The Dissertation Committee for Eita Sasaki Certifies that this is the approved version of the following dissertation:

Biosynthetic Studies of Thiosugar-Containing Natural Products, BE-7585A and Lincomycin A

Committee:

Hung-wen Liu, Supervisor

Eric V. Anslyn

Christopher Bielawski

Walter L. Fast

Christian P. Whitman

Biosynthetic Studies of Thiosugar-Containing Natural Products, BE-7585A and Lincomycin A

by

Eita Sasaki, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2011

Acknowledgements

It was early November in 2004 that I first met Dr. Liu in Tokyo, Japan. I was a second year graduate student in the master's program in pharmaceutical sciences at the University of Tokyo. Prof. Liu was an invited speaker for a special symposium held in the department. The next day, I joined a small group of graduate students who showed him around the city. I was interested in pursuing my doctoral studies in the Unites States. He kindly told me to contact with him if I was really interested in his research. A half-year later, on May 24, 2005, I arrived in Austin to join the Dr. Hung-wen (Ben) Liu's research group at the University of Texas at Austin.

I would like to thank Ben for introducing me to his wonderful lab. I really like the freedom to do research in this group as long as we seriously dedicate ourselves to answer important scientific questions. His suggestions from daily conversation and the group meetings stimulated my research interest and helped me to keep my research on the right track. I would also like to thank Yung-nan, who is extremely kind and takes care of everybody in the lab. She also helped me to express some of the proteins for my project.

When I joined Dr Liu's lab, I met nine postdocs, Chai-Lin, Ron, Sung-Ju, Qingquan, Tun-Cheng, Kenji, Gang, Feng, and Sveta, and twelve graduate students, Lin, William, Peng, Allen, Chad, Jeff, Ping-Hui, Zhihua, Ying, Jess, Chris, and Hak Joong. They were all very nice people. Especially, I would like to thank Chad and Sveta for teaching me many biochemical techniques in my first year. Without Chad, I may not have studied thiosugar biosynthesis because the BE-7585A project was initially introduced by him. I also thank Jess and Chris for kindly proofreading my progress reports and meeting abstracts. I would like to thank Ping-Hui and Hak Joong for giving me helpful materials

and information for coursework. Finally, I would like to thank Kenji, a nice neighbor, for helping me to set up life in Austin and share good times with other friends on weekends. All of the members introduced above, including Xiaotao and Bryan, who joined the group after me, have already left the group. I hope to meet them again sometime in the future.

I would like to thank current members of Dr. Liu's lab, too. We have five postdocs, Yasushi, Steve, Mark, Hui, and Jordi, and thirteen graduate students, Wei-Chen, Sei, Mel, Grace, Anthony, Chia-I, Ke-Yi, Meilan, Eta, Nam Ho, Hak, Byung-Sun, and Yeonjin. Yasushi showed me many biochemical experiments and gave me useful research suggestions. Steve, Mark, and Anthony kindly proofread this dissertation and were always helpful in discussions about my research projects. Chia-I joined the lincomycin A project and provided me the synthetic standards necessary to characterize the enzymatic products. Ke-Yi helped me to express some proteins for my project during his rotation period in this laboratory. Grace was a good bench neighbor to have—we shared some experimental materials and learned the language of each other's country. I would also like to thank an excellent undergraduate researcher, David, who assisted my research project for two years and graduated in 2010.

I am also grateful to Banyu Pharmaceutical Co. (Tokyo, Japana) for kindly providing the *Amycoloatopsis. orientalis* subsp. *vinearia* BA-07585 strain for the BE-7585A project and to Prof. Jürgen Rohr at the University of Kentucky for generously giving us urdamycin A to demonstrate C–S bond formation between the angucycline core and a thiosugar. Dr. Ting-Jen R. Chen at the Genomics Research Center of Academia Sinica (Taiwan) helped us to obtain the draft sequence of the whole genome of the *A. orientalis* strain. Lastly, I would also like to thank some people working in the core facilities of the University of Texas at Austin for providing spectral data and giving me

valuable technical suggestions, including Dr. Ben Shoulder, Mr. Steven D. Sorey, and Mr. Jim Wallin in the Nuclear Magnetic Resonance Facility and Dr. Mehdi Moini, Mr. Charles R. Cartwright, and Dr. Karin Keller in the Mass Spectrometry Facility of the Department of Chemistry & Biochemistry, Dr. Stony H. Lo in the Analytical Instrumentation Facility of the College of Pharmacy, and Mr. Cecil W. Harkey in the DNA Sequencing Facility of the Institute for Cellular and Molecular Biology.

I have been supported by many friends I met in Austin. Aiko Umeda, Taki Adachi, and Tomo Umemura are the three best friends I enjoy hanging out with when we have free time. I would also like to thank Prof. Tetsuo Nagano, Dr. Hirotatsu Kojima, and Dr. Kenjiro Hanaoka in the University of Tokyo for their long-term support and encouragement. Finally, I sincerely appreciate my parents' understanding, support, and encouragement for my decision to pursue my studies in the United States. I would like to dedicate this dissertation to my parents in return for their belief in my success.

Thank you very much. Apr. 17, 2011

Eita Sasaki

Biosynthetic Studies of Thiosugar-Containing Natural Products, BE-7585A and Lincomycin A

Publication No._____

Eita Sasaki, Ph.D.

The University of Texas at Austin, 2011

Supervisor: Hung-wen Liu

Sulfur is an essential element found ubiquitously in living systems. However, there exist only a few sulfur-containing sugars in nature and their biosyntheses have not been well understood. On the other hand, a wide variety of sugar derivatives commonly found in natural products are often vital components for the efficacy and specificity of their parent molecules. Elucidation of such unusual sugar biosyntheses is important both for understanding their intriguing chemical mechanisms and creating unnatural compounds by altering their biosynthetic machineries, which could potentially exhibit enhanced or novel biological activities. This dissertation describes biosynthetic studies of two thiosugar-containing natural products, BE-7585A and lincomycin A, produced by *Amycolatopsis orientalis* and *Streptomyces lincolnensis*, respectively. While the former possess a C-2-thiosugar-containing disaccharide moiety, the latter contains a C-1-thio substituent on a characteristic eight-carbon backbone sugar. The focus of this research is to characterize the biological pathways and mechanisms responsible for the sulfur incorporation and the unique sugar scaffolds.

BE-7585A, an angucycline-type natural product, contains the rare C-2-thiosugar moiety. PCR-based screening of a cosmid library constructed from the genomic DNA of *A. orientalis* led to the identification of the BE-7585A biosynthetic gene cluster. A gene, *bexX*, was found to be a candidate for a thiosugar synthase with moderate sequence similarity to a thiazole synthase. The gene, *bexX*, and a glycosyltransferase homologue, *bexG2*, were heterologously expressed in *Escherichia coli*. A variety of biochemical experiments provided a wealth of evidence supporting the proposed biosynthetic pathway for the C-2-thiodisaccharide moiety. Finally, whole genome sequencing and a genome mining approach led to the identification of a sulfur carrier protein to accomplish the *in vitro* enzymatic synthesis of the C-2-thiosugar for the first time.

Lincomycin A is a lincosamide antimicrobial natural product with a C-1 methylthio substituent. Although the lincomycin A biosynthetic gene cluster has been reported, biochemical verification of the biosynthetic pathway has remained elusive. In this dissertation, the complete methlthiolincosamide biosynthetic pathway including the potential C-1 sulfur incorporation mechanism was proposed. Furthermore, two early intermediates of the pathway were characterized for the first time by demonstrating the LmbR (transaldolase) and LmbN (isomerase) reactions *in vitro*.

Table of Contents

List of Tables

List of Figures

Chapter 1: Background and Significance

1.1 INTRODUCTION

Natural products are commonly referred to small organic molecules with diverse structures produced by living organisms. Natural products often exhibit biological activities and play important physiological roles in the producer organisms.[1-5] However, the exact biological functions of most natural products or secondary metabolites remain to be determined.[2] Natural products have been a rich source for drug discovery. It is well documented that extracts of natural products were used as medicine by people in Mesopotamia (2600 BC), Egypt (1500 BC), China (1100 BC), India (1000 BC) and Greece (300 BC) long time ago.[6,7] Since the discovery of penicillin in 1928, research on antibiotics produced by microorganisms has attracted much attention. Extensive screening of microorganisms during this era led to the discovery of many antibiotics, antiparasitics, antimalarials, lipid control agents, immunosuppressants, and anticancer drugs.[8] More recently, advanced technologies including combinatorial synthesis and high-throughput screening have changed the drug discovery research from screening of natural products to screening of synthetic small molecule libraries. Nevertheless, natural products and their derivatives remain an important source of finding useful lead compounds. Indeed, 34% of the new drugs developed in small molecule entities between 1981 and 2006 are natural products or their semisynthetic derivatives.[8-10]

In addition to their pharmacological applications, studies of the chemistry of natural product syntheses have been a major area in chemical research for a long time.[10] Total synthesis of complex natural products is still a major research focus in current organic chemistry. Numerous new reactions and strategies have been developed through

these studies.[11] On the other hand, in nature, construction of complex structures of natural products is mediated by enzymes. The regio- and stereospecific enzyme-catalysis ensures the production of complex natural products in a very efficient and specific manner.

In this post-genomic era, metabolic engineering and synthetic biology have become powerful tools useful for drug development. In this context, thorough understanding of natural product biosyntheses including genetic, enzymatic, and mechanistic information could enable one to manipulate these biosynthetic machineries to generate new structural entities. Such combinatorial biosynthetic approach holds high potential in drug discovery in the future.

1.2 Unusual Sugar Biosynthesis and Glycodiversification

Carbohydrates in the Biological System

Carbohydrates are one of the fundamental class of materials in the biological systems. Living organisms store sugar molecules as polysaccharides, glycogen or starch, and utilize them to generate energy in glycolysis. Carbohydrates also serve as structural components of many biological molecules. For instance, nucleic acids, DNA and RNA, are polymers of 2-deoxy-D-ribose (**1-1**) and D-ribose (**1-2**), with each monomer decorated with one of the five nucleobases, adenine, guanine, cytosine, thymine, and uracil. Equally important are glycoproteins and glycolipids in which sugar modifications play crucial roles in cellular communication and molecular recognition. In eukaryotes, nine sugars (**1-3**−**1-11**) are known to exist in these surface glycans (Figure 1-1).[12,13] In contrast, the surface polysaccharides of prokaryotes are composed of much more diverse sugar structures. Indeed, more than one hundred monosaccharide units, known as unusual sugars, have been identified.[14]

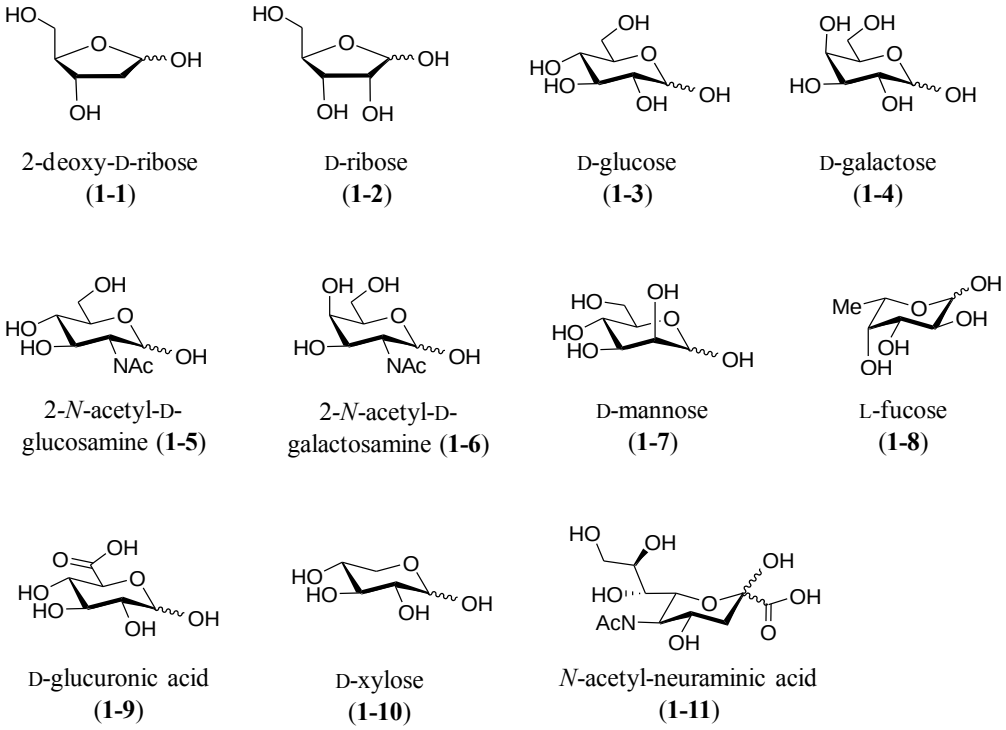


Figure 1-1. Monosaccharides found in the glycoproteins and glycolipids of eukaryotes.

Unusual Sugar Biosynthesis

The unusual sugars, including deoxy-, amino-, nitro-, and branched-sugars, are also found in many bioactive secondary metabolites produced by plants and microorganisms. Importantly, these sugar moieties often serve as the key determinants responsible for the efficacy and specificity of their biological activities.[15] Thus, elucidation of the biosyntheses of unusual sugars is important for the pathway reengineering effort to construct new glycosylated secondary metabolites having desirable biological activities, which may have clinical applications.

Deoxysugars

The biosynthetic pathways of a series of deoxysugars and the mechanisms of enzymes involved in the pathways have been extensively studied over the past two decades. In general, the sugar substrates used in unusual sugar biosynthetic pathways need to be activated at C-1 by nucleoside diphosphate (NDP) modification. This C-1-NDP group serves as the anchor which is recognized by the corresponding biosynthetic enzymes and glycosyltransferases in the pathways. Interestingly, the biosyntheses of a large portion of deoxysugars share a common intermediate, TDP-4-keto-6-deoxy-α-D-glucose (**1-14**). For example, the structurally diverse sugars, TDP-D-mycaminose (**1-17**), TDP-D-desosamine (**1-21**), TDP-L-mycarose (**1-26**), TDP-L-eremosamine (**1-30**), TDP-D-forosamine (**1-34**), and TDP-L-rhodinose (**1-36**), are all derived from **1-14** (Figure 1-2).[14]

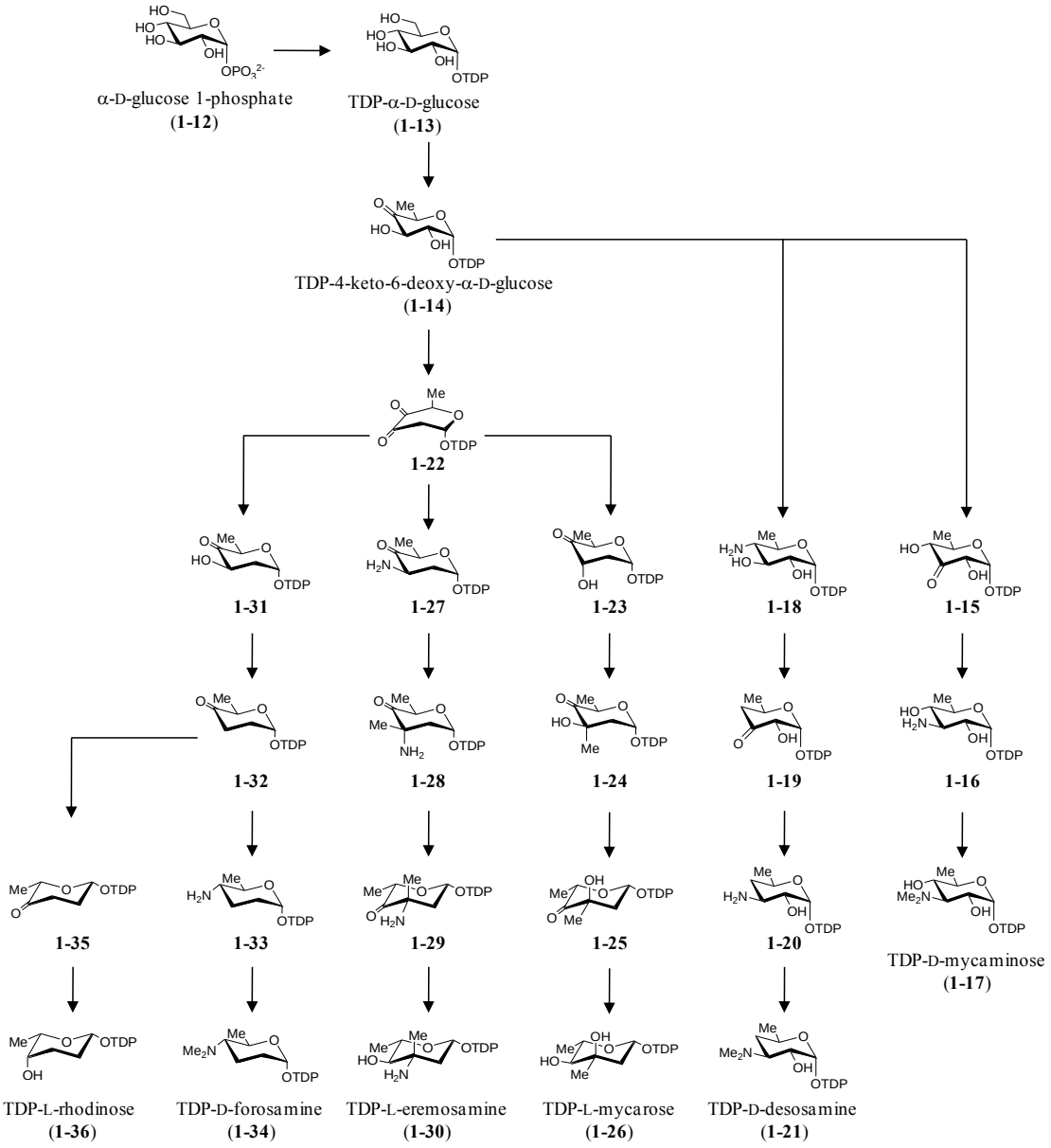


Figure 1-2. Representative examples of deoxysugar biosynthetic pathways.

Glycosyltransfer and Glycodiversification

The produced unusual sugars are transferred to their specific acceptors by glycosyltransferases. Since glycosyltransferases involved in the biosynthesis of

secondary metabolites have been shown to have relatively relaxed substrate specificity, the reengineering of the sugar biosynthetic pathways described above can be coupled with appropriate glycosyltransferase to generate new glycosylation patterns. This approach, known as glycodiversification, is a powerful strategy, possibly leading to the creation of a large number of new glycoconjugates of secondary metabolites for the screening of novel or enhanced biological activities.[16,17]

Thiosugars

A few types of thiosugars exist in nature, however, they are rare. Only a handful of thiosugar-containing natural products have been identified to date, and they are described below in more detail. Because of the rarity of their natural occurrence, knowledge regarding the biosynthesis of thiosugars has been scarce thus far. On the other hand, sulfur is seen in several primary metabolites, and its unique chemical nature is often responsible for the specific function of these biomolecules. Thus, creation of unnatural thiosugar-containing compounds by reengineering the biosynthetic pathways of the parent molecules to include the thiosugar biosynthetic genes has great potential to produce biologically active novel compounds.

1.3 SULFUR-CONTAINING BIOMOLECULES

1.3.1 Occurrence and Importance

Primary Metabolites

Sulfur is an essential element found ubiquitously in living systems. Sulfur-containing small molecules are distributed in a wide range of biomolecules in primary metabolism such as amino acids (**1-37**, **1-38**), antioxidants (**1-39**, **1-40**), nucleic acids (**1-**

41−**1-44**), metal clusters (**1-45**−**1-47**), and enzyme cofactors (**1-48**−**1-55**) (Figure 1-3).[18,19]

Cysteine (**1-37**) and methionine (**1-38**) are the two amino acids containing a sulfur atom in their structures. They are not only essential components of proteins but also utilized for the biosyntheses of other important biomolecules. For example, the antioxidants, glutathione (**1-39**) and mycothiol (**1-40**), consist of the cysteine (**1-37**) structure, and *S*-adenosyl-L-methionine (**1-49**) is directly synthesized from methionine (**1-38**). Indeed, cysteine (**1-37**) is a major sulfur donor in biological systems because the sulfur atom of cysteine (**1-37**) can be transferred to other compounds through enzymes called cysteine desulfurases, which are described below in more detail. On the basis of the chemical properties of sulfur, the thiol group of cysteine (**1-37**) can be used as an acid or base and can also readily form a disulfide bond, which is important for stabilizing many protein structures. In addition, the lone pairs of electrons on the sulfur atom can serve as a metal ligand as seen in the iron sulfur clusters (**1-45**−**1-47**), which are protein-bound cofactors serving as redox centers.

As shown in Figure 1-3, there are quite a few sulfur-containing enzyme cofactors. The characteristic chemical properties of sulfur contribute to their functions as coenzymes. For example, the two sulfur atoms of lipoic acid can be converted from the reduced dithiol form to the cyclic oxidized disulfide form during the catalysis of the oxidative decarboxylation of α-keto acids.

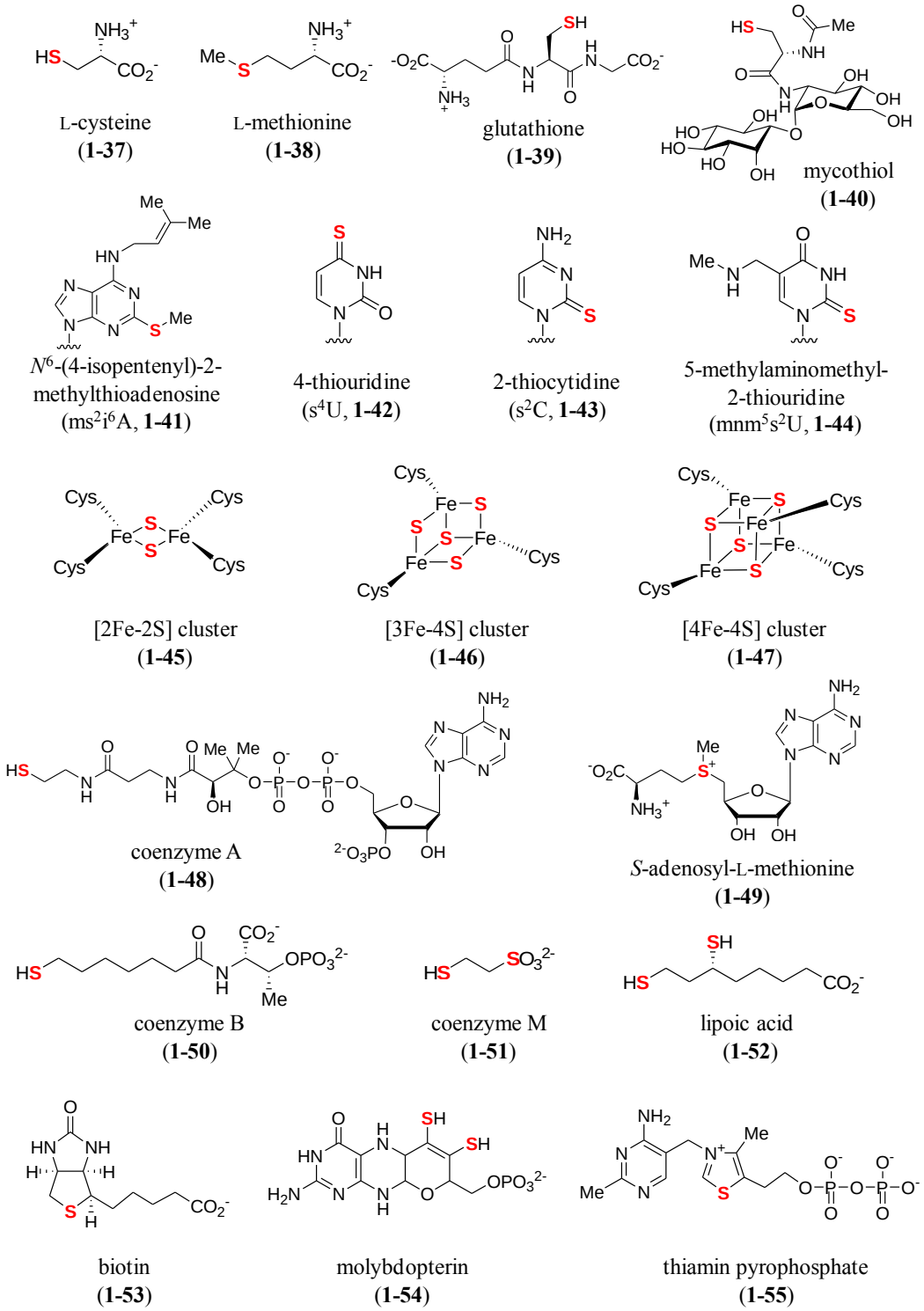


Figure 1-3. Sulfur-containing biomolecules in primary metabolism.

Secondary Metabolites

Sulfur is also found in the structures of a number of secondary metabolites. Representative examples, including clinically useful antibiotics penicillin (**1-56**) and cephalosporine (**1-57**), an antitumor antibiotic leinamycin (**1-62**), highly modified peptide metabolites thiostrepton (**1-64**), which contains multiple thiazole rings, are shown in Figure 1-4.[18,20-25]

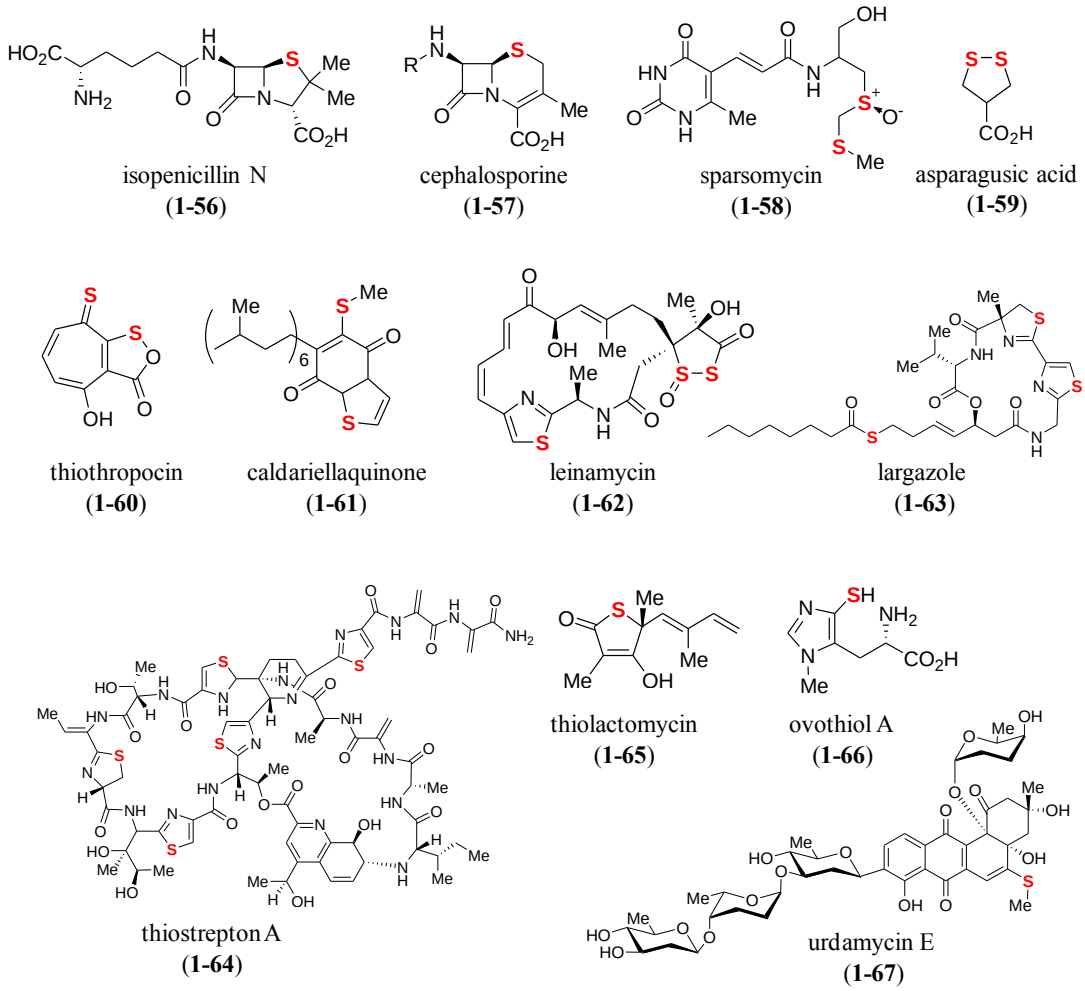
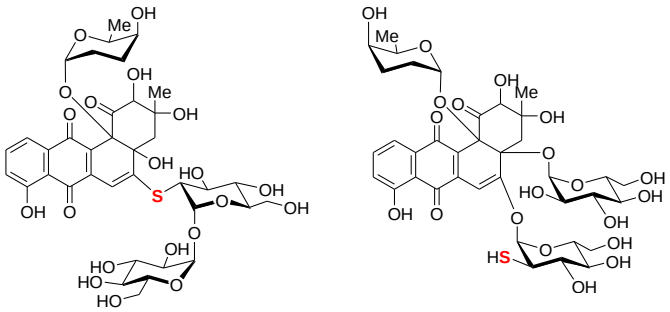


Figure 1-4. Representative structures of sulfur-containing secondary metabolites.

Thiosugar-Containing Natural Products

As described above, only a limited number of thiosugar-contaning natural products have been reported thus far. They are conveniently categorized into seven groups (group I−VII) based on their structures as described below.

Group I contains a C-2-thio functional group in a glucose-like scaffold. There are only two such natural products known thus far, BE-7585A (**1-67**) from *Amycolatopsis orientalis* subsp. *vinearia* BA-07585 and rhodonocardin A (**1-68**) from *Nocardia* sp. (Figure 1-5).[26,27] While the former shows a C-2-thioether linkage, the latter possesses a free C-2-thiol group. Interestingly, **1-67** and **1-68** are constitutional isomers to each other, and both are angucycline-type antibiotics. BE-7585A (**1-67**) was reported as a potential antitumor agent due to its thymidylate synthase inhibition activity.



BE-7585A
(**1-67**)

rhodonocardin A
(**1-68**)

Figure 1-5. Thiosugar-containing natural products (group I).

Group II contains a unique C_8-sugar backbone with a C-1-thioether. This includes lincomycin (**1-69**) from *Streptomyces lincolnensis* and celesticetin (**1-70**) from *Streptomyces caelestis* (Figure 1-6).[28,29] A few other celesticetin derivatives were also isolated from the same producer strain.[30] This class of compounds exhibit antimicrobial

activity against Gram-positive bacteria. Notably, the semi-synthetic chlorinated lincomycin derivative, 7-chloro-7-deoxylincomycin known as clindamycin (**1-71**), is a widely-used FDA approved antibiotic.

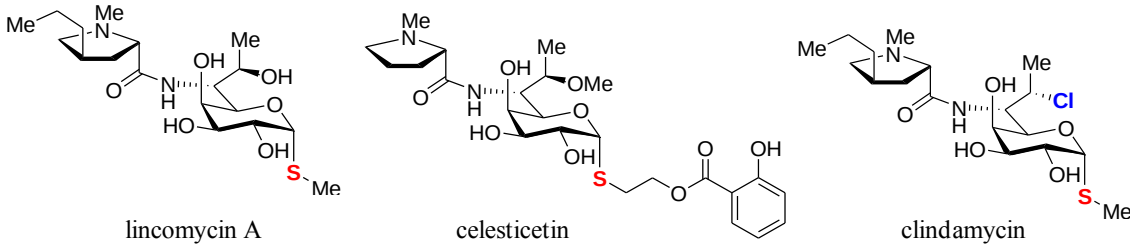


Figure 1-6. Thiosugar-containing natural products (group II).

*[1] clindamycin is a semi-synthetic compound.

Group III is composed of the endiyne class of natural products possessing a C-4-thiosugar. Calicheamicins (**1-72**) from *Micromonospora echinospora* ssp. *calichensis*, esperamicins (**1-73**) from *Actinomadura verrucosospora*, namenamicin (**1-74**) from the marine ascidian *Polysyncraton lithostrotum*, and shishijimicins (**1-75**) from the ascidian *Didemnum proliferum* are in this group (Figure 1-7).[31-34] While thio groups in compound **1-72** is involved in a ester linkage, the other three compounds (**1-73**−**1-75**) possess a methyl-thio group at the C-4 position. More interestingly, the configuration of the thio functional groups is not always equatorial (**1-72** and **1-73** and terminal sugar moiety of **1-74**) but can also be axial (internal sugar moiety of **1-74** and **1-75**). This implicates that there may be more than one type of sulfur incorporation mechanism in the biosynthesis of this class of thiosugars.

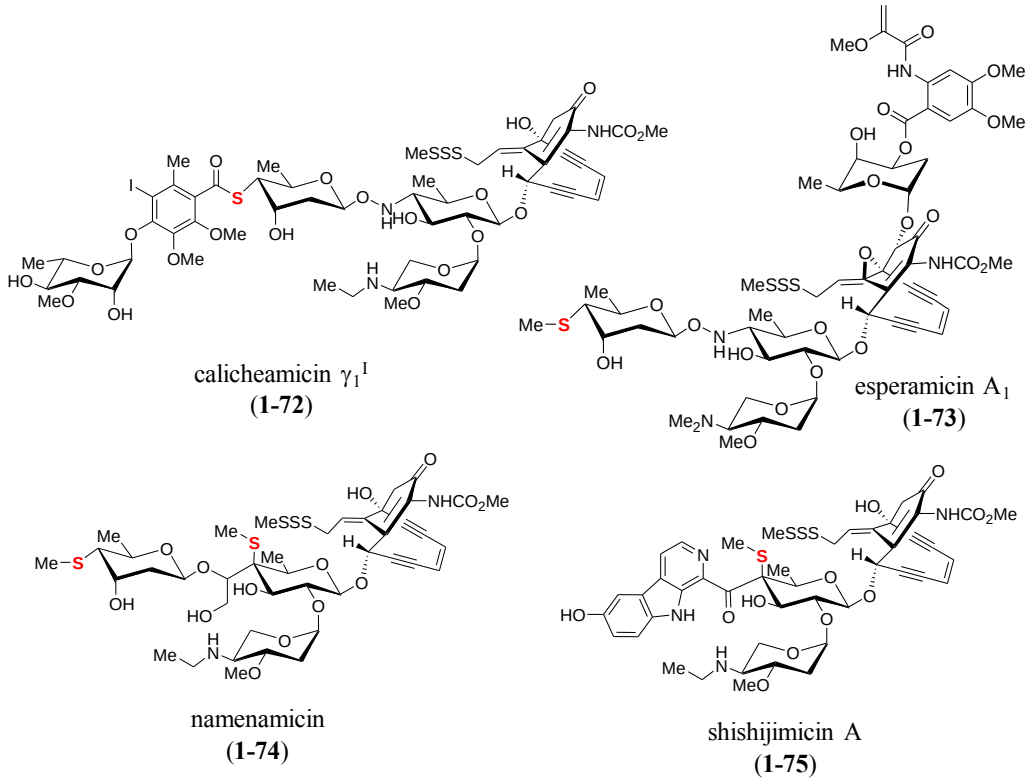
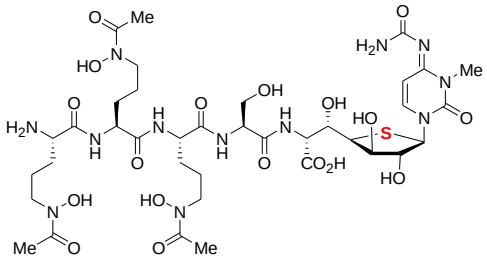
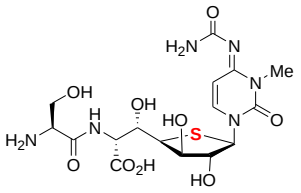


Figure 1-7. Thiosugar-containing natural products (group III).

Group IV of thiosugar-containing natural products possesses a thioribosyl pyrimidine moiety that exhibits tRNA synthetase inhibition activity. Albomycin (**1-76**) from *Streptomyces subtropicus* or several other *Streptomyces* species, and SB-217452 (**1-77**) from *Streptomyces* sp. are categorized in this group (Figure 1-8).[35,36] The former structure also contains a hydroxamate siderophore, which allows the compound to be taken up by the ferrichrome transport system of the target microorganisms such as *Escherichia coli* and related Gram-negative bacteria.[37] In this way, albomycin (**1-76**) works as a Trojan horse antibiotic.



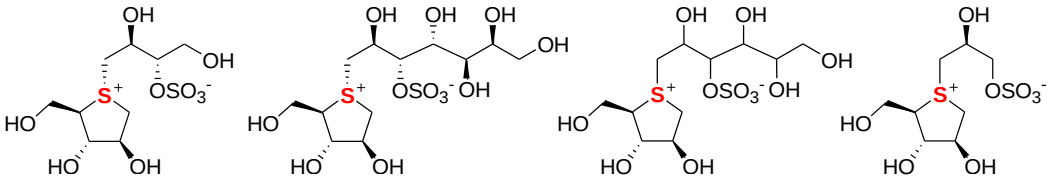
Albomycin δ_2
(**1-76**)



SB-217552
(**1-77**)

Figure 1-8. Thiosugar-containing natural products (group IV).

Group V shows a 1-deoxy-4-thio-D-arabinofranosyl cation core forming a thiosugar sulfonium sulfate inner salt with a sulfate group of the polyhydroxylated acyclic chain. This unique structure was found in plant natural products, salacinol (**1-78**), kotalanol (**1-79**), ponkaranol (**1-80**), and salaprinol (**1-81**), produced from *Salacia reticulate* (Figure 1-9).[38-41] Since this class of compounds exhibits α-glucosidase inihibition activity, they are potent antihyperglycemic drugs for type 2 diabetes.



kotalanol
(**1-79**)

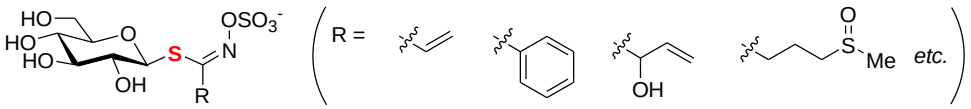
ponkaronol
(**1-80**)

salaprinol
(**1-81**)

Figure 1-9. Thiosugar-containing natural products (group V).

Group VI is composed of diverse structures of glucosinolates (**1-82**) produced by plants, mainly, but not exclusively, by the order Capparales, including the Brassicaceae, Capparaceae, and Caricaceae (Figure 1-10).[42] Approximately 120 glucosinolates are

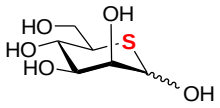
known and share this 1-thio-β-D-glucopyranose core linked via a sulfur atom to a (Z)-*N*-hydroximinosulfate ester. Upon hydrolysis of glucose moiety by a thioglucoside glucohydrolase (myrosinase), the unmasked aglycone moiety is rearranged to the chemically more active isothiocyanates, nitriles, or thiocyanate structures. These active compounds are known to be involved in plant defense. In fact, the distinct taste and flavors of cabbage, cauliflower, broccoli (Brassicaceae vegetables) and mustard, horseradish, wasabi (condiments) are mainly due to the isothiocyanate hydrolysis products.[42]



glucosinolates
(**1-82**)

Figure 1-10. Thiosugar-containing natural products (group VI).

Lastly, group VII is a 5-thio-D-mannose (**1-83**) isolated from the marine sponge *Clathria pyramida* (the actual producer might be its symbiont) (Figure 1-11).[43] Interestingly, this C-5-thiosugar is the only naturally occurring example of this class known thus far. In contrast, a number of 5-thiopyranose and 4-thiofuranose derivatives, in which the endocyclic ring oxygen of regular sugar is replaced by a sulfur atom, have been chemically synthesized as inhibitors for various enzymes.[44]



5-thio-D-mannose
(**1-83**)

Figure 1-11. Thiosugar-containing natural products (group VII).

Genes and Pathways for the Thiosugar Biosyntheses

Among the thiosugar-containing natural products (**1-67**−**1-83**) described above, the biosynthetic gene clusters for lincomycin (**1-69**), celesticetin (**1-70**), calicheamicin (**1-72**), and albomycin (**1-76**) have been identified from the producer microorganisms.[45-48] Although some of the genes found in the clusters have been experimentally investigated, the pathway and mechanisms of thiosugar biosynthesis in each case remain obscure.

In contrast to the biosynthetic gene clusters found in the bacteria, genes involved in the same biosynthetic pathway of secondary metabolites are not clustered in plants. Thus, it is more challenging to study complex biosynthetic pathways of plants' secondary metabolites. Despite this disadvantage, most enzymes involved in the biosynthesis of glucosinolates (**1-82**) have been identified (Figure 1-12).[42] Although the detailed mechanism of the key sulfur incorporation step (**1-86** → **1-87**) remains to be elucidated, cysteine (**1-37**) or glutathione (GSH, **1-39**) has been proposed to be the sulfur donor, which is added to the reactive intermediate **1-86** *via* either a glutathione-*S*-transferase-type reaction or a non-enzymatic route.[49] Recently, a γ-glutamyl peptidase (GGP1) was found to be able to hydrolyze the GSH conjugate of **1-87**.[50] This result suggests that the direct sulfur donor may be GSH (**1-39**) rather than cysteine (**1-37**), and the Cys–Gly conjugate obtained through the hydrolysis of the GSH conjugate may be the substrate for a C−S lyase (SUR1).[50,51] The ensuing sulfur incorporation into the sugar moiety (**1-88** →

1-90) forming the thioglucoside linkage is a typical glycosyltransfer reaction utilizing UDP-glucose (**1-89**) as the sugar donor.[52]

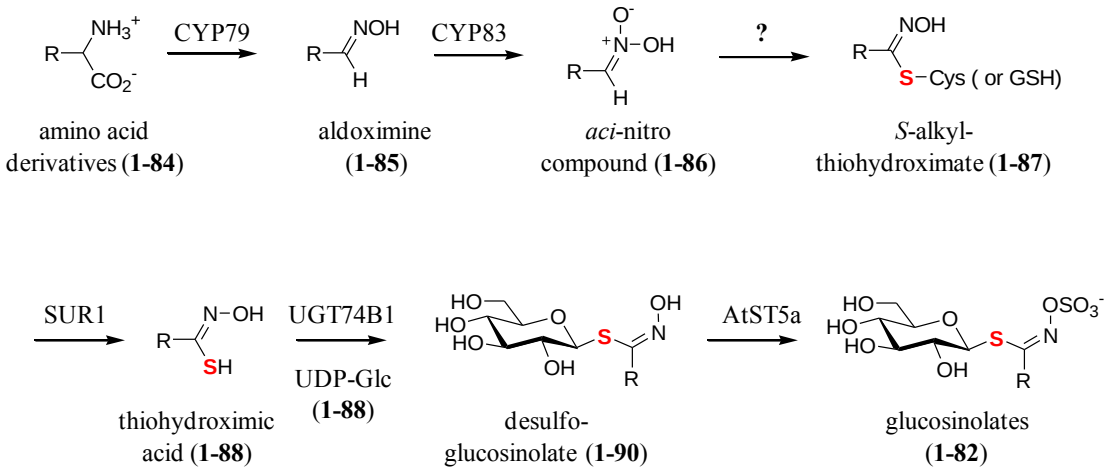
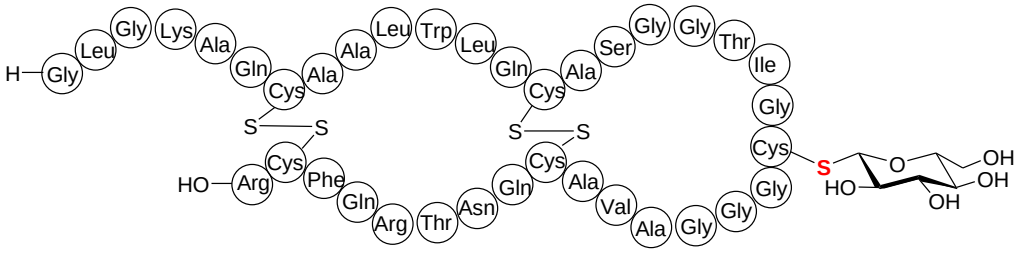


Figure 1-12. Biosynthetic pathways of glucosinolates.

Notably, a similar C-1-thioglucoside was recently identified as a new type of post-translational modification in a peptide natural product, sublancin (**1-91**), produced by *Bacillus subtilis* 168 (Figure 1-13).[53] Based on the lessons learned from studies of glucosinolates (**1-82**) and sublancin (**1-91**) biosyntheses, *S*-glycosyltransfer reaction between UDP-glucose (**1-89**) and the cysteine thiol is likely a common strategy to form a C-1-thioglycoside. Thus, it is possible that a similar glycosyltransferase is involved in the biosyntheses of C-1-thiosugars found in lincomycin A (**1-69**) and celesticetin (**1-70**) although in these cases, the sugar donor is a highly modified NDP-sugar rather than a simple UDP-glucose (**1-89**).

With the exception of the glucosinolates biosyntheses described above, sulfur incorporation pathways and mechanisms for the formation of thiosugars have not been studied at all. In view of the diverse structural categorization (group I−VII), there may be

at least seven different classes of enzymes catalyzing the sulfur incorporation reactions. In this dissertation, the biosynthetic pathways and mechanisms of a C-2-thiosugar-contaning natural product, BE-7585A (**1-67**) and a C-1-thiosugar-containing natural product, lincomycin A (**1-69**) are studied with a focus on their characteristic thiosugar structures.



sublancin
(**1-91**)

Figure 1-13. Structure of sublancin.

1.3.2 Sulfur Incorporation Mechanisms

While most biosynthetic pathways of thiosugar formation have not been studied, two types of sulfur incorporation mechanisms have been elucidated in the biosyntheses of sulfur-containing primary metabolites (Figure 1-3). The first mechanism is an ionic mechanism involving a bisulfide (HS^-), a thiocarboxylate group ($R\text{-}COS^-$) or a cysteine persulfide group ($R\text{-}S\text{-}S^-$) as sulfur donor. The second mechanism involves a radical intermediate catalyzed by radical-SAM enzymes containing an *S*-adenosyl-L-methionine (SAM, **1-49**) cofactor and iron-sulfur clusters.[18,19,54-58] The former mechanism is used in the biosyntheses of cysteine (**1-37**), thiamin (**1-55**), molybdopterin (**1-54**), 2- and 4-thiouridine (**1-42**, **1-43**). The latter is responsible for the production of biotin (**1-53**) and lipoic acid (**1-52**).

Ionic Mechanism Involving Bisulfide

Plants and microorganisms can utilize bisulfide as a direct sulfur source in cysteine (**1-37**) biosynthesis, *via* catalysis by a pyridoxal 5'-phosphate (PLP, **1-92**)-dependent enzyme, *O*-acetylserine sulfhydrylase (OASS). The reaction starts with binding of *O*-acetylserine (**1-93**) to an internal aldimine between a lysine residue of OASS and PLP (**1-92**). The resulting external Schiff base between *O*-acetylserine (**1-93**) and PLP (**1-92**) facilitates the elimination of an *O*-acetyl group of **1-93** to generate an α-aminoacrylate external Schiff base intermediate (**1-94**). The terminal alkene of this intermediate (**1-94**) is then attacked by a nucleophilic bisulfide to form the cysteine external Schiff base (**1-95**), which releases a cysteine (**1-37**) molecule *via* a transimination reaction (Figure 1-14).[59]

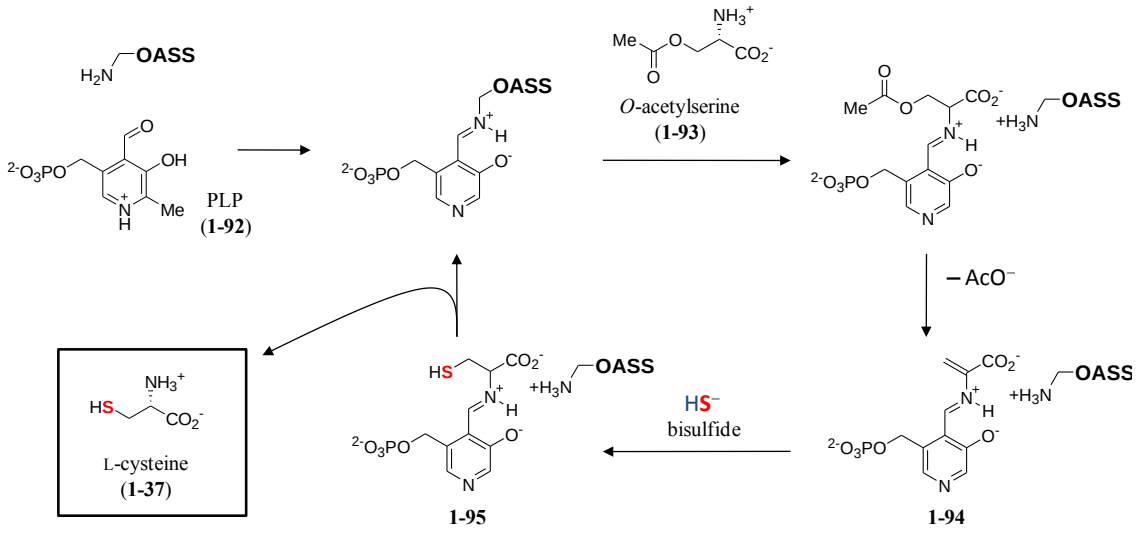


Figure 1-14. Cysteine biosynthesis catalyzed by *O*-acetylserine sulfhydrylase.

Ionic Mechanism Involving Protein Persulfide

Protein persulfides are thought to be the major source of sulfide equivalents in the synthesis of sulfur-containing biomolecules. The persulfide bound to a large protein has the advantage to react with a specific target because highly reactive free sulfide is toxic *in vivo*. Protein persulfide groups are generated by a family of cysteine desulfurases (e.g., NifS, IscS, SufS, and CsdA), which are pyridoxal 5'-phosphate (PLP, **1-92**)-dependent enzymes catalyzing the transfer of a sulfur atom from L-cysteine (**1-37**) to an active site cysteine residue of the desulfurase to yield the protein persulfide and L-alanine (**1-96**) (Figure 1-15A). The generated persulfidic sulfur (**1-97**) can be transferred to a cysteine residue of another protein to form a new protein persulfide (**1-98**) in a process known as sulfur relay (Figure 1-15B)[60] or can react with the activated carboxy terminal of a sulfur carrier protein (**1-99**) to form a protein-protein acyl disulfide intermediate (**1-100**),[61] which is then reduced to yield a protein thiocarboxylate (**1-101**; Figure 1-15C). While the protein thiocarboxylate (**1-100**) is the sulfur donor in thiamin (**1-55**) and molybdopterin (**1-54**) biosyntheses, the persulfidic sulfur of the protein persulfide (**1-96**, **1-97**) is directly introduced into the substrate in thionucleoside (**1-42**, **1-43**) biosynthesis. In either case, cysteine desulfurases are the key enzymes responsible for the sulfur incorporation directly or indirectly into the substrates.

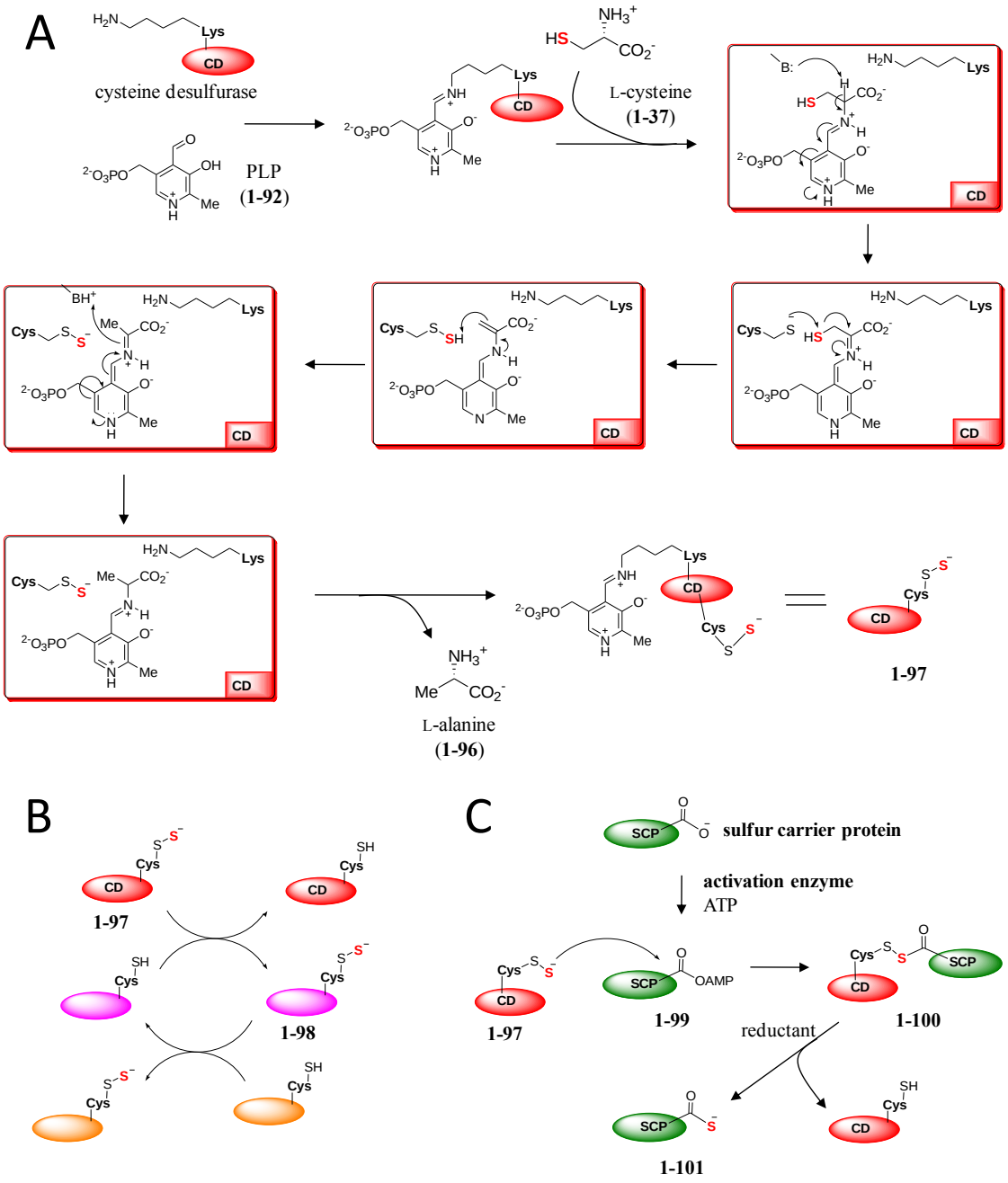


Figure 1-15. Protein persulfide and protein thiocarboxylate.

