kinetics parameters.

# CHAPTER III

# MENAQUINONES: STRUCTURE, FUNCTION AND BIOSYNTHESIS

## Introduction

Menaquinones (MK) belong to a class of compounds known as isoprenoid quinones that are composed of a polar head group and hydrophobic side chain.[74, 75] The non-polar isoprenoid side chain gives the molecules a lipid-soluble character and anchors them in membrane lipid bilayers, whereas the hydrophilic head group enables interaction with hydrophilic parts of proteins in the polar interphase region of the membranes.[76, 77]

A vast majority of biological isoprenoid quinones, which are also termed as prenylquinones, belong to either benzoquinones or naphthoquinones.[75] The two most important members of benzoquinones are ubiquinones (UQ) and plastoquinones (PQ); differing in the pattern of ring substitution, chain length of the isoprenoid side chain, degree of saturation as well as in the presence of additional groups (Figure 3.1). Some of the important naphthoquinones include menaquinones (MK, vitamin $K_2$), demethylmenaquinones (DMK) and phylloquinones (PhQ, vitamin $K_1$) (Figure 3.1). They are represented as Q-n or MK-n, where n refers to a side chain composed of varying numbers of five carbon isoprenoid residues and the number of such units is indicated.[78] Thus, Q-8 and MK-8 indicate ubiquinone and menaquinone with side chains of 40 carbon atoms, respectively. MK is considered as vitamin, while UQ is not, due to the fact that vitamin K is an essential nutrient (cannot be synthesized) in mammals, whereas UQ can be synthesized from tyrosine and hence it is not an essential nutrient.

Prokaryotes contain isoprenoid quinones of both kinds in varying amounts. Most aerobic Gram-negative bacteria contain Q as the sole quinone, while the aerobic Gram-positive bacteria contain MK as the only quinone. Most of the anaerobic bacteria, irrespective of Gram-positive or negative, contain MK as their sole quinone. However, Gram-negative facultative anaerobes such as *E.coli* contain both Q and MK (mainly UQ-8, MK-8 and DMK-8), and hence they are able to grow under both aerobic and anaerobic conditions.[74, 75] Structures of some isoprenoid quinones are shown in Figure 3.1 (Benzoquinones include: ubiquinone (UQ-8), **24** and plastoquinone (PQ-9), **25**; naphthoquiones include: menaquinone (MK-8), **26**, demethylmenaquinone (DMK-8), **27** and phylloquinone (PhQ), **28**).

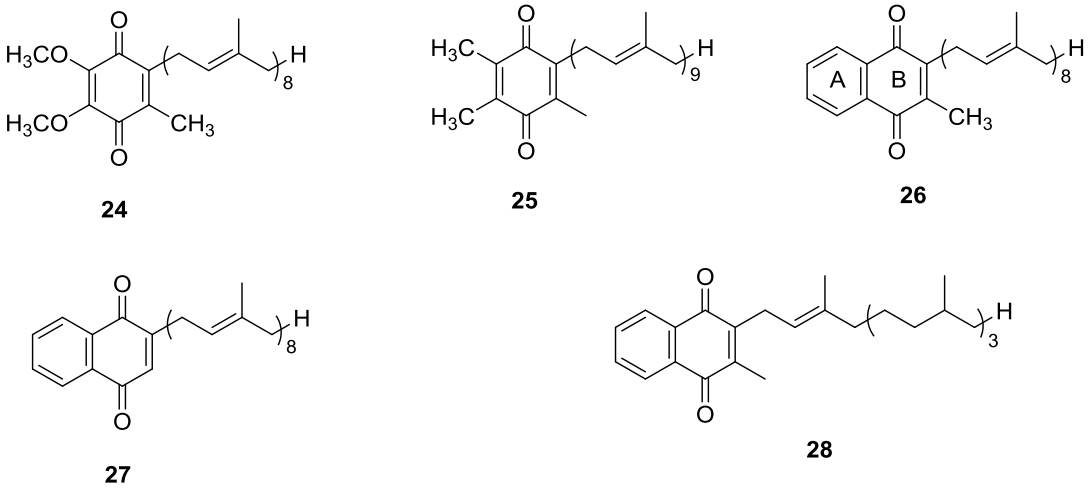



**Figure 3.1**: Structures of isoprenoid quinones.

The isoprenoid quinones are obligatory components of the electron transport chain in prokaryotes. The quinone ring can undergo two-step reversible reduction leading to the more polar quinol form which is thought to preferentially localize in polar, interphase region of membranes[77] (Figure 3.2). This ability of reversible reduction and the presence of the lipid-soluble side chain make them ideal candidates for their function as hydrogen shuttles between different protein complexes of biological membranes. Figure 3.2 shows the redox reactions of the quinone ring in which Q (**29)** is the oxidized form, $Q^{-\cdot}$ (**30**) is the charged semiquinone radical which is unstable but can be stabilized by interaction with proteins and $QH_2$ (**31**) is the stable non charged hydroquinone or quinol form.

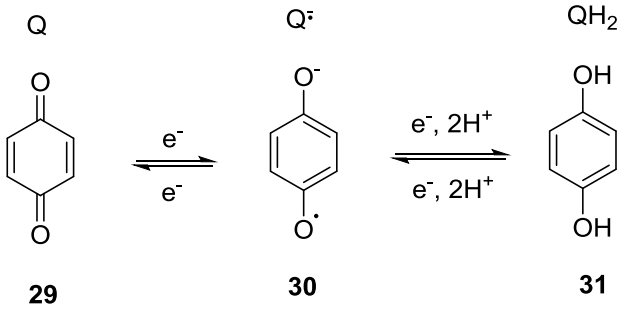



**Figure 3.2**: Redox reactions of the quinone ring.

They function as membrane-bound, mobile hydrogen carriers but some of these molecules are permanently associated with proteins and participate in electron transport within protein complexes. In *E.coli*, under anaerobic conditions, when nitrate is the electron acceptor, UQ is the preferred quinone, however, MK can substitute in the absence of UQ.[79] But when fumarate, trimethylamine N-oxide (TMAO) or dimethylsulfoxide (DMSO) serves as electron acceptors, under anaerobic conditions, the

presence of MK is essential.[80-82] Moreover, the reduced forms of the isoprenoid quinones have pronounced antioxidant properties and protect membranes from lipid peroxidation and deleterious effects of reactive oxygen species on membrane components.[75]

**Menaquinones: Occurrence and Function**

Menaquinones (MK) are the most widespread microbial respiratory quinones that can be found in many groups of archaea and bacteria such as different categories of proteobacteria, gram-positive bacteria, green sulfur bacteria, green filamentous bacteria, dienococci and flavobacteria.[83-86] The isoprenoid side chain of menaquinones in these organisms is most frequently composed of 6-10 prenyl units (often fully unsaturated), but also homologues with 1 to 14 prenyl units were also found.[74]

In microorganisms, menaquinone is a lipid soluble molecule that has low midpoint redox potential and shuttle electrons between the membrane bound protein complexes/ redox enzymes in the electron-transport chain of the respiratory and photosynthetic systems. A representation of the bacterial respiratory chain is shown along with the enzymes that use MK during electron transfer (Figure 3.3).[87]

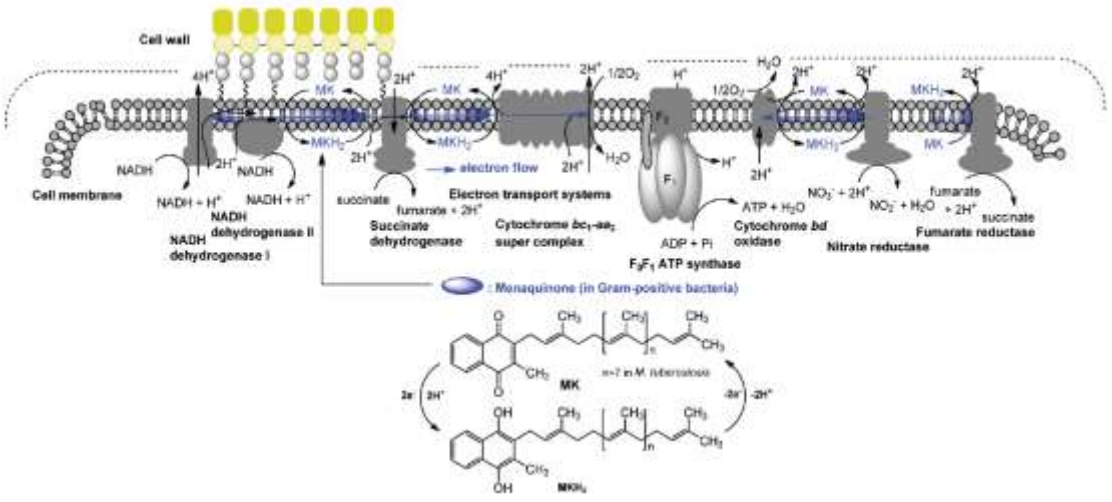



**Figure 3.3**: A pictorial representation of the bacterial electron transport chain.[87]

Many of the enzymes of the respiratory chain that use MK/$MKH_2$ system are able to generate proton gradient, which is crucial for energy conserving processes. Proton gradient can be generated by two mechanisms: pumping the protons across the membranes and the so called redox loop mechanism, where the reduction of one substrate, accompanied by proton(s) binding, occurs on one side of the membrane, whereas oxidation of the other substrate, accompanied by proton(s) release, occurs on the opposite site, while electrons are transferred through the protein.[88, 75] Some of the examples of bacterial respiratory enzymes that use MK as a cofactor include succinate: quinone oxidoreductases,[89-92], quinol: fumarate oxidoreductase,[93] NADH: MK oxidoreductases[94], hydrogenase[95], thiosulfate reductase[96], formate dehydrogenase[97], nitrate reductase[98,99], , heme:copper oxidases[100], cytochrome *bc* complexes[101] and cytochrome *bd* complexes.[102]

Menaquinones also participate in electron transfer chain in bacterial photosynthesis. MKs are found both in photosystem II (PS II)-type reaction centers of purple bacteria and green filamentous bacteria.[103,104] They also occur in photosystem I (PS I)-type reaction centers of green sulfur bacteria[105] and heliobacteria.[103,106,107]They are also found in high amounts in PS-I of certain cyanobacteria.[108]

Besides the function of MK in electron transfer, its participation in cellular signaling of *Klebsiella pneumoniae* was found, where MK takes part in regulation of nitrogen fixation. It was observed that NifL protein, acting as a corepressor in the regulation of expression of nitrogen fixation genes, can be reduced by $MKH_2$ under anerobic conditions, followed by NifL sequestration on the cellular membrane.[109]

Vitamin K plays a specific physiological role in blood coagulation in mammals. It is an essential vitamin in humans and other mammals, where it functions as a co-substrate in the carboxylation of glutamic acid residues in proteins involved in blood clotting[110] and bone formation[111]. Under normal conditions, specific glutamate residues in the N-terminal region of the proteins involved in blood coagulation and bone metabolism are carboxylated to γ-carboxyglutamates (Gla) with the help of vitamin K (Figure 3.4) which, in turn, activate them to bind $Ca^{2+}$. Anti-clotting drugs such as warfarin targets the enzyme vitamin K epoxide reductase (Figure 3.4) that converts vitamin K epoxide back to the active form of the cofactor.[112] Vitamin K is also involved in calcium metabolism, tissue calcification and cell cycle regulation.[113] Vitamin K deficiency in

humans[114, 115] leads to the synthesis of abnormal forms of coagulation factors. Other proteins in bone and dentin (such as Osteocalcin) [111], kidney and spermatozoa also contain Gla. A major portion of human requirement for vitamin K is provided by the microflora of the gut. The rest is obtained from plant sources.[115, 116] Figure 3.4 shows Vitamin-K dependent carboxylation of glutamate (Glu) residues (**32**) of blood clotting factors by γ-glutamyl carboxylase with help of reduced form of vitamin K (**33**) forming γ-carboxyglutamate, Gla (**34**) and vitamin K epoxide 2, 3-epoxide (**35**). Gla helps in binding $Ca^{2+}$ ions to initiate the blood clotting process. Anti-clotting drugs such as warfarin (**36**) inhibit vitamin K epoxide reductase which converts **35** back to **33**.

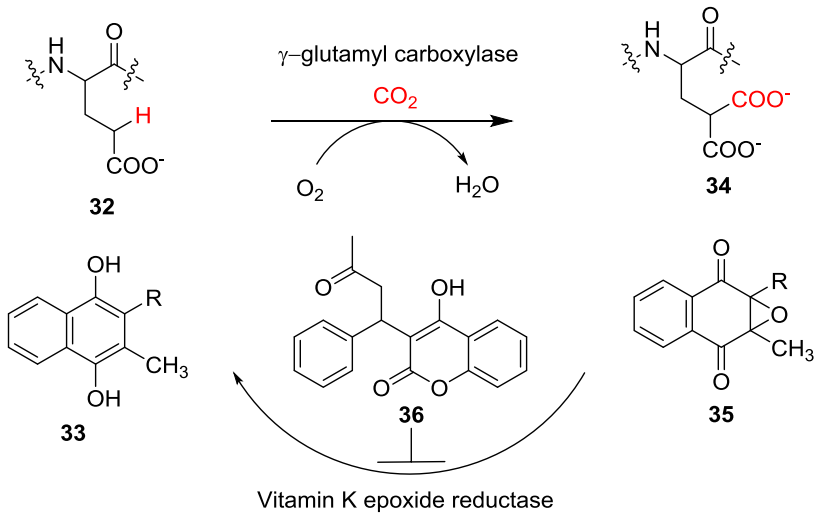



**Figure 3.4**: Vitamin-K dependent carboxylation of glutamate (Glu) residues.

**Biosynthesis of menaquinone**

During the biosynthesis of menaquinones (and other isoprenoid quinones), the head group precursor and the isoprenoid side chain are synthesized separately. Then both counterparts are condensed by an enzyme belonging to the prenyltransferase family, followed by further modifications on the resulting product.[75] The primary precursors for the isoprenoid chain are dimethylallyl diphosphate (DMAP) and isopentenyl diphosphate (IPP). The isoprenoid side chain is formed from these precursors through series of condensation reactions carried out by a specific prenyl diphosphate synthase. There are two distinct pathways leading to the biosynthesis of these precursors: mevalonate (MVA) pathway and 1-deoxy-D-xylulose-5-phosphate (DXP) pathway and different taxonomic groups utilizes different pathway.[117] For instance, majority of bacteria use DXP pathway (exceptions include *Myxococcus fulvus* which use MVA pathway and *Streptomyces* which possess both pathways), while archaea are believed to use the MVA pathway.[117] All the animals and fungi take advantage of the MVA pathway [117] while higher plants use both of them.[118]

In prokaryotes, there are two different pathways for the biosynthesis of the naphthoquinone core (polar head group) of menaquinone (MK). Both pathways are branches of the shikimate pathway that diverge at chorismate and converge at DMK.[119, 120] The vast majority of bacteria produce MK by the well-established o-succinylbenzoate (OSB) pathway, which is also known as the classical pathway.[121, 122] Recently an alternative pathway for producing menaquinone, termed as futalosine pathway,[123, 124]

was discovered in *Streptomyces coelicolor* which was later demonstrated in *Helicobacter pylori, Campylobacter jejuni* and *Thermus thermophilus.* Moreover, occurrence of this new pathway was confirmed in several Gram negatives and Archaea, mainly in chlamydia and spirochetes.[123] This pathway is distributed only in prokaryotes and absent in eukaryotes, including fungi and yeasts. However, till date, no microorganism is found that possesses both the classical and the alternative pathway.[123]

**Classical pathway (o-succinylbenzoate pathway)**

Biosynthesis of MK by this pathway, studied in *E.coli*, has been elucidated on the basis of isotopic tracer experiments, isolation of mutants blocked in different steps, isolation and identification of accumulated intermediates in the mutants and by in-vitro assays using recombinant enzymes.[121, 122] Results from the radioisotope tracer experiments led to the conclusion that the naphthoquinone portion of MK arise from shikimate derived chorismate and 2-ketoglutarate while methyl and isoprenoid side chains are derived from S-adenosylmethionine (SAM) and an isoprenyl alcohol pyrophosphate ester respectively. The isotopic labelling pattern is summarized (Figure 3.5).[125-129]

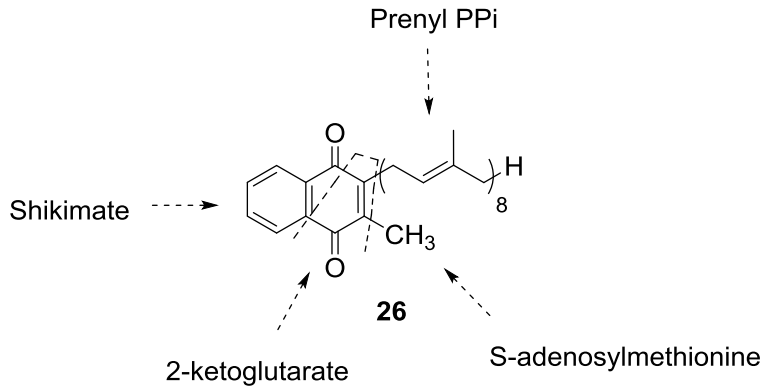



**Figure 3.5**: Primary biosynthetic precursors of MK in the classical pathway

Subsequently, it was shown that the benzenoid aromatic compound, o-succinylbenzoate (OSB)[130] and the naphthalenoid aromatic compound, 14-dihydroxy-2-naphthoate (DHNA)[131] were incorporated into the naphthoquinone ring of MK validating the intermediacy of these compounds in the pathway (Figure 3.6). This result was confirmed by the demonstration that *menB* and *menA* mutants of *E. coli* excrete OSB and DHNA into the culture medium, respectively.[132] Later on it was also shown that carbon one of glutamate (2-ketoglutarate) was lost and consequently not incorporated to OSB or finally into MK (Figure 3.6).[133]

In this classical pathway studied in *E.coli* in great detail, eight enzymes encoded by *men* genes [120, 122,] and the recently identified DHNA-CoA thioesterase[134] are involved in MK formation. In this pathway (Figure 3.6), chorismate, **39** which is derived from the shikimate pathway, is initially converted into isochorismate, **40** by the isochorismate synthase, MenF. In the next step, 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-

caroxylate (SEPHCHC), **42** is formed from **40** and 2-ketoglutarate, **41** in the reaction catalyzed by MenD, which is a thiamin pyrophosphate (TPP) dependent enzyme.

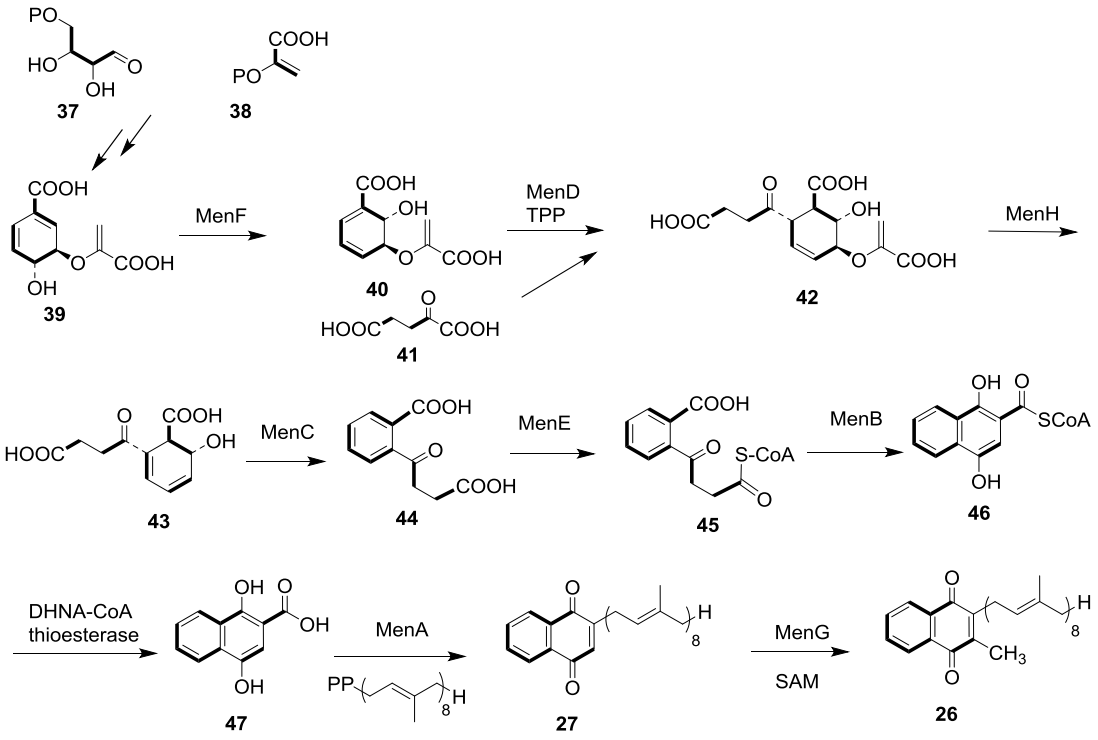



**Figure 3.6**: Classical pathway for the biosynthesis of menaquinone, **26** (MK-8)

Thereafter, MenH catalyzes the conversion of **42** into 2-succinyl-6-hydroxy-2, 4-cyclohexadiene-1-carboxylate (SHCHC), **43.** This compound is dehydrated by MenC to give the first aromatic compound of the pathway, o-succinylbenzoate (OSB), **44.** This is followed by its attachment with CoA yielding o-succinylbenzoyl-CoA, **45** in a reaction catalyzed by MenE. The latter is then converted into 1, 4-dihyroxy-2-naphthoyl-CoA, **46** (DHNA-CoA), by MenB. It is then converted to 1,4-dihydroxy-2-naphthoate, **47** (DHNA) by a recently discovered enzyme DHNA-CoA thioesterase.[134] In the last two

steps of the pathway, MK, **26** is synthesized by MenA and MenG, which catalyze prenylation (using octaprenyl pyrophosphate for MK-8) and S-adenoylmethionine (SAM) dependent methylation, respectively. The number of prenyl units in the attached isoprenoid side chain is determined by the length of isoprenoid pyrophosphate and may differ among species (e.g. *E.coli*- maily MK-8, *Bacillus firmus*-MK-7 etc.).[85] The majority of enzymes involved in MK synthesis are found to be soluble and mechanistic studies on each of these enzymes have been performed.[122]

The MK and phylloquinone (vitamin K1) biosynthetic pathways are unique to bacteria and plants, respectively, and are absent in humans and animals. Hence, there is great research interest and opportunity to discover novel chemicals that will inhibit the enzymes of the pathway and could be used as herbicides and chemotherapeutic agents against human pathogens such as multidrug-resistant *Mycobacterium tuberculosis* and methicillin-resistant *Staphylococcus aureus* (MRSA). Several studies relating to discovery of inhibitors and antibiotic compounds targeting the enzymes on the menaquinone biosynthetic pathway have been reported.[87, 135-138]

**Alternative pathway (futalosine pathway)**

Based on bioinformatics studies, it was found that some strains of *Streptomyces* such as *S. coelicolor* A3(2) and *S. avermitilis* do not possess the orthologs of the *men* genes of the classical *E.coli* pathway, except for two genes *menA* and *menG*, even though they synthesize MK.[123,139,140] In addition to them, *Epsilon* and *Delta* categories of

proteobacteria, including *Helicobacter pylori* and *Campylobacter jejuni*, which are known to cause gastric carcinoma and diarrhea, respectively, also lacked the *men* gene orthologues.[123] These facts strongly suggested that the naphthoquinone moiety of MK is synthesized via an alternative pathway in these bacteria. From the isotopic labeling studies done on *Streptomyces sp*., using $^{13}C$ glucose, it was established that the A ring of MK is derived from the shikimate pathway while the remaining carbons of the B ring are reported to be derived from phosphoenol pyruvate (PEP) and an unidentified metabolite derived from C-5 and C-6 carbons of glucose.[124] The labelling pattern is shown in Figure 3.7.

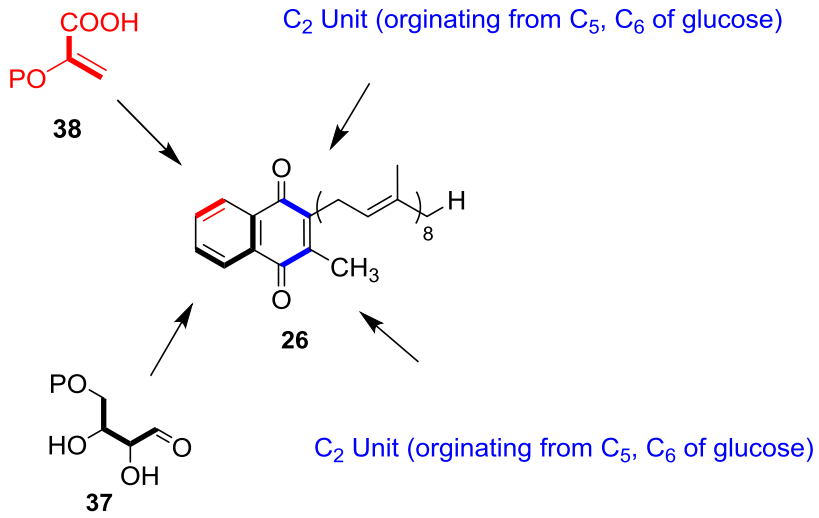



**Figure 3.7**: Labelling pattern in the alternative menaquinone biosynthetic pathway.

Furthermore, based on the distinct labelling patterns and biochemical experiments, 1, 4-dihydroxy-6-naphthoate was deduced as an intermediate of this pathway. Later on, based on *in-vivo* studies using MK auxotrophs, it was established that the branch point for this

pathway is also at chorismate derived from the shikimate pathway.[123] Finally, Tomoshige Hiratsuka *et al* were able to delineate the general outlines of an alternate menaquinone biosynthetic pathway in *S. coelicolor* A3(2) with the help of bioinformatics screening and various biochemical techniques such as gene knockouts, shotgun cloning with isolated mutants and in vitro studies with recombinant enzymes.[123, 141, 142] In this alternative pathway (Figure 3.8), designated as the futlaosine pathway, futalosine, **49** is proposed to form from chorismate and presumably inosine, **48** and phosphoenolpyruvate (PEP), **38** in a complex reaction catalyzed by MqnA (SCO4506). Recently, 6-amino-6-deoxyfutalosine is also reported as the first intermediate of this pathway in *C. jejuni* and *H. pylori*[143] in place of futalosine discovered in *S.coelicolor*.

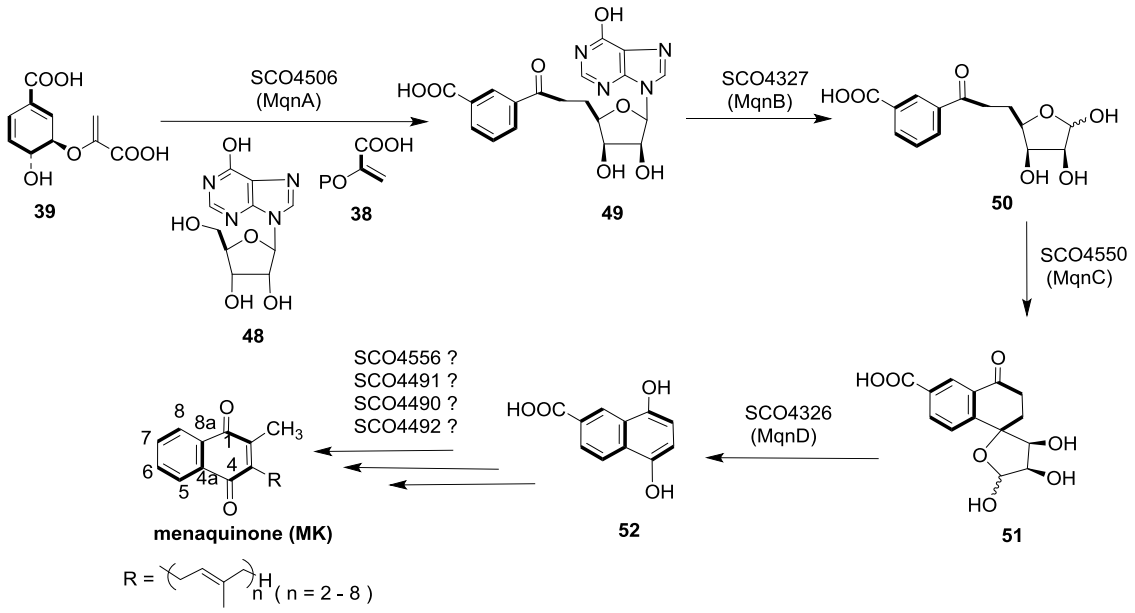



**Figure 3.8**: Alternative pathway for menaquinone biosynthesis

In the next step, **49** is converted to dehypoxanthine futalosine (DHFL), **50** by futalosine hydrolase, MqnB (SCO4327)[144], which is then transformed into cyclic dehypoxanthine futalosine (CDHFL), **51** by the radical SAM enzyme MqnC (SCO4550)[143]. MqnD (SCO4326) then converts **28** to 1, 4-dihydroxy-6-naphthoic acid, **52** which is an important intermediate on the pathway. In the later steps of the pathway, based on the annotation of the open reading frames of *S. coelicolor* A3(2), it was presumed that SCO4491 (prenylation), SCO4556 (methylation), SCO4490 (decarboxylation), and SCO4492 (decarboxylation) would be involved.[123]

**Research opportunity**

Discovery of this alternative pathway opened up new prospects in the investigation of some novel and intriguing enzymology involved. *In vitro* reconstitution using recombinant enzymes and in depth mechanistic studies on these enzymes are yet to be achieved. Moreover, since humans and commensal intestinal bacteria, such as lactobacilli, lack this pathway; hence it represents an attractive target for the development of chemotherapeutics against some human pathogens. Hence, like the classical pathway, this new pathway also represents a novel target for drug discovery, however, against some different pathogenic strains of bacteria.[123, 141] So far, one study has been reported in which an attempt has been made to find a suitable inhibitor for this pathway in *H.pylori*.[145] Our aim was to investigate the *in-vitro* reconstitution of the first step of this important menaquinone[146] biosynthesis pathway and carry out mechanistic studies on these enzymes.

# CHAPTER IV

## MENAQUINONE BIOSYNTHESIS: RECONSTITUTION, MECHANISTIC AND STRUCTURAL STUDIES ON CHORISMATE DEHYDRATASE, MqnA*#

### Introduction

In the early steps of the alternative menaquinone biosynthesis pathway discovered in *S. coelicolor* A3(2),[123] it was proposed that MqnA catalyzes the formation of futalosine from chorismate, inosine and phosphoenolpyruvate using unknown chemistry (Figure 4.1).[123,124] However, attempts to reconstitute MqnA *in vitro* were unsuccessful.[123] The aim here was to reconstitute the activity of MqnA and verify the proposed step.

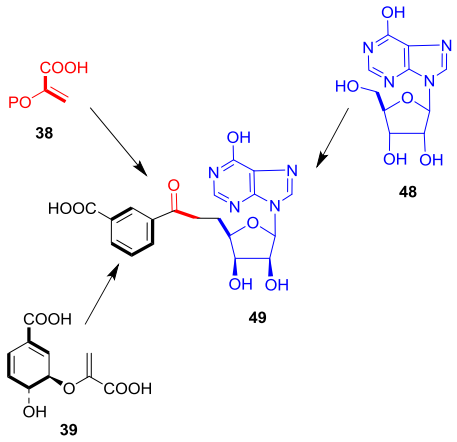



**Figure 4.1**: Proposed precursors of futalosine in the alternative pathway.

*Part of this chapter is reprinted with permission from "Menaquinone Biosynthesis: Formation of Aminofutalosine Requires a Unique Radical SAM Enzyme" by Mahanta, N.; Fedoseyenko, D.; Dairi, T.; Begley, T. P., 2013. *J. Am. Chem. Soc.*, 135, 15318-15321, Copyright [2013] by American Chemical Society.

#(Nilkamal Mahanta carried out all the biochemical studies. Dmytro Fedoseyenko synthesized compound **53**. Cynthia Kinsland, Ealick lab, cloned MqnA**.** Katherine A Hicks, Ealick lab, performed the structural studies on MqnA).

**Results and Discussion**

***HPLC analysis of the MqnA reaction mixture:*** Genetic studies previously identified MqnA as essential for futalosine formation. To identify the reaction catalyzed by MqnA, the *mqnA* gene from *T. thermophilus* HB8 (TTHA0803) was amplified by PCR, cloned into the p28nT expression vector, overexpressed in *E. coli* BL21 (DE3), and purified by Ni-NTA affinity chromatography (Figure 4.2). Incubation of chorismate with MqnA resulted in the formation of a new product as shown in the HPLC chromatogram (Figure 4.3), which was identified as 3-[(1-carboxyvinyl)- oxy]benzoic acid, **53** by NMR and MS analysis and comparison with an authentic synthetic reference compound.[147] Shown below is the SDS-PAGE analysis of MqnA in Figure 4.2 [The molecular weight is ~30.5 kDa. A) Protein molecular weight markers and B) Purified MqnA)]

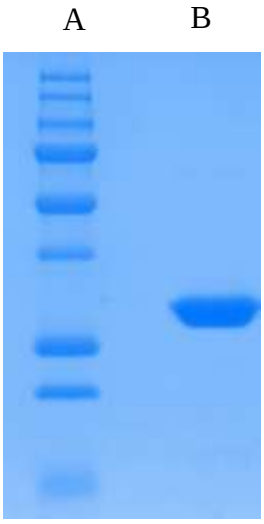



**Figure 4.2**: SDS-PAGE analysis of MqnA.

Neither adenosine/inosine nor phosphoenolpyruvate/pyruvate was accepted as substrate by MqnA in these assays as previously proposed. Formation of aminofutlaosine (or futalosine) was not observed in the reaction mixture after analysis by HPLC and LCMS.

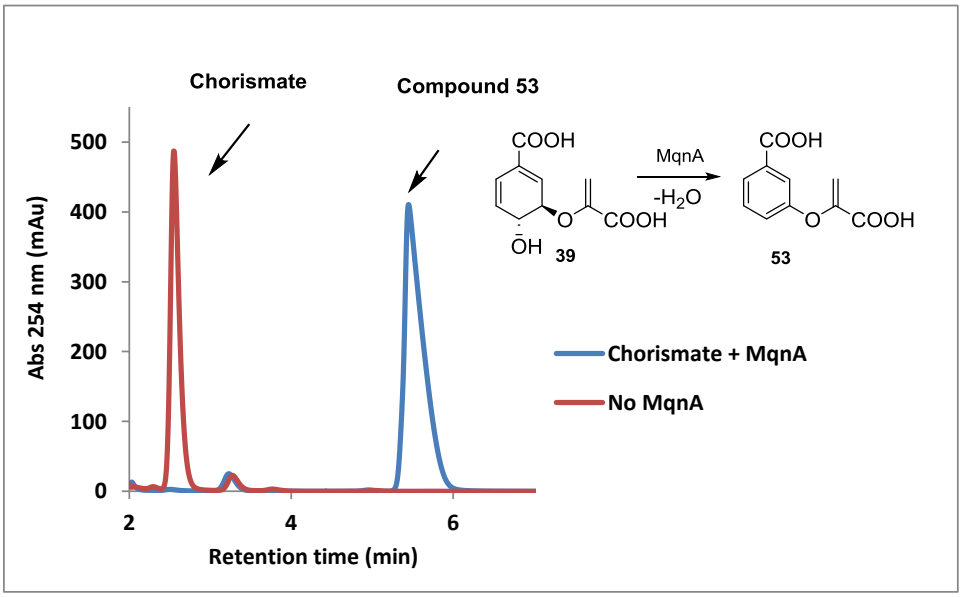



**Figure 4.3**: HPLC analysis of the reaction catalyzed by MqnA.[147]
[Blue trace: chromatogram of the full reaction mixture (MqnA + chorismate) Red trace: chromatogram of the control reaction mixture lacking MqnA. The inset shows the structure of the identified and characterized MqnA reaction product, **53]**

***Characterization of compound 53 formed in the MqnA reaction*:** The compound formed from chorismate in the MqnA catalyzed reaction was purified by HPLC and characterized by NMR (Figures 4.4-4.7, Figure 8, Appendix A) and MS (Figure 4.8). Chorismate (3 mM) was incubated with MqnA (400 µM) in 100 mM Tris-HCl buffer (pH 7.5) and the reaction was allowed to proceed for 3 hrs. The protein was removed by ultrafiltration using 10 kDa centrifugal filters. The reaction product was purified by

HPLC (eluted at 5.1 min) and characterized by NMR (Figures 4.4-4.8) and LC-MS analysis (Figure 4.9).[147]

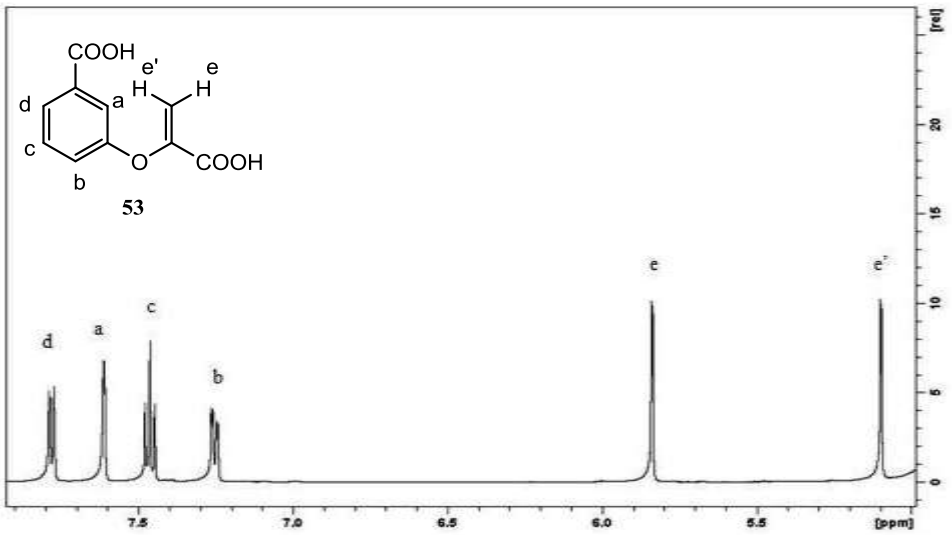



**Figure 4.4:** $^{1}H$ NMR (500 MHz, $CD_3OD$) spectrum of compound **53.**[147]

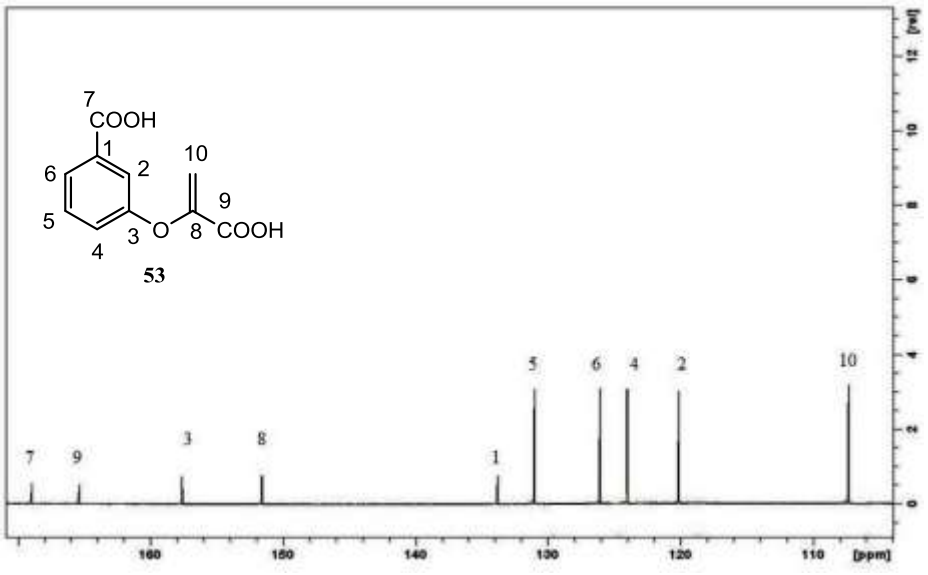



**Figure 4.5:** $^{13}C$ NMR (125 MHz, $CD_3OD$) spectrum of compound **53**.[147]

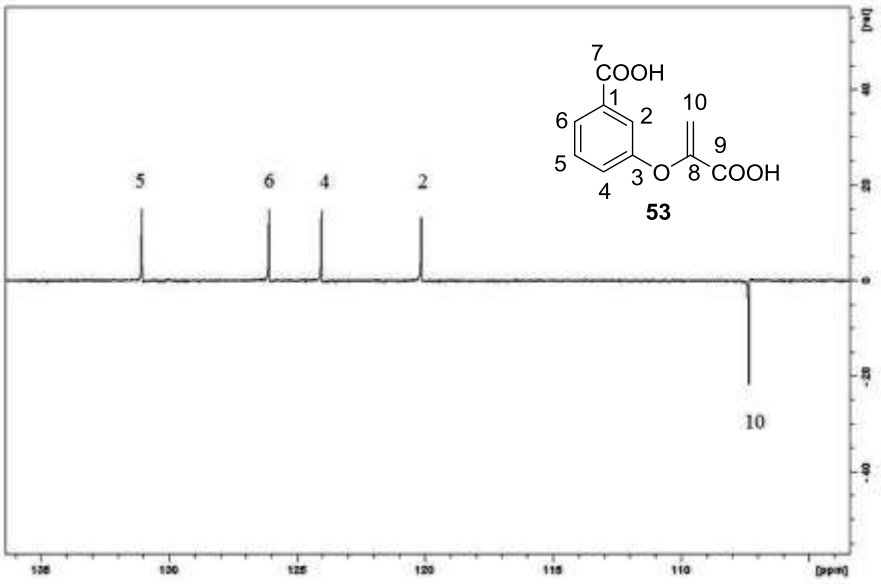



**Figure 4.6**: $^{13}C$ DEPT-135 spectrum (125 MHz, $CD_3OD$) of compound **53.**[147]

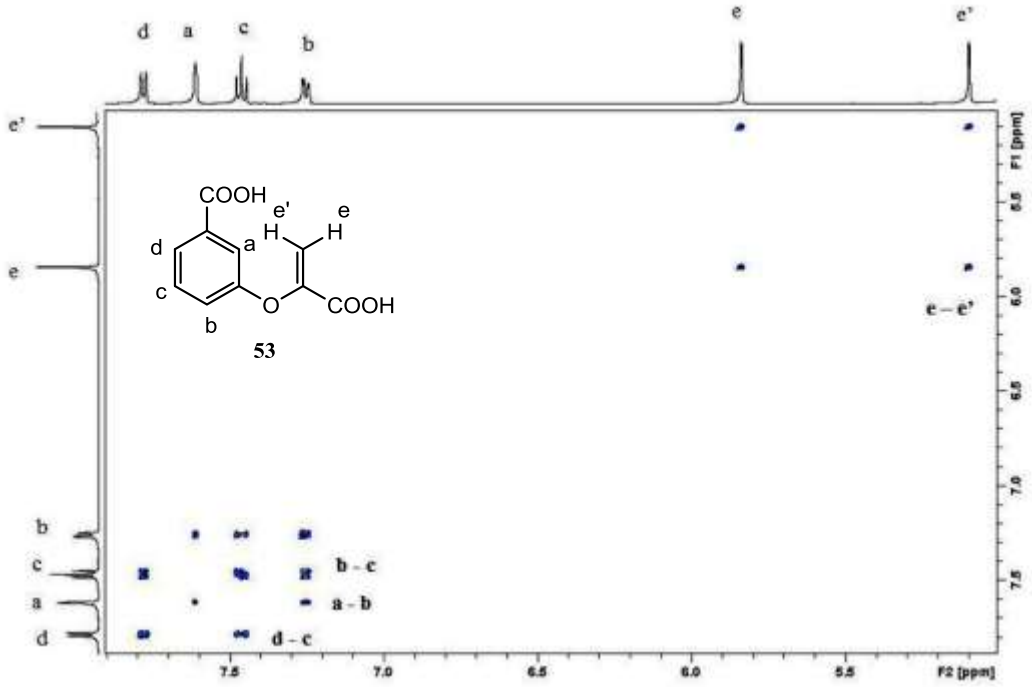



**Figure 4.7**: $^{1}H$-COSY spectrum of compound **53.**[147]

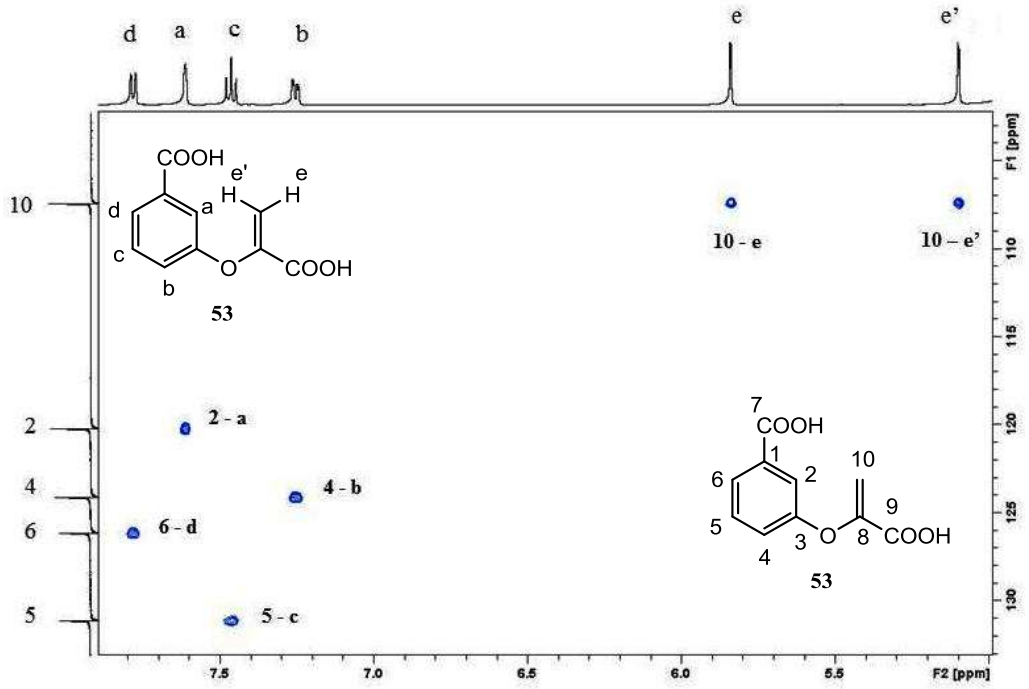



**Figure 4.8**: HSQC spectrum of compound **53**.[147]

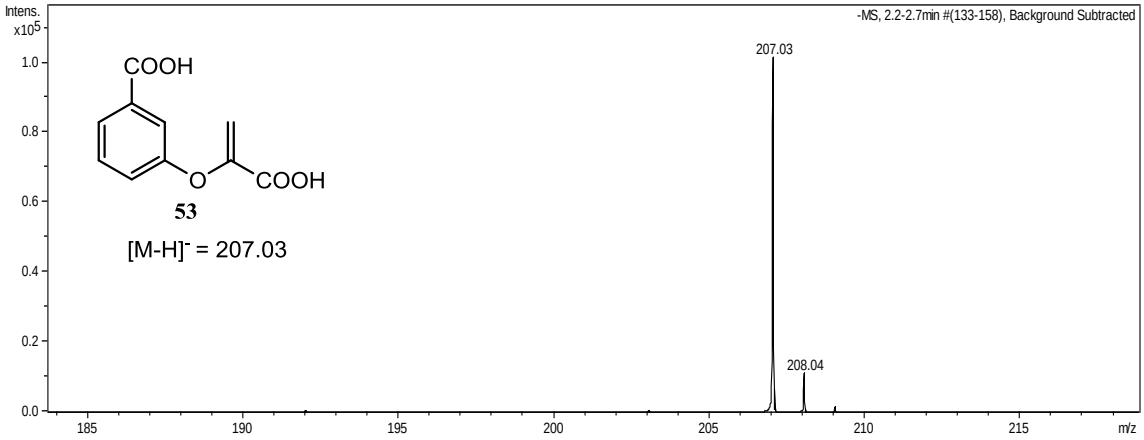
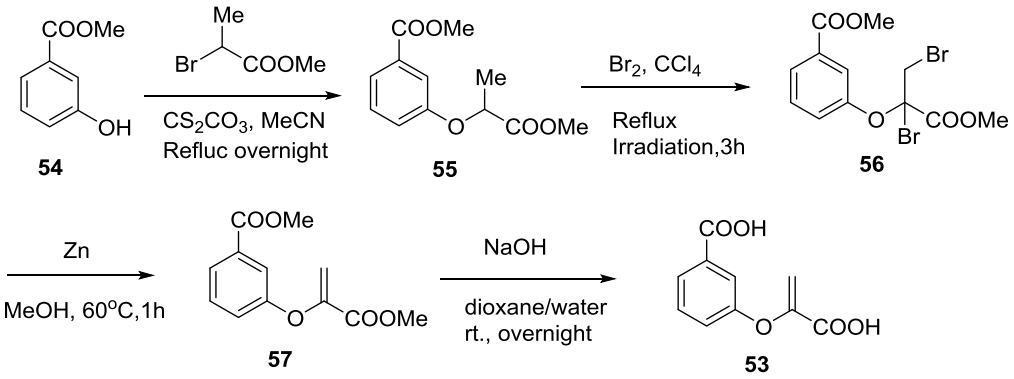



**Figure 4.9**: ESI-MS (negative ion mode) analysis of compound **53**.[147]
[In the mass spectrum, m/z 207.03, $[M-H]^-$ corresponds to 3-((1-carboxyvinyl) oxy) benzoic acid, **53** (theoretical m/z 207.03)]

***Synthesis of compound 53***:The synthetic scheme for making 53 as a authentic reference compund to complete MqnA product characterization is shown Scheme 4.1.[148, 149]




**Scheme 4.1**: Synthetic scheme for 3-((1-carboxyvinyl) oxy) benzoic acid, **53**

***Mechanistic studies on MqnA*:** Since the leaving groups in chorismate are not in an anti periplanar arrangement, the classical E2 elimination is not likely in this dehydration reaction. However, an E1 or E1cb mechanism is probable in this reaction. Following mechanistic proposal (Figure 4.10) is put forward for MqnA through the carbocation **58** (E1 mechanism) or carbanion **59** (E1cb mechanism).

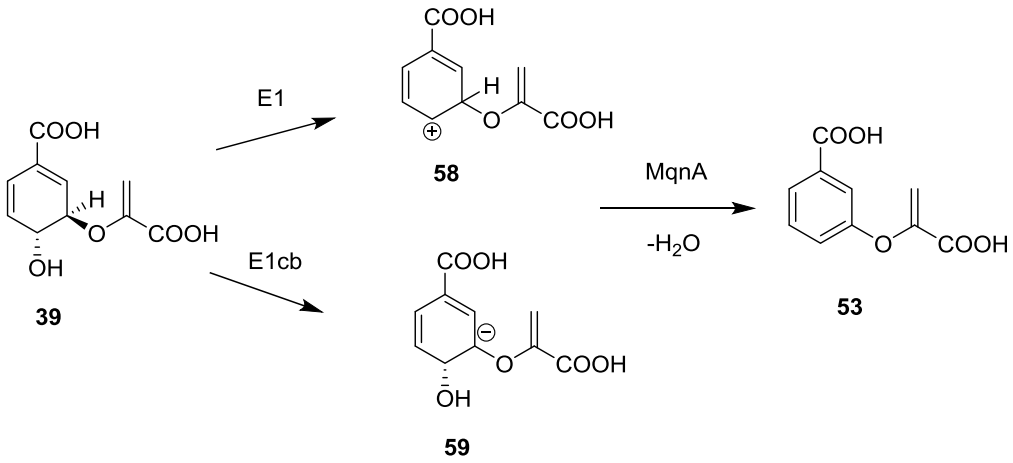



**Figure 4.10**: Mechanistic proposal for MqnA.

To differentiate between these two mechanisms, MqnA reaction was run in deuterated buffer and incorporation of deuterium was looked for in unreacted starting material chorismate by NMR spectroscopy. Incorporation of deuterium into C-3 of chorismate would provide evidence for an E1cb mechanism. The reaction ([Chorismate] = 500 µM, [MqnA]= 1 µM) was performed in 100% deuterated buffer (potassium phosphate, pH 7.5) and quenched after 5 min using 6M guanidine hydrochloride. The NMR spectra of the starting material chorismate and unreacted chorismate in the reaction mixture were compared to see if there is any difference in the peak shapes or splitting patterns of hydrogens at C-2 and C-3 due to H/D exchange (Figure 4.11 and 4.12). However, no such difference was observed in NMR spectra suggesting that H/D exchange is not occurring. This argued against the E1cb mechanism.

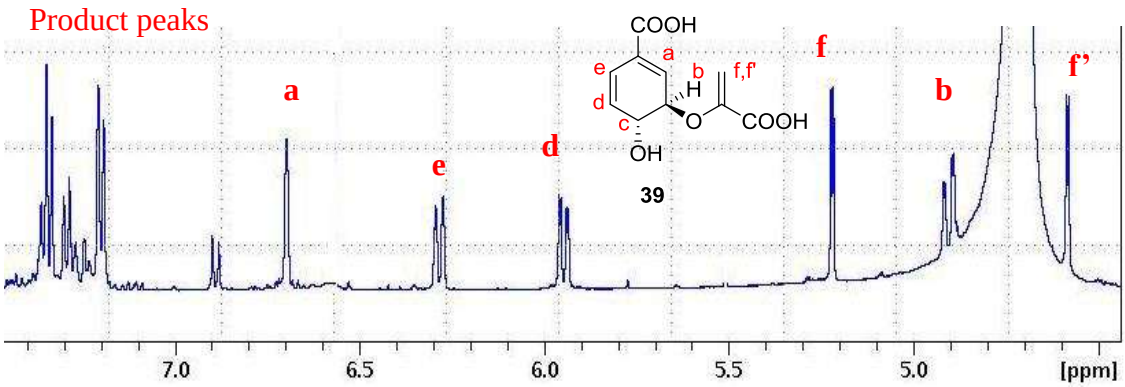



**Figure 4.11**: NMR spectrum of the MqnA reaction mixture

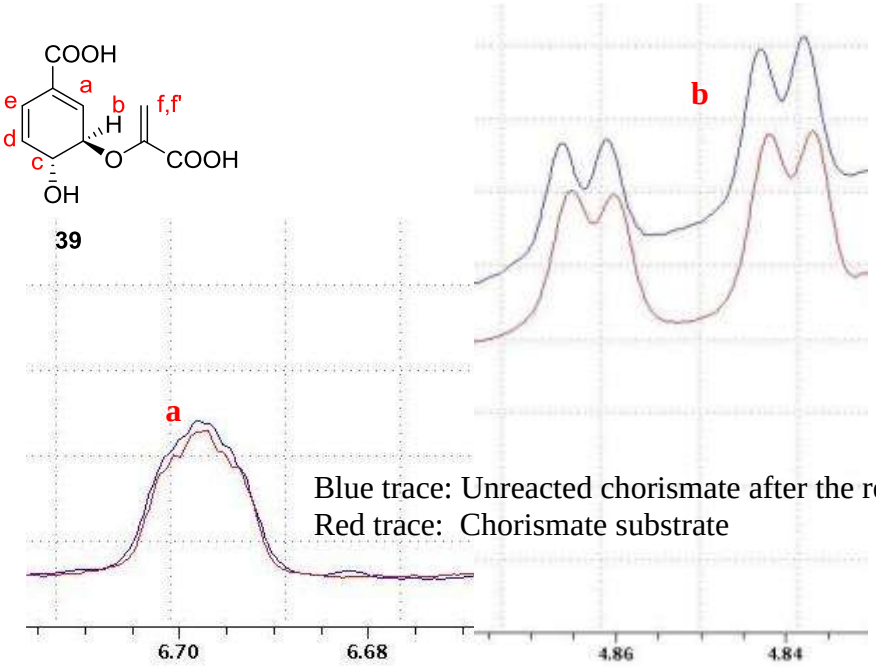



**Figure 4.12**: Comparison of NMR spectra of unreacted chorismate and substrate.

***Steady state kinetics parameters for MqnA*:** Steady state kinetics parameters for MqnA were determined by varying chorismate concentrations while keeping fixed MqnA concentration and monitoring the formation of 3-[(1-carboxyvinyl) oxy]-benzoic acid, **53** by HPLC using the method described above (The buffer used in HPLC analysis was 100 mM potassium phosphate, pH 6.6).[147] Each reaction mixture containing different concentrations of chorismate (50 µM, 100 µM, 150 µM, 250 µM, 500 µM, 750 µM, 1 mM and 1.5 mM) in 100 mM Tris-HCl buffer (pH 7.5) was divided into 90 µL and 10 µL of 1 µM MqnA was added to achieve a final concentration of 0.1 µM. For each set of chorismate concentrations, the reaction was quenched at various time points by adding an equal volume of 8M guanidine hydrochloride. Protein was removed by ultrafiltration. Each sample was analyzed by HPLC and the amount of 3-[(1-carboxyvinyl) oxy]-benzoic acid, **53** formed was plotted against time for each set of substrate concentrations. The initial linear slopes from each set of reactions were then plotted against the chorismate **39** concentrations and fit to the Michaelis Menten equation using KaleidaGraph (Synergy Software). The $k_{cat}$ and $K_m$ for MqnA were found out to be 1.27 ±.02 $s^{-1}$ and 161.6 ± 7.1 µM respectively (Figure 4.13).[147]

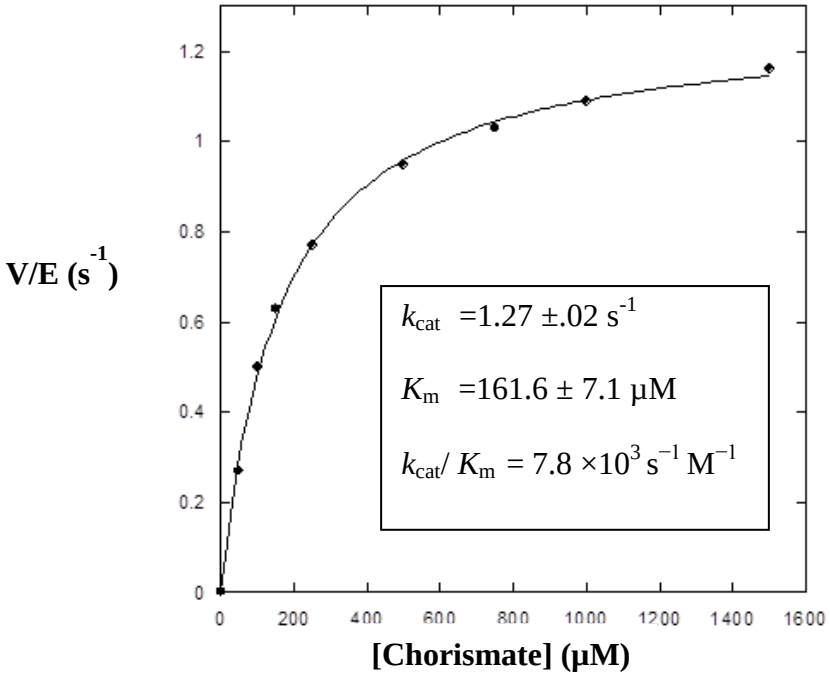



**Figure 4.13**: Michaelis-Menten plot for MqnA.[147]

***Structural studies on MqnA*:** The structural studies are performed on a *D. radiodurans* R1 ortholog (DrMqnA) of *T. thermophilus* MqnA (TTHA0803) and 3-hydroxybenzoic acid (a structural analog of the product **53**) was used in co-crystallization studies (Katherine A Hicks, Ealick lab, Cornell University, unpublished results).