(A) Mechanism of cysteine desulfurase (CD) to form protein persulfide (**1-97**). (B) Sulfur relay found in 2-thiouridine biosynthetic pathway. (C) An activation mechanism of sulfur carrier protein to form protein thiocarboxylate (**1-101**).

The representative example of sulfur incorporation using a protein persulfide is seen in the 4-thiouridine (**1-42**) biosynthesis. The reaction employs ThiI and one of the cysteine desulfurases, IscS. As shown in the proposed mechanism (Figure 1-16),[62-64] the persulfide formed on IscS (**1-102**) is first transferred to an active site cysteine residue of ThiI to generate the ThiI-persulfide (**1-103**). ThiI also catalyzes the adenylation of the substrate uridine using ATP. The persulfide group of **1-103** attacks this activated uridine to produce a disulfide intermediate (**1-104**). Upon the attack by another active site cysteine on the disulfide bond, 4-thiouridine (**1-42**) is released. The internal disulfide bond of ThiI formed during the catalysis must be reduced by a reductant to restore the active enzyme.

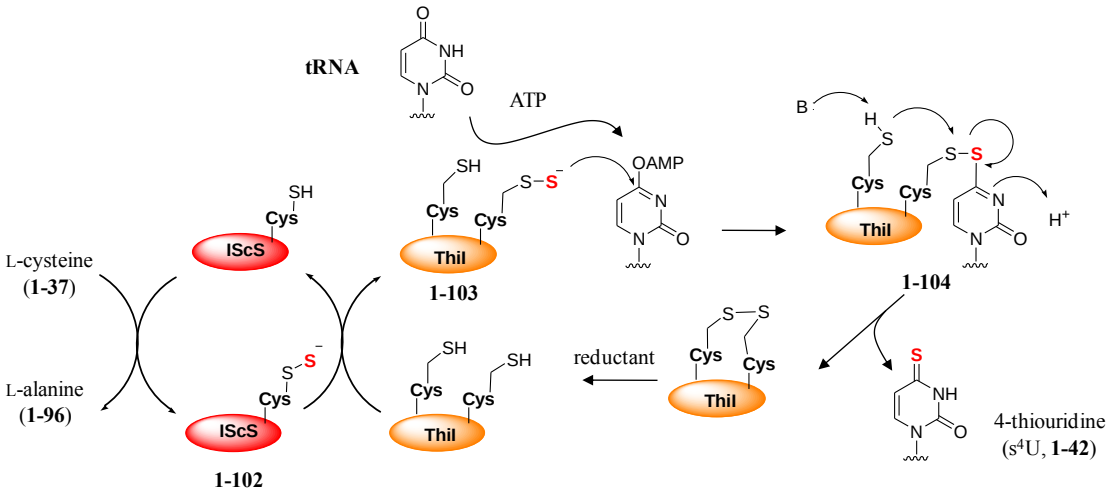


Figure 1-16. Proposed mechanism for the 4-thiouridin biosynthesis.

Ionic Mechanism Involving Protein Thiocarboxylate

The sulfur incorporation mechanism in bacterial thiamin biosynthesis is the best characterized example of a reaction involving a sulfur carrier protein-thiocarboxylate (**1-**

101; Figure 1-15C).[65] The sulfur carrier protein and its activation enzyme in this pathway are ThiS and ThiF, respectively. ThiF catalyzes the adenylation of the *C*-terminal carboxylate group of ThiS. Sulfur is then transferred to this position from a protein persulfide group of a cysteine desulfurase. The reactive thiocarboxylate generated on the *C*-terminus of ThiS attacks the substrate, 1-deoxy-D-xylulose 5-phosphate (DXP, **1-105**), which forms a covalent complex with a thiazole synthase, ThiG (**1-106**) (Figure 1-17). ^{18}O-labeling experiments revealed an *S* to *O* acyl shift (**1-107** $\rightarrow$ **1-108**) followed by the formation of a thioketone intermediate (**1-109**).[66] The acyl modification on the C-4 alcohol facilitates the elimination of the hydroxyl group from this position together with ThiS (**1-110** $\rightarrow$ **1-111**). The ThiG-bound substrate intermediate further reacts with dehydroglycine (**1-112**) to eventually form the thiazole ring (**1-113**). Recent studies using an optimized reconstitution procedure for ThiG revealed that the actual enzymatic product is a thiazole phosphate carboxylate tautomer (**1-114**) which shows unexpected stability.[67] Aromatization of **1-114** is mediated by another enzyme, TenI, to produce thiazole phosphate carboxylate (**1-115**).[68]

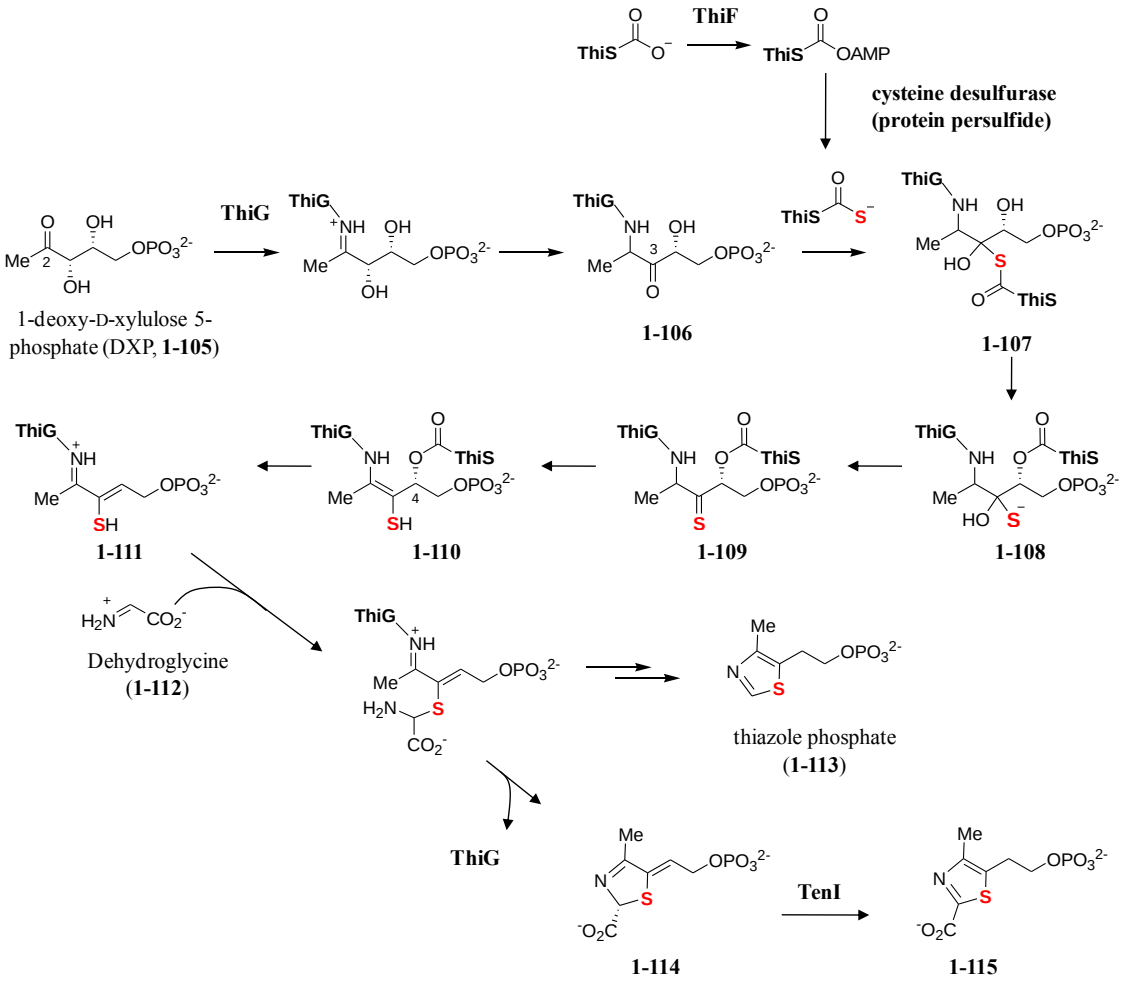


Figure 1-17. Proposed mechanism for thiazole biosynthesis.

Sulfur carrier proteins are a growing class of proteins consisting of 65–100 amino acids with a characteristic GG sequence at their *C*-termini. The *C*-terminal carboxylate of a sulfur carrier protein is activated by a ThiF-like enzyme, and the resulting acyl adenylates is attacked by a sulfur donor to form a sulfur carrier protein thiocarboxylate (Figure 1-15C). This thiocarboxylate serves as a sulfur donor in the specific biosynthetic pathway, such as thiamin, as described above. Interestingly, structure and activation mechanism of sulfur carrier proteins resemble those of eukaryotic ubiquitin.[69-71] Indeed,

recent studies have shown an evolutionary linkage between bacterial sulfur carrier proteins and eukaryotic ubiquitin or ubiquitin-like proteins.[72-77] Figures 1-18 and 1–19 summarize the sulfur carrier proteins and their activating enzymes biochemically characterized thus far, including the thiamin (**1-55**, A),[61,66] molybdopterin (**1-54**, B),[78-80] prydine dithiocarboxylic acid (pdtc, **1-116**, C),[81,82] thioquinolobactin (**1-117**, D),[83] L-cysteine (**1-37**, E),[84] L-methionine (**1-38**, F),[85] 2-thioribothymidine (**1-118**, G),[86] and 5-methoxycarbonylmethyl-2-thiouridine (mcm^5s^2U, **1-119**, H)[75,87] biosynthetic pathways, as well as the corresponding eukaryotic ubiquitination pathway **(**I).[88]

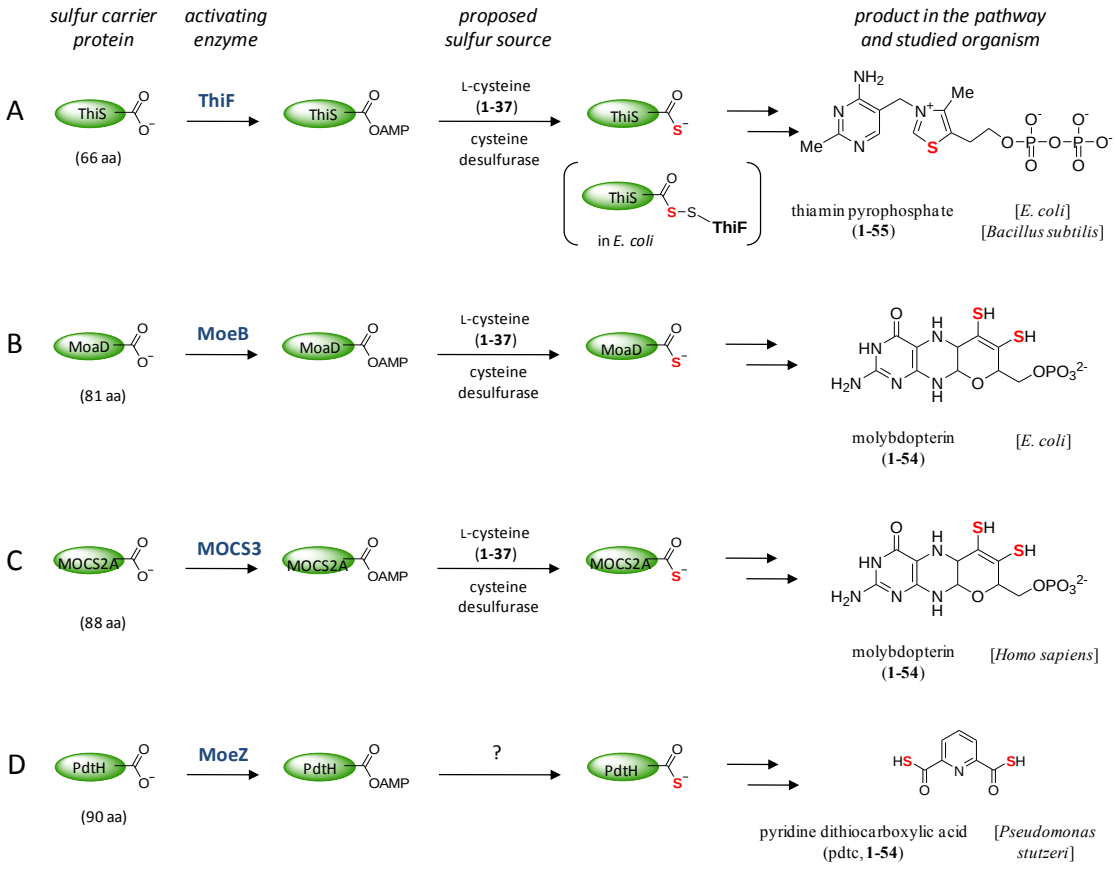


Figure 1-18. Sulfur carrier proteins and their activating enzymes (I).

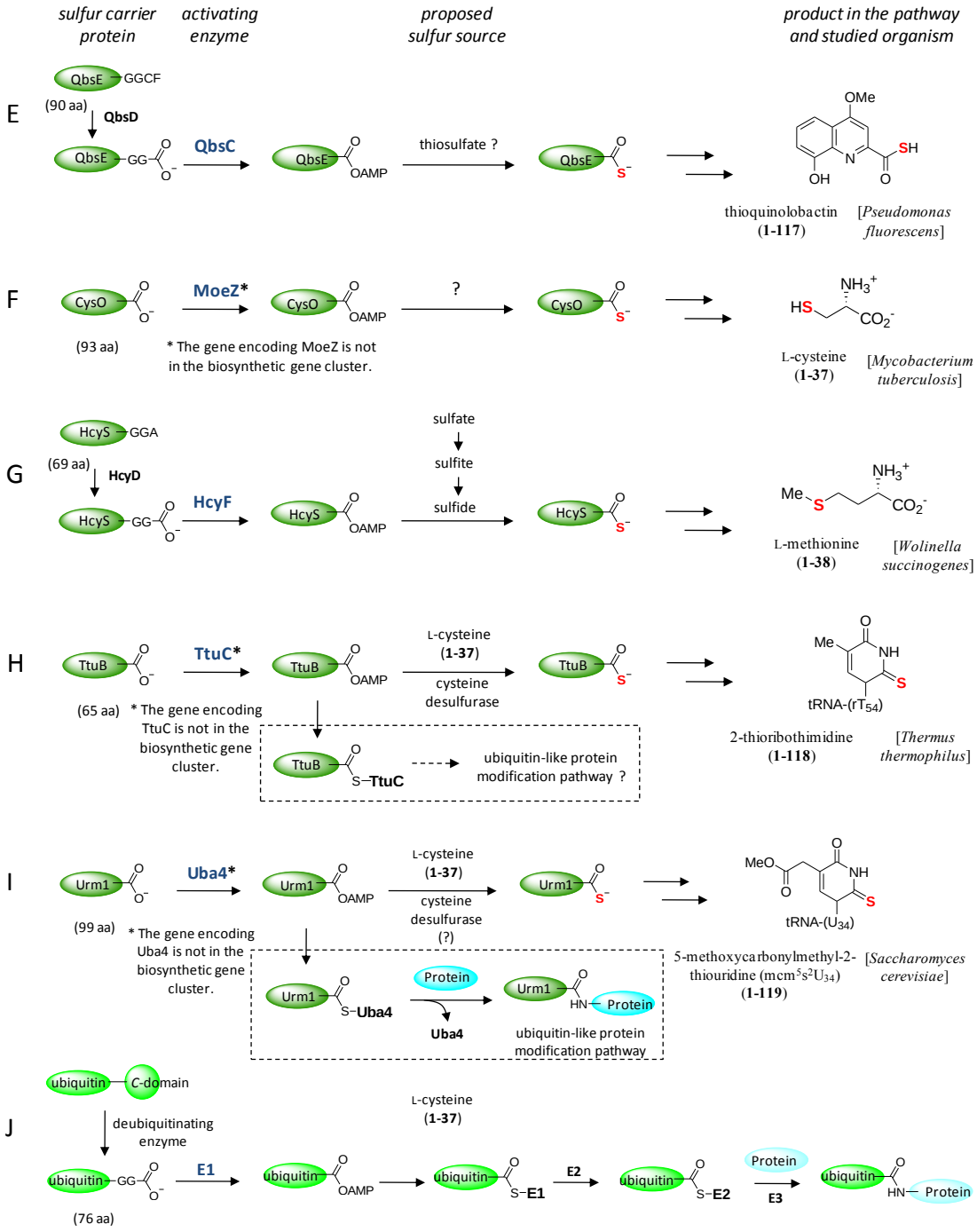


Figure 1-19. Sulfur carrier proteins and their activating enzymes (II) and the corresponding eukaryotic ubiquitination pathway.

Radical Mechanism

Sulfur can also be incorporated *via* a radical mechanism using radical-SAM enzymes. These enzymes, which contain *S*-adenosyl-L-methionine (SAM, **1-49**) and iron-sulfur clusters, catalyze a reductive cleavage of SAM to L-methionine (**1-38**) and a 5'-deoxyadenosyl 5'-radical (**1-120**; Figure 1-20A).[18,54,56,89,90] The reactive radical intermediate, **1-120**, then functions to remove a hydrogen atom from an appropriate substrate. Biotin synthase and lipoyl synthase are the two best characterized radical-SAM enzymes involving sulfur incorporation (Figure 1-20B,C). Both reactions are similarly initiated by the formation of a 5'-deoxyadenosyl 5'-radical (**1-120**) and utilizing sacrificial iron-sulfur clusters as the sulfur sources.

The proposed reaction mechanism of biotin synthase is shown in Figure 1-21.[54,91] Biotin synthase inserts a sulfur atom between carbons 6 and 9 of dethiobiotin (**1-122**) to form a thiophane ring. In this pathway, the sulfur atom is known to be derived from a nearby $[2Fe\text{-}2S]^{2+}$ cluster (**1-45**) in the active site, and the first C–S bond formation involves an electron transfer from the initially formed C-9 radical to one of the iron atoms, leading to a $Fe^{(II)}$–$Fe^{(III)}$ Fe/S species (**1-123**). In the second cycle, another molecule of 5'-deoxyadenosyl 5'-radical (**1-120**) abstracts the 6-*pro*-*S* hydrogen atom of **1-123**. Upon the attack of the resulting C-6 radical to the sulfur atom coordinated to the $Fe^{(III)}$ site, the second C–S bond is formed with reduction of the $Fe^{(II)}$–$Fe^{(III)}$ cluster to its reduced $Fe^{(II)}$–$Fe^{(II)}$ form (**1-124**).

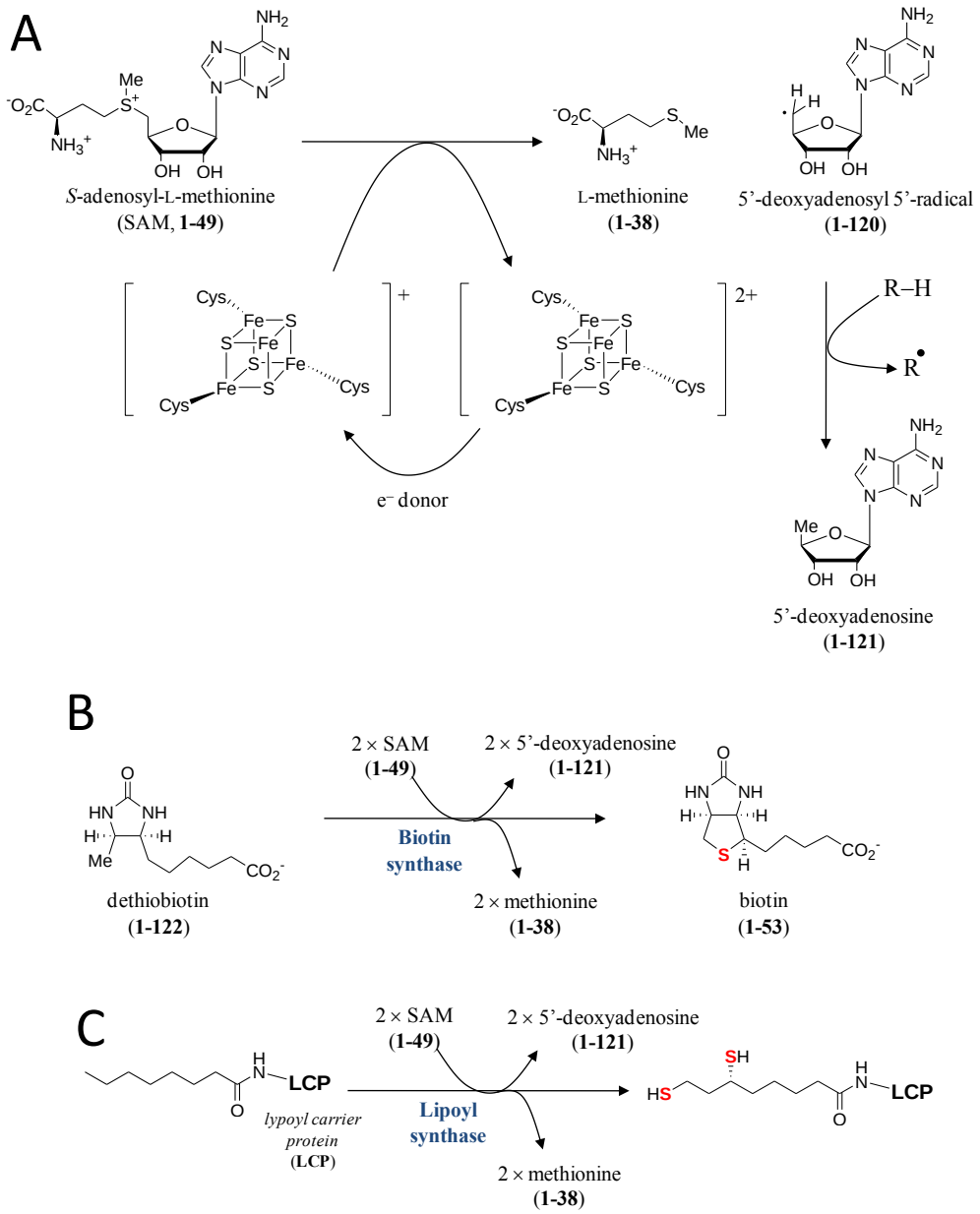


Figure 1-20. Radical-SAM enzymes involved in sulfur incorporation.

(A) Reductive cleavage of SAM to form a reactive 5'-deoxyadenoyl 5'-radical (**1-120**). (B) Biotin synthase reaction. (C) Lipoyl synthase reaction.

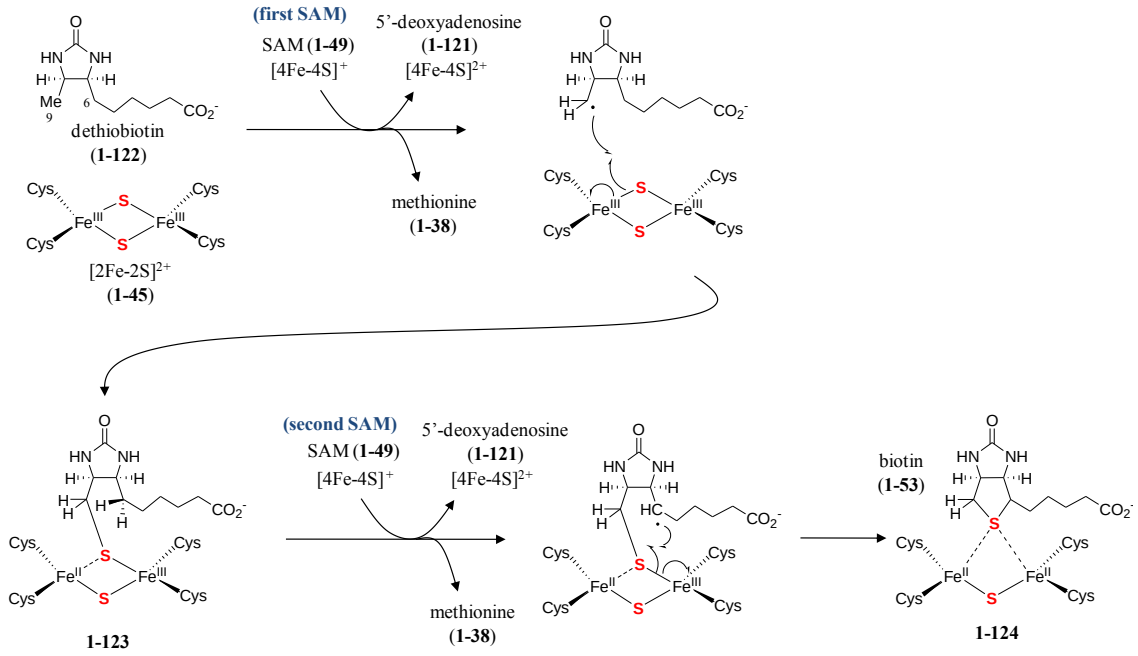


Figure 1-21. Proposed reaction mechanism of biotin synthase.

1.4 SUMMARY AND THESIS STATEMENT

Highly modified sugars commonly found in prokaryotic secondary metabolites are often the vital components determining the efficacy and specificity of the parent molecules. Thus, a thorough understanding of unusual sugar biosynthesis may enable the exploitation of the biosynthetic machineries to generate new therapeutic agents. Over the past 20 years, many deoxy-, amino- and branched-chain sugar biosynthetic pathways and their detailed mechanisms have been elucidated. However, knowledge regarding the biosynthesis of thiosugars has been scarce due to the rarity of its natural occurrence. On the other hand, mechanistic details of several biological sulfur insertion reactions involved in the biosyntheses of thiamin, molybdopterine, biotin, lipoic acid, and tRNA

modifications have recently been unraveled. Highly regulated reactive intermediates, including protein persulfides, protein thiocarboxylates, or *S*-adenosylmethionine (SAM)-dependent radicals are often generated in the enzymes catalyzing these reactions. To investigate whether similar or different mechanisms are used to form the C–S bond in thiosugar biosyntheses is the primary focus of this study.

BE-7585A, an angucycline-type natural product, consisting of an angucycline core, a deoxysugar (rhodinose), and a 2-thiodisaccharide moieties, is one of the only two known natural products containing a C-2-thiosugar moiety. Described in Chapter 2 of this dissertation is the first report of the isolation of its biosynthetic gene cluster, which contains a candidate gene responsible for 2-thiosugar biosynthesis. This genetic information as well as the feeding experiments using ^{13}C-labeled acetate led to a proposal of the biosynthetic pathway of BE-7585A. Biochemical verification of the proposed function of BexG2, a glycosyltransferase homologue, responsible for making the thiodisaccharide moiety is described in Chapter 3. The relaxed substrate specificity at the C-2 position of the glycosyl acceptor was unraveled during the course of substrate specificity assays, suggesting that BexG2 might have evolved from a trehalose 6-phosphate synthase to accept the unusual 2-thioglucose 6-phosphate as the substrate. In addition, a model reaction demonstrating a nonenzymatic C–S bond formation between the angucycline core and a thiosugar is documented in Chapter 3. Finally, detailed mechanistic studies of 2-thiosugar biosynthesis are described in Chapter 4. On the basis of a series of MS studies, a thiazole synthase homologue, BexX, was discovered to form a covalent adduct with glucose 6-phosphate with an active site lysine residue under physiological conditions. The most important finding in this study, obtained by genome mining, is that a sulfur carrier protein involved in a primary metabolite biosynthetic pathway is recruited to deliver a sulfur atom to produce the 2-thiosugar in the BexX-

catalyzed reaction. Surprisingly, despite the analogy of the proposed mechanism of the BexX reaction to thiazole synthesis, the sulfur carrier protein involved in thiamin biosynthesis is incapable of performing this function in 2-thiosugar biosynthesis. Instead, sulfur carrier proteins likely involved in cysteine and molybdopterin biosyntheses are effective in promoting sulfur addition to the C-2 keto group of the BexX–glucose 6-phosphate complex.

Lincomycin A, a lincosamide antimicrobial natural product carrying a C-1 methylthio substituent, is another thiosugar-containing natural product studied in this dissertation. Chapter 5 describes a detailed proposal of the complete biosynthetic pathway for the unique thiosugar moiety, methlthiolincosamide (MTL), in lincomycin A. A transaldolase homologue, LmbR, was identified to be responsible for the construction of the C_8 sugar backbone of MTL. Although LmbR showed relatively relaxed substrate specificity, the physiologically relevant substrate pair in this pathway was found to be D-ribose 5-phosphate and D-fructose 6-phosphate, and the enzymatic product is D-*glycero*-D-*altro*-octulose 8-phosphate. In contrast, the next enzyme, LmbN, responsible for the subsequent C-1–C-2 isomerization, showed strict substrate specificity. The product of LmbN, D-*erythro*-D-*gluco*-octose 8-phosphate, was verified by comparison with a synthetic standard. This work identified two key intermediates in the MTL biosynthetic pathway for the first time. The results are consistent with the proposed biosynthetic pathway. This work set the stage for future mechanistic studies of NDP-octose formation and its modifications including the C-1 sulfur incorporation to make MTL.

Chapter 2: Biosynthetic Studies of BE-7585A (I): Structural and Genetic Analyses for the Biosynthetic Pathway

2.1 INTRODUCTION

BE-7585A (**2-1**), produced by *Amycolatopsis orientalis* subsp. *vinearia* BA-07585, is an angucycline-type natural product possessing thymidylate synthase inhibition activity.[26] The angucycline class of natural products, which have a characteristic fused four-ring frame assembled in an angular manner, are rich in antibacterial or anticancer activities.[92] In addition to the benz[*a*]anthraquinone core, BE-7585A also contains the deoxysugar, rhodinose (**2-2**), and a disaccharide appendage (**2-4**) consisting of a glucose and a highly unusual 2-thioglucose (**2-3**). Notably, BE-7585A is one of only two 2-thiosugar-continaing natural products reported thus far. The other compound, rhodonocardin A (**2-5**),[27] is a constitutional isomer of BE-7585A (**2-1**), independently isolated from different soil bacteria (Figure 2-1, also see Chapter 1).

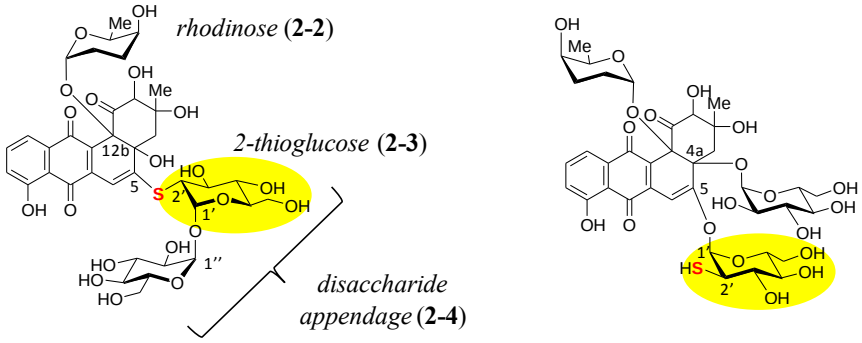


Figure 2-1. Structures of BE-7585A and rhodonocardin A.

BE-7585A (**2-1**) and rhodonocardin A (**2-5**) are closely related, having an identical benz[*a*]anthraquinone core structure and the same three sugar components, rhodinose, 2-thioglucose, and glucose. But they differ in the attachment of the 2-thioglucose and glucose moieties on the core molecule. In BE-7585A (**2-1**), the 2-thioglucose and glucose moieties are linked via their anomeric carbons, and the resulting disaccharide (**2-4**) is attached to the C-5 position of the aglycone through a thioether linkage. In contrast, the 2-thioglucose and glucose in **2-5** are separately attached to C-5 and C-4a of the angucycline core via *O*-glycosidic bonds, leaving the 2-mercapto-substituent in **2-3** as a free thiol group. The structure of rhodonocardin A (**2-5**) is fully documented, [27] but that of BE-7585A (**2-1**) is not because the structure is reported in a patent.[26] Thus, in this chapter, isolation and characterization of **2-1** from *Amycolatopsis orientalis* subsp. *vinearia* BA-07585 are first described, and the glycosydic linkage of each sugar moiety is carefully analyzed using a series of 1D- and 2D-NMR experiments to confirm that the reported structure of BE-7585A (**2-1**) is indeed different from rhodonocaridn A (**2-5**) produced from *Nocardia* sp.

The detailed NMR-characterization of BE-7585A (**2-1**) is also required in order to perform ^{13}C-labeled acetate (**2-6**) feeding experiments, which is a well-established method to study the biosynthesis of polyketide-type natural products. In general, the angucycline core structures (**2-8**) are assembled with ten acetate units catalyzed by type II polyketide synthases (PKSs) *via* a decaketide intermediate (**2-7**) (Figure 2-2).[93] For example, reported feeding experiments for vineomycin A_1 (**2-9**) from *Streptomyces matensis* subsp. *vineus*,[94] urdamycin A (or also known as kerriamycin B) (**2-10**) from *Streptomyces fradiae* (strain Tü 2717),[95] PD 116740 (**2-11**) from *Streptomyces* WP 4669,[96] kinamycin D (**2-12**) from St*reptomyces murayamaensis*,[97,98] gilvocarcin M (**2-13**) from *Streptomyces griseoflavus* Gö 3592,[99,100] and PD 116198 (**2-14**) from *Streptomyces*

phaeochromogenes WP 3688[101] led to the determination of the biosynthetic origins of their core structures (Figure 2-3). In this chapter, similar feeding experiments were carried out for BE-7585A (**2-1**) to verify the involvement of an analogous type II PKS-catalyzed angucycline core formation in its biosynthesis.

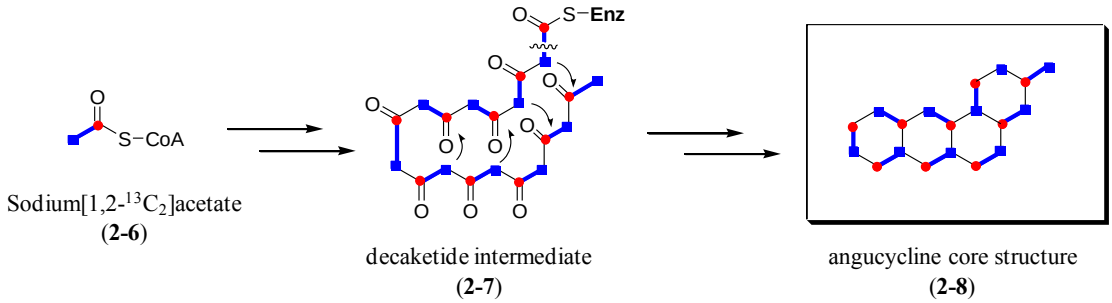


Figure 2-2. General biosynthetic pathway for an angucycline core structure.

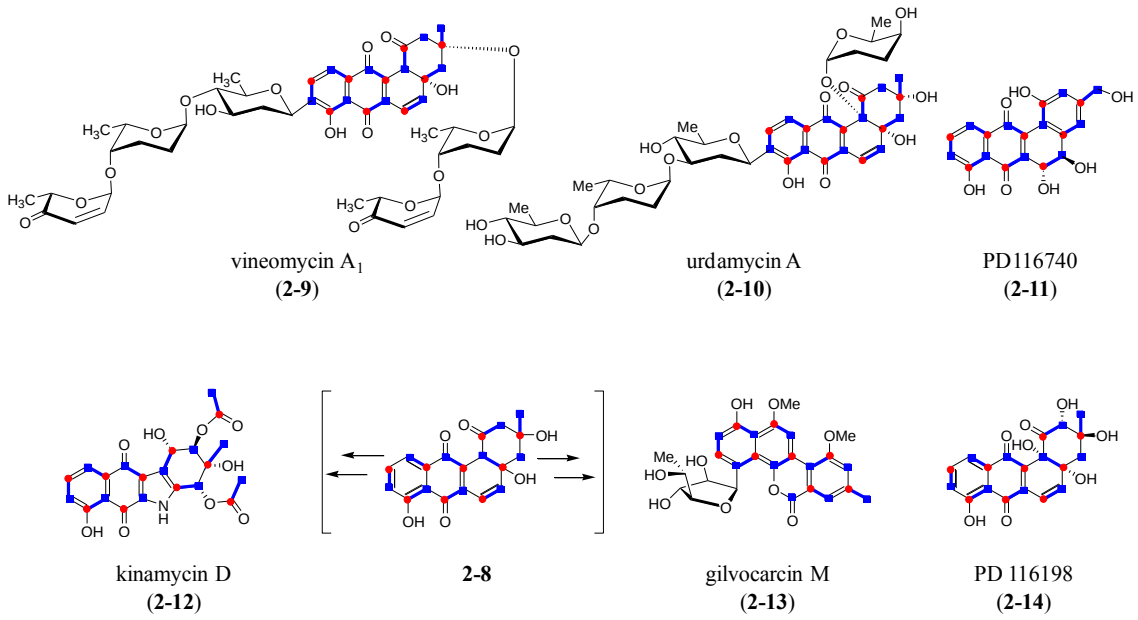


Figure 2-3. Acetate incorporation patterns of angucycline-type or angucyline core-derived natural products.

Moreover, the highly deoxygenated sugar, rhodinose (**2-2**) and its derivatives have been found in several natural products in addition to BE-7585A (**2-1**) and rhodonocrdin A (**2-5**) described above. Those of whose biosynthetic gene clusters have been identified include aclarubicin (**2-15**) from *Streptomyces galilaeus,*[102], urdamycin A (**2-10**) from *Streptomyces fradiae,*[103] β-rhodomycin IV (**2-16**) from *Streptomyces galilaeus,*[104], landomycin A (**2-17**) from *Streptomyces cyanogenus*,[105] and granaticin B (**2-18**) from *Streptomyces violaceoruber* (Figure 2-4).[106] Homologous gene analyses and gene disruption experiments of these gene clusters have led to the identification of the genes responsible for rhodinose formation (Figure 2-5). Although *in vitro* biochemical verification of the functions of the corresponding gene products in the synthesis of TDP-L-rhodinose (**2-26)** has not been performed, a similar set of sugar biosynthetic genes likely exists in the BE-7585A biosynthetic gene cluster.

On the basis of the reported studies about the structurally similar angucycline-type or rhodinose-containing natural products mentioned above, the biosynthetic pathway of BE-7585A (**2-1**) is predicted to include catalysis by type II PKS and TDP-sugar modifying enzymes in the formation of its angucyline core (**2-8**) and TDP-L-rhodinose (**2-26**), respectively. Although no information about 2-thioglucose biosynthesis is available, it is known that genes involved in post-PKS modifications are usually located around the PKS genes, forming a gene cluster responsible for the production of the targeted polyketide-derived natural product. Thus, in order to identify the unknown thiosugar biosynthetic genes, we decided to first locate the BE-7585A biosynthetic gene cluster by using probes designed based on the sequence of some predictable genes such as type II PKS or one of the TDP-sugar biosynthetic genes that are likely present in the cluster. In this chapter, identification of the BE-7585A biosynthetic gene cluster from *A. orientalis* genome is described. The approach is based on the PCR-based cosmid library

screening, targeting genes encoding β-ketoacyl synthase α subunit (KS_α), TDP-6-deoxy-4-keto-D-glucose 2,3-dehydratase, and TDP-2,6-dideoxy-4-keto-D-glucose 3-dehydrase. Detailed sequence analysis combined with mechanistic consideration leads to the complete proposal of the biosynthetic pathway for BE-7585A (**2-1**) containing the highly unusual 2-thioglucose (**2-3**) appendage.

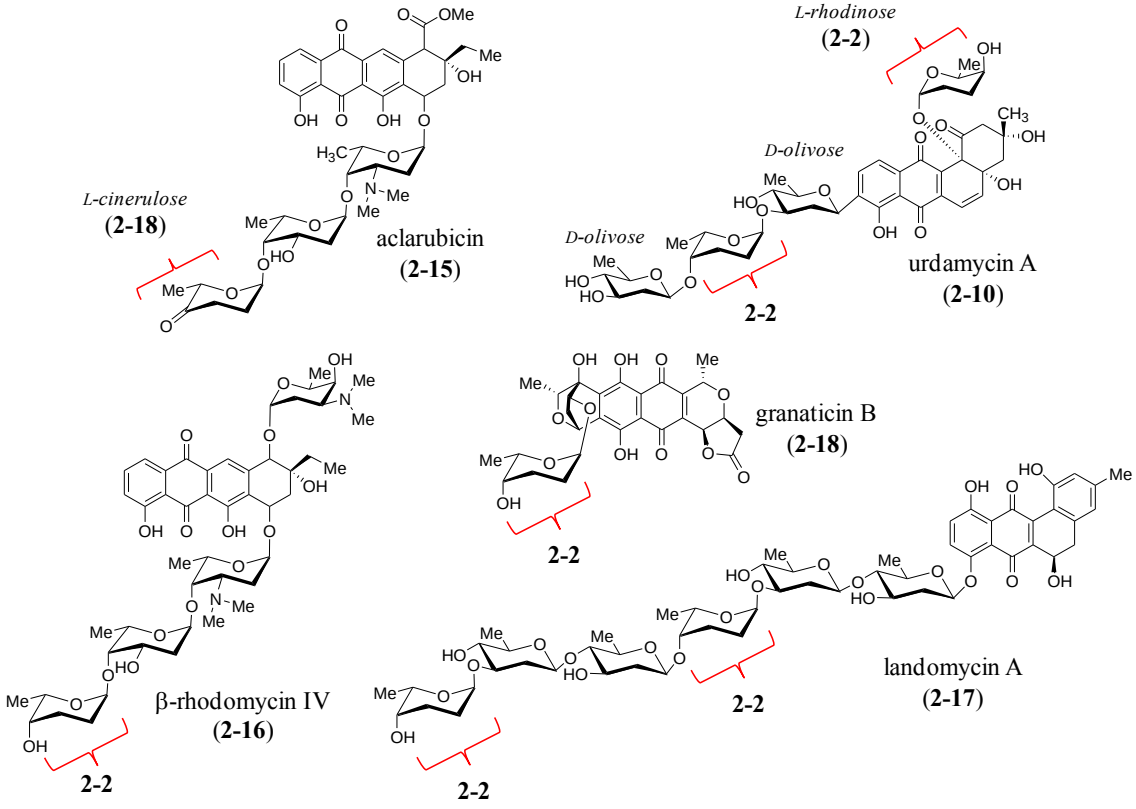


Figure 2-4. A partial list of natural products containing rhodinose or its derivatives.

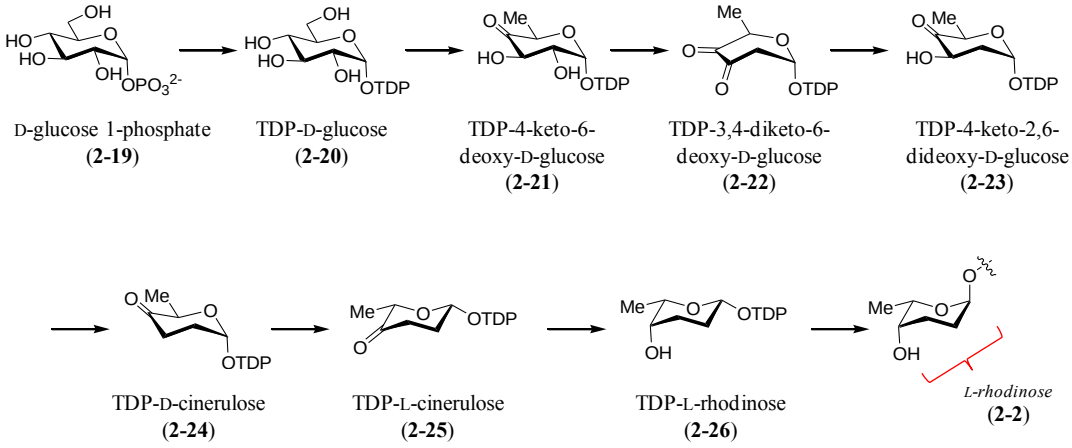


Figure 2-5. Proposed L-rhodinose biosynthetic pathway.

2.2 EXPERIMENTAL PROCEDURES

2.2.1 General

Materials

All chemicals and reagents were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA), and were used without further purification unless otherwise specified. Enzymes and molecular weight standards used for the cloning experiments were products of Invitrogen (Carlsbad, CA) or New England Biolabs (Ipswich, MA). Kits for DNA gel extraction and spin minipreps were obtained from Qiagen (Valencia, CA). Gigapack III XL packaging extract was product of Stratagene (La Jolla, CA). Deep Vent (exo–) DNA polymerase was purchased from New England Biolabs. Growth medium components were acquired from Becton Dickinson (Sparks, MD). Sterile syringe filters were products of Fisher Scientific. Analytical C_{18} HPLC columns were purchased from Varian (Palo Alto, CA). Oligonucleotide primers were prepared by Invitrogen or Integrated DNA Technologies (Coralville, IA).

Bacterial Strains and Plasmids

A. orientalis subsp. *vinearia* BA-07585 was generously provided by Banyu Pharmaceutical Co. (Tokyo, Japan). *Escherichia coli* DH5α, acquired from Bethesda Research Laboratories (Gaithersburg, MD), was used for routine cloning experiments. *E. coli* XL-1 Blue MRF', purchased from Stratagene, was employed for cosmid manipulations. Plasmids pGEM-T Easy and pUC119, used for sub- cloning, were products of Promega (Madison, WI) and Takara Bio (Shiga, Japan), respectively. Standard genetic manipulations of *E. coli* were performed as described by Sambrook et al.[107]

Instrumentation

UV−vis spectra were recorded using a Beckman DU 650 spectrophotometer. NMR spectra were acquired on a Varian Unity 500 MHz spectrometer, and chemical shifts (in ppm) are reported relative to that of the solvent peak (δ_H = 4.65 for deuterated water, δ_H = 3.30 and δ_C = 49.0 for deuterated methanol). HPLC was performed on a Beckman Coulter System Gold equipped with a UV detector. DNA concentrations were measured using a NanoDrop ND-1000 UV−vis instrument from Thermo Fisher Scientific. DNA sequencing was performed by the core facility of the Institute of Cellular and Molecular Biology at the University of Texas, Austin. Vector NTI Advance 10.1.1 from Invitrogen was used for sequence alignments. Mass spectroscopy was performed at the Mass Spectrometry core facility in the Department of Chemistry and Biochemistry at the University of Texas, Austin.

2.2.2 Production, Isolation, and Identification of BE-7585A.

Spores of *A. orientalis* subsp. *vinearia* BA-07585 were inoculated into 100 mL of International *Streptomyces* Project (ISP) medium 2 and grown in a rotary incubator at 29

°C and 250 rpm for 5 days. The resultant seed culture (25 mL) was transferred to 1 L of ISP-2 medium and grown under the same conditions for 10 days. The growth culture was centrifuged at 6000 × *g* for 10 min to remove the cells and other insoluble materials. The supernatant was applied to a column of the synthetic polymeric adsorbent Diaion HP20 (250 mL) and washed with water. The adsorbed compounds were then eluted with 500 mL of methanol. The collected red fractions were pooled and concentrated by rotary evaporation *in vacuo* and further purified on a Diaion CHP20P column (50 mL) using 5−20% of methanol in water as the eluent. The typical yield was 100−150 mg of BE-7585A (**2-1**) from 1-L of culture. The structure of **2-1** was determined by ^{1}H, ^{13}C, COSY, HSQC, HMBC, and NOESY NMR experiments and ESI-MS.

2.2.3 ^{13}C-Labeling Experiments.

Sodium [1-^{13}C]acetate Feeding Experiment

Spores of *A. orientalis* subsp. *vinearia* BA-07585 were inoculated into 50 mL of ISP-2 medium and grown in a rotary incubator at 30 °C and 250 rpm for 5 days. The resultant seed culture (25 mL) was transferred to 1 L of ISP-2 medium and grown under the same conditions for 2 days. Sodium [1-^{13}C]acetate, in an aqueous solution (4%, 5 mL) sterilized by filtration through a syringe filter (0.2 μm), was added to the culture, which had a light red color. Additional [1-^{13}C]acetate solution was added (4%, 5 mL × 4) to the growth culture every 13−20 h. The final concentration of labeled acetate was 0.1%. After a 10-day incubation period, the culture was centrifuged at 6000 × *g* for 10 min. The ^{13}C-labeled BE-7585A (**2-1**, 150 mg) was isolated from the supernatant by using Diaion HP20 and CHP20P columns as described above for the purification of the nonlabeled **2-1**. The purified compound was analyzed by ^{13}C NMR spectroscopy.

Sodium [1,2-$^{13}C_2$]acetate Feeding Experiment

The feeding experiment was also performed with sodium [1,2-$^{13}C_2$]acetate in 1-L ISP-2 medium. A 25 mL aqueous solution containing both sodium [1,2-$^{13}C_2$]acetate (0.68%) and nonlabeled sodium acetate (1.3%) was prepared and sterilized by filtration through a syringe filter. The solution (10 mL) was then added to the culture, which had been grown for 2 days and had turned light red. The remaining solution (15 mL) was added 20 h after the first addition. The final concentration of the labeled acetate was 0.017%. After a 10-day incubation period, the culture was centrifuged at 6000 × *g* for 10 min to remove cells and any insoluble materials. The ^{13}C-labeled BE-7585A (**2-1**, 80 mg) was isolated from the supernatant using Diaion HP20 and CHP20P columns and analyzed by ^{13}C NMR spectroscopy as described above.

2.2.4 Construction of the Cosmid Library.

Spores of *A. orientalis* subsp. *vinearia* BA-07585 were inoculated into 50 mL of tryptone soya broth (TSB) medium and grown in a rotary incubator at 30 °C and 250 rpm for 3 days. The mycelia were harvested by centrifugation at 5000 × *g* for 30 min. The cells were disrupted by lysozyme and SDS, and the cell lysate was treated with proteinase K. The released genomic DNA was isolated by phenol-chloroform extraction followed by sodium acetate-isopropyl alcohol precipitation. The DNA was partially digested with *Sau*3AI restriction endonuclease in a time dependent manner, and ligated into the *Bam*HI site of the pOJ446 vector.[108] The ligated DNA was packaged into phage particles using a Gigapack III XL packaging extract and then introduced into *E. coli* XL1-Blue MRF' according to the manufacturer's instructions. Recombinants were selected on Luria-Bertani (LB)-agar plates in the presence of apramycin (50 μg/mL). After incubation for 16 h, the resulting transformants were spotted within grids on new plates. This resulting genomic library contains a total of 454 cosmids.

2.2.5 PCR-Based Screening of the Cosmid Library.

Polymerase chain reaction (PCR) primers for screening of type II PKS were designed based on multiple sequence alignments of 10 known β-ketoacyl synthase α subunits (KS_α) of actinomycetes including *Streptomyces Venezuelae*[109] and *Streptomyces coelicolor* A3(2),[110,111] which are available in the National Center for Biotechnology Information (NCBI) database. Similarly, the primers for deoxysugar genes, such as TDP-6-deoxy-4-keto-D-glucose 2,3-dehydratase[112-114] and TDP-2,6-dideoxy-4-keto-D-glucose 3-dehydrase,[115] were designed based on the reported gene sequences from eight actinomycete strains including *Streptomyces fradiae*[103] and *Streptomyces cyanogenus*.[105] The sequences of the primers used for the PCR-based screening are listed in Table 1. PCR was performed using Deep Vent (exo–) DNA polymerase with these primers and the cosmid DNA from the genomic library. Amplified DNA fragments were ligated into the pGEM-T Easy vector for cloning and sequencing.

Table 2-1. Primers used for cosmid library screening

Primer name	Sequence
KS_α-forward	5'-CGACGCVCCSATCDCVCCSATC-3'
KS_α-reverse	5'-GGAANCCDCCGAABCCGCTGCC-3'
2,3-dehydratase-forward	5'-AGCTSTCSCCSACVGTBCAGC-3'
2,3-dehydratase-reverse	5'-WRGAAVCGRCCSCCYTCYTCSG-3'
3-dehydrase-forward	5'-TSAACCCGMTCVTSCAGACGG-3'
3-dehydrase-reverse	5'-CSGGRTGSCKGGTSAKGTTSCC-3'

M = AC, R =AG, W = AT, S = CG, Y = CT, K = GT, V = ACG, H = ACT, D = AGT, B = CGT, N = ACGT

2.2.6 Sequencing and Gene Cluster Identification.

Cosmids A108 and C006, which together span the entire BE-7585A biosynthetic gene cluster (*bex* cluster), were digested with either *Kpn*I, *Nco*I, or *Pst*I. The resulting 2.5−12 kb DNA fragments were subcloned into a modified pUC119 vector, whose multiple cloning site harbors extra restriction sites.[116] Both strands of the subclones were sequenced using M13 universal primers or by primer walking. Sequencing data were obtained using a capillary-based AB 3700 DNA analyzer and assembled using the Vector NTI Suite program. Open reading frame assignments were made with the assistance of FramePlot 2.3.2.[117] Homologous protein sequences were identified in the NCBI database using the basic local alignment search tool (BLAST).

2.2.7 Metyrapone-Feeding Experiments.

Spores of *A. orientalis* subsp. *vinearia* BA-07585 were inoculated into 10 mL of the ISP-2 medium and grown in a rotary incubator at 30 °C and 230 rpm for 2 days. The resultant seed culture (100 μL) was transferred to 10 mL of ISP-2 medium and grown under the same conditions for 1 day. Metyrapone (2-methyl-1,2-dipyrid-3-yl-1-propanone), in a DMSO solution (1 M, 10 μL), was added to the culture, and the cells were grown under the same conditions for 4 days. Additional metyrapone DMSO solution (1 M, 5 μL) was added to the culture, which had a red color. The final concentration of metyrapone in the culture was 1.5 mM. After a 1-day incubation period, the culture was centrifuged, and the supernatant was subjected to HPLC analysis equipped with a C_{18} analytical column (4 × 250 mm). Control sample obtained from the culture without adding metyrapone was similarly analyzed. The sample (10 μL) was eluted with a gradient of water (solvent A) and methanol (solvent B). The gradient was run from 20 to 90% B over 25 min, 90−20% B over 5 min, followed by re-equilibration at 20% B for 5 min. The flow rate was 0.8 mL/min, and the detector was set at 300 nm.

The supernatant obtained from the culture containing metyrapone was also subjected to LC-MS analysis. The sample (100 μL) was diluted with 400 μL of deionized water and analyzed by LC-MS (positive ion detection mode) equipped with a C_{18} column and a photodiode array (PDA) detector using a gradient of water and acetonitrile.