***Overall structure of DrMqnA:*** DrMqnA crystallized in the C222 space group, and contained two protomers in the asymmetric unit (Figure 4.14A). Electron density was present for 271 residues, with the exception of disordered loop regions of residues at the N-terminus consisting of residues 1-16. The protomer consists of 10 β-strands and 11 α-helices (by Katherine Hicks, Ealick lab, Cornell University).

***DrMqnA active site*:** DrMqnA was co-crystallized with a product analog 3-hydroxybenzoic acid (3-HB). This analog makes minimal interactions with *DrMqnA*. The only residue that makes significant interactions with 3-HB is Asn26 (Figure 4.14B), which hydrogen bonds with the carboxylate group on 3-HB orienting the substrate analog in the active site (by Katherine Hicks, Ealick lab, Cornell University).

It was observed form the active site structure of MqnA (Figure 4.14 B) that no protein derived acid/base residues are present near the leaving groups of bound product analog suggesting that both elimination mechanisms could be wrong. Moreover, this also explains the results of the NMR experiment in which no H/D exchange was observed in unreacted chorismate. We conclude that the carboxylic acid group of the pyruvic acid side chain of chorismate **39** is participating in the dehydration mechanism. We propose that it first protonates the hydroxyl group at C-4, making it a good leaving group, water. This leads to the formation of intermediate **60** with carbocation character at C-4, which lowers the pKa of the hydrogen at C-3 and it is being abstracted by the hydrogen bonded water (Figure 4.15).

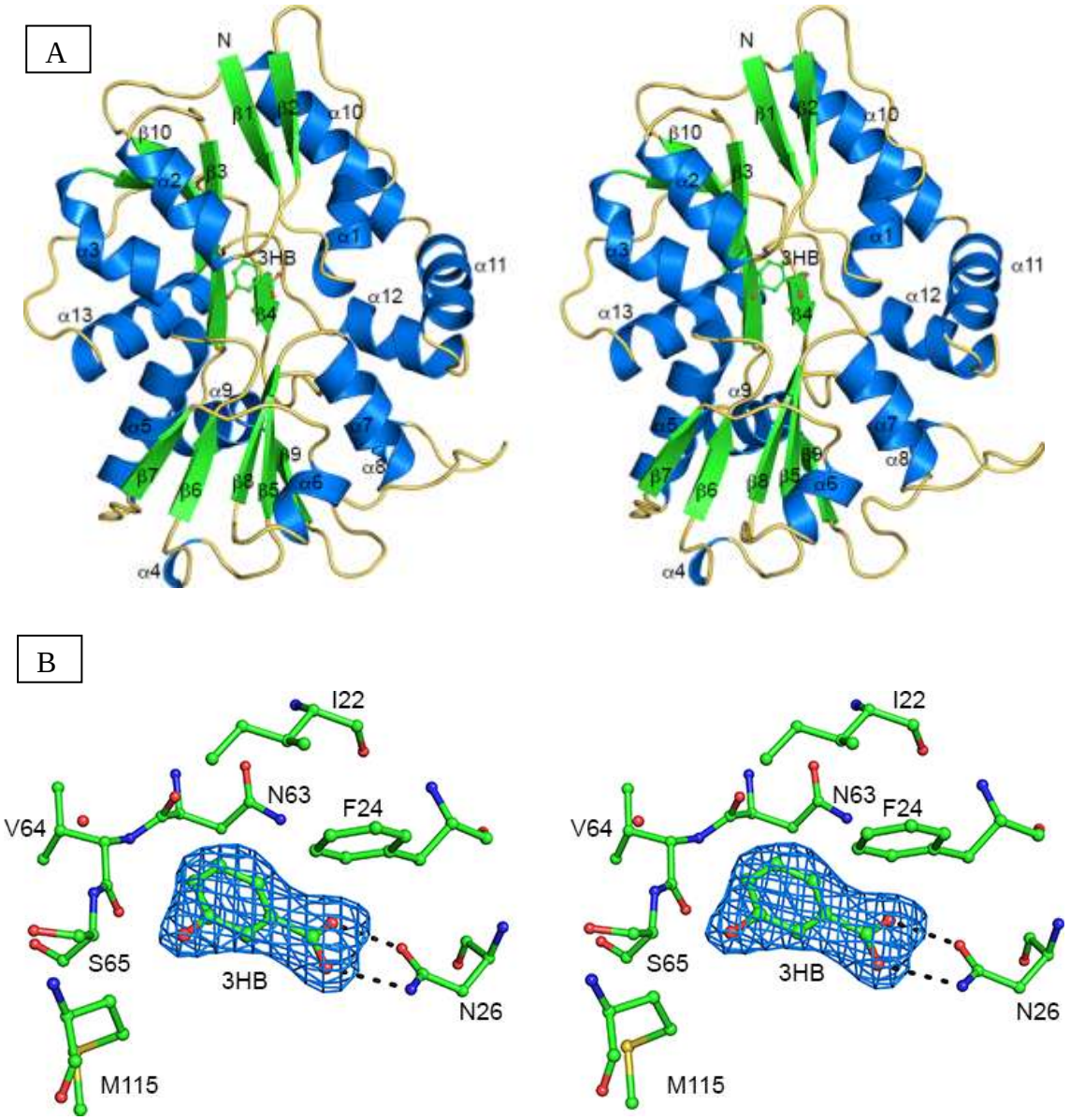



**Figure 4.14**: Crystal structure of MqnA.
[A) Overall structure B) MqnA active site with 3-HB bound]

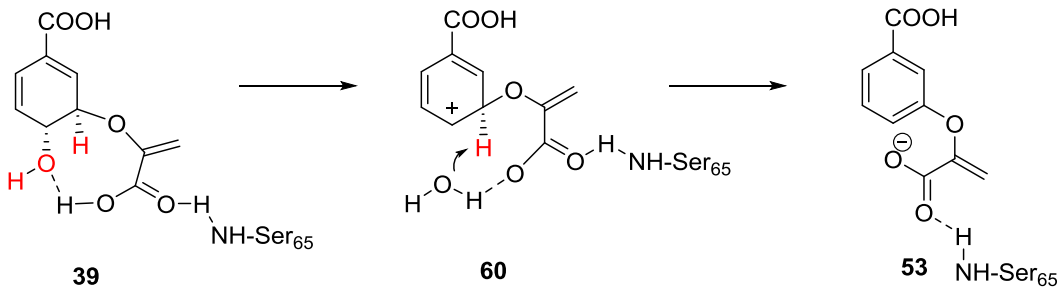



**Figure 4.15:** Proposed mechanism for MqnA

Formation of compound **53** in the MqnA reaction raised questions whether it is an on-pathway intermediate in the biosynthesis of menaquinone or whether it is just an off-pathway shunt product. Further studies are required to confirm this observation. This also proves the original proposal for futalosine formation wrong and indicates that some additional enzymes or cofactors are necessary to complete this transformation.

***D. radiodurans R1 feeding study using [U-$^{13}C_{10}$]-chorismate*:** *Dienococcus radiodurans* synthesize menaquinone by the futalosine dependent pathway.[123] Hence, in an *in vivo* experiment menaquinone biosynthetic precursor **39** ([U-$^{13}C_{10}$]-chorismate isolated from the *E.coli* KA12 mutant, a chorismate overproducing strain, Figure B1 and B2 Appendix B)[150] was fed to *D. radiodurans R1* in order to look for $^{13}C$ label incorporation in menaquinone, **26**. About 400 ml of *D. radiodurans R1* culture was grown with ~85 mg of labeled chorismate. Menaquinone was extracted from the dried cell pellets of *D.radiodurans* using 2:1 chloroform: methanol solvent mixture and analyzed by LC-MS to look for the incorporation of the $^{13}C$ label from chorismate into menaquinone (Figure 4.16).

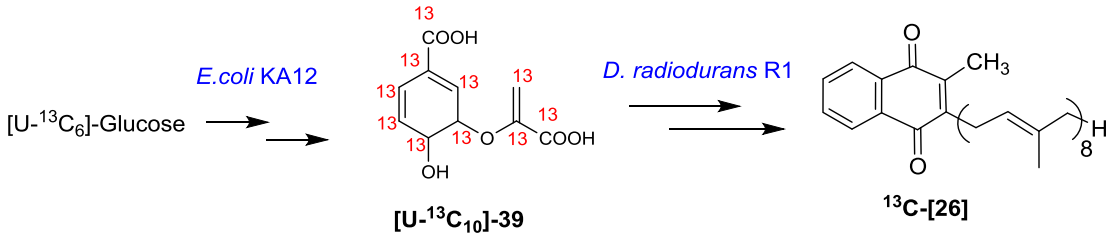



**Figure 4.16**: *D. radiodurans* R1 Feeding study using [U-$^{13}C_{10}$]-chorismate

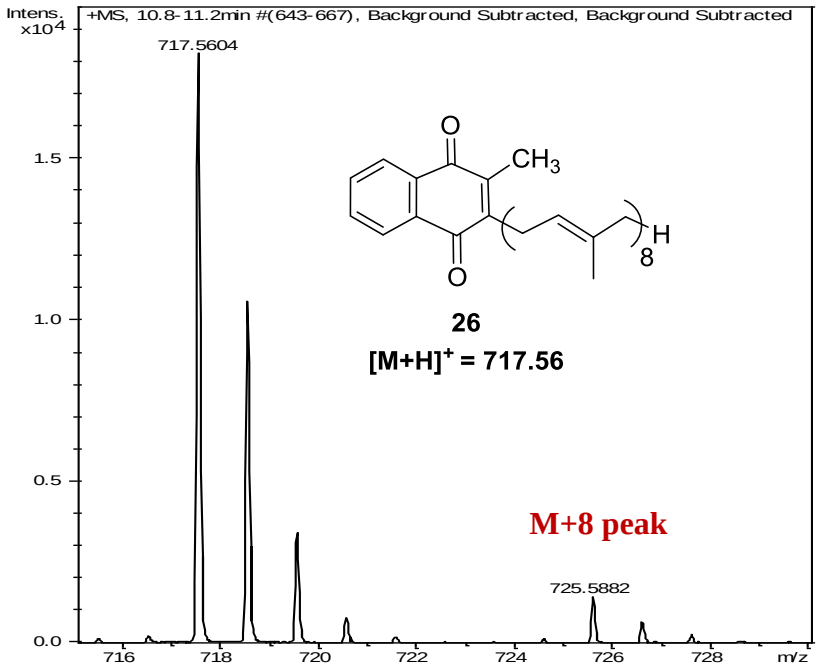



**Figure 4.17:** LC-MS chromatogram of the isolated menaquinone*

*[ESI-MS (positive ion mode) analysis of MK-8 **26**. In the mass spectrum, m/z 717.56, $[M+H]^+$ corresponds to, **26** (theoretical m/z 717.56). A [M+8] MK-8 peak is found with m/z 725.58].

In the LC-MS analysis (Figure 4.17) we have found an M+8 peak with m/z 725.58 along with the parent menaquinone peak (~10% of the parent M peak at m/z 717.56). This experimental result indicated that at least eight carbons from chorismate may have

incorporated into menaquinone (Figure 4.18). This suggests that chorismate may be incorporated into menaquinone as one single unit, probably with the loss of two carboxyl groups during this process. Involvement of inosine or phosphoenolpyruvate is probably not required as previously proposed.[123]

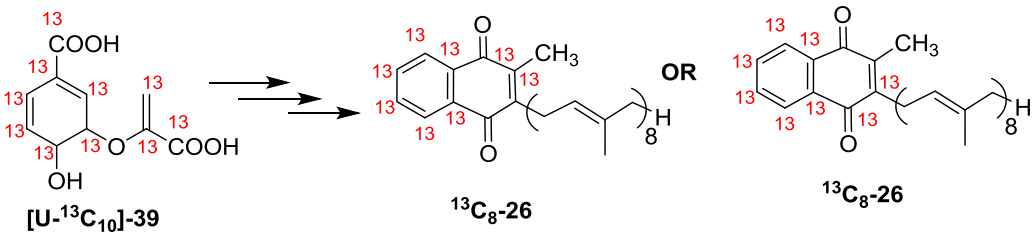



**Figure 4.18**: Probable incorporation of carbons from labeled **39** into MK-8, **26**.

## Conclusion

While non-enzymatic acid catalyzed dehydration of chorismate is well known, [151] surprisingly, enzymatic catalysis of this reaction has not been previously reported. MqnA is, to the best our knowledge, the first example of an enzyme that catalyzes the dehydration of chorismate to form the aromatic compound 3-[(1-carboxyvinyl) oxy]-benzoic acid, **53**.[147] Although, earlier we predicted that this elimination reaction is occurring either through E1 or E1cb mechanism, but absence of acid base residue pair in the substrate binding pocket led us reject this proposal. However, we believe that dehydration of chorismate to the aromatic product probably occurs with the assistance of the neighboring carboxylate group of the pyruvic moiety of chorismate and enzyme

active site perhaps provides the perfect orientation and proximity of the groups required for this dehydration. We put forward a mechanistic proposal based on this hypothesis.

Moreover, it was also found that adenosine/inosine and phosphoenolpyruvate/pyruvate are not substrates for MqnA as previously proposed.[123] Formation of futalosine was not observed when chorimate, inosine and phosphoenolpyruvate were incubated with MqnA. Hence, it is quite possible that additional cofactors and/or enzymes are required for this transformation. At this point, it is difficult to say whether the aromatic product, **53** is an on pathway intermediate or an off pathway shunt product in the menaquinone biosynthesis pathway. However, the *in-vivo* feeding study suggests that chorismate, **39** is probably directly involved in forming the naphthoquinone core structure of menaquinone **26**, which might indicate that **53** is actually an on pathway intermediate. Further studies are required to confirm our results and the mechanistic hypothesis.

**Experimental procedures**

***General methods and materials*:** Molecular biology manipulations were carried out using standard techniques.[64, 152] PCR amplifications were performed using an automatic thermocycler (Mastercycler, Eppendorf) and oligonucleotide primers were synthesized by Integrated DNA Technologies. *Phusion* DNA polymerase was obtained from New England BioLabs. The cloning vectors p28nT and pTHT are derivatives of pET-28, which was purchased from Novagen. Plasmid mini-prep and PCR purification kits were acquired from Qiagen or Fermentas. DNA fragments were purified from agarose gels

with the Zymoclean Gel DNA Recovery Kit from Zymo Research. All restriction enzymes and T4 DNA ligase were purchased from New England Biolabs. DNA sequencing was performed using the Applied Biosystems Automated 3730 DNA Analyzer at the Cornell University Life Sciences Core Laboratories Center (Cornell BioResource Center). *Escherichia coli* strains DH5α or MachI (Invitrogen) were used as recipients for transformations during plasmid construction as well as for plasmid propagation and storage. Growth media were obtained from Difco. Reagents were purchased from Sigma-Aldrich unless stated otherwise.

***Expression constructs for $His_6$-MqnA:*** Chromosomal DNA from *Deinococcus radiodurans* R1 (13939D-5), and *Thermus thermophilus* HB8 (27634D-5) was purchased from American Type Culture Collection (ATCC) and used as the template for PCR amplification of *mqnA*. The genes were amplified by PCR using the primer sequences listed below. The product containing the desired gene was excised from an agarose gel, purified and digested with the appropriate restriction enzymes, repurified, and then ligated into similarly digested and purified p28nT or pTHT. The final constructs encode the gene product with six histidine residues and a TEV protease recognition sequence (underlined font) appended to the N-terminus: p28nT, MGSSHHHHHHSSGENLYFQGH; pTHT, MGSDKIHHHHHHSSGENLYFQGH. Competent *E. coli* DH5α cells were transformed with the sample. Cells were plated on LB agar (1.5%) supplemented with 50 μg/mL kanamycin (LabScientific) to select for

positive clones. Subsequently, clones were screened by colony PCR and the desired gene products were confirmed by DNA sequencing.

*p28nT-ttmqnA.* The forward and reverse primers for *ttmqnA* gene amplification by PCR were designed with *Nde*I and *Bam*HI restriction sites (indicated in underlined font), respectively, as TtMqnA FP: 5′- gggtagcatATGAGGCCCTACGTCCTGGGCCTTCCC-3′ and TtMqnA RP: 5′-cccTAGGATCCTACTCAGGATAGAACATAGGGGCG-3′.

*pTHT-DrmqnA.* The forward and reverse primers for *drmqnA* gene amplification by PCR were designed with *Nde*I and *Xho*I restriction sites (indicated in underlined font), respectively, as
DrMqnA FP: 5′ gggtagcatatgACCAACCATCACCCATCATCTACTGGC-3′ and DrMqnA RP: 5′-ccctactcgagttaTGCTTCCTCACGCGCGCTCCGAAGG-3′.

***Expression and purification of $His_6$-MqnA*:** Electrocompetent *Escherichia coli* BL21 (DE3) (Novagen) cells were prepared using standard methods. The cells were freshly transformed with the pTHT or p28nT constructs containing the appropriate N-terminal hexa-histidine *mqnA* fusion. Single colony transformants were grown aerobically for 12-15 h at 37 °C in 20 mL of LB medium supplemented with 40 μg/mL kanamycin (LB-kan). Cells from the starter culture were added to 2 L of LB-kan, and the culture was grown aerobically at 37 °C until the $A_{600}$ was 0.6-0.8. Expression was induced by the addition of 1 mM IPTG and the culture was grown aerobically at 15 °C for an additional

~ 16 h. Cells were harvested by centrifugation at 6,300×*g* in a Beckman Coulter Avanti J-26 XPI with a JLA-8.1000 rotor for 15 min at 4 °C. The medium was discarded and the cell paste (~ 2 g per liter of cell culture) was stored at –80 °C until use. Cells were thawed, resuspended in 20 mL of Lysis Buffer (50 mM potassium phosphate, pH 7.8 at 4 °C, 150 mM NaCl), and then lysed by sonication on ice for 6 × 20 s using a Misonix Sonicator 3000 (24 W, 1.5 s pulse, 1.5 s pause). Cell debris was removed by centrifugation at 31,000×*g* in a Beckman Coulter Avanti J-E with a JA-17 rotor for 45 min at 4 °C and the supernatant was clarified through a 0.45 μm syringe filter. The enzyme was purified by nickel affinity chromatography at 4 °C using a stepwise gradient of increasing imidazole concentration with a 5 mL HiTrap chelating HP column (GE Healthcare Life Sciences). Wash Buffers were prepared by combining the Lysis Buffer and Elution Buffer (50 mM potassium phosphate, pH 7.8 at 4 °C, 150 mM NaCl, 500 mM imidazole) to give the final imidazole concentrations desired. The resin was charged with 2 CV of 0.1 M $NiSO_4$, washed with 5 CV of filtered water to remove excess nickel, and equilibrated with 2 CV of Lysis Buffer. After loading the filtered protein sample, the resin was washed with 2 CV each of Wash Buffers containing 20 mM, 50 mM, 200 mM, and 500 mM imidazole. Fractions of 10 mL were collected and analyzed for purity by 12% SDS-PAGE. $His_6$-MqnA eluted from the column in the 200 mM imidazole Wash Buffer fractions. Nickel affinity column fractions containing the desired protein were pooled and concentrated from ~ 20 mL to ~ 3 mL using a centrifugal filtering device with a 10 kDa MWCO membrane (Millipore) and the buffer was exchanged to Final Buffer (50 mM potassium phosphate, pH 7.5 at 4 °C, 150 mM KCl, 30% glycerol) via

Econo-Pac 10DG size exclusion chromatography. SDS-PAGE analysis showed the molecular weight of MqnA as ~ 30.5 kDa (Figure S1). Protein concentrations were determined using the Bradford assay and $A_{280}$ measurements. Typical yields were ~ 10 mg per liter of cell culture. Proteins were stored at –80 °C.

***Characterization of 3- ((1-carboxyvinyl) oxy) benzoic acid, 53****:* Chorismate (3 mM) was incubated with MqnA (400 µM) in 100 mM Tris-HCl buffer (pH 7.5) and the reaction was allowed to proceed for 3 hrs. The protein was removed by ultrafiltration using 10 kDa centrifugal filters. The reaction product was purified by HPLC (eluted at 5.1 min) and characterized by NMR and LC-MS analysis.

***HPLC method****:* HPLC analysis using a linear gradient, at a flow rate of 1 mL/min was used with absorbance detection at 254 nm. Solvent A is water, solvent B is 100 mM $K_2HPO_4$, pH 6.6 and solvent C is methanol: 0 min, 100% B; 5 min, 100 % B; 12 min, 48 % A, 40 % B, 12 % C; 14 min, 50 % A, 30 % B, 20 % C; 18 min, 30 % A, 10% B, 60% C; 20 min, 100 % B; 25 min 100 % B. The column used was a Supelcosil LC-18- T HPLC column (15 cm x 4.6 mm, 3 µM particle size).

The purified compound was concentrated using a speedvac concentrator (Thermo Scientific) and dissolved in $CD_3OD$ for NMR analysis. The spectroscopic data are as follows (Figures 4.4, 4.5): $^1H$ NMR (500 MHz, $CD_3OD$): 7.78 (dt, J = 7.8, 1.35 Hz, 1H), 7.60-7.625 (m, 1H), 7.46 (t, J = 7.9 Hz, 1H), 7.25 (dddd, J = 8, 2.4, 1.1 Hz, 1H),

5.83 (d, J = 2.01 Hz, 1H)**,** 5.09 (d, J = 2.01 Hz, 1H), $^{13}C$ NMR (125 MHz, $CD_3OD$): 168.9, 165.3, 157.6, 151.5, 133.8, 131.0, 126.1, 124.0, 120.1, 107.3. DEPT-135 $^{13}C$ NMR (Figure 8, Appendix A): 133.8 (CH), 131.0 (CH), 126.1 (CH), 124.0 (CH), 120.1 (CH), 107.3 ($CH_2$). The 2D NMR spectra (COSY and HSQC) of **53** are shown in Figures 4.6 and 4.7 respectively**.** LC-MS analysis of the compound, **53** (ESI, negative ion mode) yielded m/z of 207.03, $[M-H]^-$ (calculated m/z 207.03 m/z, Figure 4.8).

***LC-MS method:*** Samples were analyzed by reverse-phase HPLC on an Agilent 1200 HPLC system equipped with a diode array UV-Vis detector and a thermostatted autosampler (10 °C). The stationary phase was Supelcosil LC-18-T (15 cm × 3 mm, 3 µm particles), maintained at 22 °C. The LC eluent consisted of a gradient of 75% methanol in 25% water with 5 mM ammonium acetate buffer, pH 6.6 (0.4 mL/min flow rate). The percentages of ammonium acetate buffer (*B*) and methanol (*M*) at time *t* varied according to the following scheme: {*t*,*M*,*B*}: {0,0,100}, {2,0,100}, {10,35,65}, {15,57.5,42.5}, {18,100,0}, {23,100,0}; {26,0,100}; {33,0,100}. Compounds were detected using absorbance at 254 nm. UV-Vis spectra (190 – 640 nm) were also collected and evaluated. Mass data were collected using an in-line Bruker Daltonics micrOTOF-Q II ESI-Qq-TOF mass spectrometer (HyStar) in positive ion mode.

***Crystallization, data collection and processing***: Initial crystallization conditions were identified using the hanging drop vapor diffusion method (Crystal Screens 1 and 2, Hampton Research; Wizard Screens 1, 2, 3 and 4, Emerald Biosystems) at 18 °C.

Hanging drops were formed by mixing 1.5 μL of reservoir solution with 1.5 μL of protein sample. Optimized crystals required a microseeding step using cat whiskers and a reservoir solution contained 0.1 M acetate (pH 5.5), 13% PEG 1500, 2.5 M sodium chloride (NaCl), and 1.5% 2-methyl-2,4-pentanediol (MPD). Large rectangular shaped crystals grew to 500-600 μm × 50-100 μm in two to four days. To obtain the liganded structure, the crystals were removed from the hanging drop and placed in a sitting drop tray. A sequential addition of 3-hydroxybenzoic acid (3-HB) was added to the crystals from 2-100 mM in 10 mM increments. At each concentration, the crystals were transferred to a new drop and allowed to incubate for ~10 minutes. The crystallization buffer supplemented with 100 mM 3-HB and 26% ethylene glycol was used as the cryoprotectant. Samples were flash frozen by plunging into liquid nitrogen. Data sets were collected at the Cornell High Energy Synchrotron Source (CHESS) at beamline A1. The data was collected on a single crystal over 180° using a 1° oscillation range at a wavelength of 0.9180 Å. The crystals diffracted to 2.33Å resolution. The space group was *C*222 with average unit cell parameters of $a$ = 143 Å, $b$ = 152 Å , and $c$ = 75 Å. The asymmetric unit consisted of two chains corresponding to a Matthews's coefficient of 2.26 $Å^3$/Da solvent content of 46 %.[153] Data were indexed, integrated, and scaled using *HKL*-200[66] (by Katherine Hicks, Ealick lab, Cornell University).

***Structure determination, model building and refinement***: The structure was determined by molecular replacement using MOLREP[67] with protein DR0370 from *Deinococcus radiodurans,* Pfam DUF178 as the search model[154] (PDB entry 2I6E). Initial refinement

was carried out using Refmac 5.0 and the CCP4i interface[153, 155]. Iterative rounds of manual model building were performed with COOT[156] and later rounds of refinement were carried out using PHENIX[156, 157] (by Katherine Hicks, Cornell University)

***Isolation and purification of chorismate***: Chorismate was purified as reported before.[150] The growth media contained following components (per 1L) : Yeast extract (2g), casein hydrolysate (2g), anhydrous $MgSO_4$ (0.1g),citric acid (1.83g),$K_2HPO_4$ (10g), $Na(H)(NH_4)PO_4.4H_2O$ (3.5g), Tryptophan (41mg), glucose (1.6 g), tetracycline (12 µg/mL). The accumulation media contained following components (per 1L): Glucose (18g), $Na_2HPO_4$ (12.8g), $KH_2PO_4$ (1.36g), $NH_4Cl$ (2.74 g), $MgCl_2.6H_2O$ (20.3 mg), tryptophan (3mg). [U-$^{13}C_6$]- Glucose in both growth media and accumulation media.

***Growth procedure***: 1L sterile growth medium was inoculated with 5 mL of O/N culture of *E.coli* KA-12 strain and grown at 30°C with vigorous shaking (250-300 rpm) for ~20hrs. Then it was centrifuged to remove the growth media. The cells were re-suspended in 1L accumulation media and incubated with shaking (~250 rpm) at 30°C for 16-20 hrs. The accumulation of chorismate was checked by taking the O.D. of a 100 µL aliquot of the accumulation media. The aliquot was centrifuged and the supernatant was diluted to 1:20 in water and checked OD (275 nm) (should be ~0.3-0.7). When the O.D. reached about this value, the accumulation media was centrifuged and apply it for further purification.

***Purification procedure***: Dowex-Cl 1x8 beads (50-100 mesh, 40g) were washed with 1M HCl (~150mL), followed by thorough rinsing with water until the pH is neutral. This is followed by washing the beads with 1M NaOH (~150 mL), followed by rinsing with water until the pH is <8. Then, the accumulation media was made basic with 10M NaOH (to pH 8.5) and applied it to the column packed with the beads (pre-running with water). The column was washed column with water and chorismate is eluted with 2M $NH_4Cl$ (pH 8.5) at slower flow rate (~2.5 ml/min). 15 ml fractions (Chorismic acid elutes early with a visible dark brown color) were collected and fractions with OD 275 >0.25 (in a 1:100 dilution) were pooled for further purification. The collected fractions (130ml) were acidified to pH 1.5 with HCl and extracted 5X with 100 ml of cold diethyl ether. The ether extracts were dried with anhydrous $MgSO_4$, filtered and concentrated to <10 ml using Rota vapor. The ether concentrate was cooled to 0°C and petroleum ether was slowly added drop wise and swirled the flask until a light yellow precipitate forms and remains in the solvent (At this point adding chorismic acid helps in faster crystallization). Allowed the chorismic acid to crystallize (~3 hrs) and then added more petroleum ether to ensure complete crystallization (left at -20°C overnight). The solvent was poured off onto vacuum filter without pouring out crystals. The crystals were dried by lyophilization for ~6 hrs and stored as powder form at -80°C (~330mg of chorismate from 1L culture was obtained).

# CHAPTER V

# MENAQUINONE BIOSYNTHESIS: RECONSTITUTION AND CHARACTERIZATION OF A UNIQUE RADICAL SAM ENZYME, AMINOFUTALOSINE SYNTHASE (MqnE)*#

## Introduction

It was proposed in the futalosine dependent menaquinone biosynthesis pathway that MqnA catalyzes the formation of futalosine from chorismate, inosine and phosphoenolpyruvate (Figure 3.8).[123] Since we were unable to reconstitute the formation of futalosine or aminofutalosine using MqnA alone,[147] we considered the possible involvement of additional enzymes and/or cofactors for this transformation. Three possible pathways for the formation of aminofutalosine were considered (Figure 5.1). Paths A and B require a thiamin pyrophosphate dependent ketoacid decarboxylase, while path C requires a radical SAM enzyme. To determine which path was used, we analyzed the gene neighborhood of the identified menaquinone biosynthetic genes and our analysis yielded one promising gene (now named as *mqnE*), encoding a putative radical SAM enzyme that frequently clustered with menaquinone biosynthetic genes (SCO4494 in *S. coelicolor*). The studies described herein are aimed at characterizing this putative radical SAM enzyme and its relevance in the menaquinone biosynthesis pathway.

*Part of this chapter is reprinted with permission from "Menaquinone Biosynthesis: Formation of Aminofutalosine Requires a Unique Radical SAM Enzyme" by Mahanta, N.; Fedoseyenko, D.; Dairi, T.; Begley, T. P., 2013. *J. Am. Chem. Soc.*, 135, 15318-15321, Copyright [2013] by American Chemical Society.

# (Nilkamal Mahanta performed the biochemical studies. Dmytro Fedoseyenko synthesized compounds **53** and **50**).

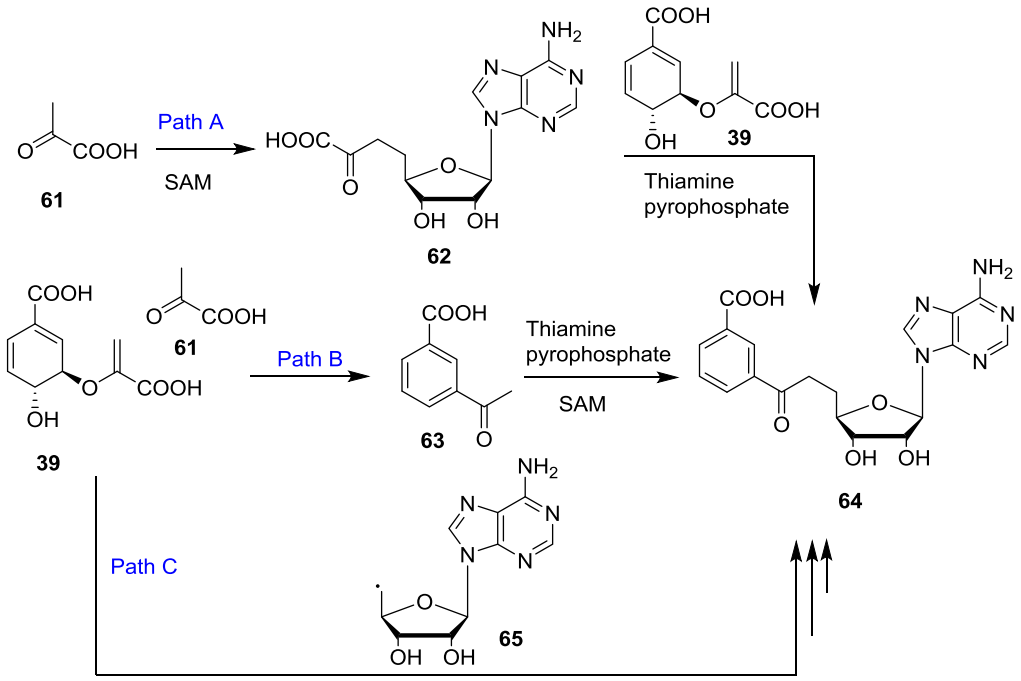



**Figure 5.1:** Three possible pathways for the formation of aminofutalosine, **64**.[147]

## Results and Discussion

***Bioinformatics analysis***: The gene neighbourhood of the identified menaquinone biosynthetic genes was analyzed for additional genes encoding a putative thiamin pyrophosphate dependent decarboxylase or a putative radical SAM enzyme using the SEED database (http://theseed.uchicago.edu/FIG/). Figure 5.2 summarizes some relevant portion of the gene cluster in the neighborhood of MqnA (MqnA is shown in black, MqnB in purple, MqnC in green and MqnD in Cyan) and other menaquinone biosynthetic genes in some of the organisms synthesizing menaquinone by the futalosine dependent pathway.[123]

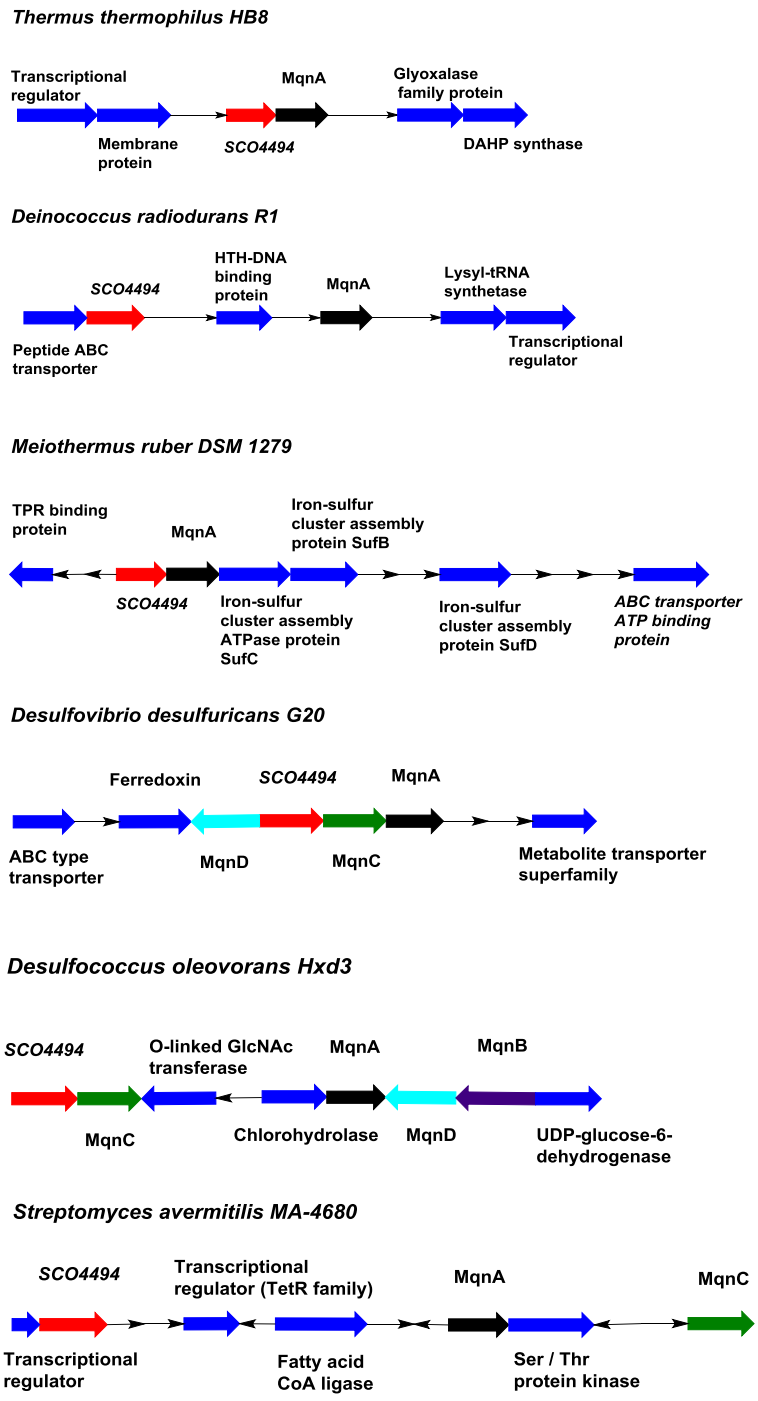



**Figure 5.2:** Biosynthetic gene cluster of some menaquinone producing organisms.

This bioinformatics analysis (Figure 5.2) shows the presence of a gene named SCO4494 (shown in red, nomenclature is based on *S.coelicolor* genome) clustered along with MqnA in some organisms (such as *T. thermophilus*) and along with the other genes of menaquinone biosynthesis in some other organisms (such as *S. avermitilis*). NCBI blastp analysis using the amino acid sequence corresponding to SCO4494 indicated homology with some known radical SAM enzymes (such as biotin synthase, lipoate synthase) and contains a characteristic $CX_3CX_2C$ motif required to bind the [4Fe-4S] cluster.[147] An amino acid sequence from a *T. thermophilus* orthologue, TTHA0804 shows this binding motif (Figure B, Appendix B). With this finding, we wanted to experimentally validate whether a putative radical SAM enzyme encoded by the gene SCO4494 could be a possible candidate involved in the formation of aminofutalosine (or futlaosine).

***Enzyme characterization (TTHA0804 from T. thermophilus HB8):*** To test the hypothesis that aminofutalosine, **64** is formed via path C (Figure 5.1) using the newly discovered gene SCO4494, we decided to use the recombinant enzyme. An orthologue from *T. thermophilus* HB8 (Figure B3, Appendix B), TTHA0804 was PCR amplified, cloned into the pQE-30 expression vector, and overexpressed in *E. coli* M15 (pREP4). The resulting $His_6$-tagged MqnE protein was purified by Ni-NTA affinity chromatography under anaerobic conditions (Figure 5.3 shows the SDS-PAGE analysis in which A is the molecular weight marker and B is the purified MqnE. The molecular weight is ~42.6 Da). The purified protein had an absorbance maximum at 415 nm, which disappeared upon treatment with excess dithionite or on exposure to oxygen, consistent

with the presence of the predicted [4Fe-4S] cluster (Figure 5.4). Iron and sulfide analysis yielded 2.4 irons and 2.8 sulfides per monomer of the enzyme, indicating incomplete reconstitution of the cluster. The cluster incorporation was improved by co-expression of an *E.coli* Fe-S cluster biosynthetic protein *suf* [158] along with TTHA0804. Iron and sulfide analysis of the purified protein yielded 3.7 irons and 4.2 sulfides per monomer of protein. We termed it as MqnE.

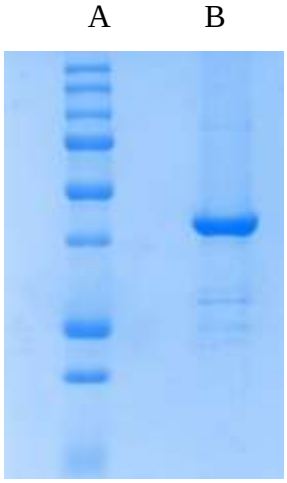



**Figure 5.3**: SDS-PAGE analysis of MqnE (TTHA0804).

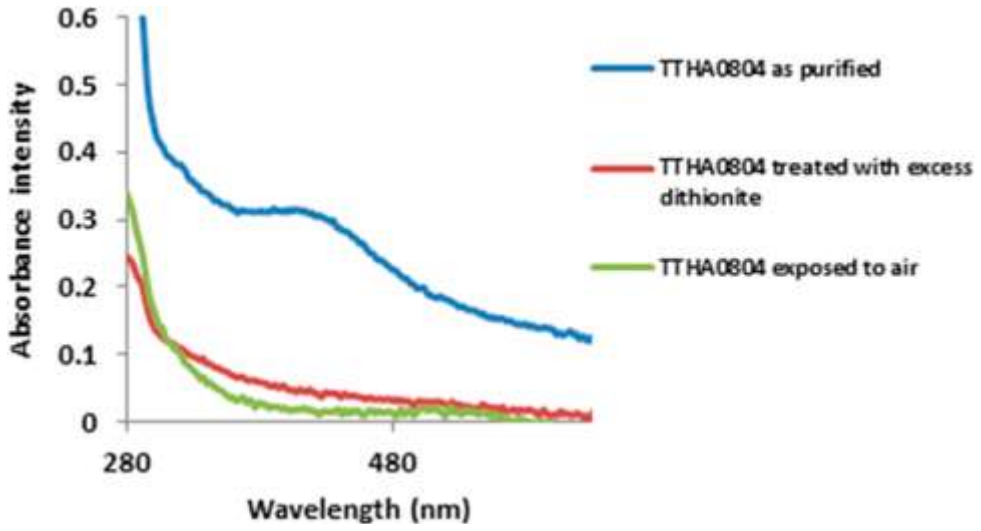



**Figure 5.4**: UV−visible spectra of MqnE (TTHA0804).[147]
[Blue trace: protein "as isolated" enzyme, Red trace: dithionite (excess) treatment of the "as-isolated" sample, Green trace: aerobic oxidation of the "as-isolated" sample.Absorbance at 415 nm was observed due to the presence of the bound Fe-S cluster]

MqnE (as isolated) was analyzed by EPR spectroscopy and it was found that the EPR spectral properties of the bound Fe-S cluster is consistent with a $[3Fe\text{-}4S]^{+}$ cluster with g value of 2.01 (Figure 5.5).[159] On *in-vitro* reconstitution using iron and sulfide and after reduction with sodium dithionite, it is converted to a catalytically active $[4Fe\text{-}4S]^{+}$ cluster. Such an observation has been made some other radical SAM enzymes (such as PqqE involved in Pyrroloquinoline quinone biogenesis) [160] and other Fe-S cluster containing enzymes (such as Aconitase).[161]

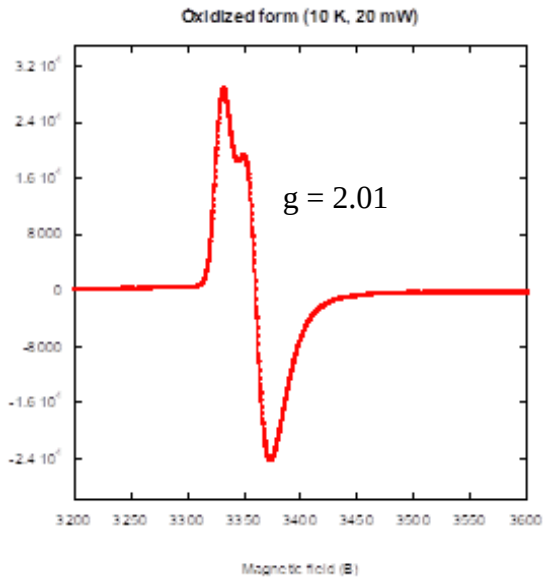



**Figure 5.5**: EPR spectrum of the Fe-S cluster bound to MqnE ( as-isolated form).

***HPLC analysis of the MqnE reaction mixture*:** To investigate the involvement of MqnE (TTHA0804) in the formation of aminofutalosine, 3-[(1-carboxyvinyl) oxy]- benzoic acid, **53** was prepared by incubating chorismate with MqnA (Figure 4.3) or by chemical synthesis (Scheme 4.1). Compound **53** (2 mM) was then incubated anaerobically with MqnE (200 μM) in the presence of S-adenosylmethionine (2 mM) and excess sodium dithionite. After 10 h, the reaction was quenched with 8 M guanidine-HCl, protein was removed by ultrafiltration, and the reaction mixture was analyzed by reverse-phase HPLC. This analysis showed the formation of a new compound in the HPLC chromatogram which elutes at 20.2 min (Figure 5.6). It was purified from reaction mixture using HPLC and characterized by 1D and 2D NMR (Figures 5.7-5.11) and MS analysis (Figure 5.12). This analysis suggested that the new compound forming in the MqnE reaction is 6-aminodeoxyfutalosine (or aminofutalosine), **64**.[147]

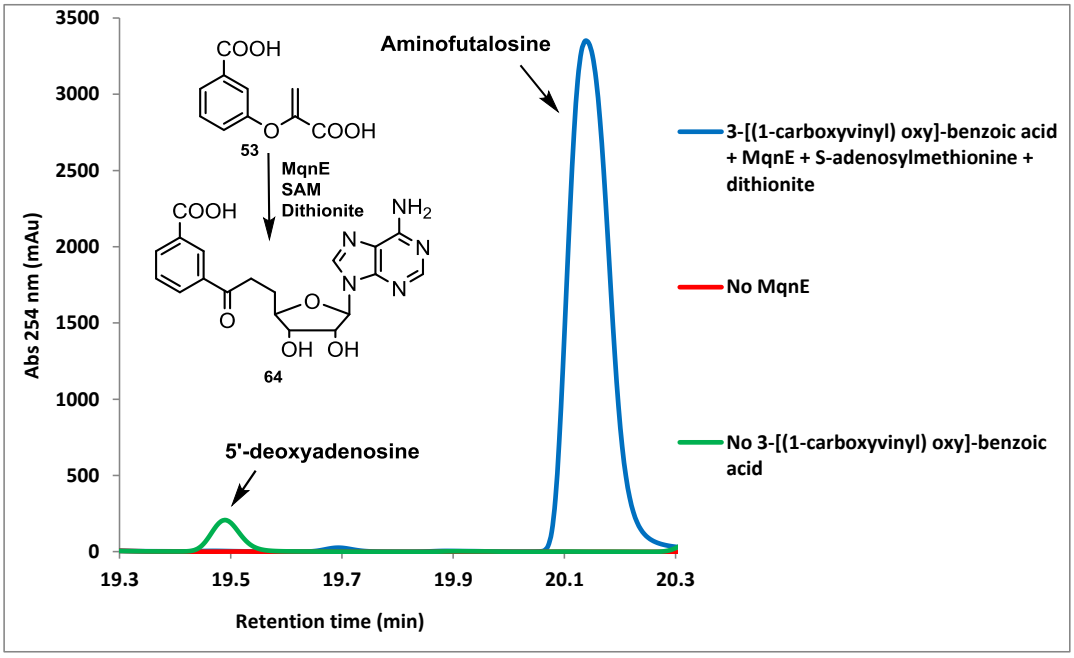



**Figure 5.6**: HPLC analysis of the MqnE reaction mixture.[147]
[Blue trace: chromatogram of the full reaction mixture (MqnE + **53** + SAM + dithionite); Red trace: chromatogram of the control reaction mixture lacking MqnE; Green trace: chromatogram of the control reaction mixture lacking substrate **53**, showing low levels of the uncoupled production of 5′-deoxyadenosine. No signals at 20.2 and 19.5 min were obtained in the control reactions individually lacking SAM or dithionite (data not shown). The inset shows the structure of the identified MqnE reaction product, **64**, which was also, characterized by 1D and 2D NMR and MS spectrometry]

***NMR and MS characterization of aminofutalosine, 64***:

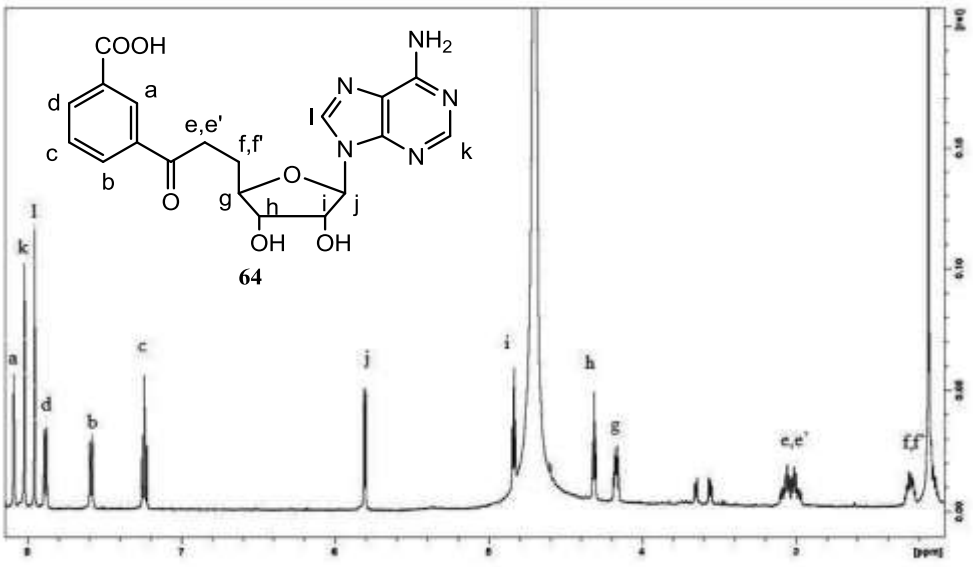



**Figure 5.7:** $^{1}$H NMR spectrum (500 MHz, $D_2O$) of aminfutalosine, **64.**[147]

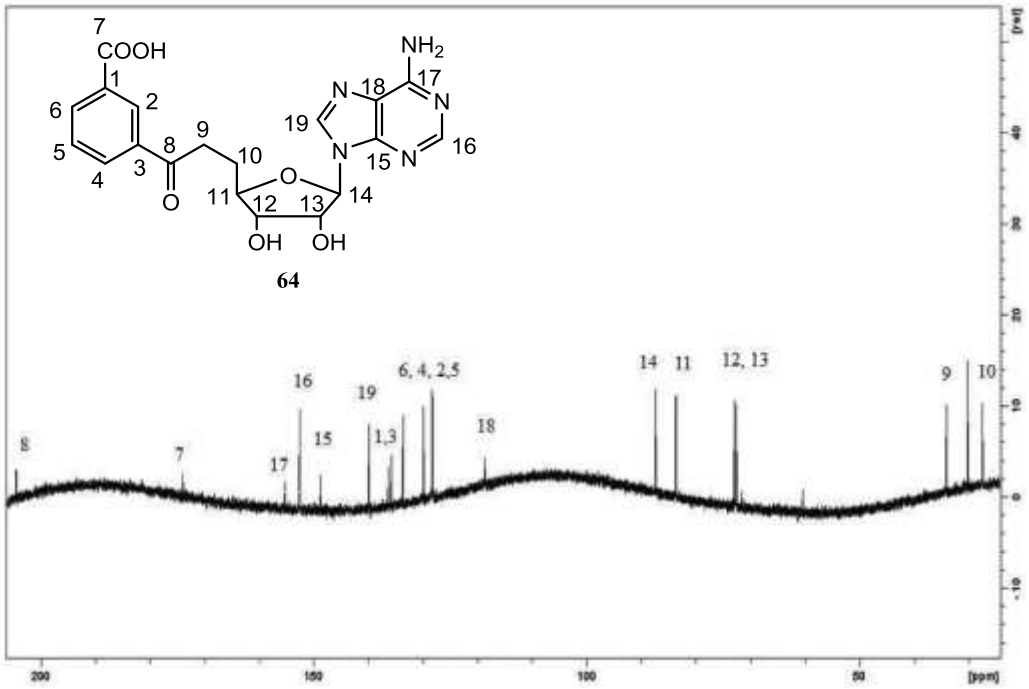



**Figure 5.8**: $^{13}$C NMR (125 MHz, $D_2O$) of aminofutalosine, **64.**[147]

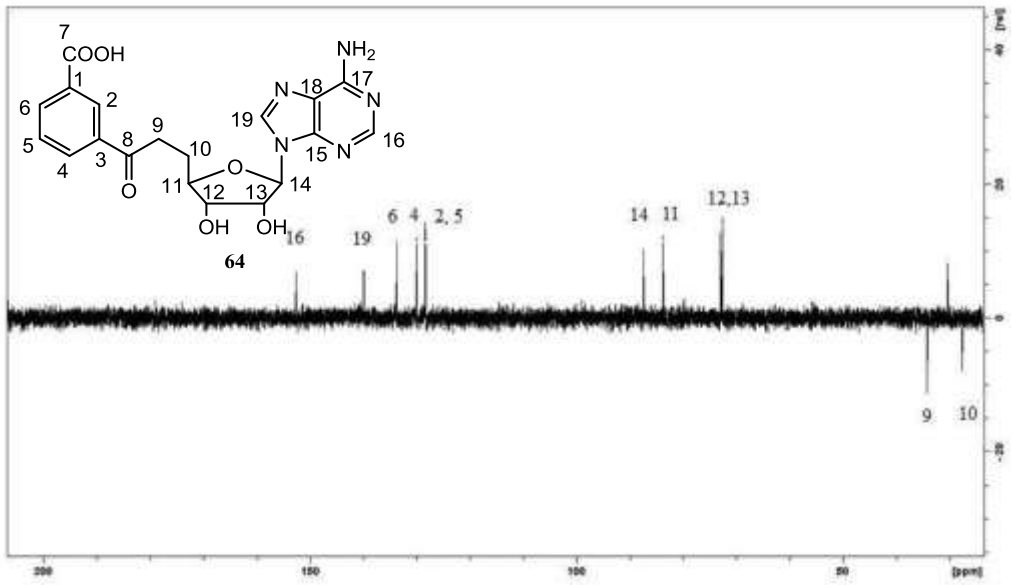



**Figure 5.9:** $^{13}C$ DEPT-135 NMR (125 MHz, $D_2O$) spectrum of aminofutalosine, **64**.[147]

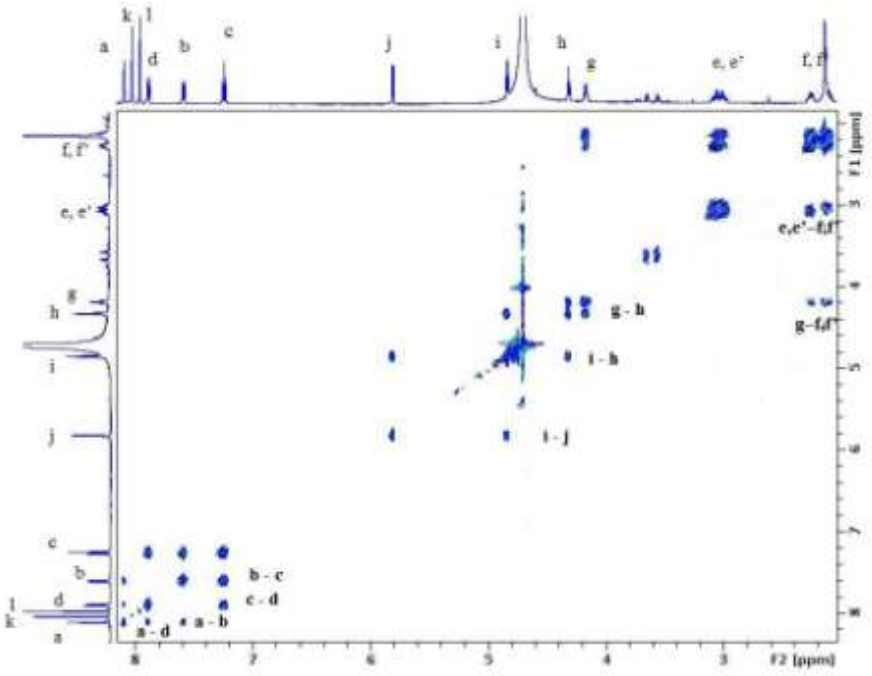



**Figure 5.10**: $^{1}H$ COSY spectrum of aminofutalosine,**64.**[147]

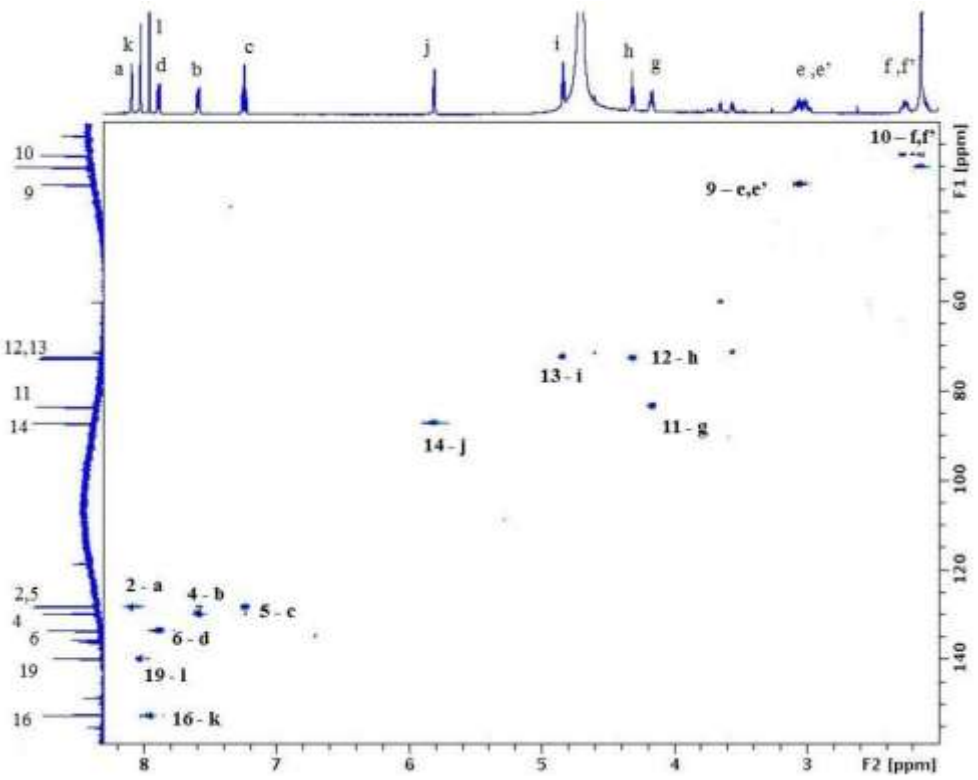



**Figure 5.11**: HSQC spectrum of aminofutalosine, **64**.[147]

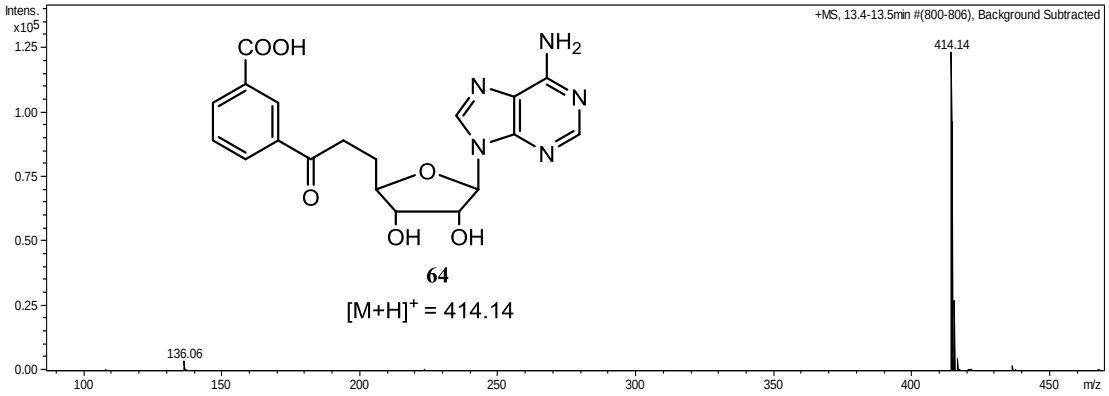



**Figure 5.12**: LC-MS (ESI, positive ion mode) analysis of aminofutalosine, **64**.[147]
*[In the mass spectrum, m/z 414.14, $[M+H]^+$ corresponds to aminofutalosine **64** (theoretical m/z 414.14) and m/z 136.06 corresponds to adenine derived from aminofutalosine (theoretical m/z 136.06)]

***Deamination of aminofutalosine 64 as a confirmation of its structure***: The 6-aminodeoxyfutalosine deaminase[162, 163] SCO5662 (100 µM) was incubated with biosynthesized aminofutalosine **64** in the presence of $Mn^{2+}$ (250 µM) for 3 hrs in 100 mM Tris-HCl buffer (pH 7.5). The reaction mixture (100 µL) was then treated with an equal volume of 8M guanidine-HCl and the protein was removed by ultrafiltration using 10 kDa centrifugal filters (Pall Life Sciences). The reaction mixture was analyzed by reverse phase HPLC. The reaction product (eluted at 19.6 min) was confirmed as futalosine **49**, by co-migration with an authentic sample isolated from an *S. coelicolor* mutant [123] (obtained from Tohru Dairi lab, Figure 5.13).[147] LC-MS (ESI, positive ion mode) analysis of the product showed m/z = 415.13 $[M+H]^+$ which corresponds to **49** (theoretical m/z = 415.13, Figure 5.14).[147]

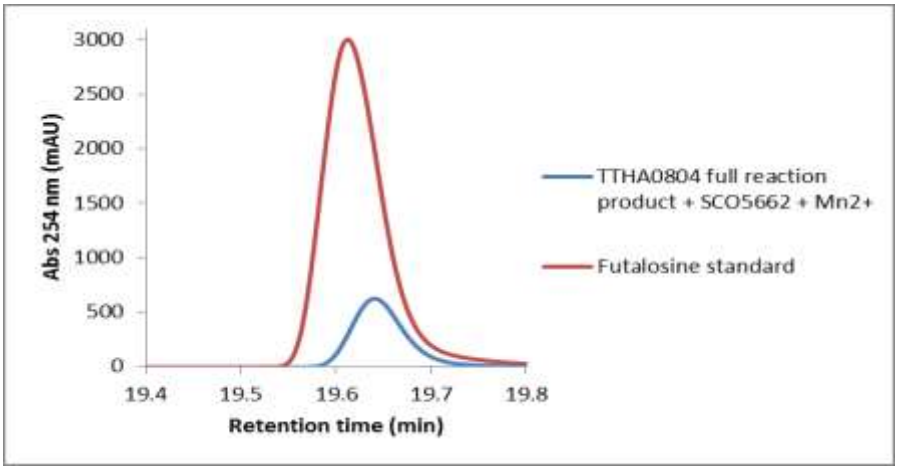



**Figure 5.13**: HPLC analysis of the aminofutalosine deaminase reaction mixture.[147]
[Blue trace: chromatogram of the full reaction mixture (TTHA0804 reaction products + SCO5662 + $Mn^{2+}$); red trace: chromatogram of the standard futalosine.[147]The substrate, aminofutalosine **64**, elutes at 20.2 min on the chromatogram].

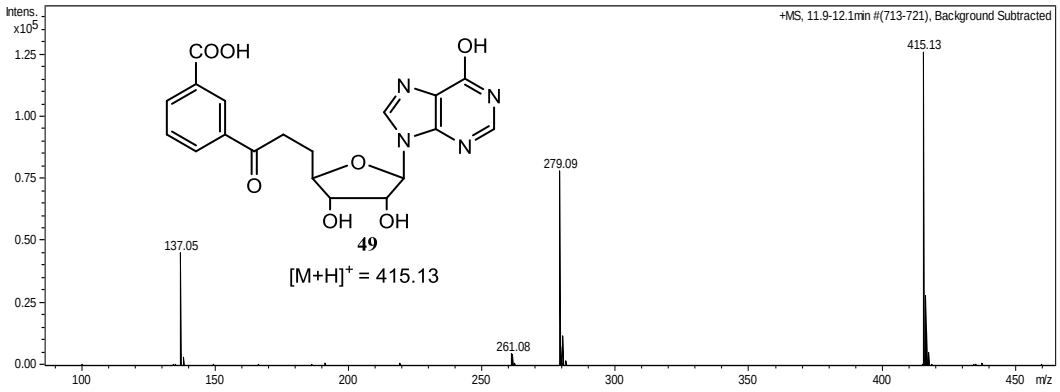



**Figure 5.14**: LC-MS analysis of the aminofutalosine deaminase reaction.[147] [In the mass spectrum, (m/z 415.13, $[M+H]^+$ corresponds to futalosine **49**, theoretical m/z = 415.13 and m/z 137.05 corresponds to hypoxanthine derived from futalosine 28, theoretical m/z 137.05)]

***Conversion of aminofutalosine, 64 to 50 as a confirmation of its structure***: The hydrolase HP0089[162] (100 µM) was incubated with biosynthesized aminofutalosine, **64** for 3 hrs. The reaction mixture (150 µL) was treated with an equal volume of 8M guanidine-HCl, protein was removed by ultrafiltration and analyzed by reverse phase HPLC. The reaction product was confirmed as compound **50**, by co-migration with an authentic synthesized sample of **50** [144] (eluted at 16.8 min, Figure 5.15) and by LC-MS analysis (Figure 5.16).[147]

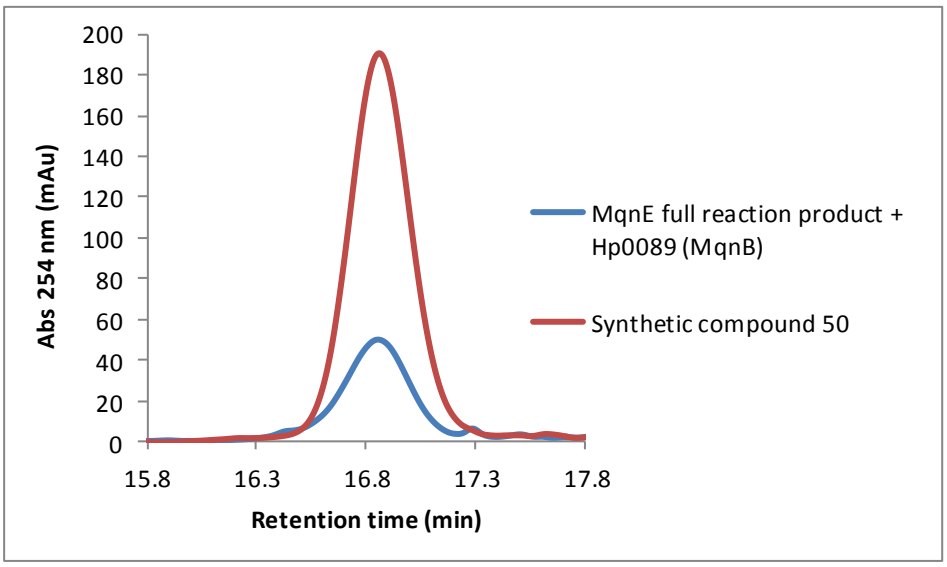



**Figure 5.15**: HPLC analysis of the aminofutalosine nucleosidase reaction mixture.[147] [Blue trace: chromatogram of the reaction mixture; red trace: chromatogram of the standard **50**.[144]The substrate, aminofutalosine, **64** elutes at 20.2 min on the chromatogram]

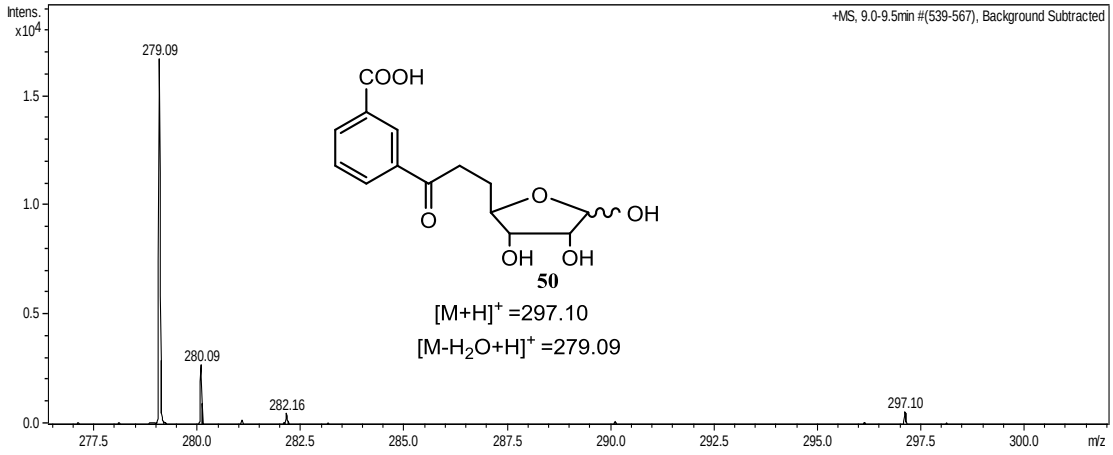



**Figure 5.16**: LC-MS analysis of the aminofutalosine nucleosidase reaction.[147] [In the mass spectrum, m/z 297.10 corresponds to **50**, theoretical m/z 297.10, $[M+H]^+$ and m/z 279.09 corresponds dehydrated **50** ( $[M-H_2O+H]^+$)]

***Catalytic capacity of MqnE:*** The MqnE reaction mixtures were also analyzed to determine if MqnE was catalytic. 3-[(1-carboxyvinyl) oxy]-benzoic acid, **53** (2 mM) was incubated with SAM (2.5 mM) and MqnE (120 µM) in the presence of excess sodium dithionite for 12 hrs. The reaction was quenched with 8M guanidine hydrochloride and MqnE was removed by ultrafiltration. The amount of aminofutalosine formed was quantified by HPLC analysis (**64** eluted at 20.2 min, Figure 5.6) to yield the product concentration of 1.7mM demonstrating that the reaction is catalytic (based on a standard curve, Figure B4, Appendix B). In other words, with the current enzyme preparations, 1 equivalent of enzyme gives 14 equivalents of aminofutalosine **64**.

***Fate of the carbon lost during the MqnE reaction****:* The MqnE reaction mixtures were also analyzed to determine the fate of the carboxylate of compound **53**. NMR analysis of the MqnE reaction mixture, run using [$^{13}$C]-**53**, demonstrated the formation of bicarbonate and not formate, suggesting that the carboxyl group of **53** is lost in a decarboxylation reaction (Figure 5.17).[147]

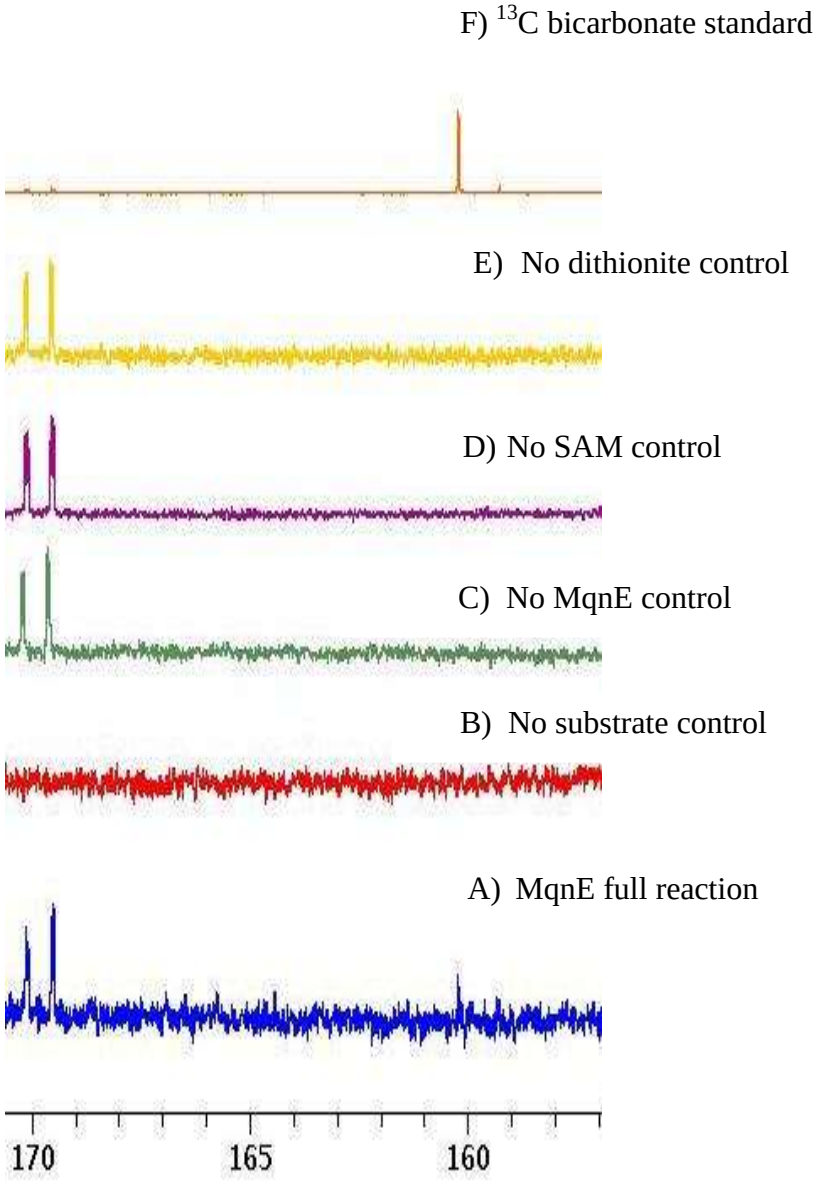



**Figure 5.17**: $^{13}C$ NMR (125 MHz, $D_2O$) analysis of the MqnE catalyzed reaction.[147] [(A) Blue trace: spectrum of the full reaction mixture (MqnE + $^{13}C$ labeled **53** + SAM, **66** + dithionite); (B) red trace: spectrum of the control reaction mixture lacking substrate, **53**; (C) green trace: spectrum of the control reaction mixture lacking MqnE ; (D) violet trace: spectrum of the control reaction mixture lacking SAM; (E) yellow trace: spectrum of the control reaction mixture lacking dithionite ; (F) brown trace: spectrum of standard $^{13}C$ sodium bicarbonate. The signal observed in spectrum A at 160.01 ppm matched with the chemical shift of the bicarbonate signal of spectrum F].

***Mechanistic proposal for the radical SAM enzyme MqnE*:** Radical SAM enzymes are among the most interesting family of enzymes that catalyze several mechanistically intriguing reactions found in biological systems.[164-168] In these enzymes, 5''deoxyadenosyl radical abstracts a hydrogen atom directly from the substrate or via a protein based radical and generates 5'-deoxyadenosine.[164-168] The resulting substrate radical then further undergoes remarkable rearrangement and/or fragmentation reactions to yield the product. Radical SAM enzymes are particularly prevalent in biosynthetic pathways of cofactors and catalyze various complex rearrangement reactions (such as ThiC in thiamin biosynthesis, MoaA in molybdopterin biosynthesis and CofG/CofH or FbiC in coenzyme $F_{420}$ bisynthesis).[169]

However, the radical SAM enzyme MqnE discovered in this pathway represents a unique example of this superfamily. On the basis of our experimental results so far, a mechanistic proposal for the aminofutalosine synthase (MqnE)-catalyzed reaction is outlined (Figure 5.18).[147] First, reductive cleavage of S-adenosylmethionine (SAM) **66** with the help of the reduced [4Fe-4S] cluster generates 5'-deoxyadenosyl radical **67** and methionine **68.** Addition of the radical **67** to the vinyl double bond of compound **53** gives the captodative radical **69**. A non-enzymatic side reaction showing a similar radical addition has been reported in a pyruvate formate lyase-derived peptide in which the active-site glycine was replaced with dehydroalanine.[170] Rearrangement of this radical via **70** gives **71**. Such rearrangements have previously been proposed but not mechanistically well characterized.[171-173] Decarboxylation of **71**, facilitated by the

alkoxy radical, gives **72**. A final electron transfer from **72** to the $[4Fe\text{-}4S]^{2+}$ cluster completes the formation of **64**.[147]

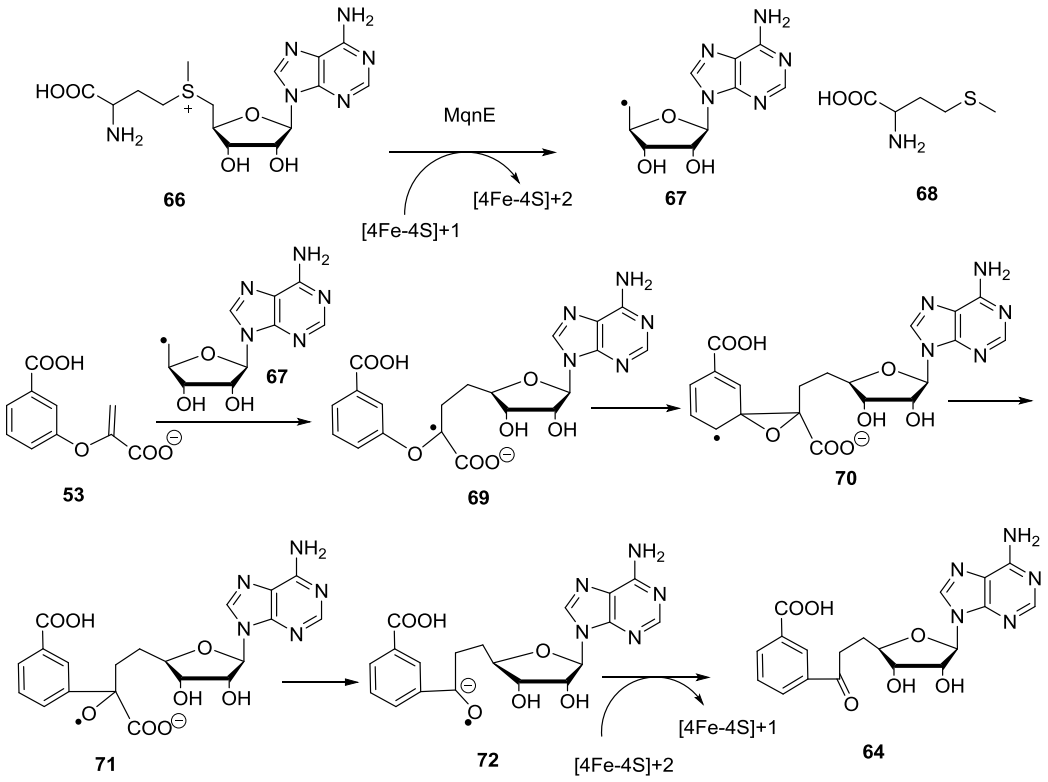



**Figure 5.18**: Mechanistic proposal for the radical SAM enzyme MqnE.[147]

## Conclusion

Previously, it was proposed in this biosynthetic pathway, MqnA (SCO4506) catalyzes the coupling of chorismate, **39**, adenosine/inosine **48** , and phosphoenolpyruvate, **38** to form aminofutalosine, **64** or futalosine, **49** using unknown chemistry.[123] However, failure to reconstitute the formation of futalosine in the MqnA reaction by this proposed route suggested that additional enzymes and cofactors could be involved in this transformation. Our discovery of the radical SAM enzyme MqnE, guided by

bioinformatics analysis followed by *in-vitro* reconstitution indicated that this was the missing enzyme on the pathway.[147] Aminofutalosine is further converted to futalosine by 6-aminodeoxyfutalosine deaminase.[147, 162, 163] Hence, what was previously believed to be a one enzyme (MqnA) catalyzed transformation[123] is actually a three enzymes (MqnA, MqnE and deaminase) mediated reaction (Figure 5.19).[147] This finding also provided the evidence that chorismate derived compound **53** in the MqnA reaction is actually an on-pathway intermediate and also supports our *in-vivo* feeding studies.

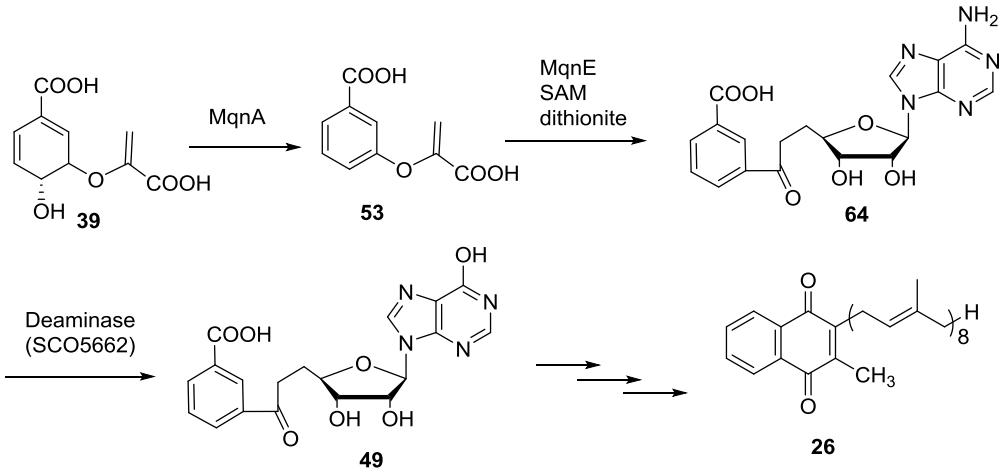



**Figure 5.19**: Revised route for futalosine formation in menaquinone biosynthesis.

In summary, we have reconstituted the formation of aminofutalosine, **64** in the new menaquinone biosynthesis pathway. This reaction involves the addition of an adenosyl radical to the enol ether double bond of chorismate-derived **53** (Figure 5.20) and represents a new catalytic motif in radical SAM enzymology, where all previously

characterized enzymes[164-168] except for diphthamide synthase[174] use the adenosyl radical to abstract a hydrogen atom from the cosubstrate or from an active-site glycine.

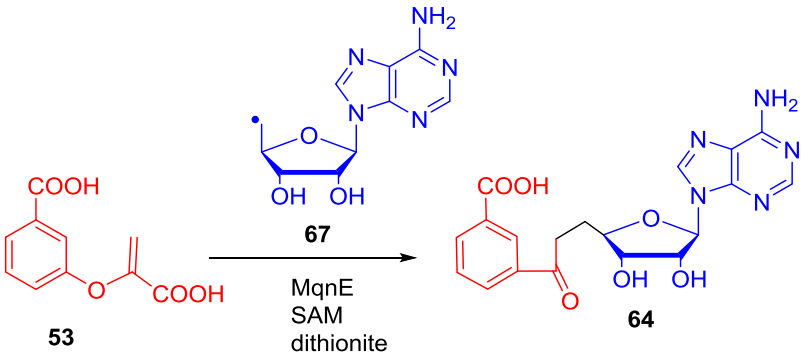



**Figure 5.20:** New reaction motif represented by the radical SAM enzyme MqnE

In other words, the MqnE-catalyzed reaction represents an efficient way to add adenosyl or ribose units to a double bond to generate an intermediate (such as **74** in Figure 5.21A) with rich biosynthetic potential. Other examples of this efficient catalytic motif are likely to emerge in natural product biosynthetic pathways (Figure 5.21A). Griseolic acid, **77** biosynthesis presents one possible example in which intermediate **76** may be formed by addition of the 5'-deoxyadenosyl radical, **67** to fumaric acid, **75** (Figure 5.21B).[175, 176]

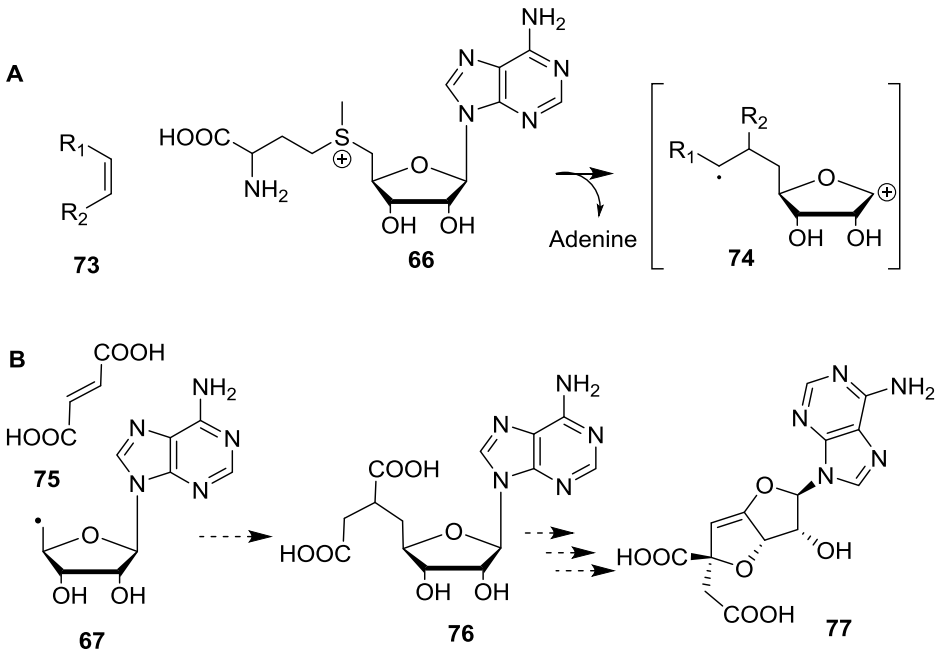



**Figure 5.21**: Potential applications and examples of MqnE type chemistry.[147]
[(A) Adenosyl radical addition to a double bond generates a new, potentially versatile biosynthon **74** enabling chemistry at both C1 and C5 of ribose. (B) Griseolic acid biosynthesis presents a possible example of this chemistry.][175, 176]

## Experimental Procedures[147]

***General methods and materials*:** Molecular biology manipulations were carried out using standard techniques.[64, 152] PCR amplifications were performed using an automatic thermocycler (Mastercycler, Eppendorf) and oligonucleotide primers were synthesized by Integrated DNA Technologies. *Phusion* DNA polymerase was obtained from New England BioLabs. Plasmid mini-prep and PCR purification kits and *E. coli* M15 (pREP4) cells were acquired from Qiagen or Fermentas. DNA fragments were purified from agarose gels with the Zymoclean Gel DNA Recovery Kit from Zymo Research. All restriction enzymes and T4 DNA ligase were purchased from New England Biolabs.

*Escherichia coli* strains DH5α or MachI (Invitrogen) were used as recipients for transformations during plasmid construction as well as for plasmid propagation and storage. Growth media were obtained from Difco. Reagents were purchased from Sigma-Aldrich unless stated otherwise.

***HPLC method***: HPLC analysis using a linear gradient, at a flow rate of 1 mL/min was used with absorbance detection at 254 nm. Solvent A is water, solvent B is 100 mM $K_2HPO_4$, pH 6.6 and solvent C is methanol: 0 min, 100% B; 5 min, 100 % B; 12 min, 48 % A, 40 % B, 12 % C; 14 min, 50 % A, 30 % B, 20 % C; 18 min, 30 % A, 10% B, 60% C; 20 min, 100 % B; 25 min 100 % B. The column used was a Supelcosil LC-18- T HPLC column (15 cm x 4.6 mm, 3 µM particle size).

***LC-MS method:*** Samples were analyzed by reverse-phase HPLC on an Agilent 1200 HPLC system equipped with a diode array UV-Vis detector and a thermostatted autosampler (10 °C). The stationary phase was Supelcosil LC-18-T (15 cm × 3 mm, 3 µm particles), maintained at 22 °C. The LC eluent consisted of a gradient of 75% methanol in 25% water with 5 mM ammonium acetate buffer, pH 6.6 (0.4 mL/min flow rate). The percentages of ammonium acetate buffer (*B*) and methanol (*M*) at time *t* varied according to the following scheme: {*t*,*M*,*B*}: {0,0,100}, {2,0,100}, {10,35,65}, {15,57.5,42.5}, {18,100,0}, {23,100,0}; {26,0,100}; {33,0,100}. Compounds were detected using absorbance at 254 nm. UV-Vis spectra (190 – 640 nm) were also collected and evaluated. Mass data were collected using an in-line Bruker Daltonics

micrOTOF-Q II ESI-Qq-TOF mass spectrometer (HyStar) in positive ion mode as indicated.

***Expression and purification of aminofutalosine synthase, MqnE:*** *Escherichia coli* M15 (pREP4) (Qiagen) harboring the pQE-30 vector containing the *ttha0804* gene (obtained from Tohru Dairi lab) was grown aerobically for 12-15 h at 30 °C in 10 mL of LB medium supplemented with 40 μg/mL kanamycin and 100 μg/mL ampicillin. This starter culture was added to 1.5 L of LB medium, supplemented with 40 μg/mL of kanamycin and 100 μg/mL of ampicillin. The cultures were grown aerobically (220 rpm) at 30 °C until the $A_{600}$ was 0.5-0.6, at which point ferrous ammonium sulfate (100 mg/L) and l-cysteine (100 mg/L) were added. The cultures were then cooled to 4 °C and protein expression was induced by the addition of 100 μM of isopropyl-1-thio-β-d-galactopyranoside (IPTG, LabScientific). The cultures were grown with slow shaking (110 rpm) at 15 °C for an additional ~ 18 h. and the cells were harvested by centrifugation at 6,300×*g* in a Beckman Coulter Avanti J-26 XPI with a JLA-8.1000 rotor (15 min at 4 °C). The medium was discarded and the cell paste (~ 1.5 g per liter of cell culture) was stored submerged in liquid nitrogen until use. Protein purification was carried out in an anaerobic chamber (Coy Laboratories) containing 5% hydrogen and 95% nitrogen. Cells were thawed and resuspended in 40 mL of Lysis Buffer (100 mM Tris-HCl, pH 7.5 at 25 °C) supplemented with 2 mM dl-dithiothreitol (DTT) and 10 mg of lysozyme. After 1.5 hours of incubation on ice with continuous stirring, the cells were further lysed by sonication, while stirring on ice, for 6 × 100 s using a Misonix Sonicator

XL-2000 (setting = 10). Cell debris was removed by centrifugation at 31,000×*g* in a Beckman Coulter Avanti J-E with a JA-17 rotor for 30 min at 4 °C and the supernatant was clarified through a 0.45 μm syringe filter. Aminofutalosine synthase was purified by nickel NTA affinity chromatography using a stepwise gradient of increasing imidazole concentration with two 5 mL HiTrap chelating HP column (GE Healthcare Life Sciences) arranged in tandem to attain a higher binding capacity and better retention efficiency of the protein of interest. The resin was charged with 2 column volumes (CV) of 0.1 M $NiSO_4$, washed with 3 CV of filtered water to remove excess nickel, and equilibrated with 2 CV of Lysis Buffer supplemented with 2 mM. After loading the filtered protein sample, the resin was washed with 5 CV each of Wash Buffers (100 mM Tris-HCl, pH 7.5 at 25 °C, 300 mM NaCl, 1 mM DTT) containing 10 mM, 30 mM, and 250 mM imidazole. Fractions of 0.5 mL were collected and analyzed for purity by 12% SDS-PAGE. $His_6$-TTHA0804 (brownish color) eluted at 250 mM imidazole. Fractions containing the desired protein were pooled and the buffer was exchanged to Final Buffer (100 mM potassium phosphate, pH 7.5 at 25 °C, 2 mM DTT, 30% glycerol) via Econo-Pac 10DG size exclusion chromatography (Bio-Rad). SDS-PAGE analysis showed the molecular weight of the protein as ~ 42.6 kDa (Figure S18). Typical yields were ~ 3.5 mg per liter of cell culture. Proteins were stored submerged in liquid nitrogen.

***Characterization of aminofutalosine, 64*:** Each reaction mixture (150 μL) contained chemically or enzymatically synthesized 3-[(1-carboxyvinyl) oxy]-benzoic acid, **53** (2 mM), S-adenosylmethionine, **66** (2 mM), purified MqnE (200 μM) and excess sodium

dithionite in 100 mM Tris-HCl buffer, pH 7.5. The reaction was allowed to proceed anaerobically for about 10 hrs. The reaction was repeated multiple times to give sufficient product for analysis. The reactions were quenched with an equal volume of 8M guanidine hydrochloride and the protein was removed by ultrafiltration using 10 kDa centrifugal filters (Pall Life Sciences). Aminofutalosine was purified from the reaction mixture by HPLC (eluted at 20.2 minute) using the HPLC method mentioned previously. The compound was buffer exchanged into 5mM ammonium acetate by HPLC. The purified product was concentrated using a speedvac concentrator (Thermo Scientific), dissolved in $D_2O$ and characterized by $^{1}H$ and $^{13}C$ NMR (Figures 5.8, 5.9).

$^{1}H$ NMR (500 MHz, $D_2O$): 8.09 (t, J = 1.9 Hz, 1H), 8.02 (s, 1H), 7.96 (s, 1H), 7.88 (dt, J = 7.9, 1.5 Hz, 1H), 7.59 (dt, J = 7.9, 1.5 Hz, 1H), 7.24 (t, J = 7.8 Hz, 1H), 5.80 (d, J = 5 Hz, 1H), 4.85 (t, J = 5 Hz, 1H), 4.32 (t, J = 5 Hz, 1H), 4.14-4.19 (m, 1H), 2.96-3.11 (m, 2H), 2.07-2.29 (m,2H). (Peaks from trace contaminant acetone in the sample also appears along with the f,f' peaks of the compound, **64).**

$^{13}C$ NMR (125 MHz, $D_2O$): 204.5, 173.9, 155.3, 152.5, 148.6, 139.9, 136.3, 135.7, 133.7, 129.9, 128.3, 128.27, 118.7, 87.4, 83.6, 72.9, 72.6, 34.2, 27.5. (Peak at 205 and 33 ppm also comes from trace acetone contaminant in the sample).

DEPT-135 $^{13}C$ NMR (Figure A11, AppendixA): 152.5 (CH), 139.9 (CH), 133.7 (CH), 129.9 (CH), 128.3 (CH), 128.27 (CH), 87.4 (CH), 83.6 (CH), 72.9 (CH), 72.6 (CH), 34.2

($CH_2$), 27.5 ($CH_2$). The 2D NMR spectra (COSY and HSQC) of aminofutalosine **64** are shown in Figures 5.10 and 5.11 respectively.

LC-MS analysis (ESI positive ion mode) of the TTHA0804 reaction mixture identified the new compound with m/z 414.14 $[M+H]^+$ as aminofutalosine, **64** (theoretical m/z 414.14) (Figure 5.12).

***Expression and purification of the aminofutalosine deaminase*:** *Escherichia coli* TB1 (NEB) harboring the pMAL-c2x vector containing the SCO5662 gene (obtained from Tohru Dairi lab) fused to the maltose binding protein [162] was grown aerobically for 12-15 h at 37 °C in 20 mL of LB medium supplemented with 100 μg/mL ampicillin. This starter culture (20 mL) was added to 2 L of LB medium supplemented with 100 μg/mL ampicillin. The cultures were grown aerobically with shaking (220 rpm) at 37 °C until the $A_{600}$ was 0.6-0.7, at which point the cultures were cooled to 4 °C and protein expression was induced by the addition of 1 mM isopropyl-1-thio-β-d-galactopyranoside (IPTG, Lab Scientific). The cultures were grown with slow shaking (180 rpm) at 15 °C for an additional ~ 16 h. and the cells were harvested by centrifugation at 6,300×*g* in a Beckman Coulter Avanti J-26 XPI with a JLA-8.1000 rotor for 15 min at 4 °C. The medium was discarded and the cell paste (~ 2 g per liter of cell culture) was stored at -80 °C until use. Protein purification was carried out using amylose resin affinity chromatography. Cells were thawed and resuspended in 20 mL of column Buffer (20 mM Tris, 200 mM NaCl, 1 mM EDTA, pH 7.4 at 25 °C) supplemented with 20 mg of

lysozyme and 1 mM dl-dithiothreitol (DTT). Cells were further lysed by sonication 6 × 30 s using a Misonix Sonicator 3000 (24 W, 1.5 s pulse, 1.5 s pause) and the cell debris was removed by centrifugation at 31,000×*g* in a Beckman Coulter Avanti J-E with a JA-17 rotor for 45 min at 4 °C and the supernatant was clarified by passage through a 0.45 μm syringe filter. Amylose resin (15 ml, NEB) was poured into a 2.5 x 10 cm column and the column was washed with 5 CV of column buffer. Then, the filtered crude lysate was loaded onto the column and allowed to flow under gravity. The column was washed with 6 CV of column buffer and the protein was eluted with 50 ml of the elution buffer (20 mM Tris, 200 mM NaCl, 1 mM EDTA, 10 mM maltose, pH 7.4) Fractions of 5 mL were collected and analyzed for purity by 12% SDS-PAGE. The fractions containing the desired protein were pooled and the buffer was exchanged to Final Buffer (100 mM Tris-HCl, pH 7.5 at 25 °C, 1 mM DTT, 30% glycerol) via Econo-Pac 10DG size exclusion chromatography (Bio-Rad). Protein concentrations were determined using the Bradford assay or $A_{280}$ measurements. Factor Xa protease was used to cleave the maltose binding protein (MBP) from the deaminase at a w/w ratio of 1% the amount of fusion protein. The deaminase in the cleavage mixture was purified by anion exchange chromatography using HiTrap QFF (GE Life Sciences) as follows: The cleavage mixture was dialyzed against 20 mM sodium phosphate buffer, 25 mM NaCl, pH 5.5 for 4 hrs. The anion exchange column (5 mL) was washed with 3CV of the same buffer and the fusion protein cleavage mixture was loaded onto the column. The column flow through (1 mL fractions) was collected and the column was washed with 5CV of the same buffer. The flow through fractions were checked by A280 measurements and fractions

containing the deaminase were pooled and concentrated using a centrifugal filtering device with a 10 kDa MWCO membrane (Millipore) and the buffer was exchanged to Final Buffer (50 mM potassium phosphate, pH 7.5 at 4 °C, 150 mM KCl, 30% glycerol) via Econo-Pac 10DG size exclusion chromatography. SDS-PAGE analysis showed the molecular weight of SCO5662 as ~ 42.4 kDa (data not shown). Typical yields were ~ 20 mg of protein per liter of cell culture. Proteins were stored at -80 °C.

***Expression and purification of the aminofutalosine nucleosidase:*** *Escherichia coli* TB1 (NEB) harboring the pMAL-c2x vector containing the *hp0089* gene [162] (obtained from Tohru Dairi lab) was grown aerobically for 12-15 h at 37 °C in 20 mL of LB medium supplemented with 100 μg/mL ampicillin. This starter culture (20 mL) added to 2 L of LB medium supplemented with 100 μg/mL ampicillin. The cultures were grown aerobically with shaking (220 rpm) at 37 °C until the $A_{600}$ was 0.6-0.7, at which point the cultures were cooled to 4 °C and protein expression was induced by the addition of 1 mM isopropyl-1-thio-β-d-galactopyranoside (IPTG, LabScientific). The cultures were grown with slow shaking (180 rpm) at 15 °C for an additional ~ 16 h. Cells were harvested by centrifugation at 6,300×*g* in a Beckman Coulter Avanti J-26 XPI with a JLA-8.1000 rotor for 15 min at 4 °C. The medium was discarded and the cell paste (~ 2 g per liter of cell culture) was stored at -80 °C until use. Protein purification was carried out using amylose resin affinity chromatography. Cells were thawed and resuspended in 20 mL of column Buffer (20 mM Tris, 200 mM NaCl, 1 mM EDTA, pH 7.4 at 25 °C) supplemented with 20 mg of lysozyme and 1 mM dl-dithiothreitol (DTT). Cells were

further lysed by sonication 6 × 30 s using a Misonix Sonicator 3000 (24 W, 1.5 s pulse, 1.5 s pause) and the cell debris was removed by centrifugation at 31,000×*g* in a Beckman Coulter Avanti J-E with a JA-17 rotor for 45 min at 4 °C and the supernatant was clarified by passage through a 0.45 μm syringe filter. 15 ml of the amylose resin (NEB) was poured into a 2.5 x 10 cm column and the column was washed with 5 CV of column buffer. Then, the filtered crude lysate was loaded onto the column and allowed to flow under gravity. The column was washed with 6 CV of column buffer and the protein was eluted with 50 ml of the elution buffer (20 mM Tris, 200 mM NaCl, 1 mM EDTA, 10 mM maltose, pH 7.4) Fractions of 5 mL were collected and analyzed for purity by 12% SDS-PAGE. The fractions containing the glycosidase were pooled and the buffer was exchanged to Final Buffer (100 mM Tris-HCl, pH 7.5 at 25 °C, 1 mM DTT, 30% glycerol) via Econo-Pac 10DG size exclusion chromatography (Bio-Rad). Protein concentrations were determined using the Bradford assay or $A_{280}$ measurements. Factor Xa protease was used to cleave the maltose binding protein (MBP) from the glycosidase at a w/w ratio of 1% the amount of fusion protein. The glycosidase was purified by anion exchange chromatography using HiTrap QFF (GE Life Sciences) as described above. SDS-PAGE analysis showed the molecular weight to be ~25.2 kDa (data not shown). Typical yields were ~ 15-20 mg of protein per liter of cell culture. Proteins were stored at -80 °C.

***Analysis of the MqnE reaction mixture of bicarbonate and formate*:** The fate of the carboxyl group of **53** was determined using $^{13}$C-NMR. Potassium phosphate buffer (60

mM, pH 8.6) in degassed water, was boiled for 30 min and then cooled and stored under a nitrogen atmosphere inside the anaerobic chamber until use. This prevented any contamination from atmospheric carbon dioxide. All reactions were carried out in this buffer in the anaerobic chamber. [U- $^{13}C_{10}$] 3-[(1-carboxyvinyl) oxy]-benzoic acid was prepared by incubating [U- $^{13}C_{10}$] chorismate (4 mM) with MqnA (500 µM) for 5 hrs. ([U- $^{13}C_{10}$] chorismate was isolated from a chorismate mutase deficient strain of *Escherichia coli* KA12 [177] by feeding [U- $^{13}C_6$] glucose according to the procedure described in Rieger *et al* [150]. ) After removing MqnA by ultrafiltration, [U- $^{13}C_{10}$] 3-[(1-carboxyvinyl) oxy]-benzoic acid was incubated with MqnE (200 µM), SAM, **66** (2.5 mM) and excess sodium dithionite for 12 hrs. Four control reactions were also set up lacking individual components. After removing MqnE by ultrafiltration, the reaction mixtures (150 µL) were transferred into 3 mm NMR tubes inside the anaerobic chamber, sealed with septa and parafilm and analyzed by $^{13}C$ NMR on a Bruker Avance III 500 MHz NMR instrument with a H-C-N Cryoprobe ($^{13}C$ 125 MHz). A new signal was detected at 160.01 ppm in the MqnE reaction containing all the components, which was not present in any of the control reactions (Figure 5.17). The chemical shift of the new signal matched exactly with a $^{13}C$ sodium bicarbonate standard and the signal intensity also increased when spiked with an authentic sample of $^{13}C$ bicarbonate. This suggests that $CO_2$, trapped as bicarbonate in the basic reaction medium, is the most likely fate of the carboxyl group of **53**. No NMR signal for formate was detected in any of the samples.

***Iron content of the aminofutalosine synthase, MqnE:*** Iron content was analyzed as previously described. [178] The following reagents were prepared: 8 M guanidine hydrochloride, 2 M HCl, 10 mM ferrozine in water (3-(2-Pyridyl)-5,6-diphenyl-1,2,4-triazine-*p,p'*-disulfonic acid monosodium salt hydrate), 100 mM l-ascorbic acid in water and saturated ammonium acetate solution. To 100 µL of enzyme solution containing 0 to 50 µM TTHA0804, 100 µL of 8 M guanidine hydrochloride and 100 µL of 2 M HCl were added and the solution was diluted to 550 µL. The resulting solution was centrifuged to remove any denatured protein. To 500 µL of the supernatant, 30 µL of the 10 mM ferrozine solution and 30 µL of 100 mM freshly prepared L-ascorbic acid were added. The resulting solution was mixed thoroughly and incubated at RT for 30 min. The absorbance of the iron ferrozine complex was recorded at 562 nm using a Varian Cary Bio 300 UV-Visible Spectrophotometer (extinction coefficient for iron ferrozine complex is 27.9 $mM^{-1}cm^{-1}$). The iron content was determined by comparing this reading to a standard curve that was generated under identical conditions using Fe $(NH_4)_2(SO_4)_2$ with a concentration range from 0 to 100 µM.

***Sulfide content of the aminofutalosine synthase, MqnE:*** Sulfide content was analyzed using the methylene blue assay. [179] To 300 µL of assay solution containing 0 to 50 µM TTHA0804, 1 ml of 1% (w/v) zinc acetate was added followed by 50 µL of 3 M NaOH. This mixture was agitated gently and 250 µL of 0.1 % N,N-dimethyl-*p*-phenylenediamine (DMPD) monohydrochloride in 5 M HCl and 50 µL of 23 mM $FeCl_3$ in 1.2 M HCl were added. The resulting solution was mixed vigorously after each 5 min

intervals over a 30 min time period. The samples were then centrifuged at 16,110×*g* in an Eppendorf 5415 R with a F45-24-11 rotor for 5 min at 25 °C. The supernatant was collected and the absorbance at 670 nm was recorded using a Varian Cary Bio 300 UV-Visible Spectrophotometer. The sulfide content was determined by comparing this reading to a standard curve that was generated under identical conditions using a fresh solution of sodium sulfide ($Na_2S$; Fisher Scientific) in 0.1 M NaOH with a concentration range from 0 to 100 µM.

# CHAPTER VI

# MENAQUINONE BIOSYNTHESIS: MECHANISTIC STUDIES ON THE RADICAL SAM ENZYME, MqnE[#]

## Introduction

Menaquinone (**26**, vitamin K) is a lipid-soluble molecule that shuttles electrons between membrane-bound redox enzymes in the bacterial respiratory chain.[75] It is also a very important vitamin in mammals where it participates as a cosubstrate in the carboxylation of glutamic acid residues in various proteins involved in blood coagulation[110] and bone morphogenesis[111]. Various steps of the alternative menaquinone biosynthetic pathway are reconstituted in *vitro*.[123,142,144,147]The pathway begins (Figure 6.1) with MqnA (SCO4506) catalyzing the transformation of chorismate, **39** into 3-[(1-carboxyvinyl) oxy]-benzoic acid, **53** and then the radical SAM enzyme MqnE (SCO4494) catalyzes the transformation of the latter into aminofutalosine **64** in the presence of S-adenosylmethionine (SAM) and dithionite.[147] Aminofutalosine is then converted to compound **50** by aminofutalosine hydrolase (MqnB, SCO4327) or by a combination of aminofutalosine deaminase (SCO5662) and MqnB[142,162]. Cyclization of **50** by another radical SAM enzyme MqnC (SCO4550) gives compound **51** [144] which is then converted to 5, 8-dihydroxy-2-naphthoic acid **52** by MqnD (SCO4326).[123] Bioinformatics studies suggests that **52** is converted to menaquinone **26** in *S. coelicolor* A3(2) in several tailoring reactions catalyzed by SCO4491 (prenylation), SCO4556 (methylation), and SCO4490 or SCO4492 (decarboxylation)(Figure 6.1).[123]

[#](Nilkamal Mahanta carried out all the biochemical studies reported here. Dmytro Fedoseyenko synthesized compounds **83**, **88**, **96**, **101a**, **101b**, **101c**, **109**, **114, 117**,**118**)

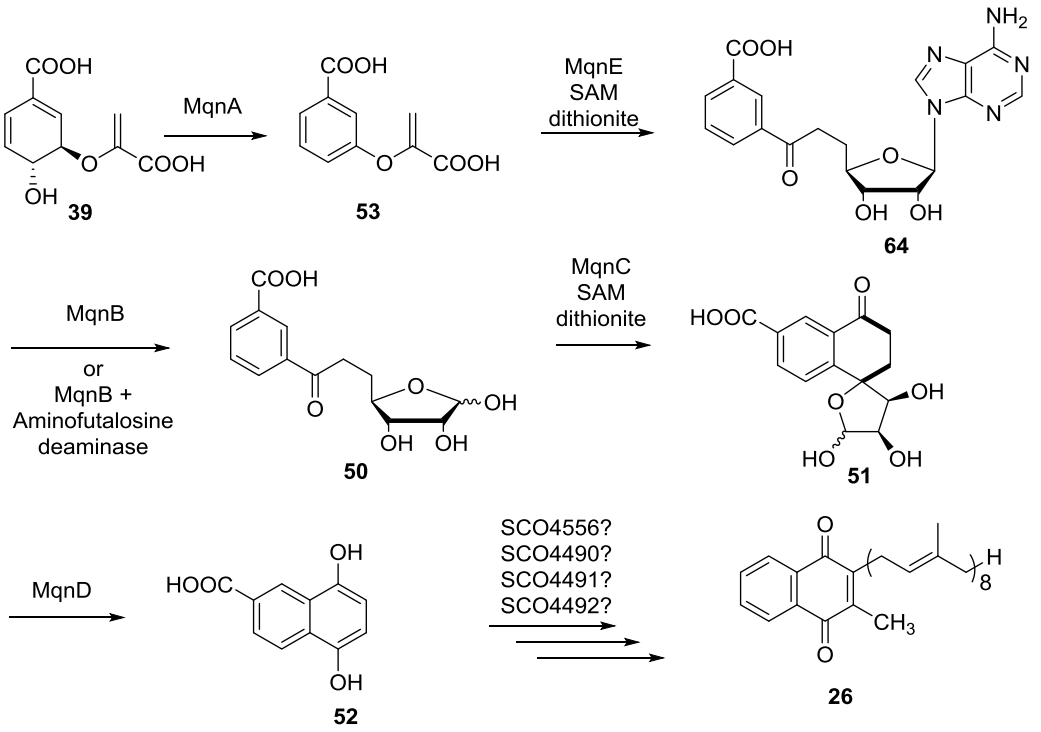



**Figure 6.1**: Aminofutalosine dependent pathway for the biosynthesis of menaquinone.

One of the most interesting enzymes on this pathway, MqnE catalyzes the addition of 5'-deoxyadenosyl radical to the enol ether double bond of 3-[(1-carboxyvinyl) oxy]-benzoic acid **53** to finally yield aminofutalosine **64**. It represents a unique catalytic motif in radical SAM enzymology, where all previously reported radical SAM enzymes[164-168] use the 5'-deoxyadenosyl radical to abstract a hydrogen atom from the substrate or from the active site glycine residue to carry out the radical mediated transformations. A mechanism for the MqnE catalyzed reaction has been proposed (Figure 6.2).[147] Various biochemical experiments that are performed to test our mechanistic proposal are described in this chapter.