2.3 RESULTS AND DISCUSSION

2.3.1 Production, Isolation, and Identification of BE-7585A.

To determine that the *A. orientalis* subsp. *vinearia* BA-07585 strain can produce BE-7585A (**2-1**) in our hand and to verify the assigned structure, the bacterial strain was initially grown in the vegetative medium reported in the patent, which is composed of glycerin (2%), maltsyrup (1%), meat extract (0.2%), cotton seed meal (0.5%), yeast extract (0.1%), wheat germ (0.5%), $MgSO_4 \cdot 7H_2O$ (0.2%), NaCl (0.2%), $CaCO_3$ (0.2%), KH_2PO_4 (0.1%), $(NH_4)_2SO_4$ (0.1%), $FeSO_4 \cdot 7H_2O$ (0.001%), and $ZnSO_4 \cdot 7H_2O$ (0.001%), pH 6.7.[26] Despite several trials, we could not detect the target compound, which should exhibit the maximum absorbance (λ_{max}) at near 500 nm. Thus, ISP-2 medium, which is one of the typical media for *Actinomycetes* species composed of glucose (0.4%), yeast extract (0.4%) and malt extract (1%), pH 7.3, was tested. Interestingly, this bacterial culture growing in the ISP-2 medium gradually turned to red. The major product was then purified and subjected to spectroscopic analyses.

Absorbance spectrum of the purified compound in an aqueous solution shows the λ_{max} at 492 nm, consistent with that reported for BE-7585A (**2-1**) (Figure 2-6). In addition, the HRMS results (HRMS (ESI+) calculated for $C_{37}H_{46}O_{20}NaS$ $[M + Na]^+$ 865.2201, found 865.2208; HRMS (ESI−) calculated for $C_{37}H_{45}O_{20}S$ $[M - H]^-$ 841.2225, found 841.2222) are also consistent with the structure of BE-7585A. Further verification

of the structure of **2-1** relied on a series of NMR experiments (Figures 2-7−15). Based on the 2D-NMR data, the ^{1}H and ^{13}C NMR signals could be completely assigned (Table 2-2).

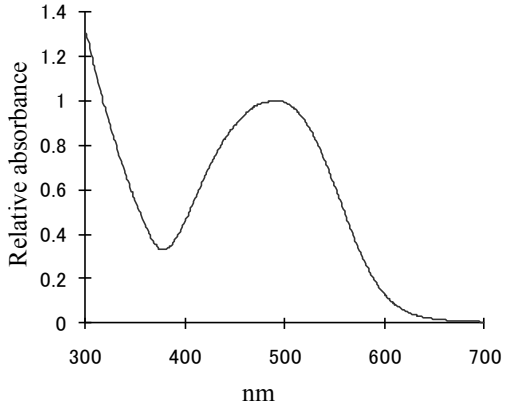


Figure 2-6. Absorbance spectrum of BE-7585A in H_2O.

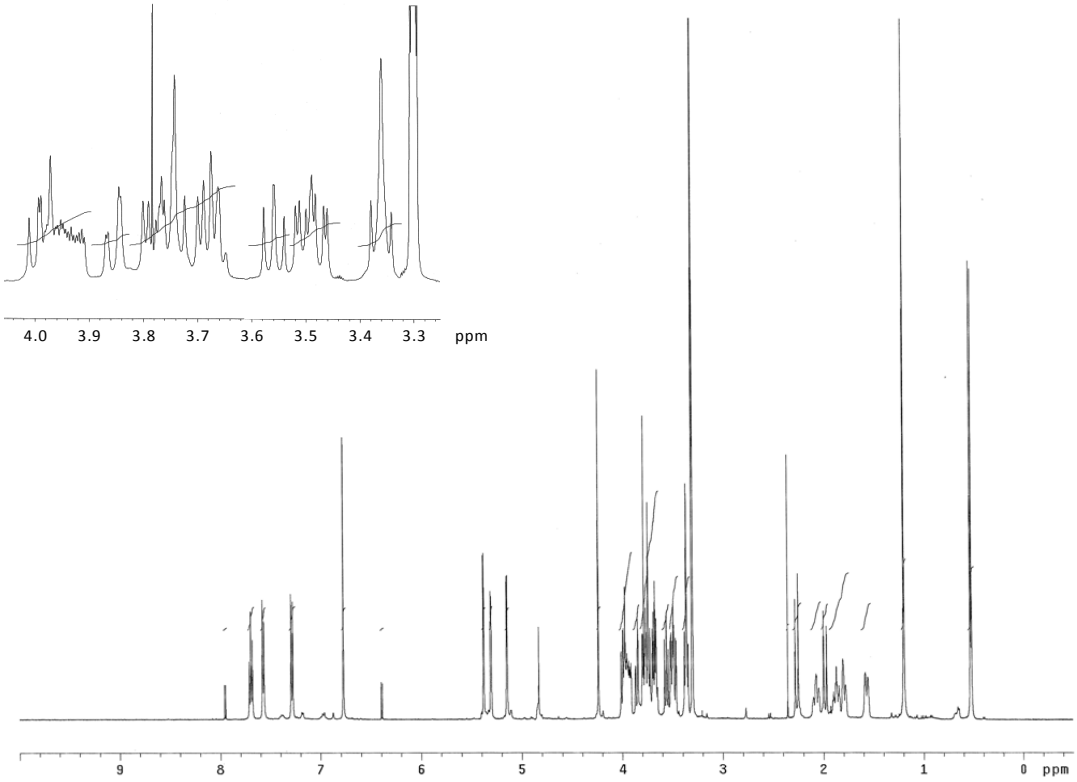


Figure 2-7. ^{1}H NMR spectrum (500 MHz, CD_3OD) of BE-7585A.

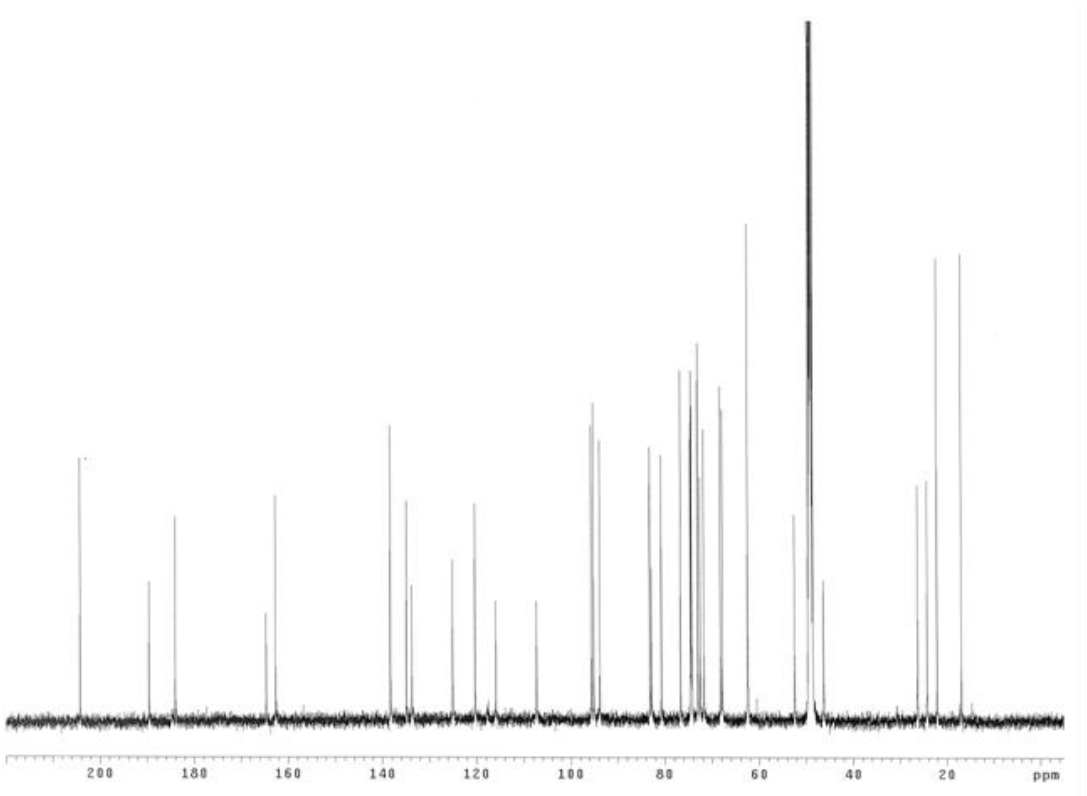


Figure 2-8. ^{13}C NMR spectrum (125 MHz, D_2O) of BE-7585A.

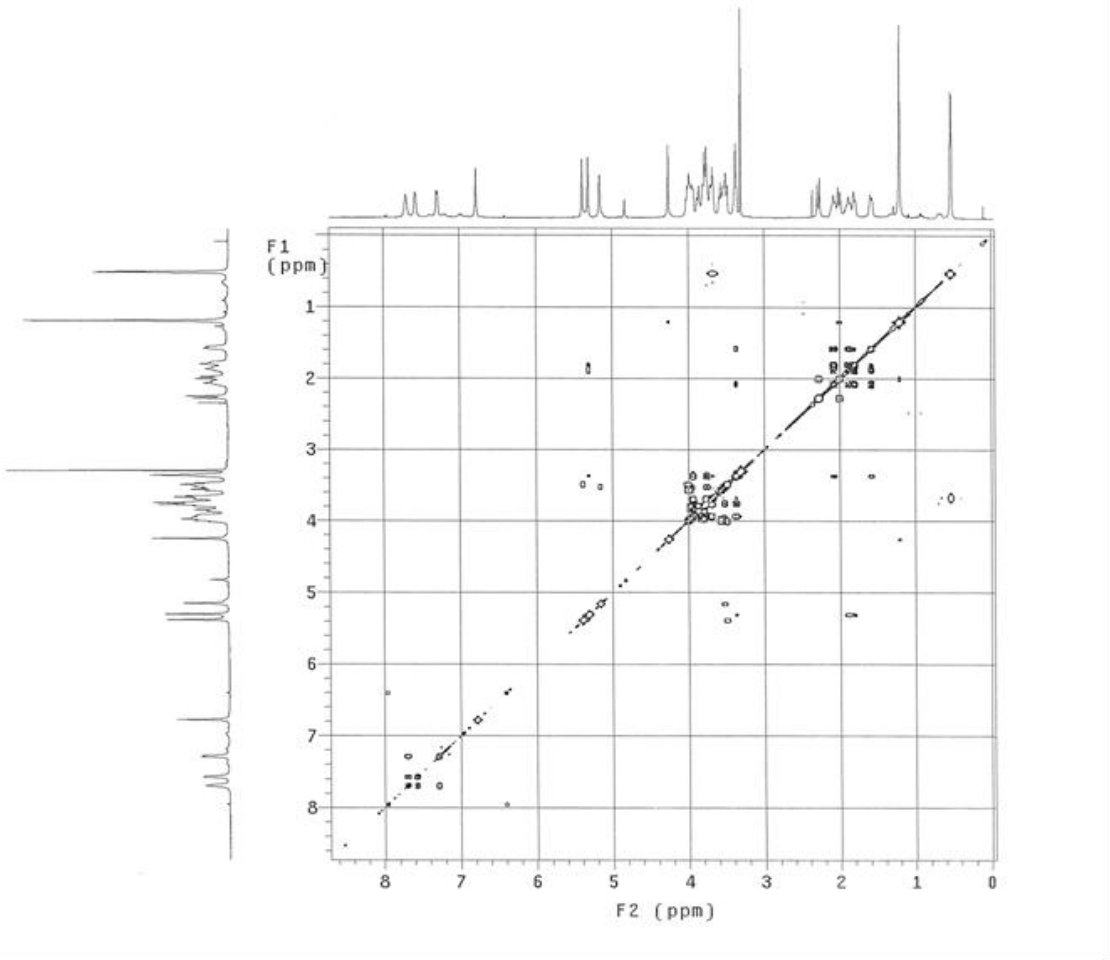


Figure 2-9. ^{1}H-^{1}H COSY (correlated spectroscopy) spectrum (CD_3OD) of BE-7585A.

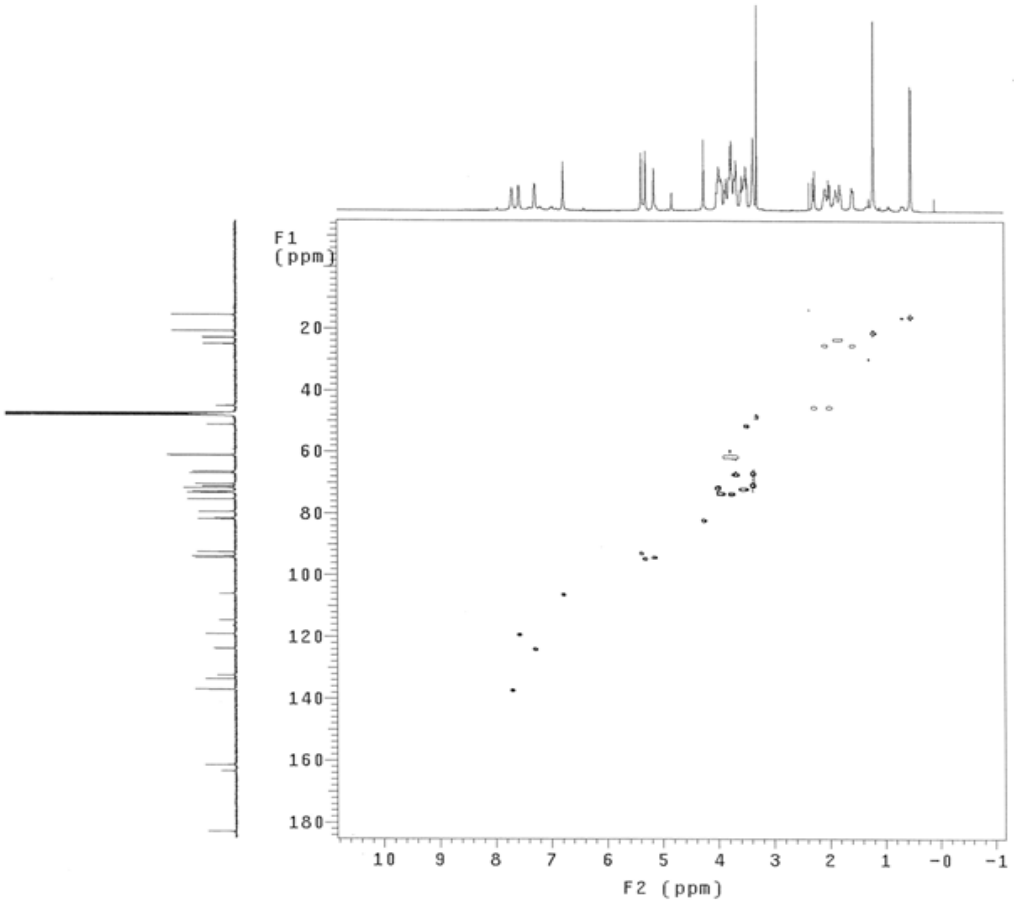


Figure 2-10. HSQC (heteronuclear single quantum coherence) spectrum of BE-7585A.

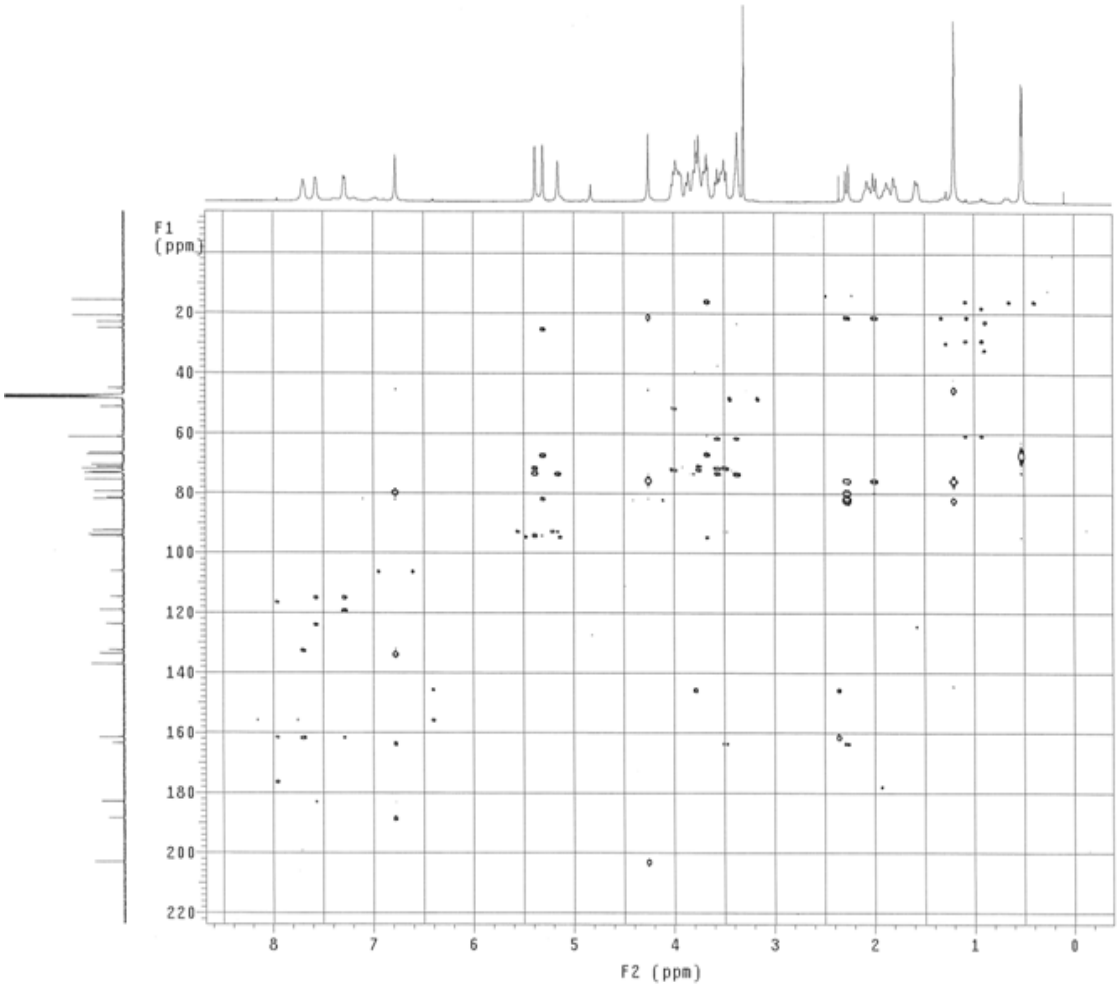


Figure 2-11. HMBC (heteronuclear multiple bond correlation) spectrum of BE-7585A.

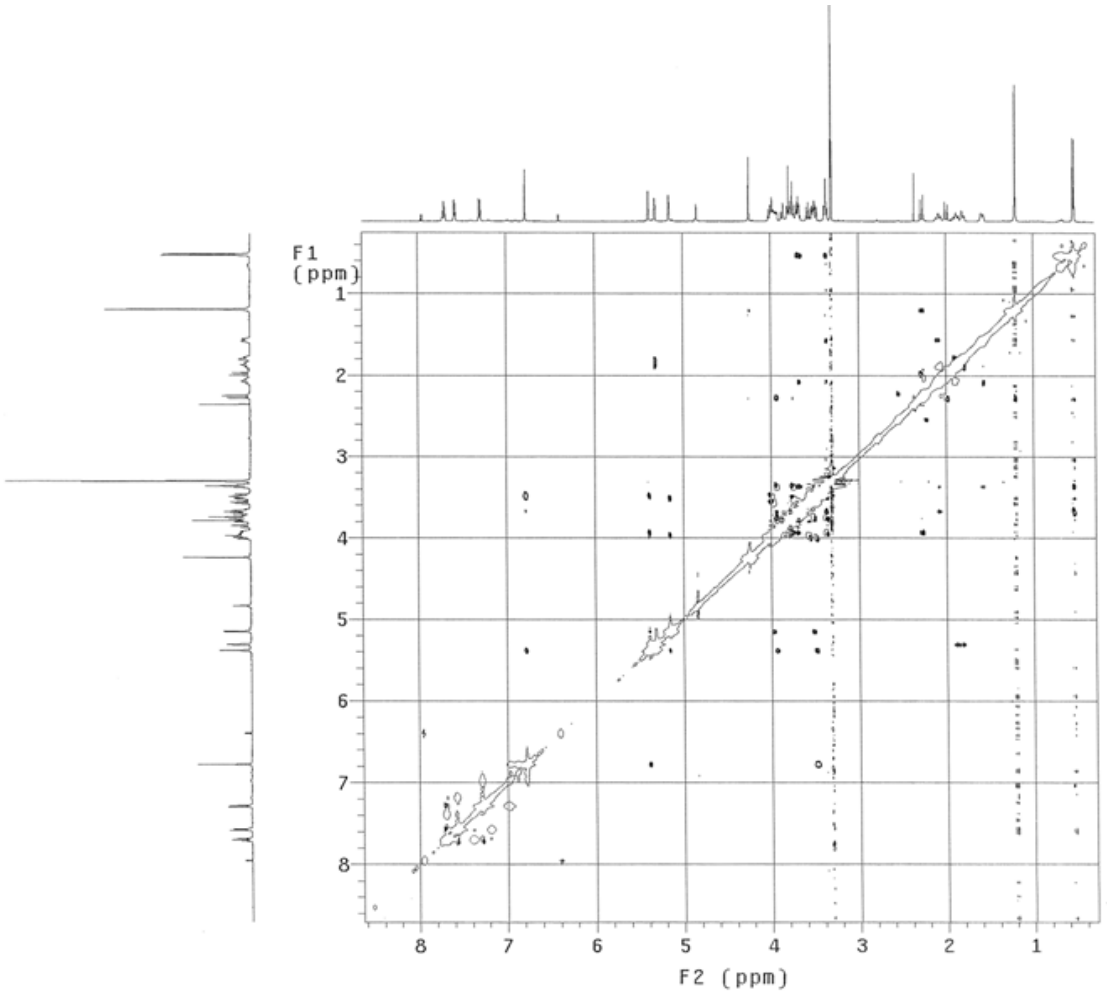
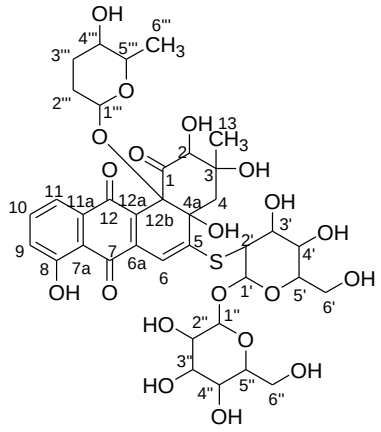


Figure 2-12. NOESY (nuclear Overhauser enhancement spectroscopy) spectrum of BE-7585A.

Table 2-2. ^{1}H and ^{13}C NMR data of BE-7585A



Measured in CD_3OD at 500 MHz (^{1}H) and 125 MHz (^{13}C)

Assignment	^{1}H δ (ppm)	^{13}C δ (ppm)	Assignment	^{1}H δ (ppm)	^{13}C δ (ppm)
1	-	204.2	1'	5.38 (1H, d, *J* = 3.1 Hz)	93.8
2	4.24 (1H, s)	83.1	2'	3.48 (1H, dd, *J* = 11.1 Hz, 3.1 Hz)	52.4
3	-	76.6	3'	3.99 (1H, m)	72.5
4	1.98 (1H, d, *J* = 14.7 Hz)	46.3	4'	3.56 (1H, t, *J* = 9.3 Hz)	73.0
	2.26 (1H, d, *J* = 14.7 Hz)		5'	3.96 (1H, m)	74.2
4a	-	80.7	6'	3.78 (1H, m)	62.4
5	-	164.6		3.86 (1H, dd, *J* = 11.7 Hz, 1.7 Hz)	
6	6.77 (1H, s)	107.3	1''	5.15 (1H, d, *J* = 3.5 Hz)	95.1
6a	-	138.3	2''	3.51 (1H, dd, *J* = 9.8 Hz, 3.5 Hz)	73.0
7	-	189.6	3''	3.74 (1H, m)	74.4 or 74.5
7a	-	115.9	4''	3.36 (1H, m)	71.7
8	-	162.7	5''	3.92 (1H, m)	74.4 or 74.5
9	7.29 (1H, d, *J* = 8.0 Hz)	125.0	6''	3.68 (1H, m)	62.4
10	7.70 (1H, t, *J* = 8.0 Hz)	138.3		3.77 (1H, m)	
11	7.57 (1H, d, *J* = 8.0 Hz)	120.3	1'''	5.31 (1H, br d, *J* = 2.7 Hz)	95.6
11a	-	133.7	2'''	1.79 (1H, br)	24.2
12	-	184.0		1.86 (1H, br)	
12a	-	134.8	3'''	1.57 (1H, br)	26.2
12b	-	82.8		2.07 (1H, br)	
13	1.10 (3H, s)	22.1	4'''	3.36 (1H, m)	67.8
			5'''	3.68 (1H, m)	68.2

The ^{13}C chemical shift of C-2' at δ 52.4 is further upfield than what is typically expected for an *O*-methine carbon and, thus, argues for the presence of a thiol substituent at this position rather than a hydroxyl group. The connectivity of the three sugar moieties

in the structure is established by HMBC analysis. The results clearly reveal the presence of a 1',1''-*O*-glycosyl bond between the 2-thioglucose and the glucose moieties (Figure 2-13). The $^{3}J_{H\text{-}H}$ values (3.1−3.5 Hz) of the anomeric protons (H-1', H-1'', and H-1''') and the NOESY experiment established an α,α-linkage between the disaccharide, and an anomeric α-configuration of the rhodinose moiety (**2-2**) to C-12b of the aglycone (Figure 2-14). In addition, the observed HMBC correlation between H-2' and C-5 implicates a thioether linkage between the 2-thiosugar and the aglycone (Figure 2-13). These results clearly distinguish the structure of BE-7585A (**2-1**) from its constitutional isomer, rhodonocardin A (**2-5**), and, thus, exclude the possibility that **2-1** is identical to rhodonocardin A (**2-5**).

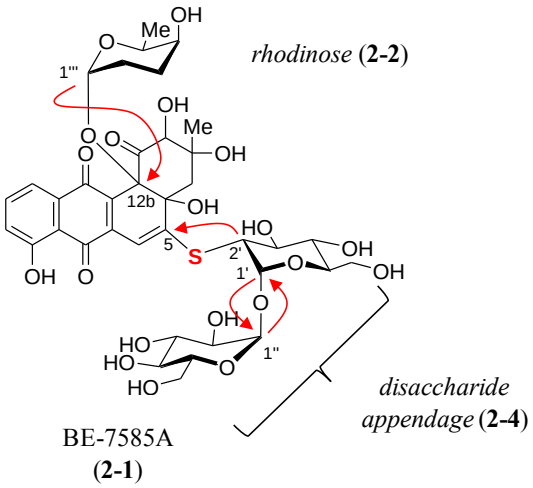


Figure 2-13. Selected HMBC results for BE-7585A.

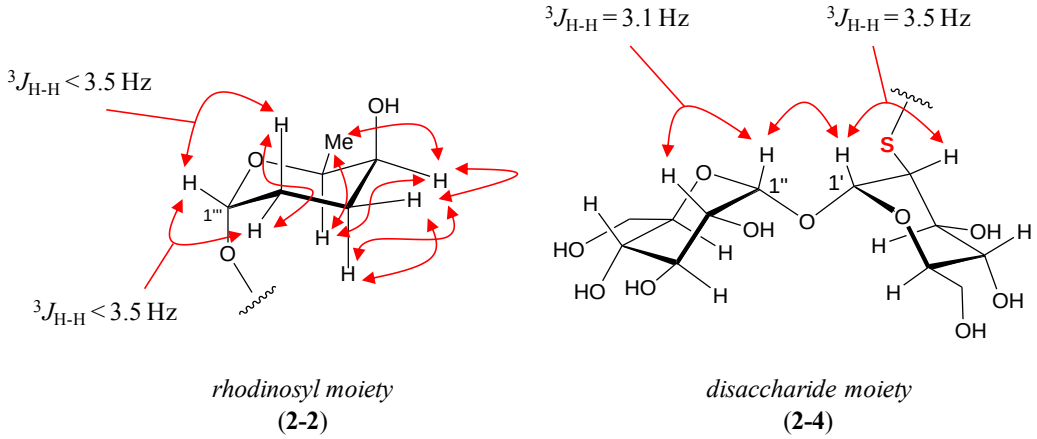


Figure 2-14. Selected NOESY results for rhodinosyl and disaccharide moieties of BE-7585A.

2.3.2 ^{13}C-Labeling Experiments.

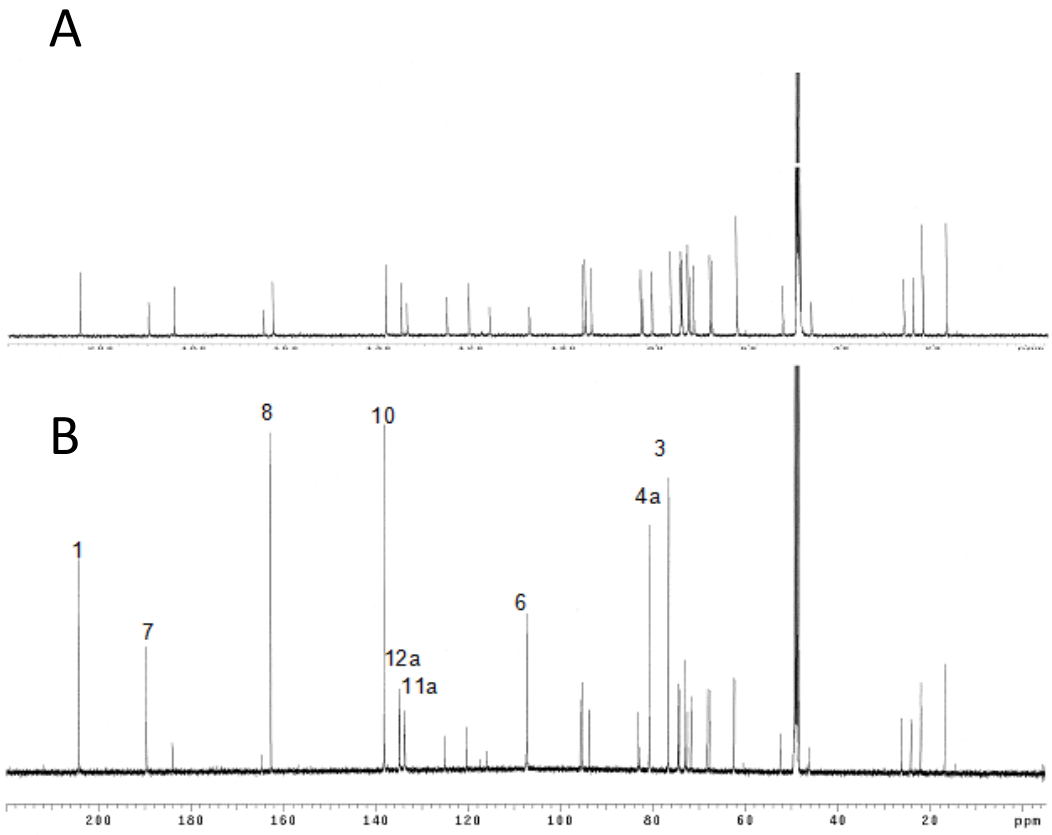
The angucycline core structures are generally derived from the assembly of 10 acetate units catalyzed by type II polyketide synthases (PKSs) as described above (Figure 2-2).[93] To confirm the biosynthetic origin of the ring carbons of BE-7585A (**2-1**), [1-^{13}C]acetate (**2-27**) was added into a growing *A. orientalis* culture. The ^{13}C-labeled product was purified and analyzed by ^{13}C NMR spectroscopy. Enhanced signals for C-1, C-3, C-4a, C-6, C-7, C-8, C-10, C-11a, and C-12a of the ring structure of **2-1** are observed, indicating these carbons are derived from C-1 of acetate (Figure 2-15). To determine the orientation of these acetate units during assembly, [1,2-$^{13}C_2$]acetate (**2-6**) was also utilized in the feeding experiment. The ^{13}C-^{13}C spin couplings were observed for C-1/C-2, C-3/C-13, C-4a/C-12b, C-5/C-6, C-6a/C-7, C-7a/C-8, C-9/C-10, C-11/C-11a, and C-12/C-12a (Figure 2-16, Table 2-3), indicating that these pairs of carbons are derived from intact acetate units. This assignment was further confirmed by the INADEQUATE (incredible natural abundance double quantum transfer experiment)

spectrum (Figure 2-17). Thus, the position and orientation of acetate units in the angucycline core structure of **2-1** were unambiguously determined.

Interestingly, the acetate incorporation pattern observed here is unusual for angucycline class natural products. As exemplified by the biosynthesis of urdamycin A (**2-10**),[95] vineomycin A1 (**2-9**),[94,118] PD116740 (**2-11**),[118] kinamycin D (**2-12**),[97,119,120] gilvocarcins (**2-13, 2-30**),[99,100] in the typical ring closure reaction, the linear decaketide (**2-7**) formed through the action of PKS is cyclized into the benz[*a*]anthracene backbone (**2-8**) in a manner in which the head (C-20) and the tail (C-2) of the decaketide chain coincide with C-13 and C-2, respectively, in the cyclic product (**2-7** $\rightarrow$ **2-8**) (Figure 2-18, pathway A). A decarboxylation of the last acetate unit (at C-2 to release C-1) is a necessary step in this pathway to account for the incorporation pattern.

While such a mode of skeletal assembly involving simple folding and condensation is found for most augucycline-type natural products studied thus far, an alternative pathway has been found in the biosynthesis of PD116198 (**2-14**).[101] In this case, chain elongation and ring condensation lead to a linearly fused tetracyclic anthracyclinone intermediate (**2-28**). Subsequent oxidative bond cleavage between C-10a and C-11 followed by rearrangement and recylcization are proposed to give the angular angucyclinone core (**2-28** $\rightarrow$ **2-29**) (Figure 2-18, pathway B). Since BE-7585A (**2-1**) displays the same acetate incorporation pattern as PD116198 (**2-14**), an analogous ring reconstruction via an enzyme-catalyzed Baeyer-Villiger oxidation may also be operative in the biosynthesis of **2-1**. Similar oxidative C−C bond cleavages of polyketide-derived tetracyclic substrates have also been proposed for the biosynthesis of gilvocarcin V (**2-30**) and jadomycin A (**2-31**) based on gene disruption and gene complementation experiments.[121,122] Enzymes responsible for this type of reaction are flavin-dependent monooxygenases. Notably, a flavin monooxygenase (MtmOIV) catalyzing the Baeyer-

Villiger ring-opening in the biosynthesis of mithramycin (**2-32**) was recently characterized and its crystal structure determined (Figure 2-19).[123,124]

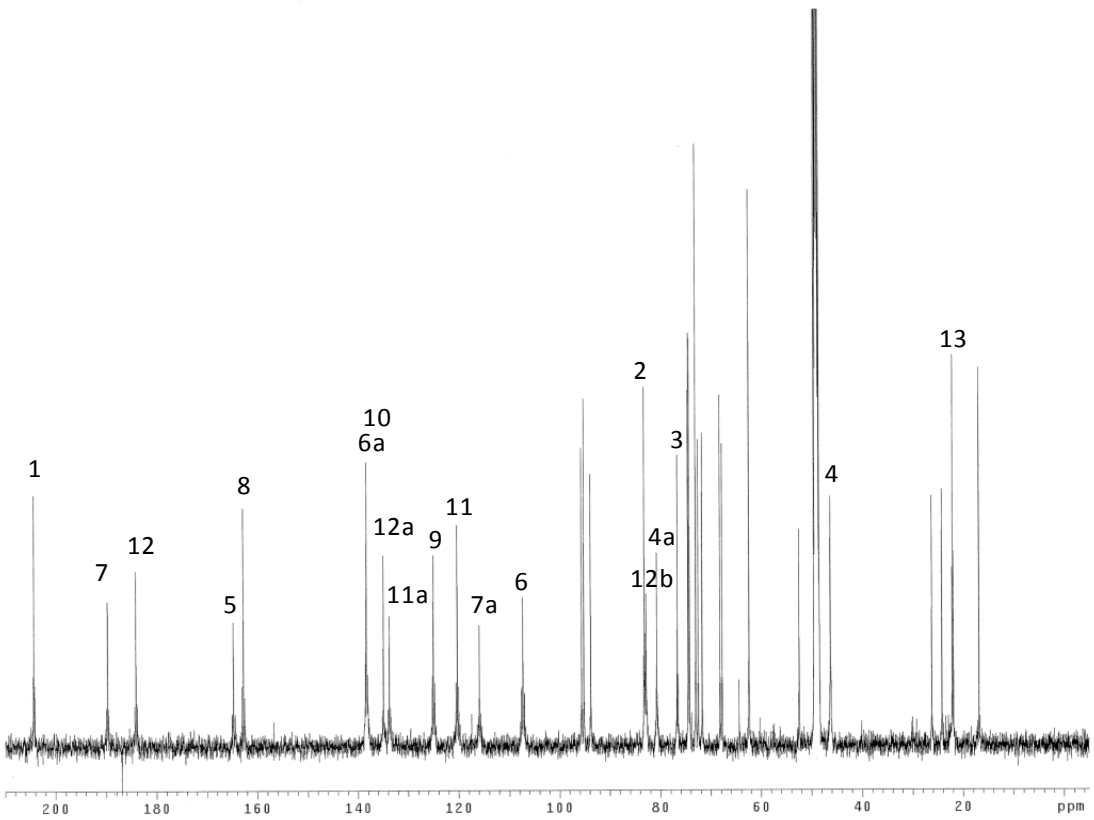


BE-7585A
(**2-1**)

Sodium[1-^{13}C]acetate
(**2-27**)

Figure 2-15. Sodium [1-^{13}C]acetate feeding experiment.

(A) ^{13}C NMR spectrum of **2-1** (unlabeled). (B) ^{13}C NMR spectrum of **2-1** labeled with ^{13}C derived from **2-27**. Incorporation ratio of **2-27** was estimated as 4%. The spectra were measured in CD_3OD at 125 MHz.



BE-7585A
(**2-1**)

Sodium[1,2-$^{13}C_2$]acetate
(**2-6**)

Figure 2-16. Sodium [1,2-$^{13}C_2$]acetate feeding experiment.

^{13}C NMR spectrum was measured in CD_3OD at 125 MHz.

Table 2-3. ^{13}C NMR data of BE-7585A obtained from sodium [1,2-$^{13}C_2$]acetate feeding experiment.

Assignment	^{13}C (ppm)	J_{CC} (Hz)	Assignment	^{13}C (ppm)	J_{CC} (Hz)
1	204.2	40.6	1'	93.8	-
2	83.2	40.6	2'	52.5	-
3	76.6	39.5	3'	72.5	-
4	46.3	-	4'	73.0	-
4a	80.7	38.6	5'	74.3	-
5	164.7	71.6	6'	62.5	-
6	107.3	71.6	1"	95.2	-
6a	138.3	51.1	2"	73.0	-
7	189.7	51.1	3"	74.4 or 74.6	-
7a	115.9	63.0	4"	71.8	-
8	162.7	63.0	5"	74.4 or 74.6	-
9	125.0	57.5	6"	62.5	-
10	138.2	57.5	1"'	95.6	-
11	120.3	61.5	2"'	24.2	-
11a	133.7	61.5	3"'	26.3	-
12	184.1	57.0	4"'	67.8	-
12a	134.9	57.0	5"'	68.2	-
12b	82.8	38.6	6"'	16.9	-
13	22.1	39.5			

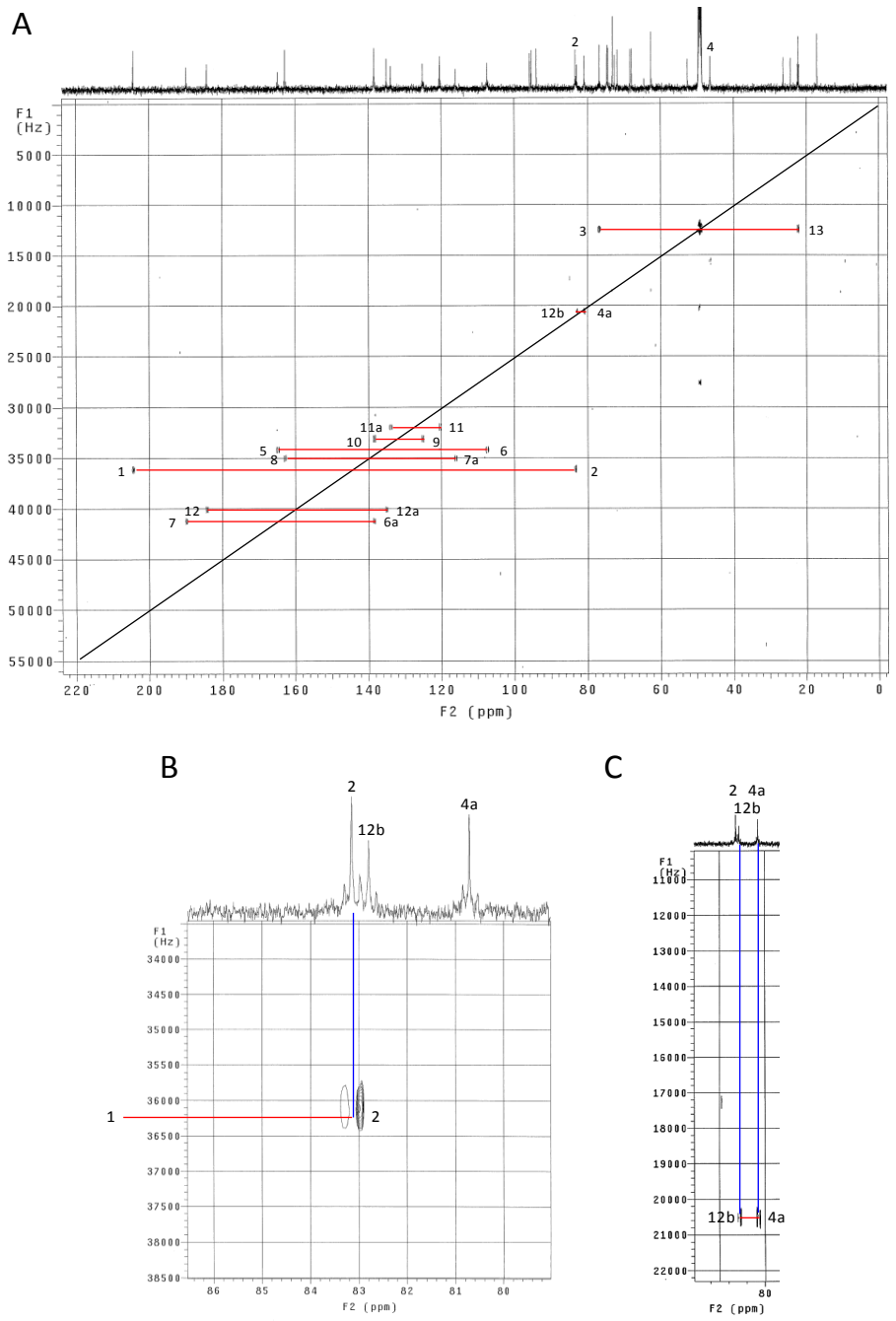


Figure 2-17. INADEQUATE spectrum of BE-7585A labeled with ^{13}C.

Sodium [1,2-$^{13}C_2$]acetate feeding experiment**.** (A) Entire spectrum. (B) Magnified spectrum showing the C1−C2 correlation. (C) Magnified spectrum showing the C-4a−C-12b correlation.

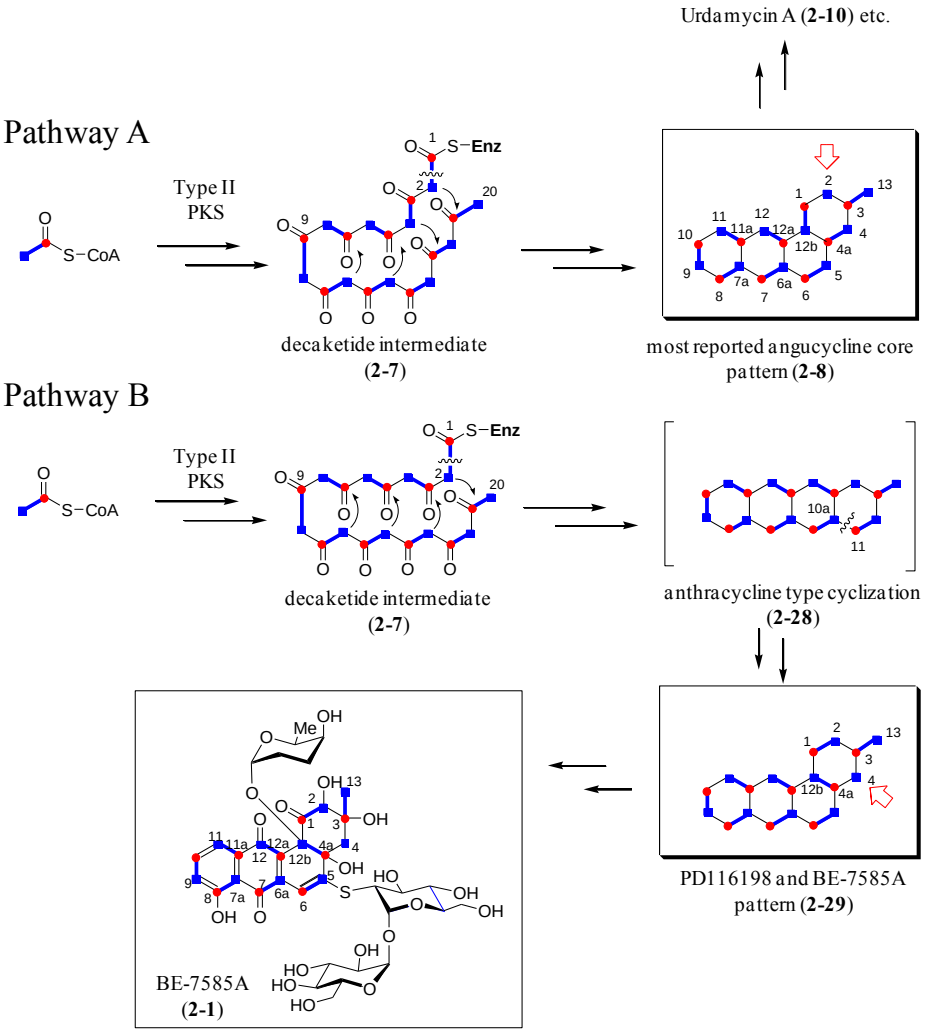
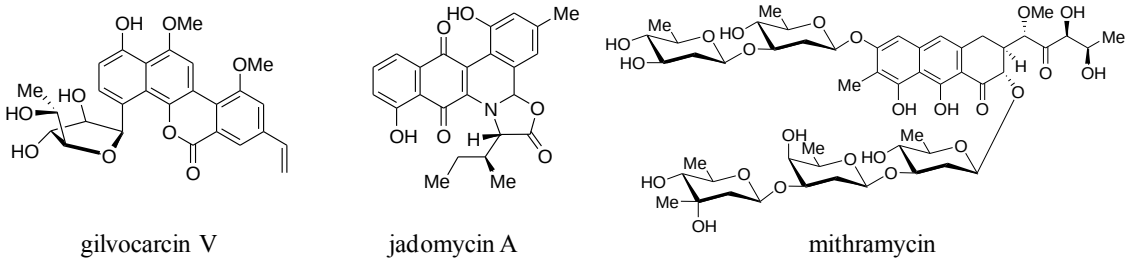


Figure 2-18. Two biosynthetic pathways for angucycline core formation.



gilvocarcin V
(**2-30**)

jadomycin A
(**2-31**)

mithramycin
(**2-32**)

Figure 2-19. Structures of gilvocarcin V, jadomycin A and mithramycin.

2.3.3 Identification of the Gene Cluster.

As mentioned above, urdamycin A (**2-10**)[125] is an angucycline class antibiotic having a similar structure as BE-7585A (**2-1**). It is composed of a benz[*a*]anthraquinone core, two L-rhodinose (**2-2**), and two D-olivose moieties, with one of the rhodinosyl groups attached at the C-12b position, as seen in the structure of **2-1** (Figure 2-4). The gene cluster of **2-10** has been identified and sequenced, and the biosynthetic pathway of **2-10** has been studied.[126,127] In view of the close resemblance between **2-10** and **2-1**, homologous genes encoding the type II polyketide synthase (PKS), the aglycone tailoring enzymes, TDP-rhodinose biosynthetic enzymes, and a rhodinosyltransferase found in the urdamycin gene cluster are expected to exist in the gene cluster of **2-1** as well.

To locate the BE-7585A biosynthetic gene cluster, a cosmid library was constructed from the genomic DNA of *A. orientalis*. This library was screened using PCR probes designed based on multiple sequence alignment of several KS_{α}s, TDP-hexose 2,3-dehydratases and TDP-hexose 3-dehydrases found in the gene clusters of urdamycin A (**2-10**) and several other related compounds. Five cosmids were found to harbor these genes. Sequencing of these cosmids led to a contiguous DNA sequence of 48.4 kb, from which 42-open reading frames (ORFs) (*orf1-42*) were identified (Figure 2-20). Among

them, *orf9−36* are believed to be involved in BE-7585A biosynthesis based on BLAST searches (Table 2-4). As predicted, the cluster includes type II PKSs, post-PKS aglycone tailoring genes, and the deoxysugar biosynthetic genes. The remaining ORFs (*orf1-8* and *orf37-42*) are likely boundary ORFs, since they appear to be the biosynthetic genes for primary metabolites (Table 2-5). Overall, the BE-7585A biosynthetic gene cluster (*bex* cluster) spans a region of 28.9 kb and contains 28 ORFs. The nucleotide sequences reported herein have been deposited into the GenBank database under the accession number HM055942.

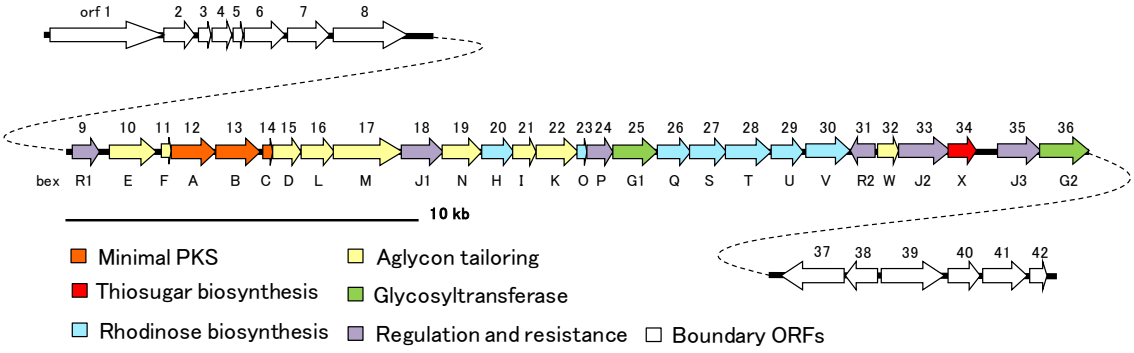


Figure 2-20. Organization of the BE-7585A biosynthetic gene cluster (*bex* cluster). The gene cluster spans 28.9 kbp and contains 28 open reading frames.

Table 2-4. Proposed functions of ORFs in the BE-7585A biosynthetic gene cluster (*bex* cluster).

gene	proposed function	protein homologue and origin	identity / similarity (%)	protein accession number
bexR1	regulator	repressor-response regulator [*Streptomyces* sp. AM-7161]	56 / 68	BAC79018
bexE	oxygenase	putative oxygenase (Aur1A) [*Streptomyces aureofaciens*]	75 / 83	AAX57188
bexF	cyclase	putative cyclase (UrdF) [*Streptomyces fradiae*]	78 / 85	CAA60568
bexA	ketosynthase α subunit	putative ketoacyl synthase (UrdA) [*Streptomyces fradiae*]	92 / 85	CAA60569
bexB	ketosynthase β subunit	keto-acyl synthase beta [*Streptomyces antibioticus*]	74 / 83	CAG14966
bexC	acyl carrier protein	putative acyl carrier protein (SimA3) [*Streptomyces antibioticus*]	60 / 74	AAK06786
bexD	ketoreductase	ketoreductase [*Streptomyces antibioticus*]	81 / 89	CAG14968
bexL	cyclase	putative aromatase [*Streptomyces* sp. SCC 2136]	70 / 78	CAH10113
bexM	oxygenase-reductase	oxygenase-reductase (UrdM) [*Streptomyces fradiae*]	62 / 71	AAF00206
bexJ1	transporter	Transporter (UrdJ2) [*Streptomyces fradiae*]	47 / 61	AAF00207
bexN	carboxyl transferase	decarboxylase (JadN) [*Streptomyces sviceus* ATCC 29083]	43 / 54	ZP_05020741
bexH	sugar epimerase / reductase	predicted nucleoside-diphosphate sugar epimerase [*Streptosporangium roseum* DSM 43021]	48 / 65	ZP_04471118
bexI	oxygenase	putative anthrone monooxygenase (GilOII) [*Streptomyces griseoflavus*]	27 / 38	AAP69583
bexK	hydroxylase	cytochrome P-450-like enzyme [*Saccharopolyspora erythraea* NRRL 2338]	41 / 54	YP_001107923
bexO	ferredoxin	ferredoxin [*Streptomyces coelicolor* A3(2)]	40 / 59	NP_625075
bexP	methyltransferase	putative methyltransferase [*Saccharopolyspora erythraea* NRRL 2338]	37 / 52	YP_001105373
bexG1	glycosyltransferase	glycosyl transferase (UrdGT1a) [*Streptomyces fradiae*]	49 / 62	AAF00214
bexQ	sugar 3-ketoreductase	sugar 3-ketoreductase (KijD10) [*Actinomadura kijaniata*]	52 / 65	ACB46498
bexS	sugar 4,6-dehydratase	putative dNDP-glucose 4,6-dehydratase [*Streptomyces aureofaciens*]	72 / 80	ACK77743
bexT	sugar 3-dehydrase	NDP-hexose 3,4-dehydratase homolog [*Streptomyces cyanogenus*]	78 / 86	AAD13547
bexU	TDP-glucose synthase	putative TDP-glucose synthase [*Streptomyces nogalater*]	69 / 82	AAF01820
bexV	Sugar 2,3-dehydratase	NDP-hexose 2,3-dehydratase [*Frankia* sp. CcI3]	58 / 72	YP_483211
bexR2	regulator	TetR family transcriptional regulator [*Saccharopolyspora erythraea* NRRL 2338]	50 / 61	YP_001103502
bexW	reductase	putative reductase [*Streptomyces griseoflavus*]	67 / 74	AAP69587
bexJ2	transporter	EmrB/QacA family drug resistance transporter [Frankia sp. EAN1pec]	36 / 53	YP_001507963
bexX	thiosugar synthase	Thiazole biosynthesis protein (ThiG) [*Stigmatella aurantiaca* DW4/3-1]	58 / 75	ZP_01459245
bexJ3	transporter	major facilitator superfamily transporter [*Mycobacterium vanbaalenii* PYR-1]	36 / 50	YP_956695
bexG2	glycosyltransferase	α,α-trehalose-phosphate synthase [*Methanothermobacter thermautotrophicus* str. Delta H]	30 / 47	NP_276863

Table 2-5. BLAST analysis of the boundary ORFs of the *bex* cluster.

orf	Protein with the highest sequence similarity	Identity / Similarity (%)	Protein accession number
1	carbamoyl-phosphate synthase large subunit [*Saccharomonospora viridis* DSM 43017]	88 / 94	ZP_04506718
2	orotidine 5'-phosphate decarboxylase, subfamily 2 [*Saccharomonospora viridis* DSM 43017]	72 / 80	ZP_04506717
3	hypothetical protein SvirDRAFT_1577 [*Saccharomonospora viridis* DSM 43017]	93 /97	ZP_04506716
4	guanylate kinase [*Saccharomonospora viridis* DSM 43017]	70 / 81	ZP_04506715
5	DNA-directed RNA polymerase omega subunit [*Saccharopolyspora erythraea* NRRL 2338]	96 / 98	YP_001104332
6	phosphopantothenoylcysteine synthase/decarboxylase [*Saccharomonospora viridis* DSM 43017]	80 / 87	ZP_04506713
7	methionine adenosyltransferase 1 [*Saccharopolyspora erythraea* NRRL 2338]	87 / 92	YP_001104334
8	primosomal protein N' [*Saccharomonospora viridis* DSM 43017]	71 / 80	ZP_04506710
37	secreted protein [*Saccharopolyspora erythraea* NRRL 2338]	48 / 60	YP_001105521
38	lysozyme M1 precursor [*Saccharopolyspora erythraea* NRRL 2338]	68 / 78	YP_001109221
39	putative gamma-glutamyltranspeptidase [*Streptomyces roseosporus* NRRL 15998]	66 / 77	ZP_04697280
40	methionyl-tRNA formyltransferase [*Saccharopolyspora erythraea* NRRL 2338]	75 / 86	YP_001104344
41	tRNA/rRNA cytosine-C5-methylase [*Saccharomonospora viridis* DSM 43017]	70 / 81	YP_003133431
42	flavoprotein [*Streptomyces* sp. AA4]	80 / 86	ZP_05478729

2.3.4 Functional Predication of Genes Involved in Angucycline Core Formation.

The deduced gene products of *bexA*, *bexB*, and *bexC* display high sequence similarities to the type II minimal PKS, KS_α, β-ketoacyl synthase β subunit (KS_β), and acyl carrier protein (ACP), respectively. The minimal PKS (BexA, BexB, and BexC) assembles one acetyl-CoA and 10 malonyl-CoAs from 10 acetate units into a linear decaketide chain (**2-7**). To accommodate the unusual labeling pattern observed in the ^{13}C-

labeling experiments (Figure 2-18), **2-7** should be cyclized to a linear structure **2-28** rather than an angular structure (i.e., **2-8**). On the basis of the established biosynthesis of other angucycline type compounds, a ketoreductase homologue, BexD, and two cyclase homologues, BexL and BexF, are predicted to act together with PKS to construct the ring structure.[93] Specifically, the reduction of the 9-keto group by BexD likely initiates the ring formation reaction catalyzed by BexL. Subsequent decarboxylation and cyclization by BexF yields **2-34** as a key intermediate (Figure 2-21). The resulting ring structure **2-34** would then be tailored by a set of oxygenases to the angular angucyclinone core (**2-40**).

Recent studies indicated that multioxygenase complexes are required to perform the oxidative C−C bond cleavage and ring rearrangement in both gilvocarcin (**2-30**) and jadomycin (**2-31**) biosyntheses.[122,128] Since a similar oxidative ring-opening process (**2-34** → **2-40**) likely occurs in the formation of BE-7585A, a complex of oxygenases may also be involved. A few possible candidates include a flavin-dependent oxygenase homologue, BexE, a bifunctional oxygenase-reductase homologue, BexM, and the anthrone monooxygenase homologue, BexI, which displays sequence similarity to GilOII and JadG from the gilvocarcin and jadomycin pathways, respectively.[122,129] Although it is difficult to define the cascade of the oxidation events solely based on sequence analysis, the identification of genes encoding these oxygenases (BexE, BexI, and BexM) in the *bex* cluster provides important leads for future biochemical experiments to address these tailoring oxidation reactions in the biosynthesis of BE-7585A (**2-1**).

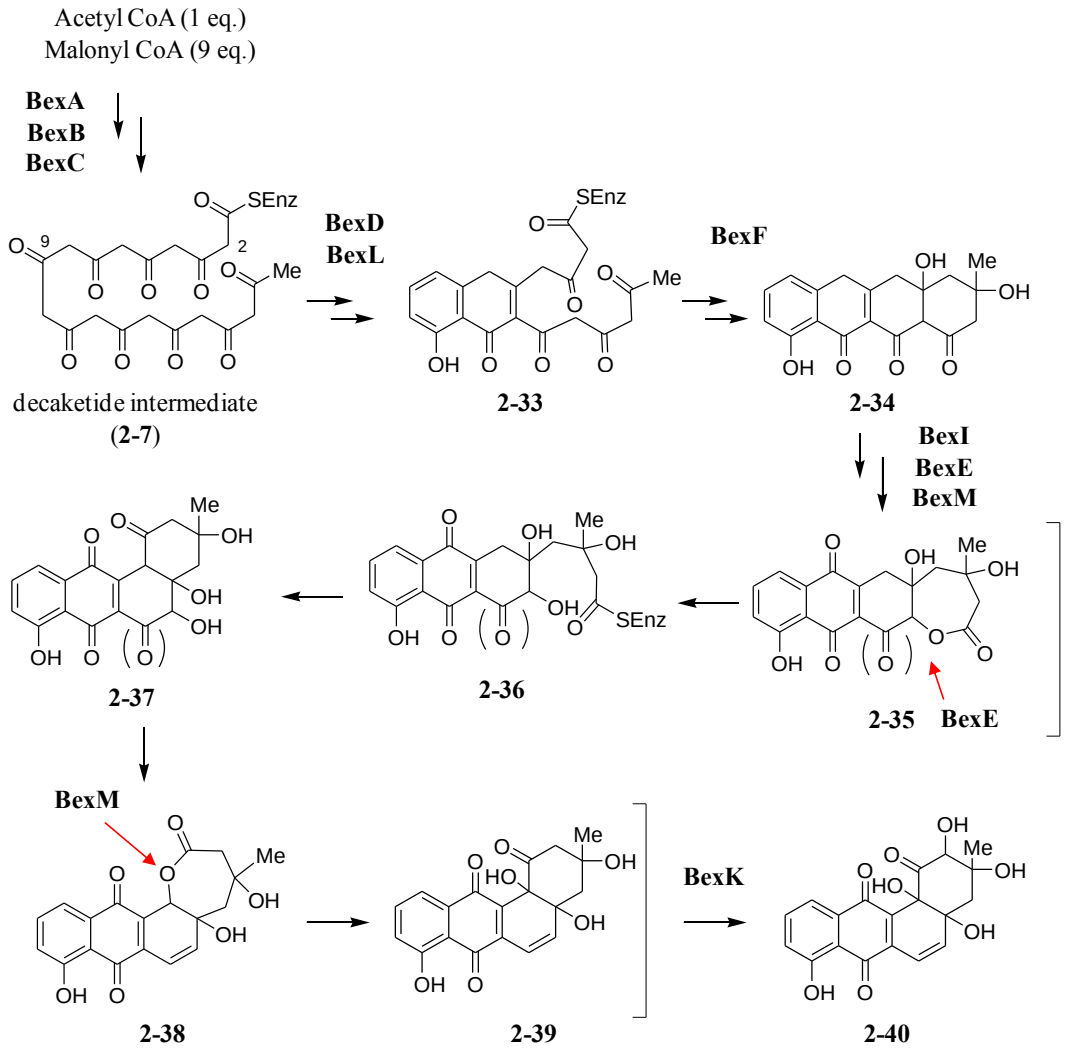


Figure 2-21. Proposed pathway of the angucycline core formation in biosynthesis of BE-7585A.

Other genes likely involved in aglycone biosynthesis are *bexK*, *bexN*, and *bexW*. BexK is a cytochrome P450 hydroxylase homologue, which may be responsible for hydroxylation at C-2 of **2-39** to produce **2-40**. This assignment is supported by an early observation that a cytochrome P450 inhibitor blocked C-2 hydroxylation in PD116198 (**2-14**) biosynthesis, implicating C-2 hydroxylation in both pathways as a P450-dependent

reaction.[130] Similar *in vivo* experiments using a P450 inhibitor, metyrapone, is described below. BexN is an acetyl-CoA carboxylase homologue, which may be involved in supplying malonyl-CoA in PKS-catalyzed reactions. Precedence is found in the jadomycin pathway where a BexN homologue, JadN, is believed to play a similar role.[131] BexW is a NADPH-dependent FMN reductase homologue, which may serve as the reductase for the flavin-dependent oxygenases involved in the aglycone biosynthesis.

2.3.5 Functional Prediction of Genes Involved in Rhodinose Formation and Rhodinosyltransfer.

Rhodinose and rhodinose derivatives are found in several natural products whose biosynthetic gene clusters have been isolated and sequenced (Figure 2-4).[102,103,105,106,132] Sequence comparison of the putative sugar biosynthetic genes in these clusters and those in the *bex* cluster allowed the identification of the rhodinose biosynthetic genes. Figure 2-22 shows the correlation of their proposed functions and the chemical transformation steps required to convert glucose 1-phosphate (**2-19**), which is the biosynthetic precursor for most deoxyhexoses, to TDP-L-rhodinose (**2-26**).[14,16] The pathway is initiated by the activation of **2-19** to TDP-glucose (**2-20**), catalyzed by a TDP-glucose synthase homologue, BexU. The second step is the conversion of **2-20** to **2-21** catalyzed by a TDP-glucose 4,6-dehydratase homologue, BexS. Subsequent C-2 deoxygenation (**2-21** $\rightarrow$ **2-23**) is catalyzed by a 2,3-dehydratase homologue, BexV, and a 3-ketoreductase homologue, BexQ. The resulting product **2-23** then undergoes C-3 deoxygenation to **2-24** catalyzed by a [2Fe-2S]-containing and PMP-dependent BexT, which displays good sequence homology to hexose 3-dehydrase enzymes (78% identity to LanQ from S. cyanogenus; 73% identity to SpnQ from *Saccharopolyspora spinosa*).[115,133] While this reaction typically requires an iron-sulfur containing flavoprotein reductase for coenzyme regeneration,[134,135] no gene encoding such a specific reductase can be located in the bex

cluster. It is possible that a general cellular reductase or BexO, which is a ferredoxin homologue, assumes this role.[115,136]

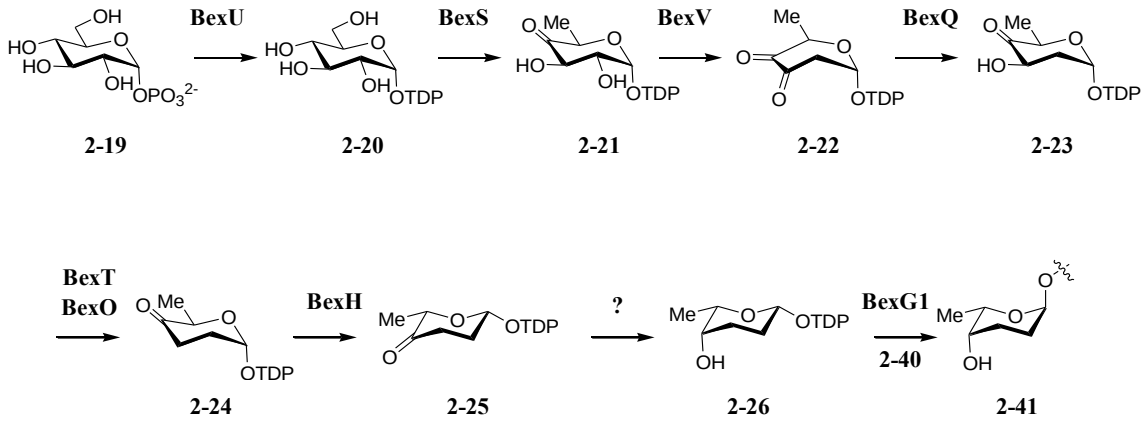


Figure 2-22. Proposed pathway of the rhodinose formation and its transfer to the aglycone in biosynthesis of BE-7585A.

The *bex* cluster also contains the gene bexH that is believed to encode a sugar epimerase. Thus, a 5-epimerization reaction (**2-24** → **2-25**) followed by 4-ketoreduction (**2-25** → **2-26**) is proposed to complete the biosynthesis of TDP-L-rhodinose (**2-26**). However, no hexose 4-ketoreductase gene can be located in the *bex* cluster. Whether the 4-ketoreduction of **2-25** is mediated by a nonspecific reductase in the cellular pool or BexH functions as a bifunctional epimerase-ketoreductase[137,138] will be determined in future biochemical studies. Further support for the assignment of **2-26** as an L-sugar comes from another gene product in the *bex* cluster, BexG1. This glycosyltransferase homologue exhibits high sequence similarity to the TDP-L-rhodinosyltransferase, UrdGT1a, from the urdamycin pathway having 49% identity and 62% similarity.[139] Therefore, BexG1 is predicted to have a similar function in catalyzing the transfer of the

L-rhodinosyl group of **2-26** to the C-12b position of the angucycline core (**2-40**) to generate **2-41**. Since the UrdGT1a reaction proceeds with inversion of stereochemistry at the anomeric center of rhodinose and both BexG1 and UrdGT1a are members of the GT-1 glycosyltransferase family of enzymes,[140] BexG1 is expected to be an inverting enzyme as well. This implies that the TDP leaving group of the sugar donor should be β-oriented as shown in **2-26**, because the glycosidic linkage of the rhodinose moiety in **2-41** is in the α-configuration. An epimerization at C-5 of **2-24** is, thus, necessary to fix the anomeric configuration as β in **2-25** and **2-26**.

2.3.6 Biosynthesis of the Thiosugar Moiety.

Among all ORFs in the *bex* cluster, the *bexX* gene, which displays good sequence similarity to thiazole synthases, ThiG (58% identity to ThiG from *Stigmatella aurantiaca* DW4/3-1;[141] 38% identity to ThiG from *Bacillus subtilis* subsp. *subtilis* str 168[142]) is likely important for the biosynthesis of the thiosugar (**2-3**). ThiG is a key enzyme involved in the formation of thiazole of thiamine. Studies of thiamine biosynthesis in *B. subtilis* showed that the ThiG reaction is initiated by the Schiff base formation between one of lysine residues (Lys96) of ThiG and the 2-keto group of its substrate, 1-deoxy-D-xylulose-5-phosphate (DXP) (see Chapter 1, Figure 1-16). The sulfur atom is transferred from the thiocarboxylate at the *C*-terminus of a sulfur carrier protein, ThiS, and incorporated into the ThiG-DXP adduct.[65,143] Because of the presence of the *bexX* gene in the cluster, an analogous mechanism of sulfur incorporation can be proposed for the biosynthesis of 2-thioglucose (Figure 2-23).

Accordingly, the putative substrate, glucose 6-phosphate (**2-42**), may form an imine adduct (**2-43**) with the active site lysine of BexX at the C-1 position. Deprotonation of C-2 of **2-43** followed by tautomerization via an enamine **2-44** generates a 2-keto

intermediate **2-45**, which can then react with a nucleophilic sulfur group to incorporate a sulfur atom at C-2 in **2-46**. The reactive sulfur donor is likely a ThiS-thiocarboxylate or a protein persulfide equivalent. Dehydration of **2-46** to **2-47** followed by C-1 deprotonation gives an enamine intermediate **2-48**, which can tautomerize to C-1 imine **2-49**. Upon hydrolysis, the BexX-bound product **2-50** is released to give 2-thioglucose 6-phosphate (**2-51**).

The sulfur atom of ThiS-thiocarboxylate has been shown to be derived from cysteine, mediated by a cysteine desulfurase.[65] In *B. subtilis* (and *E. coli)*, this sulfur loading reaction also requires ThiF, which plays a role in ThiS activation.[61,144] Unfortunately, neither a ThiS homologue nor a cysteine desulfurase homologue is found in the *bex* cluster. However, since the substrate and the function of ThiG and BexX are analogous, it is possible that the endogenous ThiS used in thiamin biosynthesis may be recruited to serve as the sulfur donor for thiosugar biosynthesis in *A. orientalis*. Further investigation of the sulfur delivery process is performed in Chapter 4.

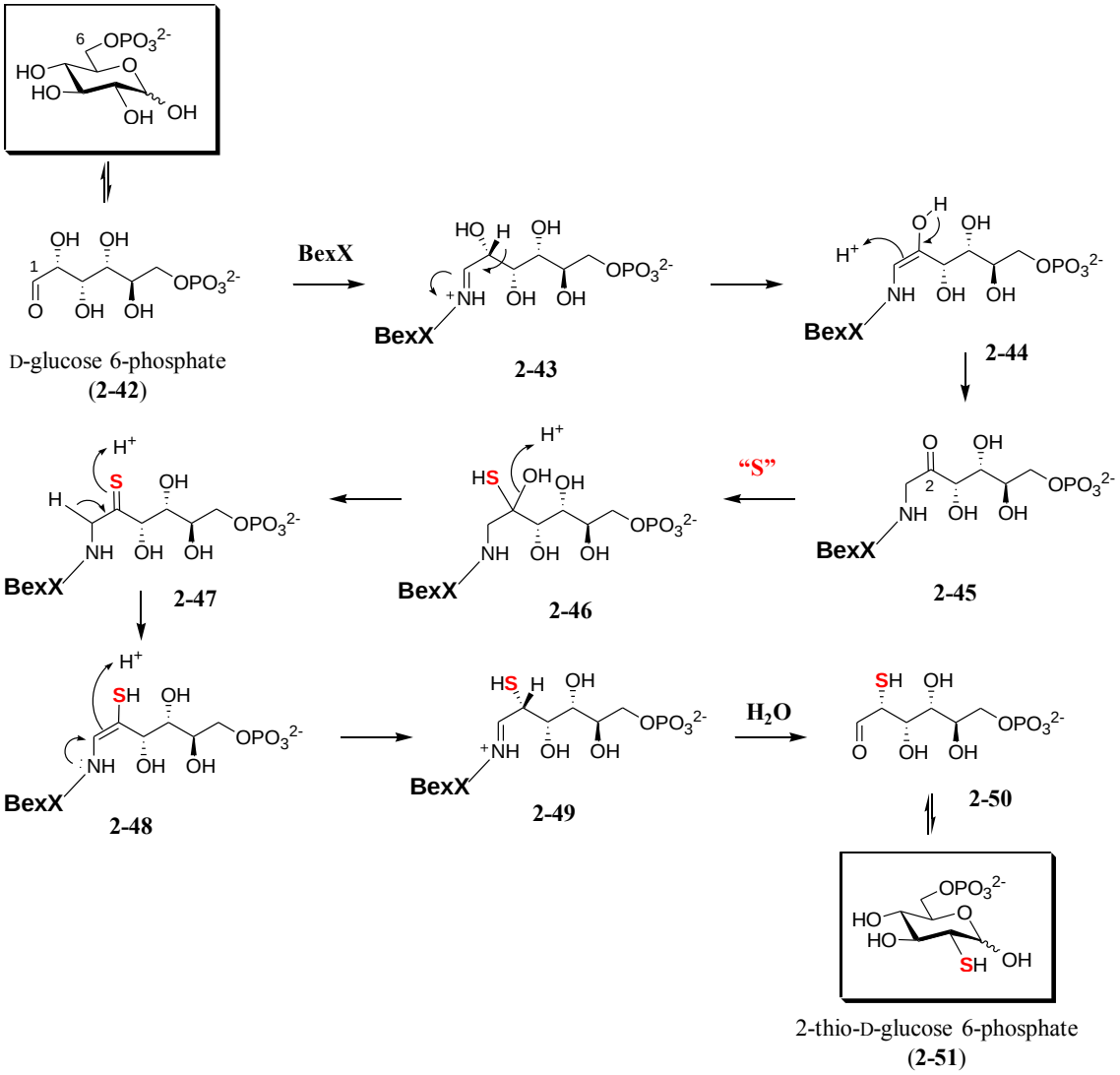


Figure 2-23. Proposed pathway and mechanism of 2-thiosugar formation in the biosynthesis of BE-7585A.

2.3.7 Pathway of the Disaccharide Formation and Its Transfer to the Aglycone.

Once formed, the BexX product, 2-thioglucose 6-phosphate (**2-51**), is expected to accept a glucose moiety from UDP-glucose (**2-52**) to form 2-thiotrehalose 6-phosphate (**2-53**). This disaccharide moiety of BE-7585A (**2-1**) is structurally related to trehalose. The latter is commonly formed by linking the glucose moiety of UDP-glucose (**2-52**) to

glucose 6-phosphate (**2-42**), and the reaction is catalyzed by trehalose 6-phosphate synthases (TPSs).[145] Since the *bexG2*-encoded protein exhibits sequence similarity to TPSs (30% identity and 47% similarity to TPS from *Methanothermobacter thermautotrophicus* str. Delta H[146]), BexG2 is likely the 2-thiotrehalose 6-phosphate synthase catalyzing the coupling of **2-51** and **2-52** to form **2-53**, whose 6-phosphate group is removed at a later stage (Figure 2-24).

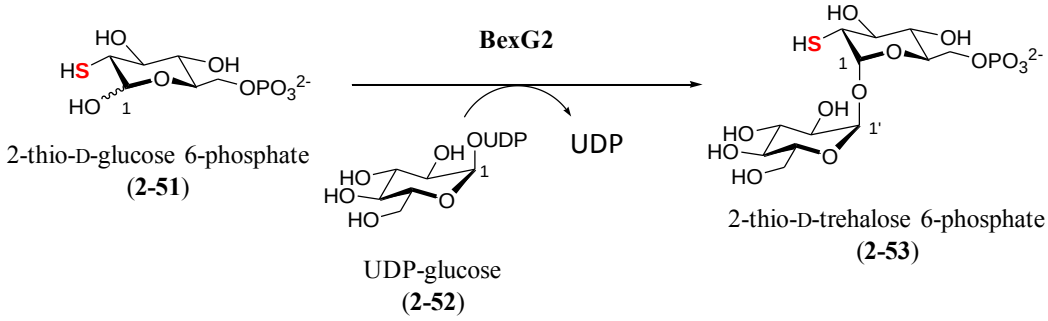


Figure 2-24. Proposed pathway of disaccharide formation in the biosynthesis of BE-7585A.

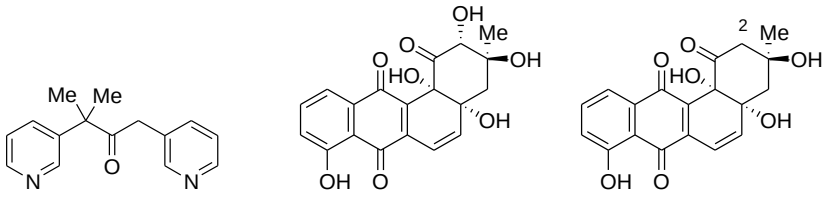
How the thiodisaccharide (**2-53** or its hydrolyzed product without 6-phosphate group) is attached to C-5 of the benz[*a*]anthraquinone core is not clear. It is conceivable that the thioether bond formation may simply be a result of a nonenzymatic attack by the nucleophilic 2-thiol group of the disaccharide at the C-5 position of benz[*a*]anthraquinone (**2-40** or **2-41**). Functional characterization of BexG2 and the possible nonenzymatic C−S bond formation are described in Chapter 3.

2.3.8 Regulatory and Resistant Genes.

The remaining genes in the BE-7585A gene cluster are mainly regulatory and resistance genes. Among them, BexR1 and BexR2 display sequence similarity to transcriptional regulators. BexJ1, BexJ2, and BexJ3 are transporter homologues and are possibly involved in self-resistance mechanisms.[147] A methyltransferase homologue, BexP, may also be involved in the self-resistance mechanism by catalyzing methylation of rRNA.

2.3.9 Metyrapone Feeding Experiments.

As described above, sequence analysis displayed that BexK is a putative cytochrome P450 hydroxylase homologue, which may be responsible for hydroxylation at C-2 of **2-39** to produce **2-40**. Similar hydroxylation was predicted in PD 116198 (**2-14**) biosynthesis based on results from blocking the hydroxylation by feeding a cytochrome P450 inhibitor, metyrapone (**2-54**). The resulting cell culture of *Streptomyces phaeochromogenes* WP 3688 produced a less oxygenated compound **2-55** (Figure 2-25)[130] Thus, similar feeding experiments using metyrapone was performed with the BE-7585A (**2-1**) producer strain, *A. orientalis.*



metyrapone (**2-54**) PD 116198 (**2-14**) **2-55**

Figure 2-25. Structures of metyrapone, PD 116198 and its less oxygenated analogue.

Interestingly, HPLC analysis of the supernatant of *A. orientalis* culture containing metyrapone (**2-54**) showed clear suppression of the BE-7585A (**2-1**) production level compared to the control culture without metyrapone. More importantly, a large peak (**2-56**), which is also seen as a very minor product in the control culture, was eluted about 2 min after **2-1** (Figure 2-26). While the mass of this peak (**2-56**) was confirmed to be 16-Da less than that of BE-7585A (**2-1**), the acquired UV-vis spectra from LC were almost identical for both **2-1** and **2-56** (Figure 2-27). These results are consistent with the predicted inhibition of C-2 hydroxylation in the biosynthesis of BE-7585A (**2-1**) and the resulting formation of less oxygenated 2-deoxy BE-7585A (**2-56**). Since BexK is the only P450 type hydroxylase homologue in the *bex* cluster, the results shown here strongly support the proposed function of BexK (C-2 hydroxylation of **2-39** to produce **2-40**, Figure 2-20).

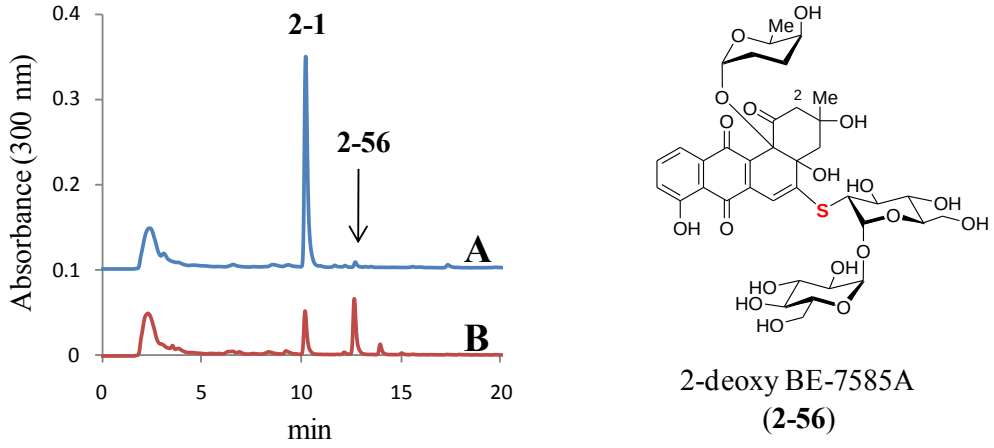


Figure 2-26. Metyrapone Feeding Experiments (HPLC analysis).

HPLC analyses of the supernatants from *A. orientalis* culture and the predicted structure of the product (**2-56**). **A**. Control culture without using metyrapone. **B**. Metyrapone-feeding culture.

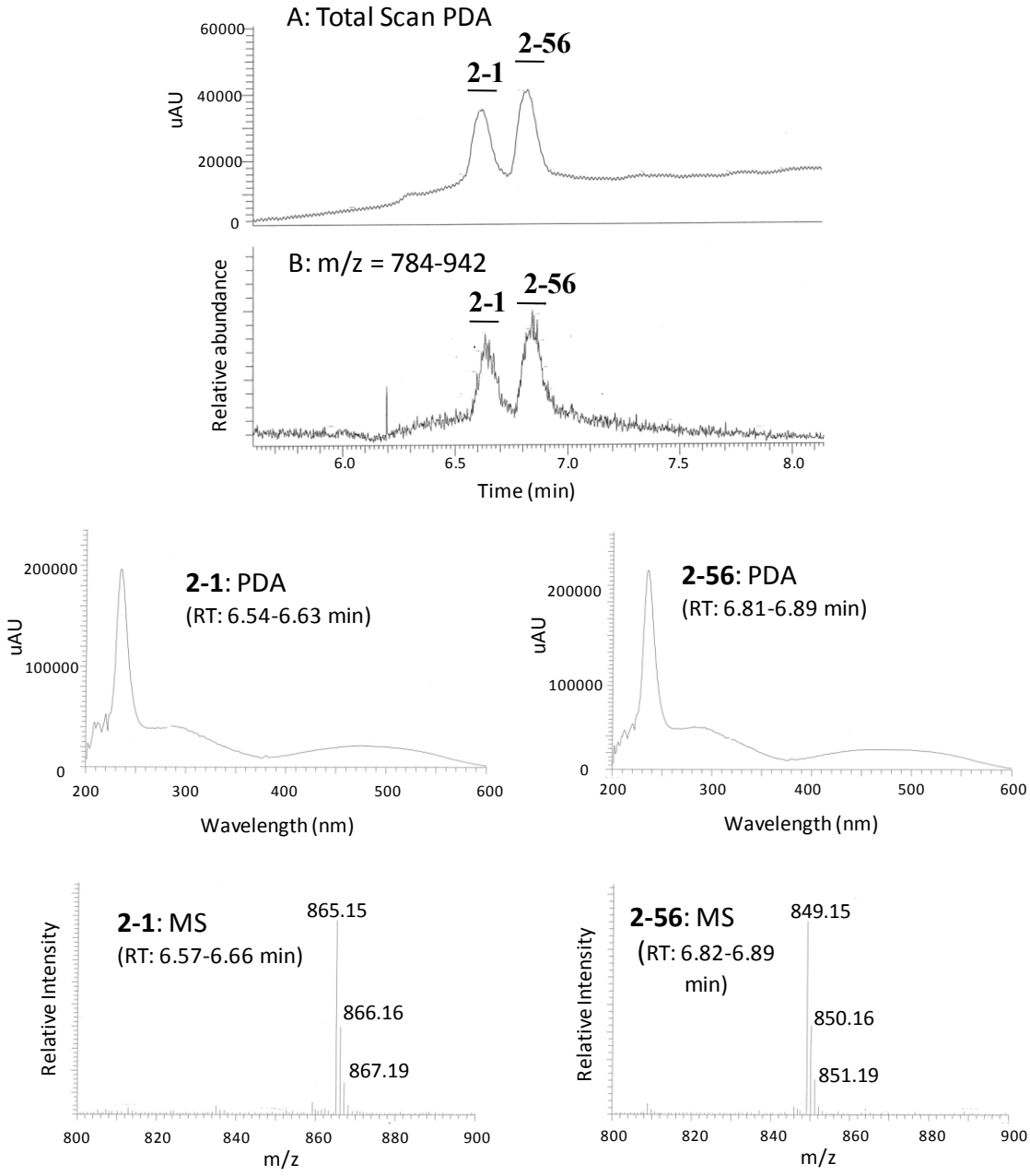


Figure 2-27. Metyrapone Feeding Experiments (LC-MS analysis).

LC-MS/PDA analyses of the supernatants from *A. orientalis* culture with metyrapone. BE-7585A (**2-1**) MS calcd for $C_{37}H_{46}O_{20}NaS^+$ $[M + Na]^+$ 865.2, found 865.2. **2-56** MS calcd for $C_{37}H_{46}O_{19}NaS^+$ $[M + Na]^+$ 849.2, found 849.2.

2.4 CONCLUSIONS

In this chapter, the structure of BE-7585A (**2-1**) from *A. orientalis* was first characterized by a series of 1D- and 2D-NMR experiments. The determined three glycosydic linkages clearly show that **2-1** is a constitutional isomer of rhodonocardin A (**2-5**) but not an identical compound. The BE-7585A biosynthetic gene cluster (*bex* cluster) containing 28 ORFs was then identified by PCR-based cosmid screening of the genomic DNA of the producing strain. The cluster harbors genes typical for type II polyketide synthesis. Also contained in the cluster is a thiazole synthase (ThiG) homologue, BexX, which may catalyze the key reaction producing 2-thioglucose 6-phosphate (**2-51**) from glucose 6-phoshate (**2-42**). On the basis of sequence analysis, a biosynthetic pathway is proposed for BE-7585A including the benz[*a*]anthraquinone core formation, rhodinose formation, and 2-thiosugar-containing disaccharide formation. Further isotopic tracer experiments using [1-^{13}C] and [1,2-$^{13}C_2$]acetate revealed that the construction of the angucycline skeleton involves an unusual oxidative ring-opening and rearrangement. Moreover, feeding experiments using a P450 hydroxylase inhibitor provided good evidence that the C-2 hydroxylation on the angucycline core is likely catalyzed by BexK.

Overall, this chapter described the structural and genetic analyses of BE-7585A (**2-1**) and proposed its biosynthetic pathway. These results set the stage to study the predicted functions of identified genes in the later chapters. Proposed disaccharide formation and C−S bond formation are described in Chapter 3. Studies of the molecular mechanism of the sulfur incorporation reaction are described in Chapter 4. Although the direct sulfur donor was not identified in the *bex* cluster, an endogenous sulfur carrier protein, such as ThiS used in thiamin biosynthesis, may be recruited for the thiosugar production, and this is also described in Chapter 4.

Chapter 3: Biosynthetic Studies of BE-7585A (II): Investigation of the Glycosyltransferase, BexG2, and C−S Bond Formation

3.1 INTRODUCTION

A putative biosynthetic gene cluster (*bex* cluster) for the 2-thiosugar-containing natural product, BE-7585A (**3-1**), was identified in Chapter 2. A biosynthetic pathway of the unique disaccharide moiety of **3-1** was proposed involving reactions catalyzed by a putative 2-thioglucose 6-phosphate synthase (*bexX*) and a putative 2-thiotrehalose 6-phoshate syntahse (*bexG2*) found in the gene cluster. The expected product, 2-thiotrehalose 6-phosphate (**3-4**), is then transferred to the aglycone (**3-2** → **3-3** → **3-4** → **3-5**, Figure 3-1, solid line). The last coupling step might be nonenzymatic since besides *bexG1* the only other glycosyltransferase gene found in the *bex* cluster is a putative rhodinosyl transferase gene (*bexG1*), which is likely responsible for transferring the rhodinosyl moiety to the aglycone at the C-12b position.

As discussed in the previous chapter, BexG2 exhibits moderate sequence similarity to trehalose 6-phosphate synthase (TPS, 30% identity and 47% similarity to TPS from *Methanothermobacter thermautotrophicus* str. Delta H). Thus, one may consider that BexG2 acts as a "regular" TPS to produce trehalose 6-phosphate (**3-6**) in the BE-7585A biosynthetic pathway. In this case, a sulfur atom must be introduced at the C-2 position of the resulting disaccharide (**3-6**) to construct the 2-thiotrehalose moiety found in BE-7585A (**3-1**). However, the proposed sulfur incorporation mechanism using BexX, which is analogous to ThiG-catalyzed thiazole formation, cannot account for the latter transformation because the reaction mechanism of BexX requires a reducing sugar substrate to form a schiff base with an active site lysine residue of the enzyme (Figure 2-23). Therefore, this second pathway (**3-2** → **3-6** → **3-4** → **3-5**) is less likely unless the

introduction of the sulfur atom into **3-6** can be catalyzed by BexX in a novel mechanism different from what has been proposed.

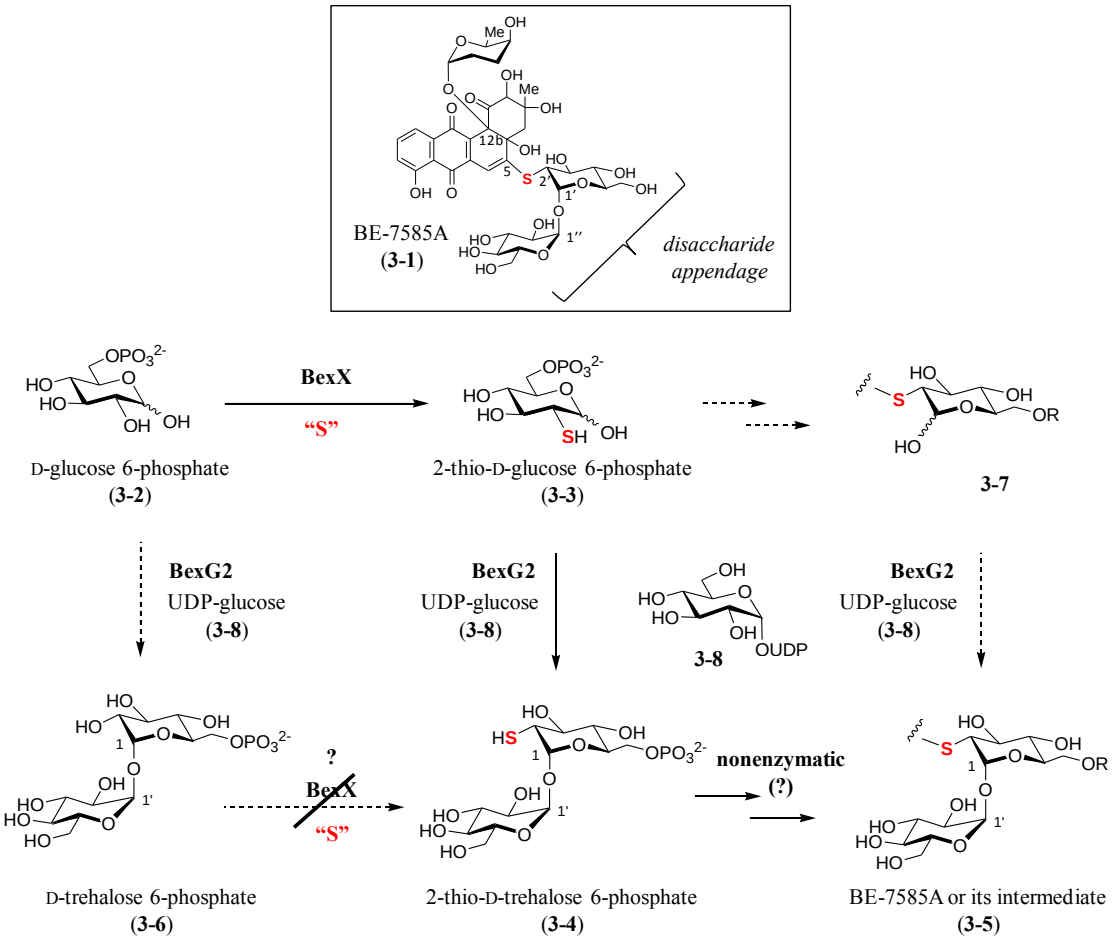


Figure 3-1. Possible biosynthetic pathways of the thiosugar-containing disaccharide moiety of BE-7585A and its attachment to the aglycone.

On the other hand, C−S bond formation between the generated thiosugar and C-5 of the angucycline core may not be enzyme-catalyzed. In the previous chapter, a possible nonenzymatic attack by the nucleophilic 2-thiol group of the disaccharide at the C-5

position of benz[*a*]anthraquinone was discussed (**3-4** → **3-5**). This hypothesis is based on the fact that the C-5 site of the core structure is electrophilic and, thus, can function as a Michael acceptor. In fact, a similar pathway has been proposed in urdamycin E (**3-10**) biosynthesis (Figure 2-2).[148] Urdamycin E (**3-10**), which is a minor product from *Streptomyces fradiae* (strain Tü 2712), possesses a methylthio substituent at C-5 on the angucycline core of urdamycin A (**3-9**). It was reported that the feeding experiment using 200 mg/L L-methionine (**3-11**) to the fermentation medium showed a clear increase of the yield of **3-10** from 3% to 30%. It was also confirmed that urdamycin A (**3-9**) can readily react with sodium methanethiolate (**3-12**) to yield urdamycin E (**3-10**) in a dimethyl sulphoxide (DMSO) [or tetrahydrofuran (THF)]−phosphate buffer (1 : 1) at pH 7.2 (Figure 3-2).[148]

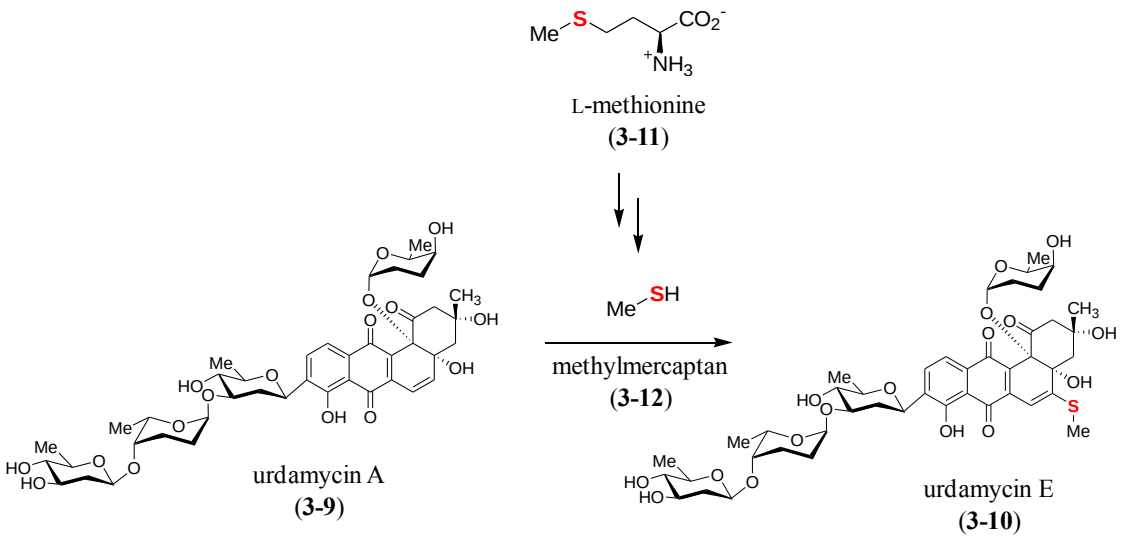


Figure 3-2. Proposed biosynthetic pathway of urdamycin E involving a nonenzymatic C–S bond formation.

It is notable that though the observed thiolate addition to **3-9** occurred spontaneously at a neutral pH to produce urdamycin E (**3-10**), this C−S bond formation might be controlled by some enzyme(s) *in vivo* since **3-10** is the only product possessing a thioether substituent at C-5 of the angucyline core isolated from *S. fradiae,* and no adducts derived from other thiol-containing biomolecules in the cellular pool were reported. If this process is truly nonenzymatic in the biosyntehsis of BE-7585A (**3-1**), both 2-thioglucose 6-phosphate (**3-3**) and 2-thiotrehalose 6-phosphate (**3-4**) may attack the C-5 position of the benz[*a*]anthraquinone core. Although a monosaccharide thioadduct on the angucycline core (**3-7**) has not been isolated from *A. orientalis,* it is possible that formation of the disaccharide moiety catalyzed by BexG2 may take place after the C−S bond is formed (Figure 3-1, **3-3** → **3-7** → **3-5**).

In this chapter, functional characterization of a TPS homologue, BexG2, is described. To test the proposed enzymatic reaction *in vitro*, the recombinant enzyme was heterologously expressed in *E. coli,* and the predicted substrate, 2-thioglucose 6-phosphate (**3-3**), was chemoenzymatically synthesized. The substrate specificity of the BexG2-catalyzed reaction was investigated and several possible biosynthetic pathways as well as the timing of coupling to produce the disaccharide moiety *in vivo* were also sudied. Finally, the proposed C−S bond formation is tested with a model compound, urdamycin A (**3-8**), and a synthetic 2-thiosugar.

3.2 EXPERIMENTAL PROCEDURES

3.2.1 General

Materials

All chemicals and reagents were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA), and were used without further purification unless otherwise specified. Most materials used for molecular cloning were identical to those described in Chapter 2. In addition, PfuUltra DNA polymerase was purchased from Stratagene (La Jolla, CA). Ni-NTA agarose was obtained from Qiagen (Valencia, CA). Amicon YM-10 ultrafiltration membranes were purchased from Millipore (Billerica, MA). Reagents for sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) were purchased from Bio-Rad (Hercules, CA), with the exception of the protein molecular weight markers, which were obtained from Invitrogen or New England Biolabs. The CarboPac PA1 high-performance liquid chromatography (HPLC) column was acquired from Dionex (Sunnyvale, CA). Analytical C_{18} HPLC columns were products of Varian (Palo Alto, CA).

Bacterial Strains and Plasmids

Cosmid C006 was obtained from the *A. orientalis* genomic library as described in Chapter 2. *E. coli* DH5α, acquired from Bethesda Research Laboratories (Gaithersburg, MD), was used for routine cloning experiments. The protein overexpression host *E. coli* BL21 star (DE3) was obtained from Invitrogen (Carlsbad, CA). Vector pET28b(+) for protein over-expression was purchased from Novagen (Madison, WI). Standard genetic manipulations of *E. coli* were performed as described by Sambrook et al.[107]

Instrumentation

UV−vis spectra were recorded using a Beckman DU 650 spectrophotometer. NMR spectra were acquired on a Varian Unity 500 MHz spectrometer, and chemical shifts (in ppm) are reported relative to that of the solvent peak (δ_H = 4.65 for deuterated water, δ_H = 3.30 and δ_C = 49.0 for deuterated methanol) or an internal standard (dioxane, δ_C = 67.4) for spectra taken in D_2O solvent. HPLC was performed on a Beckman Coulter System Gold equipped with a UV detector or Corona CAD (charged aerosol detector). DNA concentrations were measured using a NanoDrop ND-1000 UV−vis instrument from Thermo Fisher Scientific. DNA sequencing was performed by the core facility of the Institute of Cellular and Molecular Biology at the University of Texas, Austin. Mass spectroscopy was performed at the Mass Spectrometry core facility in the Department of Chemistry and Biochemistry at the University of Texas, Austin.

3.2.2 Cloning of *BexG2*.

The *bexG2* gene was PCR-amplified from cosmid C006 using primers with engineered *Nde*I and *Hind*III restriction sites. The sequences of the primers are shown in Table 3-1. The PCR-amplified gene fragments were purified, digested with *Nde*I and *Hind*III, and ligated into pET28b(+) vector digested with the same enzymes. The resulting plasmid, *bexG2*/pET28b(+), was used to transform *E. coli* BL21 star (DE3) strain for protein overexpression. The BexG2 enzyme was expressed as an *N*-terminal-His_6-tagged protein.

Table 3-1. Primers used for constructing *bexG2*/pET28b(+)

Primer name	Sequence
bexG2 for pET28b(+)-forward	5'-GGTTAGG**CATATG**AGGATTCTGACCTGCAGC-3'
bexG2 for pET28b(+)-reverse	5'-GTAG**AAGCTT**AGG*TCA*GAGCCCGAAAACCCC-3'

The engineered restriction sites are shown in bold, the start codon is shown in bold and also underlined, and the stop codon is italic.

3.2.3 Expression and Purification of BexG2.

An overnight culture of *E. coli* BL21 star (DE3)-*bexG2*/pET28b(+), grown in the LB medium (10 mL) containing 50 μg/mL of kanamycin at 37 °C, was used to inoculate 1L of the same growth medium. The culture was incubated at 37 °C with shaking (230 rpm) until the OD_{600} reached ~0.5. Protein expression was then induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM, and the cells were allowed to grow at 18 °C and 125 rpm for an additional 24 h. The cells were harvested by centrifugation at 4500 × *g* for 15 min and stored at −80 °C until lysis. All purification steps were carried out at 4 °C using Ni-NTA resin according to the manufacturer's protocol with minor modifications. Specifically, the thawed cells (~3 g) were resuspended in the lysis buffer (15 mL) containing 10% (v/v) glycerol and 10 mM imidazole. After incubation with lysozyme (15 mg) for 30 min, the cells were disrupted by sonication using 10 × 10-s pulses with a 30-s cooling pause between each pulse. The resulting lysate was centrifuged at 10,000 × *g* for 20 min, and the supernatant was subjected to Ni-NTA chromatography. Bound protein was eluted using buffer containing 10% glycerol and 250 mM imidazole. The collected protein solution was dialyzed against 3 × 1-L of 50 mM Tris·HCl buffer (pH 8.0) containing 300 mM NaCl and 15% glycerol. The protein solution was then flash-frozen in liquid nitrogen and stored at −80 °C until

use. Protein concentration was determined by the Bradford assay[149] using bovine serum albumin as the standard. The yield of BexG2 was 40 mg from 1 L culture. The molecular mass and purity of BexG2 were estimated by SDS-PAGE analysis.

3.2.4 Synthesis of 2-Thio-D-glucose.

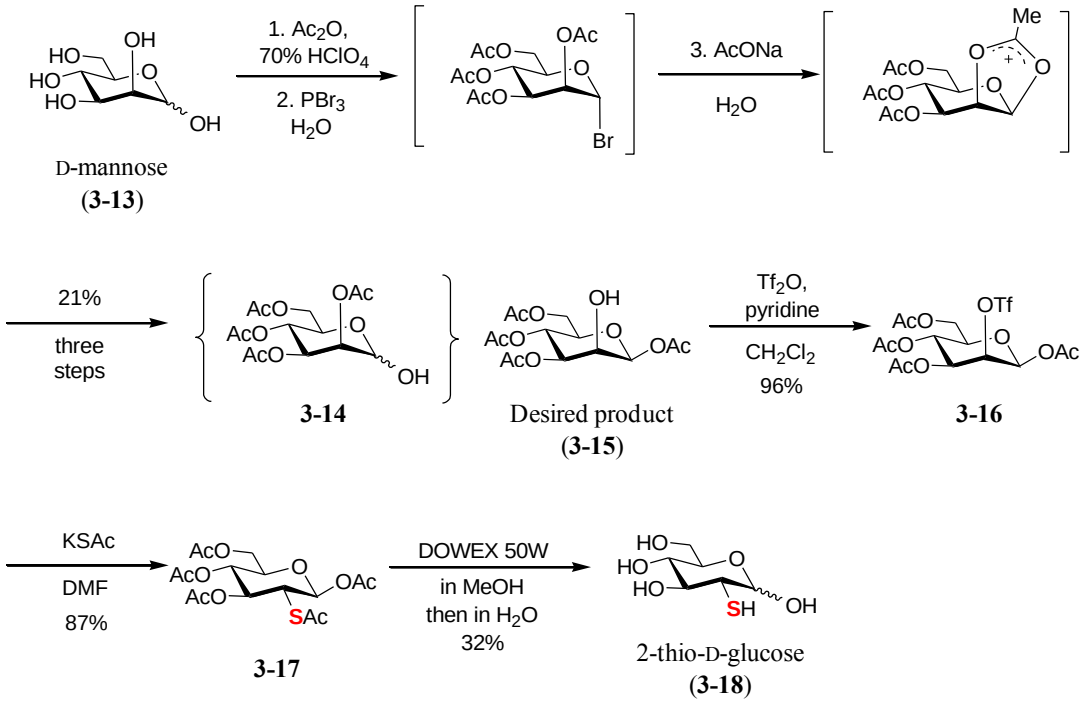


Figure 3-3. Synthetic scheme of 2-thio-D-gluose.

1,3,4,6-Tetra-O-Acetyl-β-D-Mannopyranose (3-15)

As shown in Figure 3-3, compound **3-15** was synthesized from D-mannose (**3-13**, 16.5 g, 92 mmol) following a previously reported procedure with slight modifications.[150,151] Specifically, to a small portion of **3-13** (~30 mg) in acetic anhydride (63 mL) was added 70% perchloric acid (5 drops). The remaining **3-13** was slowly added to the reaction mixture over a 20 min period, keeping the reaction temperature at

40−45 °C. The resulting mixture was stirred at room temperature for 1 h and then cooled to 15 °C. Phosphorus tribromide (13 mL) was added dropwise to the reaction with the reaction temperature kept at 20−25 °C. Cold water (8.3 mL) was slowly added to the resulting mixture, and the reaction was kept at room temperature for 1.5 h. After cooling to 10 °C, a chilled solution of sodium acetate trihydrate (50 g, 370 mmol) in water was slowly added with the temperature kept at 25−30 °C. The reaction mixture was stirred at room temperature for 20 min and poured onto ice. The resulting mixture was extracted with methylene chloride (100 mL × 2) and washed sequentially with cold saturated sodium bicarbonate and cold water. The solution was dried over anhydrous Na_2SO_4 and evaporated *in vacuo*. The obtained crude mixture, which contained 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranose (**3-14**) as a byproduct, was thoroughly washed with anhydrous diethyl ether to give the desired product **3-15** as a white powder (6.7 g, 19 mmol, 21%). The ^{1}H NMR spectral data of **3-15** are consistent with those previously reported.[151]

1,3,4,6-Tetra-O-Acetyl-2-O-Trifluoromethylsulfonyl-β-D-Mannopyranose (3-16)

Compound **3-16** was synthesized from **3-15** using a previously reported procedure with minor modifications.[152,153] Specifically, to a solution of **3-15** (3.5 g, 10 mmol) in anhydrous methylene chloride (100 mL) was added anhydrous pyridine (1.8 mL, 22 mmol). The solution was cooled to −20 °C and treated with trifluoromethane-sulfonic anhydride (3.6 mL, 22 mmol), which was added dropwise in 1 h under N_2. The resulting suspension was stirred at room temperature for an additional 1 h. The mixture was washed sequentially with cold water, saturated sodium bicarbonate, and water; dried over anhydrous Na_2SO_4; and concentrated *in vacuo*. The product **3-16** was obtained as a pale yellow solid (4.6 g, 9.6 mmol, 96%). The ^{1}H NMR spectral data of **3-16** match those reported previously.[153]

1,2,3,4,6-Penta-O,S,O,O,O-Acetyl-2-Thio-β-D-Glucopyranose (3-17)

Compound **3-17** was prepared from **3-16** following a reported procedure with minor modifications.[154,155] Accordingly, to a solution of **3-16** (4.1 g, 8.5 mmol) in anhydrous dimethylformamide (170 mL) was added potassium thioacetate (9.7 g, 85 mmol). The resulting mixture was stirred at room temperature for 1 h under N_2. After evaporation of the solvent *in vacuo*, the residue was mixed with methylene chloride and water. The organic layer was separated and washed sequentially with water and saturated sodium chloride, dried over anhydrous Na_2SO_4, and concentrated *in vacuo*. The residue was purified by silica gel column chromatography (hexanes−EtOAc, 2:1) and recrystallized from hexanes−EtOAc. The product **3-17** was obtained as a white powder (3.0 g, 7.4 mmol, 87%). The ^{1}H NMR spectrum of **3-17** is consistent with the literature report.[154]

2-Thio-D-Glucose (3-18)

To a solution of **3-17** (510 mg, 1.25 mmol) in anhydrous methanol (15 mL) was added Dowex 50W X8 (~200 mg). The reaction was gently refluxed for 15 h. The resulting solution was filtered and the filtrate was concentrated under reduced pressure to give an α/β mixture of 1-*O*-methyl-2-thio-D-glucopyranose (α/β) 1:3, determined by ^{1}H NMR spectroscopy). The mixture was dissolved in water (20 mL) and stirred at 80 °C in the presence of Dowex 50W X8 for 20 h under N_2. The resulting solution was filtered and lyophilized (260 mg). A part of the residue (100 mg) was used for further purification by silica gel column chromatography (water−acetonitrile, 1:50). The product **3-18** was obtained as a white powder (30 mg, 0.15 mmol, 32%). In solution, **3-18** exists as a mixture of α and β isomers (α/β) 9:10, under the ^{1}H NMR conditions). ^{1}H NMR (500 MHz, D_2O) δ 5.13 (d, J = 3.3 Hz, 1H, 1_α-H), 4.55 (d, J = 8.6 Hz, 1H, 1_β-H), 3.75 (dd, J = 12.3, 2.3 Hz, 1H, $6_{\beta1}$-H), 3.74 (m, 1H, 5α-H), 3.69 (dd, J = 12.3, 2.4 Hz, 1H, $6_{\alpha1}$-H), 3.62

(dd, J = 12.3, 5.1 Hz, 1H, $6_{\alpha 2}$-H), 3.58 (dd, J = 12.3, 5.9 Hz, 1H, $6_{\beta 2}$-H), 3.48 (dd, J = 10.7, 9.0 Hz, 1H, 3α-H), 3.34 (ddd, J = 9.7, 5.9, 2.3 Hz, 1H, $5_{\beta 2}$-H), 3.27 (m, 1H, 3_{β}-H), 3.26 (m, 1H, 4_{α}-H), 3.23 (m, 1H, 4_{β}-H), 2.75 (dd, J = 10.7, 3.3 Hz, 1H, 2_{α}-H), 2.57 (dd, J = 10.2, 8.6 Hz, 1H, 2_{β}-H); ^{13}C NMR (125 MHz, D_2O) δ 98.2 (C-1_{β}), 94.5 (C-1_{α}), 77.2 (C-3_{β}), 76.7 (C-5_{β}), 74.6 (C-3_{α}), 72.8 (C-5_{α}), 71.5 (C-4_{α}), 71.2 (C-4_{β}), 61.6 (C-6_{β}), 61.4 (C-6_{α}), 48.5 (C-2_{β}), 45.8 (C-2_{α}); HRMS (ESI+) calcd for $C_6H_{12}O_5NaS^+$ $[M + Na]^+$ 219.0298, found 219.0303. (ESI−) calcd for $C_6H_{11}O_5S^-$ $[M - H]^-$ 195.0333, found 195.0330. NMR spectra are shown in Figure 3-4−3-10.