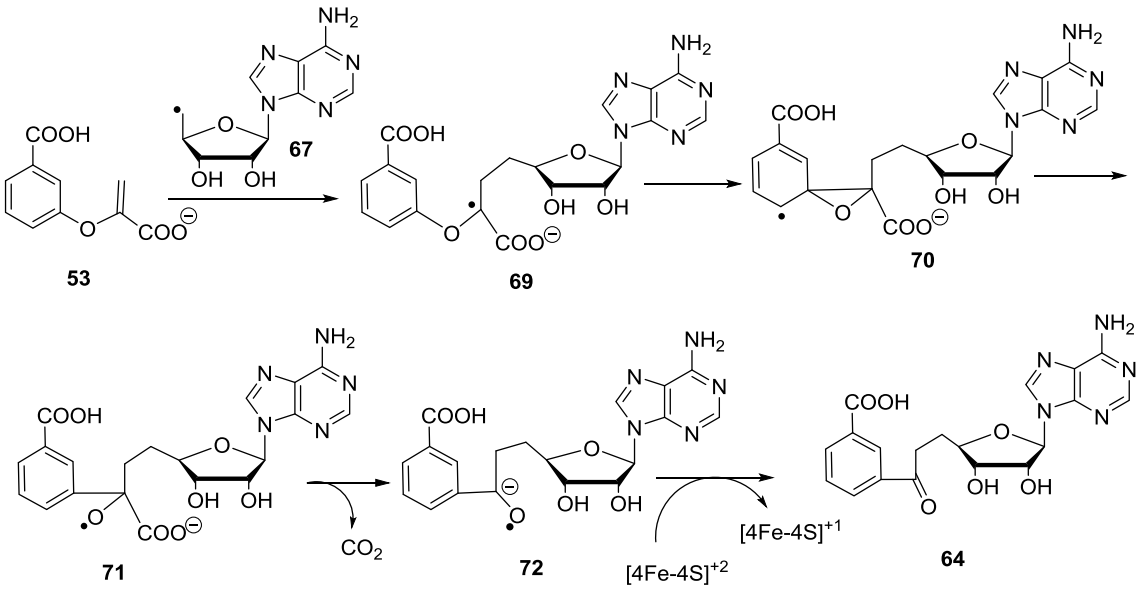



**Figure 6.2**: Mechanistic proposal for the MqnE catalyzed reaction.

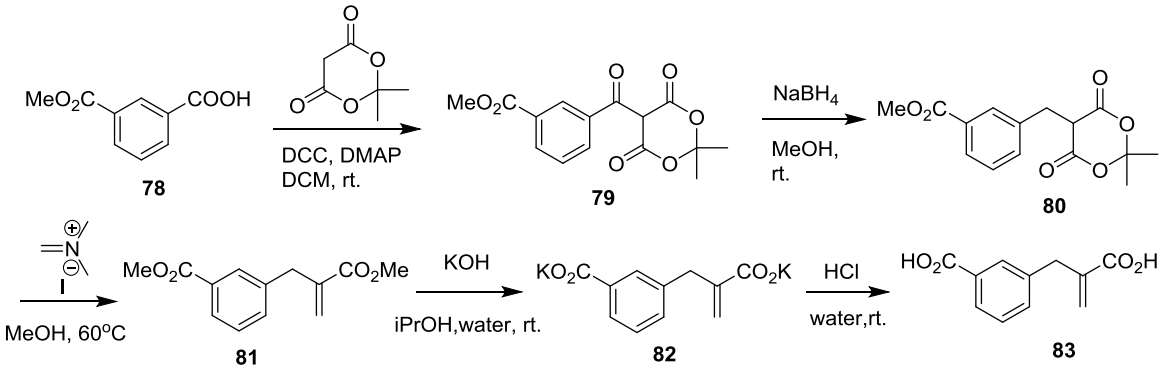
## Results and discussion

***Mechanistic studies*:** We have previously proposed a mechanism for the MqnE catalyzed reaction (Figure 6.2). In this proposal, addition of the radical **67** to the vinyl double bond of compound **53** gives the captodative radical **69**. Rearrangement of this radical via **70** gives **71**. Decarboxylation of **71**, facilitated by the alkoxy radical, gives **72**. A final electron transfer from **72** to the $[4Fe\text{-}4S]^{2+}$ cluster completes the formation of **64**. In this study, we have used different substrate analogs to trap intermediate radical species to provide support for the various steps of the mechanism.

***Trapping of the captodative radical 69***: To probe the first step of the mechanistic proposal (Figure 6.2), a substrate analog **83** was synthesized (Scheme 6.1) in which the oxygen in substrate **53** is replaced with a methylene group. To test whether compound

**83** is a substrate, MqnE (100 uM) was incubated with compound **83** (1.5 mM) anaerobically in the presence of SAM (2 mM) and excess sodium dithionite. After 10 hrs., the reaction was quenched with 8M guanidine hydrochloride, the protein was removed by ultrafiltration and the reaction mixture was analyzed by reverse phase HPLC. This analysis showed the formation of a new compound at 19.3 min in the HPLC chromatogram (Figure 6.3) which was identified as compound **84** by NMR spectroscopy (Figures B5-B10, Appendix B) and MS spectrometry (Figure 6.4). Formation of this compound was not observed in any of the control reactions. Moreover, formation of 5'-deoxyadenosine was also not observed in the HPLC analysis of the full reaction mixture.

***Synthesis of compound 83:*** Compound **83** was synthesized using the scheme shown below:




**Scheme 6.1**: Synthetic scheme for compound **83**

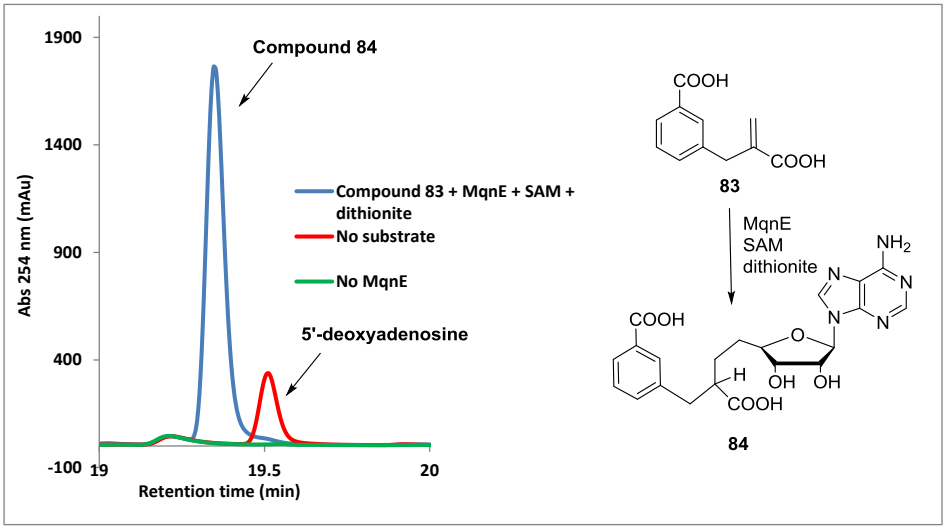



**Figure 6.3**: HPLC analysis of the MqnE reaction mixture with substrate analog **83**. [Blue trace: chromatogram of the full reaction mixture (MqnE + **83**+ SAM+ dithionite). Red trace: chromatogram of the control reaction mixture lacking **83**. Uncoupled production of 5'-deoxyadenosine was observed at 19.5 min in this control reaction. Green trace: chromatogram of the control reaction mixture lacking MqnE. No signal at 19.3 min was observed in control reactions individually lacking SAM or dithionite (data not shown). The inset shows the structure of the product **84**].

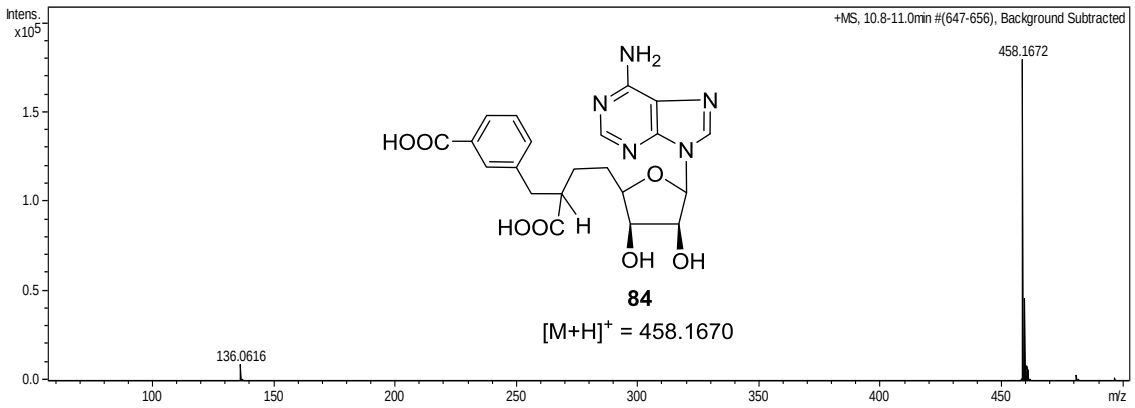



**Figure 6.4:** LC-MS analysis of the MqnE reaction mixture with substrate analog **83**. [In the mass spectrum, m/z 458.1672, $[M+H]^+$ corresponds to compound **84** (theoretical m/z 458.1670) and m/z 136.0616 corresponds to adenine derived from aminofutalosine (theoretical m/z 136.0620)]

Formation of compound **84** suggested the addition of 5'-deoxyadenosyl radical to the vinyl double bond of **83** to yield the radical intermediate **85**, which upon abstraction of a hydrogen atom, led to the formation of compound **84** (Figure 6.5).

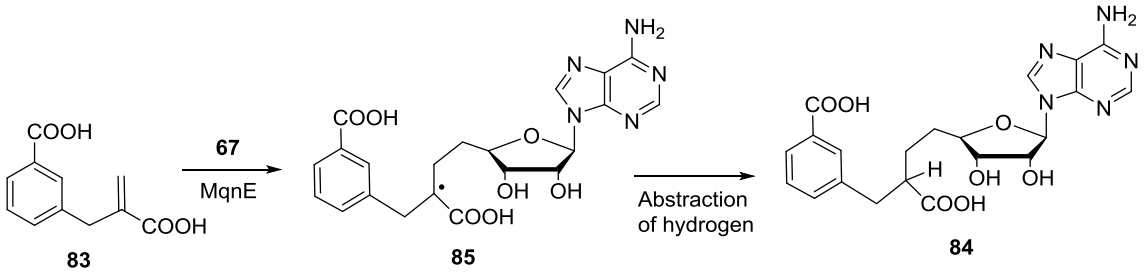



**Figure 6.5**: Mechanistic proposal for formation of compound **84**.

When MqnE reaction was run using **83** as substrate in deuterated buffer, incorporation of deuterium was observed into **84** (Figure B11, Appendix B) suggesting that the hydrogen is probably derived from a buffer component or from an exchangeable amino acid residue in the active site. Moreover, analysis of MqnE by intact protein mass spectrometry did not detect any change in the mass of the enzyme after the reaction which rules out any enzyme-substrate covalent adduct formation (data not shown). Analysis of MqnE after the reaction with **83** by SDS-PAGE (Figure B12, Appendix B) failed to detect any degradation of the enzyme suggesting that radical **85** probably does not abstract a hydrogen atom from the α- carbon of the backbone peptide which, in turn, may lead to a radical mediated cleavage of the backbone altering the mass of the protein.

We provide two possible mechanistic explanations for the formation of compound **84**. First, we propose that replacing the bridging oxygen of **53** with a methylene group as in **83** might slow down the decarboxylation of radical **87** significantly (Figure 6.6) as it would yield a less stable alkene radical anion as compared to a stable ketyl radical anion **72** (Figure 6.2) when **53** is used as the substrate. In such as case, the carboxyl stabilized radical **85** which probably has the highest concentration in this equilibrium (Figure 6.6) would capture a hydrogen atom to give **84**. The turn over number is close to one.

Secondly, since ring opening of the epoxide radical **70** (Figure 6.2) is likely to be accelerated by acid catalysis, we believe that substituting the bridging oxygen of **53** with a carbon might block the rearrangement of the radical **85** to radical **87** through the spiro intermediate **86** as it cannot undergo acid catalyzed ring opening (Figure 6.6). Radical **85** would ultimately capture a hydrogen atom to yield compound **84**. However, a crystal structure of MqnE bound with **53** or **83** would be required to support this hypothesis. Identification of compound **84** provides evidence for the first step of the MqnE reaction mechanism, where 5'-deoxyadenosyl radical is shown to be involved in a radical addition reaction with the substrate in radical SAM enzymology.

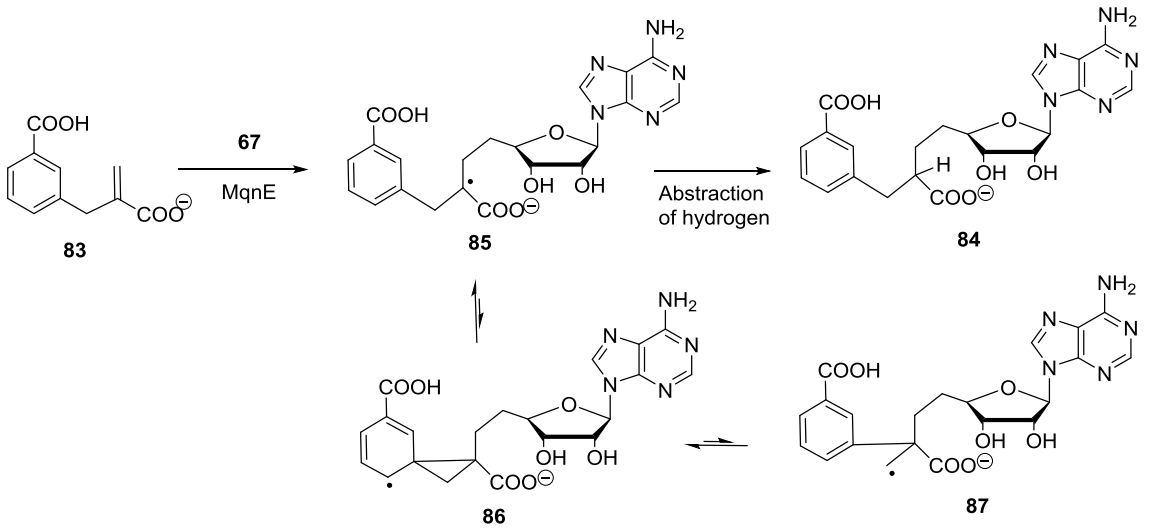
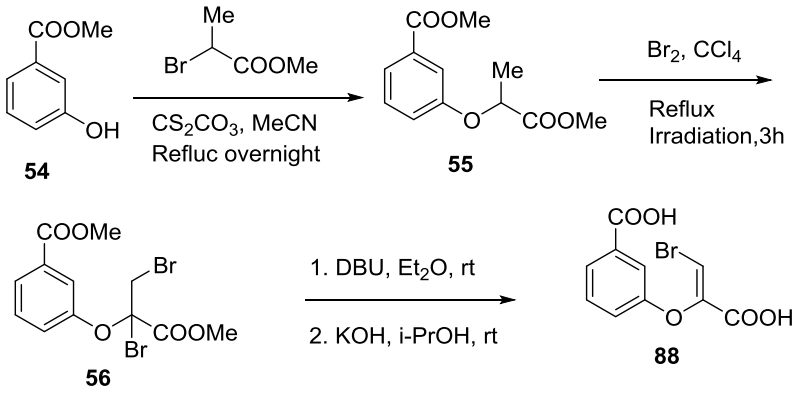



**Figure 6.6:** Probable mechanism for the formation of compound **84**.

***Alternate strategy for trapping the captodative radical 69:*** As an alternative strategy to trap the captodative radical **69** proposed in the mechanism (Figure 6.2), another substrate analog **88** was designed and synthesized in which a halogen group (Br) was introduced in the vinyl double bond terminus of the substrate **53** (Scheme 6.2).

Carbon halogen bonds, beta to a radical center are known to undergo very rapid fragmentation reactions (C-Br $k_{cleavage} = 3x10^{8}s^{-1}$)[180, 181]. Hence using **88** as the substrate might help in trapping the captodative radical as compound **90** after the loss of the bromine radical which would also serve as a mechanistic evidence for the 5'-deoxyadenosyl radical addition step in the MqnE catalyzed reaction (Figure 6.7).[147]

***Synthesis of compound 88:*** Compound **88** was synthesized using the scheme shown below.




**Scheme 6.2:** Synthetic scheme for substrate analog **88**

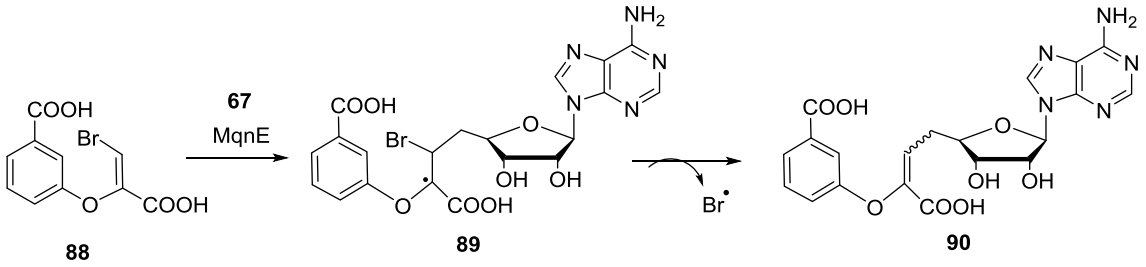



**Figure 6.7**: Mechanistic proposal for trapping the captodative radical **89**.

To test this hypothesis, MqnE (100 uM) was incubated with compound **88** (2 mM) anaerobically in the presence of SAM (2 mM) and excess sodium dithionite. After 10 hrs., the reaction was quenched with 8M guanidine hydrochloride, the protein was removed by ultrafiltration and the reaction mixture was analyzed by reverse phase HPLC. This analysis showed the formation of a new compound at 17.6 min in the HPLC

chromatogram (Figure 6.8) which was identified as compound **90** by 1D and 2D NMR spectroscopy (Figures 6.9-6.13) and MS spectrometry (Figure 6.14). This compound has resulted from the cleavage of the C-Br bond beta to the radical center in **89** generated by the addition of 5'-deoxyadenosyl radical to the enol ether double bond of the compound **88** (Figure 6.6). This observation is consistent with our mechanistic proposal (Figure 6.2) where the addition of the 5'-deoxyadenosyl radical on the substrate has been implicated in the formation of C-C bond in the product.

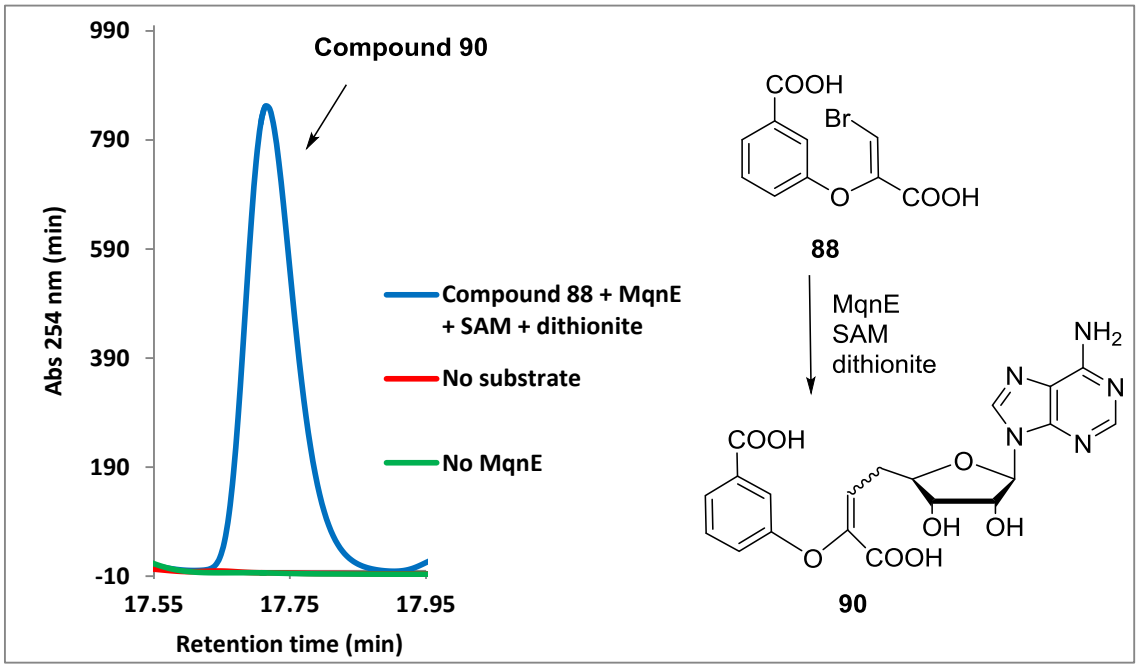



**Figure 6.8:** HPLC analysis of the MqnE reaction mixture with compound **88** (A).
[Blue trace: chromatogram of the full reaction mixture (MqnE + **88**+ SAM+ dithionite). Red trace: chromatogram of the control reaction mixture lacking **88**. Green trace: chromatogram of the control reaction mixture lacking MqnE. No signal at 17.6 min was observed in control reactions individually lacking SAM or dithionite (data not shown). The inset shows the structure of the identified reaction product **90** which was characterized by 1D and 2D NMR (Figures 6.9-6.13) and MS analysis (Figure 6.14)].

***NMR and MS characterization of compound 90*:**

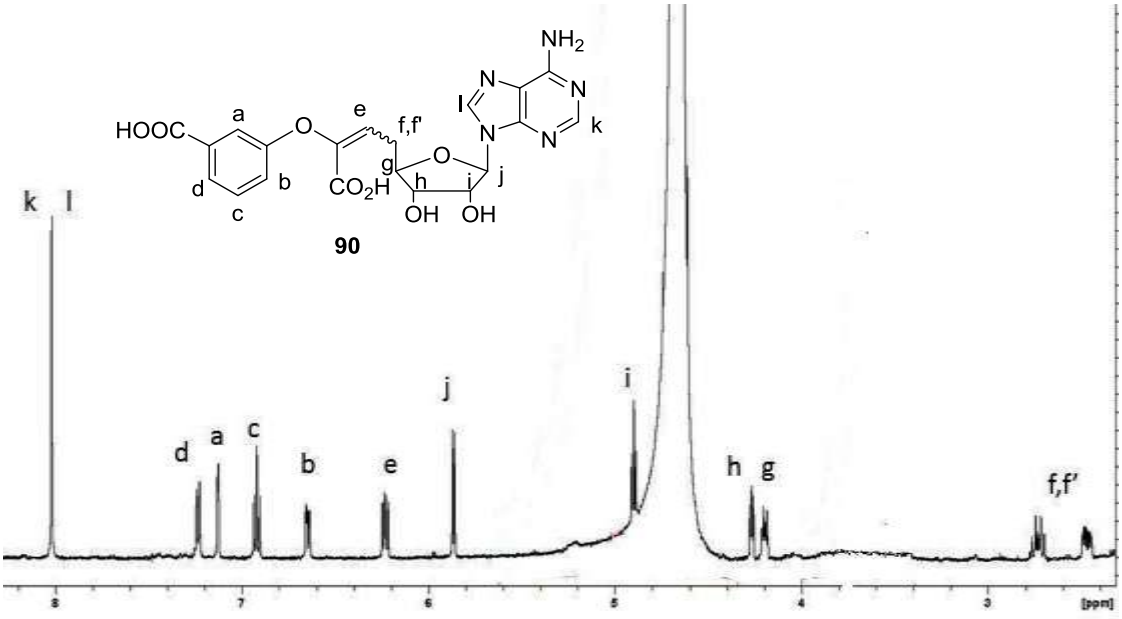



**Figure 6.9:** $^{1}$H NMR (500 MHz, $D_2O$) spectrum of compound **90**.

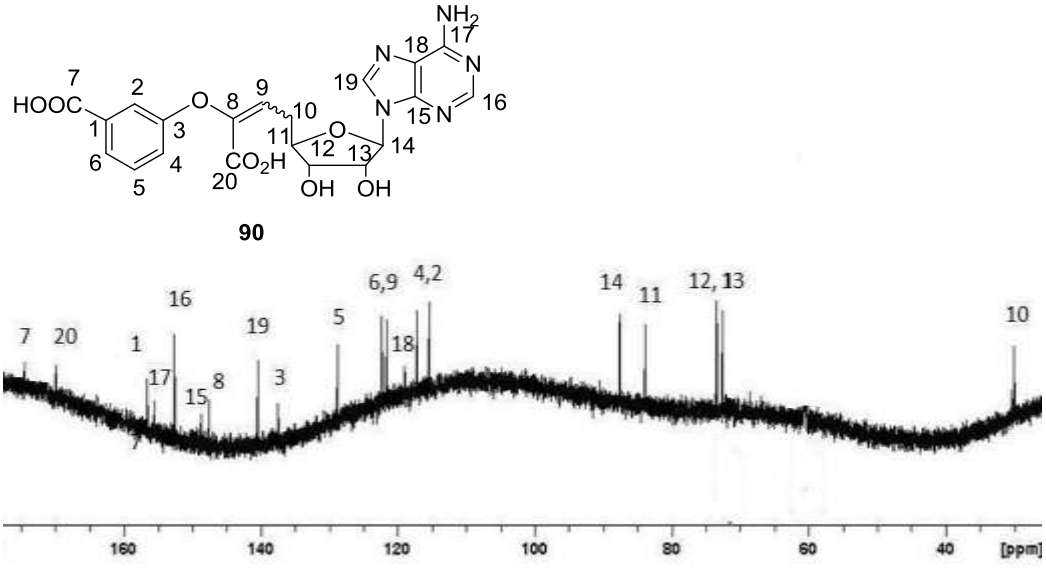



**Figure 6.10:** $^{13}$C NMR (125 MHz, $D_2O$) spectrum of compound **90**.

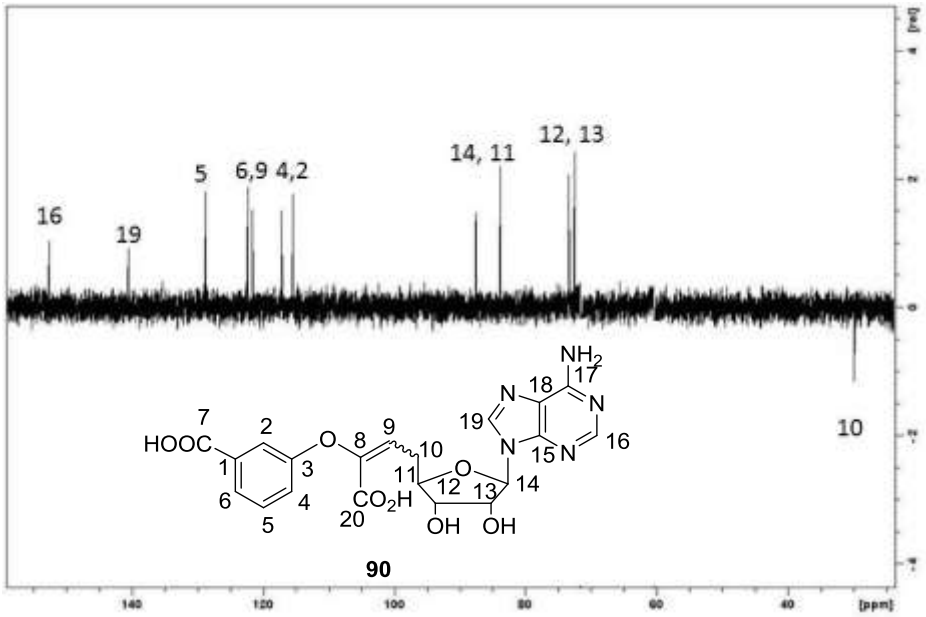



**Figure 6.11**: $^{13}C$ DEPT-135 NMR (125 MHz, $D_2O$) spectrum of compound **90**

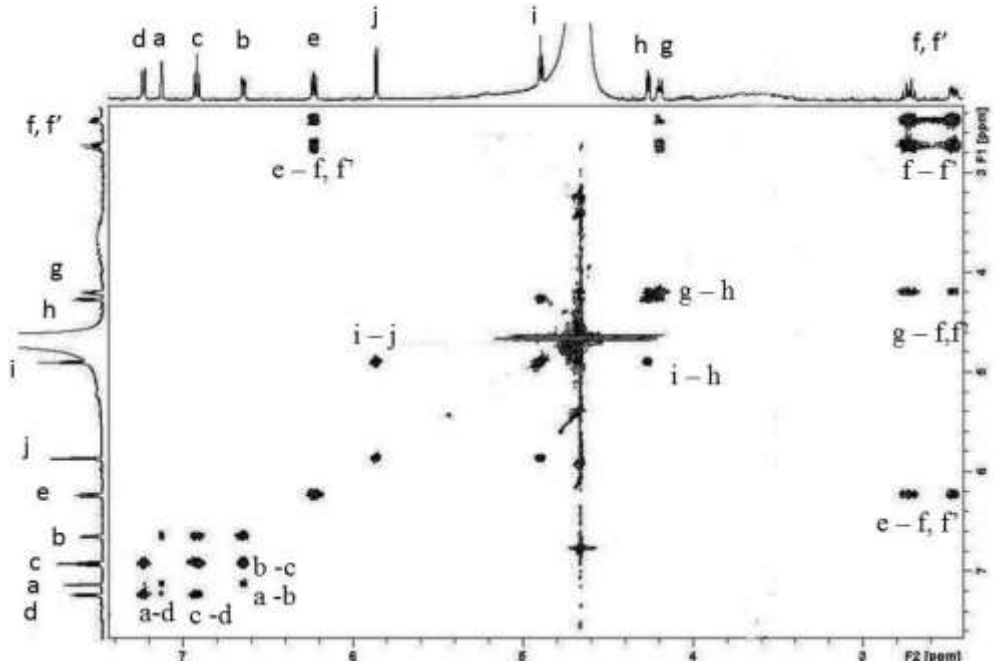



**Figure 6.12:** $^{1}H$-COSY spectrum of compound **90.**

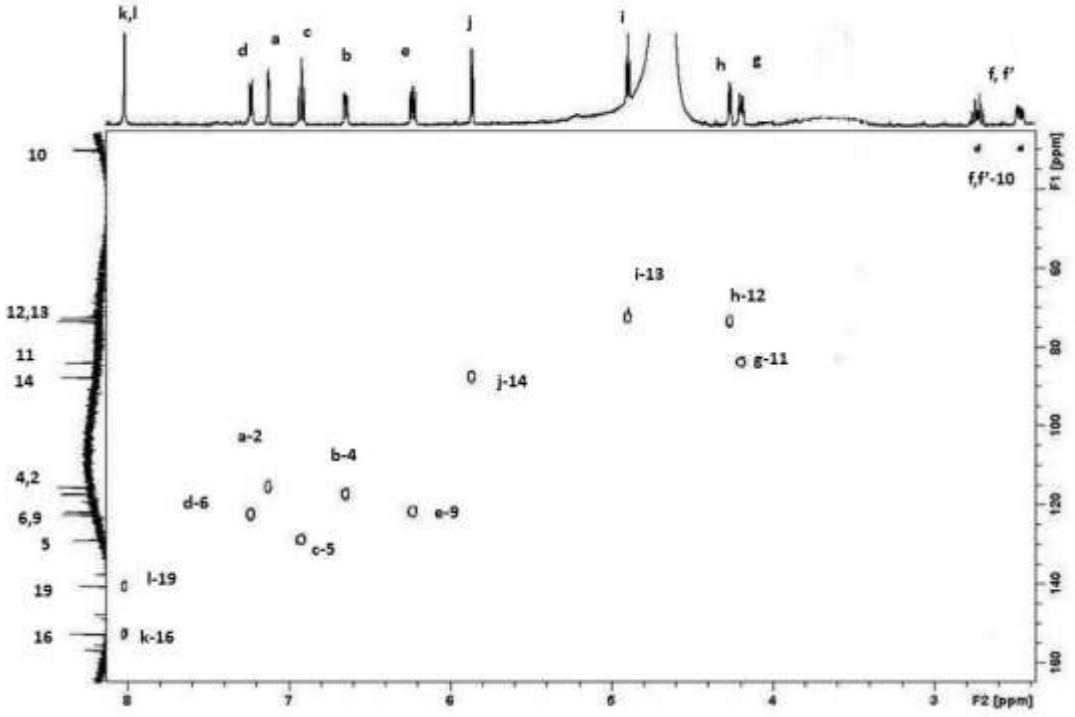



**Figure 6.13:** HSQC spectrum of compound **90**.

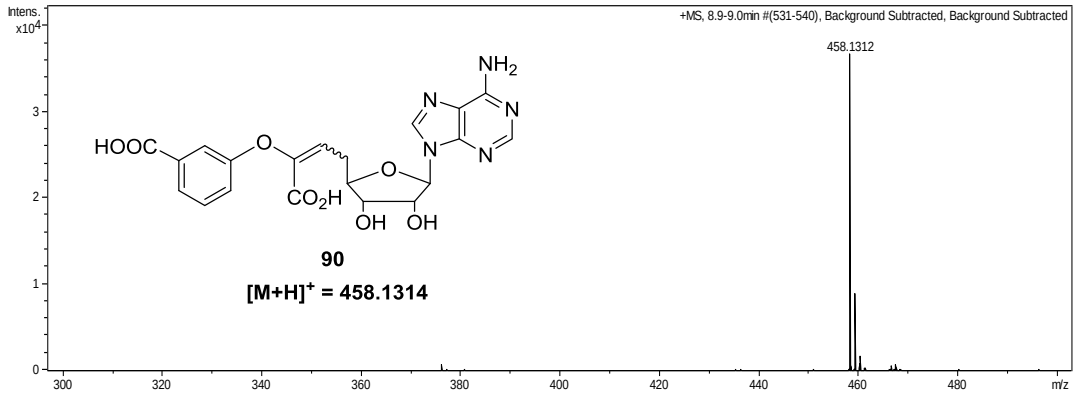



**Figure 6.14:** LC-MS (ESI, positive ion mode) analysis of compound **90**.
[In the mass spectrum, m/z 458.1312, $[M+H]^+$ corresponds to compound **90** (theoretical m/z 458.1314)].

Careful observation of the HPLC chromatogram led to the finding of another new product in the reaction mixture in low intensity at 19.25 min (Figure 6.15) which was identified to be aminofutalosine **64** by NMR and MS analysis.[147]

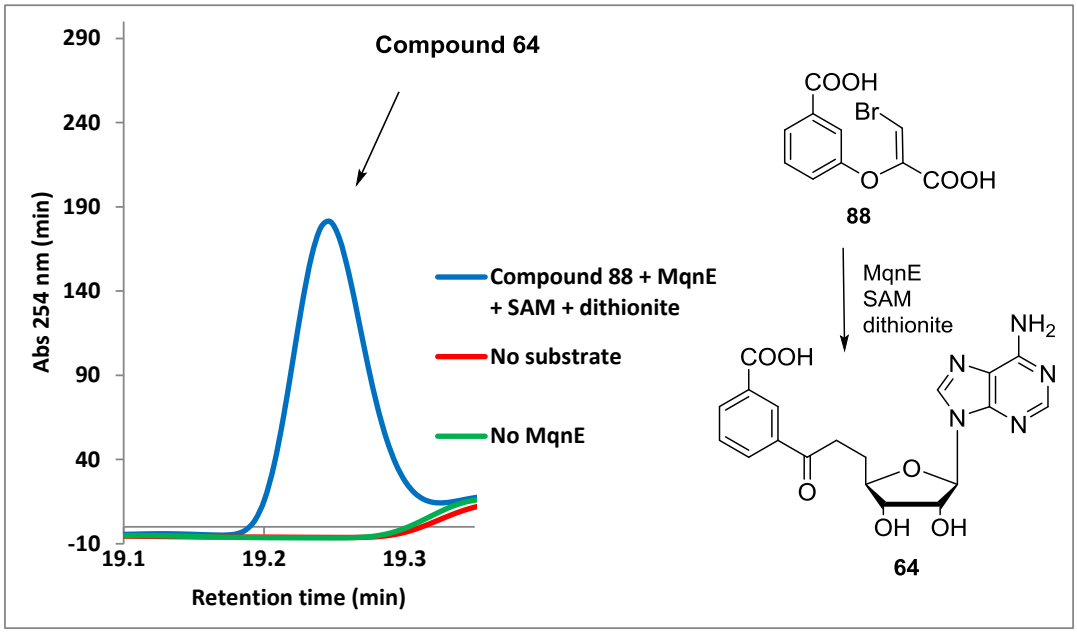



**Figure 6.15**: HPLC analysis of the MqnE reaction mixture with compound **88** (B). [Blue trace: chromatogram of the full reaction mixture (MqnE + **88**+ SAM+ dithionite). Red trace: chromatogram of the control reaction mixture lacking **88**. Green trace: chromatogram of the control reaction mixture lacking MqnE. No signal at 19.2 min was observed in control reactions individually lacking SAM or dithionite (data not shown). The inset shows the structure of aminofutalosine **64** characterized before (Chapter V)].

Observation of the formation of aminofutalosine in this reaction was mechanistically intriguing. We propose that aminofutalosine, **64** might have resulted from the captodative radical intermediate **89** via the spirocylization reaction which seems to be in competition with the beta scission reaction as mentioned above (Figure 6.7) suggesting

that the spirocyclization reaction is considerably faster. A mechanistic proposal for the formation of aminofutalosine, **64** from substrate analog **88** is outlined in Figure 6.16. Addition of 5'-deoxyadenosyl radical **67** to substrate analog **88** generates the captodative radical **89** which rearranges through the spirocyclic radical **91** to give radical **92**. Radical **92** loses carbon dioxide to generate **93** which upon loss of a bromine radical followed by protonation yields aminofutalosine, **64**.

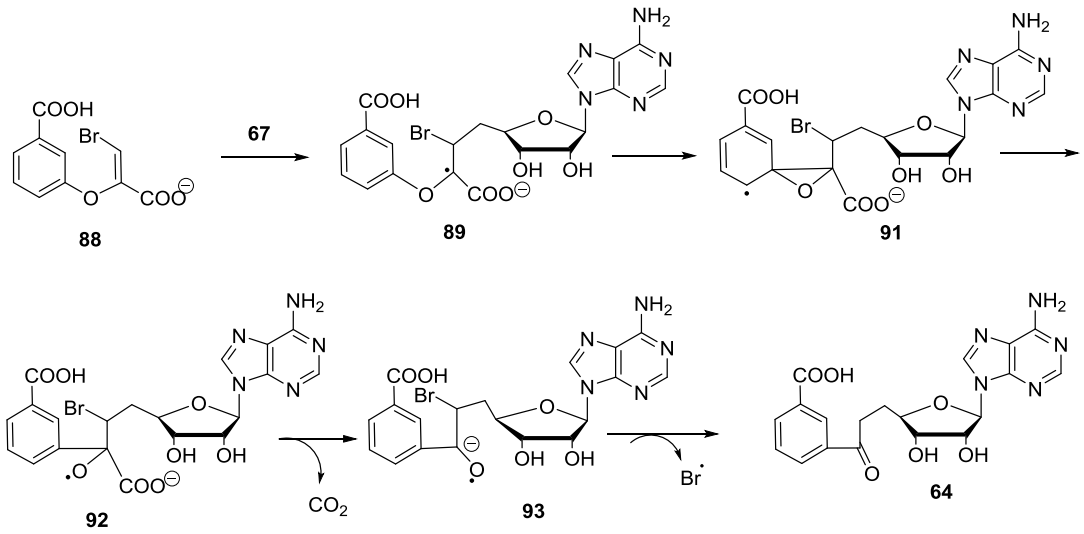



**Figure 6.16**: Mechanistic proposal for the formation of aminofutalosine **64**.

***Alternate mechanistic proposals for MqnE:*** In our study, we also considered two alternate mechanistic possibilities for the MqnE catalyzed reaction. In mechanism A (Figure 6.17), oxidation of **69** to the oxocarbenium ion **69a** would be followed by electrophilic addition to the benzene ring to generate the intermediate **69b**. However, this mechanism seems unlikely because subsequent epoxide ring opening in **69b** would generate a high energy two-electron deficient oxene intermediate. We propose addition

of an electron will convert **69b** to spirocyclic radical **70** which will then proceed as proposed (Figure 6.2). In mechanism B, a high energy aryl radical **94** is proposed to be formed from the captodative radical **69** on cleaving the C-O bond by beta scission. This is followed by radical addition from **94** onto the carbonyl group of the intermediate **95** to form the alkoxy radical **71** which loses carbon dioxide and an electron to form aminofutalosine **64** (Figure 6.17).

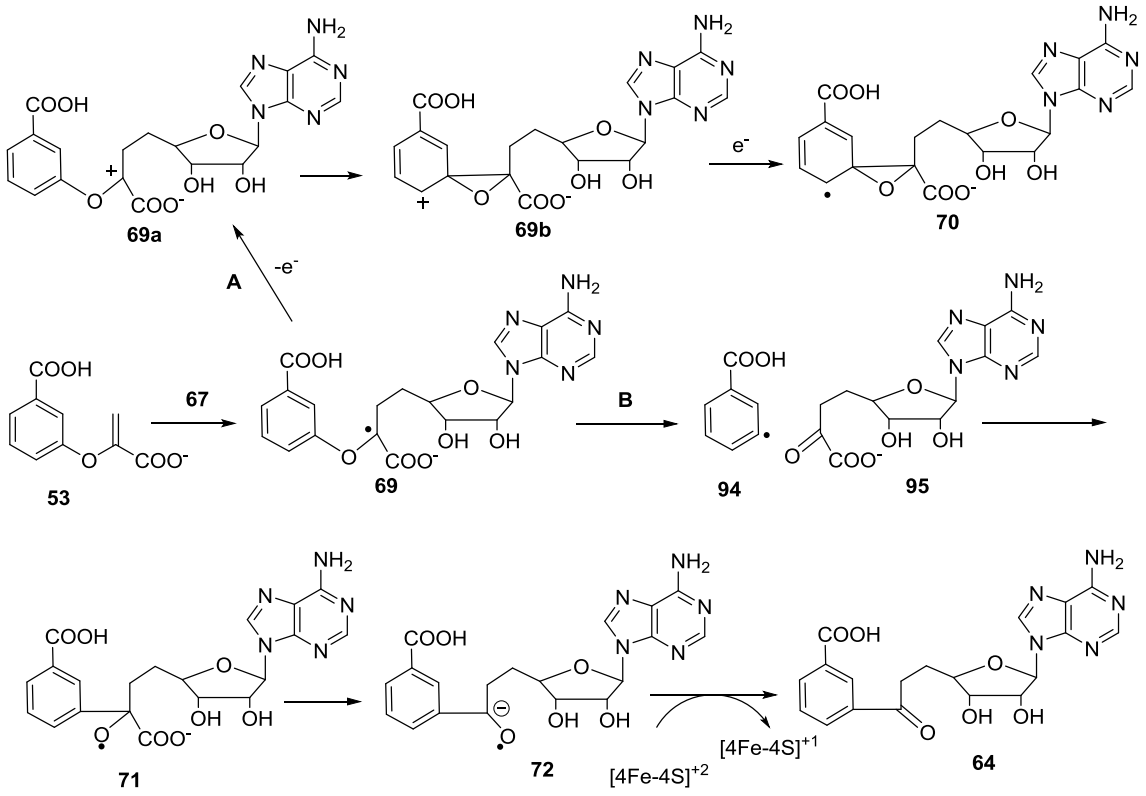



**Figure 6.17:** Alternate mechanistic proposals for the MqnE catalyzed reaction.

***Measurement of substituent effects on the radical addition to benzene ring:*** In this alternative strategy for the captodative radical trapping, observation of these two competing reactions with substrate analog **88** provided us with an opportunity to gain further mechanistic insights to determine the most probable mechanism. As it was clear from this experiment that the spirocyclization reaction is surprisingly rapid, hence using the radical trap incorporated at the double bond terminus of the substrate, substituent effects on the radical addition to the benzene ring could be determined. If the spirocyclization is occurring as proposed (Figure 6.2), then radical stabilizing substituents at the sites of high spin density of **70** would accelerate the rate of spirocyclization while substituents at positions of low spin density would have a smaller effect. Moreover, since in our system, the partitioning of **89** between spirocyclization and radical trapping is primarily influenced by the competition between bromine atom loss and the radical addition to the benzene ring, hence the binding effects arising from the substitution is likely to cancel out. Substituting the benzoic acid ring at carbons 4 and/or 6 positions with hetero atom containing substituents such as amino / hydroxyl or fluoride group in compound **88** should accelerate the spirocyclization because these groups are known to stabilize radicals[182-184]. This effect could be explained by a stabilizing two center-three electron interaction between the singly occupied orbital on carbon of the radical center and nonbonding electrons on nitrogen, oxygen or fluoride[182, 183]. Hence, in order to probe the proposed radical rearrangement through the spirocyclic intermediate, amino and fluoro substituted substrate analogs (Figure 6.18) were synthesized (Scheme B1, Appendix B). Observation of substituent effects on the

spirocyclization step will be consistent for the proposed mechanism of MqnE (Figure 6.2) and will not support the mechanism B through the intermediacy of a highly energetic aryl radical intermediate **94** (Benzene C-H bond dissociation energy = 472 kJ/mol, Figure 6.16).[185]

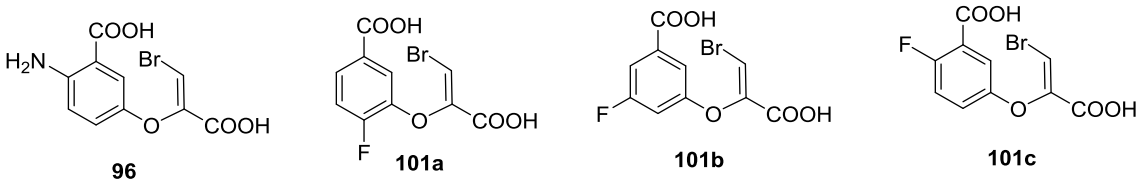



**Figure 6.18**: Substituted substrate analogs for MqnE

When substrate analog **96** was used in the MqnE catalyzed reaction, (1.5 mM substrate analog, 2mM SAM, 100 µM MqnE and excess dithionite) formation of two compounds were observed in the HPLC chromatogram which were absent in the control reactions (at 16.5 min and 19.8 min, Figures 6.19 and 6.21). Compounds **98** and **100** formed in this reaction were characterized by 1D and 2D NMR spectroscopy (Figures B13-B17 and B18- B20 Appendix B, respectivley) and mass spectrometry (Figure 6.20 and Figure 6.22). Compound **98** (Figure 6.19) resulted from the C-Br bond cleavage radical clock reaction (defined by $k_{trap}$, Figure 6.23) while compound **100** (Figure 6.21) resulted from the spirocylization reaction (defined by $k_{spiro}$, Figure 6.23). By using the relation $k_{spiro} = k_{trap}$ ([**100**] / [**98**]), we can determine the relative rate of spirocyclization and hence the effect of the $NH_2$ substituent on the radical addition step (Figure 6.23).

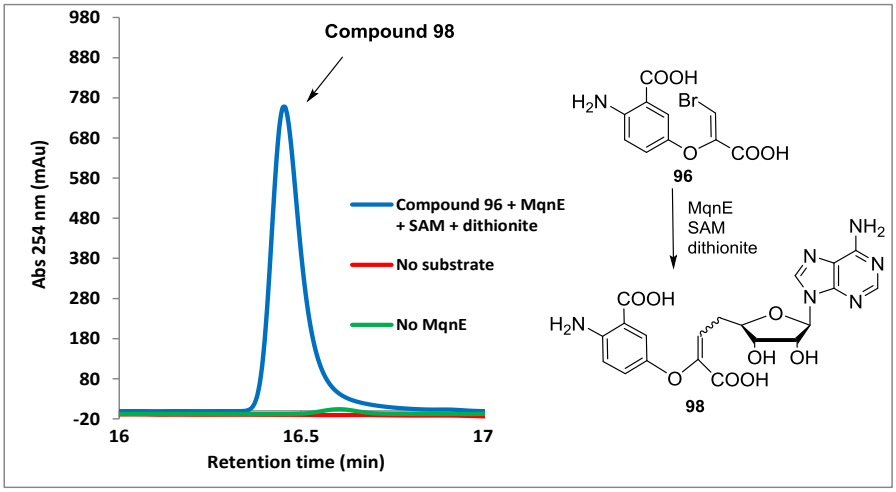



**Figure 6.19:** HPLC analysis of the MqnE reaction mixture using compound **96** (A). [Blue trace: chromatogram of the full reaction mixture (MqnE + **96**+ SAM+ dithionite). Red trace: chromatogram of the control reaction mixture lacking **96**. Green trace: chromatogram of the control reaction mixture lacking MqnE. The inset shows the structure of identified reaction product **98**].

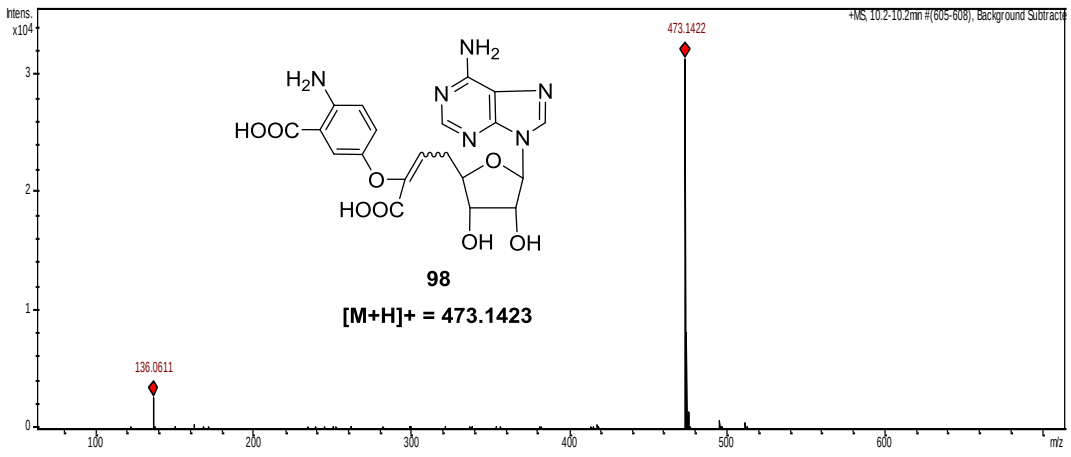



**Figure 6.20:** LC-MS (ESI, positive ion mode) analysis of compound **98**. [In the mass spectrum, m/z 473.1422, $[M+H]^+$ corresponds to compound **98** (theoretical m/z 473.1423)]

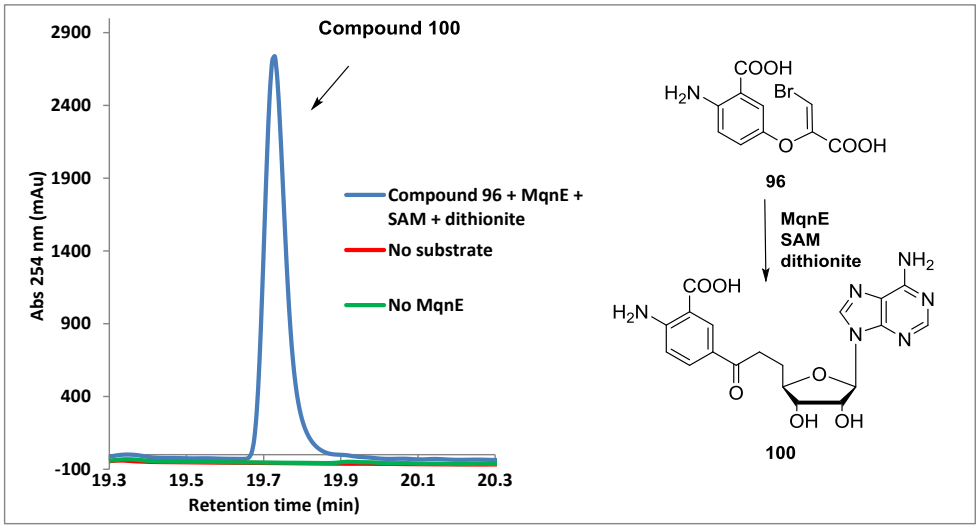



**Figure 6.21**: HPLC analysis of the MqnE reaction mixture using compound **96** (B). [Blue trace: chromatogram of the full reaction mixture (MqnE + **96**+ SAM+ dithionite). Red trace: chromatogram of the control reaction mixture lacking **96**. Green trace: chromatogram of the control reaction mixture lacking MqnE. The inset shows the structure of the identified reaction product **100**].

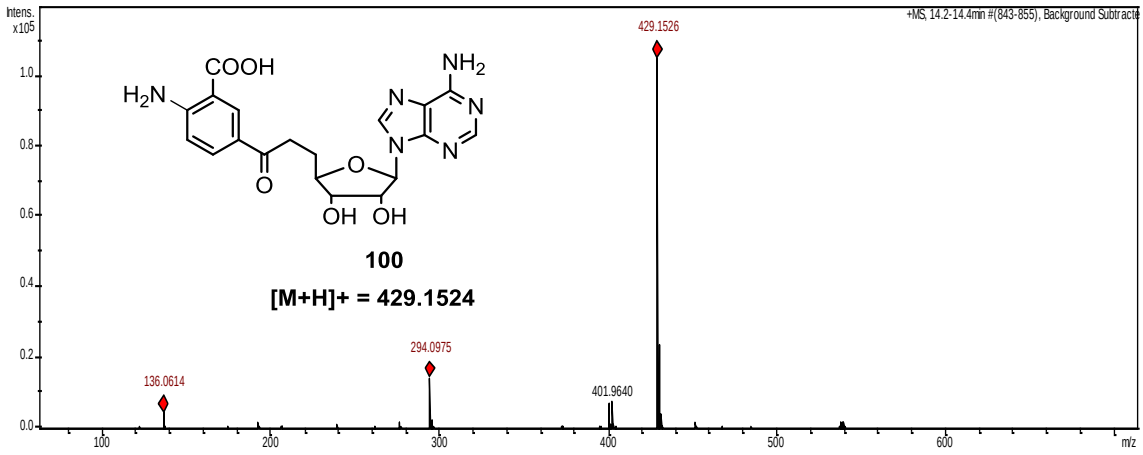



**Figure 6.22**: LC-MS (ESI, positive ion mode) analysis of compound **100**. [In the mass spectrum, m/z 429.1526, $[M+H]^+$ corresponds to compound **100** (theoretical m/z 429.1524)].

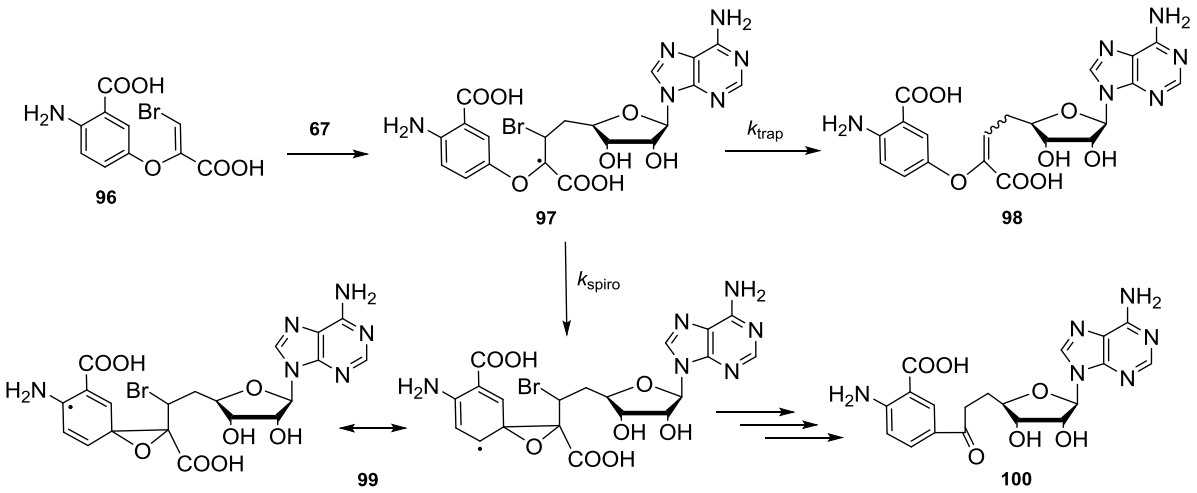



**Figure 6.23:** Substituent effects measurement using substrate analog **96.**

Compounds **98** and **100** produced in the reaction mixture were purified using reverse phase HPLC and treated separately with 5M hydrochloric acid and heated at 90°C for 3-4 hrs in order to release adenine. Each acid treated sample was neutralized and analyzed immediately using HPLC. The amount of adenine formed from each compound was determined based on a standard curve of adenine (Figure B21, Appendix B). Then the ratio [**100**]/ [**98**] (Figure 6.22) was calculated for substrate analog **96** based on the amounts of adenine formed (directly proportional to the amounts of **98** and **100**) to give an estimate of the effect of the amino substituent on the radical addition step (Table 6.1).

Formation of two compounds of similar structure as that of **98 (103a, 103b, 103c)** and **100 (105a, 105b, 105c)** were also observed (Figure 6.24) when fluoro substituted compounds **101a**, **101b** and **101c** (Figure 6.18) were used as substrates in the MqnE reaction. These compounds were analyzed by HPLC and characterized by 1D and 2D

NMR and mass spectrometry [Figures B22-B24 (**103a**), B25-B27 (**103b**), B28-B33 (**103c**), B34-B38 (**105a**), B39-B43 (**105b**), and B44-B49 (**105c**) Appendix B].

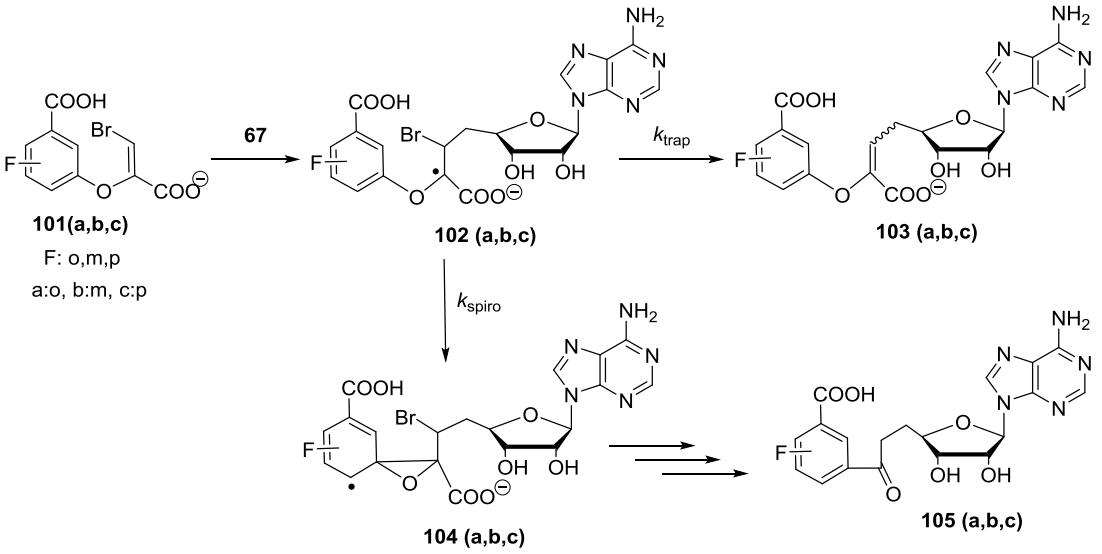



**Figure 6.24:** Substituent effects measurement using substrate analogs **101(a,b,c)**.

Similar experiments were performed for the quantification of the final compounds formed from each of the fluoro substituted substrate analogs (Figure 6.24) and the product ratio ([**105a**]/ [**103a**], [**105b**]/ [**103b**], ([**105c**]/ [**103c**]) in each case was determined using the method described above. Compounds **90** and **64** formed when substrate analog **88** was used as the substrate (Figures 6.8 and 6.15) were also quantified in a similar manner. Thereafter, the ratios of the compounds formed for the substituted substrate analogs were compared with the ratio of [**64**]/ [**90**] for the analogue **88** in order to determine a relative rate of spirocyclization and hence the effect of substituents on the step involving radical addition to the benzene ring (Figure 6.25 and Table 6.1).

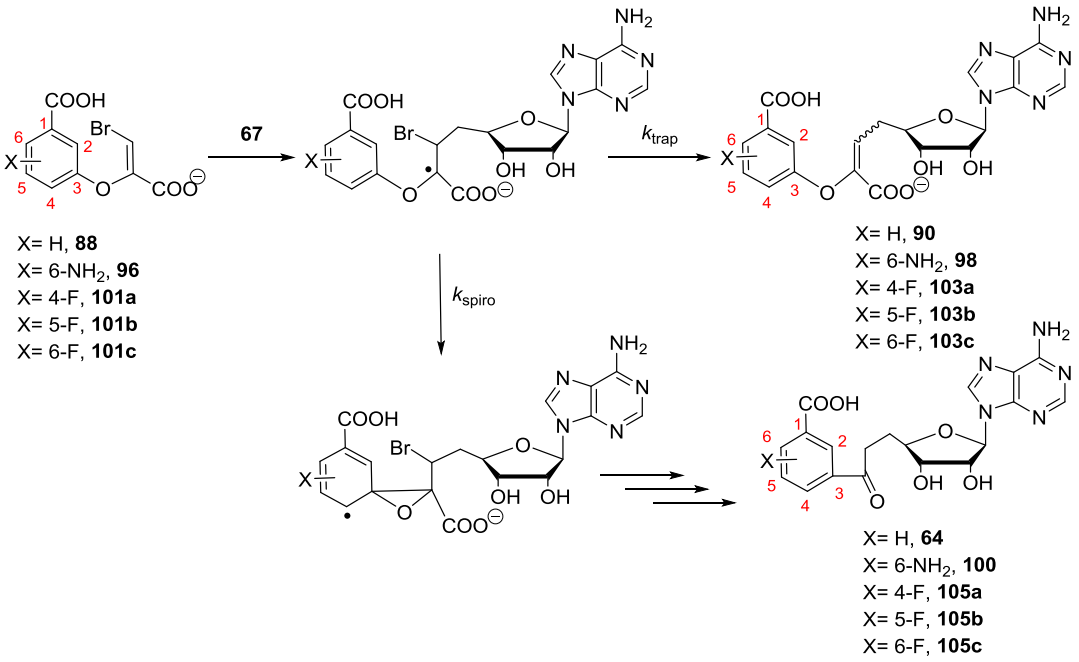



**Figure 6.25**: Substituent effects measurement in the MqnE reaction

**Table 6.1**: Ratio of the two compounds formed in the MqnE reaction

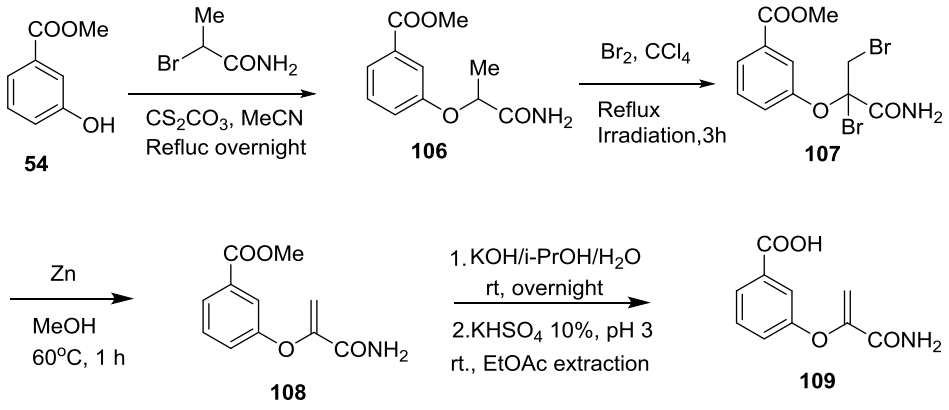
| Substrate analog | $[k_{spiro}]/[k_{trap}]$ | Relative ratio |
|---|---|---|
| **88** | 0.11 ± 0.02 | 1 |
| **96** | 0.89 ± 0.04 | 8.1 |
| **101a** | 0.45 ± 0.03 | 4.1 |
| **101b** | 0.22 ± 0.02 | 2.0 |
| **101c** | 0.42 ± 0.05 | 3.8 |

It is evident from these results that the presence of the radical stabilizing amino substituent (analogue **96**) at the site of high spin density of **99** accelerates the spirocyclization (Figure 6.23) as compared to the unsubstituted analog **88** (indicated by about 8 times larger ratio of [$k_{spiro}$]/[$k_{trap}$]). On the other hand, presence of lesser radical stabilizing fluoride substituents at 4 and/or 6 positions (analogs **101a** and **101c**) in the aromatic ring have a relatively lower relative rate of spirocyclization with respect to the unsubstituted analog **88** (about 4 times).The substrate analog with the fluoride substituent at position 5 in the ring (analog **101b**) has even smaller effect**.** These results indicated that transient radical character at 4 and/or 6 positions in the aromatic ring of the substrate during the MqnE catalytic cycle is more stabilized by the presence of hetero atom containing substituents at these positions (Figure 6.25).[182-184] The relative ratio of spirocyclization was also found to be independent of the time of the reaction. These substituent effect measurements suggested that radical addition to the benzene ring during the rearrangement of the captodative radical **69** to alkoxy radical **71** via the spirocyclic radical intermediate **70** is the most probable mechanistic route for MqnE (Figure 6.2). Observation of these substituent effects also rules out an alternate mechanistic proposal (Figure 6.17) via a highly energetic aryl radical intermediate in which such an effect would not have been observed.

***Trapping of the alkoxy radical 71***: Another important insight for the mechanistic proposal for MqnE could be obtained from characterization of the alkoxy radical **71** formed after the rearrangement of the captodative radical via the spirocyclic radical intermediate (Figure 6.2). In order to trap the alkoxy radical **71,** a substrate analog **109**

was synthesized (Scheme 6.3) in which the carboxyl group of the pyruvic side chain of the substrate **53** is replaced with an amide group. Replacingt the carboxyl group of the substrate by the amide group should block the decarboxylation reaction (Figure 6.26) because isocyanic acid resulting from removal of the amide group is much less stable than the released carbon dioxide as proposed before (Figure 6.2). This is expected to lead to the formation of compound **113** from the alkoxy radical intermediate **112** on abstraction of hydrogen from solvent or from an active site residue (Figure 6.26). Detection of compound **113** in the reaction mixture would be consistent with the intermediacy of radical **71** in the proposed mechanism of MqnE (Figure 6.2).

***Synthesis of compound 109:*** Compound **109** was synthesized using the scheme shown below.




**Scheme 6.3**: Synthetic scheme for the substrate analog **109**

To test this hypothesis, compound **109** was used as a substrate in the MqnE reaction using flavodoxin (FldA) / flavodoxin reductase (FldR) as the reductant system ([MqnE] = 100 uM, [SAM]= 2 mM, [compound **109**]= 1 mM, [Flavodoxin]= [Flavodoxin reductase]= 50 uM, [NADPH]= 2 mM). The reaction was allowed to proceed for about 8hrs, and the protein was removed by ultracentrifugation after quenching. The reaction mixture was analyzed by HPLC and LC-MS.

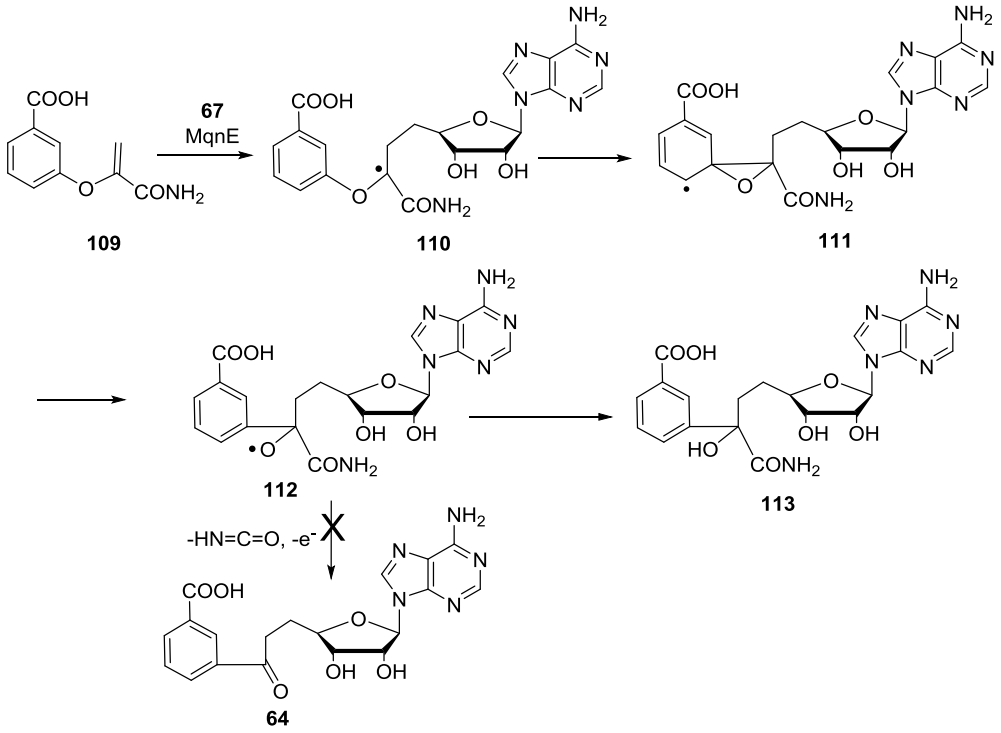



**Figure 6.26:** Trapping the alkoxy radical **112** using the amide substrate analog, **109.**

In the LC-MS chromatogram, formation of a new compound was observed in the full reaction (Figure 6.27) which was absent in the control reactions. From the MS spectrum (Figure 6.28) and MS/MS spectrum, the new compound formed was determined to be

**113** as predicted above (Figure 6.26). However, comparison with an authentic reference compound would be required to confirm its identity. Currently, synthesis of **113** is in progress. Moreover, formation of aminofutalosine **64** was not observed in the reaction mixture suggesting that isocyanic acid is not released from radical **112** (Figure 6.26). The trapping of the alkoxy radical **112** is as compound **113** is consistent with the intermediacy of radical **71** in our mechanistic proposal for MqnE.

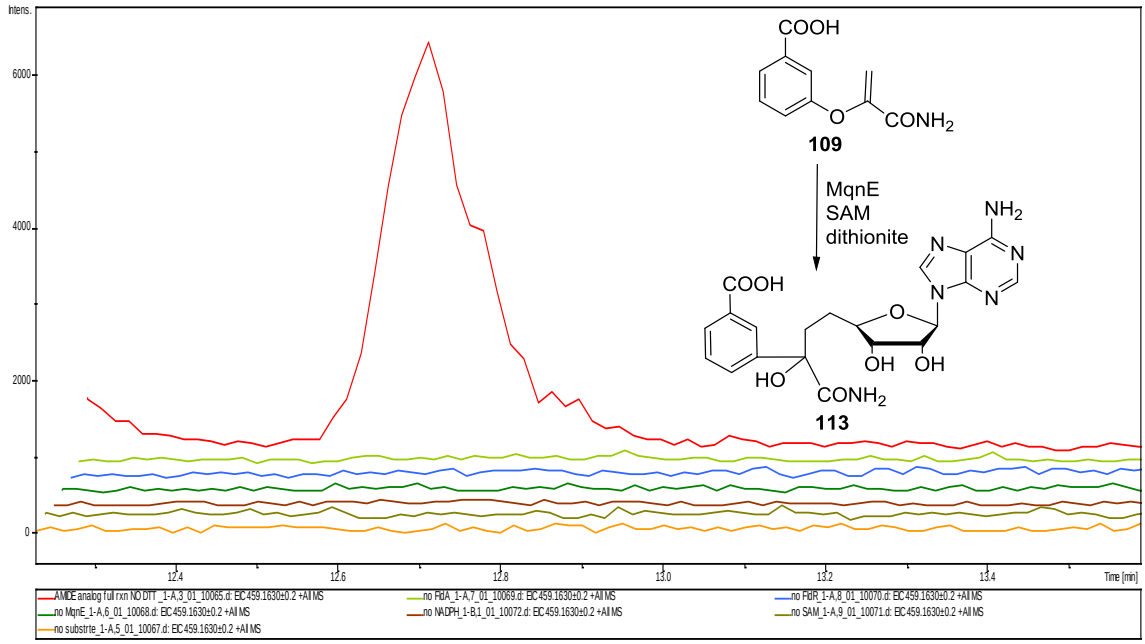



**Figure 6.27**: LCMS analysis of the MqnE reaction mixture using compound **109**. [(EIC: 459.163 m/z) Red trace: chromatogram of the full reaction mixture (MqnE + **109**+ SAM+ FldA + FldR + NADPH). No signal at 12.7 min was observed in control reactions individually lacking substrate **109**, MqnE, SAM, FldA, FldR or NADPH (other traces). The inset shows the structure of the identified reaction product **113** which was characterized by MS analysis].

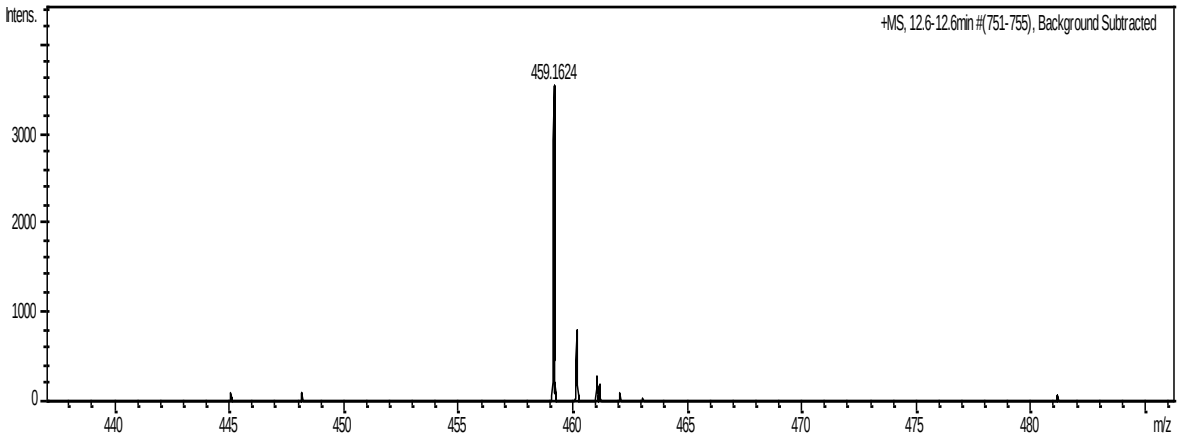
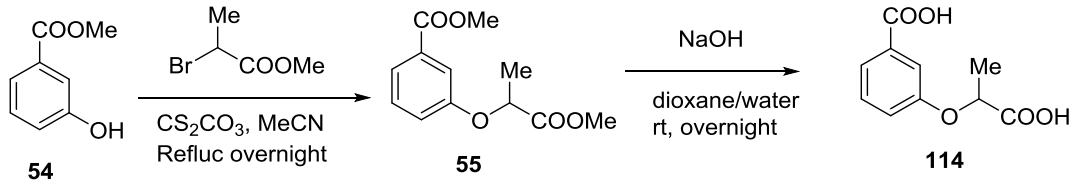



**Figure 6.28:** LC-MS (ESI, positive ion mode) analysis of compound 113. [In the mass spectrum, m/z 459.1624, $[M+H]^+$ corresponds to compound **113** (theoretical m/z 459.1630)].

So far, we have been able to trap the captodative radical **69** and the alkoxy radical **71** (Figure 6.2) using substrate analog based mechanistic strategy. Moreover, observation of substituent effects on the step involving radical addition to the aryl ring also suggest the intermediacy of the proposed spirocyclic radical intermediate **70.** A structure of MqnE with the substrate **53** or a substrate analogue bound at the active site would be required to understand whether our proposal is consistent with the stereo-electronic requirements of the spirocyclization reaction. Currently, structural studies on MqnE are in progress in collaboration with Juan Fontecilla group, IBS, Grenoble, France.

***A new reaction of MqnE***: An additional activity of MqnE was observed when **114** was used as substrate analogue. Substrate analog **114** was synthesized (Scheme 6.4) in which the vinyl double bond of the substrate **53** is replaced with a methyl group.




**Scheme 6.4**: Synthetic scheme for compound **114**

Analog **114** was used as a substrate in the MqnE reaction ([MqnE] = 100 uM, [SAM] = 2 mM, [compound **114**] = 1.5 mM, dithionite in excess). Protein was removed by ultracentrifugation and the reaction mixture was analyzed by HPLC and LCMS methods. From the analysis of the reaction mixture, formation of two new compounds was observed as suggested by the appearance of two new signals at 17.4 min and 19.4 min in the HPLC chromatogram (Figure 6.29). The compound eluting at 16 min was identified to be compound **117** (Figure 6.28) from mass spectrometric studies (Figure 6.30 and co-elution with a reference compound (Figures B50-B52, Scheme B2, Appendix B) and the compound eluting at 19.4 min was identified to be 5'-deoxyadenosine, **116** (Figure 6.28) by mass spectrometric analysis (Figure 6.31) and comparing with a reference standard (Figure B53, Appendix B).

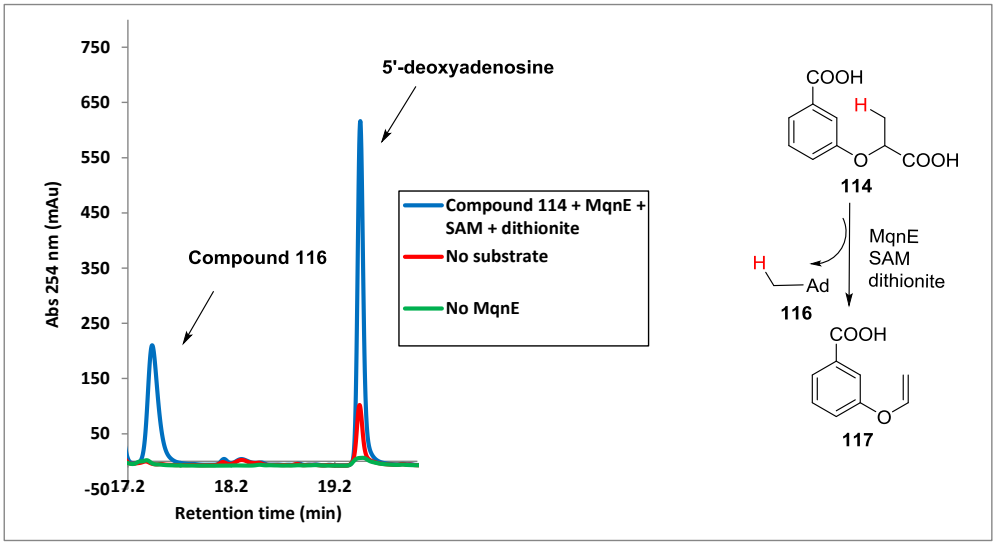



**Figure 6.29**: HPLC analysis of the MqnE reaction mixture using compound **114**. [Blue trace: chromatogram of the full reaction mixture (MqnE + **114** + SAM+ dithionite). Formation of both **117** and 5'-deoxyadenosine, **116** were observed at 17.4 min and 19.4 min respectively. Red trace: chromatogram of the control reaction mixture lacking **114**. Uncoupled formation 5'-deoxyadenosine was observed at 19.4 min. Green trace: chromatogram of the control reaction mixture lacking MqnE. No signals at 17.4 min and 19.4 min were observed in control reactions individually lacking SAM or dithionite (data not shown). The inset shows the structure of the identified reaction products **117** and 5'-deoxyadenosine **116**]

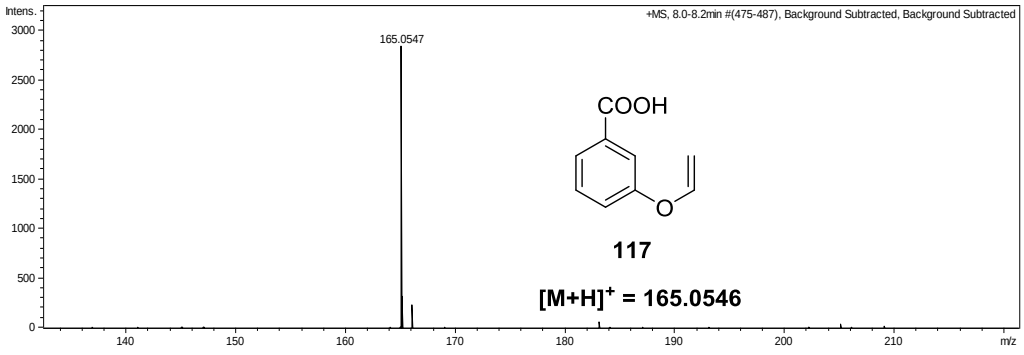



**Figure 6.30**: LC-MS (ESI, positive ion mode) analysis of compound **117**. [In the mass spectrum, m/z 165.0547, $[M+H]^+$ corresponds to compound **117** (theoretical m/z 165.0546)].

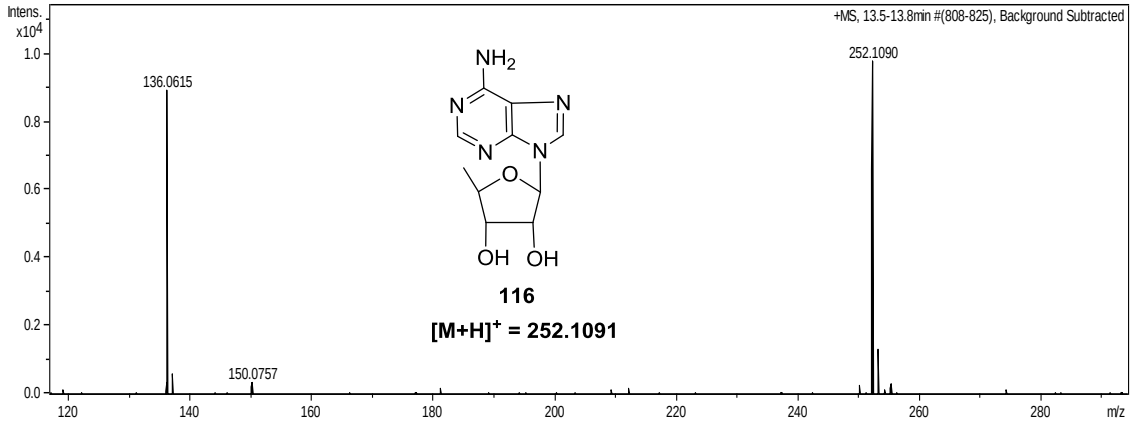



**Figure 6.31**: LC-MS (ESI, positive ion mode) analysis of compound **116**. [In the mass spectrum, m/z 252.1090, $[M+H]^+$ corresponds to 5'-deoxyadenosine, **116** (theoretical m/z 252.1091) and m/z 136.06 corresponds to adenine derived from **116** (theoretical m/z 136.06)].

Formation of these compounds can be described as shown in the mechanistic proposal (Figure 6.32). 5'-deoxyadenosyl radical, **67** abstracts a hydrogen atom from the methyl group of **114** to generate 5'-deoxyadenosine **116** and radical **115** which upon decarboxylation and oxidation gives compound **117**. This observation suggested that when the vinyl group of the substrate is absent, instead of the radical addition, the 5'-deoxyadenosyl radical **67** is able to abstract a hydrogen atom from the methyl group of the substrate, a mode of reactivity observed for almost all the reported radical SAM enzymes[163-167]. This is a new activity observed for MqnE and it probably indicates that the hydrogen abstraction capability by the reactive 5'-deoxyadenosyl radical from substrate or protein active site amino acid residue is inherent in radical SAM enzymes; however, in some instances they evolve to function differently as observed in the MqnE reaction in menaquinone biosynthesis.[147]

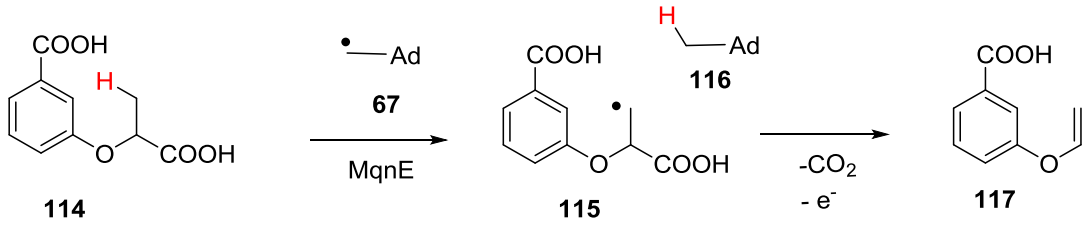



**Figure 6.32**: Mechanistic proposal for the formation of **117**.

An important aspect of the MqnE reaction is the facial selectivity of the addition of 5'-deoxyadenosyl radical **67** on the double bond of substrate **53** to generate the captodative radical **69** (Figure 6.2). Since the methyl group of **114** must be pointing at the 5'-deoxyadenosyl radical **67** for hydrogen abstraction to occur, hence identifying the correct enantiomer of **114** that is the substrate for MqnE might tell us about the facial selectivity of the adenosyl radical addition on the double bond in the native reaction. The stereochemical experiment is described in Appendix B (Figures B54- B58).

Currently studies are in progress to obtain further stereochemical information in the radical addition step in the MqnE reaction. However, experiments using substrate analogs with a constrained double bond would be required to determine the correct conformation of the bound substrate in the active site. This stereochemical information would facilitate our structural studies with MqnE, particularly in modelling studies if the resolution of the structure is low or suitable substrate/SAM bound crystal structure of MqnE is not obtained.

***Inhibition studies on MqnE:*** Several human pathogenic organisms such as *Helicobacter pylori* (causes gastric carcinoma), *Campylobacter jejuni* (causes gastroenteritis), *Chlamydia* (causes urethritis and respiratory tract infections), *Spirochetes* (causes syphilis / Lyme disease) synthesize menaquinone by the alternative pathway.[123,141] Moreover, this pathway is absent in humans and commensal intestinal bacterial species.[123] Hence the inhibitors and antibiotics targeting the enzymes on this pathway could prevent the growth of several important human pathogens without damaging the intestinal bacterial flora. In the hunt for suitable chemotherapeutics against these pathogenic organisms, previous attempts have been made to find a suitable inhibitor for this pathway.[145] It was found that branched chain fatty acids inhibit the biosynthesis of menaquinone in *H.pylori* and the most probable enzyme target was predicted to be the prenyl transferase enzyme involved in the latter steps of the biosynthetic pathway.[145]

We found that on incubation of MqnE (100 µM) with compound **83** (1.5 mM) for about an hour, the enzyme almost completely lost its activity as the activity was not recovered on addition of the native substrate **53** (1.5 mM)**,** SAM (2 mM) and dithionite (Figure 6.33). We previously reported that in MqnE catalyzed reaction, compound **83** is converted to **84** (Figure 6.3).

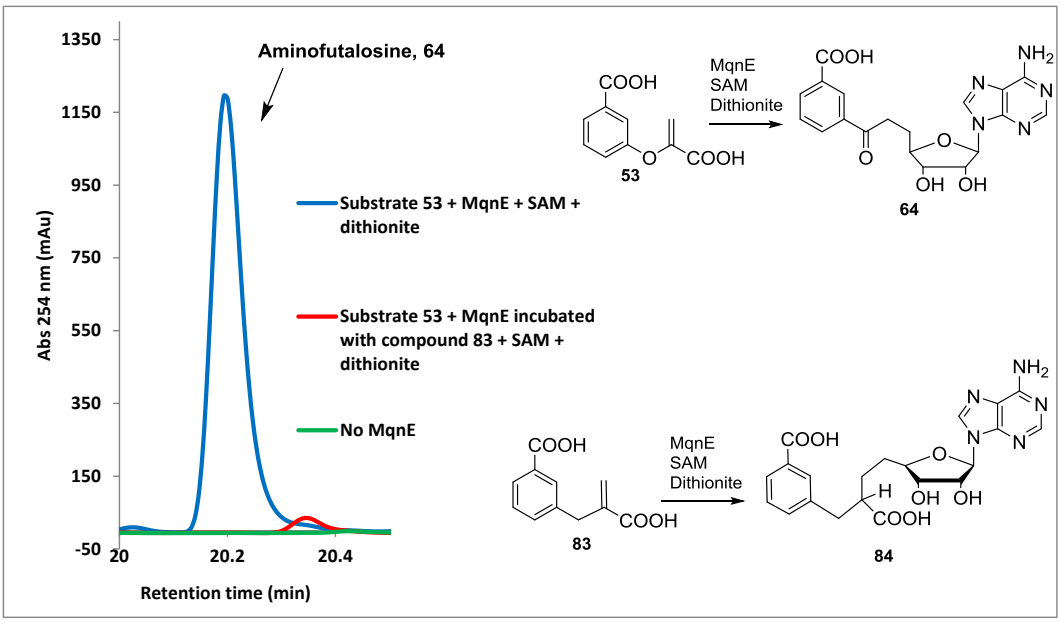



**Figure 6.33**: HPLC analysis of the MqnE reaction mixture showing inhibition. [Blue trace: chromatogram of the full reaction mixture (MqnE + **53** + SAM+ dithionite). Formation of aminofutalosine was observed at 20.2 min. Red trace: chromatogram of the reaction mixture using MqnE preincubated with compound **83,** followed by addition of substrate **53**, SAM and dithionite. Formation of aminofutalosine was not observed. MqnE retained about 5% of the activity when a 1:1 mixture of **53** and **83** was used as the substrate mixture (data not shown). Green trace: chromatogram of the control reaction mixture lacking MqnE].

We predict that this could be an irreversible mechanism based inactivation[186, 187] of the enzyme wherein tight binding of compound **84,** a reaction intermediate analogue and a bisubstrate analogue formed in the reaction coordinate of **83** may inactivate the enzyme. This renders the enzyme unable to bind the native substrate and hence loses its activity. However, a structure of MqnE with the substrate analog **83** bound at the active site would be required to clarify this assumption and provide the mechanistic insight for the observed inhibition of MqnE. No other substrate analogs (such as **88** and **114**) were

observed to exhibit such an inhibitory activity when incubated with MqnE (Figure B59, Appendix B).

A simple kinetic model could be put forward for this type of irreversible inhibition.[186] The inactivator initially forms a reversible enzyme-inhibitor (E-I) complex defined by the binding constant $K_I$, followed by the transformation of the inactivator to a different form (I') still bound to the enzyme (E-I'), which is defined by the rate constant $k_{inact}$ (Figure 6.34). This modified form of the inactivator I' binds very tightly to the enzyme and is released very slowly from the enzyme active site (Figure 6.34).

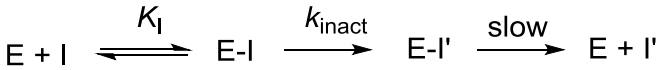



**Figure 6.34**: Kinetic model for the mechanism based inactivation.

To obtain the kinetics of inactivation, we employed a strategy previously reported for mechanism based inactivators.[186, 187] The activity of MqnE remaining after its incubation with the inhibitor **83** was determined. MqnE was incubated with different concentrations of **83** (ranging from 10 nm to 50 µM) for certain amount of time and then the enzyme was allowed to react with the native substrate, **53** in the presence of SAM and dithionite. An aliquot of the reaction mixture (50 µL) was quenched at different time intervals using 8M guanidine hydrochloride. Protein was removed by ultracentrifugation and the reaction mixture was analyzed by HPLC for the formation of aminofutalosine, **64**. The

amount of aminofutalosine formed was quantified based on a standard curve (Figure B4, Appendix B) and the log of % enzyme activity remaining was plotted against time for each of the concentrations of substrate **83**. The slopes of these lines represent $k_{app}$. Then the reciprocal of $k_{app}$ is plotted against the reciprocal of inhibitor concentration (I) to construct the Kitz-Wilson plot[186] (Figure 6.35). From the slope and intercept of this plot, $K_I$ and $k_{inact}$ were determined to be 1.07 ± 0.05 µM and 1.28± 0.08 $min^{-1}$ respectively.

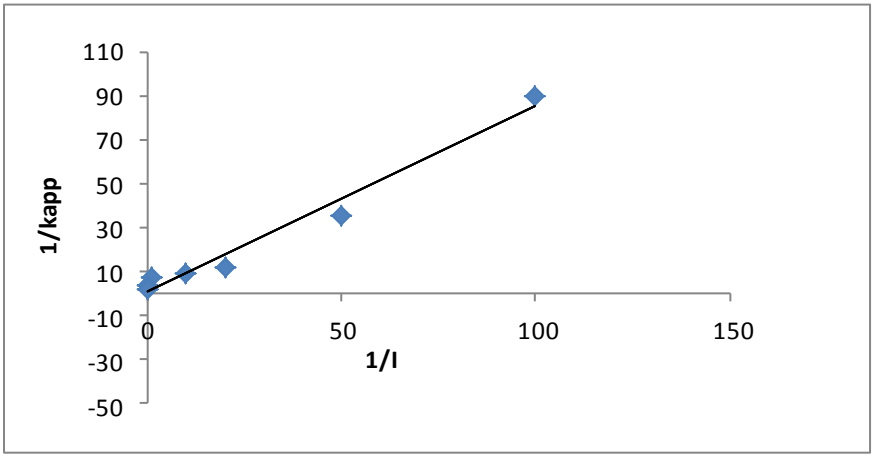



**Figure 6.35**: Kitz –Wilson plot for the inhibitor **83**

From these kinetic parameters it is evident that the inhibitor **83** binds relatively tightly in the active site of MqnE and perhaps the compound **84** formed during the catalytic cycle of MqnE binds even tighter and is released very slowly. In such a case, native substrate **53** is unable to bind and formation of aminofutalosine is not observed. However, a crystal structure of MqnE with **83** bound at the active site would be very

informative in understanding the mechanism of inactivation. Moreover, it was also found in an *in-vivo* study (by Dinesh Simkhada, Figure B55, Appendix B) that compound **83** inhibit the growth of *Dienococcus radiodurans* R1, an organism that synthesizes menaquinone by the futalosine dependent pathway with an $IC_{50}$ value of 6.4mM. This large value of $IC_{50}$ may be due to the fact that compound **83** cannot get into the cells very easily. It was also found that its growth could be recovered by the addition of menaquinone to the growth medium suggesting that this compound is a pathway specific inhibitor with the most likely target being MqnE (Figure B60, Appendix B). However, additional experiments would be required to determine the potency of this inhibitor using a pathogenic strain of bacteria that synthesizes menaquinone by the new pathway. Further attempts are in progress in understanding the inhibition mechanism of compound **83** and its derivatives using biochemical and structural studies.

**Conclusion**

In summary, in this mechanistic study, several biochemical experiments were performed to test our proposed mechanism for MqnE (Figure 6.2). The captodative radical **69** formed in the MqnE catalyzed reaction by the addition of the 5'-deoxyadenosyl radical on the double bond of the substrate was trapped using substrate analog **83** (Figure 6.5) and **88** (Figure 6.7). The support for the spirocylization step in the proposed mechanism is obtained by exploiting the competition reaction in which C-Br bond cleavage competes with reaction involving radical addition to the benzene ring using suitably designed substrate analogs **88**, **96**, **101a**, **101b**, **101c** (Figure 6.25). Observation of substituent effects (Table 6.1) on the step involving radical addition to benzene ring

suggested intermediate radical character in the ring at positions 4 and/or 6 is consistent with the proposed spirocyclic intermediate **70** (Figure 6.2).

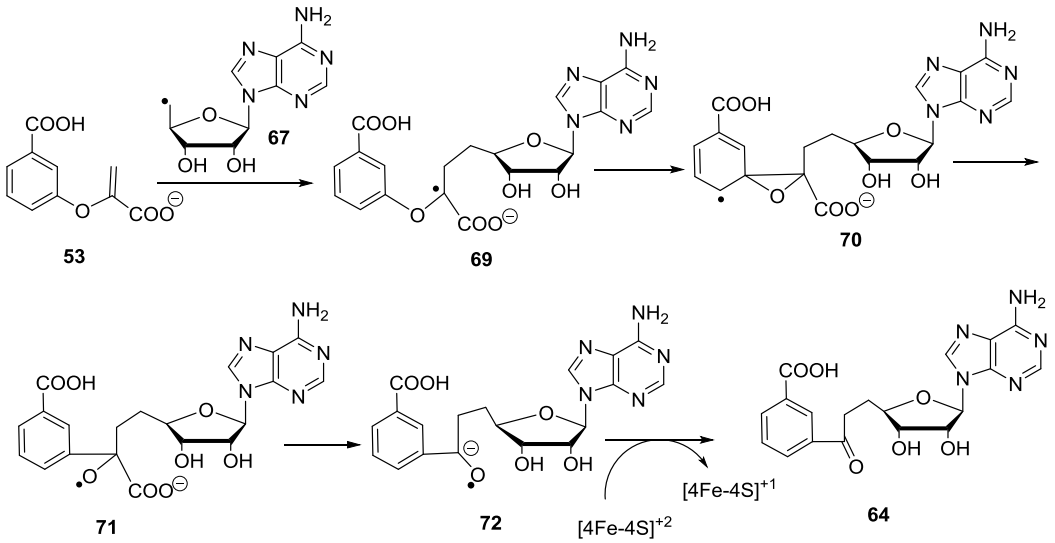



**Figure 6.2**: Mechanistic proposal for the MqnE catalyzed reaction.

Moreover, the alkoxy radical **71** proposed in the reaction mechanism after the radical rearrangement is also trapped by using the substrate analog **109** (Figure 6.26). Compound **83** was found to be an irreversible mechanism based inactivator of MqnE (Figure 6.33). The experimental results from the substrate analogs based mechanistic study were consistent with our proposed mechanism for this remarkable radical SAM enzyme reaction. In this regard, MqnE represents a unique motif in radical SAM enzymology in which the reactive 5'-deoxyadeonsyl radical is used as a substrate partner to form the C-C bond in the product aminofutalosine via a spirocyclic radical rearrangement mechanism.

**Experimental procedures**

***General methods and materials***: All chemicals were purchased from sigma Aldrich (St. Louis, MO) unless otherwise stated and used without further purification. LB growth medium was obtained from Difco. Kanamycin, ampicillin, chloramphenicol antibiotics and isopropyl-β-D-thiogalactopyranoside (IPTG) were purchased from LabScientific Inc. Ni-Nitrilotriacetic acid (NTA) resin was from GE life sciences. *E. coli* M15 (pREP4) competent cells were obtained from Qiagen. Molecular biology manipulations were carried out as previously reported.[64] Centrifugal filters were obtained from Pall Life Sciences. Analytical HPLC (Agilent) was carried out using a Supelcosil LC-18- T HPLC reverse phase column (15 cm x 4.6 mm, 3 µM particle size). HPLC grade solvents were obtained from Fisher Scientific.

***Reverse phase HPLC method***: HPLC analysis using a linear gradient, at a flow rate of 1 mL/min was used with absorbance detection at 254 nm. Solvent A is water, solvent B is 100 mM $K_2HPO_4$, pH 6.6 (in some cases 5 mM ammonium acetate, pH 6.6 was used), and solvent C is methanol: 0 min, 100% B; 5 min, 100 % B; 12 min, 48 % A, 40 % B, 12 % C; 14 min, 48 % A, 30 % B, 22 % C; 18 min, 32 % A, 10% B, 58% C; 20 min, 100 % B; 25 min 100 % B.

***Chiral HPLC method***: The chiral column used was Chiralcel OD HPLC column in which the bulk stationary phase consisted of 20 µM silica support onto which the polymeric chiral selector, cellulose tris (3,5-dimethylphenylcarbamate) has been physically coated. The mixture of racemic methylated derivatives (**44**) were extracted to

ethylether after the diazomethane treatment and evaporated into dryness before dissolving in isopropanol to run on the chiral HPLC column. The mobile phase used was 5% isopropanol and 95% hexane (isocratic solution) for a time period of about 25 min.

***LC-MS method:*** Samples were analyzed by reverse-phase HPLC on an Agilent 1200 HPLC system equipped with a diode array UV-Vis detector and a thermostatted autosampler (10 °C). The stationary phase was Supelcosil LC-18-T (15 cm × 3 mm, 3 µm particles size), maintained at 22 °C. The LC eluent consisted of a gradient of 75% methanol in 25% water with 5 mM ammonium acetate buffer, pH 6.6 (0.4 mL/min flow rate). The percentages of ammonium acetate buffer (*B*) and methanol (*M*) at time *t* varied according to the following scheme: {*t*,*M*,*B*}: {0,0,100}, {2,0,100}, {10,35,65}, {15,57.5,42.5}, {18,100,0}, {23,100,0}; {26,0,100}; {33,0,100}. Compounds were detected using absorbance at 254 nm. UV-Vis spectra (190 – 640 nm) were also collected and evaluated. Mass data were collected using an in-line Bruker Daltonics micrOTOF-Q II ESI-Qq-TOF mass spectrometer (HyStar) in positive or negative ion mode as indicated.

***NMR analysis***: All NMR analyses were performed using a Bruker Avance III 500 MHz instrument with H−C−N Cryoprobe ($^{1}$H 500 MHz, $^{13}$C 125 MHz) in 3 mm Wilmad labglass 535-PP-7 high precision NMR tubes at room temperature (22 °C) using standard instrument settings and processing parameters.

***Expression and purification of MqnE ($His_6$-TTHA0804):*** Heterologous expression in and purification from *E. coli* M15 (pREP4) cells (Qiagen) was performed as previously

described[147] except that the expression construct also contained *p*SUF[188] (co-transformed with pQE-30 vector containing the *ttha0804* gene) and 34 µg/mL of chloramphenicol was also added along with 40 µg/mL kanamycin and 100 µg/mL ampicillin during growth. The protein expression was initiated by adding IPTG to a final concentration of 100 µM along with the supplementation of ferrous ammonium sulfate (100 mg/L) and l-cysteine (100 mg/L). Protein purification was carried out in an anaerobic chamber (Coy Laboratories) containing 5% hydrogen and 95% nitrogen. Purified protein was buffered exchanged into the desalting Buffer (100 mM potassium phosphate, pH 7.5 at 25 °C, 2 mM DTT, 30% glycerol) via Econo-Pac 10DG size exclusion chromatography (Bio-Rad). Proteins were submerged in liquid nitrogen until use.

***MqnE activity assays***: Reaction mixtures were set up as previously described.[147] Briefly, the mixtures (100 µL) consisted of 50 µL solutions of 100 mM Tris-HCl buffer (pH 7.5), 1.5 mM substrate analog, 2mM SAM, and 100 µM of MqnE (excess sodium dithionite). In some cases, biological reducing system (50 µM of FldA, 50 µM FldR and 2 mM NADPH) was also used in place of dithionite. The reaction was allowed to proceed for about 8-10 hrs, quenched with 8M guanidine hydrochloride and protein was removed by ultracentrifugation using Pall centrifugal devices before analyzing the reaction mixtures by HPLC and LCMS methods.

***Purification of the compounds formed in the substrate analogs reaction***: Each reaction mixture (150 µL) contained chemically synthesized substrate analog, **88** (2 mM), S-adenosylmethionine, **66** (2 mM), purified MqnE (200 µM) and excess sodium dithionite

in 100 mM Tris-HCl buffer, pH 7.5. The reaction was allowed to proceed anaerobically for about 10 hrs. The reaction was repeated multiple times to give sufficient product for analysis. The reactions were quenched with an equal volume of 8M guanidine hydrochloride and the protein was removed by ultrafiltration using 10 kDa centrifugal filters (Pall Life Sciences). Compounds **90** and **64** was purified from the reaction mixture by HPLC (eluted at 17.7 minutes and 19.2 minutes) using the HPLC method mentioned previously. The compound was buffer exchanged into 5mM ammonium acetate by HPLC. The purified product was concentrated using a speedvac concentrator (Thermo Scientific), dissolved in $D_2O$ and characterized by $^{1}H$ and $^{13}C$ NMR (Appendix B). Same procedure is applied for the purification of the compounds formed in the reactions using substrate analogs **96**, **101a, 101b** and **101c**.

***Quantification of the compounds formed in the substrate analogs reaction***: The compounds **90** and **64** formed in the reaction of the analog **88** were purified out from the reaction mixtures using HPLC and each sample were treated separately with 5M HCl and heated at 90°C for about 3 hrs. Upon this treatment, adenine was formed as a result of the breaking of the N-glycosyl bond which is common in both the structures. Hence, the amounts of adenine formed are directly proportional to the amounts of these compounds. The acid treated samples were neutralized and injected into the HPLC and the amounts of adenine formed from each compound were determined based on a standard curve of adenine (Appendix A). The quantification was repeated three times. Same procedure is applied for other substrate analogs **96**, **101a, 101b** and **101c**.

# CHAPTER VII

# SUMMARY AND CONCLUSIONS

## Pseudouridine catabolism

***Summary:*** The pseudouridine (Ψ) catabolic pathway occurs in two steps: first pseudouridine is phosphorylated at the 5'-hydroxyl group by a kinase (YeiC) to form pseudouridine monophosphate (ΨMP) and then the latter is cleaved to form uracil and ribose 5-phosphate by the ΨMP glycosidase (YeiN). Structural studies performed on the ΨMP glycosidase provides four snapshots of the reaction coordinate which gave key mechanistic insights into the enzymatic hydrolysis of C-glycosides. Most importantly, the structural studies suggested that the reaction involves an intermediate imine with Lys166 in the enzyme active site. The intact protein mass spectrometric studies, followed by the analysis of the trypsin digested fragments led to the confirmation of this finding. These studies allowed us to propose a reasonable mechanism for this glycosidase. The steady state kinetics analysis of the wild type YeiN and the active site mutants helped us to assign roles for specific amino acid residues that might be involved during catalysis.

***Conclusions:*** The biochemical and structural studies performed on the ΨMP glycosidase (YeiN) catalyzed transformation demonstrates that it utilizes a new kind of mechanism involving ribose ring opening and subsequent covalent linkage between C1′ of sugar moiety of ΨMP and an active site lysine (K166), thus setting the stage for a facile C-glycosyl bond fragmentation reaction by a novel retro aldol type reaction mechanism.

Moreover, this enzyme also requires divalent metal ion cofactors for its function. In contrast, most glycosidases, characterized till date, utilize a dissociative mechanism to reversibly cleave C−N and C−O glycosidic bonds through the intermediacy of an oxocarbonium ion. During such a mechanism, the ring opening of the sugar moiety has never been observed and there is also no requirement for metal ion cofactors. In all these respects, ΨMP glycosidase deviates from the reported glycosidase chemistry so far. To the best our knowledge, this is the first mechanistically characterized example of a glycosidase that catalyzes the cleavage of a C-C glycosidic bond and probably many more examples of such kind are likely to emerge in the future providing new insights into their mechanisms and adding to the repertoire of rich glycosidase chemistry.

**Menaquinone Biosynthesis**

***Summary:*** Menaquinone (vitamin K) is a lipid-soluble molecule that shuttles electrons between membrane-bound redox enzymes in the bacterial respiratory chain. It is also a very important vitamin in mammals where it participates as a cosubstrate in the carboxylation of glutamic acid residues in various proteins involved in blood coagulation and bone morphogenesis. Besides the classical pathway in *Escherichia coli*, a new pathway for menaquinone biosynthesis was reported in *Streptomyces coelicolor* A3 (2). Except for the last two steps, all the other biosynthetic steps are completely different in this pathway from the classical pathway. We have recently reconstituted the early steps of this new pathway *in vitro* and revised the biosynthetic steps for the formation of futalosine, the first intermediate on this pathway. The pathway begins with MqnA catalyzing the transformation of chorismate into 3-[(1-carboxyvinyl) oxy]-benzoic acid

and then the newly discovered radical SAM enzyme MqnE catalyzes the transformation of the latter into aminofutalosine using radical mediated chemistry.

We have proposed a mechanism for chorismate dehydratase, MqnA based on structural and biochemical studies. We believe that the carboxyl group of the pyruvate side chain of substrate chorismate is involved in this transformation. We have proposed a mechanism for the radical SAM enzyme MqnE catalyzed reaction. It involves addition of the reactive 5'-deoxyadenosyl radical on the enol ether double bond of the substrate to generate the captodative radical which further undergoes rearrangement to give the alkoxy radical via the spirocyclic radical intermediate. The alkoxy radical then loses carbon dioxide and an electron back to the [4Fe-4S] cluster to give aminofutalosine. In our mechanistic study, we have obtained several experimental results which are consistent with our proposed mechanism. We have trapped the captodative radical formed in the MqnE catalyzed reaction using substrate analog based chemistry. Moreover, evidence for the spirocylization step in the proposed mechanism is also obtained by exploiting the competition reaction in which carbon-halogen bond cleavage competes with the reaction involving radical addition to the benzene ring using a suitably designed substrate analog. Observation of substituent effects using substituted substrate analogs suggested intermediate radical character in the aryl ring at positions 4 and/or 6 providing support for the intermediacy of a spirocyclic radical. Moreover, the alkoxy radical proposed in the reaction mechanism is also trapped by using the substrate analog based strategy. Currently, studies are underway to characterize the spirocyclic radical by electron paramagnetic resonance (EPR) spectroscopy (in collaboration with

David Britt lab, UC Davis) using N-oxide substrate analogs. Works are also in progress to determine the facial selectivity of radical addition. A probable mechanism based inactivator of MqnE is also uncovered during this study. Structural studies on MqnE are in progress (in collaboration with Juan Fontecilla lab, IBS, Grenoble, France) to visualize the active site architecture of this enzyme and relative spatial orientation of the native substrate or analogs and SAM in the active site to get further insights into the mechanism.

***Conclusions*:** In the study reported here, a new radical SAM enzyme is discovered to be involved in the alternative menaquinone biosynthesis pathway which led us to clarify the early steps of this pathway in contrast to the previous proposal. The mechanistic investigation on MqnE provided great insights into this remarkable radical SAM enzyme reaction. In this regard, MqnE represents a unique motif in radical SAM enzymology in which the reactive 5'-deoxyadeonsyl radical is used as a substrate partner to form the carbon-carbon bond in the product aminofutalosine. This is a new reaction motif in biological chemistry that employs a radical transposition mechanism via a spirocyclic radical intermediate during its catalytic cycle.

## REFERENCES


1. Charette, M.; Gray, M. W. *IUBMB Life* **2000**, *49*, 341.
2. Hoang, C.; Ferré-D'Amaré, A. R. Cell **2001**, *107*, 929.
3. Uliel, S.; Liang, X.-h.; Unger, R.; Michaeli, S. *Int. J. Parasitol*. **2004**, *34*, 445.
4. Auffinger, P.; Westhof, E.: Effects of Pseudouridylation on tRNA Hydration and Dynamics: a Theoretical Approach. In *Modifications and Editing of RNA* (Grosjean, H.; Benne, R.; eds.) **1998**, pp. 103, ASM Press, Washington, DC.
5. Davis, D. R.; Poulter, C. D. *Biochemistry*, **1991**, 30, 4223.
6. Davis, D. R.; Veltri, C. A.; Nielsen, L. *J. Biomol. Struct. Dyn*. **1998**, 15, 1121.
7. Yarian, C. S.; Basti, M. M.; Cain, R. J.; Ansari, G.; Guenther, R. H.; Sochacka. E.; Czerwinska, G.; Malkiewics, A.; Agris, P. F. *Nucleic Acids Res*. **1999**, 27, 3543.
8. Durant, P. C.; Davis, D. R. *J. Mol. Biol*. **1999**, 285, 115.
9. Davis, D. R. *Nucleic Acids. Res.* **1995**, 23, 5020.
10. Dao, V.; Guenther, R.; Malkiewicz, A.; Nawrot, B.; Sochacka. E.; Kraszewski, A.; Jankowska, J.; Everett, K.; Agris, P. F. *Proc. Natl. Acad. Sci. USA*, **1994**, 91, 2125.
11. Harrington, K. M.; Nazarenko, I. A.; Dix, D. B.; Thompson, R. C.; Uhlenbeck, O. C. *Biochemistry*, **1993**, 32, 7617.
12. Ofengand, J.; Fournier, M. J.: The Pseudouridine Residues of rRNA: Number, Location, Biosynthesis, and Function. In *Modifications and Editing of RNA*

(Grosjean, H.; Benne, R.; eds.) **1998**, pp. 229, ASM Press, Washington, DC.

13. Cunningham, P. R.; Richard, R. B.; Weitzmann, C. J.; Nurse, K.; Ofengand, J. *Biochimie* **1991**, 73, 789.

14. Ofengand, J.; Bakin, A.; Wrzesinski, J.; Nurse, K.; Lane, B. G. *Biochem. Cell Biol.* **1995**, 73, 915.

15. Lane, B. G.; Ofengand, J.; Gray, M. W. *Biochimie* **1995**, 77, 7.

16. Raychaudhuri, S.; Conrad, J.; Hall, B. G.; Ofengand, J. *RNA 4* **1998**, 1407.

17. Philips, B.; Billin, A. N.; Cadwell, C.; Buchholz, R.; Erickson, C.; Merriam, J. R.; Carbon, J.; Poole, S. J. *Mol. Gen. Genet*. **1998**, 260, 20.

18. Lafontaine, D. L. J.; Bousquet-Antonelli, C.; Henry, Y.; Caizergues-Ferrer, M.; Tollervey, D. *Genes Dev.* **1998**, 12, 527.