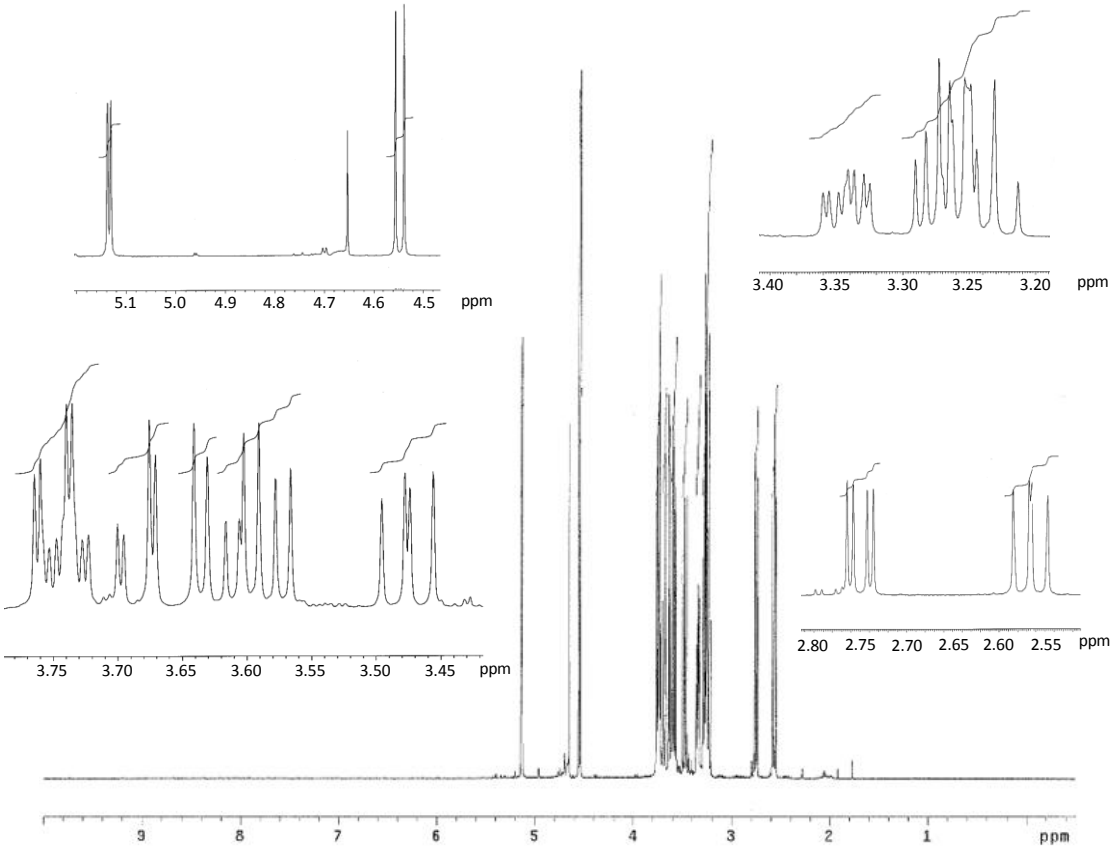


Figure 3-4. ^{1}H NMR spectrum (500 MHz, D_2O) of 2-thio-D-glucose.

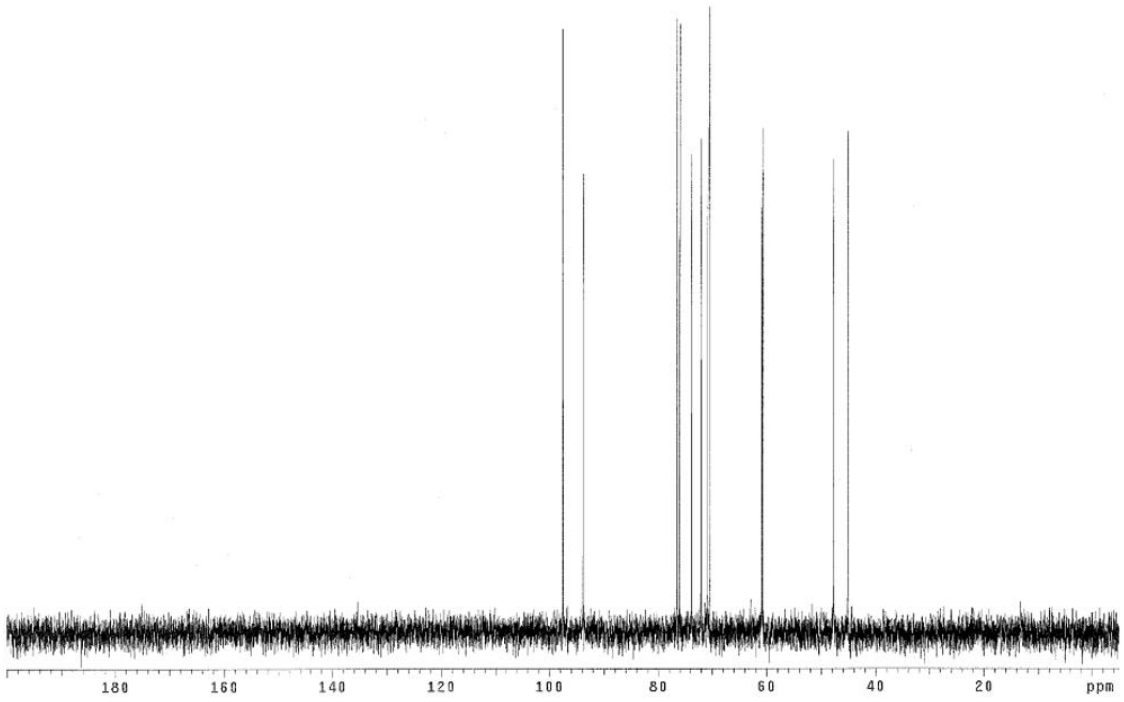


Figure 3-5. ^{13}C NMR spectrum (125 MHz, D_2O) of 2-thio-D-glucose.

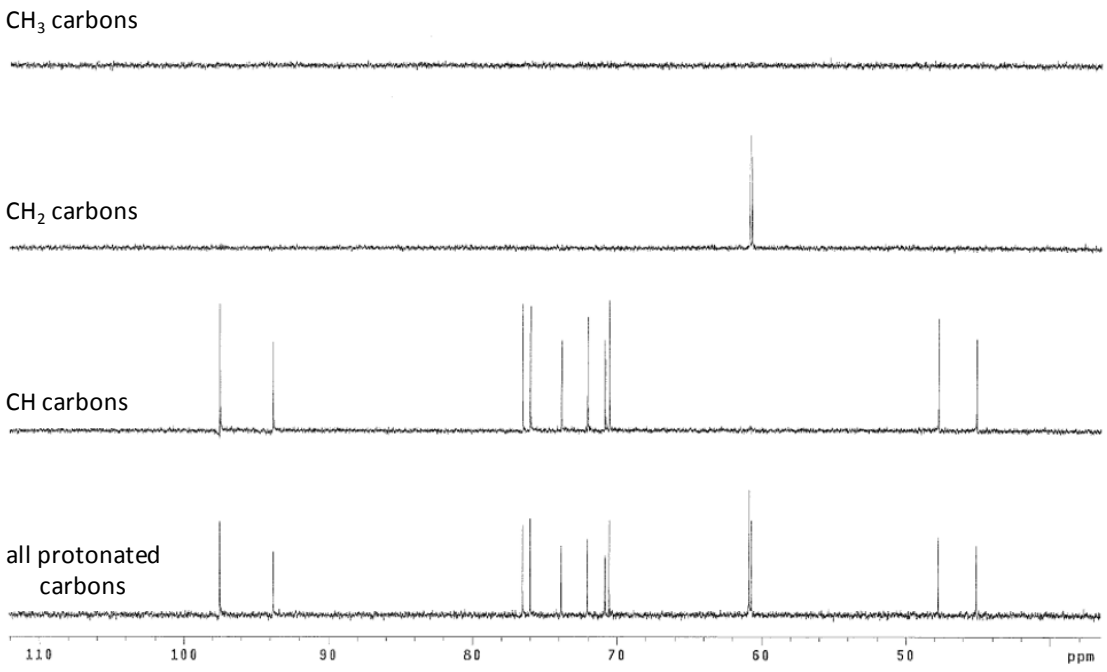


Figure 3-6. DEPT spectrum of 2-thio-D-glucose.

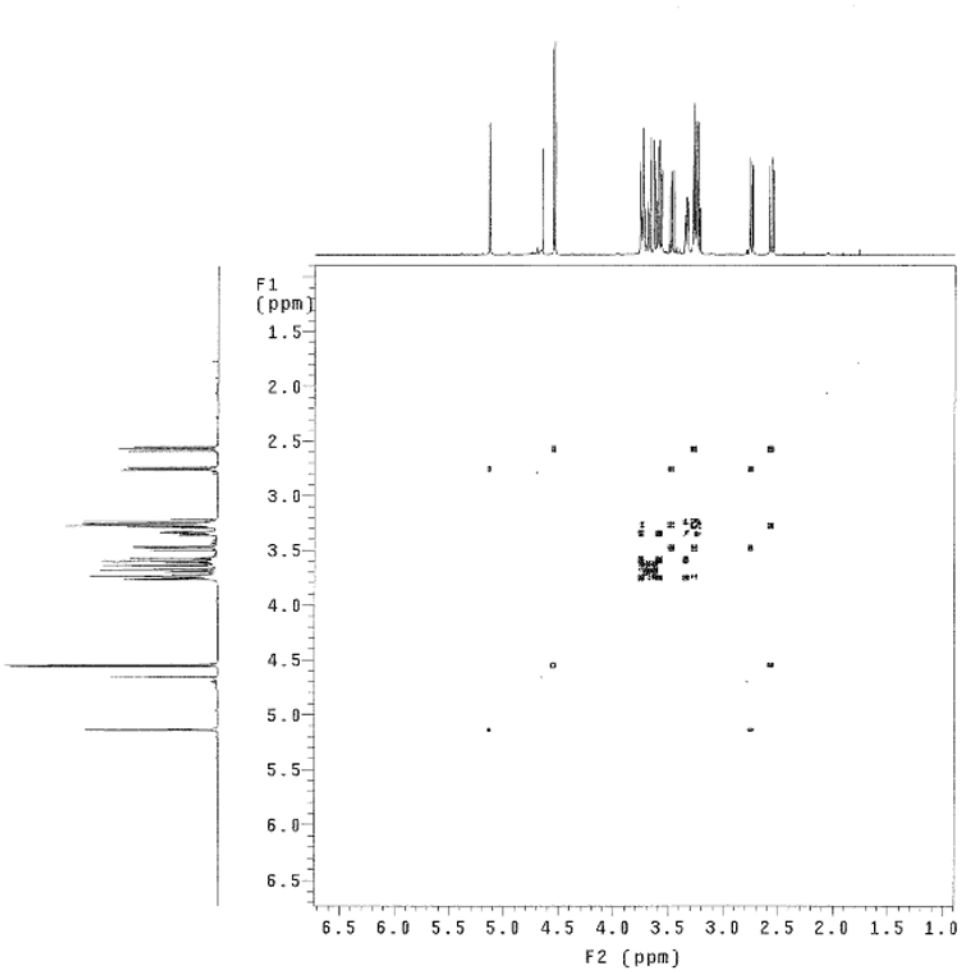


Figure 3-7. ^{1}H-^{1}H COSY spectrum of 2-thio-D-glucose.

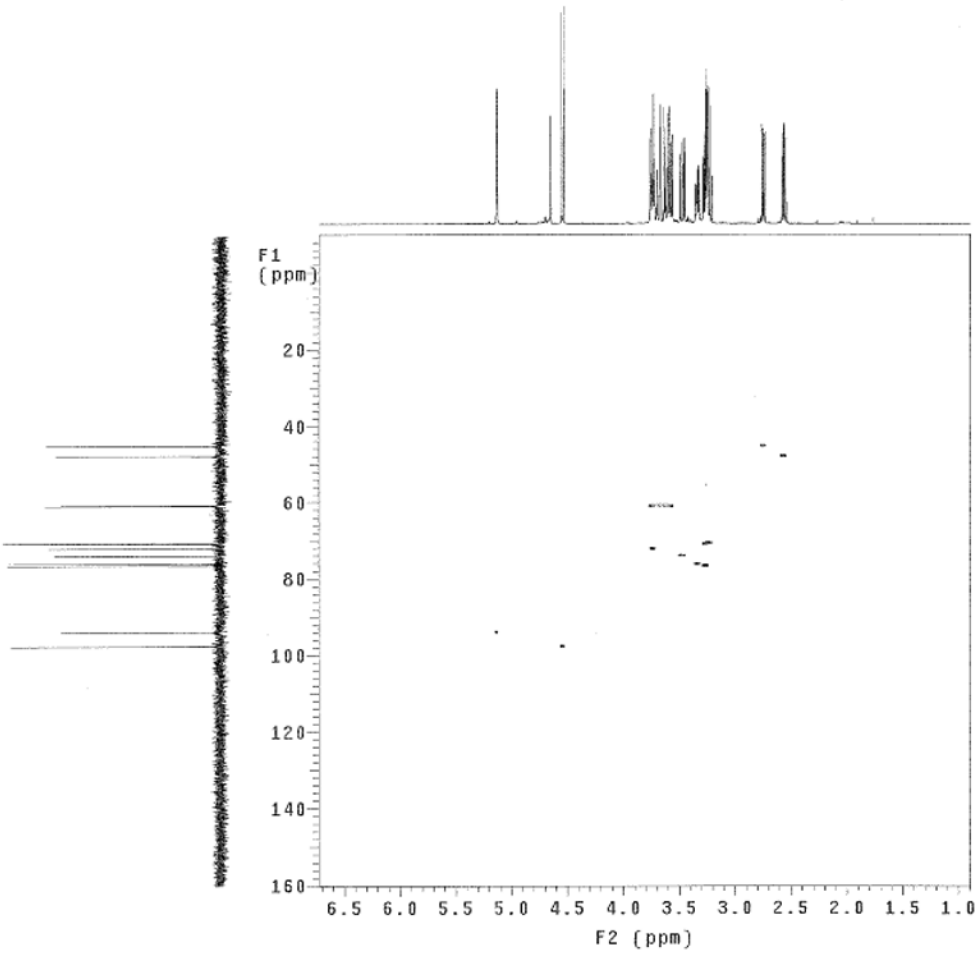


Figure 3-8. HSQC spectrum of 2-thio-D-glucose.

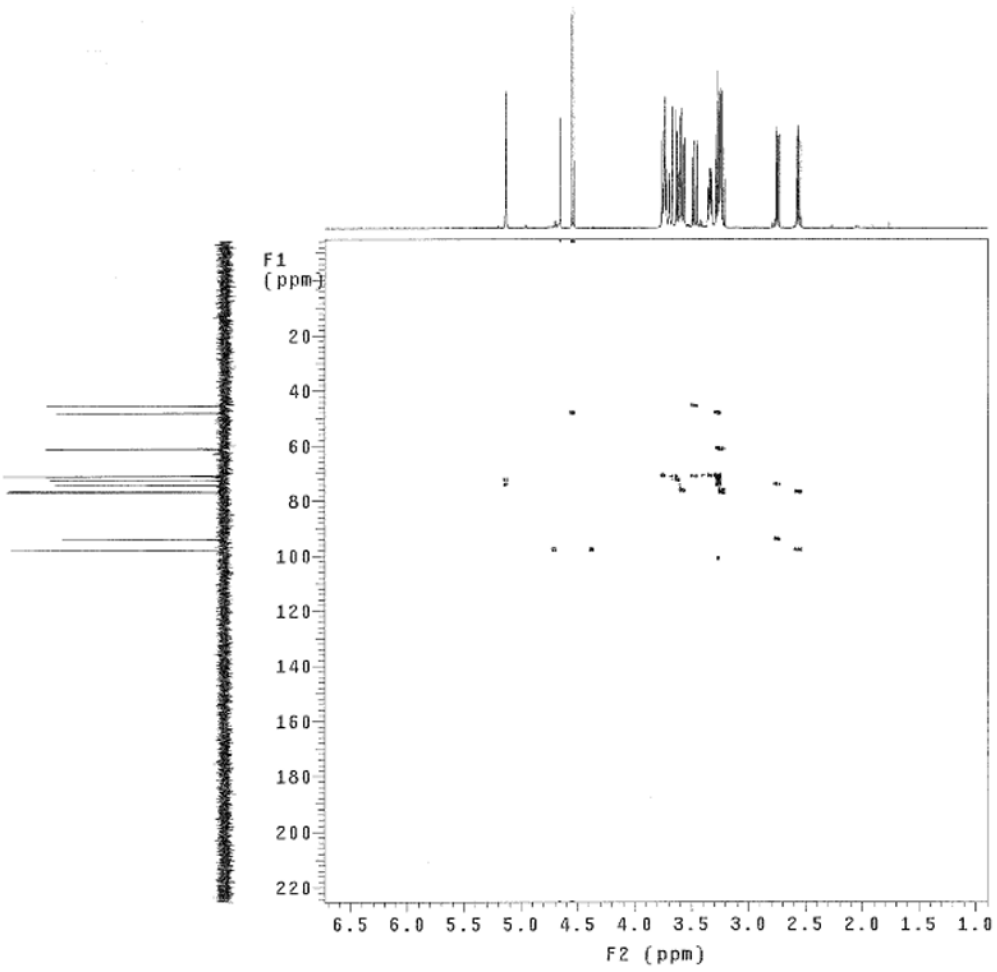


Figure 3-9. HMBC spectrum of 2-thio-D-glucose.

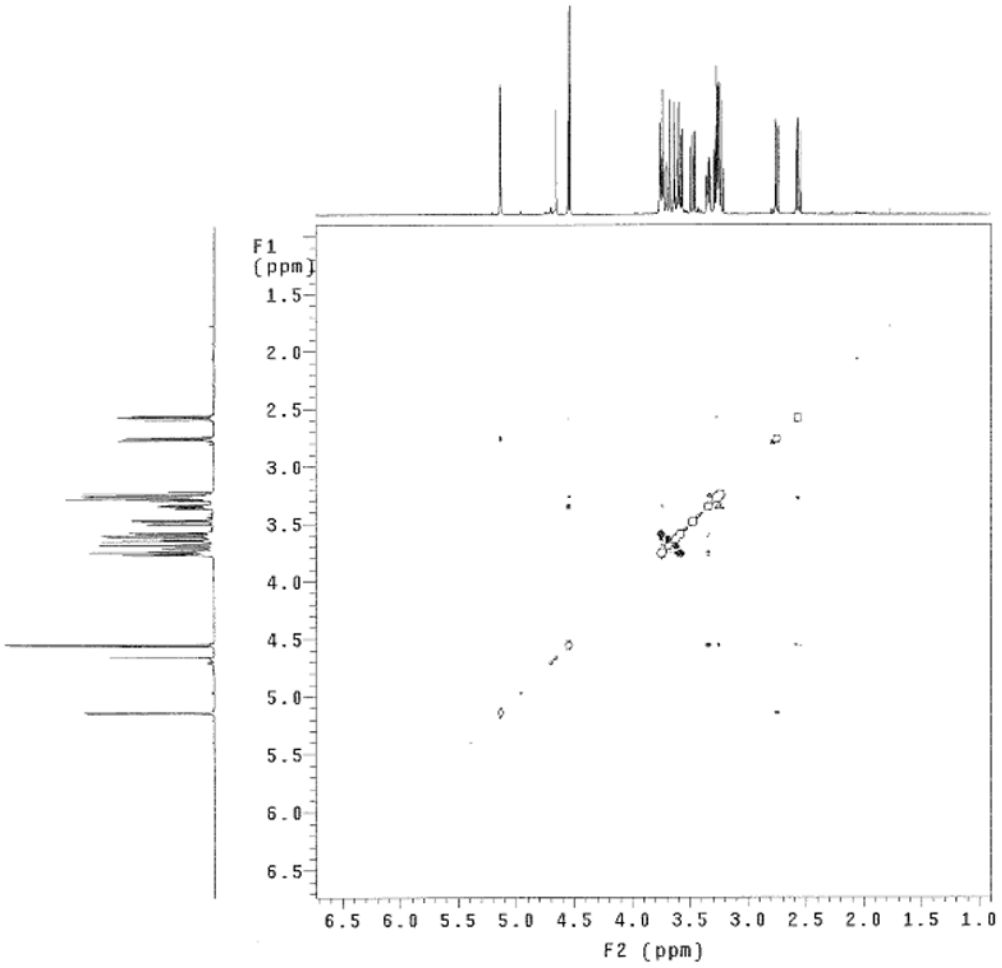


Figure 3-10. NOESY spectrum (500 MHz, D_2O) of 2-thio-D-glucose.

3.2.5 BexG2 *in Vitro* Activity Assay.

As depicted in Figure 3-11, a typical BexG2 assay mixture (0.1 mL) contained 2-thio-D-glucose (**3-18**, 1 mM), ATP (1 mM), UDP-D-glucose (**3-8**, 1 mM), $MgCl_2$ (10 mM), hexokinase (5 μg), and BexG2 (16 μM) in 50 mM Tris·HCl buffer, pH 8.0. The substrate for BexG2, 2-thio-D-glucose 6-phosphate (**3-3**), was produced *in situ* from **3-18** and ATP by hexokinase. The reaction mixture was incubated at 30 °C for 1.5 h and filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the

proteins. HPLC analysis was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm). The sample was eluted with a gradient of water (solvent A) and 1 M NH_4OAc (solvent B). The gradient was run from 30 to 50% B over 10 min, 50−80% B over 5 min, with a 5 min wash at 80% B, and 80−30% B over 5 min, followed by re-equilibration at 30% B for 10 min. The flow rate was 1 mL/min, and the detector was set at 276 nm.

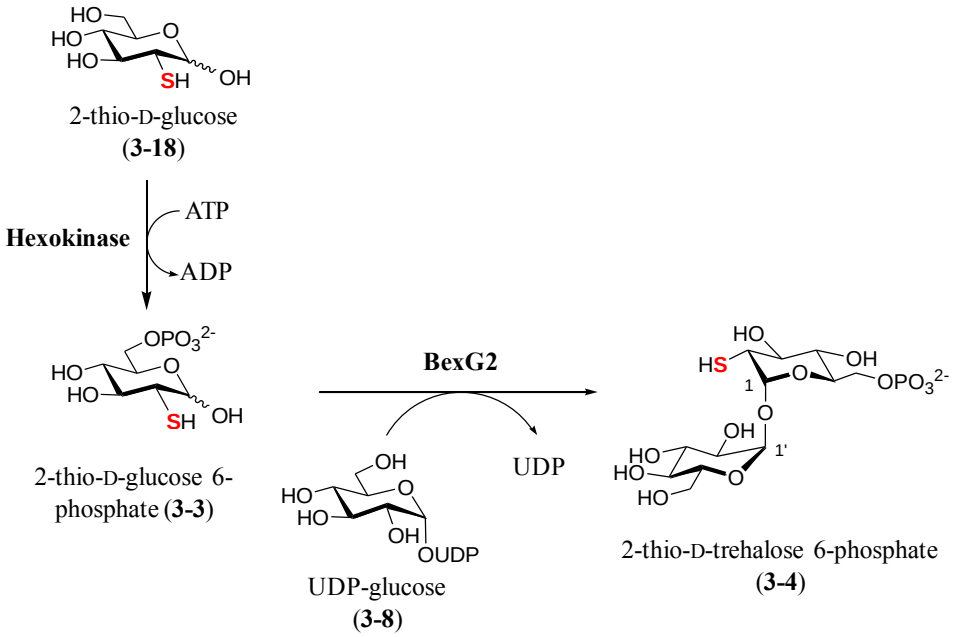


Figure 3-11. BexG2-catalyzed 2-thiotrehalose 6-phosphate formation.

3.2.6 Isolation and Characterization of BexG2 Product.

A large-scale BexG2 reaction (1 mL) containing 2-thio-D-glucose (**3-18**, 10 mM), ATP (12 mM), UDP-D-glucose (**3-8**, 12 mM), $MgCl_2$ (3 mM), hexokinase (20 μg), and BexG2 (30 μM) in 100 mM Tris·HCl buffer, pH 8.0, was carried out. The reaction mixture was incubated at 30 °C for 6.5 h and filtered through a YM-10 membrane using

an Amicon ultrafiltration unit to remove the proteins. HPLC analysis and purification were performed using a Dionex CarboPac PA1 semipreparative column (9 × 250 mm). The sample (330 μL × 3) was eluted with 0.1 M NH_4OAc at 5 mL/min. The resulting flow through was split into a collection fraction (4.85 mL/min) and a detection fraction (0.15 mL/min) by a splitter. The elution was monitored by a Corona CAD. The retention time of **3-4** was 10.4 min under the HPLC conditions. The fraction containing **3-4** was collected and lyophilized. ^{1}H NMR (500 MHz, D_2O) δ 5.11 (d, *J* = 3.5 Hz, 1H, 1-H), 5.06 (d, *J* = 3.8 Hz, 1H, 1'-H), 3.96 (m, 2H, $6_{1,2}$-H), 3.93 (m, 1H, 5'-H), 3.85 (m, 1H, 5-H), 3.76 (dd, *J* = 12.5, 2.2 Hz, 1H, $6'_1$-H), 3.68 (m, 1H, 3'-H), 3.66 (m, 1H, 3-H), 3.65 (m, 1H, $6'_2$-H), 3.53 (dd, *J* = 9.9, 3.8 Hz, 1H, 2'-H), 3.44 (dd, *J* = 9.9, 9.2 Hz, 1H, 4-H), 3.34 (dd, *J* = 10.1, 9.2 Hz, 1H, 4'-H), 2.90 (dd, *J* = 10.8, 3.5 Hz, 1H, 2-H); ^{13}C NMR (125 MHz, D_2O) δ 94.8 (C-1), 93.4 (C-1'), 73.8 (C-3), 72.9 (C-5'), 72.7 (C-3'), 71.9 (C-5), 71.1 (C-2'), 70.4 (C-4), 69.8 (C-4'), 63.8 (C-6), 60.8 (C-6'), 44.5 (C-2); HRMS (ESI−) calcd for $C_{12}H_{22}O_{13}SP^-$ $[M - H]^-$ 437.0524, found 437.0520. NMR spectra are shown in Figure 3-12−3-18.

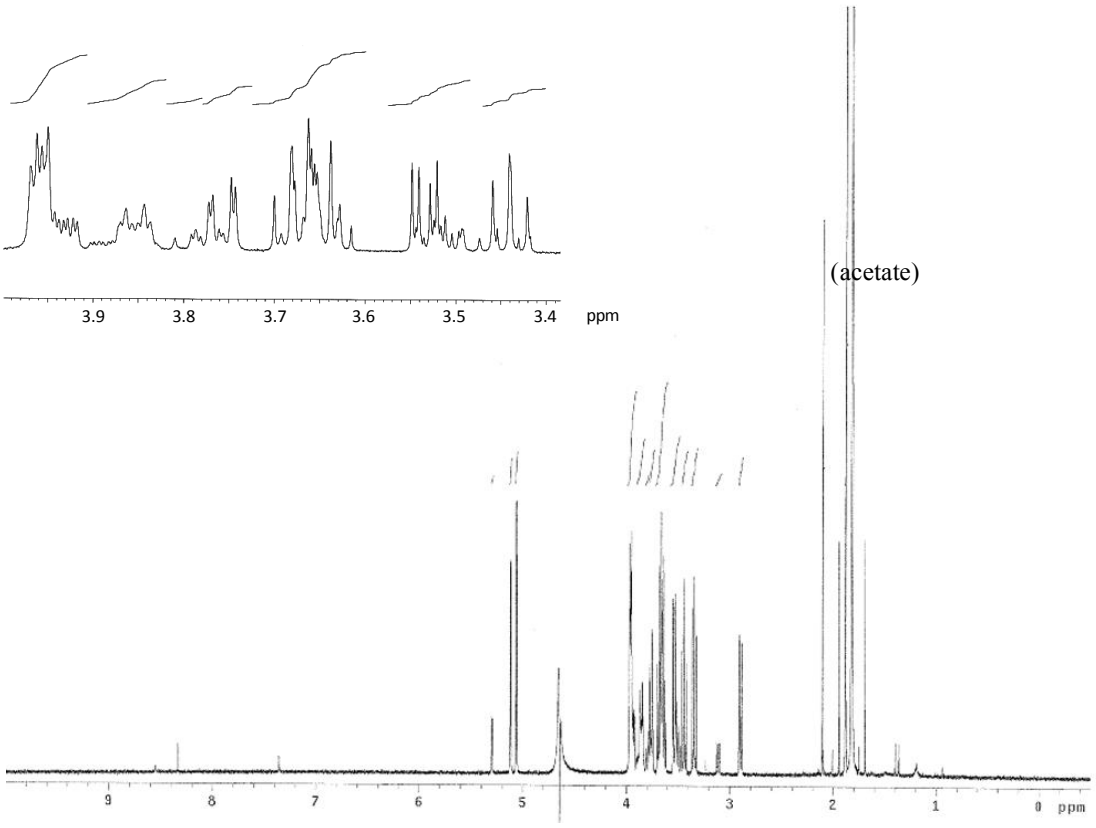


Figure 3-12. ^{1}H NMR spectrum (500 MHz, D2O) of 2-thio-D-trehalose 6-phosphate.

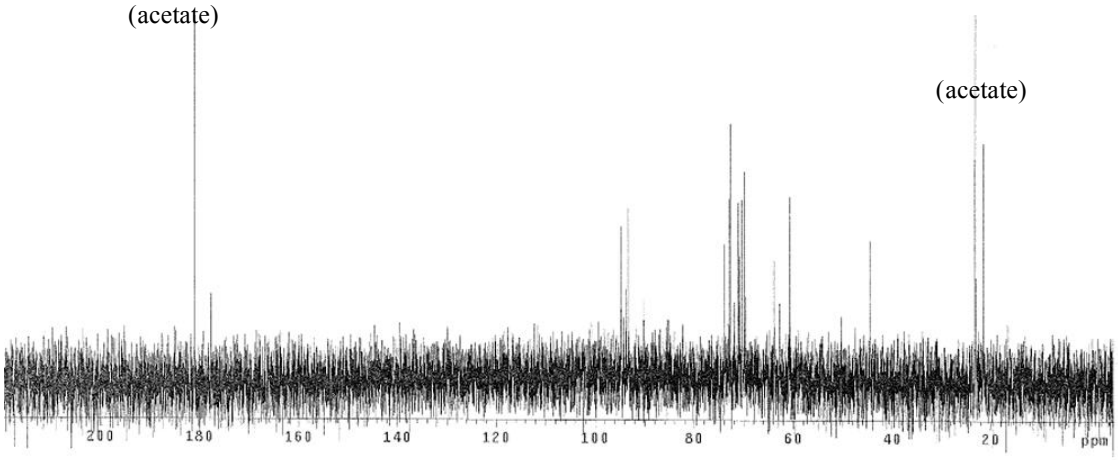


Figure 3-13. ^{13}C NMR spectrum (125 MHz, D_2O) of 2-thio-D-trehalose 6-phosphate.

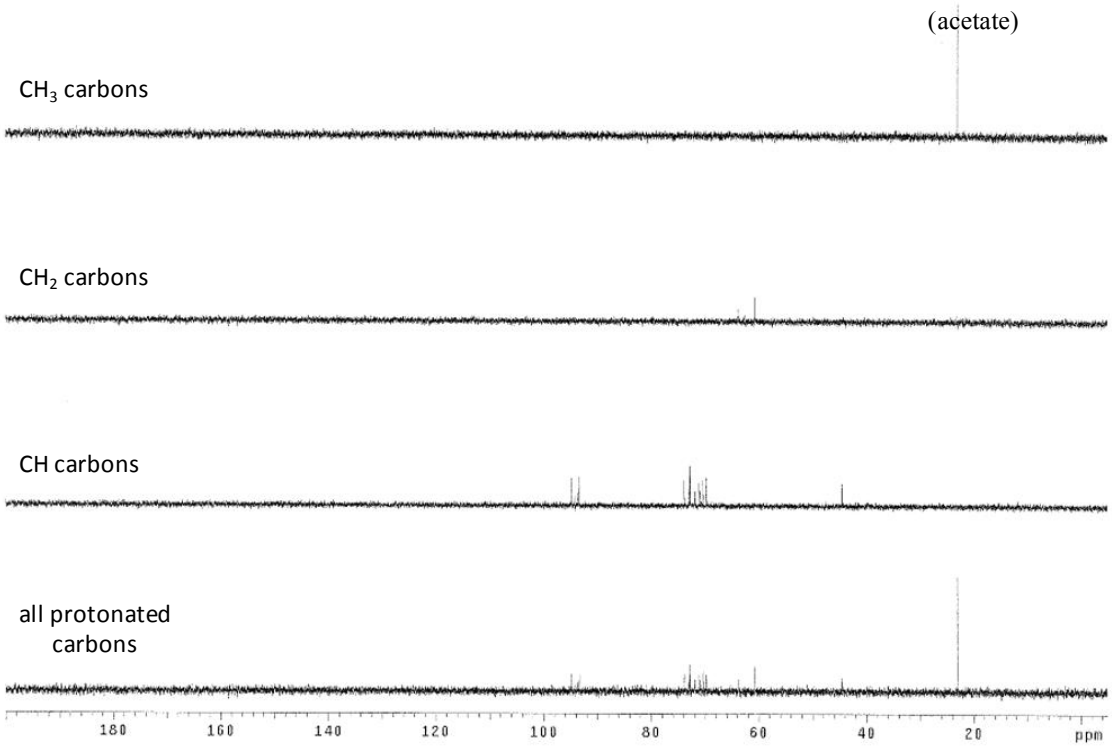


Figure 3-14. DEPT spectrum of 2-thio-D-trehalose 6-phosphate.

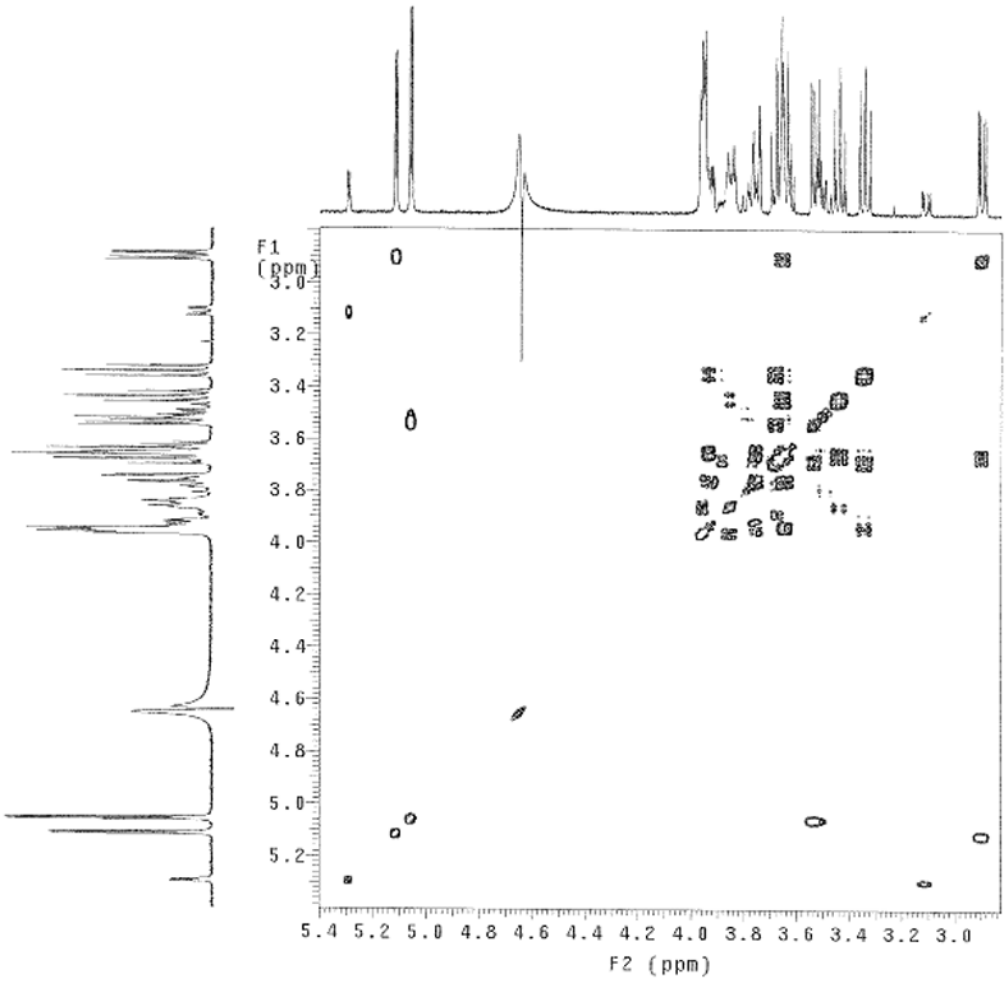


Figure 3-15. ^{1}H-^{1}H COSY spectrum of 2-thio-D-trehalose 6-phosphate.

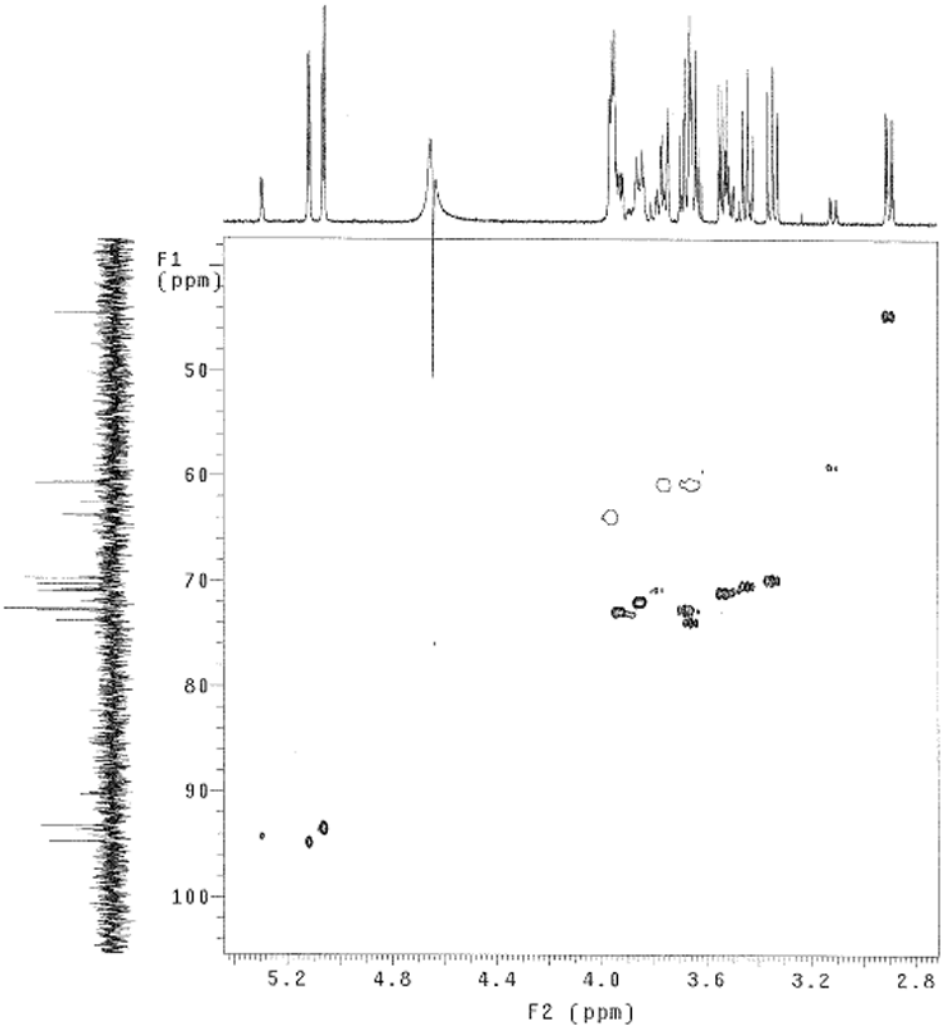


Figure 3-16. HSQC spectrum of 2-thio-D-trehalose 6-phosphate.

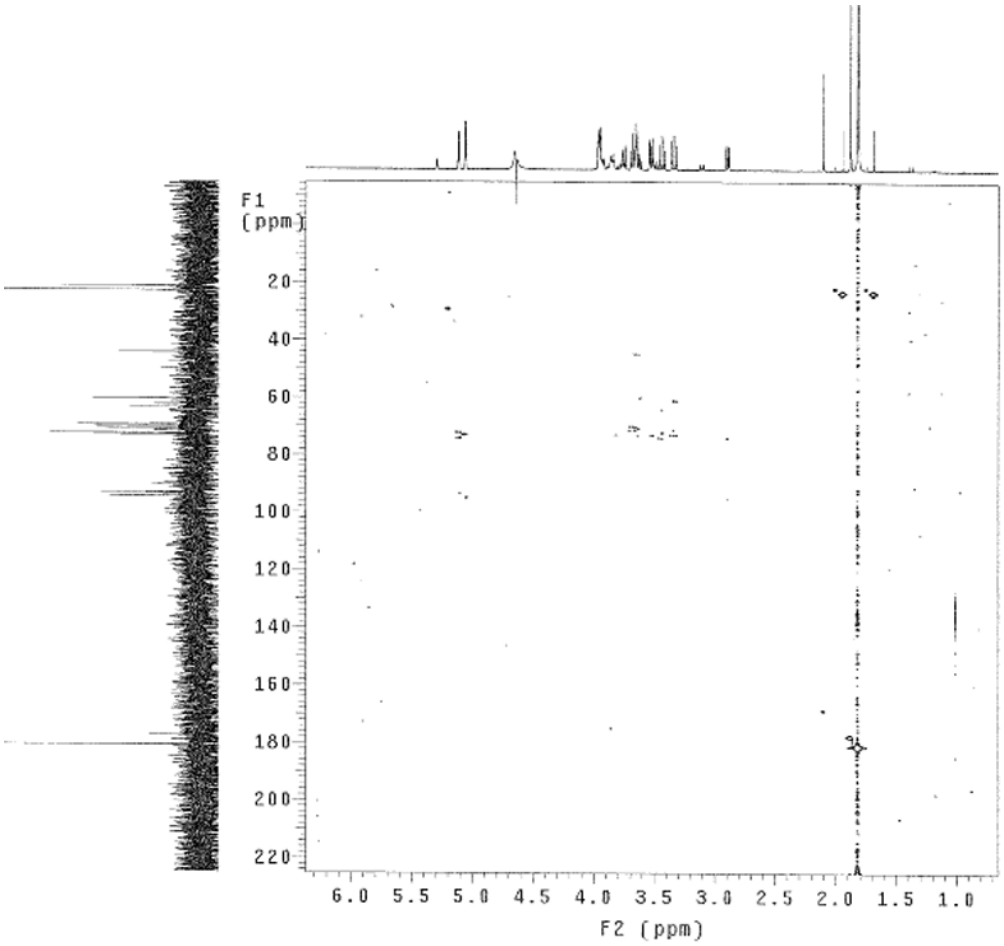


Figure 3-17. HMBC spectrum of 2-thio-D-trehalose 6-phosphate.

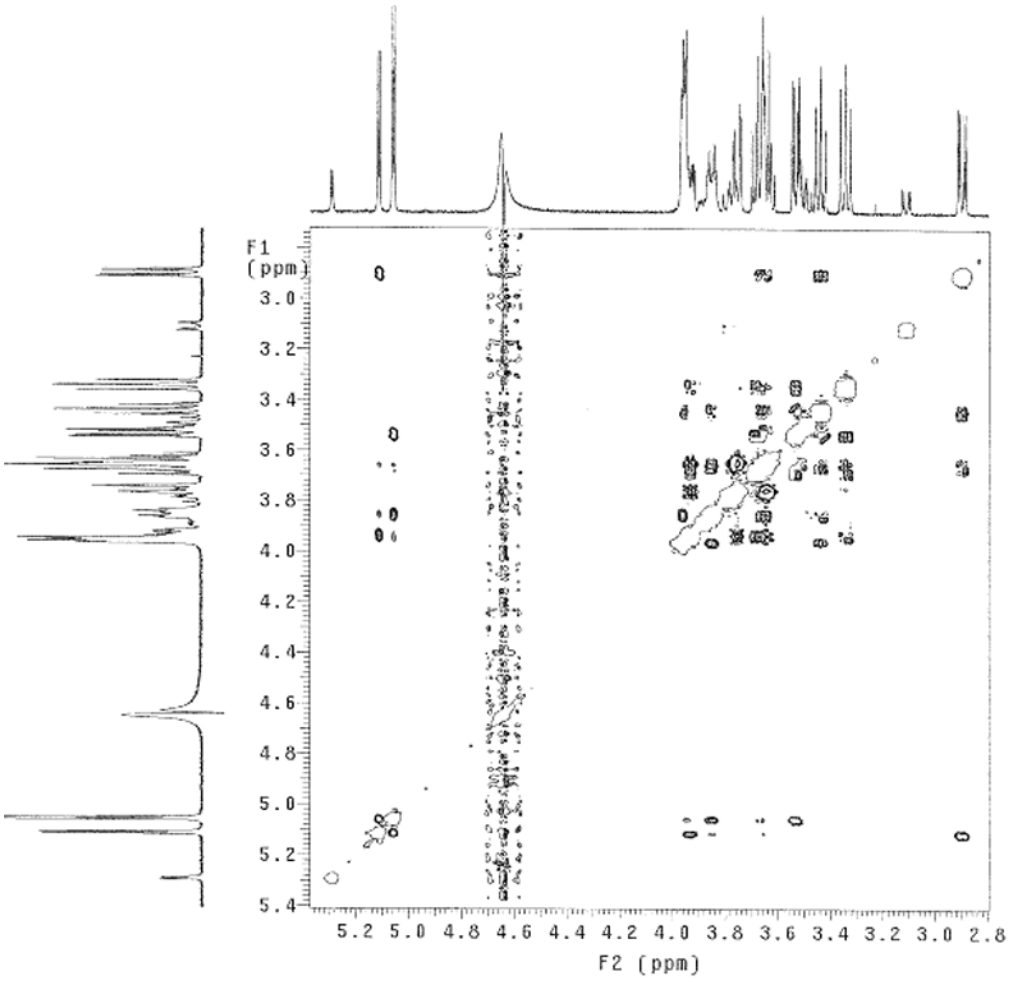


Figure 3-18. NOESY spectrum of 2-thio-D-trehalose 6-phosphate.

3.2.7 Substrate Specificity of BexG2-Catalyzed Glycosyltransfer Reaction.

D-Glucose 6-Phosphate as the Glycosyl Acceptor

The reaction solution (0.5 mL × 3) contained D-glucose 6-phosphate (**3-2**, 10 mM), UDP-D-glucose (**3-8**, 10 mM), $MgCl_2$ (20 mM), and BexG2 (29 μM) in 50 mM $NH_4{\cdot}HCO_3$ buffer, pH 7.9. The protein stock solution of BexG2 was dialyzed against $NH_4{\cdot}HCO_3$ buffer containing $MgCl_2$ (20 mM) prior to use to remove excess glycerol. The

reaction mixture was incubated at 30 ºC for 1, 2.5, or 5.8 h and quenched in boiling water for 5 min. Each reaction mixture was then centrifuged, lyophilized and redissolved in D_2O for the NMR analysis. The resulting solution was also subjected to MS analysis: ESI−MS (low resolution) calculated for D-trehalose 6-phosphate (**3-6**), $C_{12}H_{22}O_{14}P^-$ $[M - H]^-$ 421.1, found 421.3. As a separate experiment, a similar reaction solution (1 mL) containing D-glucose 6-phosphate (**3-2**, 10 mM), UDP-D-glucose (**3-8**, 10 mM), $MgCl_2$ (20 mM), and BexG2 (16 μM) in 50 mM $NH_4 \cdot HCO_3$ buffer, pH 7.9 was prepared. The reaction was incubated at 30 ºC for 5 h, and then 37 ºC for 1 h after addition of calf intestinal alkaline phosphatase (CIP) (1 μL, 10 units). After centrifugation to remove the precipitate, the solution was lyophilized, redissolved in D_2O, and subjected to NMR analysis. To confirm the identities of the anomeric protons, a small amount of D-glucose (**3-19**) standard and D-trehalose (**3-20**) standard were added and subjected to NMR analysis successively.

Investigation of Other Glycosyl Donors

Various NDP-sugars were substituted for UDP-glucose (**3-8**) to test the substrate specificity of BexG2 towards the glycosyl donors. NDP-sugars tested were UDP-glucose (**3-8**, positive control), UDP-xylose (**3-21**), UDP-glucuronic acid (**3-22**), UDP-galactose (**3-23**), TDP-glucose (**3-24**), ADP-glucose (**3-25**) and GDP-glucose (**3-26**). The BexG2 assay mixture (50 μL) contained D-glucose 6-phosphate (**3-6**, 5 mM or 1 mM for the reaction with TDP-glucose), one of the NDP-sugars (1 mM for TDP-glucose, 5 mM for the others), $MgCl_2$ (20 mM), and BexG2 (36 μM) in 100 mM Tris·HCl buffer, pH 8.0. The reaction mixture was incubated at 30 °C for 5 h and filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the proteins. HPLC analysis was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm). The

sample (5 μL or 10 μL for the reaction with TDP-glucose) was eluted with a gradient of water (solvent A) and 1 M NH_4OAc (solvent B). The gradient was run from 30 to 50% B over 10 min, 50−80% B over 5 min with a 5 min wash at 80% B, and 80−30% B over 5 min, followed by re-equilibration at 30% B for 10 min. The flow rate was 1 mL/min, and the detector was set at 260 nm.

Investigation of Other Glycosyl Acceptors

Various sugar-monophosphates were also substituted for the natural substrate, 2-thioglucose 6-phosphate (**3-3**), to test the substrate specificity of BexG2 for the glycosyl donors. Glycosyl donors tested are D-glucose 6-phosphate (**3-6**, positive control), 2-deoxy-D-glucose 6-phosphate (**3-27**), D-glucosamine 6-phosphate (**3-28**), and D-mannose 6-phoshate (**3-29**). The BexG2 assay mixture (50 μL) contained one of the glycosyl acceptors (5 mM), UDP-glucose (**3-8**), 5 mM), $MgCl_2$ (20 mM), and BexG2 (36 μM) in 100 mM Tris·HCl buffer, pH 8.0. The reaction mixture was incubated at 30 °C for 5 h and filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the proteins. HPLC analysis was similarly performed as described above.

3.2.8 Competitive Assay of BexG2-Catalyzed Glycosyltransfer Reaction.

The reaction solution (0.3 mL) containing 2-thio-D-glucose (**3-18**, 10 mM), D-glucose (**3-19**, 10 mM), ATP (24 mM), $MgCl_2$ (20 mM), and hexokinase (0.05 mg/mL) was incubated in 100 mM Tris·HCl (pH 8.0) at 30 °C for 1 h. Complete conversion of **3-18** and **3-19** to 2-thio-D-glucose 6-phosphate (**3-3**) and D-glucose 6-phosphate (**3-2**), respectively, was confirmed for similar reactions performed in a separate experiment. The resulting solution was filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the proteins. BexG2 and various concentration of UDP-glucose (**3-8**) were then added to an aliquot of the solution (30 μL). This reaction mixture

(60 μL) contained **3-3** (5 mM expected), **3-2** (5 mM expected), ATP (1 mM expected), ADP (10 mM expected), $MgCl_2$ (10 mM expected), BexX (26 μM), **3-8** (0, 2.5, 3.75, 5, 6.25, 7.5, 10, or 12.5 mM) in 50 mM Tris·HCl (pH 8.0). The reaction was incubated at 30 °C for 8 h and filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the proteins. HPLC analysis was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm). The sample (10 μL) was analyzed by two separate methods. The first method was designed for detecting the sugar monophosphates, in which the sample was eluted with 120 mM NH_4OAc at 1 mL/min and monitored by a Corona CAD. The second method was designed for detecting ATP, ADP, UDP, and UDP-glucose (**3-8**). The sample (5 μL) was eluted with a gradient of water (solvent A) and 1 M NH_4OAc (solvent B). The gradient was run from 30 to 50% B over 10 min, 50−80% B over 5 min with a 5 min wash at 80% B, and 80−30% B over 5 min, followed by re-equilibration at 30% B for 10 min. The flow rate was 1 mL/min, and the detector was set at 276 nm.