19. Heiss, N. S.; Knight, S. W.; Vulliamy, T. J.; Klauck S. M.; Wiemann, S.; Mason, P. J.; Poustka, A.; Dokal, I. *Nat. Genet*. **1998**, 19, 32.

20. Zebarjadian, Y.; King, T.; Fournier, M. J.; Clarke, L.; Carbon, *J. Mol. Cell. Biol*. **1999**, 19, 7461.

21. Giordano, E.; Peluso, I.; Senger, S.; Furia, M. *J. Cell Biol*. **1999**, 144, 1123.

22. Knight, S. W.; Heiss, N. S.; Vulliamy, T. J.; Greschner, S.; Stavrides, G.; Pai, G. S.; Lestringant, G.; Varma, N.; Mason, P. J.; Dokal, I.; Poustka, A. *Am. J. Hum. Genet.* **1999**, 65, 50.

23. Grosjean, H.; Benne R. *Modifications and Editing of RNA*, ASM Press, Washington, D.C. **1998**.

24. Konnin, E. V. *Nucleic Acids Res*. **1996**, 24, 2411.

25. Gustafsson, C. Reid, R.; Greene, P. J.; Santi, D. V. *Nucleic Acids Res*. **1996**, 24, 3756.

26. Kaya, Y.; Ofengand, J. *RNA 9*, **2003**, 711.

27. Mueller, E. G.; Ferre-D Amare, A. R.: Pseudouridine Formation, The Most Common Transglycosylation in RNA. *In DNA and RNA Modification Enzymes: Structure, Mechanism, Function and Evolution* (Grosjean, H. Ed.) Landes Bioscience: Austin, TX **2009**, pp. 363.

28. Gu, X. R.; Liu, Y. Q.; Santi, D. V. *Proc. Natl. Acad. Sci. USA*, **1999**, 96, 14270.

29. Huang, L. X.; Pookanjanatavip, M.; Gu, X. G.; Santi, D. V. *Biochemistry*, **1998**, 37, 344.

30. Miracco, E. J.; Mueller, E. G. *J. Am. Chem. Soc.* **2011**, 133, 11826.

31. Furmanek, A.; Hofsteenge, J. *Acta Biochim. Pol.* **2000**, 47, 781.

32. Vassiliadis, G.; Peduzzi, J.; Zirah, S.; Thomas, X.; Rebuffat, S.; Destoumieux-Garzon, D. *Antimicrob. Agents Chemother.* **2007**, 51, 3546.

33. Ward, D. C.; Cerami, A.; Reich, E.; Acs, G.; Altwerger, L. *J. Biol. Chem*. **1969**, 244, 3243.

34. Breitman, T. R. *J. Bacterol.* **1968**, 96, 1873.

35. Solomon, L. R.; Breitman, T. R. *Biochem. Biophys. Res. Commun*. **1971**, 44, 299.

36. Breitman, T. R. *J. Bacteriol*. **1970**, 103, 263.

37. Preumont, A.; Snoussi, K.; Stroobant, V.; Collet, J. F.; Van Schaftingen, E. V. *J. Biol. Chem.* **2008**, *283*, 25238.

38. Feng, B.; Zheng, M. H.; Zheng, Y. F.; Lu, A. G.; Li, J. W.; Wang, M. L.; Ma, J.

J.; Xu, G. W.; Liu, B. Y.; Zhu, Z. G. *J. Gastroenterol. Hepatol.* **2005**, *20*, 1913.

39. Bililign, T.; Griffith, B. R.; Thorson, J. S. *Nat. Prod. Rep*. **2005,** 22, 742.

40. Gopaul, D. N.; Meyer, S. L.; Degano, M.; Sacchettini, J. C.; and Schramm, V. L. *Biochemistry* **1996**, 35, 5963.

41. Duerre, J. A. *J. Biol. Chem.* **1962**, 237, 3737.

42. Hurwitz, J.; Heppel, L. A.; Horecker, B. L. *J. Biol. Chem.* **1957**, 226, 525.

43. Pugmire, M. J.; Ealick, S. E. *Structure 6* **1998**, 1467.

44. Flaks, J. G.; Erwin, M. J.; Buchanan, J. M. *J. Biol. Chem.* **1957**, 228, 201.

45. Olsen, A. S.; Milman, G. *J. Biol. Chem.* **1974**, 249, 4030.

46. Short, S. A.; Armstrong, S. R.; Ealick, S. E.; Porter, D. J. T. *J. Biol. Chem*. **1996**, 271, 4978.

47. Fromme, J. C.; Verdine, G. L. *EMBO J.* **2003**, 22, 3461.

48. Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. *Annu. Rev. Biochem.* **2008**, 77, 521.

49. Zechel, D. L.; Withers, S. G. *Acc. Chem. Res.* **2000**, *33*, 11.

50. Kharel, M. K., Pahari, P., Shepherd, M. D., Tibrewal, N., Nybo, S. E., Shaaban, K. A., and Rohr, J. *Nat. Prod. Rep*. **2012**, 29, 264.

51. Nolan, E. M.; Fischbach, M. A.; Koglin, A.; Walsh, C. T. *J. Am. Chem. Soc.* **2007**, 129, 14336.

52. Fischbach, M. A.; Lin, H.; Liu, D. R.; Walsh, C. T. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, 102, 571.

53. Brazier-Hicks, M.; Evans, K. M.; Gershater, M. C.; Puschmann, H.; Steel, P. G.;

Edwards, R. *J. Biol. Chem.* **2009**, 284, 17926.

54. Siitonen, V.; Claesson, M.; Patrikainen, P.; Aromaa, M.; Maentsaelae, P.; Schneider, G.; Metsae-Ketelae, M. *ChemBioChem* **2012,** 13, 120.

55. Oja, T.; Klika, K. D.; Appassamy, L.; Sinkkonen, J.; Mantsala, P.; Niemi, J.; Metsa-Ketela, M. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, 109, 6024.

56. Dumitru, R. V.; Ragsdale, S. W. *J. Biol. Chem*. **2004**, 279, 39389.

57. White, R. H. *Biochemistry* **2011**, 50, 6041.

58. Mittler, M.; Bechthold, A.; Schulz, G. E. *J. Mol. Biol.* **2007**, 372, 67.

59. Chang, A.; Singh, S.; Phillips, G. N., Jr.; Thorson, J. S. *Curr. Opin. Biotechnol.* **2011**, 22, 800.

60. Levin, I.; Miller, M. D.; Schwarzenbacher, R.; McMullan, D.; Abdubek, P.; Ambing, E.; Biorac, T.; Cambell, J.; Canaves, J. M.; Chiu, H. J.; Deacon, A. M.; DiDonato, M.; Elsliger, M. A.; Godzik, A.; Grittini, C.; Grzechnik, S. K.; Hale, J.; Hampton, E.; Han, G. W.; Haugen, J.; Hornsby, M.; Jaroszewski, L.; Karlak, C.; Klock, H. E.; Koesema, E.; Kreusch, A.; Kuhn, P.; Lesley, S. A.; Morse, A.; Moy, K.; Nigoghossian, E.; Ouyang, J., Page, R., Quijano, K., Reyes, R., Robb, A., Sims, E.; Spraggon, G.; Stevens, R. C.; van den Bedem, H.; Velasquez J.; Vincent, J.; Wang, X.; West, B.; Wolf, G.; Xu, Q.; Zagnitko, O.; Hodgson, K. O.; Wooley, J.; Wilson, I. A. *Proteins* **2005**, 59, 864.

61. Huang, S.; Mahanta, N.; Begley, T. P.; Ealick, S. E. *Biochemistry*, **2012**, 51, 9245.

62. Rebelo, J.; Auerbach, G.; Bader, G.; Bracher, A.; Nar, H.; Hosl, C.; Schramek,

N.; Kaiser, J.; Bacher, A.; Huber, R.; Fischer, M. *J. Mol. Biol.* **2003**, 326, 503.

63. Chung, S. J.; Verdine, G. L. *Chem. Biol.* **2004**, 11, 1643.

64. Ausubel, F. M.; Brent, F. *Current Protocols in Molecular Biology*, John Wiley and Sons, New York. **1987**.

65. Vita, A.; Huang, C. Y.; Magni, G. *Arch. Biochem. Biophys.* **1983**, 226, 687.

66. Otwinowski, Z.; Minor, W. *Methods Enzymol*. **1997**, 276, 307.

67. Vagin, A.; Teplyakov, A. *Acta Crystallogr*.D 2000, D56, 1622.

68. Stein, N. *J. Appl. Crystallogr*. **2008**, 41, 641.

69. Emsley, P.; Cowtan, K. *Acta Crystallogr.* **2004**, D60, 2126.

70. Murshudov, G. N.; Vagin, A. A.; Lebedev, A.; Wilson, K. S.; Dodson, E. J. *Acta Crystallogr.* **1999**, D55, 247.

71. Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. *Acta Crystallogr.* **2002**, D58, 1948.

72. Holm, L.; Rosenstrom, P. *Nucleic Acids Res. 38* (Suppl.), **2010**, W545.

73. Geoghegan, K. F.; Dixon, H. B.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; LeMotte, P. K.; McColl, A. S.; Kamath, A. V.; Stroh, J. G. *Anal. Biochem*. **1999**, 267, 169.

74. Collins, M. D.; Jones, D. *Microbiol. Rev*. **1981**, 45, 316.

75. Nowicka, B.; Kruk, J. *Biochim. Biophys. Acta, Bioenerg*. **2010**, 1797, 1587.

76. Lenaz, G.; Fato, R.; Formiggini, G.; Genova, M. L. *Mitochondrion* **2007**, 7, 8.

77. Fiorini, R.; Ragni, L.; Ambrosi, S.; Littarru, G. P.; Gratton, E.; Hazlett, T.

*Photochem. Photobiol*. **2008**, 84, 209.

78. IUPAC-IUB, *Biochim. Biophys. Acta* **1965**, 107, 5.

79. Kroger, A.: Phosphorylative electron transport with fumarate and nitrate as terminal hydrogen acceptors. In *Microbial Energetics* (Haddock, B. A.; Hamilton, W. A. Eds.); Cambridge University Press: Cambridge, **1977**, pp. 61.

80. Guest, J. R. *J. Gen. Microbiol*. **1979**, 115, 259.

81. Meganathan, R. *FEMS Microbiol. Lett*. **1984**, 24, 57.

82. Miguel, L.; Meganathan, R. *Curr. Microbiol.* **1991**, 22, 109.

83. Nitschke, W.; Kramer, D. M.; Riedel A.; Liebl, U.: From Naphtho-to Benzoquinones-(r) evolutionary reorganizations of electron transfer chains. In *Photosynthesis: From Light to Biosphere, Vol I,* (P. Mathis Eds.); Kluwer Academic publishers, Dordrecht, **1995**, pp. 945.

84. Lubben, M.; *Biochim. Biophys. Acta* **1995**, 1229, 1.

85. Soballe, B.; Poole R. K. *Microbiology* **1999**, 145, 1817.

86. Schoepp-Cothenet, B.; Lieutaud, C.; Baymann, F.; Vermeglio, A.; Friedrich, T.; Kramerd, D. M.; Nitschke, W.; *Proc. Natl. Acad. Sci. USA* **2009**, 106, 8549.

87. Debnath, J.; Siricilla, S.; Wan, B.; Crick, D. C.; Lenaerts, A. J.; Franzblau, S. G.; Kurosu, M. *J. Med. Chem*. **2012**, 55, 3739.

88. Richardson, D.; Sawers, G. *Science,* **2002**, 295, 1842.

89. Lancaster, C. R. D.; Simon, J. *Biochim. Biophys. Acta* **2002**, 1553, 84.

90. Kurokawa, T.; Sakamoto, J. *Arch. Microbiol.* **2005**, 183, 317.

91. Fernandes, A. S.; Konstantinov, A. A.; Teixeira, M.; Pereira, M. M. *Biochem.*

*Biophys. Res. Commun.* **2005**, 565.

92. Madej, M. G.; Nasiri, H. R.; Hilgendorff, N. S.; Schwalbe, H.; Unden, G.; Lancaster, C. R. D. *Biochemistry* **2006**, 45, 15049.

93. Lancaster, C. R. D. *FEBS Lett*. **2003**, 555, 21.

94. Nantapong, N.; Otofuji, A.; Migita, C. T.; Adachi, O.; Toyama, H.; Matsushita, K. *Biosci. Biotechnol. Biochem.* **2005**, 69, 149.

95. Kroger, A.; Biel, S.; Simon, J.; Gross, R.; Unden, G.; Lancaster, C. R. D. *Biochim. Biophys. Acta* **2002**, 1553, 23.

96. Pires, R. H.; Lourenc, A. I.; Morais, F.; Teixeira, M.; Xavier, A. V.; Saraiva, L. M.; Pereira, I. A. C. *Biochim. Biophys. Acta* **2003**, 1605, 67.

97. Jormakka, M.; Tornroth, S.; Byrne, B.; Iwata, S. *Science* **2002**, 295, 1863.

98. Pinho, D.; Besson, S.; Silva, P. J.; de Castro, B.; Moura, I. *Biochim. Biophys. Acta* **2005**, 1723, 151.

99. Giordani, R.; Buc. J. *Eur. J. Biochem*. **2004**, *271*, 2400.

100. Kusumoto, K.; Sakiyama, M.; Sakamoto, J.; Noguchi, S.; Sone, N. *Arch, Microbiol.* **2000**, 173, 390

101. Trumpower, B. L. *Microbiol. Mol. Biol. Rev*. **1990**, 54, 101

102. Junemann, S. *Biochim. Biophys. Acta* **1997**, 1321, 107.

103. Ke, B.: Photosynthesis: An Overview. In *Photosynthesis: Photobiochemistry and Photobiophysics: Advances of Photosynthesis and Respiration*, Vol. 10 (Govindjee Series Ed.), Springer, Dordrecht, **2001**, pp. 1.

104. Hale, M. B.; Blankenship, R. E.; Fuller, R. C. *Biochim. Biophys. Acta* **1983**, 723,

376.

105. Hauska, G.; Schoedl, T.; Remigy, H.; Tsiotis, G. *Biochim. Biophys. Acta* **2001**, 1507, 260.

106. Neerken, S.; Ames, J. *Biochim. Biophys. Acta* **2001**, 1507, 278.

107. Oh-oka, H. *Photochem. Photobiol***. 2007**, 83, 177.

108. Mimuro, M.; Tsuchiya, T.; Inoue, H.; Sakuragi, Y.; Itoh, Y.; Gotoh, T.; Miyashita, H.; Bryant, D. A.; Kobayashi, M. *FEBS Lett.* **2005**, 579, 3493.

109. Thummer, R.; Klimmek, O.; Schmitz, R. A. *J. Biol. Chem*. **2007**, 282, 12517.

110. Cranenburg, E. C. M.; Schurgers, L. J.; Vermeer, C. *Thromb. Haemost*. **2007**, 98, 120.

111. Plaza, S. M.; Lamson, D. W. *Altern. Med. Rev*. **2005**, 10, 24.

112. Tie, J.-K.; Stafford, D. W. *Vitam. Horm*. **2008**, 78, 103.

113. Lamson, D. W.; Plaza, S. M. *Altern. Med. Rev*. **2003**, 8, 203.

114. Olson, R. E. *Annu. Rev, Nutr*. 1984, 4, 281.

115. Suttie, J. W. *Annu. Rev. Nutr.* 1995, 15, 399.

116. Fernandez, F.; Collins, M. D.; Hill, M. J. *Biochem. Soc. Trans*. 1985, 13, 223.

117. Lange, B. M.; Rujan, T.; Martin, W.; Croteau, R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, 97, 13172.

118. Lichtenthaler, H. K. *Annu. Rev. Plant Physiol. Plant. Mol. Biol.* **1999**, 50, 47.

119. Dairi, T. *Methods Enzymol.* **2012**, 515, 107.

120. Dairi, T.; Kuzuyama, T.; Nishiyama, M.; Fujii, I. *Nat. Prod. Rep.* **2011**, 28, 1054.

121. Bentley, R.; Meganathan, R. *Microbiol. Rev.* **1982**, 46, 241.

122. Meganathan, R. *Vitam. Horm*. **2001**, 61, 173.

123. Hiratsuka, T.; Furihata, K.; Ishikawa, J.; Yamashita, H.; Itoh, N.; Seto, H.; Dairi, T. *Science* **2008**, 321, 1670.

124. Seto, H.; Jinnai, Y.; Hiratsuka, T.; Fukawa, M.; Furihata, K.; Itoh, N.; Dairi, T. *J. Am. Chem. Soc.* **2008**, 130, 5614.

125. Cox, G. B.; Gibson, F. *Biochim. Biophys. Acta* **1964**, 93, 204.

126. Campbell, I. M.; Coscia, C. J.; Kelsey, M.; Bentley, R. *Biochem. Biophys. Res. Commun.* **1967**, 28, 25.

127. Campbell, I. M. *Tetrahedron Lett.* **1969**, 1969, 4777.

128. Robins, D. J.; Campbell, I. M.; Bentley, R. *Biochem. Biophys. Res. Commun.* **1970**, 39, 1081.

129. Robins, D. J.; Bentley, R. *J. Chem. Soc. Chem. Commun*. **1972**, 232.

130. Dansette, P.; Azerad, R. *Biochem. Biophys. Res. Commun.* **1970**, 40, 1090.

131. Robins, D. J.; Yee, R. B.; Bentley, R. *J. Bacteriol*. **1973**, 116, 965.

132. Young, I. G. *Biochemistry*, **1975**, 14, 399.

133. Meganathan, R.; Bentley, R. *Biochemistry*, **1981**, 20, 5336.

134. Windhalm, J. R.; van Oostende, C.; Furt, F.; Basset, G. J. C. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, 106, 5599.

135. Kurosu, M.; Begari, E. *Molecules*, **2010**, 15, 1531.

136. Lu, X.; Zhang, H.; Tonge, P. J.; Tan, D. S. *Bioorg. Med. Chem. Lett.* **2008,** 18, 5963.

137. Li, X.; Liu, N.; Zhang, H.; Knudson, S. E.; Li, H.-J.; Lai, C. T.; Simmerling, C.;

Slayden, R. A.; Tonge, P. J. *ACS Med. Chem. Lett*. **2011**, 2, 818.

138. Lu, X.; Zhou, R.; Sharma, I.; Li, X.; Kumar, G.; Swaminathan, S.; Tonge, P. J.; Tan, D. S. *ChemBioChem* **2012**, 13, 129.

139. Bentley, S. D. et al *Nature,* **2002**, 417, 141.

140. Dairi, T. *J. Antibiot*. **2005**, 58, 227.

141. Dairi, T. *J. Antibiot*. **2009**, 62, 347.

142. Hiratsuka, T.; Itoh, N.; Seto, H.; Dairi, T. *Biosci. Biotechnol. Biochem.* **2009**, 73, 1137.

143. Li, X.; Apel, D.; Gaynor, E. C.; Tanner, M. E. *J. Biol. Chem.* **2011**, 286, 19392.

144. Cooper, L. E.; Fedoseyenko, D.; Abdelwahed, S. H.; Kim, S.-H.; Dairi, T.; Begley, T. P. *Biochemistry* **2013**, 52, 4592.

145. Tanaka, R.; Kunisada, T.; Kushida, N.; Yamada, K.; Ikeda, S.; Noike, M.; Ono, Y.; Itoh, N.; Takami, H.; Seto, H.; Dairi, T. *J. Antibiot*. **2011**, *64*, 151.

146. Adams, J.; Pepping, J. *Am. J. Health Syst. Pharm.* **2005**, 62, 1574.

147. Mahanta, N.; Fedoseyenko, D.; Dairi, T.; Begley, T. P. *J. Am. Chem. Soc.* **2013**, 135, 15318.

148. Peddi, S.; Patel, M.V.; Rohde, J.J. *U.S. Patent 20100267738,* October 21, **2010.**

149. Majer, P.; Hin, B.; Stoermer, D.; Adams, J.; Xu, W.; Duvall, B.R.; Delahanty, G.; Liu, Q.; Stathis, M.J.; Wozniak, K.M.; Slusher, B.S.; Tsukamoto, T. *J. Med. Chem*. **2006**, *49*, 2876.

150. Reiger, C. E.; Turnbull, J. L. *Prep. Biochem. Biotechnol*. **1996**, 26, 67.

151. Young, I. G.; Gibson, F.; Macdonald, C. G. *Biochim. Biophys. Acta., General*

*subjects* **1969**, 192, 62.

152. Sambrook, J.; Russell, D. W. *Molecular Cloning: A Laboratory Manual*; 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, **2001.**

153. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr. D* **2011**, *67*, 235.

154. Tyagi, R.; Burley, S.; Swaminathan, S. *BMC Structural Biology* **2007**, *7* , 62.

155. Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355.

156. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr. D Biol. Crystallogr.* 2010, *66* , 486.

157. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr. D Biol. Crystallogr.* 2010, *66*, 213.

158. Hanzelmann, P.; Harnandez, H. L.; Menzel, C.; Garcia-Serres, R.; Huynh, B. H.; Johnson, M. K.; Mendel, R. R.; Schindelin, H. *J. Biol. Chem.* **2004**, 279, 34721.

159. Cammak, R.; Patil, D. S.; Fernandez, V. M. *Biochem. Soc. Trans.* **1985**, 13, 572.

160. Wecksler, S. R.; Stoll, S.; Tran, H.; Magnusson, O. T.; Wu, S.; King, D.; Britt, R.

D.; Klinman, J. P. *Biochemistry*, **2009**, 48, 10151.

161. Beinert, H.; Kennedy, M. C.; Stout, C. D. *Chem. Rev*. **1996**, 96, 2335.

162. Arakawa, C.; Kuratsu, M.; Furihata, K.; Hiratsuka, T.; Itoh, N.; Seto, H.; Dairi, T. *Antimicrob. Agents Chemother*. **2011**, 55, 913.

163. Goble, A. M.; Toro, R.; Li, X.; Ornelas, A.; Fan, H.; Eswaramoorthy, S.; Patskovsky, Y.; Hillerich, B.; Seidel, R.; Sali, A.; Shoichet, B. K.; Almo, S. C.; Swaminathan, S.; Tanner, M. E.; Raushel, F. M. *Biochemistry*, **2013**, 52, 6525.

164. Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E. Nucleic Acids Res. 2001, 29, 1097.

165. Frey, P. A.; Hegeman, A. D.; Ruzicka, F. J. *Crit. Rev. Biochem. Mol. Biol*. **2008**, 43, 63.

166. Roach, P. L. *Curr. Opin. Chem. Biol.* **2011**, 15, 267.

167. Vey, J. L.; Drennan, C. L. *Chem. Rev.* **2011**, 111, 2487.

168. Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M. *Chem. Rev.* **2014**, 114, 4229.

169. Mehta, A. P.; Abdelwahed, S. H.; Mahanta, N.; Fedoseyenko, D.; Philmus, B.; Cooper, L. E.; Liu, Y.; Jhulki, I.; Ealick, S. E.; Begley, T. P. *J. Biol Chem.* **2015,** 290, 3980.

170. Wagner, A. V. F; Demand, J.; Schilling, G.; Pils, T.; Knappe, J. *Biochem. Biophys. Res. Commun.* **1999**, 254, 306.

171. Bietti, M.; Calcagni, A.; Cicero, D. O.; Martella, R.; Salamone, M. *Tetrahedron Lett.* **2010**, 51, 4129.

172. Baroudi, A.; Alicea, J.; Flack, P.; Kirincich, J.; Alabugin, I. V. *J. Org. Chem.* **2011**, 76, 1521.

173. Baroudi, A.; Flack, P.; Alabugin, I. V. *Chem.-Eur. J.* **2010**, 16, 12316.

174. Zhang, Y.; Zhu, X.; Torelli, A. T.; Lee, M.; Dzikovski, B.; Koralewski, R. M.; Wang, E.; Freed, J.; Krebs, C.; Ealick, S. E.; Lin, H. *Nature* **2010**, 465, 891.

175. Miyakoshi, S.; Haruyama, H.; Shioiri, T.; Takahashi, S.; Torikata, A.; Yamazaki, M. *J. Antibiot.* **1992**, 45, 394.

176. Cooper, R.; Mierzwa, R.; Patel, M.; Pramanik, B.; Puar, M. S.; Troyanovich, J.; Gullo, V. *J. Nat. Prod*. **1988**, 51, 1207.

177. Grisostomi, C.; Kast, P.; Pulido, R.; Huynh, J.; Hilvert, D. *Biorg. Chem.* **1997**, *25*, 297.

178. Hidalgo, E.; Bollinger, J. M. Jr; Bradley, T. M.; Walsh, C. T. ; Demple, B. *J. Biol. Chem.* **1995**, *270*, 20908.

179. Brumby, P. E.; Miller, R. W. Massey, V. *J. Biol. Chem.* **1965**, *240*, 2222.

180. Wagner, P.J.; Lindstrom, M.J.; Sedon, J.H.; Ward, D. R. *J. Am. Chem. Soc*. **1981**, *103*, 3842.

181. Burdi, D.; Begley, T.P. *J. Am. Chem. Soc.* **1991**, *113*, 7768.

182. Creary, X. *Acc. Chem. Res*. **2006**, *39*, 761.

183. Henry, D. J.; Parkinson, C. J.; Mayer, P. M.; Radom, L. *J. Phys. Chem. A*, **2001**, *105*, 6750.

184. Mayer, P.M.; Glukhovtsev, M.N.; Gauld, J. W.; Radom. L. *J. Am. Chem. Soc*. **1997**, *119*, 12889.

185. Haynes, W. M. CRC Handbook of Chemistry and Physics; 95th ed.; CRC Press, 2014.

186. Kitz. R; Wilson, I. B. *J. Biol. Chem*. **1962**, *237*, 3245

187. Walsh, C. T. *Tetrahedron*, **1982**, *38(7*), 871.

188. Hanzelmann, P.; Schindelin, H. *Proc. Natl. Acad. Sci. USA* **2004**, *101),* 12870.

APPENDIX A

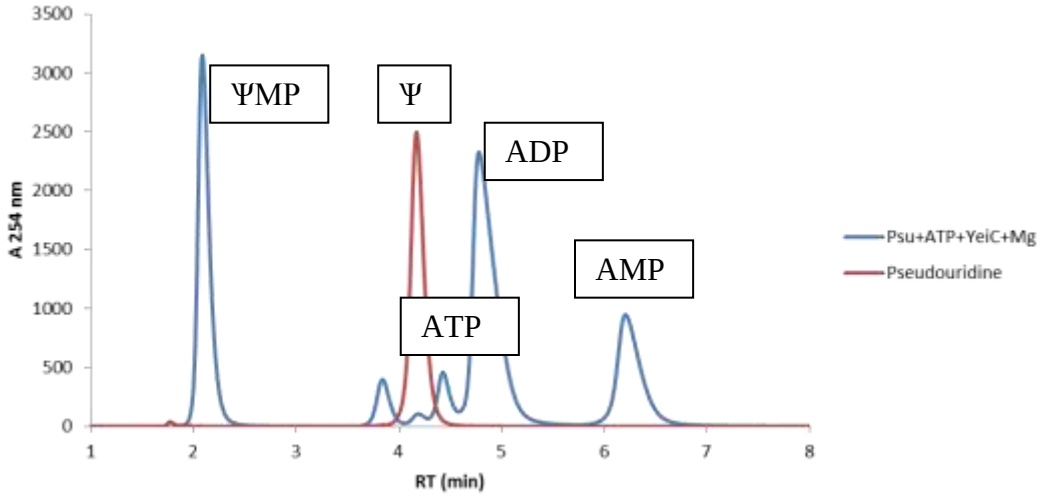



**Figure A1**: HPLC analysis of the Pseudouridine kinase (YeiC) reaction mixture. Blue trace: Chromatogram of the full reaction (Ψ,Psu + ATP + YeiC + $Mg^{2+}$) Formation of ΨMP is observed at 2.1 min. Red trace: Chromatogram of the authentic standard of pseudouridine, Ψ.

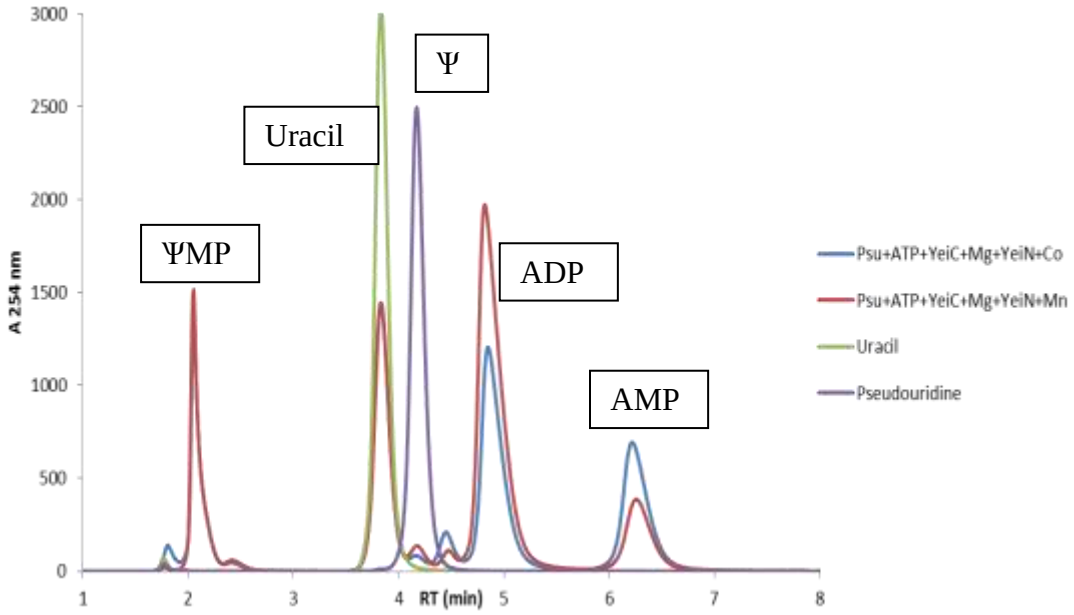


**Figure A2**: HPLC chromatogram of the coupled YeiC and YeiN assay. Blue trace: Chromatogram of the full reaction (Ψ,Psu + ATP + YeiC + $Mg^{2+}$ + YeiN + $Co^{2+}$). Red trace: Chromatogram of the full reaction (Ψ,Psu + ATP + YeiC + $Mg^{2+}$ + YeiN + $Mn^{2+}$). Formation of uracil is observed at 3.9 min. Green trace: Chromatogram of the authentic standard of uracil. Purple trace: Chromatogram of the authentic standard of pseudouridine, Ψ.

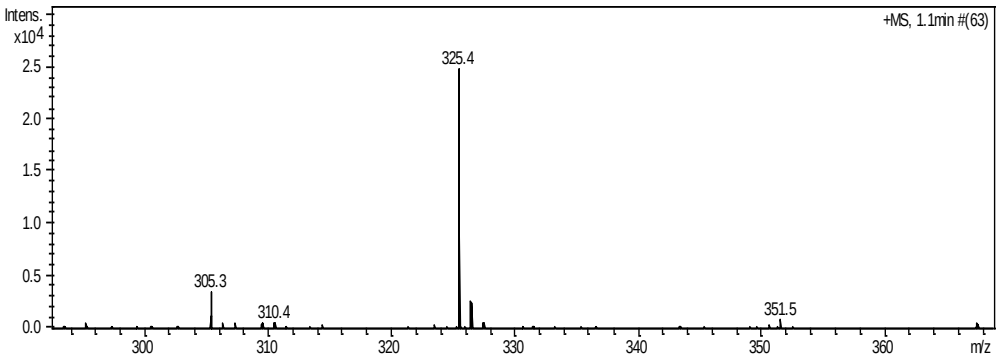



**Figure A3**: LC-MS spectrum of ΨMP, **9**

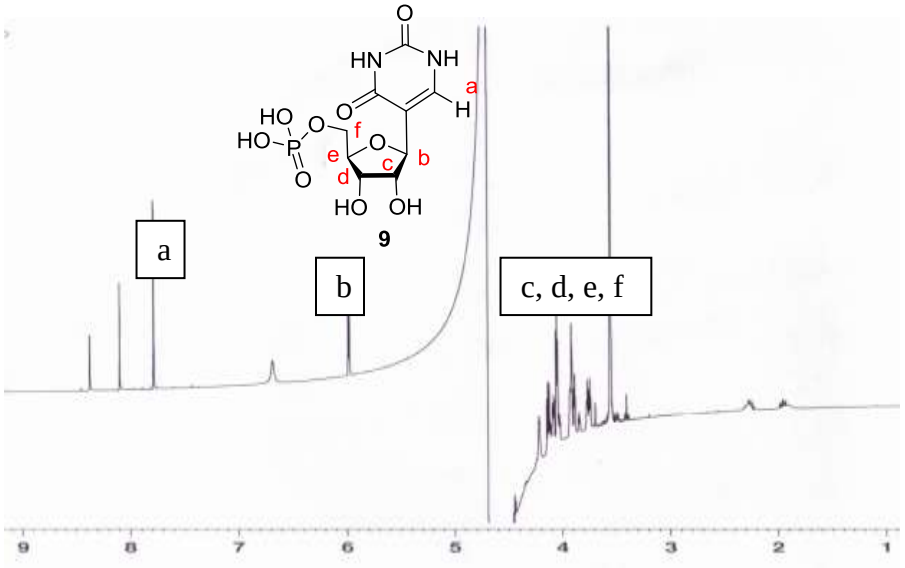



**Figure A4**: NMR spectrum of ΨMP, **9**

**Michaelis-Menten polts for YeiN mutants**

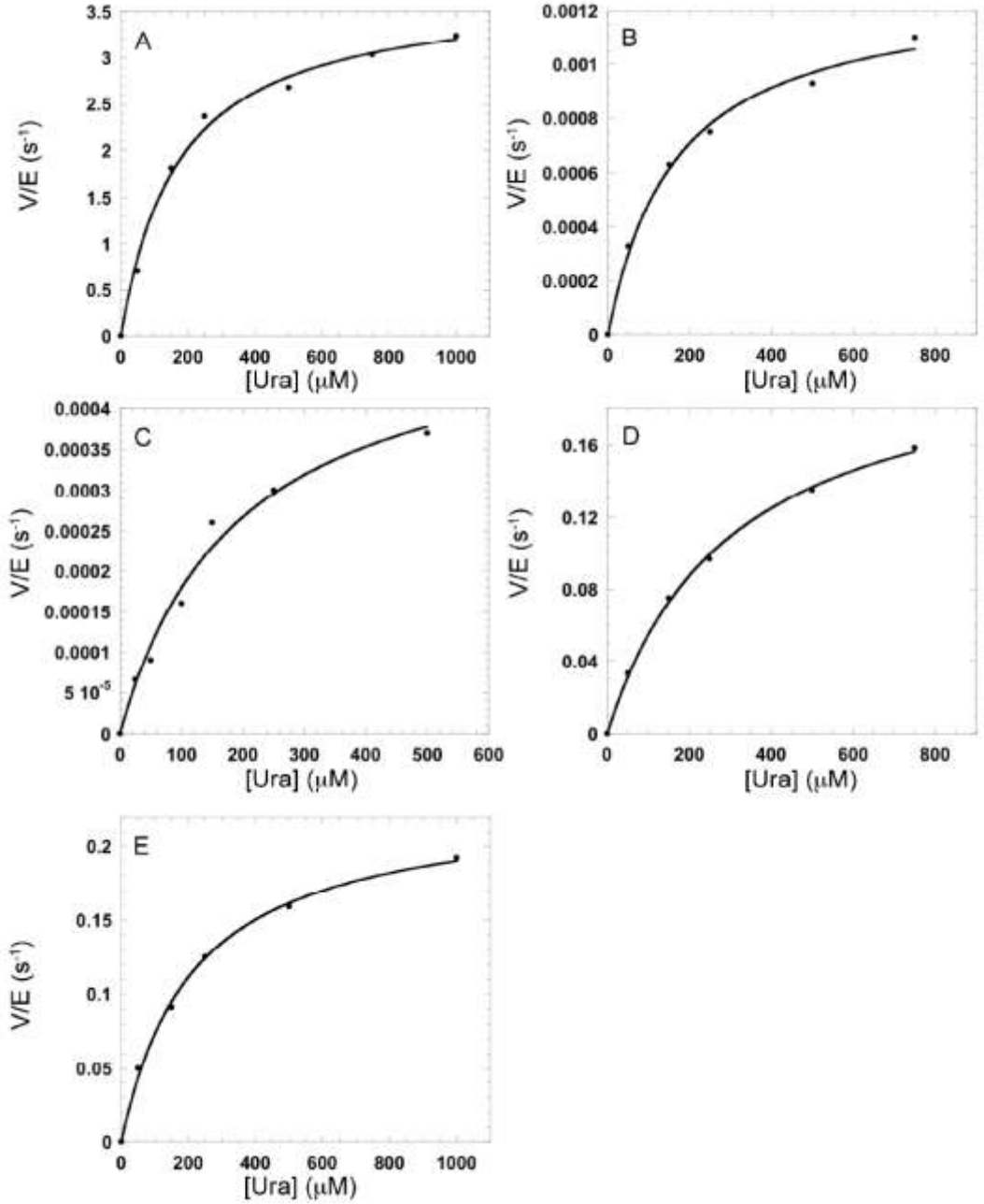



**Figure A5**: Michaelis Menten plots ΨMP glycosidase and its mutants: YeiN (A), K166A B), E31A(C), K93A (D), and N289A (E

Table A1: Designed primers for YeiN mutants.

**Supplementary Table 1**. In addition to the forward and reverse primers required to introduce the mutation, a third primer was designed to screen for the presence of the mutation by colony PCR. The primers designated "sF" were paired with the T7T primer (5′-GCTAGTTATTGCTCAGCGG-3′) and the primers designated "sR" were screened with the T7Plac primer (5′-TATAGGGGAATTGTGAGCGG-3′).

| Mutant | Primer Sequences* | |
|---|---|---|
| D149A | D149A-F | ccttcgatatttctgccgCGttgcaagaactggcaaatac |
| | D149A-R | gtatttgccagttcttgcaaCGcggcagaaatatcgaagg |
| | D149A-sF | ccttcgatatttctgccgCG |
| E31A | E31A-F | caaaaaaccggttgtggcgctggCGtcgaccattatttctcac |
| | E31A-R | gtgagaaataatggtcgaCGccagcgccacaaccggttttttg |
| | E31A-sF | aaccggttgtggcgctggCG |
| E31L | E31L-F | caaaaaaccggttgtggcgctgCTGtcgaccattatttctcac |
| | E31L-R | gtgagaaataatggtcgaCAGcagcgccacaaccggttttttg |
| | E31L-sF | aaccggttgtggcgctgCTG |
| K93A | K93A-F | cgtgaagggcataacgtgacccGCGgttagtcgtcgcgatttacc |
| | K93A-R | ggtaaatcgcgacgactaacCGCggtcacgttatgcccttcacg |
| | K93A-sF | aagggcataacgtgacccGCG |
| K93E | K93E-F | cgtgaagggcataacgtgacccGaGgttagtcgtcgcgatttacc |
| | K93E-R | ggtaaatcgcgacgactaacCtCggtcacgttatgcccttcacg |
| | K93E-sF | aagggcataacgtgacccGaG |
| K166A | K166A-F | ccgttgtttgtgccggggcgGCGtctattctcgatttagg |
| | K166A-R | cctaaatcgagaatagaCGCcgccccggcacaaacaacgg |
| | K166A-sF | ttgtttgtgccggggcgGCG |
| H137A | H137A-F | ccgggggaattggtggtgtgGCGcgcggggcggaacataccttcg |
| | H137A-R | cgaaggtatgttccgccccgcgCGCcacaccaccaatcccccgg |
| | H147A-sF | ggggaattggtggtgtgGCG |

| | | |
|---|---|---|
| H137N | H137N-F | ccgggggaattggtggtgtgAaCcgcggggcggaacatacctteg |
| | H137N-R | cgaaggtatgttccgccccgcgGtTcacaccaccaattccccegg |
| | H137N-sF | ggggaattggtggtgtgAaC |
| N289A | N289A-F | cggtgacagcctgaaatccGCGatccagctggtgttcaacaacg |
| | N289A-R | cgttgttgaacaccagctggatCGCggatttcaggctgtcaccg |
| | N289A-sR | ttgaacaccagctggatCGC |