3.2.9 Non-Enzymatic C−S Bond Formation Between a Thiosugar and Aglycon.

HPLC and MS Analysis

A reaction solution (50 μL) containing urdamycin A (**3-9**, 0.5 mM) and 2-thio-D-glucose (**3-18**, 1 mM) in 50 mM Tris·HCl buffer, pH 8.0 was incubated at room temperature for 3 min and subjected to HPLC analysis. As a control, D-glucose (**3-19**) was used instead of **3-18**. HPLC analysis was performed using a C_{18} analytical column (4 × 250 mm). The sample (10 μL) was eluted with a gradient of water (solvent A) and methanol (solvent B). After 5 min at 20% B, the gradient was run from 20 to 80% B over 10 min with a 5 min wash at 80% B, 80−20% B over 5 min, followed by re-equilibration at 20% B for 10 min. The flow rate was 1 mL/min, and the detector was set at 300 nm.

The product peak (**3-30**) was isolated and subjected to MS analysis. HRMS (ESI−) calculated for $C_{49}H_{65}O_{22}S^-$ $[M - H]^-$ 1037.3694, found 1037.3689.

Absorbance Spectra Analysis

The absorbance spectrum (220−650 nm) of urdamycin A (**3-9**, 0.2 mM) was measured in 50 mM glycylglycine buffer (pH 8.0). To this solution, 1 μL of a 10 mM 2-thio-D-glucose stock solution (**3-18**, final concentration 0.1 mM) was added, and the absorbance spectrum was immediately measured. After a 2 min incubation at room temperature, the spectrum was measured again. Similarly, an additional 0.5 μL of 100 mM 2-thio-D-glucose stock solution (**3-18**, final total concentration 0.6 mM) was added, and the absorbance spectrum was immediately measured. After 2 min and 4 min incubations at room temperature, the spectra were measured repeatedly.

3.3 RESULTS AND DISCUSSION

3.3.1 BexG2 Activity Assay.

To verify the proposed pathway for disaccharide formation (**3-3** → **3-4**, Figure 3-1), the putative glycosyltransferase, BexG2, was expressed and purified to near homogeneity (Figure 3-19). The standard BexG2 assay contained 2-thioglucose (**3-18**), ATP, UDP-glucose (**3-8**), $MgCl_2$, hexokinase, and BexG2 in Tris·HCl buffer. In this assay, the predicted substrate, 2-thioglucose 6-phosphate (**3-3**), was produced *in situ* from chemically synthesized **3-18** using hexokinase (Figure 3-11). HPLC analysis of the reaction mixture showed the consumption of ATP and **3-8** and the production of ADP and UDP (Figure 3-20A, trace c). This observation indicated that the tandem reactions carried out by hexokinase and BexG2 were successful. In the absence of BexG2, only consumption of ATP and the production of ADP were observed (Figure 3-20A, trace b).

Therefore, the conversion of **3-8** to UDP must be BexG2-dependent. When both enzymes were omitted in a control, no reaction occurred (Figure 3-20A, trace a).

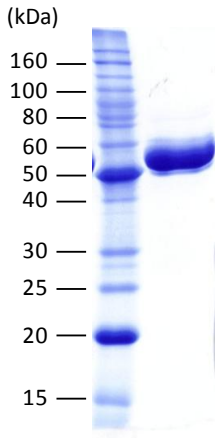


Figure 3-19. SDS-PAGE gel of the purified *N*-His_6 tagged BexG2

Calculated molecular weight of *N*-His_6-BexG2 (482 aa) is 53.6 kDa. The molecular weight marks are 220, 160, 120, 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, and 10 kDa (top to bottom).

To confirm the identity of the reaction product, a Corona CAD was used to detect product formation. As shown in Figure 3-20B, a new peak was observed in the hexokinase-BexG2 reaction. The corresponding fraction was collected, and the isolated product was subjected to NMR and MS analyses. All data are consistent with the assigned structure, **3-4**, which is linked by an α,α-1',1''-glycosidic bond (Figure 3-21).

In archaea, bacteria, fungi, plants, and animals, the dephosphorylation of trehalose 6-phosphate to trehalose is catalyzed by trehalose 6-phosphate phosphatase (TPP).[145] In fact, the TPS- and TPP-encoding genes are clustered in *Streptomyces avermitilis* and *Saccharopolyspora erythraea* genomes.[156,157] Although no TPP homologous gene is found in the *bex* cluster or in the boundary region, the phosphate group of **3-4** may be

hydrolyzed by an endogenous phosphatase encoded elsewhere in the genome. The hydrolysis may occur on the disaccharide **3-4** or after the disaccharide has been attached to the aglycone. Recently, a new trehalose biosynthetic route (catalyzed by TreTs), which utilizes UDP-glucose (**3-8**) and glucose instead of glucose 6-phosphate (**3-2**), was also found.[158-161] Although the poor sequence alignment results between BexG2 and these trehalose glycosyltransferring synthases (TreTs) renders this route less likely for the BexG2 reaction, we tested whether 2-thioglucose (**3-18**) could be processed by BexG2.

The reaction conditions were the same as described above except for the omission of hexokinase and ATP. HPLC analysis of the reaction showed that no UDP was produced (Figure 3-20A, trace d). Thus, it appears that BexG2 cannot process **3-18** as a substrate, and phosphorylation at C-6 position of 2-thioglucose is necessary for the BexG2 reaction. This is consistent with the sequence analysis of BexG2, which exhibits higher sequence similarity to trehalose phosphate synthases (TPSs) than to trehalose glycosyltransferring synthases (TreTs) as discussed above.

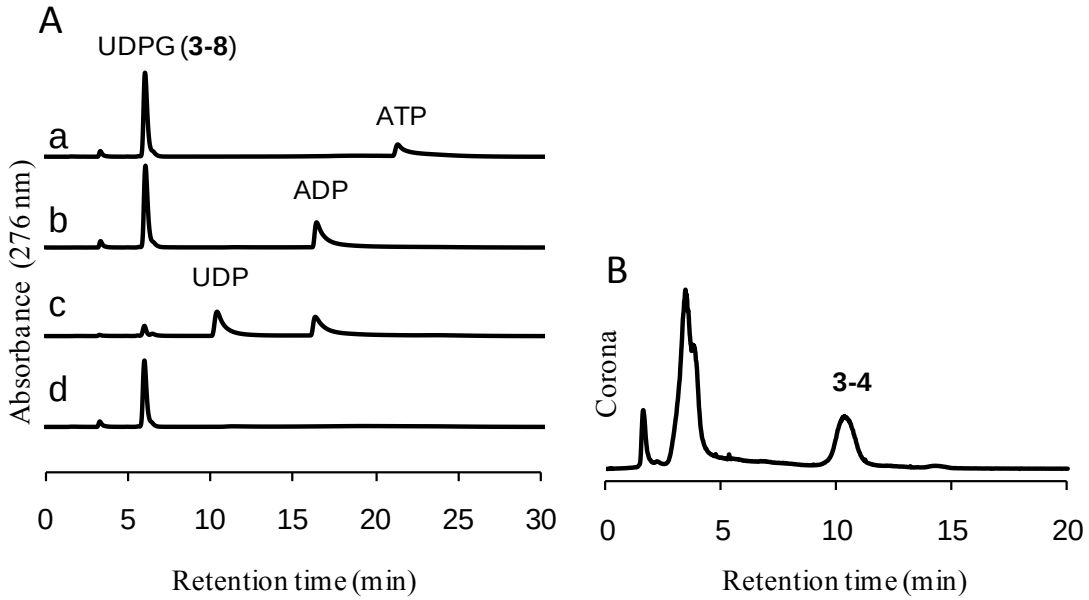
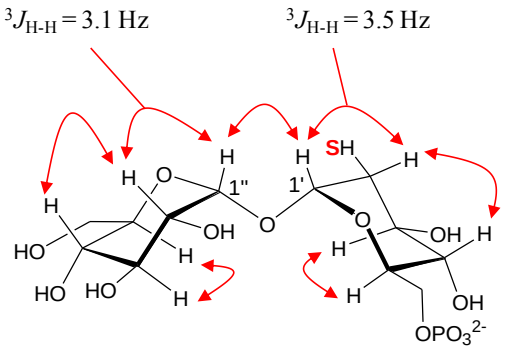


Figure 3-20. BexG2 activity and substrate specificity assay

(A) *In situ* production of 2-thioglucose 6-phosphate (**3-3**) from 2-thioglucose (**3-18**) using hexokinase was monitored by the consumption of ATP (21.4 min) and the production of ADP (16.5 min) at 276 nm. The conversion of **3-3** to **3-4** catalyzed by BexG2 was detected by the consumption of UDP-glucose (**3-8**) (6.1 min) and the production of UDP (10.5 min) at 276 nm. (a) Control reaction, no enzymes; (b) control reaction, no BexG2; (c) the complete reaction containing **3-18** (1 mM), ATP (1 mM), **3-8** (1 mM), $MgCl_2$ (10 mM), hexokinase (5 μg), and BexG2 (16 μM) in 50 mM Tris·HCl buffer, pH 8.0 (0.1 mL); (d) control reaction to investigate the substrate specificity, no ATP and hexokinase. (B) Isolation of BexG2 product **3-4**. The large-scale reaction contained **3-18** (10 mM), ATP (12 mM), **3-8** (12 mM), $MgCl_2$ (3 mM), hexokinase (20 μg), and BexG2 (30 μM) in 100 mM Tris·HCl buffer, pH 8.0 (1 mL). The resulting product **3-4** (10.4 min, monitored by the Corona CAD) was collected and characterized.



2-thio-D-trehalose 6-phosphate
(**3-4**)

Figure 3-21. Selected NOESY results for 2-thiotrehalose 6-phosphate produced from BexG2-catalyzed reaction.

3.3.2 BexG2 Substrate Specificity Assay.

D-Glucose 6-Phosphate as the Substrate

Based on the sequence similarity between BexG2 and trehalose 6-phosphate synthases (TPSs), the second possible pathway to form 2-thiotrehalose 6-phosphate can be considered (**3-2** → **3-6** → **3-4**, Figure 3-1). Although the transformation from **3-6** to **3-4** is not consistent with the proposed function and mechanism of BexX (Figure 2-23), it is interesting to test if BexG2 can utilize glucose 6-phosphate (**3-2**) in addition to 2-thioglucose 6-phosphate (**3-3**) to produce trehalose 6-phosphate (**3-6**). To address this question, the reaction solution containing **3-2** and UDP-glucose (**3-8**) was incubated with BexG2. The reaction was quenched in a time dependent manner (at 1 h, 2.5 h and 3.8 h) and analyzed by ^{1}H NMR. As expected, the decrease of signals corresponding to the anomeric protons of glucose 6-phosphate (Figure 3-22, a and b) was observed along with an increase of the intensity of the new signals corresponding to the anomeric protons of trehalose 6-phosphate (Figure 3-22, c and d). To confirm the identity of the product, the

reaction solution was treated with alkaline phosphatase, and the resulting dephosphorylated compounds were co-analyzed with D-glucose (**3-19**) and D-trehalose (**3-20**) standards by ^{1}H NMR. Successive addition of each standard compound enhanced the NMR signals obtained from the reaction mixture (Figure3-23). These results clearly show that BexG2 can accept glucose 6-phosphate (**3-2**) as the substrate and can produce trehalose 6-phospahte (**3-4**).

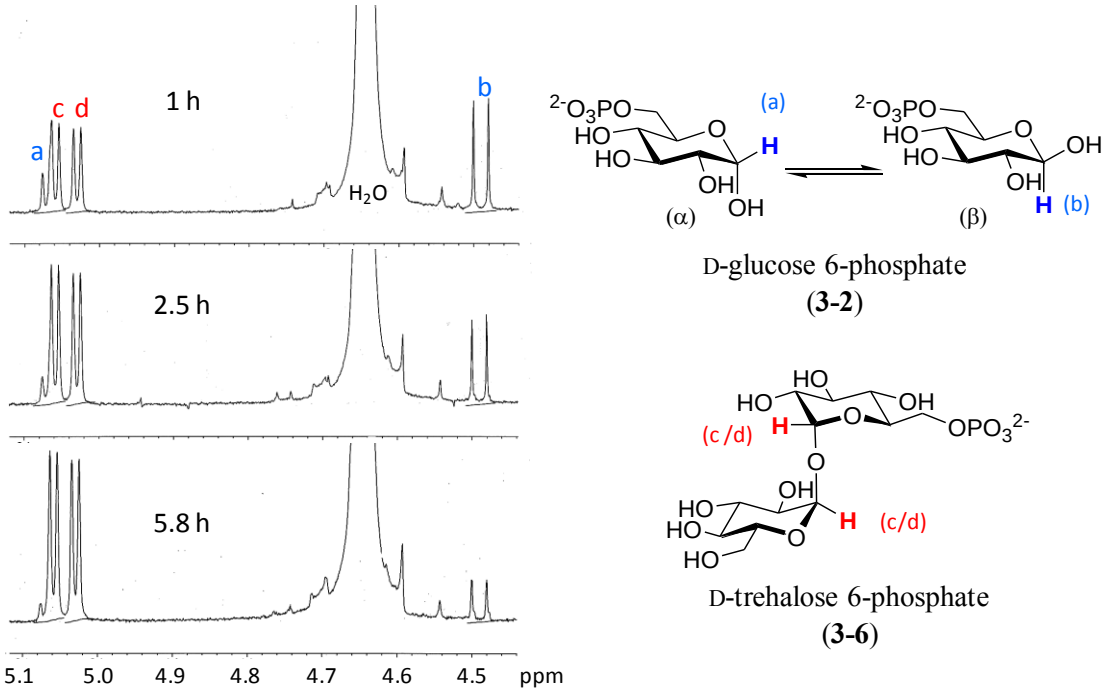


Figure 3-22. ^{1}H NMR analysis of the BexG2-catalyzed reaction using glucose 6-phosphate and UDP-glucose as the substrates.

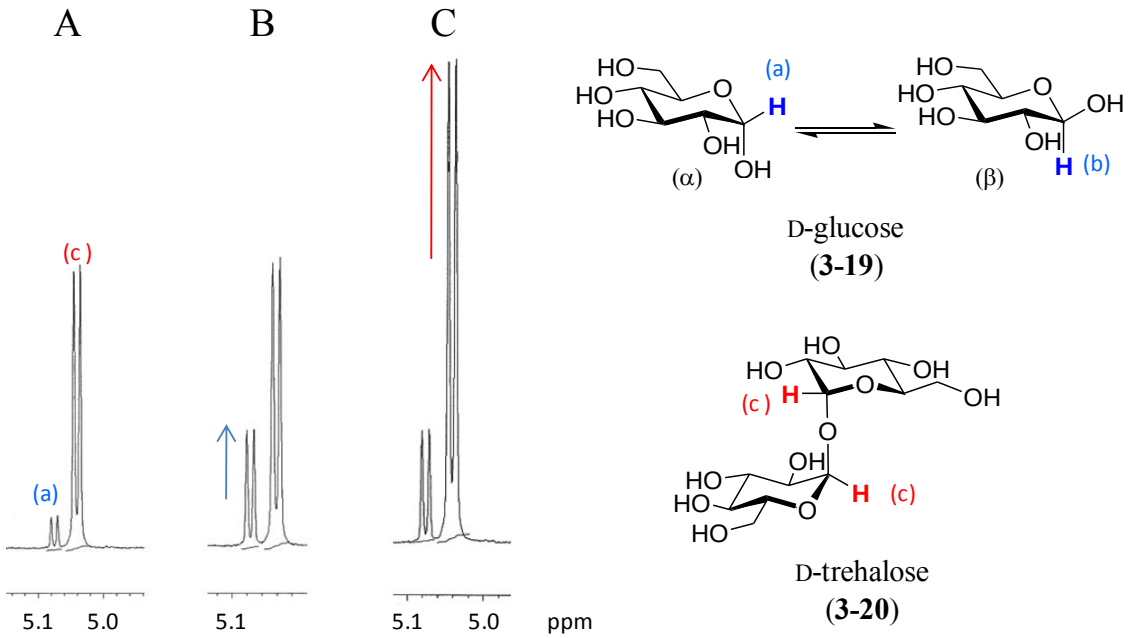


Figure 3-23. [1]H NMR analysis of the BexG2-catalyzed reaction product using D-glucose 6-phosphate and UDP-glucose as the substrates.

The dephosphorylated compounds (A) were added D-glucose (**3-19**) standard (B) and D-trehalose (**3-20**) standard (C) successively.

Investigation of Other Glycosyl Donors

Although UDP-glucose (**3-8**) was found to be an effective glycosyl donor for the BexG2-catalyzed reaction as described above, other NDP-sugars (UDP-xylose (**3-21**), UDP-glucuronic acid (**3-22**), UDP-galactose (**3-23**), TDP-glucose (**3-24**), ADP-glucose (**3-25**), and GDP-glucose (**3-26**)) were also used to test the substrate specificity of the enzyme (Figure 3-24). These NDP-sugars were incubated with glucose 6-phosphate (**3-2**) and BexG2, and the reaction was analyzed by HPLC. Successful glycosyl transfer was judged by detection of the released NDP. Among the tested NDP-sugars, only ADP-glucose yielded the expected product, and the other NDP-sugars did not release NDP except for the positive control, UDP-glucose (Figure 3-25). The results suggest that

BexG2 possesses relatively strict substrate specificity for the glycosyl donor. In the literature, more relaxed substrate specificity for the trehalose 6-phosphate synthases (TPSs) OtsA from *Thermus thermophilus* RQ-1 and TPS from *Mycobacterium tuberculosis* were reported.[162,163] The former enzyme can utilize UDP-glucose (**3-8**), TDP-glucose (**3-24**), ADP-glucose (**3-25**), and GDP-glucose (**3-26**) while it cannot take ADP-mannose, GDP-mannose, UDP-mannose, and UDP-galactose (**3-23**) as the substrate. The latter enzyme can take CDP-glucose in addition to the four NDP-glucoses mentioned above (**3-8**, **3-24**, **3-25**, and **3-26**).

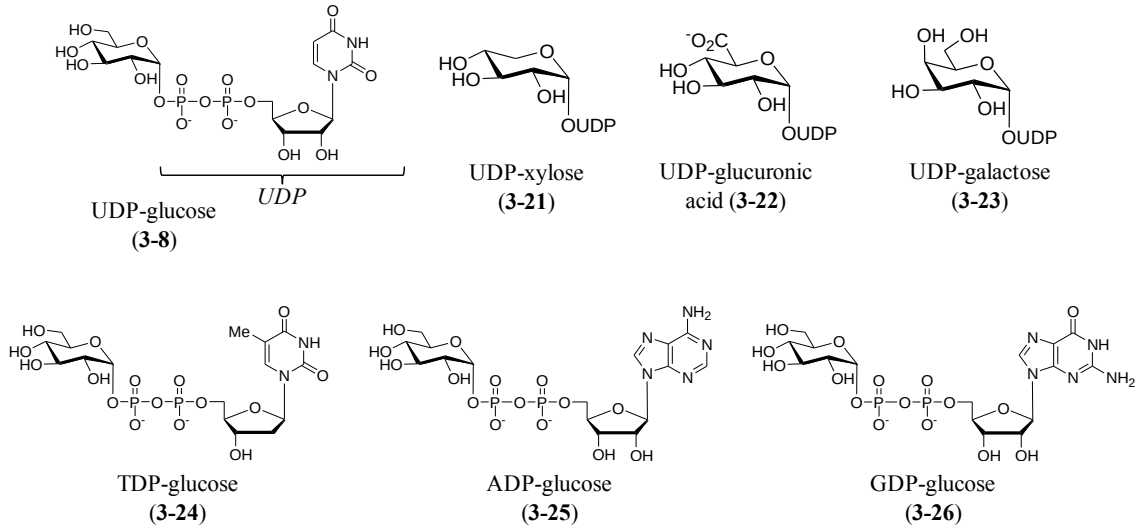


Figure 3-24. Glycosyl donor candidates tested for BexG2-catalyzed reaction.

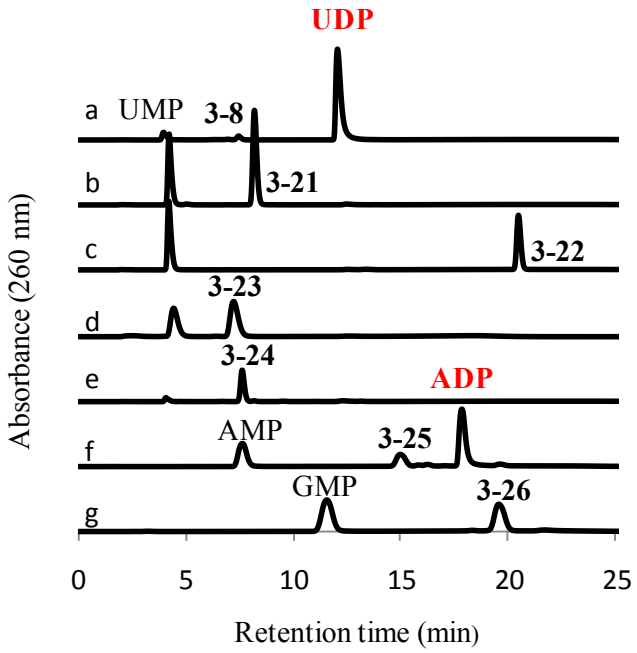


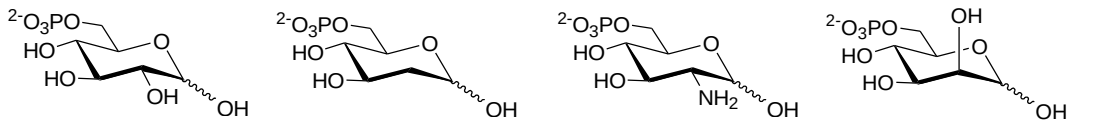
Figure 3-25. BexG2 glycosyl donor specificity assay.

(a) UDP-glucose (**3-8**, positive control) (b) UDP-xylose (**3-21**) (c) UDP-glucuronic acid (**3-22**) (d) UDP-galactose (**3-23**) (e) TDP-glucose (**3-24**) (f) ADP-glucose (**3-25**) (g) GDP-glucose (**3-26**). Produced NDP were shown in red color (**a** and **f**). Experiments **a**–**c** and **e** were done on the same day. Experiment **d** was done on a separate day. Experiments **f** and **g** were also done on a different day. Slight difference of the retention times on separate analyses was observed although the HPLC conditions were identical. Gradual degradation of NDP-sugars to NMP was observed under the tested condition.

Investigation of Other Glycosyl Acceptors

Substrate specificity of BexG2-catalyzed reaction for the glycosyl acceptor was similarly examined. Various sugar monophosphates (Figure 3-26) were incubated with UDP-glucose (**3-8**) and BexG2, and the reaction were analyzed by HPLC. Interestingly, release of UDP from **3-8** was observed in all reactions (Figure 3-27). This indicates that 2-deoxy-D-glucose 6-phosphate (**3-27**), D-glucosamine 6-phosphate (**3-28**), and D-mannose 6-phoshate (**3-29**) in addition to D-glucose 6-phosphate (**3-2**) and the natural

substrate, 2-thio-D-glucose 6-phosphate (**3-3**), are all possible glycosyl acceptors in the BexG2-catalyzed reaction. These results clearly suggest that modification of C-2 position of the glycosyl acceptor is tolerated for this enzyme while a phosphate group at C-6 is required as described above (Figure 3-20A, trace d). In the case of the reported trehalose 6-phosphate synthase (TPS) OtsA from *T. thermophilus* RQ-1, D-glucose 6-phosphate (**3-2**) could not be replaced by D-mannose 6-phoshate (**3-29**).[162] The more relaxed substrate specificity of BexG2 at the C-2 position is interesting because it suggests the enzyme has evolved to accept 2-thio-D-glucose 6-phosphate (**3-3**) instead of D-glucose 6-phosphate (**3-2**) as the natural substrate.



D-glucose 6-phosphate (**3-2**) | 2-deoxy-D-glucose 6-phosphate (**3-27**) | D-glucosamine 6-phosphate (**3-28**) | D-mannose 6-phosphate (**3-29**)

Figure 3-26. Glycosyl acceptor candidates tested for BexG2-catalyzed reaction.

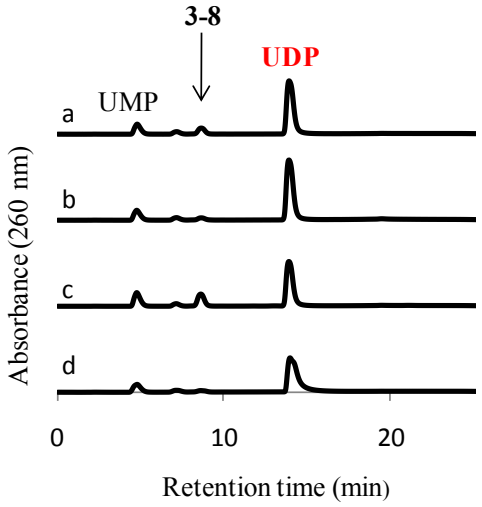


Figure 3-27. BexG2 glycosyl acceptor specificity assay

(a) D-glucose 6-phosphate (**3-2**, positive control) (b) 2-deoxy-D-glucose 6-phosphate (**3-27**) (c) D-glucosamine 6-phosphate (**3-28**) (d) D-mannose 6-phoshate (**3-29**) Produced UDP was shown in red color. Experiments **a**–**c** were done on the same day. Experiment **d** was done on a different day. Slight difference of the retention times on separate analyses was observed although the HPLC conditions were identical.

3.3.3 Competitive Assay of BexG2-Catalyzed Glycosyltransfer Reaction.

Although BexG2 can utilize both D-glucose 6-phosphate (**3-2**) and 2-thio-D-glucose 6-phosphate (**3-3**), **3-3** was expected to be a more suitable substrate based on the predicted function of BexX (Figure 3-1). To test this hypothesis, a competitive assay using both **3-2** and **3-3** was designed. The substrates **3-2** (5 mM) and **3-3** (5 mM) were incubated with various concentration (0−12.5 mM) of UDP-glucose (**3-8**) in the presence of BexG2, and the reaction mixture was analyzed by HPLC with a Corona CAD detector (Figure 3-28).

When low concentration of UDP-glucose (**3-8**, less than 1 equivalent of each glycosyl acceptor) was used, appearance of the signal corresponding to 2-thio-D-trehalose

6-phosphate (**3-4**) was observed. Interestingly, the D-glucose 6-phosphate (**3-2**) peak was barely decreased under these conditions. This peak (**3-2**) was converted to D-trehalose 6-phosphate (**3-6**) only when a higher concentration of UDP-glucose (**3-8**, more than 1 equivalent of each glycosyl acceptor) was used. The peak integration of each compound was then converted to the corresponding concentration and replotted against the initial concentration of UDP-glucose (Figure 3-29). This figure clearly shows that BexG2 utilizes 2-thio-D-glucose 6-phosphate (**3-3**) more preferably than D-glucose 6-phosphate (**3-2**) when both substrates exist in the solution. This substrate selectivity may explain the contradictory observation that BexG2 can take **3-2** as the substrate even though the product **3-6** is not likely involved in the biosynthetic pathway of BE-7585A (Figure 3-1). Moreover, under *in vivo* conditions D-glucose 6-phosphate (**3-2**) may be a much better substrate for BexX than for BexG2 and will thus be taken up by BexX and be converted to 2-thio-D-glucose 6-phosphate (**3-3**), which is a more preferential substrate for BexG2. Although kinetic experiments are required to investigate this substrate selectivity of BexG2 in more detail, the results shown here strongly support the function of the enzyme as the 2-thio-D-trehalose 6-phosphate synthase rather than an ordinary trehalose 6-phosphate synthase (TPS).

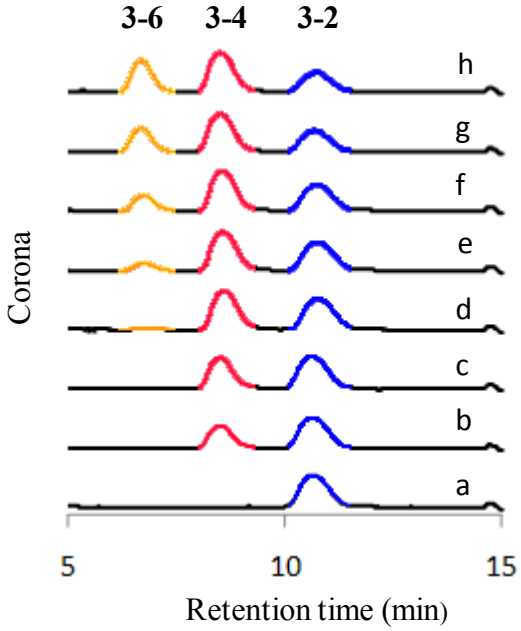


Figure 3-28. BexG2 competitive assay I

Peaks observe are D-glucose 6-phosphate (**3-2**), D-trehalose 6-phosphate (**3-6**), and 2-thio-D-trehalsoe 6-phosphate (**3-4**). One of the substrates, 2-thio-D-glucose 6-phosphate (**3-3**), was not shown because it could not be detected by Corona CAD under the tested conditions. Initial concentrations of UDP-glucose are (a) 0, (b) 2.50, (c) 3.75, (d) 5.00, (e) 6.25, (f) 7.50, (g) 10.0, (h) 12.5 mM.

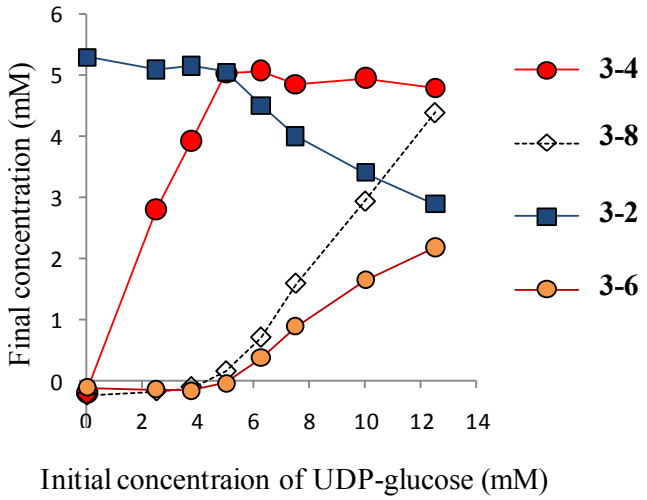


Figure 3-29. BexG2 competitive assay II

Data shown are D-glucose 6-phosphate (**3-2**), D-trehalose 6-phosphate (**3-6**), 2-thio-D-trehalsoe 6-phosphate (**3-4**), and UDP-glucose (**3-8**).

3.3.4 Non-Enzymatic C−S Bond Formation Between a Thiosugar and Aglycon

How the thiodisaccharide (**3-4** or its hydrolyzed product without 6-phosphate group) is attached to C-5 of the benz[*a*]anthraquinone core is not clear. A notable example is found in the literature addressing similar C−S bond formation in urdamycin E (**3-10**, Figure 3-2). [148] Their experiments showed that C-5 position of urdamycin A (**3-9**) is electrophilic enough to be nonenzymatically attacked by thiol-containing compounds such as methylmercaptan (**3-12**) or glutathione at physiological pH.

To test this nonenzymatic Michael-type addition model, urdamycin A (**3-9**) was incubated with 2-thio-D-glucose (**3-18**) in an aqueous buffer at pH 8.0. Interestingly, the pale yellow color of the urdamycin A (**3-9**) solution was changed to a pale pink color within one minute after addition of **3-18**. This color change was not observed when **3-18** was replaced by D-glucose (**3-19**), suggesting that the thiol functional group is important. The reaction mixtures were then analyzed by HPLC. As expected, the solution containing

3-18 showed a new peak (**3-30**) while incubation with **3-19** did not affect the original urdamycin A (**3-9**) peak (Figure 3-30). MS analysis of the isolated peak (**3-30**) was consistent with the expected structure with a 2-thioglucose attached to urdamycin A. As a control, BE-7585A (**3-1**) was also incubated with 2-thio-D-glucose (**3-18**) under the same conditions. HPLC analysis showed that no conjugate was formed in this reaction. Since C-5 position of BE-7585A (**3-1**) already has the thiodisaccharide moiety installed, the result shown here is consistent with the proposed nucleophilic attack at the C-5 position of urdamycin A (**3-9**).

The observed color change was then analyzed using a UV-vis spectrophotometer. Upon addition of 2-thio-D-glucose (**3-18**) into an urdamycin A (**3-9**) solution, an increase of absorbance and red shift of the wavelength of maximum absorbance (λ_{max}) were observed (Figure 3-31). When a sufficient amount (3 equivalents of **3-9**) of **3-18** was added, the spectrum exhibited broad absorbance over 400−600 nm with a λ_{max} at 480 nm, which closely resembles that of BE-7585A (**3-1**, Figure 2-6). The similarity of the two absorbance spectra is consistent with the expected thiosugar addition at the same position of the benz[*a*]anthraquinone core, which influences the π conjugation system on the aromatic ring and yields the unique absorbance property.

A proposed mechanism of the Michael-type addition of the thiosugar to the C-5 site of the benz[*a*]anthraquinone core is shown in Figure 3-31. The C-5 site is expected to be electrophilic and can accept the nucleophilic attack from the thiodisaccharide. Subsequent deprotonation at C-5 leads to a hydroquinone intermediate (**3-31**), which can be readily oxidized to the quinone form (**3-32**) by molecular oxygen.

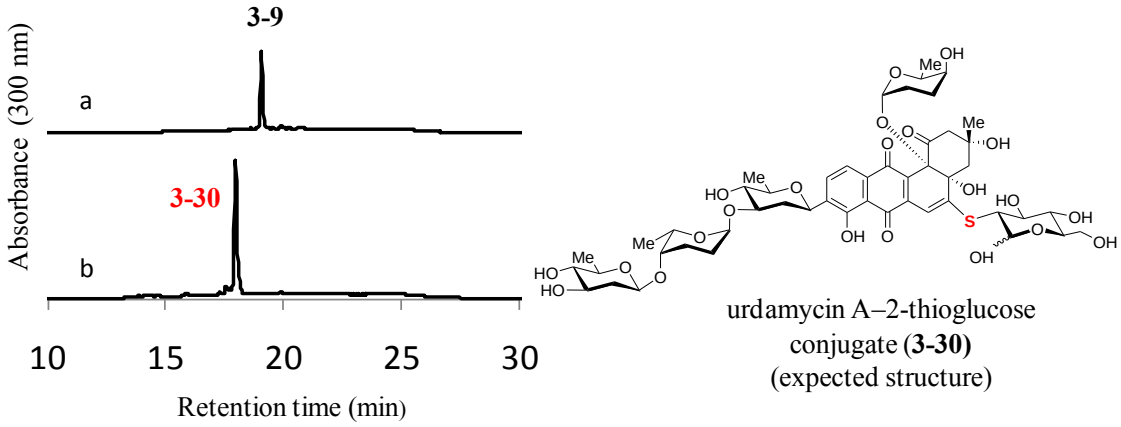


Figure 3-30. HPLC analysis of the nonenzymatic C–S bond formation

(a) Control reaction containing urdamycin A (**3-9**) and D-glucose (**3-19**). (b) Reaction containing **3-9** and 2-thio-D-glucose (**3-18**). Expected structure of the urdamycinA–2-thioglucose conjugate (**3-30**) is shown in right.

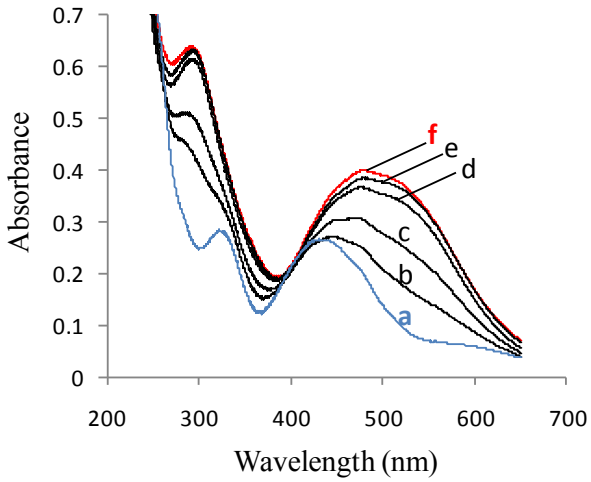


Figure 3-31. Absorbance analysis of the nonenzymatic C–S bond formation

(a) urdamycin A (**3-9**, 0.2 mM) (b) addition of 2-thio-D-glucose (**3-18**, 0.1 mM) (c) repeat of **b** (2 min after the addition) (d) addition of 2-thio-D-glucose (**3-18**, 0.6 mM final) (e) repeat of **d** (2 min after the addition) (f) repeat of **d** (4 min after the addition)

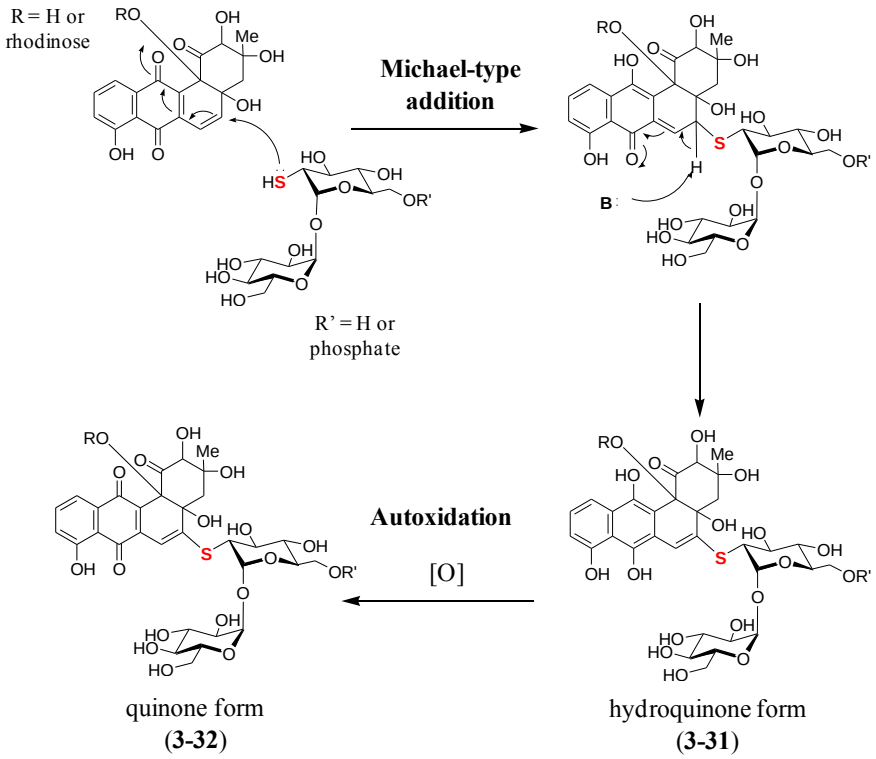


Figure 3-32. Proposed mechanism of the nonenzymatic C–S bond formation.

Although it was confirmed that the C−S bond between a benz[*a*]anthraquinone core and a thiosugar was easily formed nonenzymatically, whether this process is controlled by an enzyme or not under *in vivo* condition remains unresolved question. Based on the BE-7585A biosynthetic gene cluter (*bex*) described in Chapter 2, no candidate gene possibly responsible for this transformation can be found, indicating a nonenzymatic process. On the other hand, BE-7585A derivatives containing other thioconjugates or no thioadduct are not found in *A. orientalis*. This suggests that there may be some regulating mechanism *in vivo* to prevent other thionucleophiles to compete with **3-4** for the benz[*a*]anthraquinone core. One possible mechanism is that a specific interaction exists between protein(s) responsible for angucycline core biosynthesis and those responsible for thiosugar biosynthesis including BexG2. This interaction may allow

2-thiotrehalose 6-phosphate (**3-4**) to be delivered directely to the site where the benz[*a*]anthraquinone core is produced, leading to spontaneous bond formation before they diffuse in the cell to react with other thiol-containing compounds.

Even in this case, however, 2-thioglucose 6-phosphate (**3-4**) released from BexX must be captured by BexG2 rapidly before it reacts with the core structure unless BexG2 can accept the benz[*a*]anthraquinone−monothiosugar conjugate as the substrate (Figure 3-1, **3-7** → **3-5**). Although the observed relaxed substrate specificity at C-2 position of the glycosyl acceptor in the BexG2-catalyzed reaction supports this possibility, further experiments are clearly necessary to verify this hypothesis. To address this question, a more appropriate angucycline core structure may be required rather than the model compound, urdamycin A (**3-8**). To obtain the substrate involved in the thio-coupling step in the BE-7585A biosynthetic pathway, enzymatic total synthesis of benz[*a*]anthraquinone core moiety (using BexA, BexB, BexC, BexD, BexL, BexF, BexI, BexE, BexM and BexK, see Figure 2-21) and the rhodinose moiety (using BexU, BexS, BexV, BexQ, BexT, BexO, BexH and BexG1, see Figure 2-22) are in progress (all genes have been cloned, many enzymes have been expressed and purified, and some of their functions have been confirmed). In addition, structural analysis is a powerful approach to examine how a large modification at C-2 of the glucose 6-phosphate moiety can be accommodated by the enzyme. The observed substrate selectivity for 2-thio-D-glucose 6-phosphate (**3-4**) over D-glucose 6-phosphate (**3-2**) may be explained if the structural data is obtained and compared with the reported data of a trehalose 6-phosphate synthase (TPS) OtsA from *E. coli.*[164-167]

3.4 CONCLUSIONS

In this chapter, BexG2, one of the key enzymes responsible for the production of the unique thiodisaccharide moiety of BE-7585A (**3-1**), was heterologously expressed in *E. coli,* and its glycosyl transfer activity was investigated using a variety of substrates. The predicted natural substrate 2-thio-D-glucose 6-phosphate (**3-3**) was chemoenzymatically prepared and utilized for the reaction with UDP-glucose (**3-8**). The product was isolated and fully characterized using spectroscopic methods to confirm the identity of 2-thio-D-trehalose 6-phosphate (**3-4**). The BexG2-catalyzed reaction was found to exhibit relatively strict substrate specificity for the glycosyl donor as compared to other reported TPSs from *T. thermophilus* or *M. tuberculosis*.[162,163] In contrast, remarkable substrate tolerance at the C-2 position of the glycosyl acceptor was found. All C-2 modification examined, including 2-hydroxy- (both equatorial and axial), 2-thio- (equatorial), 2-amino- (equatorial), and 2-deoxysugars were accepted by the enzyme although a phosphate group at C-6 position is essential. The extended substrate specificity at C-2 position of BexG2 as compared to a reported TPS from *T. thermophilus*[162] suggests that the BexG2 active site must have evolved to accept the unusual thio-modification at that position. Whether this enzyme can accept more sterically demanding substrates, such as benz[*a*]anthraquinone−thioglucose 6-phosphate, is not clear at this point. However, the apparent substrate selectivity observed for 2-thio-D-glucose 6-phosphate (**3-3**) over D-glucose 6-phosphate (**3-2**) strongly supports the enzyme's function as a 2-thio-D-trehalose 6-phosphate synthase.

A thioether bond connecting the benz[*a*]anthraquinone core to thiodisaccharide moiety is another unnsual structure observed in BE-7585A (**3-1**). Although a similar thioether bond found in urdamycin E (**3-10**) had been studied before,[148] we first confirmed that a thiosugar (**3-18**) can spontaneously form a conjugate with the model

compound urdamycin A in a buffer (pH 8.0) without any organic cosolvent. These results, including additional control experiments using D-glucose (**3-19**) or BE-7585A (**3-1**) and the absorbance property of the conjugate (**3-30**), are all consistent with the proposed nonenzymatic formation of the C−S bond between the benz[*a*]anthraquinone core at C-5 and thiol functional group of **3-18**. Whether this nonenzymatic process is physiologically relevant or not remains to be verified. Neverthless, we have proposed that specific protein−protein interactions among the BE-7585A biosynthetic enzymes may help enrich the local concentrations of the angucycline core and thiosugar to an appropriate level to prevent formation of other thioadducts. Verification of this hypothesis requires further experiments including the functional investigation of the enzymes involved in the construction of the benz[*a*]anthraquinone core. Preparation of the enzymes involved in the biosyntheses of the angucycline core and the rhodinose moiety is in progress.

Chapter 4: Biosynthetic Studies of BE-7585A (III): Mechanistic Studies of the 2-Thiosugar Synthase, BexX

4.1 INTRODUCTION

In the previous chapters, the biosynthetic pathway of BE-7585A (**4-1**) was proposed based on genetic, structural, and enzymatic functional analyses (Figure 4-1). For the 2-thiosugar biosynthesis, a key enzyme, BexX, which shares sequence similarity with thiazole synthases (ThiGs), was found to be a putative 2-thio-D-glucose 6-phosphate synthase. The expected thiosugar product (**4-3**) of BexX would then be coupled with a glucose moiety by BexG2 to form the unique thiodisaccharide structure of BE-7585A (**4-1**).

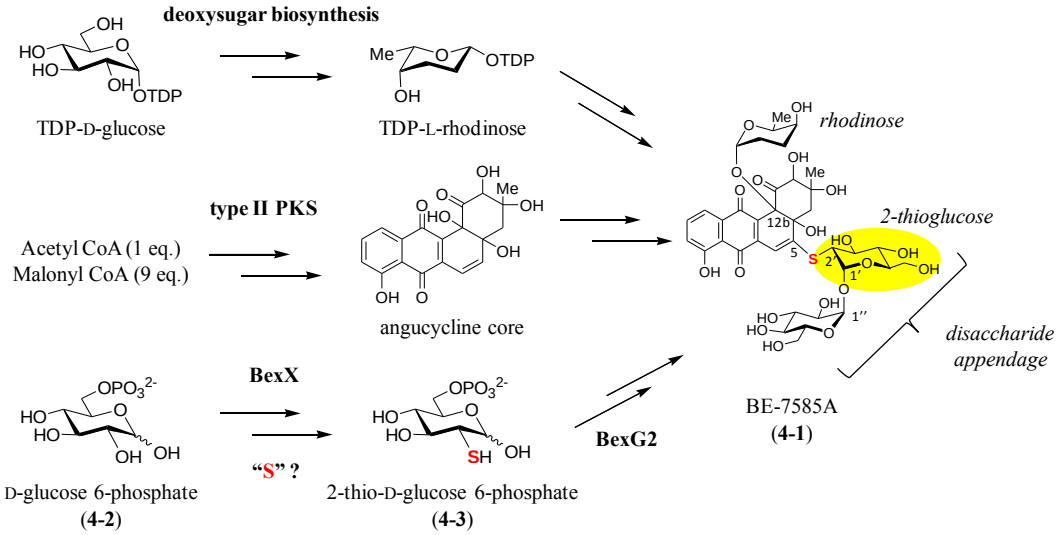


Figure 4-1. Proposed biosynthetic pathway of BE-7585A.

The proposed reaction for BexX-catalyzed thiosugar formation is analogous to that of ThiG during thiazole biosynthesis and is based on the sequence similarity of BexX and ThiG (Chapter 2). ThiG from *B. subtilis* initiates the reaction by forming an imine

complex between the active site lysine residue and the C-2 keto group of 1-deoxy-D-xylulose-5-phosphate (DXP, **4-4**; Figure 4-2).[168] A sulfur carrier protein, ThiS, then acts as the sulfur donor for sulfur incorporation at C-3 of the DXP−lysine−ThiG adduct (**4-5**).[66] Since ThiG is a key enzyme in the biosynthesis of the thiazole phosphate (**4-6**) moiety of thiamin, we predicted that the *bexX* gene may also play a key role in the biosynthesis of the 2-thiosugar moiety of BE-7585A in *A. orientalis*.

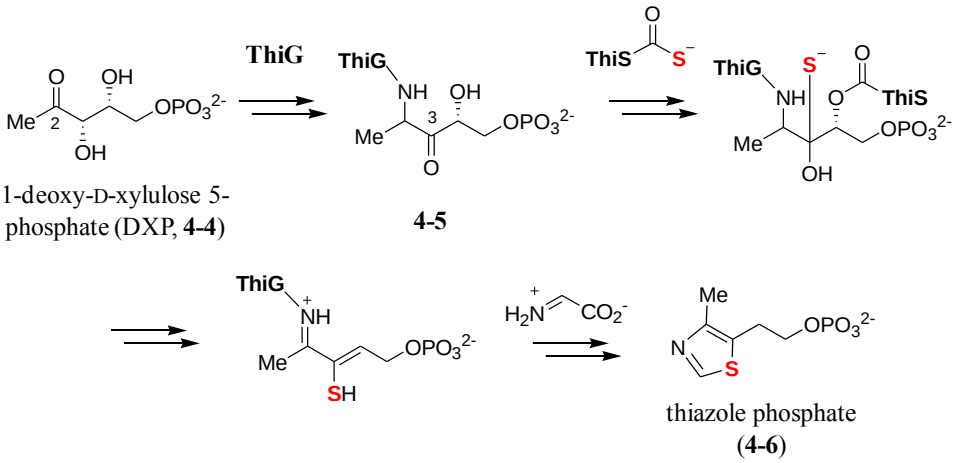


Figure 4-2. ThiG-catalyzed thiazole phosphate formation.

One hypothesis is BexX utilizes D-glucose 6-phosphate (**4-2**) as the substrate with the mechanism of sulfur incorporation to form the 2-thiosugar resembling that of the ThiG reaction in thiamin biosynthesis (Figure 4-3, pathway A). However, an alternative hexose monophosphate substrate may be D-fructose 6-phosphate (**4-7**), which could form a Schiff base with the active-site lysine at C-2, thereby directly activating the target position for sulfur incorporation (**4-14**; Figure 4-3, pathway B). In fact, an analogous C-2 activation has been demonstrated for the D-glucosamine 6-phosphate synthase (GlmS)-catalyzed introduction of an amino group at C-2 of fructose 6-phosphate (**4-7**).[169,170]

GlmS from *E. coli* is composed of a glutaminase domain that supplies ammonia from L-glutamine (**4-19**) and a synthase domain that catalyzes the amination and isomerization of **4-7** to D-glucosamine 6-phosphate (**4-21**; Figure 4-4). Structural analysis has revealed the existence of a channel by which ammonia passes between the active sites of the two GlmS domains.[169] Although BexX shares no sequence similarity with GlmS and possess no additional domain, endogenous cysteine desulfurase-type proteins, which transfer bisulfide to several *in vivo* targets such as iron-sulfur proteins, may act as the sulfur donor for BexX catalyzed reaction.

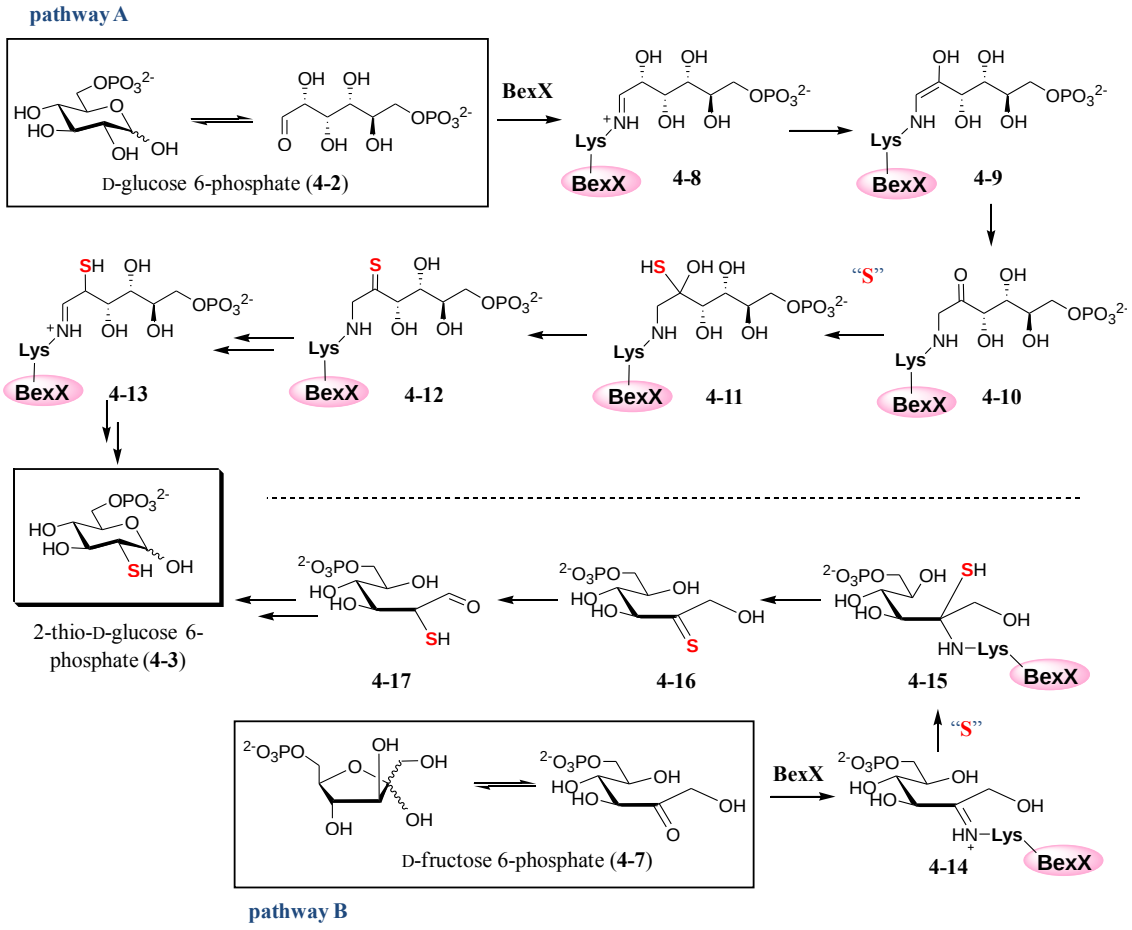


Figure 4-3. Two proposed mechanisms of BexX-catalyzed 2-thioglucose 6-phosphate formation.

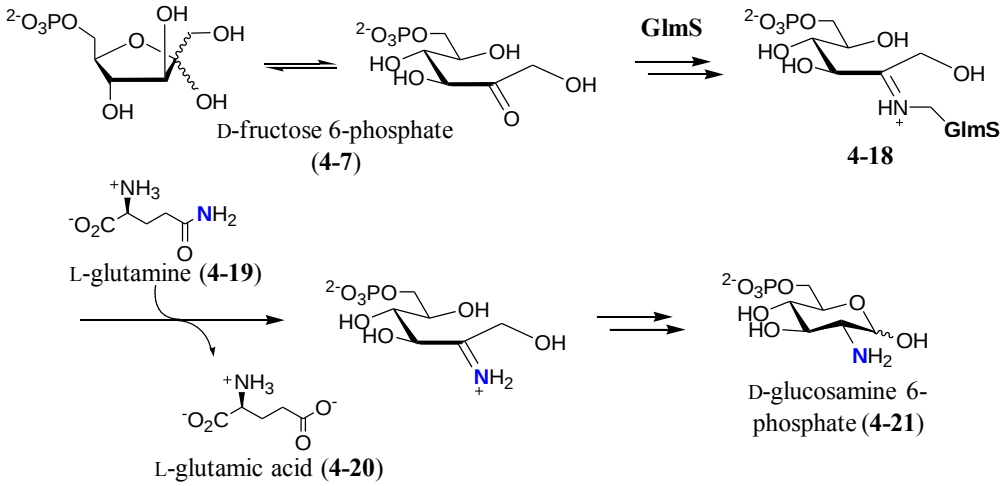


Figure 4-4. GlmS-catalyzed glucosamine 6-phosphate formation.

In addition to mechanistic studies of 2-thiosugar biosynthesis, identification of the sulfur donor is another focus of this chapter. In the pathway proposed in Figure 4-3, sulfur incorporation is expected to proceed *via* an ionic mechanism. This is because the C-2 position exists as a ketone or an imine, it is activated and susceptible to nucleophilic attack (**4-10** in the pathway A or **4-14** in the pathway B). Potential sulfur donors for this pathway may be an inorganic bisulfide or a protein persulfide (see Figure 4-5). Protein persulfides have been identified as a part of cysteine desulfurase-type proteins, in which they originate from L-cysteine. Protein persulfides are also found in rhodanese-type proteins, in which they are generated from thiosulfate. A sulfur carrier protein, which harbors a reactive thiocarboxylate at its *C*-terminus as ThiS in the thiamin biosynthetic pathway, could likewise play the role of a nucleophilic sulfur donor. Although no genes related to any of these sulfur delivery systems are found in the BE-7585A biosynthetic (*bex*) gene cluster, it is possible that an endogenous sulfur transfer protein from pathways of primary metabolism may be recruited for production of the 2-thiosugar (Figure 4-5).

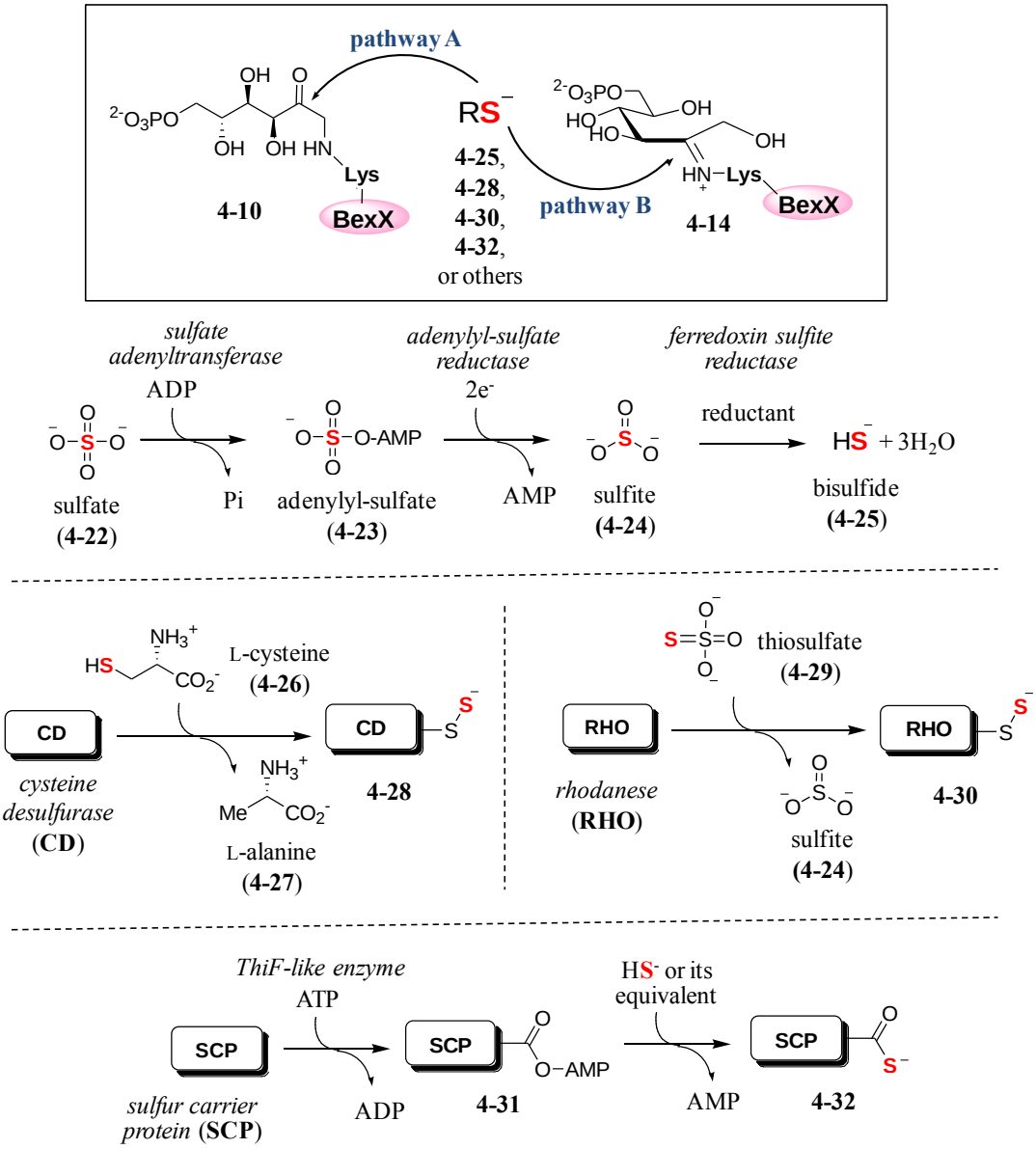


Figure 4-5. Possible sulfur donors for the 2-thiosugar biosynthesis.

In this chapter, MS analyses of recombinant BexX from *E. coli* are described. A substrate molecule that forms a covalent bond with BexX was identified. The amino acid residue involved in this covalent linkage was determined by mutagenesis studies and LC-MS/MS analysis of the trypsin digested BexX fragments. Strong evidence for the

proposed reaction mechanism was obtained by detecting the stable keto form of the BexX−substrate complex using carbonyl-directing chemical reagents. Genome mining of the entire sequenced genome of *A. orientalis* was also performed to identify potential sulfur donors. The protein candidates responsible for the endogenous sulfur transfer reaction were individually examined *in vitro* for their ability to incorporate a sulfur atom into the BexX−substrate complex. Lastly, the *in vitro* enzymatic synthesis of the 2-thiosugar was accomplished for the first time.

4.2 EXPERIMENTAL PROCEDURES

4.2.1 General.

Materials

All chemicals and reagents were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA), and were used without further purification unless otherwise specified. Sequencing grade modified trypsin was a product of Promega (Madison, WI). Adenylate kinase (AK), pyruvate kinase (PK), and lactate dehydrogenase (LDH) were purchased from Sigma-Aldrich Chemical Co. Calf intestinal alkaline phosphatase (CIP) was a product of New England Biolabs (Ipswich, MA). Materials used for molecular cloning were identical to those described in Chapter 2 except the PureLink Genomic DNA Mini Kit from Invitrogen (Carlsbad, CA) used for genomic DNA extraction. Materials used for protein expression and purification were identical to those described in Chapter 3 with the exception of Pre-cast gels (NuPAGE Novex 4-12% Bis-Tris Gel from Invitrogen) used for the pull-down assay. The CarboPac PA1 high-performance liquid chromatography (HPLC) column was obtained from

Dionex (Sunnyvale, CA). Analytical C_{18} HPLC columns were purchased from Varian (Palo Alto, CA).

Bacterial Strains and Plasmids

A. orientalis subsp. *vinearia* BA-07585 was generously provided by Banyu Pharmaceutical Co. (Tokyo, Japan). Cosmid C006 was obtained from an *A. orientalis* genomic library as described in Chapter 2. *E. coli* DH5α, acquired from Bethesda Research Laboratories (Gaithersburg, MD), was used for routine cloning experiments. The protein overexpression hosts *E. coli* BL21 star (DE3) and *E. coli* Rosetta 2 (DE3) were obtained from Invitrogen (Carlsbad, CA) and Novagen (Madison, WI), respectively. Vectors pET24b(+) and pET28b(+) for protein over-expression were purchased from Novagen. Standard genetic manipulations of *E. coli* were performed as described by Sambrook et al.[107]

Instrumentation

.UV−vis spectra were recorded using a Beckman DU 650 spectrophotometer. Evaporation of solution samples in 1.5 mL microcentfiguge-tubes was performed using a Savant Speedvac SC100 equipped with a Savant refrigerated condensation trap RT100 and Gel Pump GP100. HPLC was performed on a Beckman Coulter System Gold equipped with a UV detector. DNA concentrations were measured using a NanoDrop ND-1000 UV−vis instrument from Thermo Fisher Scientific. DNA sequencing was performed by the core facility of the Institute of Cellular and Molecular Biology at the University of Texas, Austin. Vector NTI Advance 10.1.1 from Invitrogen was used for sequence alignments. Mass spectroscopy was performed at the Mass Spectrometry core facility in the Department of Chemistry and Biochemistry and in the College of Pharmacy at the University of Texas, Austin.

4.2.2 Cloning of *BexX*.

The *bexX* gene was PCR-amplified from cosmid C006 using primers with engineered *Nde*I and *Hind*III restriction sites. The sequences of the primers are shown in Table 4-1. The PCR-amplified gene fragments were purified, digested with *Nde*I and *Hind*III, and ligated into both pET24b(+) and pET28b (+) vectors, which had also been digested with the same restriction enzymes. The resulting plasmids, *bexX*/pET24b(+) and *bexX*/pET28b(+), were sequenced using the T7 or T7 terminal universal primer and used to transform the *E. coli* BL21 star (DE3) strain for protein overexpression.

Table 4-1. Primers used for constructing *bexX*/pET24b(+) and *bexX*/pET28b(+)

Primer name	Sequence
bexX for pET24/28b(+)-forward	5'-TCTTAAA**CATATG**ACCATTCCGCACATCGGC-3'
bexX for pET24b(+)-reverse	5'-ATAT**AAGCTT**GACCGCCGCTGCTACCTC-3'
bexX for pET28b(+)-reverse	5'-TTAC**AAGCT*T***CAGACCGCCGCTGCTACCTC-3'

The engineered restriction sites are shown in bold, the start codon is shown in bold and also underlined, and the stop codon is italic.

4.2.3 Expression and Purification of *C*-His_6-BexX and *N*-His_6-BexX.

An overnight culture of *E. coli* BL21 star (DE3)-*bexX*/pET24b(+) or *E. coli* BL21 star (DE3)-*bexX*/pET28b(+), grown in the LB medium (10 mL) containing 50 μg/mL of kanamycin at 37 °C, was used to inoculate 1 L of the same growth medium. The culture was incubated at 37 °C with shaking (230 rpm) until the OD_{600} reached ~0.5. Protein expression was then induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM, and the cells were allowed to grow at 18 °C and 125 rpm for an additional 24 h. The cells were harvested by centrifugation at 4500 ×

g for 15 min and stored at –80 °C until lysis. All purification steps were carried out at 4 °C using Ni-NTA resin according to the manufacturer's protocol with minor modifications. Specifically, the thawed cells (~5 g) were resuspended in the lysis buffer (20 mL) containing 10% (v/v) glycerol and 10 mM imidazole. After incubation with lysozyme (20 mg) for 30 min, the cells were disrupted by sonication using 10 × 10-s pulses with a 30-s cooling pause between each pulse. The resulting lysate was centrifuged at 20,000 × *g* for 20 min, and the supernatant was subjected to Ni-NTA chromatography. Bound protein was eluted using 250 mM imidazole buffer containing 10% glycerol. The collected protein solution was dialyzed against 3 × 1-L of 50 mM Tris·HCl buffer (pH 8) containing 300 mM NaCl and 15% glycerol. The protein solution was then flash-frozen in liquid nitrogen and stored at −80 °C until use. Protein concentration was determined by the Bradford assay using bovine serum albumin as the standard.[149] The yield of *C*-His_6-BexX and *N*-His_6-BexX were approximately 50 mg and 30 mg from 1 L cultures, respectively. The molecular mass and purity (> 90%) of the protein were estimated by SDS-PAGE analysis.

4.2.4 Construction of K110A-*BexX* Mutant.

The site-specific K110A-*bexX* mutant was constructed according to the QuikChange site-directed mutagenesis protocol (Stratagene) using the plasmid, *bexX*/pET24b(+) or *bexX*/pET28b(+) as the template. The PCR primers are shown in Table 4-2. The resulting plasmids K110A-*bexX*/pET24b(+) and K110A-*bexX*/pET28b(+) were used to transform the *E. coli* BL21 star (DE3) strain for protein overexpression.

Table 4-2. Primers used for constructing K110A-*bex*X/pET24b(+)

Primer name	Sequence
K110A-*bexX*-forward	5'-GGCATCGAGATCCTGGCCCTGGACGTGCGC-3'
K110A-*bexX*-reverse	5'-GCGCACGTCCAGGGCCAGGATCTCGATGCC-3'

The codon for the mutated residue is underlined.

4.2.5 Expression and Purification of the *C*-His$_6$-K110A-BexX Mutant.

Overexpression and protein purification of the *C*-His$_6$-K110A-BexX mutant were performed using the plasmid K110A-*bexX*/pET24b(+) according to the same procedure described above for the wild type enzyme. Expression of the *N*-His$_6$-K110A-BexX mutant using the plasmid *bexX*/pET24b(+) was not very successful; and thus, the corresponding protein was not purified. The yield of mutant protein, *C*-His$_6$-K110A-BexX, was approximately 120 mg per 1 L culture. The molecular mass and purity (> 90%) of the protein were estimated by SDS-PAGE analysis.

4.2.6 Construction of Other BexX Mutants.

The site-specific D112A-*bexX,* E194A-*bexX,* D112E-*bexX,* and E194D-*bexX* mutants were similarly constructed according to the QuikChange site-directed mutagenesis protocol (Stratagene) using the plasmids *bexX*/pET24b(+) and *bexX*/pET28b(+) as the template. In addition, a D112E-E194D-*bexX* double mutant was also prepared using D112E-*bexX*/pET24b(+) and D112E-*bexX*/pET28b(+) as templates. The PCR primers used in this work are shown in Table 4-3. The resulting plasmids were used to transform the *E. coli* BL21 star (DE3) strain or the *E. coli* Rosetta 2(DE3) strain for protein overexpression.

Table 4-3. Primers used for constructing the plasmids for various BexX mutants

Primer name	Sequence
D112A-*bexX*-forward	5'-CGAGATCCTGAAGCTGGCGGTGCGCGGCG-3'
D112A-*bexX*-reverse	5'-CGCCGCGCACCGCCAGCTTCAGGATCTCG-3'
E194A-*bexX*-forward	5'-CGGTGGTCGTCGCGGGCGGGATCGGCAGC-3'
E194A-*bexX*-reverse	5'-GCTGCCGATCCCGCCCGCGACGACCACCG-3'
D112E-*bexX*-forward	5'-CGAGATCCTGAAGCTGGAGGTGCGCGGCG-3'
D112E-*bexX*-reverse	5'-CGCCGCGCACCTCCAGCTTCAGGATCTCG-3'
E194D-*bexX*-forward	5'-CGGTGGTCGTCGACGGCGGGATCGGCAGC-3'
E194D-*bexX*-reverse	5'-GCTGCCGATCCCGCCGTCGACGACCACCG-3'

The codon for the mutated residue is underlined.

4.2.7. Expression and Purification of the Other BexX Mutants.

The overexpression and protein purification of the various BexX mutants were performed using the mutated plasmids *bexX*/pET24b(+) according to the same procedure described above for the wild type enzyme except for the presence of 50 μg/mL chloramphenicol in the overnight culture (10 mL) when the *E. coli* Rosetta 2(DE3) strain was used. The majority of the mutant constructs did not express well or the expressed proteins resulted in formation of inclusion bodies. In these cases they were not further purified. One exception was the *C*-His_6-D112A-BexX mutant, which was partially purified.

4.2.8 ESI-MS Analysis of *C*-His_6-BexX.

The purified *C*-His_6-BexX protein stock solution was added to 50 mM Tris·HCl buffer (pH 7.5) containing 10 mM $MgCl_2$ (final concentration of protein was 13 μM) and incubated at 30 °C for 2 h. The resulting solution (200 μL) was cooled on ice. To this solution was added 20 μL of cold $NaBH_4$ solution in methanol (530 mM; final concentration of $NaBH_4$ was 50 mM). The mixture was incubated on ice for 1 h. The

excess salts and glycerol were removed at 4 °C using an Amicon ultrafiltration unit through a YM-10 membrane. The resulting protein solution was subjected to ESI-MS analysis.

4.2.9 ESI-MS Analysis of *C*-His$_6$-BexX with Various Sugars.

Similarly, the purified *C*-His$_6$-BexX protein (13 μM final) was mixed with 5 mM of either D-glucose 6-phosphate (**4-2**), D-fructose 6-phosphate (**4-7**), D-glucose (**4-33**), or 1-deoxy-D-xylose 5-phosphate (DXP, **4-4**) in 50 mM Tris·HCl buffer (pH 7.5) containing 10 mM $MgCl_2$. The resulting solution (200 μL) was incubated at 30 °C for < 2 h and cooled on ice. A shorter incubation time (0.5 h) was used for **4-33**- and **4-4**-containing mixtures because protein started to precipitate beyond this time period. An aliquot of 20 μL of cold $NaBH_4$ solution in methanol (530 mM; final concentration of $NaBH_4$ was 50 mM) was added to each reaction, and the resulting mixture was incubated on ice for 1 h. The excess salts and glycerol were removed at 4 °C using Amicon ultrafiltration units through YM-10 membranes. The resulting protein solutions were subjected to ESI-MS analysis.

4.2.10 ESI-MS Analysis of the *C*-His$_6$-K110A-BexX Mutant.

Similar ESI-MS analysis was performed for the *C*-His$_6$-K110A-BexX mutant. The purified mutant protein (30 μM final) was incubated with or without 5 mM of D-glucose 6-phosphate (**4-2**) in 50 mM Tris·HCl buffer (pH 7.5) containing 10 mM $MgCl_2$. To prevent protein precipitation, the mixture (400 μL each) was incubated at room temperature for 10 min (similar conditions as for the wild-type enzyme caused significant precipitation of the K110A mutant). An aliquot of 40 μL of a cold solution of $NaBH_4$ in methanol (530 mM; final concentration of $NaBH_4$ was 50 mM) was added to each reaction, and the resulting mixture was incubated on ice for 1 h. The excess salts and

glycerol were removed by using an Amicon ultrafiltration unit through a YM-10 membrane at 4 °C. The resulting protein solution was subjected to ESI-MS analysis.

4.2.11 Protein Sequence Alignment and Active Site Prediction.

The protein sequence of BexX was aligned with ThiG from *Bacillus subtilis* using Vector NTI Advance 10.1.1. The sequence of ThiG was obtained from the National Center for Biotechnology Information (NCBI) database. (Gene ID: 936422).[171]

4.2.12 Trypsin Digestion and LC-MS/MS Analysis of the Protein−Substrate Complex.

The purified *C*-His_6-BexX protein (13 μM final) was incubated with or without 5 mM of D-glucose 6-phosphate (**4-2**) in 50 mM Tris·HCl buffer (pH 7.5) containing 10 mM $MgCl_2$ at 30 °C for 1 h. An aliquot of 20 μL of cold $NaBH_4$ solution in methanol (530 mM; final concentration of $NaBH_4$ was 50 mM) was added to 200 μL of each protein solution, and the resulting mixture was incubated on ice for 1 h. The excess salts and glycerol were removed by using an Amicon ultrafiltration unit through a YM-10 membrane at 4 °C. An aqueous dithiothreitol solution (14 μL, final concentration 10 mM), acetonitrile (15 μL), and trypsin stock solution (5 μL containing 2 μg of trypsin) were sequentially added to each of the protein solutions and the resulting mixtures were incubated at 37 °C overnight. The reactions were then frozen and stored at –20 °C prior to analysis. Frozen samples were submitted to the Mass Spectrometry core facility and subjected to LC-ESI-MS/MS analysis on a nanoACQUITY UPLC (Waters) in conjunction with a Q-TOF premier mass spectrometer (Waters).

4.2.13 Preparation of the 2,4-Dinitrophenylhydrazine Derivative of the BexX–Substrate Complex.

The trypsin digests of *C*-His_6-BexX–D-glucose-6-phosphate were prepared according to the procedure described above. The digest mixture (40 μL) was added to a 2,4-dinitrophenylhydrazine (DNPH, **4-36**) solution in DMSO (160 μL, final concentration of DNPH was 80 mM). The resulting mixture was acidified to pH ~ 4 by adding 8 μL of 10% trifluoroacetic acid (TFA) and incubated at 37 °C overnight. The reaction solution was then subjected to LC-ESI-MS analysis as described above.

4.2.14 Preparation of the Hydroxylamine Derivative of the BexX–Substrate Complex.

The purified *C*-His_6-BexX protein (75 μM final concentration) was incubated with 5 mM of D-glucose 6-phosphate (**4-2**) in 50 mM Tris·HCl buffer (pH 7.5) containing 10 mM $MgCl_2$ at 30 °C for 0.5 h. An aqueous solution of hydroxylamine (NH_2OH) (10 μL, final concentration 20 mM) was added to the protein mixture (90 μL) and the resulting mixture was incubated at room temperature for 1 h. After cooling on ice, the excess salts and glycerol were removed by using an Amicon ultrafiltration unit through a YM-10 membrane at 4 °C. The resulting protein solution was then subjected to ESI-MS analysis.

4.2.15 Investigation of Bisulfide, Thioacetate, and Glutathione as the Direct Sulfur Donor.

Bisulfide and Thioacetate

The purified *C*-His_6-BexX protein (13 μM final concentration) was incubated with 5 mM D-glucose 6-phosphate (**4-2**) in 100 mM Tris·HCl buffer (pH 7.5) containing 10 mM $MgCl_2$ at 30 °C for 0.5 h. The buffer of the solution (400 μL) was then exchanged with 100 mM Tris·HCl buffer (pH 7.5) containing 20 μM of D-glucose 6-phosphate (**4-2**)

by using an Amicon ultrafiltration unit equipped with a YM-10 membrane at 4 °C. To this solution was added dithiothreitol (DTT, 1 mM) and sodium bisulfide (**4-25**, 50 mM) or potassium thioacetate (**4-39**, 50 mM). The resulting mixture (~400 μL) was incubated at 30 °C for 1 h and then cooled on ice. A 40 μL aliquot of cold $NaBH_4$ solution in methanol (530 mM; final concentration of $NaBH_4$ was 50 mM) was added to each reaction, and the resulting mixture was incubated on ice for 2 h. The excess salts and glycerol were removed by using an Amicon ultrafiltration unit through a YM-10 membrane at 4 °C. The resulting protein solution was subjected to ESI-MS analysis.

Glutathione

For glutathione (GSH, **4-40**), a mixture of the purified *C*-His_6-BexX and *N*-His_6-BexX proteins (100 μM each) was incubated with 5 mM of D-glucose 6-phosphate (**4-2**) in 100 mM Tris·HCl buffer (pH 7.5) at 4 °C for 1 h. To this solution (90 μL) was added 10 μL of **4-40** (0.2 M; final concentration of **4-40** was 20 mM), and the mixture was incubated at room temperature for 1 h. This mixture was analyzed by ESI-MS without further purification.

4.2.16 H/D exchange experiment in D_2O.

The purified *C*-His_6-BexX and *N*-His_6-BexX proteins (100 μM each final) were mixed with 50 mM $NH_4 \cdot HCO_3$ buffer (pH 8.0) containing 1 mM D-glucose 6-phosphate (**4-2**) prepared in D_2O. To reduce the amount of H_2O from the protein stock solution, the solution (200 μL) was concentrated to half of the original volume (100 μL) using an Amicon ultrafiltration unit equipped with a YM-10 membrane. The same volume of the buffer containing **4-2** in D_2O was added to the concentrated protein solution. This cycle was repeated 10 times (theoretical ratio of D_2O / H_2O in the solution was over 1000 in the final solution). The resulting protein solution was incubated at room temperature for 12 h.

The protein was removed by YM-10 membrane filtration, and the filtrate containing D-glucose 6-phosphate was lyophilized, redissolved in H_2O, and subjected to ESI-MS analysis.

4.2.17 Pull-Down Assay.

Preparation of the A. Oreintalis Cell Free Extracts

Cell free extracts from *A. orinetalis* subsp. *vinearia* BA-07585 were prepared by a previously reported method with some modification.[172] Spores of *A. orientalis* were inoculated into 10 mL of International Streptomyces Project (ISP) medium 2 and grown in a rotary incubator at 30 °C and 220 rpm for 2 days. The resultant seed culture (3 mL) was transferred to 1 L of ISP-2 medium and grown under the same conditions for 6 days. The growth culture, which had a light red color, was centrifuged at 4,500 × *g* for 10 min at 4 °C, and the cells were washed with 300 mL of 50 mM potassium phosphate buffer (pH 7.5) containing 30% (v/v) glycerol. After another centrifugation, the cells (~40 g) were resuspended in 50 mL of protein-extraction buffer, which is composed of potassium phosphate buffer (50 mM, pH 7.5), 30% (v/v) glycerol, β-mercaptoethanol (2 mM), phenylmethylsulfonyl fluoride (0.1 mM), benzamidine hydrochloride (0.1 mM), and ethylenediaminetetraacetic acid (EDTA, 1 mM). The resulting cell suspension was dropped into liquid N_2 using Pasteur pipette to make small frozen balls and stored at −80 ºC until use. The frozen balls (~5 g) were cryogenically ground to a fine powder by using a Mixer Mill MM 300. The ground cells were gently resuspended in the protein-extraction buffer for 2 h at 4 °C. The mixture was centrifuged at 20,000 × *g* for 20 min at 4 °C, and the supernatant was dialyzed against 2 × 0.5 L of 50 mM potassium phosphate buffer (pH 7.5) and used as the cell fee extracts.

Pull-down assay

The purified *C*-His_6-BexX or *N*-His_6-BexX protein solution (50 μL; final concentration of the protein was 25 μM) was mixed with 580 μL of the cell free extract and 0.1 mL of Ni-NTA suspension pre-washed with the lysis buffer (50 mM phosphate buffer, pH 7.5 containing 10% (v/v) glycerol and 10 mM imidazole) in 50 mM potassium phosphate buffer. The resulting mixture (1 mL) was incubated at 4 ºC for 1 or 4 h. Following centrifugation, the supernatant was discarded. The resin was washed with wash buffer containing 20 mM imidazole and 10% (v/v) glycerol. The bound proteins were eluted with 100 μL of elution buffer containing 250 mM imidazole and 10% (v/v) glycerol. The eluted solution was subjected to SDS-PAGE and transblotted onto a polyvinylidene fluoride (PVDF) membrane. A protein band shown at < 10 kDa on the membrane was further subjected to *N*-terminal protein sequencing analysis.

4.2.18 Whole Genome Sequencing and Analysis.

Spores of *A. orientalis* were inoculated into 10 mL of soybean-casein digest (TSB) medium and grown in a rotary incubator at 30 °C and 250 rpm for 2 days. The resultant seed culture (4 mL) was transferred to 100 mL of TSB medium and grown under the same conditions for 2 days. The growth culture (25 mL) was centrifuged at 5,000 × *g* for 20 min at 4 °C, and the cells were washed with 25 mL of 10 mM EDTA. After another centrifugation, the cells were stored at −80 ºC until use. The cells were resuspended in 5 mL of 1 mM EDTA, and the suspension was divided into 1.5 mL-tubes (0.4 mL each). The genomic DNA was extracted using the PureLink Genomic DNA Mini Kit according to the manufacturer's instructions. The resultant DNA solution (0.5 μg/μL, 50 μL from each tube) was frozen and sent to the Genomics Research Center of Academia Sinica (Taiwan). The *de novo* sequencing was performed using Roche 454 and the open reading frames were predicted by using Glimmer (Gene Locator and

Interpolated Markov ModelER) version 3.0.[173] Homologous protein sequences were identified in the NCBI database using the basic local alignment search tool (BLAST).

4.2.19 Cloning of *ThiS, MoaD, CysO, MoaD2, MoeZ, Cd1, Cd2, Cd3, Cd4, and Cd5*.

The *thiS (orf13974)*, *moaD (orf13839), cysO (orf06461), maD2 (orf10102), moeZ (orf02110), cd1 (orf10706), cd2 (orf11099), cd3 (orf14916), cd4 (orf04763), and cd5 (orf09299)* genes were PCR-amplified from *A. oreintalis* genomic DNA using primers with engineered *Nde*I and *Hind*III, or *Xho*I for *cd1*, restriction sites. The sequences of the primers are shown in Table 4-4. The PCR-amplified gene fragments were purified, digested with *Nde*I and *Hind*III, or *Nde*I and XhoI for *cd1*, and ligated into pET28b (+) vector that was also digested with the same enzymes. The resultant plasmids, *thiS*/pET28b(+), *moaD*/pET28b(+), *cysO*/pET28b(+), *moaD2*/pET28b(+), *moeZ*/pET28b(+), *cd1*/pET28b(+), *cd2*/pET28b(+), *cd3*/pET28b(+), *cd4*/pET28b(+), and *cd5*/pET28b(+) were sequenced using the T7 or T7 terminal universal primer and used to transform *E. coli* BL21 star (DE3) strain for protein overexpression.

Table 4-4. Primers used for constructing *thiS*/pET28b(+), *moaD*/pET28b(+), *cysO*/pET28b(+), *moaD2*/pET28b(+), *moeZ*/pET28b(+), *cd1*/pET28b(+), *cd2*/pET28b(+), *cd3*/pET28b(+), *cd4*/pET28b(+), and *cd5*/pET28b(+)

Primer name	Sequence
thiS for pET28b(+)-forward	5'-TCTATAA**CATATG**GAGATCAAGCTCAACGG-3'
thiS for pET28b(+)-reverse	5'-TTAC**AAGCTT***CA*GCCTCCCTGGACGG-3'
moaD for pET28b(+)-forward	5'-TCTATAA**CATATG**ACCACCCTGACCATCG-3'
moaD for pET28b(+)-reverse	5'-TTAC**AAGCTT**AGAGAAGCC*TCA*GCCGC-3'
cysO for pET28b(+)-forward	5'-TCTATAA**CATATG**GCCGTGACCGTCTCC-3'
cysO for pET28b(+)-reverse	5'-TTAC**AAGCTT***CA*GCCACCGGCCACG-3'
moaD2 for pET28b(+)-forward	5'-TCTATAA**CATATG**ACCGTGCGGATCACC-3'
moaD2 for pET28b(+)-reverse	5'-TTAC**AAGCTT***CA*GCCTCCCGCGACC-3'
moeZ for pET28b(+)-forward	5'-TCTATAA**CATATG**GACGCGAACCGCGCGC-3'
moeZ for pET28b(+)-reverse	5'-TTAC**AAGCTT***CA*GTAGGTCGGCAGGCTCGG-3'
cd1 for pET28b(+)-forward	5'-TCTATAA**CATATG**TACCTGGACTCCGGAGC-3'
cd1 for pET28b(+)-reverse	5'-TCAA**CTCGAG**GT*TCA*CTGCGCCACC-3'
cd2 for pET28b(+)-forward	5'-TCTATAA**CATATG**ACTCTCGCCATCGACC-3'
cd2 for pET28b(+)-reverse	5'-TTAC**AAGCTT***CA*GCAGAACGGCGGG-3'
cd3 for pET28b(+)-forward	5'-TCTATAA**CATATG**AGTTCCGCGCTCGGC-3'
cd3 for pET28b(+)-reverse	5'-TTAC**AAGCTT***CA*GCGGGGGAGCGACG-3'
cd4 for pET28b(+)-forward	5'-TCTATAA**CATATG**ACCTATCTCGACCACGCG-3'
cd4 for pET28b(+)-reverse	5‘-TTAC**AAGCTT***TA*CACCTCTTGCTTCTGGG-3'
cd5 for pET28b(+)-forward	5'-TCTATAA**CATATG**GCGTTCGACGTCGC-3'
cd5 for pET28b(+)-reverse	5'-TTAC**AAGCTT***TA*GCGGAGTTCTTCCAGC-3'

The engineered restriction sites are shown in bold, the start codon is shown in bold and also underlined, and the stop codon is italic.

4.2.20 Expression and Purification of *N*-His$_6$-ThiS, *N*-His$_6$-MoaD, *N*-His$_6$-CysO, *N*-His$_6$-MoeZ, *N*-His$_6$-CD1, *N*-His$_6$-CD2, *N*-His$_6$-CD3, *N*-His$_6$-CD4, and *N*-His$_6$-CD5.

The overexpression and protein purification of the *N*-His$_6$-tagged sulfur carrier proteins (ThiS, MoaD, CysO, and MoaD2), MoeZ, and cysteine desulfurases (CD1, CD2, CD3, CD4, and CD5) were performed using the plasmids constructed above according to the same procedure described for *N*-His$_6$-BexX enzyme. The yields of *N*-His$_6$-BexX, *N*-His$_6$-CysO, and *N*-His$_6$-MoaD2 were 30 mg, 15 mg, and 20 mg, respectively, from 1 L

culture. Protein expression of *N*-His_6-MoaD was much poorer than the other sulfur carrier proteins, and the yield was < 5 mg from 1 L culture. The yield of *N*-His_6-MoeZ was 40 mg from 1 L culture. The yields of cysteine desulfurases, *N*-His_6-CD1, *N*-His_6-CD2, *N*-His_6-CD3, and *N*-His_6-CD4 were 50 mg, 60 mg, 30 mg, and 100 mg, respectively, from 1 L culture. Protein *N*-His_6-CD5 was expressed as inclusion bodies and, thus, was not purified. The molecular mass and purity (> 90% except *N*-His_6-MoaD) of the proteins were estimated by SDS-PAGE analyses.

4.2.21 MoeZ-Catalyzed Activation of ThiS and its ESI-MS Analyses.

ESI-MS Analysis of ThiS

The purified *N*-His_6-ThiS protein (90 μM) in 5 mM Tris·HCl buffer (pH 8.0) was subjected to ESI-MS analysis.

ESI-MS Analysis of ThiS with MoeZ

The purified *N*-His_6-ThiS (90 μM) was incubated with *N*-His_6-MoeZ (80 μM) and ATP (5 mM) in 100 mM Tris·HCl buffer (pH 8.0) containing $MgCl_2$ (5 mM) at 30 ºC for 0.5 h. The resultant solution was subjected to ESI-MS analysis. Additionally, to aliquotes of the above solution was added sodium sulfide (NaSH, 10 mM) or sodium thiosulfate ($Na_2S_2O_3$, 10 mM) and the resulting mixture was subjected to ESI-MS analysis after incubation at 30 ºC for 0.5 h. Finally, a control reaction containing only *N*-His_6-ThiS (90 μM) and NaSH (5 mM) in 100 mM Tris·HCl buffer (pH 8.0) was similarly incubated and subjected to ESI-MS analysis.

4.2.22 Investigation of Sulfur Transfer from ThiS-Thiocarboxylate to the BexX−sugar complex by ESI-MS Analysis.

The purified protein mixture containing *C*-His_6-BexX (100 μM), *N*-His_6-ThiS (90 μM), and *N*-His_6-MoeZ (7 μM) was incubated with ATP (5 mM) and NaSH (10 mM) in

50 mM Tris·HCl buffer (pH 8.0) containing $MgCl_2$ (5 mM) at 30 ºC for 0.5 h. The resultant solution was subjected to ESI-MS analysis.

4.2.23 MoeZ-Catalyzed Activation of MoaD, CysO, and MoaD2 and their ESI-MS Analyses.

ESI-MS Analysis of MoaD, CysO and MoaD2

The purified *N*-His_6-MoaD (50 μM), *N*-His_6-CysO (60 μM), or *N*-His_6-MoaD2 (90 μM) in 100 mM Tris·HCl buffer (pH 8.0) was subjected to ESI-MS analysis.

ESI-MS Analysis of MoaD, CysO and MoaD2 with MoeZ

The purified *N*-His_6-MoaD (50 μM), *N*-His_6-CysO (60 μM), or *N*-His_6-MoaD2 (90 μM) was incubated with *N*-His_6-MoeZ (7 μM), ATP (5 mM), and NaSH (10 mM) in 50 mM Tris·HCl buffer (pH 8.0) containing $MgCl_2$ (5 mM) at 30 ºC for 0.5 h. The resultant solution was subjected to ESI-MS anayslsis.

4.2.24 Investigation of Sulfur Transfer from CysO- or MoaD2-Thiocarboxylate to the BexX−sugar complex by ESI-MS Analysis.

The purified protein *N*-His_6-CysO (60 μM) or, *N*-His_6-MoaD2 (90 μM) was incubated with *C*-His_6-BexX (100 μM), *N*-His_6-MoeZ (7 μM), ATP (5 mM), and NaSH (10 mM) in 50 mM Tris·HCl buffer (pH 8.0) containing $MgCl_2$ (5 mM) at 30 ºC for 0.5 h. The resultant solution was subjected to ESI-MS analysis.

4.2.25 Coupled Enzymatic Assay to Detect the Formation of AMP Upon Activation of the Sulfur Carrier Proteins.

AMP formation upon activation of sulfur carrier proteins was detected spectrophotometrically using commercially available adenylate kinase (AK), pyruvate kinase (PK), and lactate dehydrogenase (LDH) in the presence of ATP, phosphoenolpyruvate (PEP), and nicotinamide adenine dinucleotide (reduced form,

NADH). The typical reaction solution (99 μL) using *N*-His_6-MoaD2 (9 μM) contained *C*-His_6-BexX (10 μM), *N*-His_6-MoeZ (3 μM), G6P (5 mM), NaSH (5 mM), PEP (0.3 mM), NADH (0.1 mM), AK (2 units), PK (0.3 units), and LDH (0.5 units), in 50 mM Tris·HCl buffer (pH 8) containing $MgCl_2$ (5 mM). Absorbance at 340 nm of the resulting solution was monitored at room temperature using UV−vis spectrophotometer. The reaction was initiated by adding 1 μL of 100 mM ATP stock solution (final concentration of ATP was 1 mM). Similarly, *N*-His_6-MoaD2 was substituted in a series of analogous experiments with *N*-His_6-ThiS (10 or 20 μM), *N*-His_6-CysO (6 μM), or *N*-His_6-MoaD (9 μM) and the absorbance at 340 nm was monitored. Various control reactions were also performed by omitting one of the coupling enzymes or substrates.

4.2.26 Enzymatic 2-Thiosugar Formation and its Detection.

BexX Reaction Using Bisulfide

The typical BexX reaction mixture (50 μL) contained *C*-His_6-BexX (100 μM), sulfur carrier proteins (*N*-His_6-CysO (30 μM), *N*-His_6-MoaD2 (45 μM), or *N*-His_6-ThiS (45 μM)), *N*-His_6-MoeZ (15 μM), ATP (2 mM), G6P (2 mM), NaSH (5 mM) in 50 mM $NH_4{\cdot}HCO_3$ buffer (pH 8.0) containing $MgCl_2$ (5 mM). The resulting reaction mixture was incubated at room temperature for 8 h and stored at −20 ºC for < 5 days until use.

Direct mBBr Derivatization and HPLC and ESI-MS Analysis

To the reaction mixture (10 μL) prepared above was then added 10 μL of 10 mM mBBr (**4-66**) methanol solution (final concentration of mBBr was 5 mM) and incubated at room temperature for 5 min. The mixture was centrifuged at 16,000 × *g* for 5 min to remove the precipitant, and 10 μL of the supernatant was transferred to a new tube. The solution was evaporated *in vacuo* using a Speedvac SC100 (Savant). The resulting compound mixture was redissolved in 100 μL of 50 mM $NH_4{\cdot}HCO_3$ buffer (pH 8.0) and

subjected to HPLC analysis using a Dionex CarboPac PA1 analytical column (4 × 250 mm). The sample was eluted with a gradient of water (solvent A) and 1 M NH_4OAc (solvent B). The gradient was run from 5 to 15% B over 5 min, 15−30% B over 15 min, 30−100% B over 7 min with a 5 min wash at 100% B, and 100−5% B over 3 min, followed by re-equilibration at 5% B for 5 min. The flow rate was 1 mL/min, and the detector was set at 260 nm. The peak corresponding to the enzymatic reaction product was isolated and subjected to ESI-MS analysis.

CIP Treatment, mBBr Derivatization, and HPLC and ESI-MS Analysis

The BexX-reaction mixture prepared above was also analyzed by HPLC using a C_{18} column along with an authentic standard. The reaction mixture stored at −20 ºC was thawed and treated with calf intestinal alkaline phosphatase (CIP) (0.2 μL, 2units) and incubated at 37 ºC for 1 h. The precipitant that appeared during the incubation was removed by centrifugation at 16,000 × *g* for 2 min, and 2 μL of a 100 mM monobromobimane (mBBr, **4-66**) methanol solution was added to the reaction solution (final concentration of mBBr was 5 mM). The resulting mixture was incubated at room temperature for 5 min, and the supernatant (5 μL) was diluted with deionized water (95 μL) prior to HPLC analysis using an analytical C_{18} column (4 × 250 mm). The sample (20 μL) was eluted with a gradient of water (solvent A) and 80% acetonitrile (solvent B). The gradient was run from 5 to 30% B over 15 min, 30−80% B over 5 min, 80−5% B over 5 min, followed by re-equilibration at 5% B for 10 min. The flow rate was 1 mL/min, and the detector was set at 260 nm. The 2-thio-D-glucose-bimane (**4-69**) standard (0.1 mM) was prepared from the chemically synthesized 2-thio-D-glucose (**4-68**, see Chapter 3 for the detail of synthesis) incubated with mBBr (**4-66**) at room

temperature for 5 min. The peak corresponding to the enzymatic reaction product was also isolated and subjected to ESI-MS analysis.

BexX Reaction Using Thiosulfate

BexX reaction was also prepared using sodium thiosulfate (NaS_2O_3, 5 mM) and dithiothreitol (DTT, 5 mM) instead of NaSH. The resultant reaction mixture was similarly analyzed by HPLC after derivatization.

BexX Reaction Using Cysteine Desulfurases

BexX reaction was also prepared using a cysteine desulfrase (*N*-His_6-CD1 (40 μM), *N*-His_6-CD2 (40 μM), *N*-His_6-CD3 (40 μM), or *N*-His_6-CD4 (40 μM)), L-cysteine (**4-26**, 3 mM), pyridoxal-5'-phosphate (PLP, 0.1 mM), and DTT (3 mM) instead of NaSH. The resultant reaction mixture was derivatized with mBBr (**4-66**) and subjected to HPLC analysis as described above.

4.2.27 Protein Sequence Analyses of Sulfur Carrier Proteins.

The protein sequences of sulfur carrier proteins, ubiquitin, and ubiquitin-like proteins were aligned using Vector NTI Advance 10.1.1. The protein sequences were obtained from the National Center for Biotechnology Information (NCBI) database. The four sulfur carrier proteins from *A. orientalis* and a PdtH homologue from *Stigmatella aurantiaca* DW4/3-1 (ZP_01459250) were submitted to I-TASSER online for the structural prediction.[174,175]

4.3 RESULTS AND DISCUSSION

4.3.1 Identification of the BexX Substrate.

In order to verify the predicted function of the *bexX* gene, the corresponding protein, BexX, was heterologously overexpressed in *E. coli* and purified as a *C*-terminal His_6-tagged protein (Figure 4-6, lane 2). Interestingly, electrospray ionization mass spectrometry (ESI-MS) of the isolated protein demonstrated two main peaks. The dominant peak corresponded to the expected His_6-tagged BexX (calculated, 28488 Da; observed, 28485 Da), whereas the minor peak showed a mass increase of ~240 Da relative to the dominant peak. Upon treatment with $NaBH_4$, the intensity of the latter peak of the reduced enzyme became more apparent (observed, 28729 Da) (Figure 4-7A). The observed mass increase of 244 Da is consistent with formation of a reduced dehydration adduct between BexX and a hexose monophosphate substrate (e.g., reduced **4-8** or **4-14**, Figure 4-3). It was thus hypothesized that a fraction of the purified BexX contains a reaction intermediate covalently bound in substrate−enzyme complex as an imine that can be stabilized by borohydride reduction.

To test this hypothesis, the purified BexX was incubated with the putative substrate D-glucose 6-phosphate (**4-2**) and treated with $NaBH_4$ prior to MS analysis. As shown in Figure 4-7B, the addition of **4-2** resulted in complete conversion of the peak at 28485 Da to the new peak at 28729 Da. In contrast, no significant change in the peak distribution was discernible when D-fructose 6-phosphate (**4-7**), D-glucose (**4-33**), or DXP (**4-4**) was used instead of **4-2** (Figure 4-8). Thus, the observed covalent modification of BexX is clearly D-glucose 6-phosphate (**4-2**)-specific. These results indicate that **4-2** likely serves as the substrate for BexX and that formation of the imine intermediate **4-8** (Figure 4-3, pathway A) with an active-site lysine residue may be one of the initiating steps of the BexX reaction.

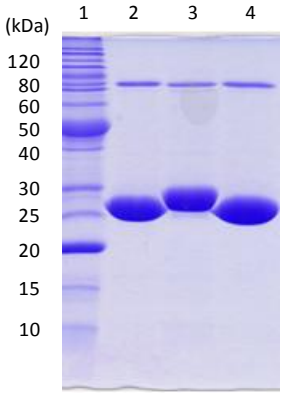


Figure 4-6. SDS-PAGE gel of purified *C*-His_6-tagged BexX (268 aa, 28.5 kDa, lane 2), *N*-His_6-tagged BexX (275 aa, 29.1 kDa, lane 3), and *C*-His_6-tagged K110A BexX (268 aa, 28.4 kDa, lane 4)

The molecular weight marks are 220, 160, 120, 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, and 10 kDa (top to bottom, lane 1). The protein band of ~ 80 kDa is a minor impurity from endogenous *E. coli* proteins.

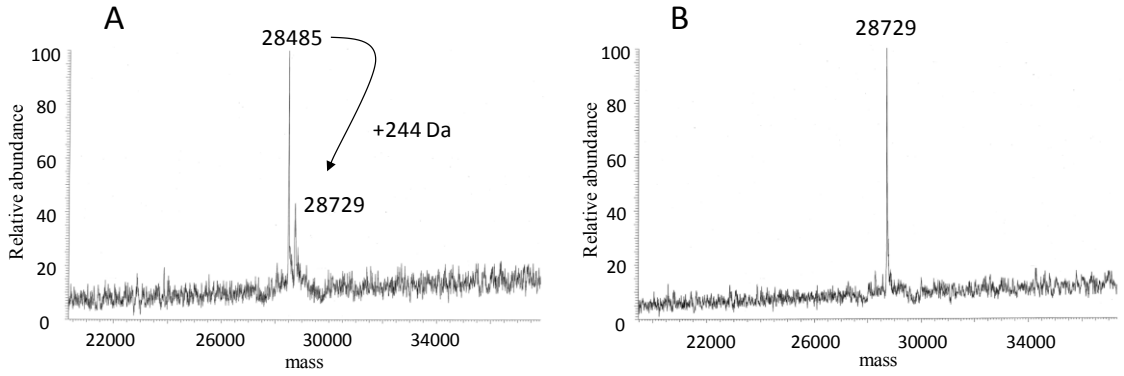


Figure 4-7. ESI-MS analysis of BexX with glucose 6-phosphate

(A) Deconvoluted ESI-MS of *C*-His_6-tagged BexX treated with $NaBH_4$. (B) ESI-MS of *C*-His_6-tagged BexX incubated with D-glucose 6-phosphate prior to $NaBH_4$ treatment. The calculated molecular weights of *C*-His_6-BexX (268 aa) and *C*-His_6-BexX−D-glucose-6-phosphate (reduced) are 28488 and 28732 Da, respectively.

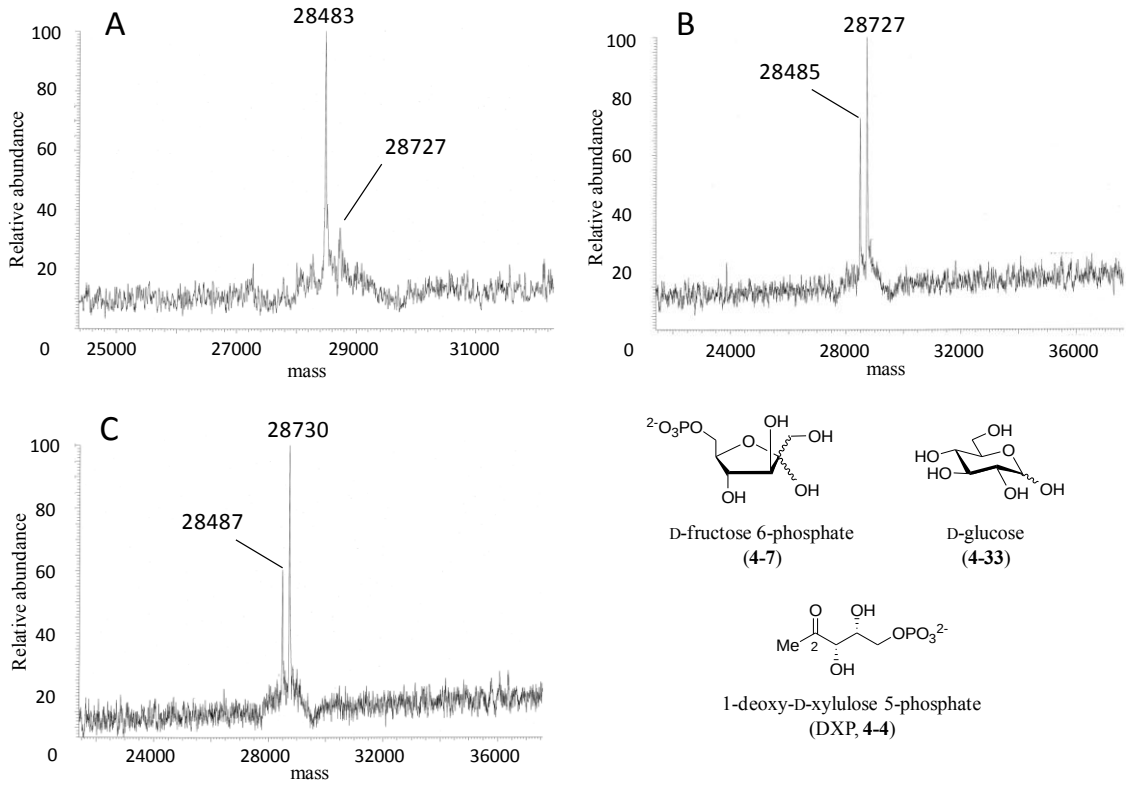


Figure 4-8. ESI-MS analysis of BexX with various sugars.

Deconvoluted ESI-MS of *C*-His_6-tagged BexX and each sugar mixture: (A) D-fructose 6-phosphate (**4-7**), (B) D-glucose (**4-33**), and (C) DXP (**4-4**). Calculated molecular weights of *C*-His_6-BexX (268 aa) and *C*-His_6-BexX−D-glucose-6-phosphate (reduced) are 28488, and 28732, respectively. Calculated molecular weights of D-glucose-6-phosphate (**4-2**), **4-7**, **4-33**, and **4-4** are 260, 260, 180 and 214, respectively. Since part of the as-isolated BexX contains varied amount of bound **4-2** (such as **4-11**, Figure 4-3), which upon reduction, gives m/z signals at 28727 or 28730 (molecular weight of the reduced BexX−D-glucose-6-phosphate complex).

4.3.2 Identification of the Modification Site.

Sequence alignment revealed that the catalytic lysine residue (Lys96) in ThiG from *B. subtilis* has a counterpart, Lys110, in BexX (Figure 4-9).[142,171] To investigate whether Lys110 plays a direct role in BexX catalysis, the K110A mutant was constructed, heterologously expressed in *E. coli*, and purified as a *C*-terminal His_6-tagged

protein (Figure 4-6, lane 4). ESI-MS of the purified mutant protein exhibited only one peak, the mass of which matched the calculated molecular mass for the unmodified K110A mutant (calculated, 28431 Da; observed, 28430 Da). No change in molecular mass was noted even after treatment with D-glucose 6-phosphate (**4-2**) and $NaBH_4$ (Figure 4-10). This finding strongly implicated Lys110 as the residue responsible for covalent adduct formation with D-glucose 6-phosphate (**4-2**).

To verify that the modification occurs at Lys110, the wild-type BexX−D-glucose 6-phosphate complex was reduced by $NaBH_4$ and then subjected to proteolysis with trypsin. The tryptic-digested peptide fragments were analyzed by LC-MS/MS, and the peptide fragment D102–R114 [calculated, m/z 1714.89 ($[M + H]^+$); observed, m/z 1714.92] contained a D-glucose 6-phosphate (**4-2**) adduct (**4-34**, Figure 4-11). Further analysis of the observed b and y ions revealed the attachment of the hexose phosphate at K110. These results unambiguously demonstrate that the sugar modification indeed occurs at a specific lysine residue (i.e., Lys110), consistent with the proposed mechanism shown in Figure 4-3.

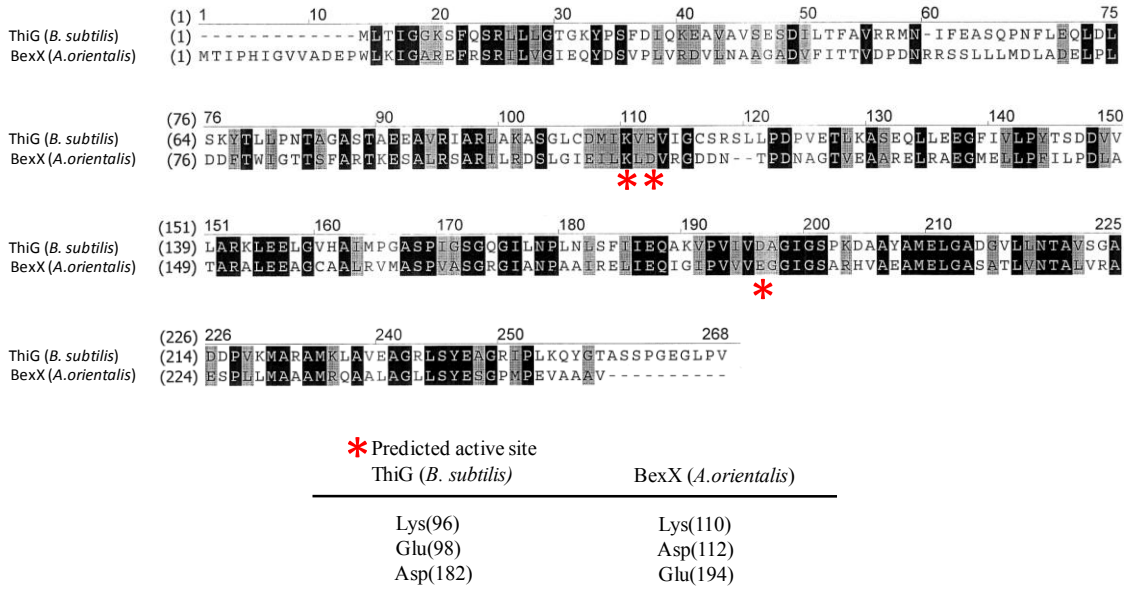


Figure 4-9. Protein sequence alignment of BexX and ThiG from *Bacillus subtilis*.