APPENDIX B

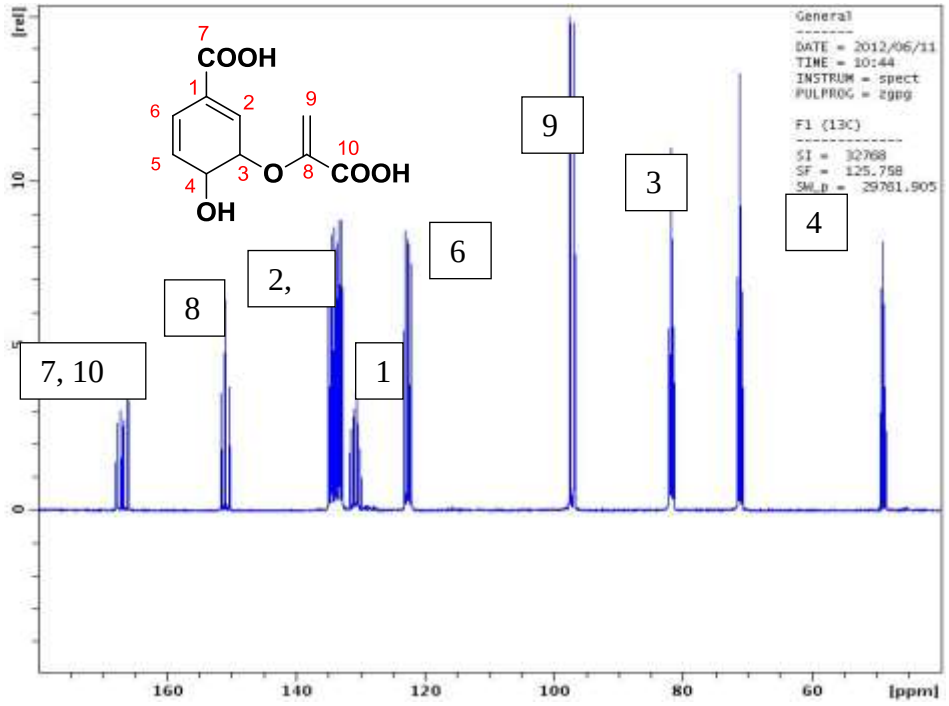



**Figure B1**: $^{1}H$ NMR spectra of [U-$^{13}C_{10}$]-chorismate isolated from *E.coli* KA-12.

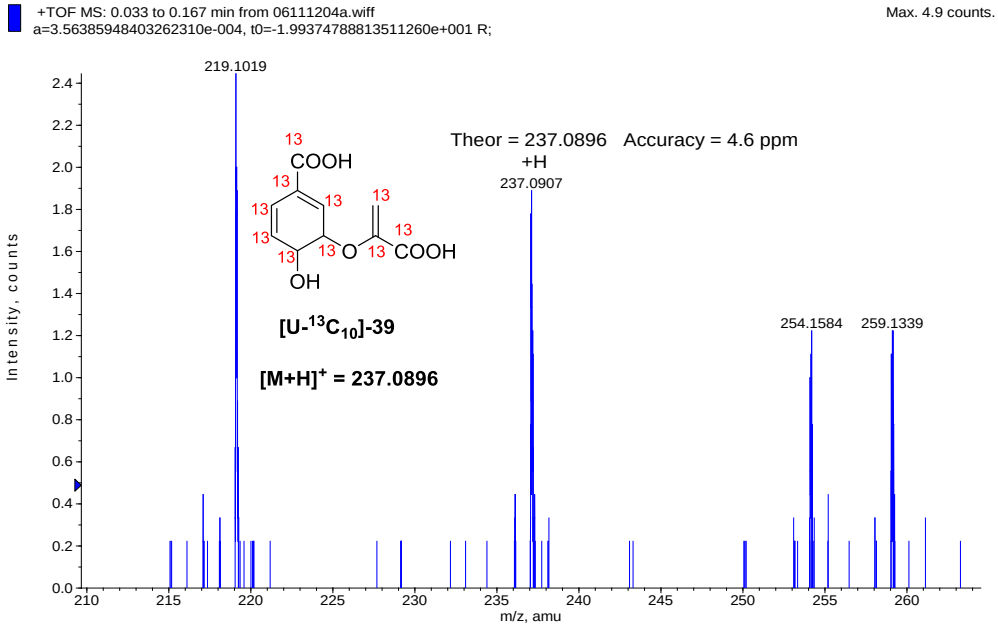



**Figure B2**: ESI MS (positive mode) of [U-$^{13}C_{10}$]-chorismate

MRGIRDPRLIPIAEKVMEGKRLSFEDGLVLYQTKDLPTLMRLANLVRERKHGHK

TYFVHSIRVSQTNI**CYVGCTFC**AFQRRFGEEGAWDWDVDEVVAWVKERYQPG

LTEIHLTAGHHPKRPFAYYLDLVRALKENFPGVQVKAWTAAEIHHFSKIARLPY

REVLKALKEAGLDAMPGGGAEIFAERVRRKIARAKVSAEGWLEIHRTAHELGIP

TNATMLYGHIETLEERLDHMDRLRRLQDETGGFMSFIPLAFQPDGNQLARELGK

KEFTTGLDDLRNLAVARLYLDNFPHIKGYWATLTPELAQVSLDWGVTDVDGTL

IEERIVHMAGSPTPQGLTKRELARIILMAGRIPVERDALYREVRVWDRVEA

**Figure B3**: Sequence of TTHA0804 protein (MqnE) from *T. thermophilus* HB8.

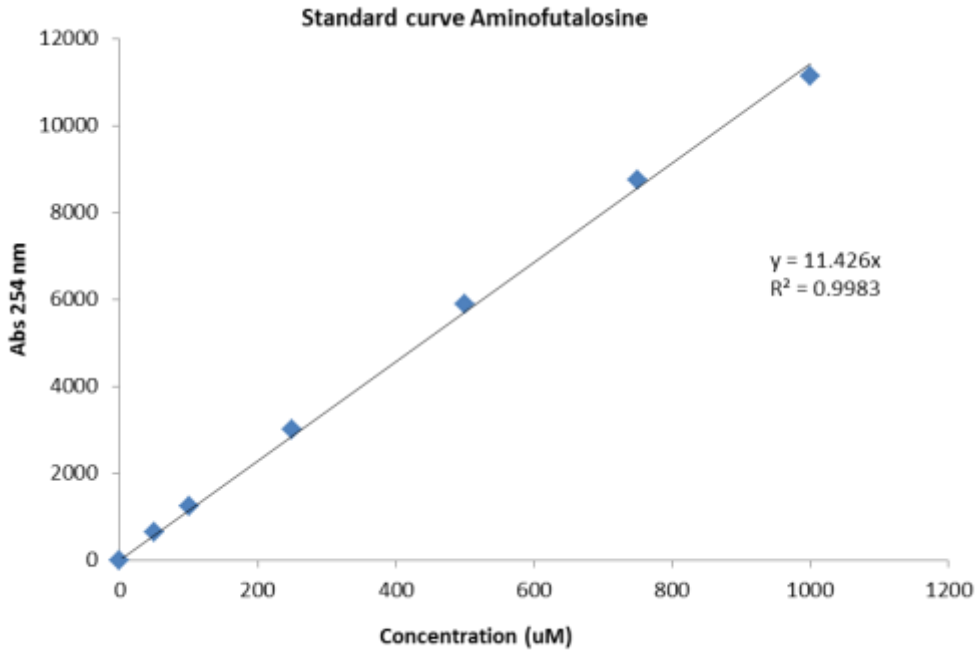



**Figure B4**: Standard curve for aminofutalosine, **64**.

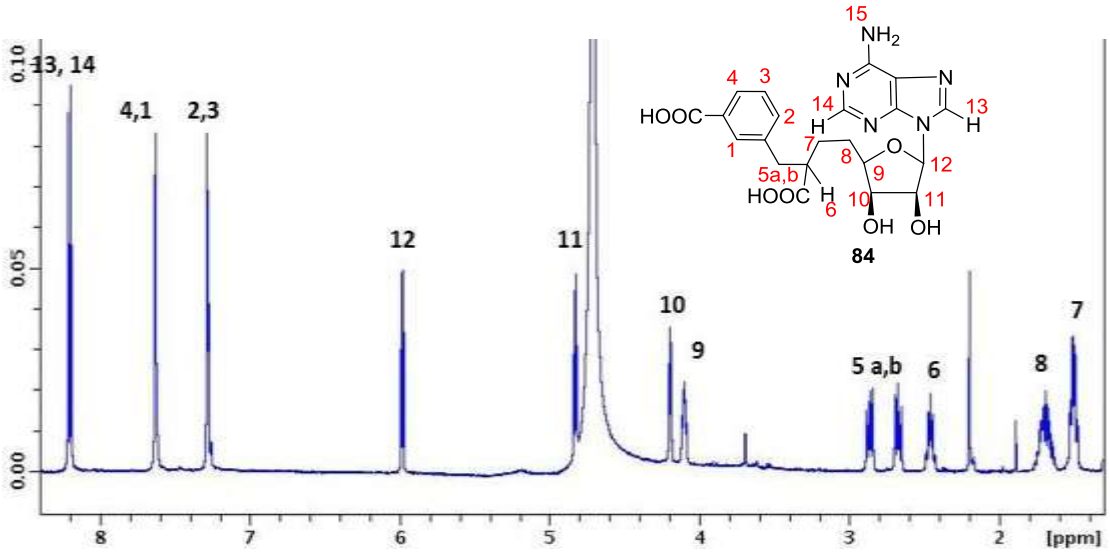



**Figure B5**: $^{1}$H NMR (500 MHz, $D_2O$) of compound **84**.

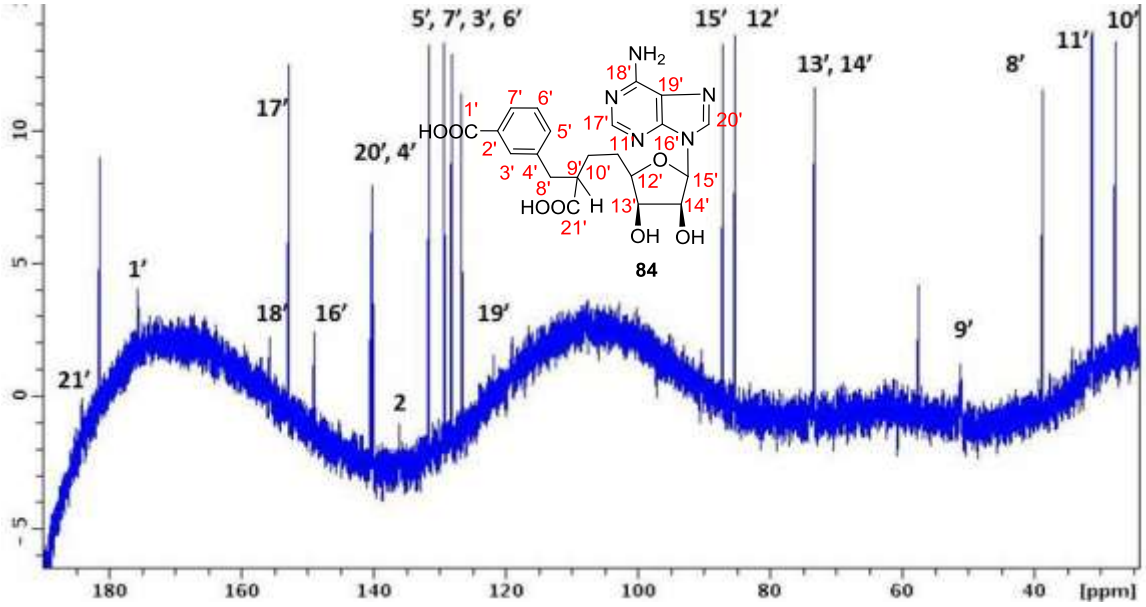



**Figure B6**: $^{13}$C NMR (125 MHz, $D_2O$) spectrum of compound **84**.

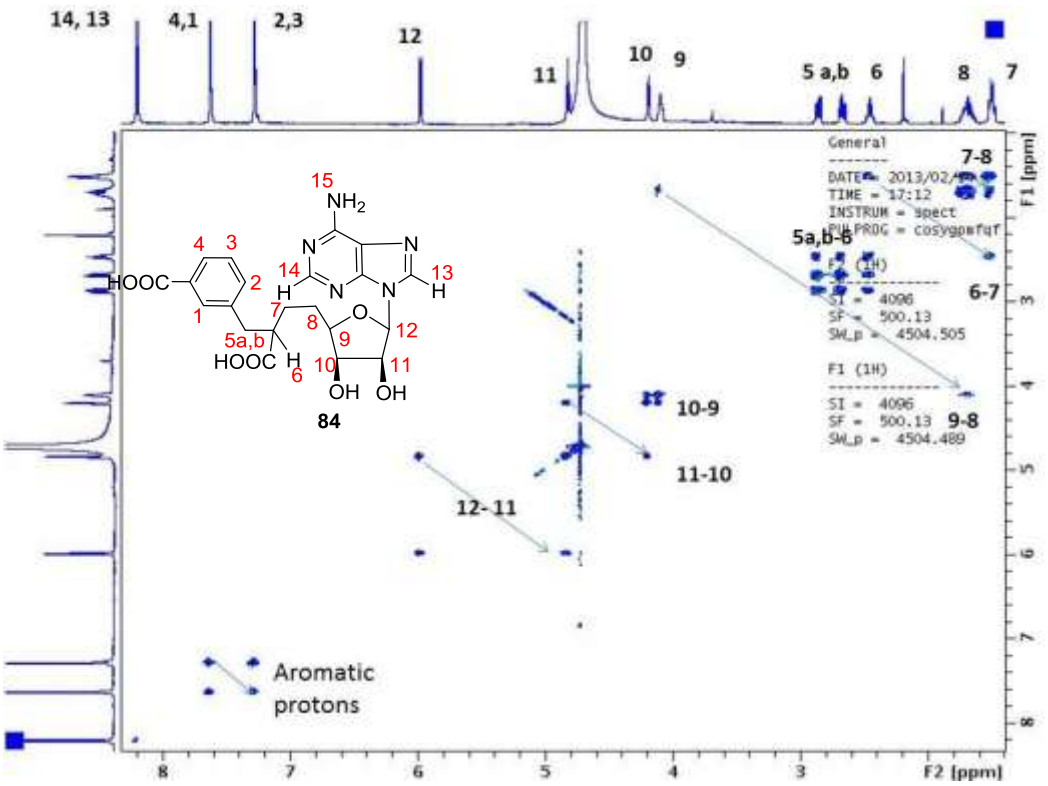



**Figure B7**: $^{1}$H COSY spectrum of compound **84**.

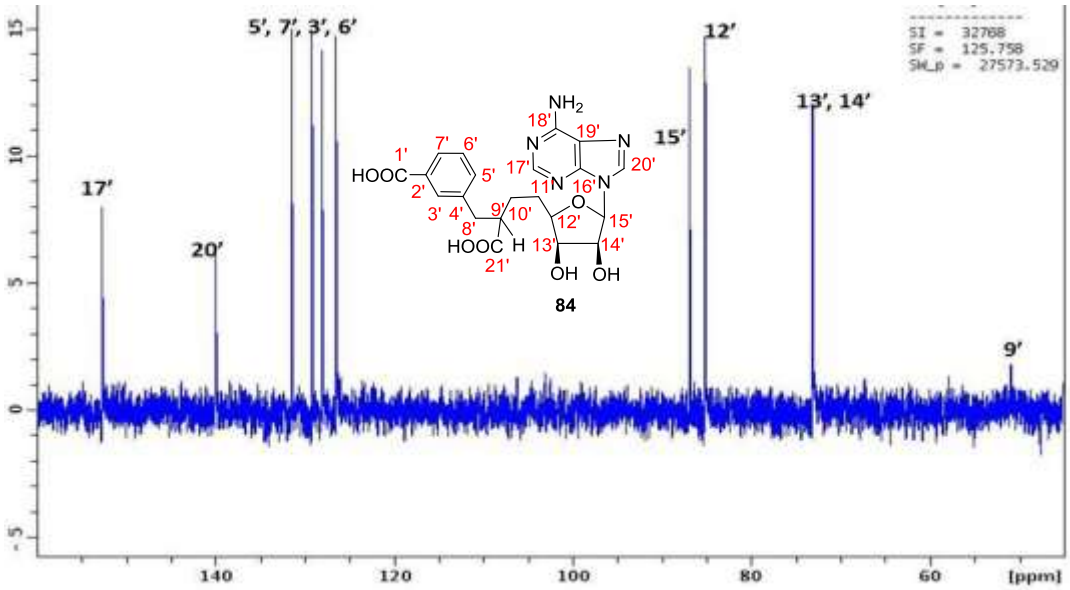



**Figure B8**: $^{13}$C DEPT-90 NMR (125 MHz, $D_2O$) spectrum of compound **84**.

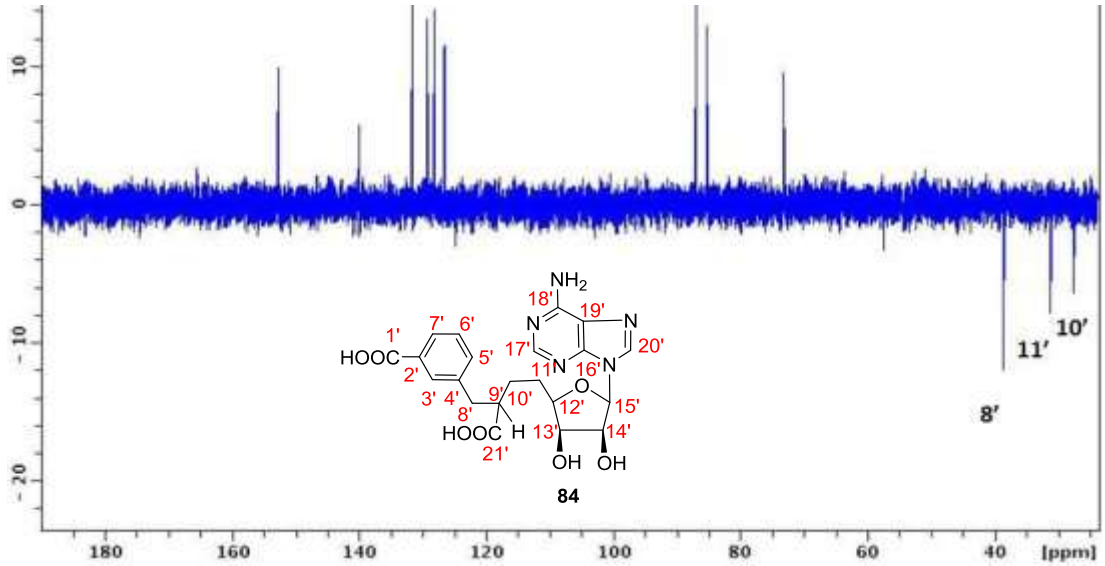



**Figure B9**: $^{13}$C DEPT-135 NMR (125 MHz, $D_2O$) spectrum of compound **84**.

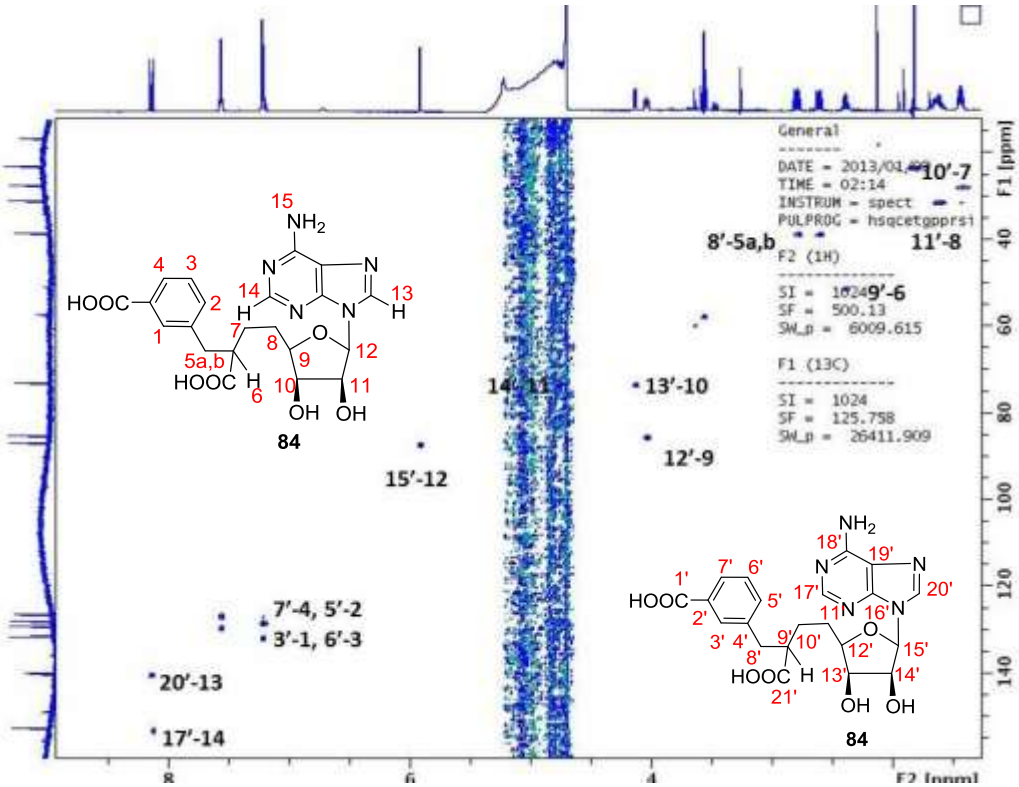



**Figure B10**: HSQC spectrum of compound **84**

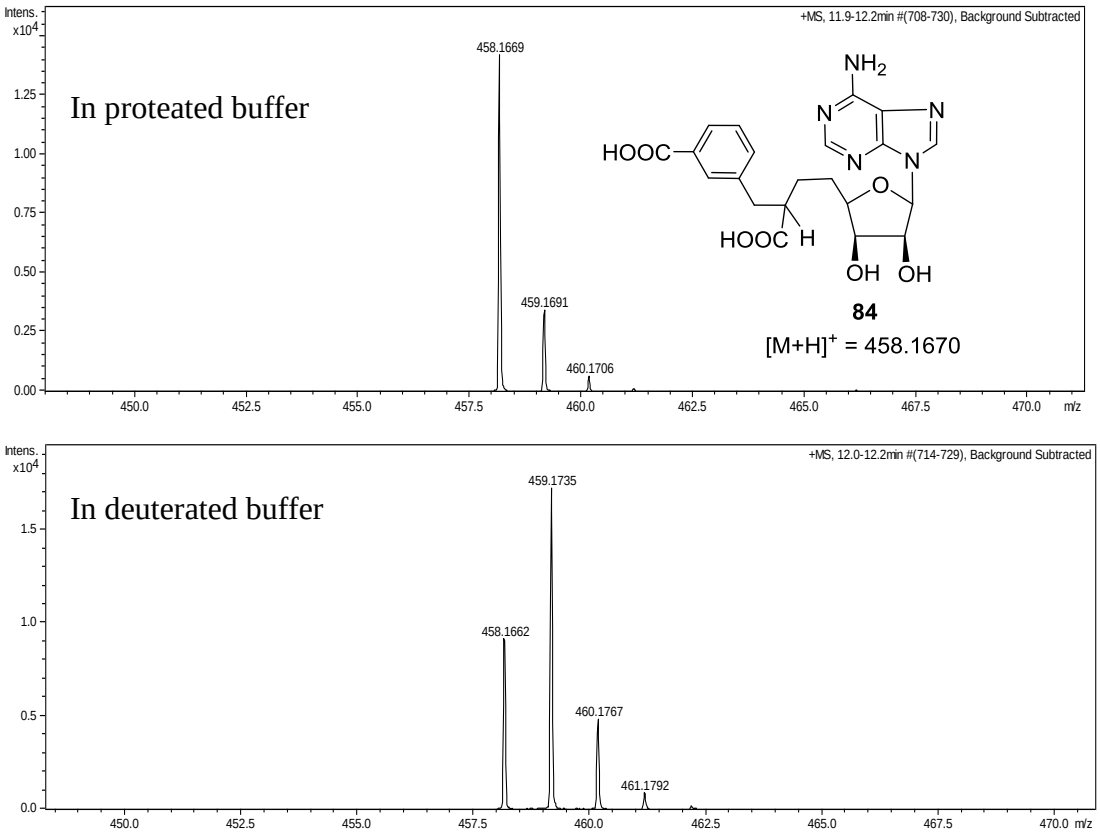



**Figure B11**: Incorporation of deuterium into compound **84**

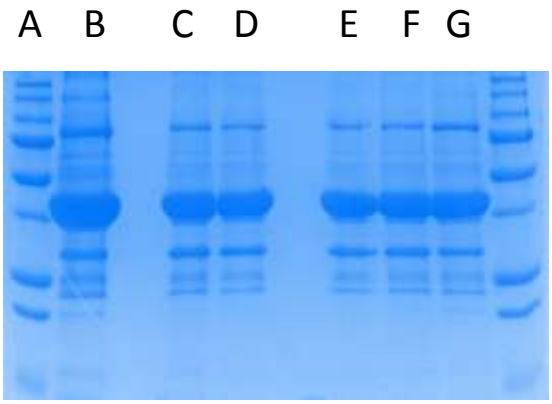
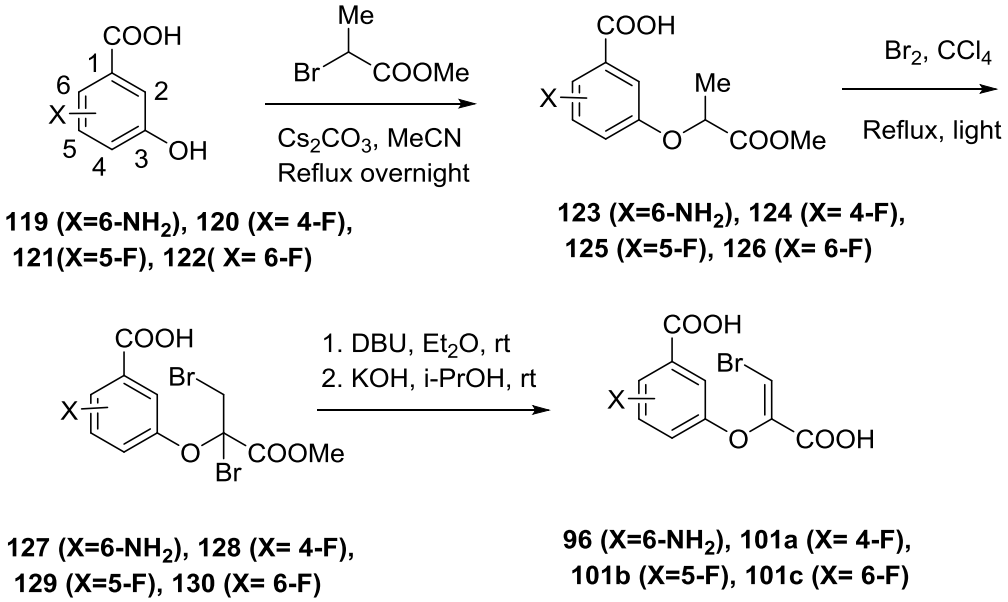


**Figure B12**: SDS-PAGE gel MqnE: A) Protein maeker B) Wild type MqnE C) and D) MqnE after reaction with **83,** E) MqnE in no substrate control F) MqnE in no SAM control G) MqnE in no dithionite control.




**Scheme B1**: Synthetic scheme for the substituted substrate analogs.

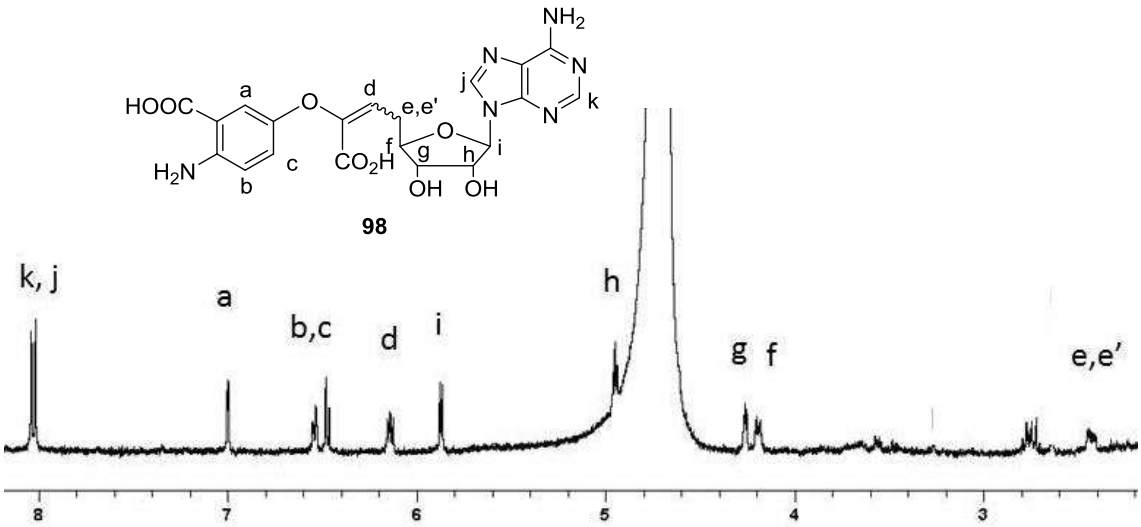



**Figure B13**: $^{1}$H NMR (500 mHz, $D_2O$) of compound **98**

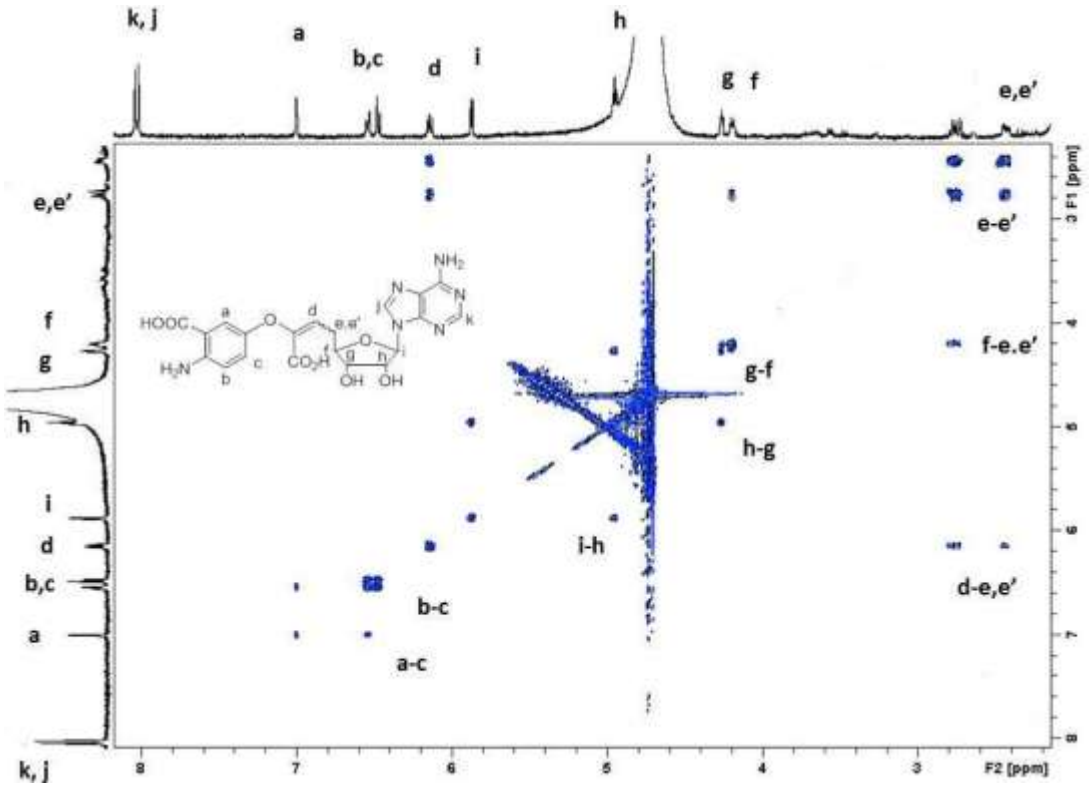



**Figure B14:** COSY spectrum of compound 98.

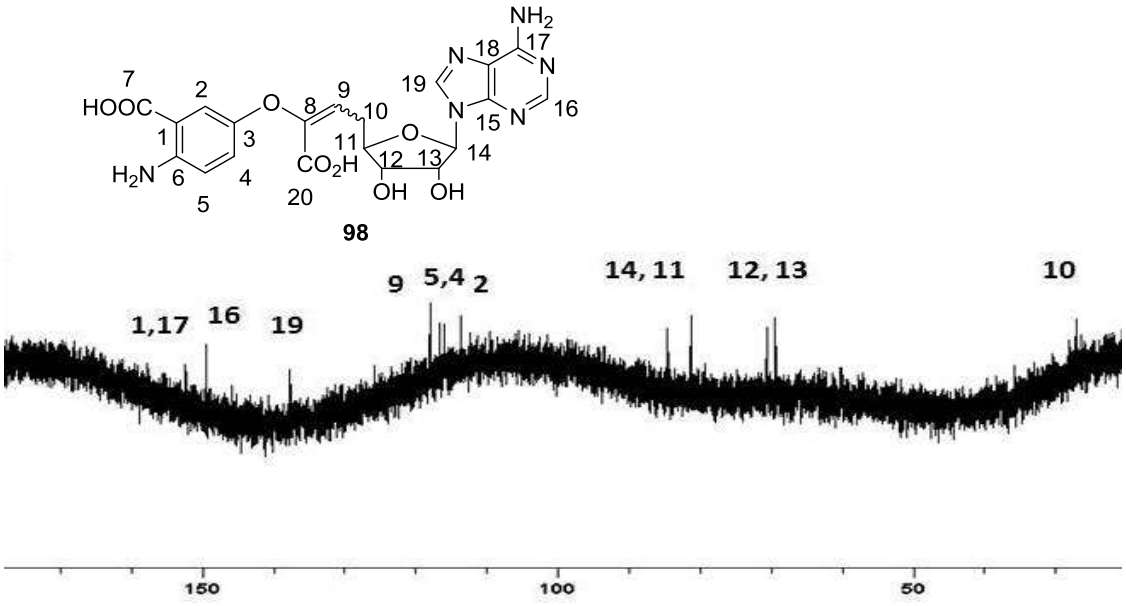



**Figure B15**: $^{13}C$ NMR (125 mHz, $D_2O$) of compound **98**

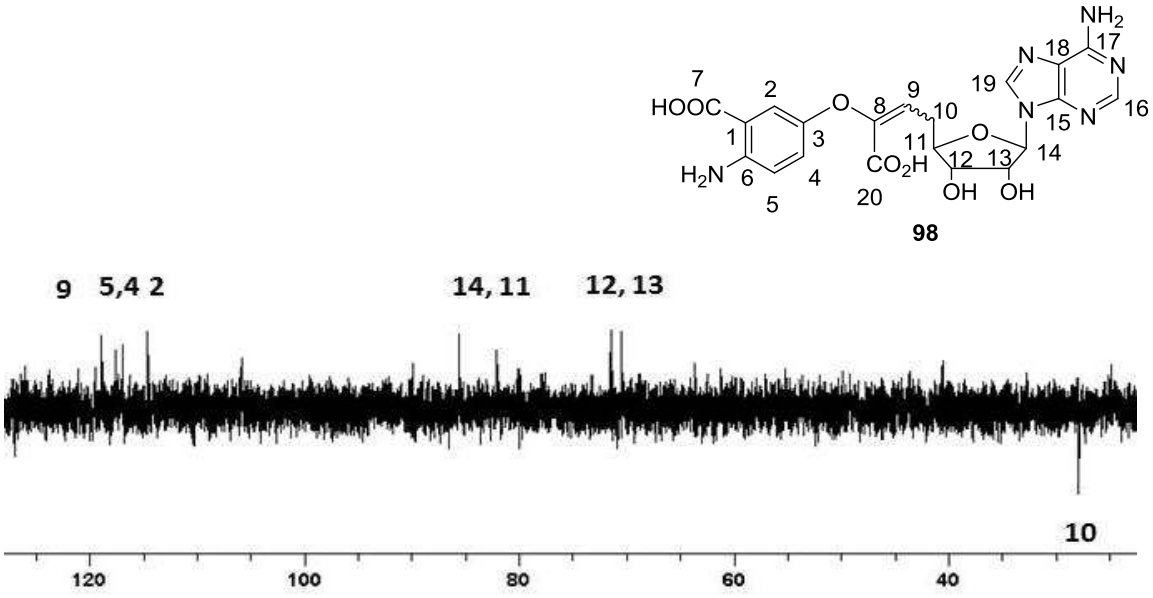



**Figure B16**: $^{13}C$ DEPT-NMR (125 mHz, $D_2O$) of compound **98**

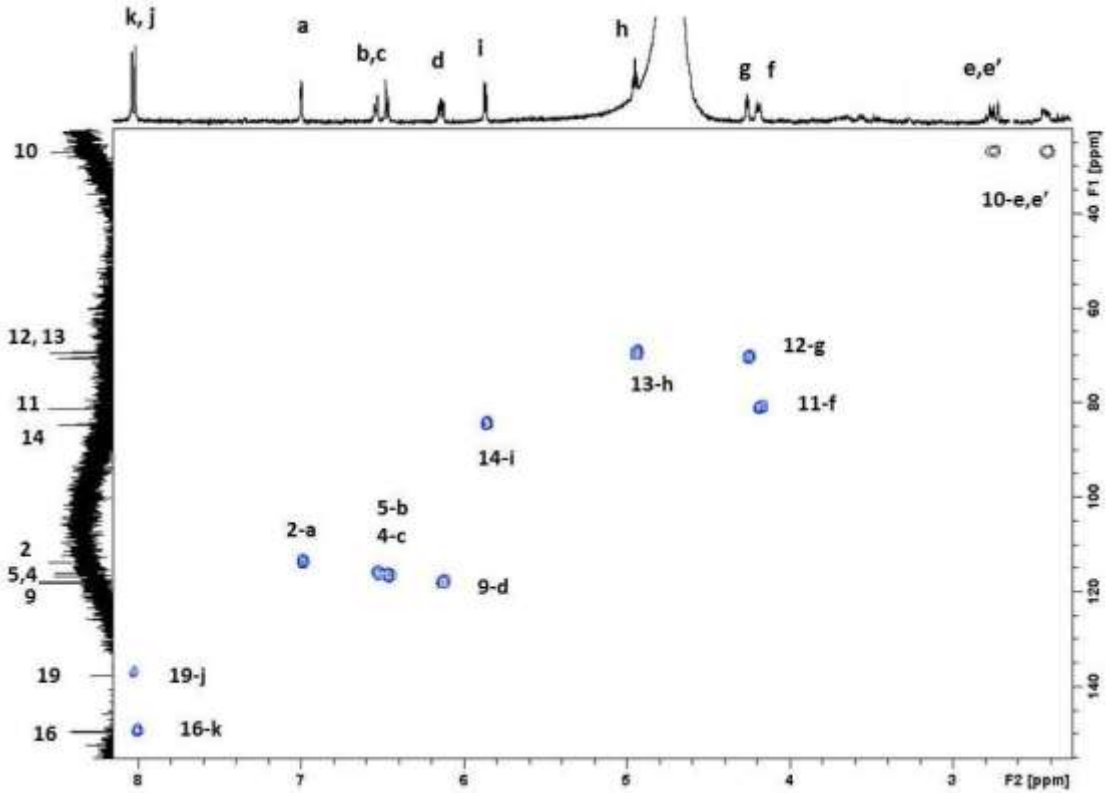



**Figure B17**: HSQC spectrum of compound **98.**

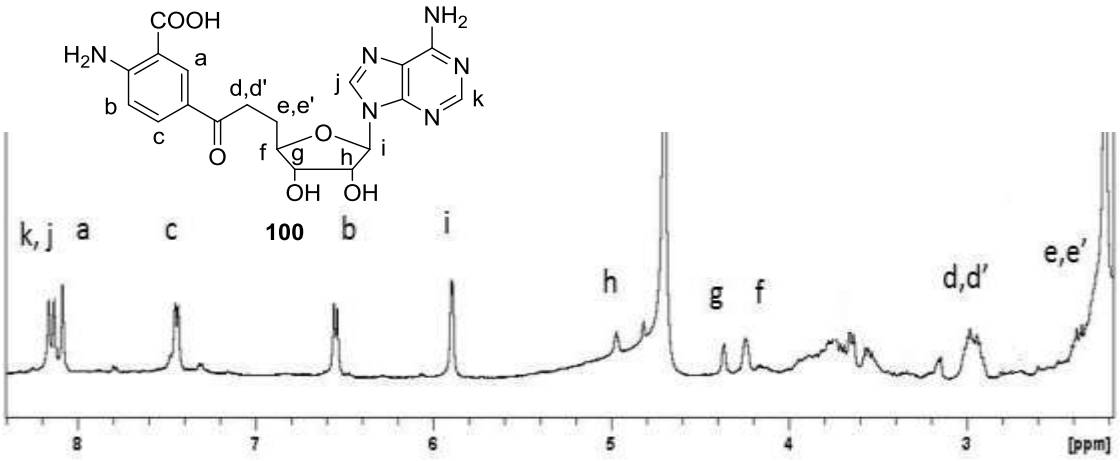



**Figure B18**: $^{1}H$ NMR (500 mHz, $D_2O$) of compound **100**

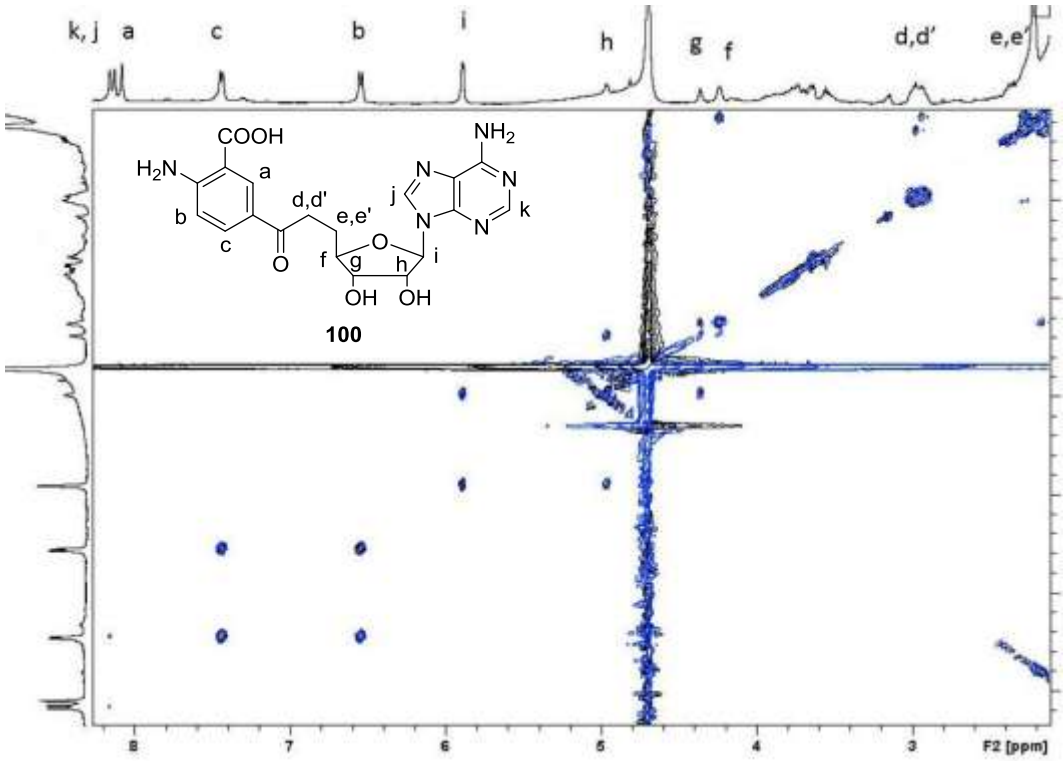



**Figure B19:** COSY spectrum of compound **100.**

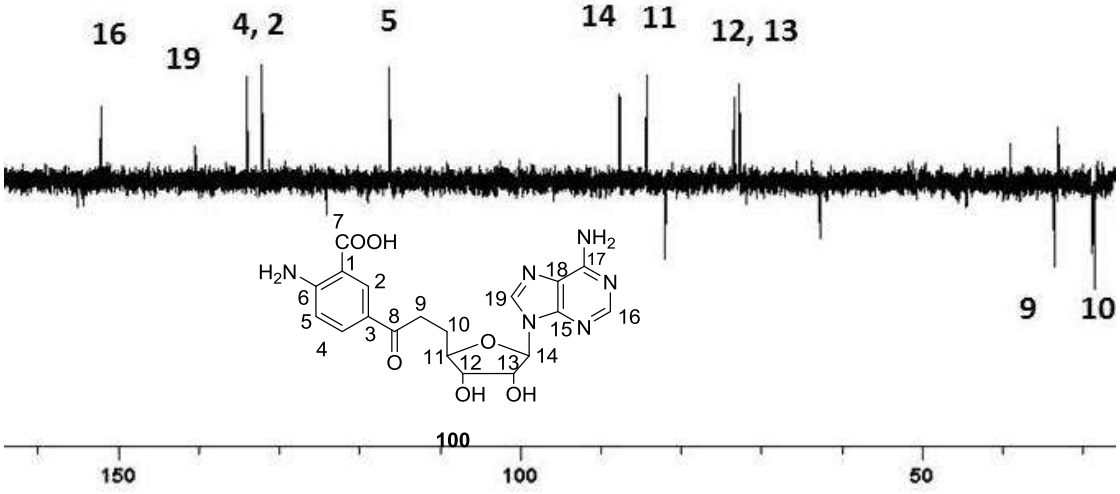



**Figure B20**: DEPT-135 spectrum of compound **100**

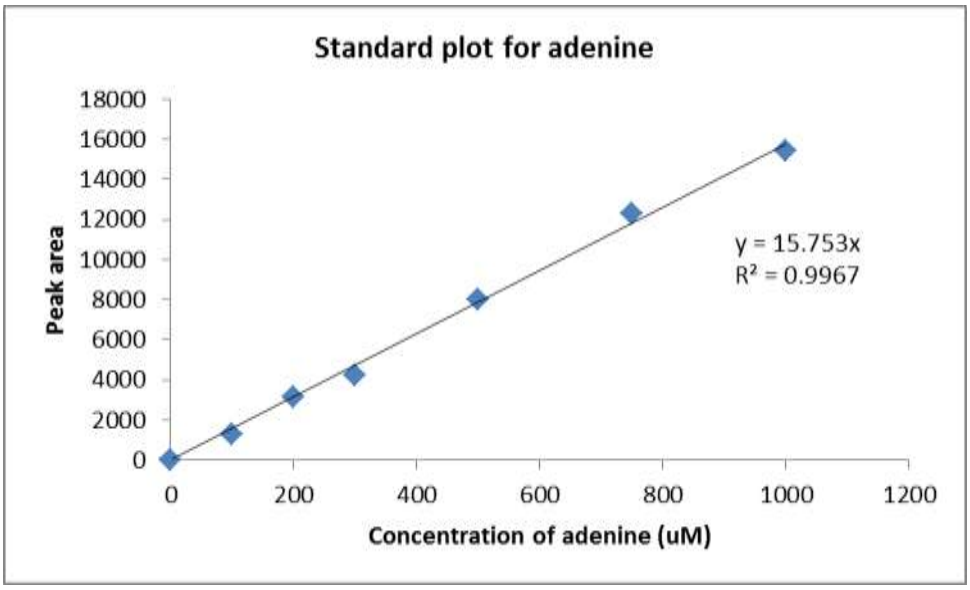



**Figure B21**: Standard curve for adenine.

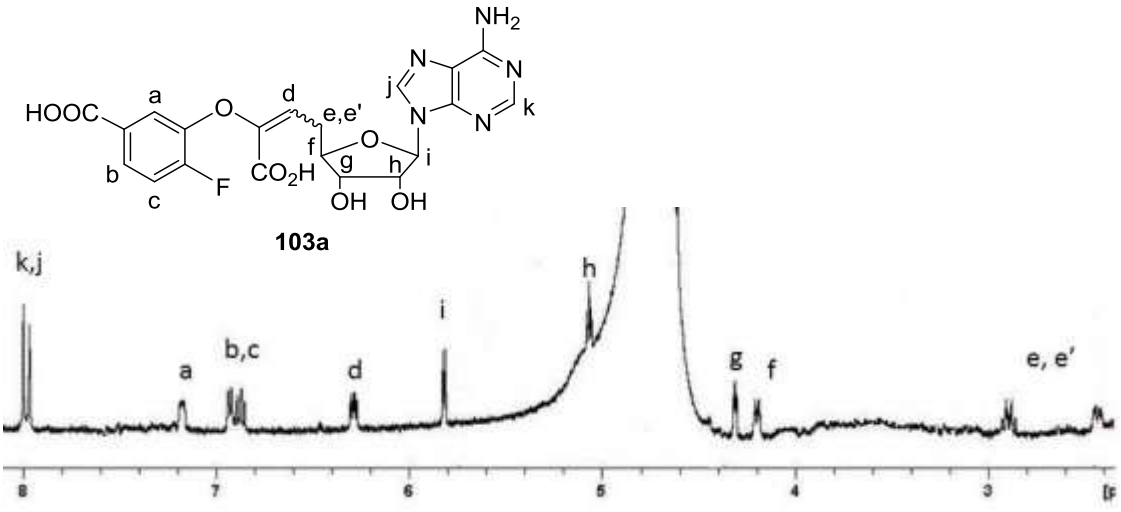



**Figure B22:** $^{1}$H NMR (500 MHz, $D_2O$) spectrum of compound **103a**.

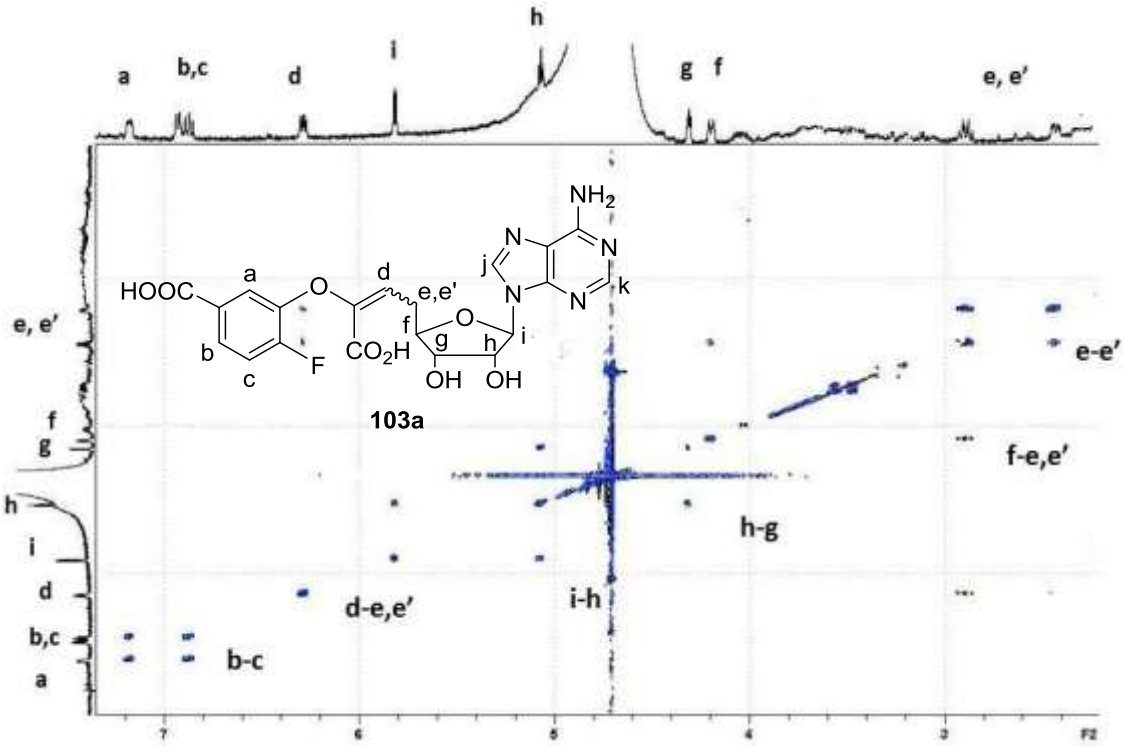



**Figure B23:** $^{1}$H-COSY spectrum of compound **103a**

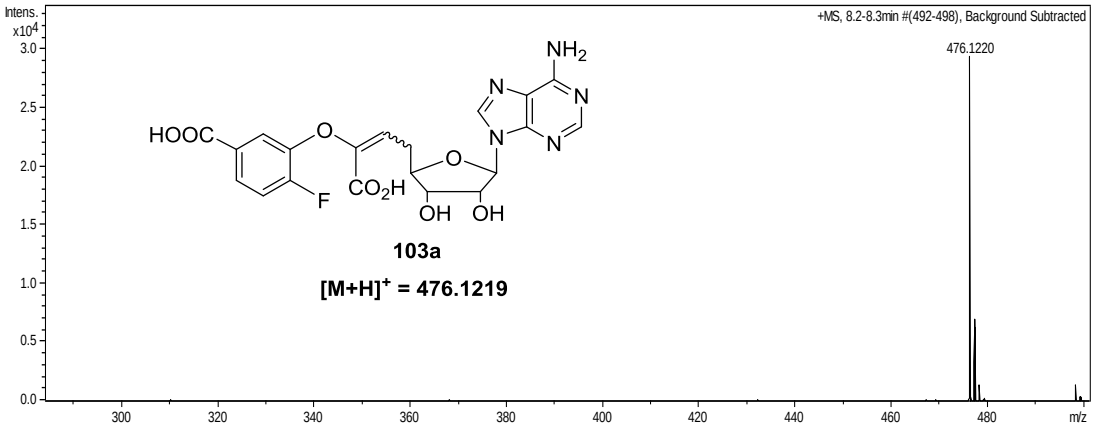



**Figure B24**: LC-MS (ESI, positive ion mode). In the mass spectrum, m/z 476.1219, $[M+H]^{+}$ corresponds to compound **103a** (theoretical m/z 476.1219)

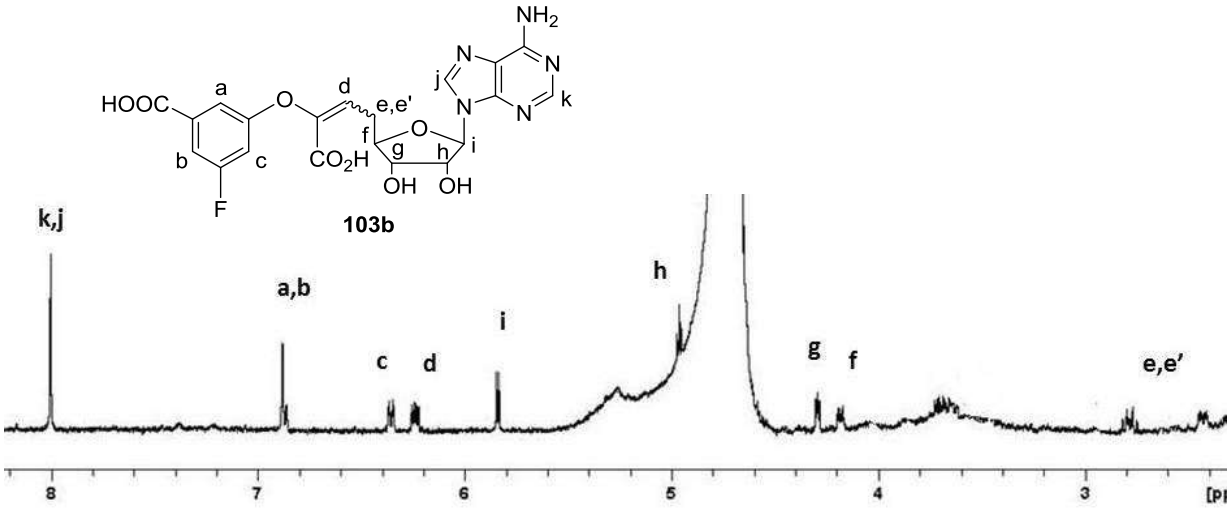



**Figure B25:** $^{1}$H NMR (500 MHz, $D_2O$) spectrum of compound **103b.**

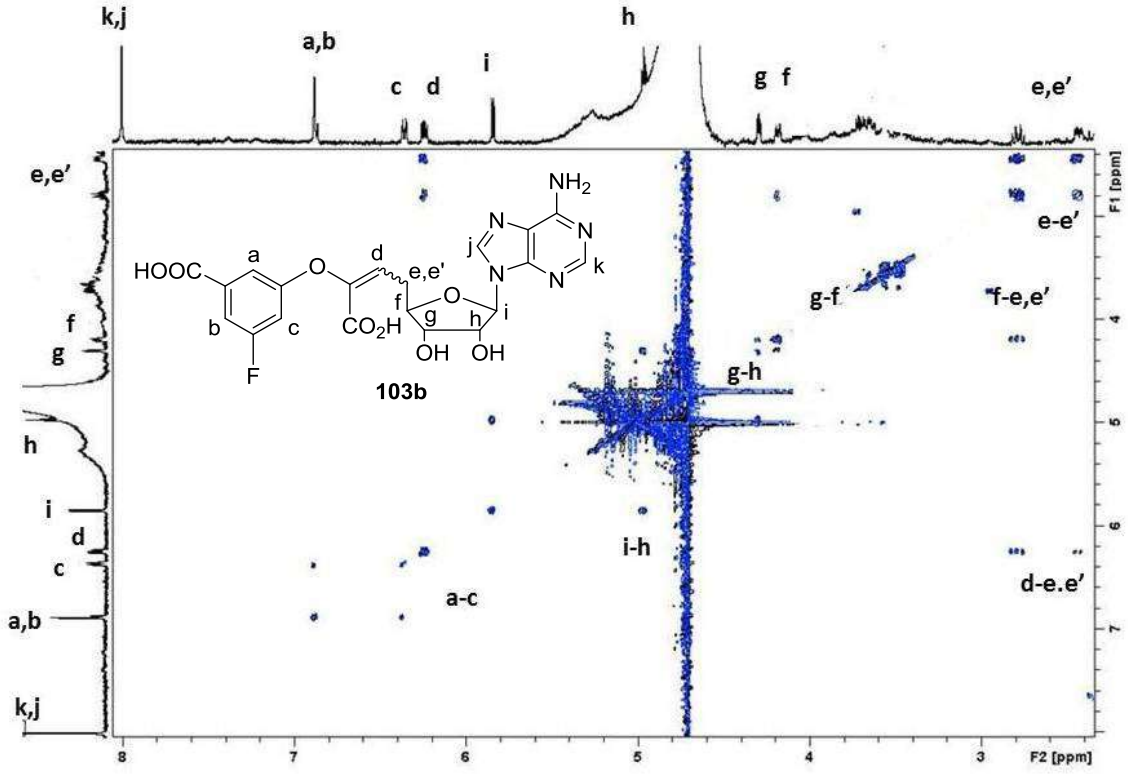



**Figure B26:** $^{1}$H-COSY spectrum of compound **103b**

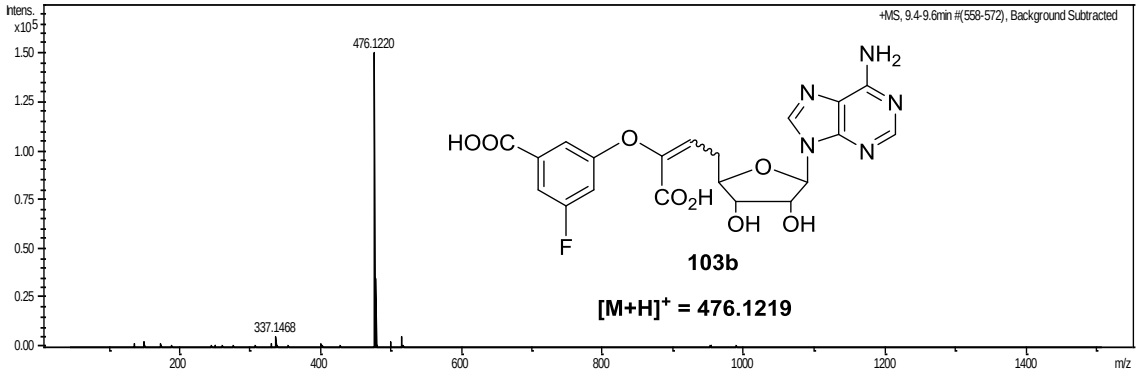



**Figure B27:** LC-MS (ESI, positive ion mode). In the mass spectrum, m/z 476.1219, $[M+H]^+$ corresponds to compound **103b** (theoretical m/z 476.1219)

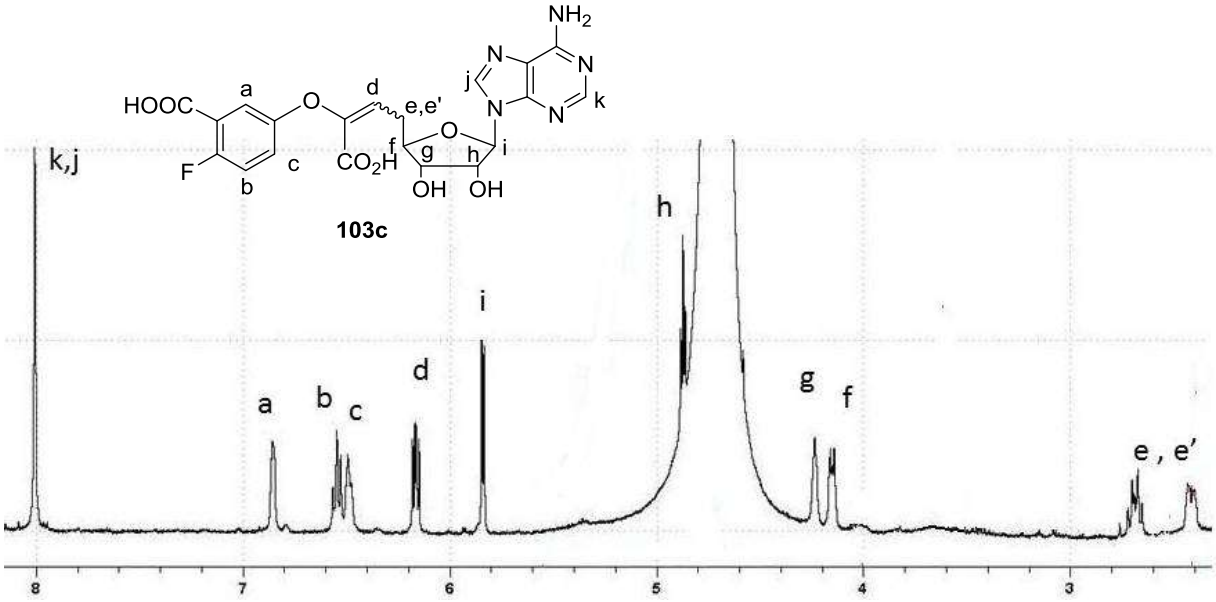



**Figure B28:** $^1H$ NMR (500 MHz, $D_2O$) spectrum of compound **103c**.

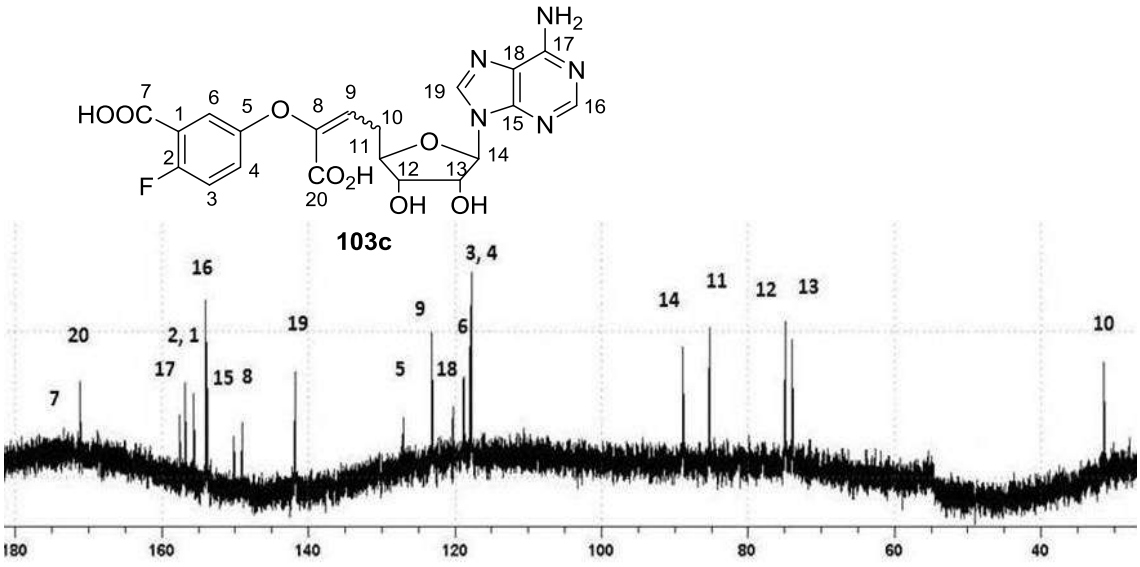



**Figure B29:** $^{13}C$ NMR (125 MHz, $D_2O$) spectrum of compound **103c**

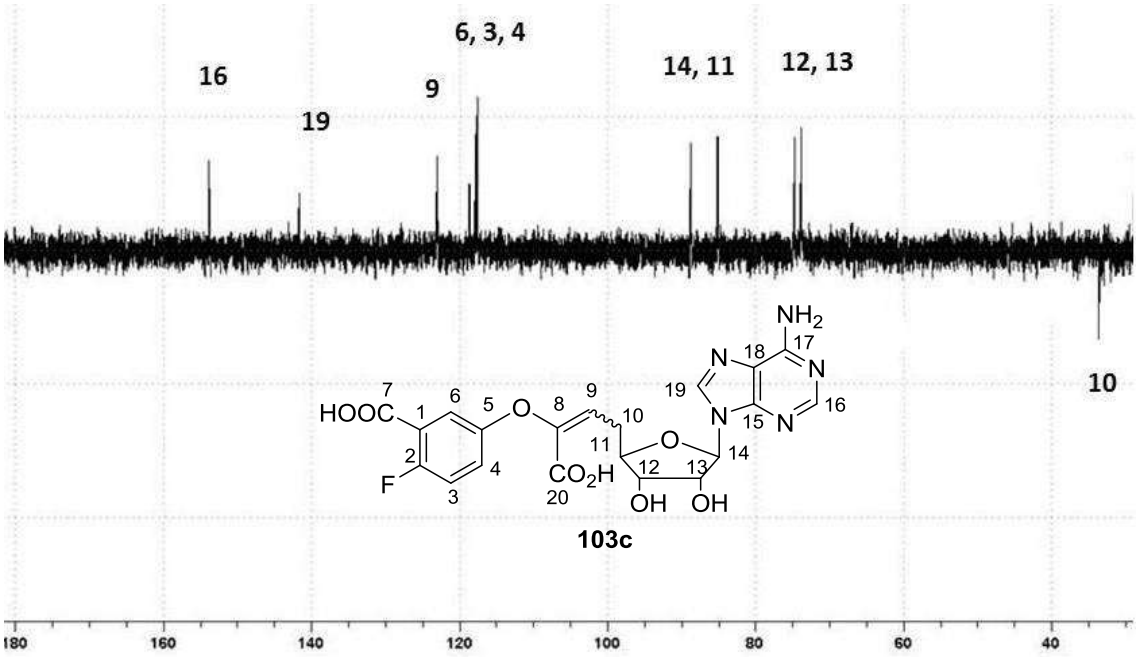



**Figure B30:** $^{13}C$ DEPT-135 NMR (125 MHz, $D_2O$) spectrum of compound **103c**.

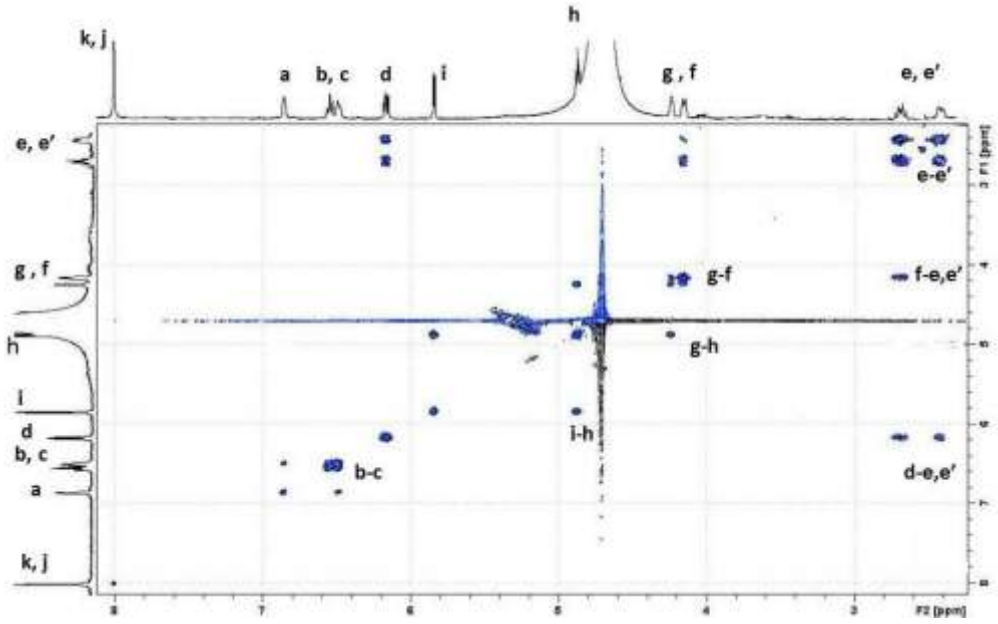



**Figure B31:** $^{1}$H-COSY spectrum of compound **103c**

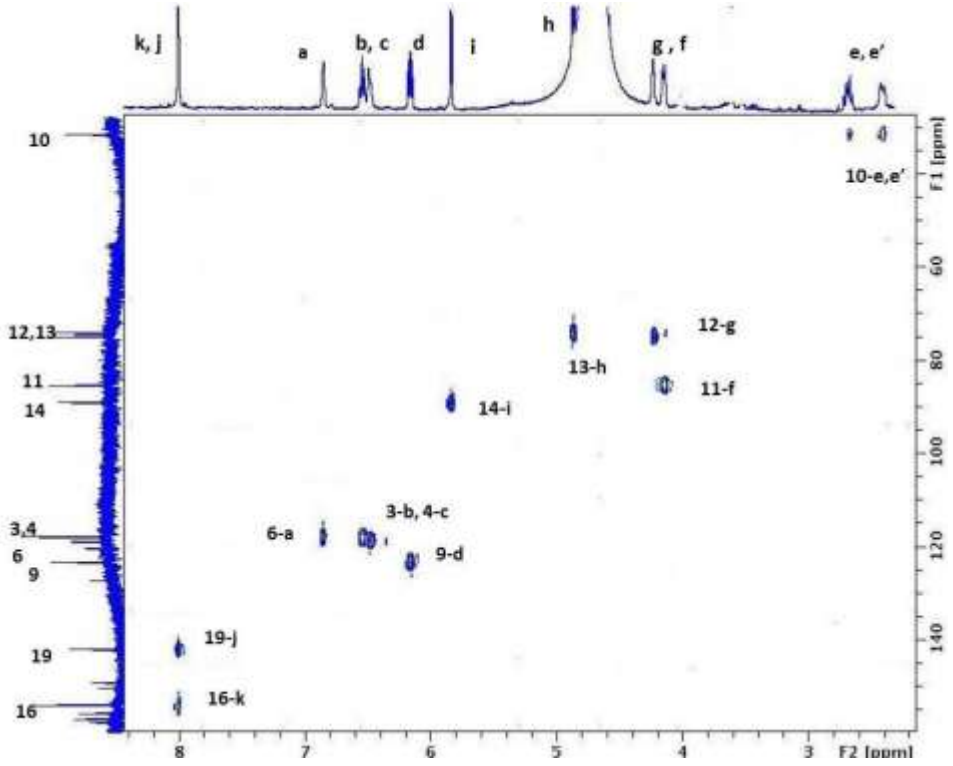



**Figure B32:** HSQC spectrum of compound **103c.**

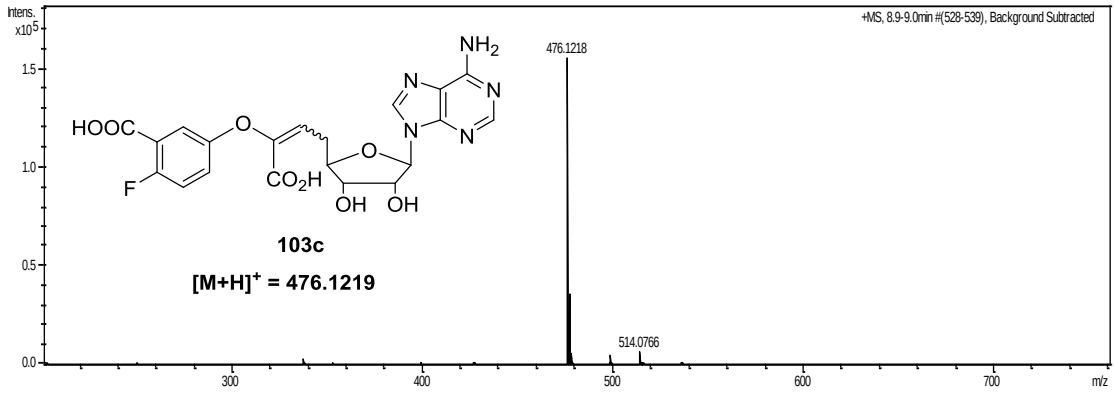



**Figure B33:** LC-MS (ESI, positive ion mode). In the mass spectrum, m/z 476.1218, $[M+H]^+$ corresponds to compound **103c** (theoretical m/z 476.1219)

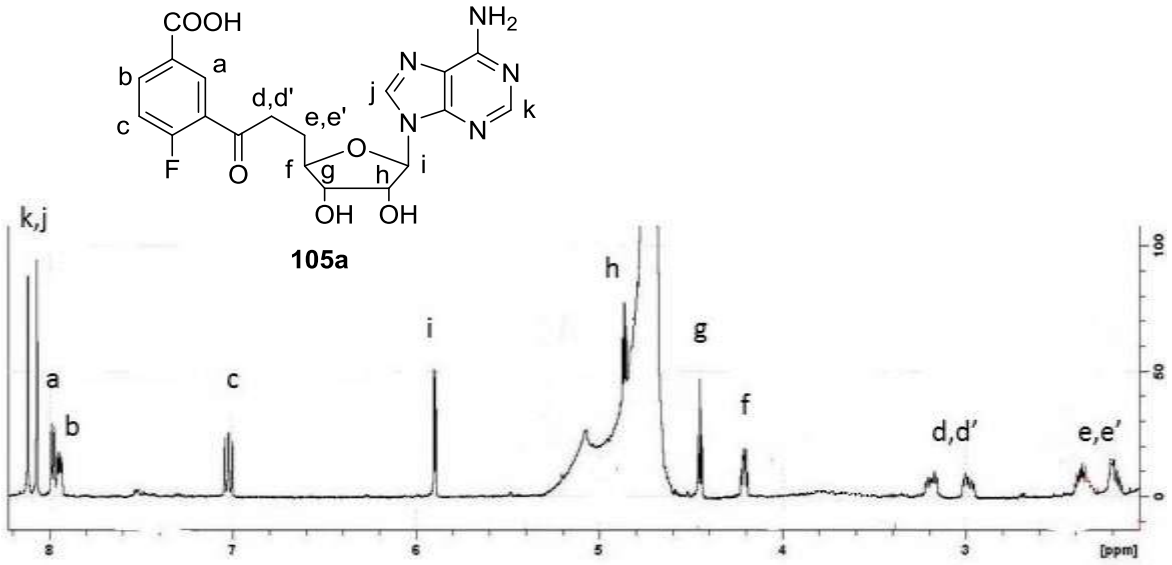



**Figure B34:** $^1H$ NMR (500 MHz, $D_2O$) spectrum of compound **105a**.

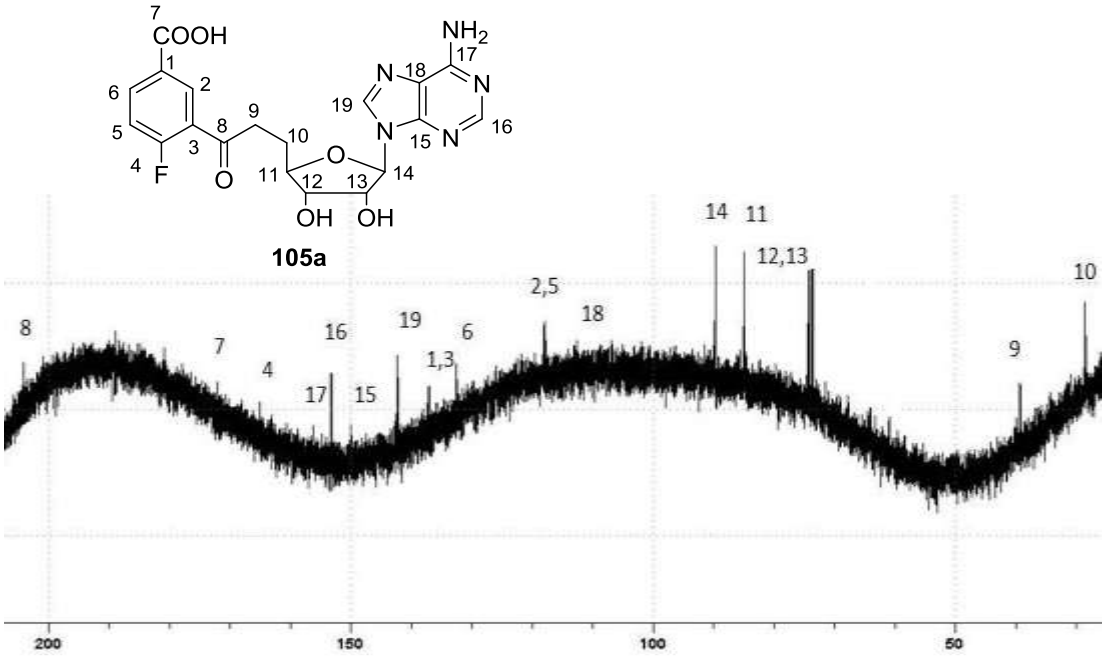



**Figure B35:** $^{13}$C NMR (125 MHz, $D_2O$) spectrum of compound **105a**

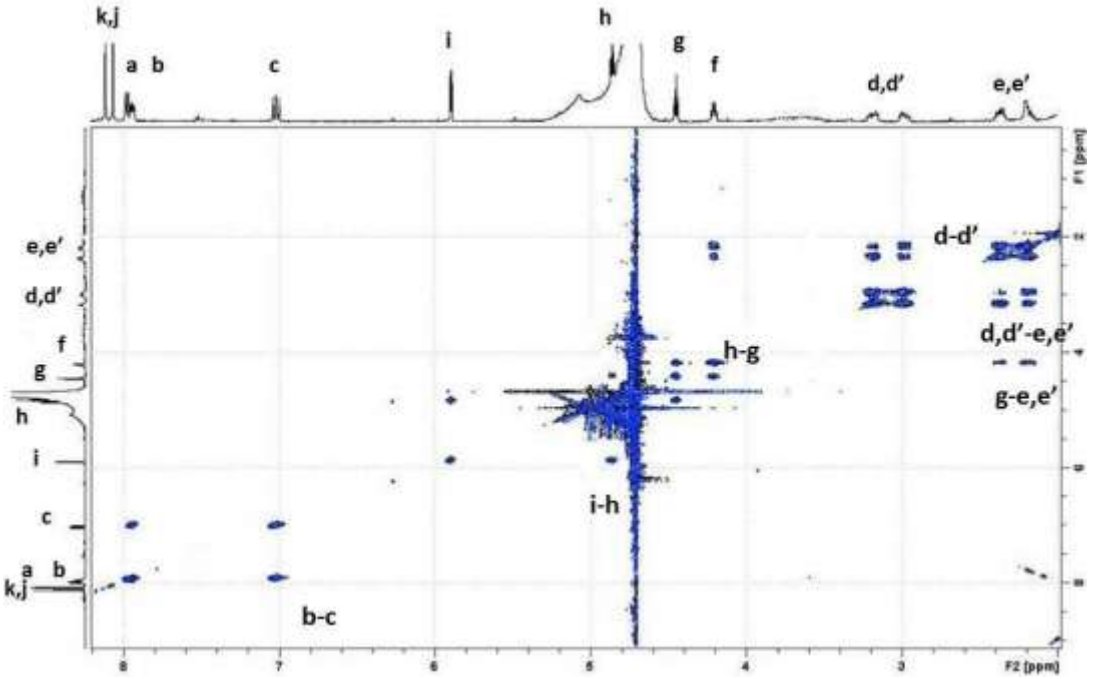



**Figure B36:** $^{1}$H-COSY spectrum of compound **105a**

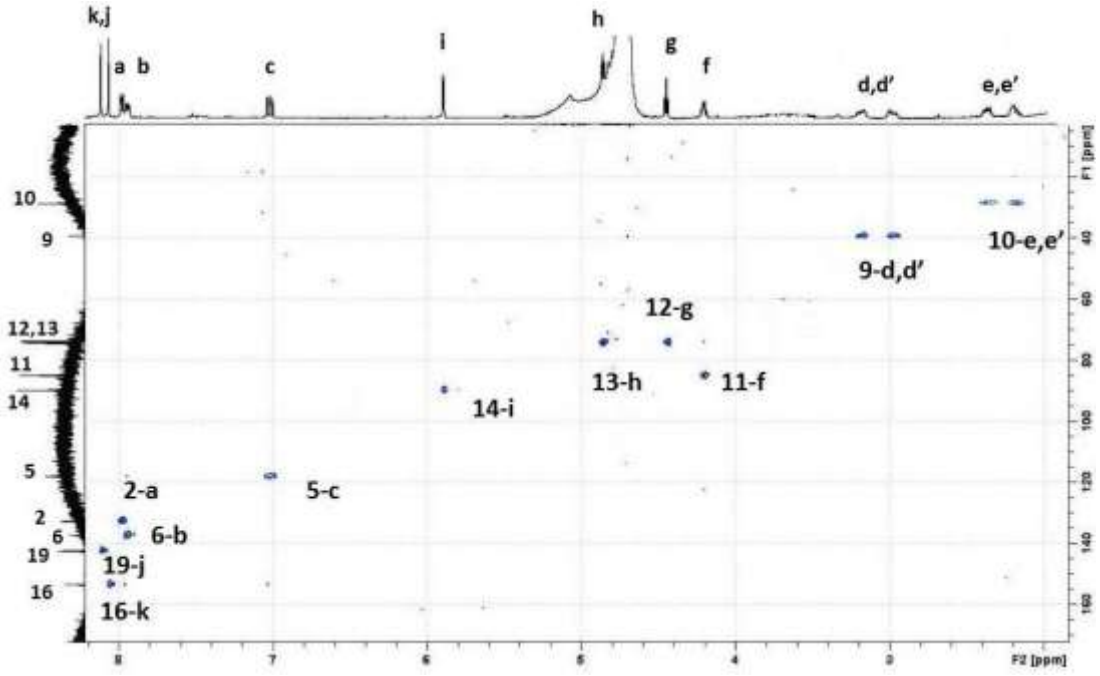



**Figure B37:** HSQC spectrum of compound **105a.**

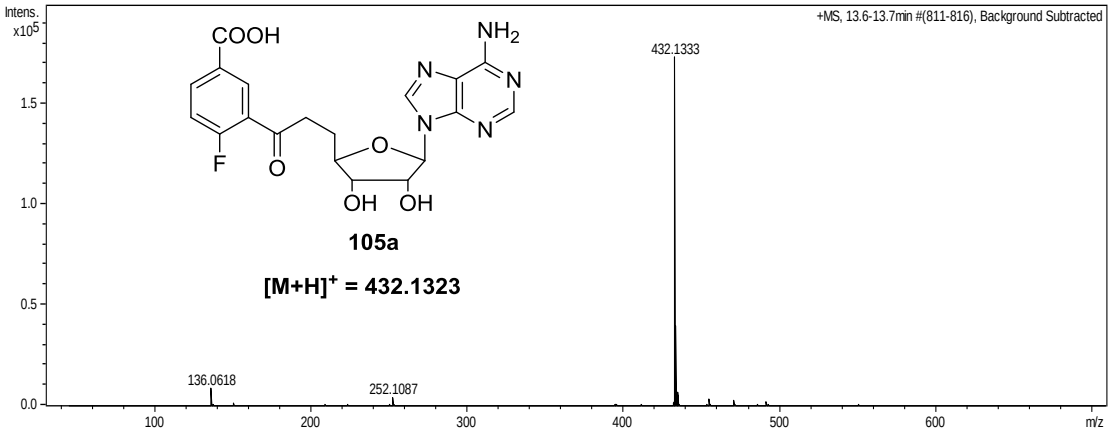



**Figure B38:** LC-MS (ESI, positive ion mode). In the mass spectrum, m/z 432.1333, $[M+H]^+$ corresponds to compound **105a** (theoretical m/z 432.1323)

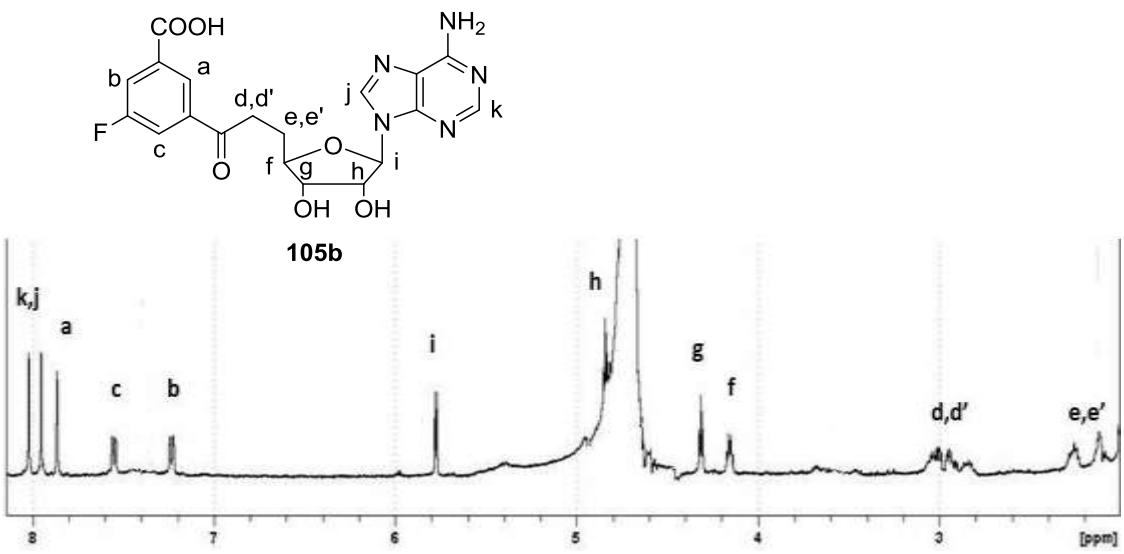



**Figure B39:** $^{1}$H NMR (500 MHz, $D_2O$) spectrum of compound **105b**.

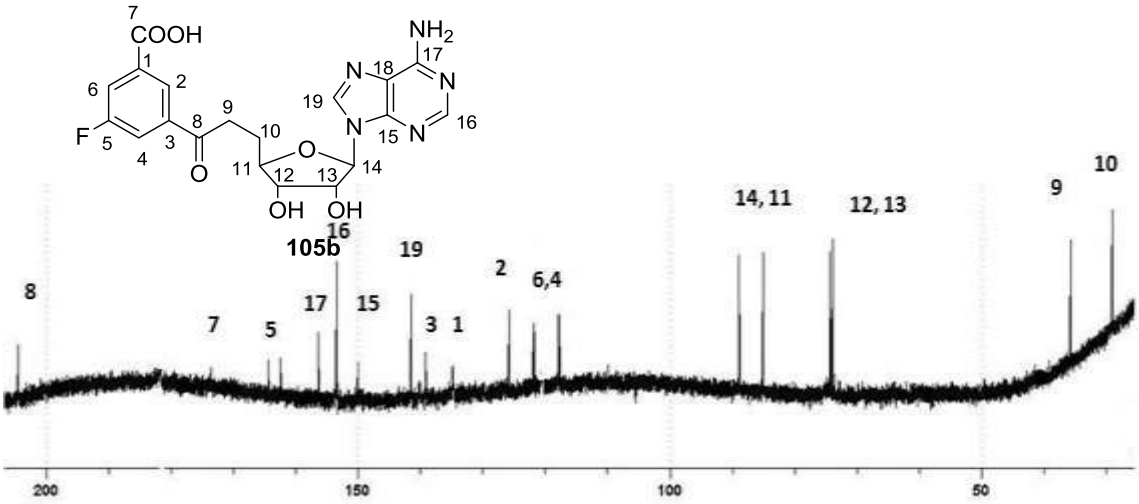



**Figure B40:** $^{13}$C NMR (125 MHz, $D_2O$) spectrum of compound **105b**

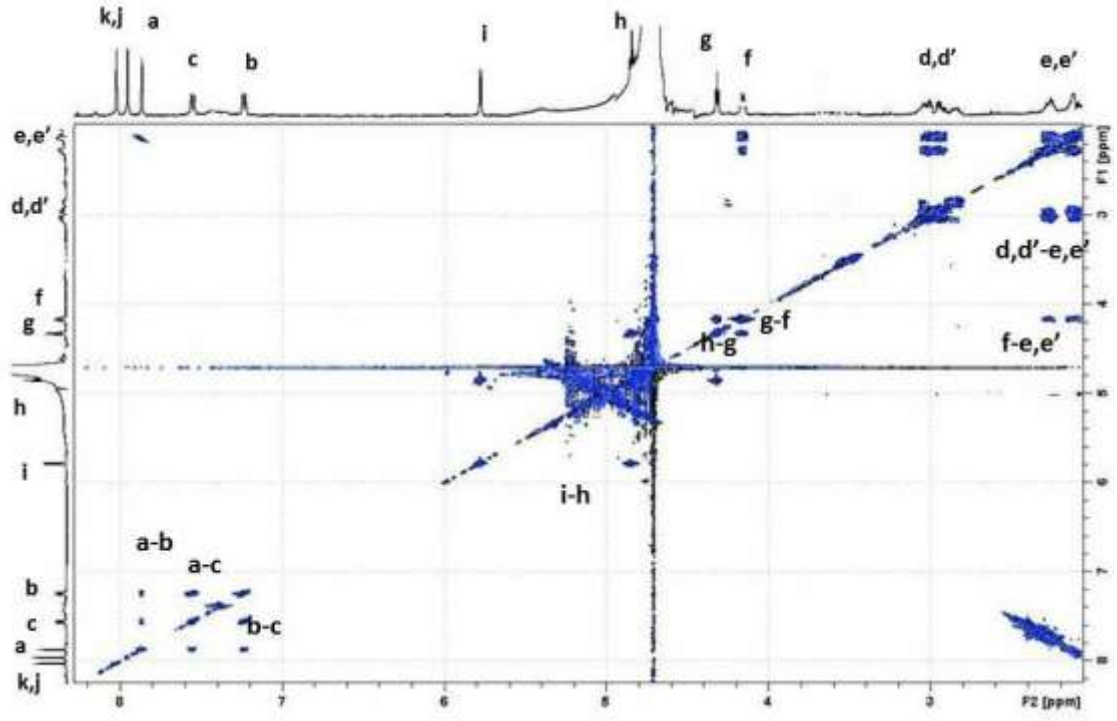



**Figure B41:** $^{1}$H-COSY spectrum of compound **105b**

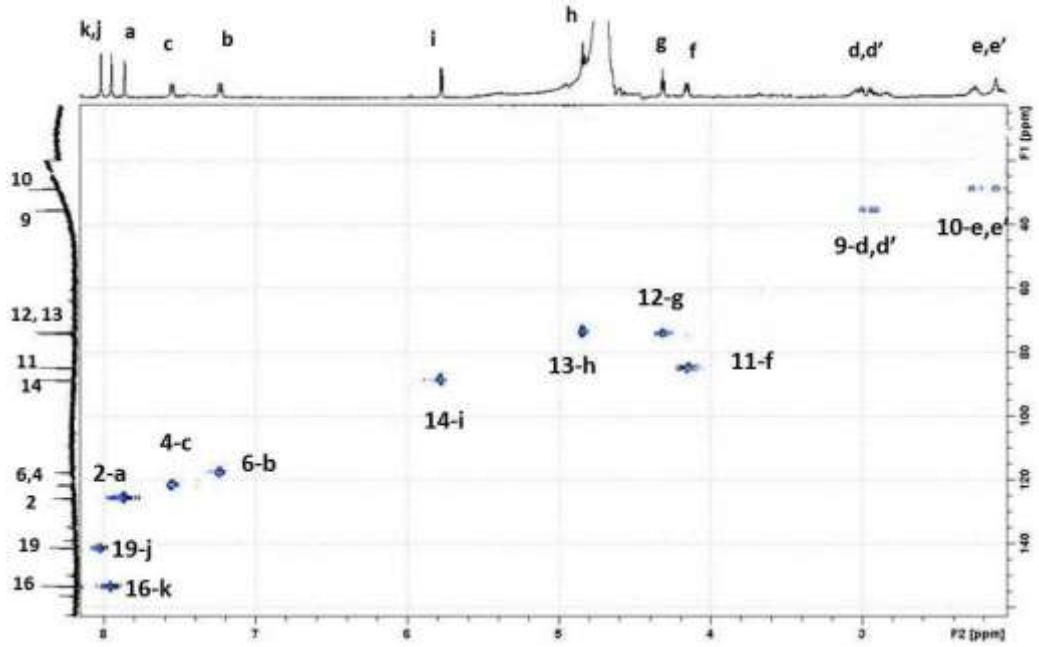



**Figure B42:** HSQC spectrum of compound **105b.**

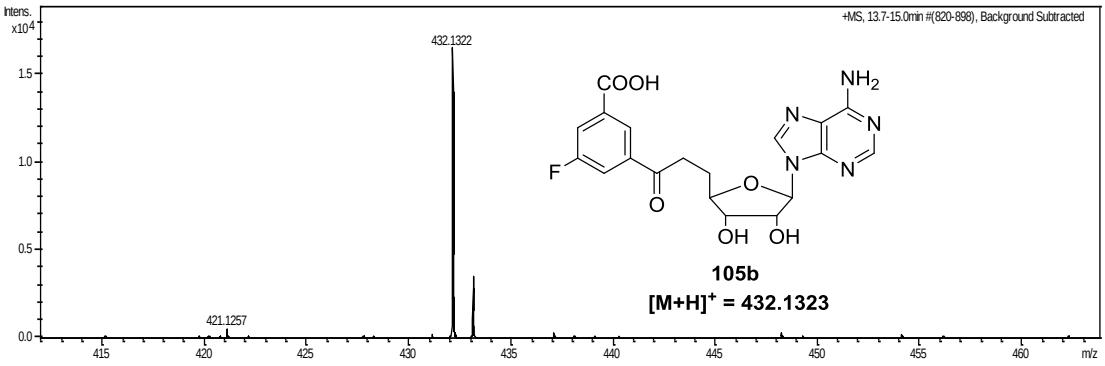



**Figure B43:** LC-MS (ESI, positive ion mode). In the mass spectrum, m/z 432.1322, $[M+H]^+$ corresponds to compound **105b** (theoretical m/z 432.1323)

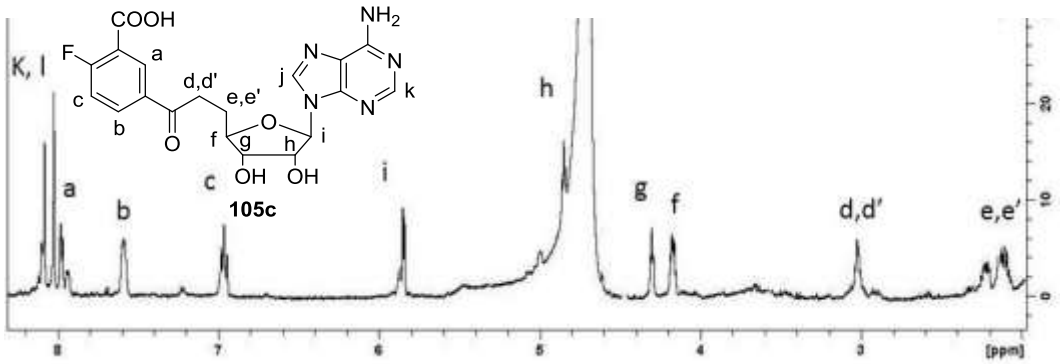



**Figure B44:** $^1$H NMR (500 MHz, $D_2O$) spectrum of compound **105c**.

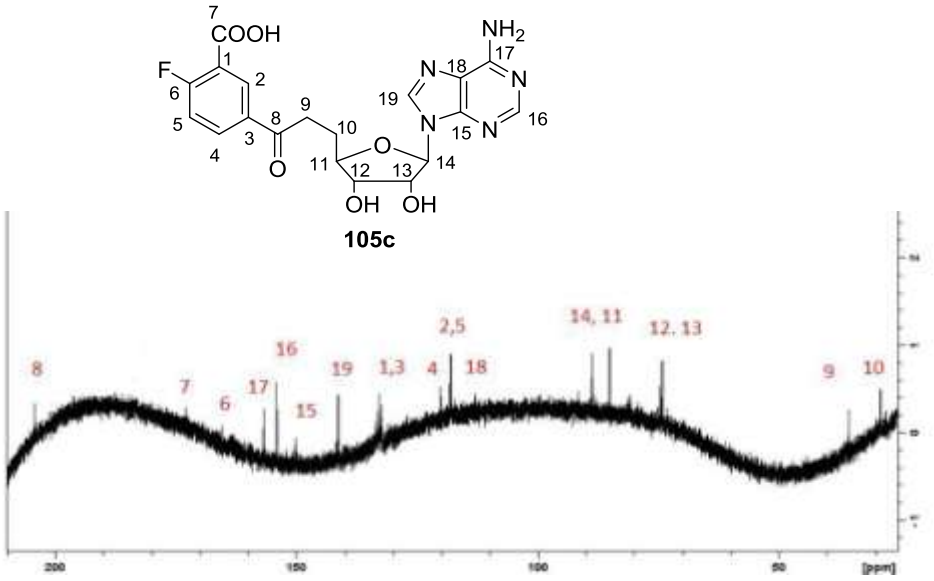



**Figure B45:** $^{13}C$ NMR (125 MHz, $D_2O$) spectrum of compound **105c**

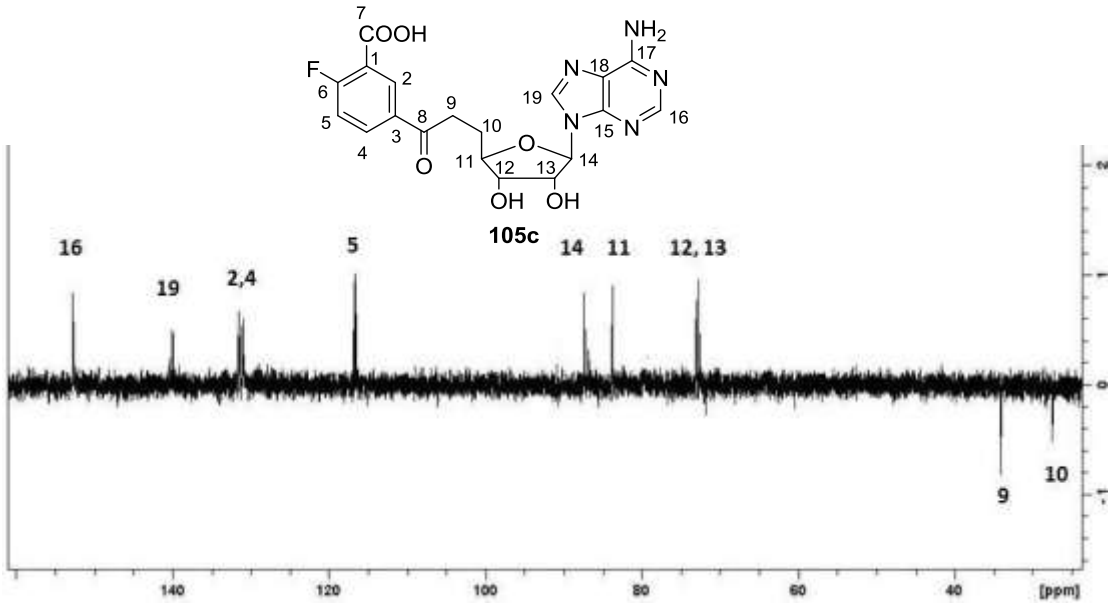



**Figure B46:** $^{13}C$ DEPT 135 NMR spectrum of compound **105c**

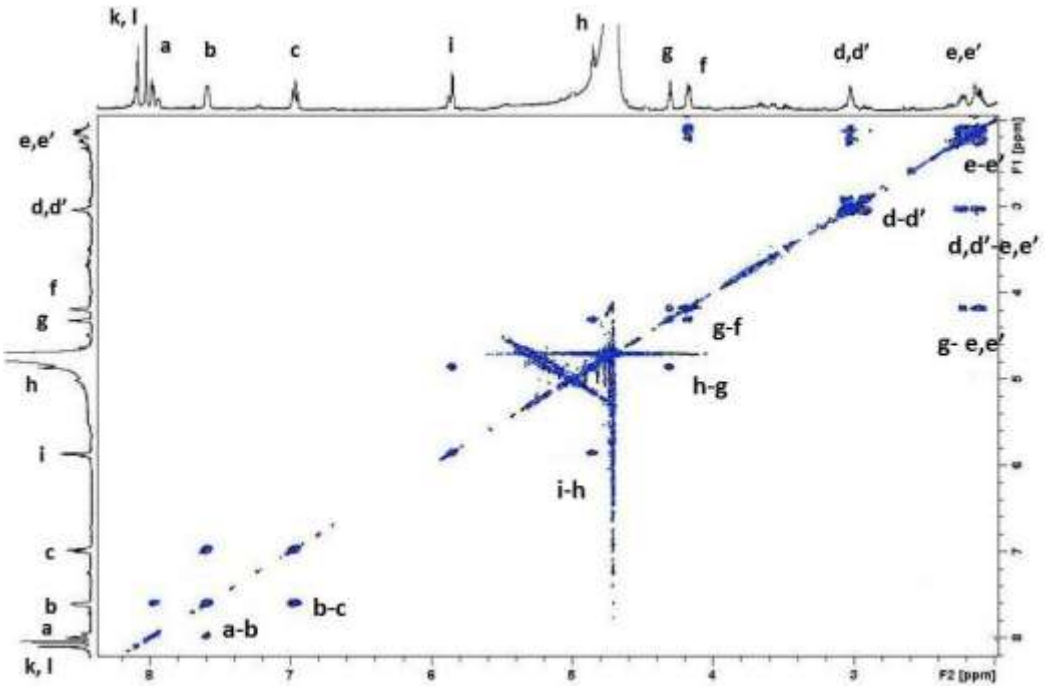



**Figure B47:** $^{1}$H-COSY spectrum of compound **105c**

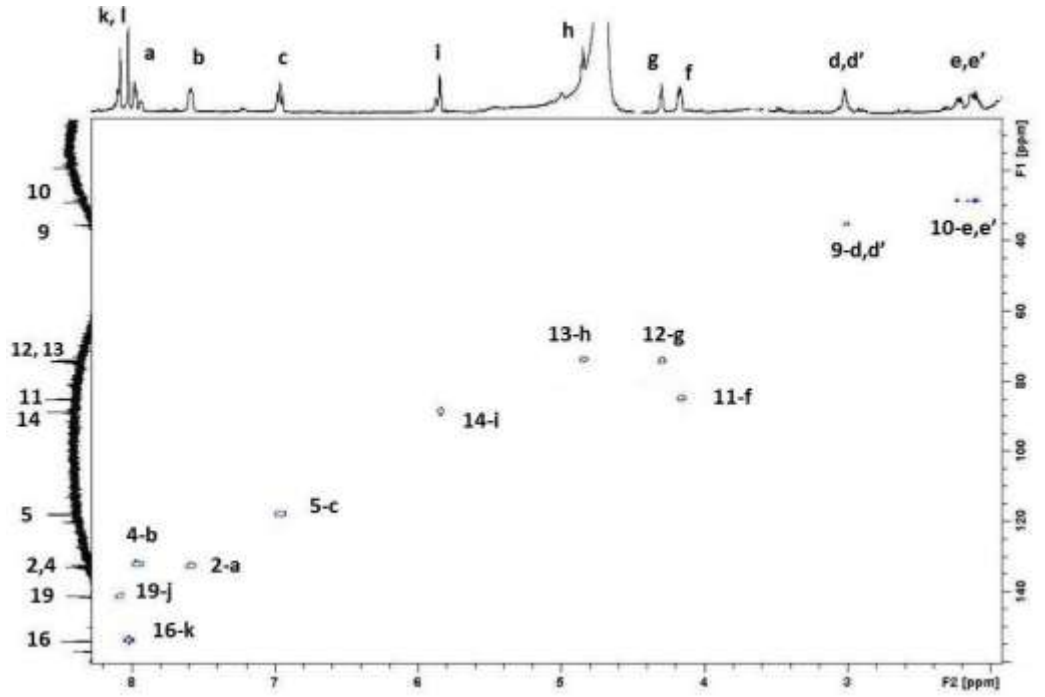



**Figure B48:** HSQC spectrum of compound **105c.**

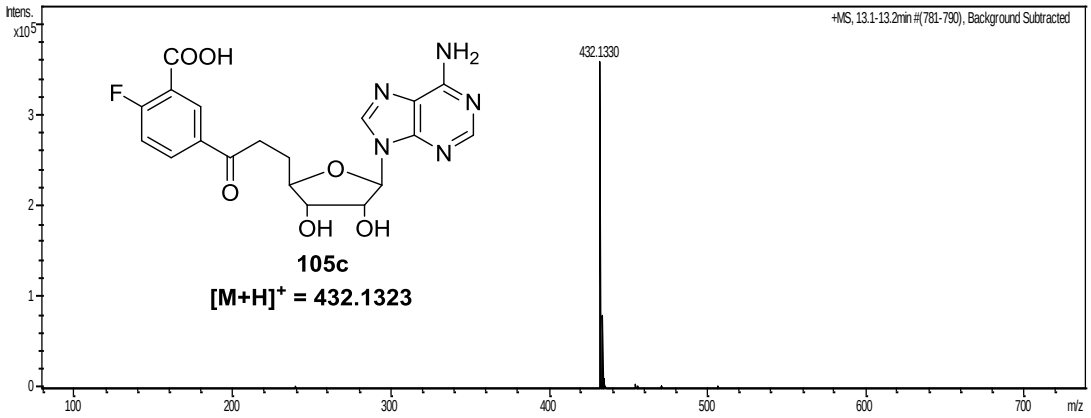



**Figure B49:** LC-MS (ESI, positive ion mode). In the mass spectrum, m/z 432.1330, $[M+H]^+$ corresponds to compound **105c** (theoretical m/z 432.1323).

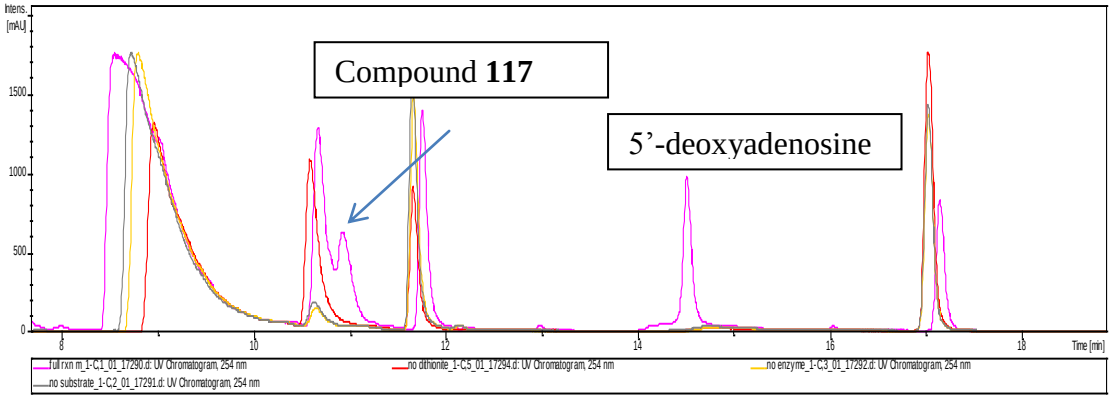



**Figure B50**:LC-MS chromatogram of the MqnE reaction mixture using **114** as substrate. Pink trace: Full reaction (114 + MqnE+ SAM + dithionite), Red trace: no dithionite control, Yellow trace: No enzyme control, Gray trace: no substrate control reactions (The compound eluting at 14.5 min is 5'-deoxyadenosine,**116**)

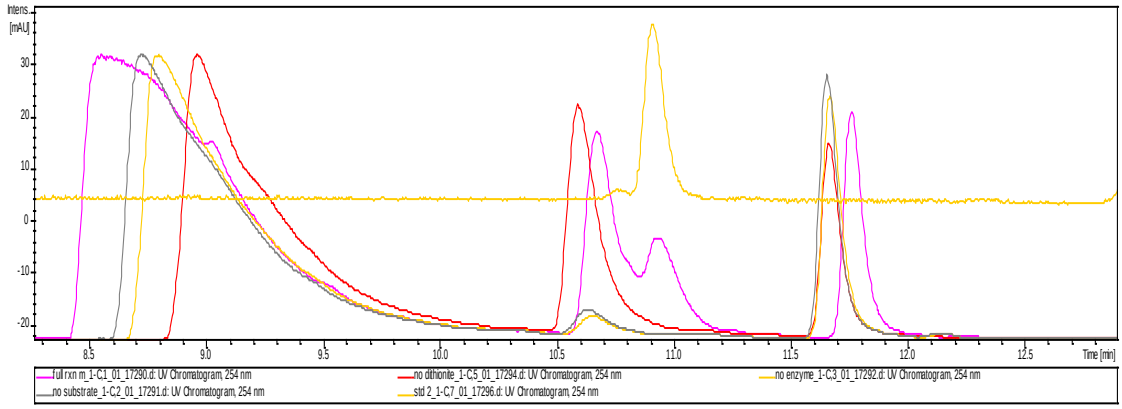



**Figure B51**:LC-MS chromatogram of the MqnE reaction mixture using **114** as substrate (The compound eluting at 11 min is identified as **117** by comparing with a reference standard as shown by the yellow trace)

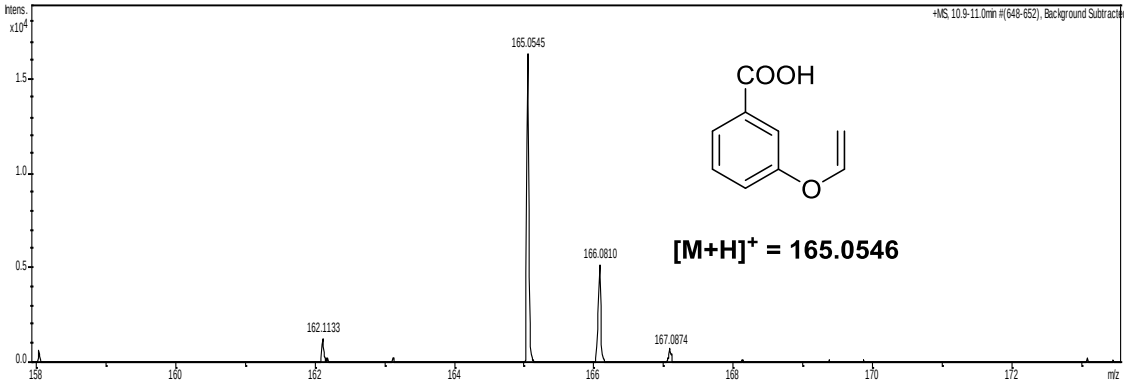
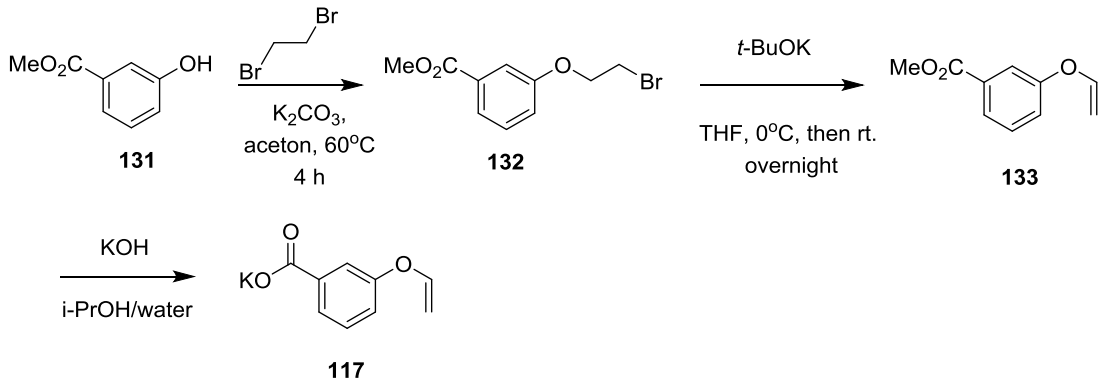



**Figure B52**: LC-ESI MS of reference standard (m/z 165.0547 $[M+H]^+$)




**Scheme B2**: Synthetic scheme for compound **117**.

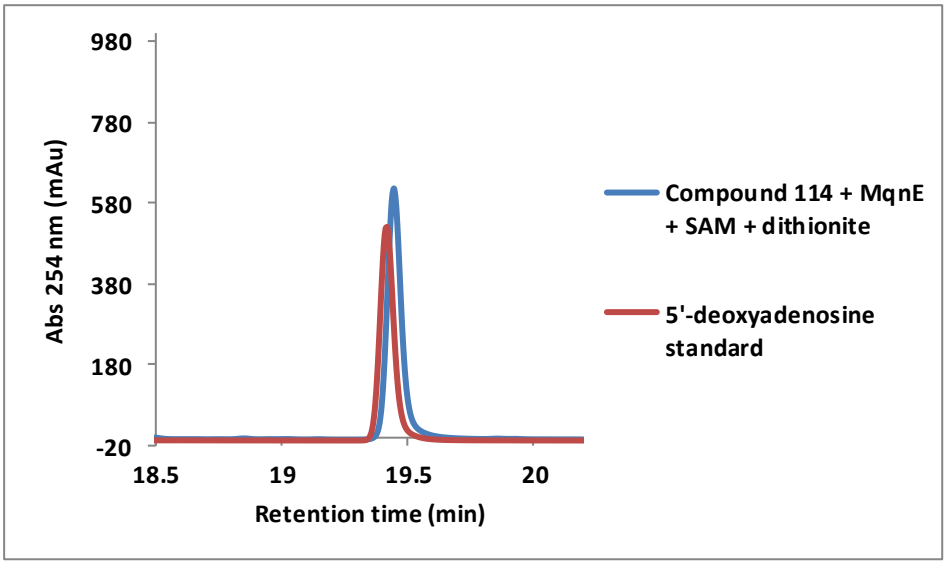



**Figure B53**: Co-migration with an authentic standard of 5'-deoxyadenosine in the MqnE assay with substrate analogue, **114**.

**Facial selectivity of the adenosyl radical addition**:

Since the methyl group of **114** must be pointing at the 5'-deoxyadenosyl radical **67** for hydrogen abstraction to occur, hence identifying the correct enantiomer of **114** (*R* or *S*) that is the substrate for MqnE would tell us if radical **67** adds to the *top* or *bottom* face of the double bond of substrate **53** in the native reaction (Figure B54).

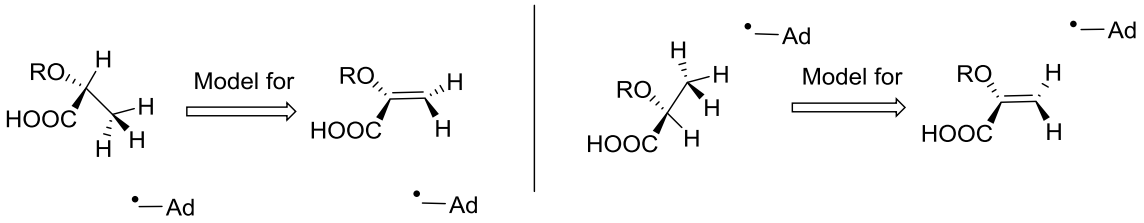



**Figure B54**: A stereochemical model for the radical addition reaction

To investigate this hypothesis, racemic **114** were treated with diazomethane in diethyl ether to convert it to the methylated derivatives, **118** and the resulting protected enantiomers were separated on a normal phase chiral HPLC column, the enantiomers eluting at 7 min and 11.5 min in 1:1 ratio (Figure B55).

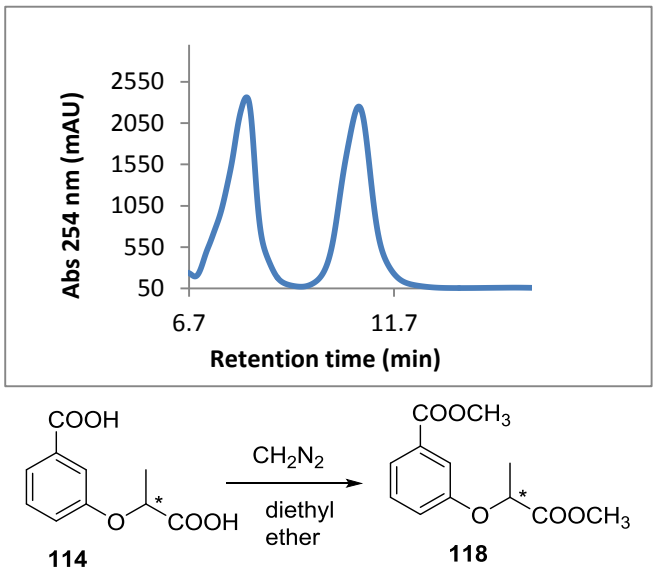



**Figure B55**: Separation of enantiomers of the racemic methylated derivatives, **118**.

In the next experiment, the racemate substrate **114** (1.5 mM) was treated with MqnE (100 uM) in presence of SAM (2 mM) and excess dithionite and reaction was allowed to proceed for about 8 hrs. The enzyme was removed by ultracentrifugation and the unreacted substrate **114** was separated from the reaction mixture by HPLC purification. This unreacted substrate was then treated with diazomethane to convert it to the methyl esters and the resulting racemic protected derivatives mixture was then run on the normal phase chiral column The ratio of enantiomers of **118** was observed to be 0.5:1 in this case(Figure B56). This suggested that the enantiomer of substrate **114** whose methyl ester elutes at 7 min in the HPLC chromatogram was preferentially consumed in the MqnE reaction. By comparing with the HPLC chromatogram of the enantiomers of the synthesized methylated derivatives (Figure B57 and B58), it was found that the *S* isomer was the preferred substrate for MqnE.

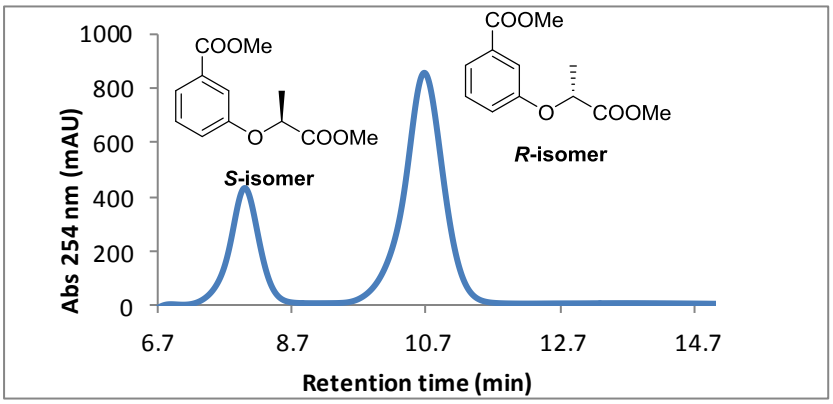



**Figure B56**: Separation of enantiomers of the methylated derivatives of the unreacted substrate after the MqnE reaction

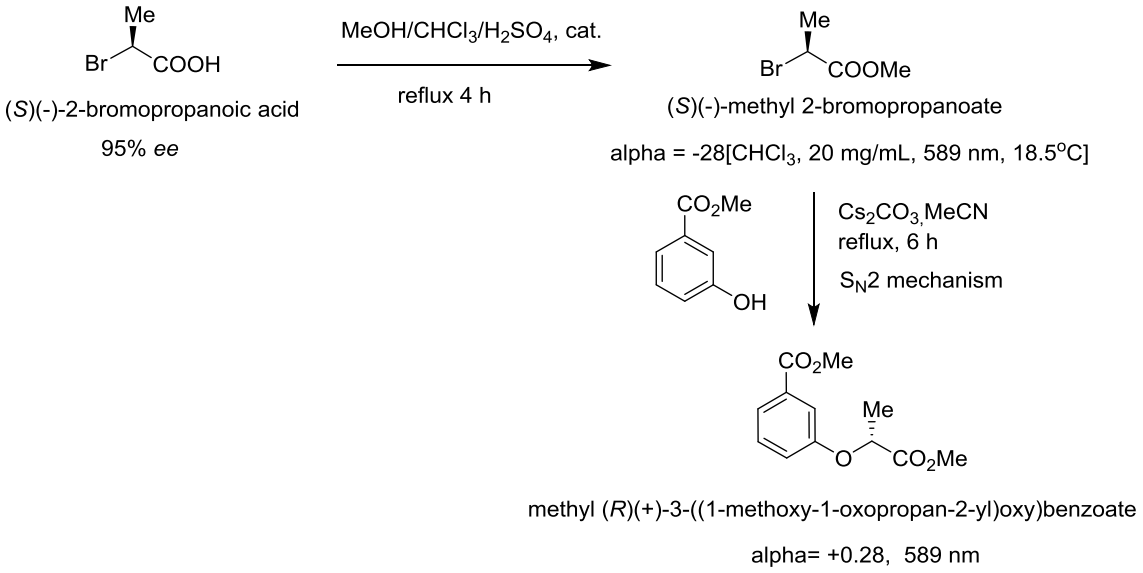



**Figure B57**: Synthetic scheme for enantiomeric methylated derivatives.

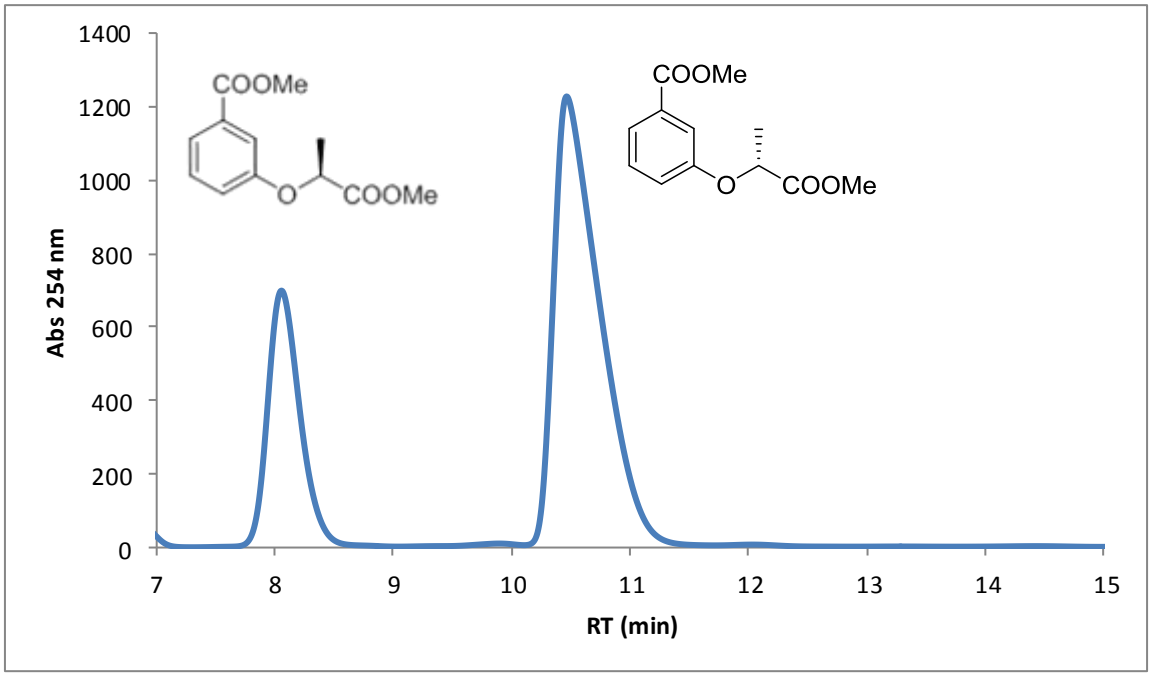



**Figure B58**: Separations of mixture enantiomers of synthesized methylated derivatives of **114** with *R* isomer being in excess (eluting at 10.8 min)

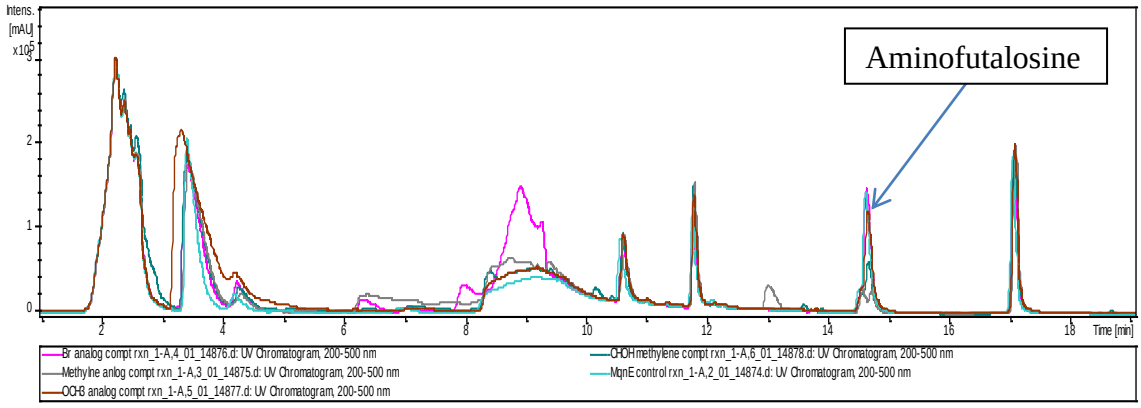



**Figure B59**: MqnE reaction with 1:1 mixture of **53** and **88** (Pink trace), **53** and **114** (Brown trace), **53** and **83** ( Grey trace) and only **53** (Blue trace). Inhibition was observed only in the case of **83** as indicated by the levels of aminofutalosine produced.

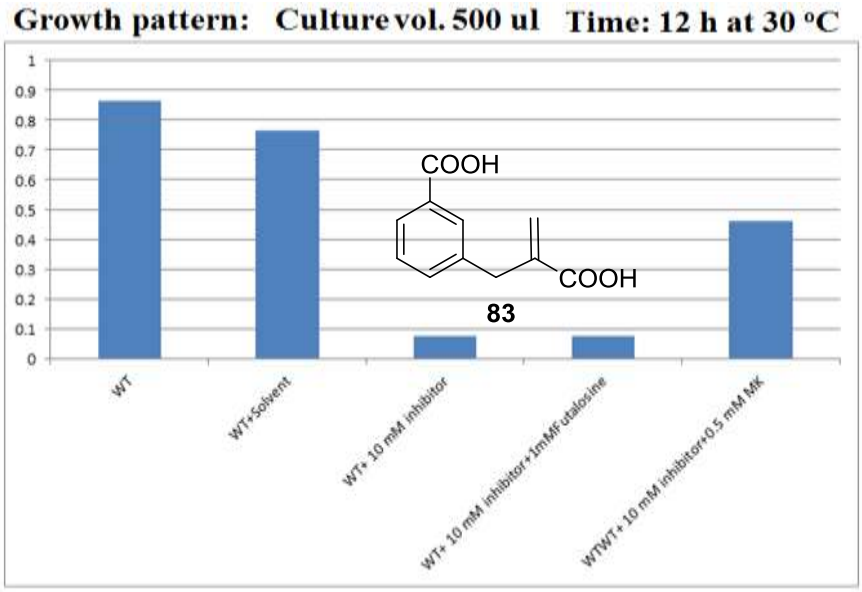



**Figure B60**: *D. radiodurans* R1 growth study (Tryptic Soy Broth, Temperatures used 30°C and 37°C) in presence of the inhibitor **83.** The growth can be recovered by addition of menaquinone (MK-4) to the culture medium (by Dinesh Simkhada, Begley lab, unpublished results).