Identity and similarity of the sequences are 37% and 59%, respectively.

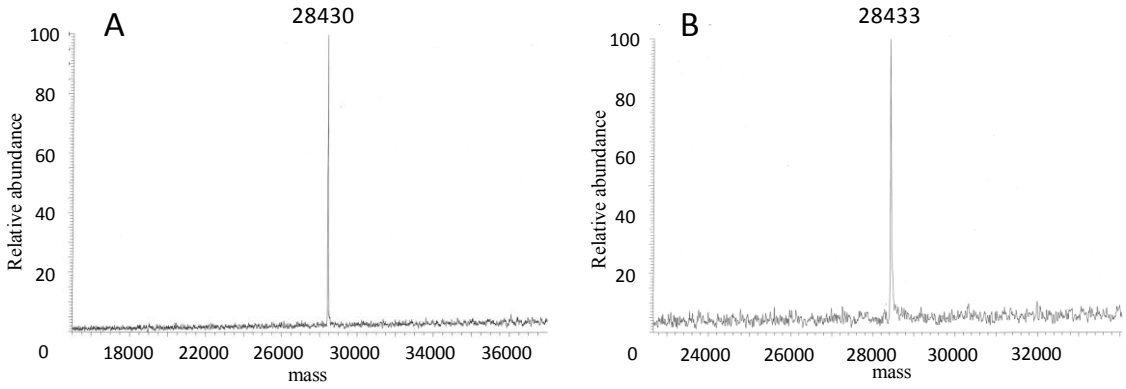


Figure 4-10. Deconvoluted ESI-MS of K110A BexX mutant.

K110A BexX mutant was incubated with (A) or without (B) D-glucose 6-phosphate prior to the $NaBH_4$ treatment. Calculated molecular weight of *C*-His_6-tagged K110A BexX (268 aa) is 28431.

A (*N*-term) (M)TIPHIGVVADEPWLK / IGAR / EFR / SR / ILVGIEQYDSVPLVR / DVLNAAGADVFITTVDPDNR / R / SSLLLMDLADELPLDDFTWIGTTSFAR / TK / ESALR / SAR / ILR / DSLGIEILK* LDVR / GDDNTPDNAGTVEAAR / ELR / AEGMELLPFILPDLATAR / ALEEAGCAALR / VMASPVASGRGIANPAAIR / ELIEQIGIPVVVEGGIGSAR / HVAEAMELGASATLVNTALVR / AESPLLMAAAMR / QAALAGLLSYESGPMPEVAAAVK / LAAALEHHHHHH (*C*-term)

B

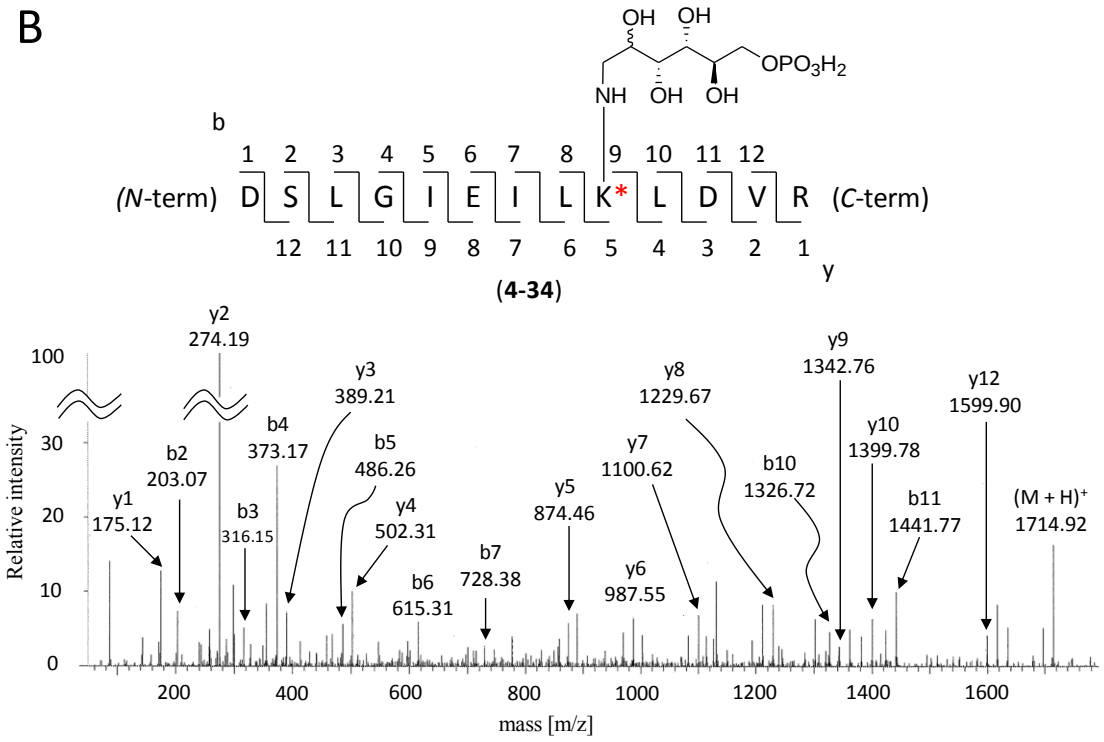


Figure 4-11. Trypsin treatment and LC-MS/MS analysis.

(A) The predicted digestion pattern of *C*-His$_6$-tagged BexX. Lysine residues are shown in bold. The predicted active site lysine (K110) is shown with an asterisk. The detected peptide fragments by LC-ESI-MS are underlined. (B) LC-MS/MS of the identified peptide (D102-R114) with a D-glucose 6-phosphate (**4-2**) coupled to the active-site lysine (K110) (**4-34**). The calculated m/z value for **4-34** was 1714.89 ($[M + H]^+$), and the observed value was 1714.92.

4.3.3 Identification of the Ketone Intermediate.

It was noted during the studies of the trypsin-digested BexX fragments that the BexX−D-glucose 6-phosphate adduct could be detected even without the addition of any

reducing reagent ($NaBH_4$) to the reaction mixture [calculated, m/z 856.94 ($[M + 2H]^{2+}$, z = 2); observed, m/z 856.84] (Figure 4-12B). This observation was surprising because it is unlikely that the proposed imine intermediate **4-8**, without prior reduction, could survive the trypsin digestion as well as the conditions used for the LC-MS/MS analysis. This prompted us to reconsider the chemical nature of the trapped intermediate in the active site of BexX. The initially formed iminium intermediate **4-8** is predicted to equilibrate with **4-9** and **4-10** (Figure 4-3). A likely candidate for the stable covalent BexX−D-glucose-6-phosphate adduct is compound **4-10**, which does not readily hydrolyze in solution. Accordingly, the tryptic fragment **4-34** detected in the LC-MS/MS analysis could result from hydride reduction of the imine moiety in **4-8** and/or the 2-keto group in **4-10** [calculated, m/z 857.95 ($[M + 2H]^{2+}$, z = 2); observed, m/z 857.84] (Figure 4-12A). However, without prior reduction, the modified D102-R114 fragment shown in Figure 4-12B likely corresponds to **4-35** (Figure 4-13).

To test this hypothesis, the trypsin-digested BexX−D-glucose 6-phosphate sample was treated with a carbonyl-specific labeling reagent, 2,4-dinitrophenylhydrazine (DNPH, **4-36**), and the resulting mixture was analyzed by LC-MS (Figure 4-14A). The observed m/z 946.96 ion ($[M + 2H]^{2+}$, z = 2) is consistent with the modified peptide D102-R114 coupled with D-glucose 6-phosphate (**4-2**) and DNPH (**4-37**; Figure 4-13). In a separate experiment, the undigested BexX−D-glucose-6-phosphate complex was treated with another carbonyl-reactive reagent, NH_2OH, and analyzed by ESI-MS. The observed signal at 28742 (Figure 4-14B) matches the predicted molecular mass of a BexX−D-glucose 6-phosphate−NH_2OH imine adduct (**4-38**; Figure 4-14). These results strongly support the presence of keto intermediate **4-10** in the protein−substrate complex.

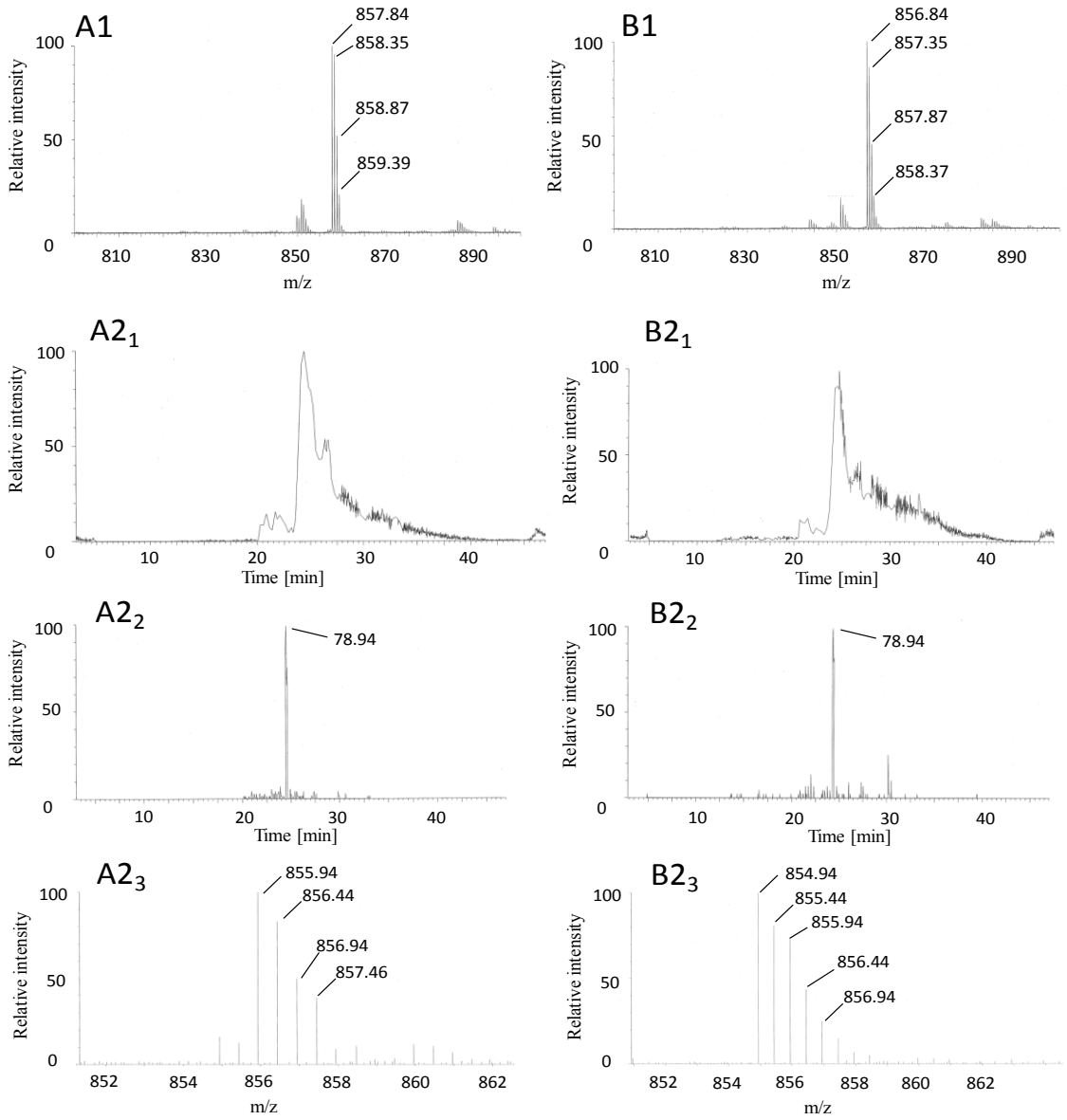


Figure 4-12. LC-MS of trypsin-digested BexX-D-glucose-6-phosphate.

(A) with and (B) without $NaBH_4$ treatment (**1**; positive-, **2**; negative-mode analysis). The calculated m/z values for **4-34** and its unreduced form **4-35** are 1; 857.95 ($[M + 2H]^{2+}$, z = 2), 2; 855.93 ($M - 2H^+$, z = 2) and 1; 856.94 ($[M + 2H]^{2+}$, z = 2), 2; 854.92 ($M - 2H^+$, z = 2), respectively. **A2$_1$** and **B2$_1$** show the target peak intensity chromatogram (854.9 ± 0.5 Da and 856.0 ± 2 Da, respectively, negative-mode). **A2$_2$** and **B2$_2$** show the extracted LC-MS/MS chromatogram containing the target peak, 79 ± 1 Da, which corresponds to PO^{3-} (negative-mode). Calculated monoisotopic mass of PO^{3-} is 78.96.

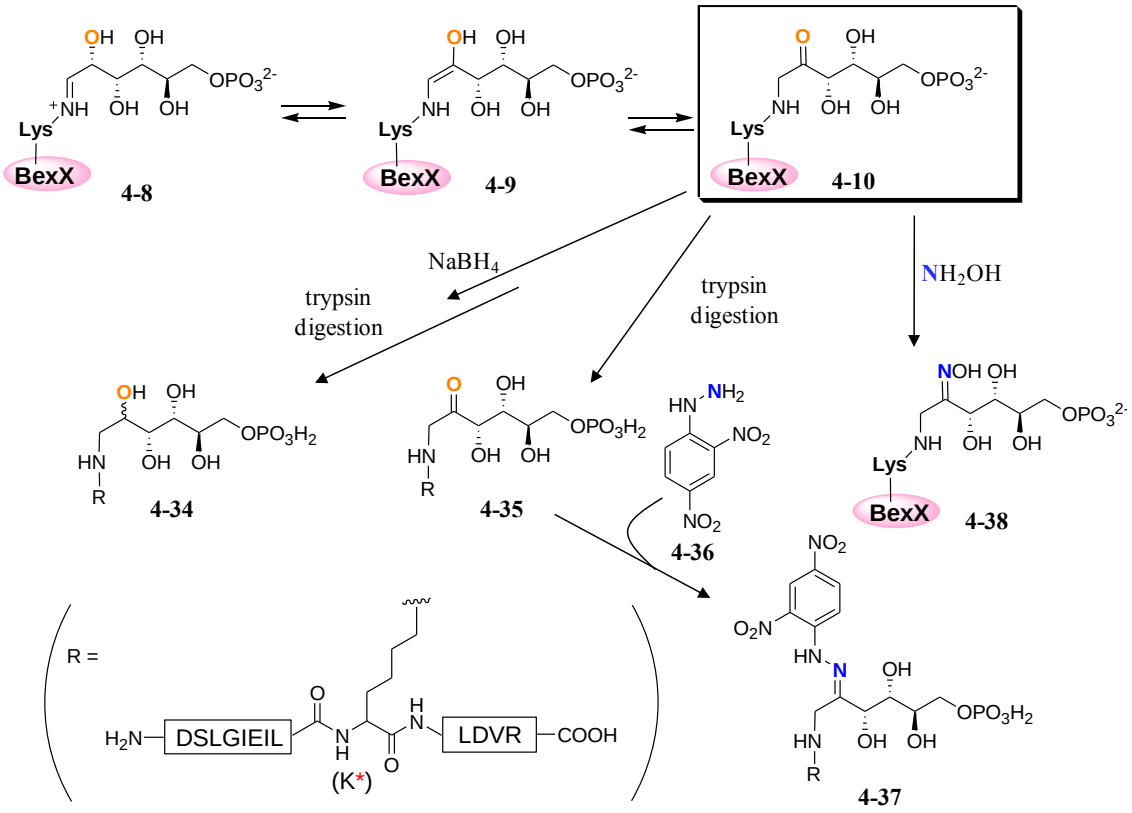


Figure 4-13. Identification of the ketone intermediate (4-10).

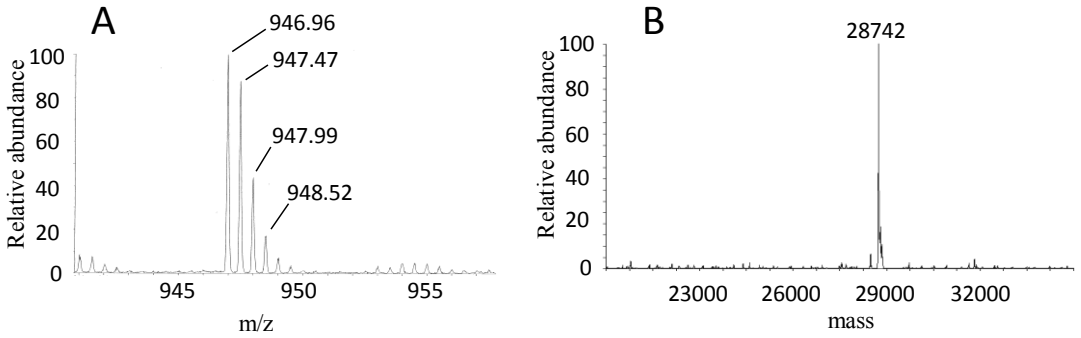


Figure 4-14. C-2-Keto intermediate trapping experiments.

(A) LC-MS of the modified peptide **4-37** derived from trypsin digested BexX−D-glucose-6-phosphate after DNPH (**4-36**) treatment. The calculated m/z value of **4-37** is 946.95 ($[M + 2H]^{2+}$, Z =2). (B) Deconvoluted ESI-MS of the *C*-His_6-BexX−D-glucose-6-phosphate ketone intermediate trapped using NH_2OH. The calculated molecular weight of **4-38** is 28745 Da.

4.3.4 Bisulfide is not the Direct Sulfur Donor.

In mechanism A of Figure 4-3, ketone intermediate (**4-10**) accepts sulfur from a donor such as bisulfide (**4-25**), a protein persulfide (**4-28** or **4-30**), or a protein thiocarboxylate (**4-32**; Figure 4-5). In the case of thiamin biosynthesis in *B. subtilis* (Figure 4-2), the physiological sulfur donor, ThiS-thiocarboxylate, could be substituted with bisulfide (1 mM) *in vitro*.[142] Bisulfide (5 mM) has also been utilized as a substitute for the sulfur carrier protein, HcyS-thiocarboxylate, from *Wolinella succinogenes* in the methionin biosynthetic pathway.[85] Such a substitution is also possible with ThiI persulfide from *E. coli*, which acts as a sulfur donor in the biosynthesis of 4-thiouridine.[63] Based on these reports, we hypothesized that bisulfide may also act as a sulfur donor for the BexX-bound sugar substrate although it may not necessarily be the physiological sulfur donor.

To test this hypothesis, a solution containing the *C*-His_6-BexX−D-glucose 6-phosphate ketone complex (**4-10,** 13 μM) and free D-glucose-6-phosphate (**4-2,** 20 μM) was incubated with an excess amount of bisulfide (**4-25**, 50 mM) followed by treatment with $NaBH_4$ prior to MS analysis. If incubation with bisulfide leads to incorporation of a sulfur atom at C-2, then the molecular weight of the protein should be altered (either free enzyme or enzyme−thiosugar complex should be found). However, the observed MS signal from this mixture was identical to that of the reduced **4-10** (Figure 4-15A). Two other sulfur-containing compounds, thioacetate (**4-39**, 50 mM) and glutathione (GSH, **4-40**, 20 mM), were also tested, but similarly failed to produce significant changes in the molecular weight of the protein–substrate complex. (Figure 4-15B, C). Since NH_2OH could react with **4-10** and produced BexX−D-glucose-6-phosphate−NH_2OH imine adduct (**4-38**; Figure 4-14), the smaller bisulfide (**4-25**) should be able to access the C-2 ketone intermediate (**4-11**; Figure 4-3). Thus, the results shown here suggest that sulfur transfer

is likely protein-mediated and sulfur incorporation in the presence of a sulfur delivery protein is essential to push the intermediate **4-11** forward to release the expected thiosugar product (**4-3**). Otherwise, the transiently formed **4-11** by reacting with bisulfide may readily goes back to the more stable ketone intermediate (**4-10**).

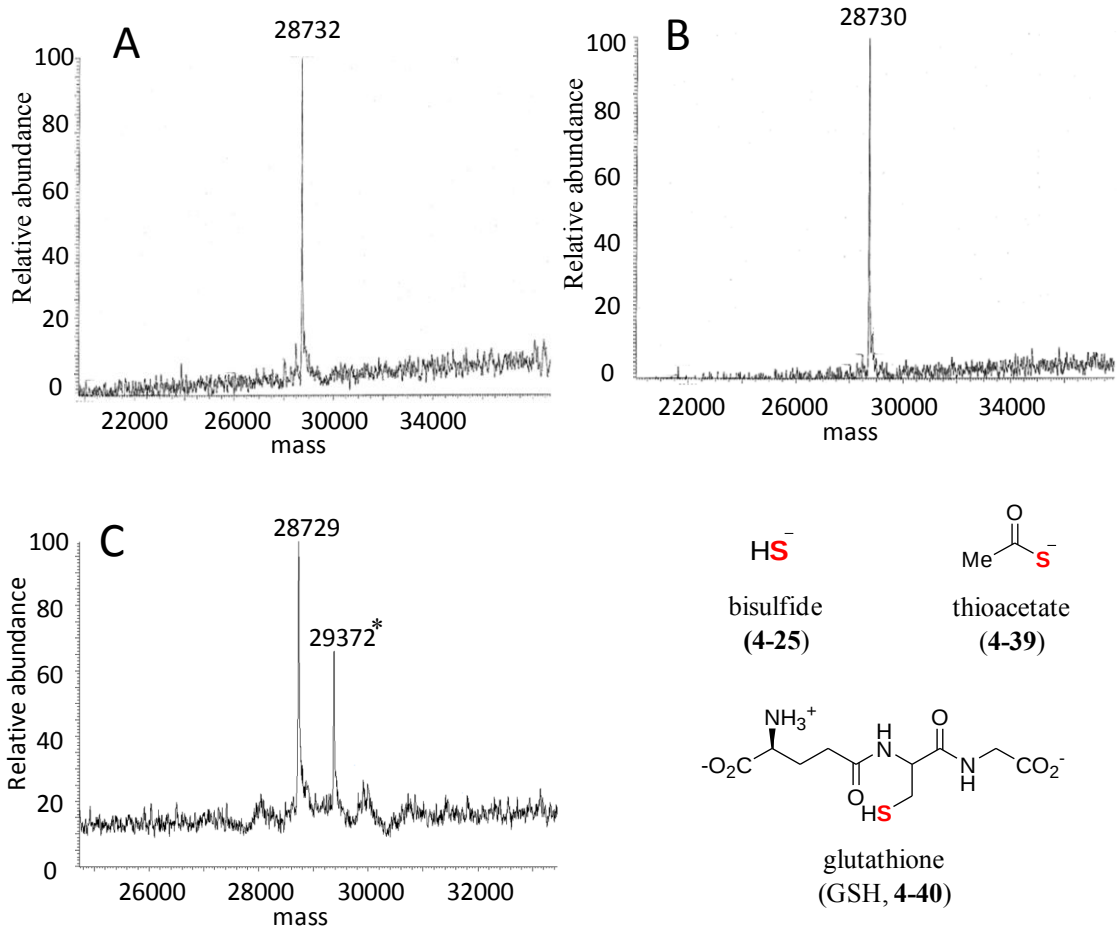


Figure 4-15. C-2-Keto Intermediate with sulfur-containing small molecules.

Deconvoluted ESI-MS of the *C*-His$_6$-BexX−D-glucose 6-phosphate ketone intermediate (**4-10**) with (A) bisulfide (**4-25**), (B) thioacetate (**4-39**), or (C) GSH (**4-40**). *The reaction mixture of (C) contains both *C*-His$_6$- and *N*-His$_6$-enzymes. The calculated molecular weights of *C*-His$_6$- and *N*-His$_6$-**4-10** (reduced) are 28730, and 29374 Da, respectively.

4.3.5 Expected H/D exchange at the C-2 position.

The apparent inertness of the *C*-His_6-BexX−D-glucose-6-phosphate ketone intermediate (**4-10**) in the presence of high concentration of bisulfide (**4-25**) was surprising. It suggests that sulfur incorporation must be coupled with the enzyme-mediated product sugar release to drive the reaction to completion. The proposed mechanism of the thiosugar release from the predicted thioketone intermediate (Figure 4-3; **4-12** → **4-13** → **4-3**) is the reverse reaction of substrate binding to form the corresponding ketone (**4-2** → **4-8** → **4-9** → **4-10**). This process involves a series of isomerization reactions, which are supposed to be in equilibrium, and thus is reversible. In fact, similar reversibility had been demonstrated in ThiG-catalyzed reaction in which H/D exchange at C-3 position of DXP (**4-4**) was observed when the incubation was carried out in a solution containing D_2O.[168]

To test the reversibility of the early isomerization steps of the BexX-catalyzed reaction, D-glucose-6-phosphate (**4-2**, 1 mM) was incubated with *C*-His_6-BexX and *N*-His_6-BexX (0.1 mM each) in a buffer solution prepared with D_2O. The expected H/D exchange at C-2 position is shown in Figure 4-16. The solution was then subjected to MS analysis, however, the spectrum displayed only the signal of **4-2** (ESI^- calculated for $C_6H_{12}O_9P^-$ [M – H^+] 259.0, observed 259.2), and no deuterium incorporation was observed. This result is not expected, but appears to be consistent with the aforementioned observation that BexX can not produce thiosugar product in the presence of bisulfide. It suggests that the enzyme-ketosubstrate adduct (**4-10**) is a thermodynamically stable species whos formation from 4-8 is practically irreversible. This is in contrast to the analogous ThiG-catalyzed reaction.

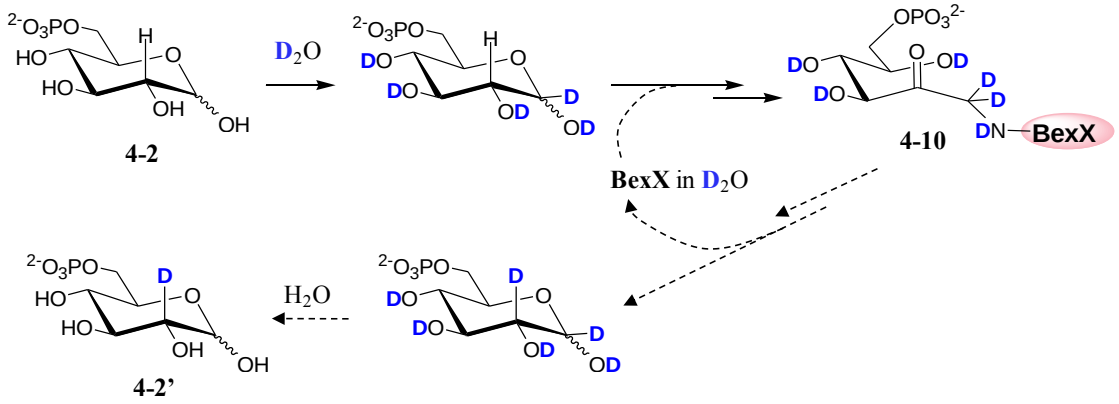


Figure 4-16. Expected H/D exchange at C-2 of D-glucose 6-phosphate.

4.3.6 Efforts to Find a Sulfur Transfer Protein by Pull-Down Assay.

Since BexX shares sequence similarity to ThiG, ThiS or a ThiS-like protein carrying a thiocarboxylate group at *C*-terminal is likely a candidate of sulfur donor in the thiosugar biosynthetic reaction. ThiG is known to form a stable complex with ThiS such that they can be copurified.[143]. If BexX also employs a protein partner as sulfur donor with which it forms a similarly stable complex, then such a complex formation could potentially be exploited to isolate the corresponding sulfur carrier protein using a pull-down assay.

In this attempt, both *C*-His_6-BexX and *N*-His_6-BexX were used as the bait proteins for the pull-down assay. The proteins were incubated with cell free extracts prepared from *A. orientalis* and purified using Ni-NTA resin. The proteins forming complexes with BexX could then be coeluted with BexX from the Ni-NTA column. Indeed, SDS-PAGE analysis showed multiple protein bands when *N*-His_6-BexX was used as the bait (Figure 4-17). The protein of particular interest are in the small molecular weight region (<10 kDa), because the ThiS protein is small with a molecular weight of ~7 kDa. Thus, *N*-terminal protein sequencing of the intense band cut from this region of

SDS-PAGE was performed. However, the sequence obtained corresponded to the His-tag (*N*-GSSHHHHHH). This result indicates that BexX was degraded during the incubation with the cell free extracts even though the mixture contained protease inhibitors. Other bands shown in the gel are also likely the degradation products of proteins, and thus, were not subjected to further analysis.

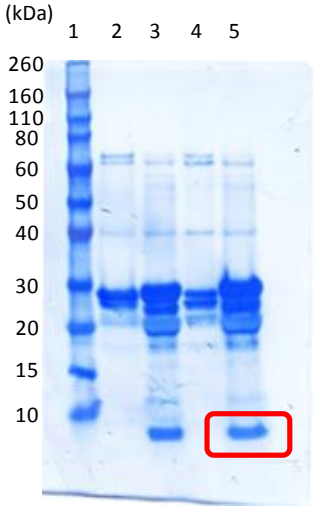


Figure 4-17. SDS-PAGE (PVDF membrane) analysis of the pull-down assay

lane1; molecular weight marks, lane 2; *C*-His_6-tagged BexX (268 aa, 28.5 kDa) as the bait, 4 h incubation, lane 2; *N*-His_6-tagged BexX (275 aa, 29.1 kDa) as the bait, 4 h incubation, lane 3; *C*-His_6-tagged BexX as the bait, 1 h incubation, lane 4; *N*-His_6-tagged BexX as the bait, 1 h incubation. The protein band subjected to *N*-terminal protein sequencing is marked with a red box.

4.3.7 Whole Genome Sequencing and its Analysis.

Whole Genome Sequencing

Since the pull-down approach was unsuccessful to identify the sulfur donor, it became necessary to individually express and purify the candidate proteins including

ThiS as well as other sulfur carrier proteins, cysteine desulfurases, and rhodanese-type proteins, and then test their ability to incorporate sulfur into the BexX-D-glucose 6-phosphate complex *in vitro*. For this purpose, whole genome sequencing of *A. orientalis* was performed.

The genome analysis of *A. orientalis* identified 9210 coding open reading frames from approximately 9.8 Mb genomic DNA (71% GC content), which is slightly larger than the reported genome sizes of other *Actinomycetes* (8.7 Mb for *Streptomyces coelicolor* A3(2),[176] 9.0 Mb for *Streptomyces avermitilis*,[156] 8.2 Mb for *Saccharopolyspora erythraea* NRRL23338,[157] and 8.5 Mb for *Streptomyces griseus* IFO13350[177]).

Genes Related to Sulfur Carrier Proteins

We first focused on sulfur carrier proteins such as ThiS. Sulfur carrier proteins are small proteins (65-100 aa) containing a characteristic GG sequence at their C-terminus. They are known to be involved in several biosynthetic pathways of sulfur-containing biomolecules such as thiamin, and are also known to be evolutionally related to eukaryotic ubiquitin or ubiquitin-like proteins (UBLs).[73,74] The BLAST analyses using the obtained *A. orientalis* genomic DNA led to four genes encoding putative sulfur carrier proteins. One of these is found in the putative thiamin biosynthetic gene cluster (*thiS*, Table 4-4), and two others are in the putative molybdopterin biosynthetic gene cluster (*moaD*, Table 4-5), or the putative cysteine biosynthetic gene cluster (*cysO*, Table 4-6). The fourth gene is similar to *moaD* from the molybdopterin biosynthetic pathway, and is thus named as *moaD2*. The *moaD2* gene is found in the absence of nearby genes related to any particular biosynthetic pathway (Table 4-6). The proposed biosynthetic pathways

for each product in *A. orientalis* are shown in Figure 4-18−20.[65,68,84,178,179] The gene organization near the *moaD2* gene is shown in Figure 4-21.

Surprisingly, unlike the well studied thiamin biosynthetic pathway of *E. coli* or *B subtilis*,[65] the thiamin biosynthetic gene cluster of *A. orientalis* does not contain *thiF*, the product of which is necessary to activate the *C*-terminus of ThiS to acquire a thiocarboxylate group. Similary, gene encoding the MoaD-activating enzyme, MoeB, is not present in the molybdopterin biosynthetic gene cluster of *A. orientalis.* Instead, the BLAST analysis identified a gene encoding an enzyme, MoeZ, composed of an *N*-terminal ThiF_MoeB_HesA_family domain and a *C*-terminal rhodanese homology domain (Figure 4-22). Notably, MoeZ from *Mycobacterium tuberculosis* can be utilized to activate CysO in the cysteine biosynthetic pathway.[84] Thus, we proposed that MoeZ may be the endogenous enzyme responsible for the activation of all sulfur carrier proteins found in *A. orientalis* (Figure 4-18−20).

Table 4-5. BLAST analysis of the genes near *thiS* in *A. orientalis* genome

gene number	proposed function in thiamin biosynthesis	protein homologue and origin	identity / similarity (%)	protein accession number
13945	ThiC	thiamine biosynthesis protein ThiC [*Amycolatopsis mediterranei* U32]	96 / 98	YP_003770668
13947	ThiD	phosphomethylpyrimidine kinase [*Amycolatopsis mediterranei* U32]	93 / 96	YP_003770669
13950		No significant similarity found		
13952		No significant similarity found		
13954		transcriptional regulator, HxlR family protein [*Streptomyces bingchenggensis* BCW-1]	60 / 74	ADI04334
13957		alcohol dehydrogenase zinc-binding domain-containing protein [*Stackebrandtia nassauensis* DSM 44728]	63 / 75	YP_003511248
13960	ThiD	phosphomethylpyrimidine kinase [*Amycolatopsis mediterranei* U32]	96 / 97	YP_003770671
13961		lipoprotein [*Amycolatopsis mediterranei* U32]	95 / 97	YP_003770670
13964		voltage-gated sodium channel [*Saccharopolyspora erythraea* NRRL 2338]	57 / 73	ZP_06564131
13966		NmrA family protein [*Catenulispora acidiphila* DSM 44928]	78 / 89	YP_003115375
13967		transcriptional regulatory protein [*Streptomyces* sp. Mg1]	85 / 93	ZP_04996370
13971		MarR family transcriptional regulator [*Amycolatopsis mediterranei* U32]	97 / 99	YP_003770672
13972	ThiG	thiamine biosynthesis protein ThiG [*Amycolatopsis mediterranei* U32]	97 / 99	YP_003770673
13974	ThiS	thiamine biosynthesis protein ThiS [*Amycolatopsis mediterranei* U32]	88 / 93	YP_003770674
13975	ThiO	glycine oxidase [*Amycolatopsis mediterranei* U32]	89 / 93	YP_003770675
13976	ThiE	tthiamine-phosphate pyrophosphorylase [*Amycolatopsis mediterranei* U32]	89 / 93	YP_003770676

The ThiS homologue is shown in red. The other putative thiamin biosynthetic proteins are shown in blue.

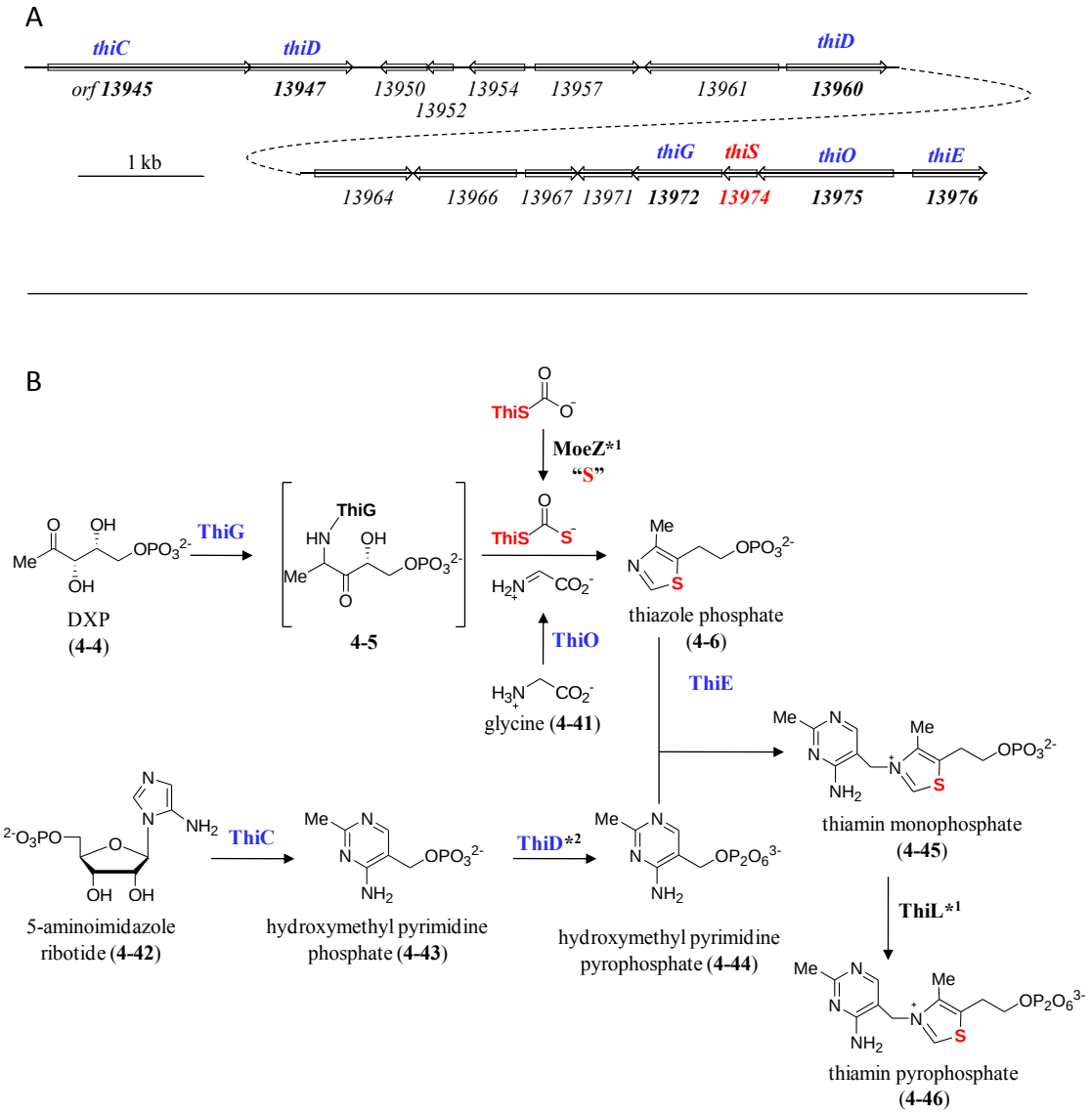


Figure 4-18. Proposed thiamin biosynthetic pathway in *A. orientalis.*

(A) Organization of the putative thiamin biosynthetic gene cluster found in the *A. orientalis* genome. (B) Proposed thiamin biosynthetic pathway in *A. orientalis.* *[1] The genes encoding MoeZ and ThiL are not found in the gene cluster. The gene encoding the ThiS-activating enzyme, ThiF, is also absent in the genome. *[2] Two genes encoding proteins homologous to ThiD are found in the gene cluster.

Table 4-6. BLAST analysis of the genes near *moaD* in *A. orientalis* genome

gene number	proposed function in molybdopterin biosynthesis	protein homologue and origin	identity / similarity (%)	protein accession number
13829		ABC transport system ATP-binding protein [*Amycolatopsis mediterranei* U32]	91 / 94	YP_003770609
13830		L- permease component of ABC-type molybdate transport system [*Amycolatopsis mediterranei* U32]	65 / 98	YP_003770610
13831		eriplasmic substrate-binding component of ABC-type molybdate transport system [*Amycolatopsis mediterranei* U32]	84 / 89	YP_003770611
13832		molybdate ABC molybdate transporter substrate-binding protein [*Streptomyces* sp. AA4]	98 / 99	ZP_07282910
13835		MarR family transcriptional regulator [*Amycolatopsis mediterranei* U32]	89 / 92	YP_003770613
13837	MoaA	molybdenum cofactor biosynthesis protein A [*Amycolatopsis mediterranei* U32]	93 / 98	YP_003770614
13839	MoaD	ThiS/MoaD family protein [*Amycolatopsis mediterranei* U32]	82 / 90	YP_003770615
13840		secreted protein [*Amycolatopsis mediterranei* U32]	86 / 90	YP_003770616
13842	MogA-MoaE	molybdenum cofactor biosynthesis protein B [*Amycolatopsis mediterranei* U32]	92 / 97	YP_003770618
13843	MoaC	molybdenum cofactor biosynthesis protein C [*Amycolatopsis mediterranei* U32]	96 / 97	YP_003770619
13846		prevent-host-death family protein [*Frankia* symbiont of Datisca glomerata]	48 / 70	ZP_06476571
13848		PilT protein domain protein [*Frankia* symbiont of Datisca glomerata]	50 / 68	ZP_06473004
13851		malic enzyme [*Amycolatopsis mediterranei* U32]	97 / 99	YP_003770620
13854		hypothetical protein AMED_8524 [*Amycolatopsis mediterranei* U32]	88 / 99	YP_003770621
13856		hypothetical protein AMED_8525 [*Amycolatopsis mediterranei* U32]	64 / 97	YP_003770622
13859		hypothetical protein AMED_8526 [*Amycolatopsis mediterranei* U32]	65 / 78	YP_003770623

The MoaD homologue is shown in red. The other putative molybdopterin biosynthetic proteins are shown in blue.

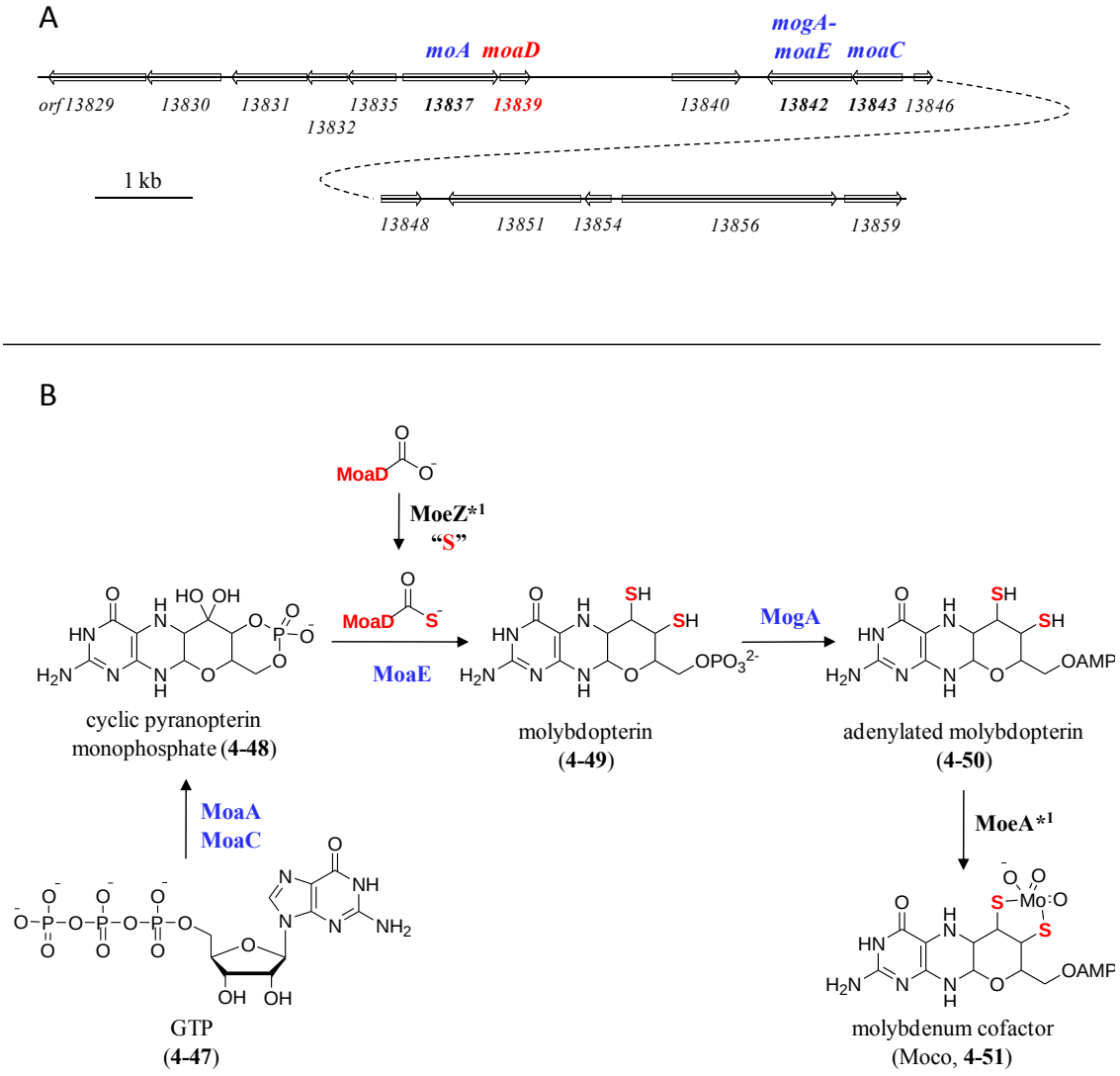


Figure 4-19. Proposed molybdenum cofactor biosynthetic pathway in *A. orientalis*.

(A) Organization of the putative molybdopterin biosynthetic gene cluster found in the *A. orientalis* genome. (B) Proposed molybdenum cofactor biosynthetic pathway in *A. orientalis*. *[1] The genes encoding MoeZ and MoeA were not found in the gene cluster. The gene encoding the MoaD-activating enzyme, MoeB, is also absent in the genome.

Table 4-7. BLAST analysis of the genes near *cysO* in *A. orientalis* genome

gene number	proposed function in cysteine biosynthesis	protein homologue and origin	identity / similarity (%)	protein accession number
06453		transcriptional regulatory protein [*Streptomyces* sp. AA4]	92 / 96	ZP_07282073
06454		L-aminopeptidase/D-esterase DmpA [*Amycolatopsis mediterranei* U32]	94 / 96	YP_003769825
06456		No significant similarity found		
06457	Mec$^+$	Mov34/MPN/PAD-1 family protein [*Amycolatopsis mediterranei* U32]	65 / 98	YP_003769823
06461	CysO	ThiS/MoaD family protein [*Amycolatopsis mediterranei* U32]	97 / 100	YP_003769822
06462	CysM	cysteine synthase [*Amycolatopsis mediterranei* U32]	98 / 99	YP_003769821
06464		aspartate racemase [*Amycolatopsis mediterranei* U32]	89 / 93	YP_003769819
06465		hypothetical protein RHA1_ro01207 [*Rhodococcus jostii* RHA1]	50 / 64	YP_701191
06467		hypothetical protein AMED_7707 [*Amycolatopsis mediterranei* U32]	89 / 95	YP_003769817

The CysO homologue is shown in red. The other putative cysteine biosynthetic proteins are shown in blue.

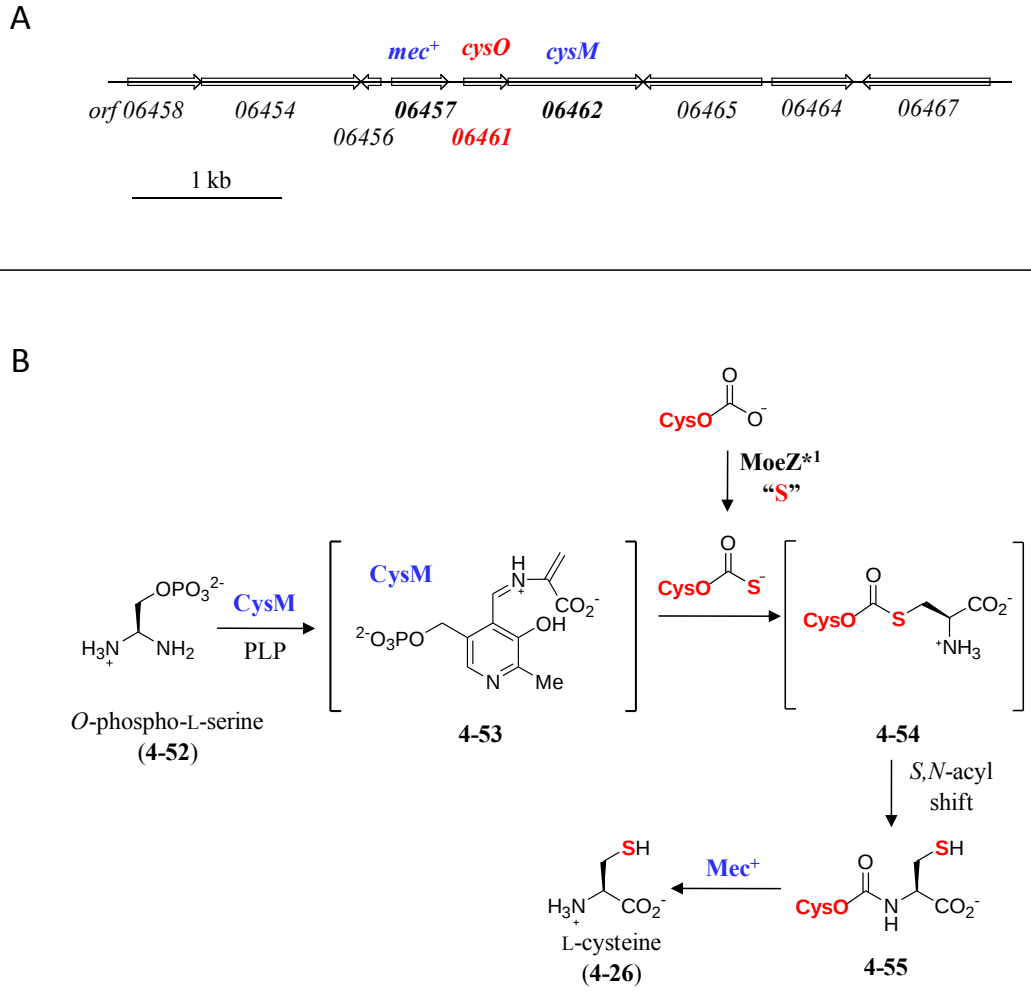


Figure 4-20. Proposed cysteine biosynthetic pathway in *A. orientalis*.

(A) Organization of the putative cysteine biosynthetic gene cluster found in the *A. orientalis* genome. (B) Proposed cysteine biosynthetic pathway in *A. orientalis*. *[1] The gene encoding MoeZ is not found in the gene cluster.

Table 4-8. BLAST analysis of the genes near *moaD2* in *A. orientalis* genome

gene number	sulfur carrier protein	protein homologue and origin	identity / similarity (%)	protein accession number
10086		enoyl-CoA hydratase [*Amycolatopsis mediterranei* U32]	88 / 93	YP_003764230
10087		No significant similarity found		
10090		cold shock protein CspA [*Amycolatopsis mediterranei* U32]	99 / 99	YP_003767860
10092		hypothetical protein AMED_2018 [*Amycolatopsis mediterranei* U32]	85 / 93	YP_003764228
10093		RNA polymerase ECF-subfamily sigma factor [*Amycolatopsis mediterranei* U32]	91 / 93	YP_003764229
10094		hypothetical protein AMED_2017 [*Amycolatopsis mediterranei* U32]	82 / 90	YP_003764227
10095		XRE family transcriptional regulator [*Amycolatopsis mediterranei* U32]	94 / 100	YP_003764226
10096		NADPH:quinone reductase and related Zn-dependent oxidoreductase [*Amycolatopsis mediterranei* U32]	84 / 89	YP_003764225
10099		hypothetical protein AMED_2012 [*Amycolatopsis mediterranei* U32]	88 / 96	YP_003764222
10100		hypothetical protein AMED_2011 [*Amycolatopsis mediterranei* U32]	90 / 95	YP_003764221
10102	**MoaD2**	ThiS/MoaD family protein [*Amycolatopsis mediterranei* U32]	91 / 94	YP_003764220
10106		MerR family transcriptional regulator [*Amycolatopsis mediterranei* U32]	92 / 98	YP_003764218
10107		intradiol ring-cleavage dioxygenase [*Geodermatophilus obscurus* DSM 43160]	45 / 52	YP_003408862
10108		major facilitator transporter [*Amycolatopsis mediterranei* U32]	85 / 90	YP_003764219
10110		oxidoreductase [*Streptomyces bingchenggensis* BCW-1]	73 / 80	ADI09540
10111		No significant similarity found		
10113		hypothetical protein Mvan_5028 [*Mycobacterium vanbaalenii* PYR-1]	40 / 54	YP_955806
10114		conserved hypothetical protein [*Streptomyces pristinaespiralis* ATCC 25486]	25 / 38	ZP_06913769
10116		hypothetical protein AMED_2005 [*Amycolatopsis mediterranei* U32]	81 / 85	YP_003764215
10117		hypothetical protein AMED_2004 [*Amycolatopsis mediterranei* U32]	86 / 92	YP_003764214
10118		No significant similarity found		

The MoaD homologue is shown in red.

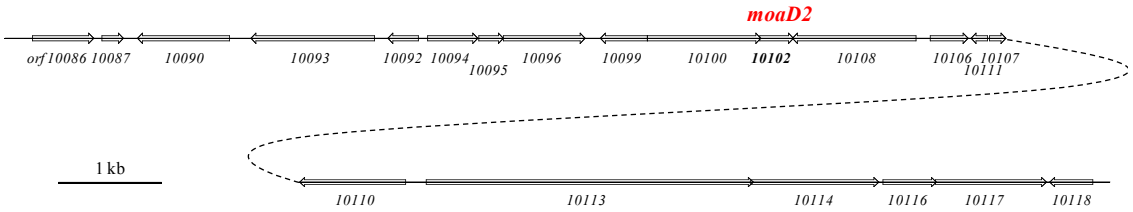


Figure 4-21. Organization near the *moaD*, homologue, *moaD2*, found in the *A. orientalis* genome.

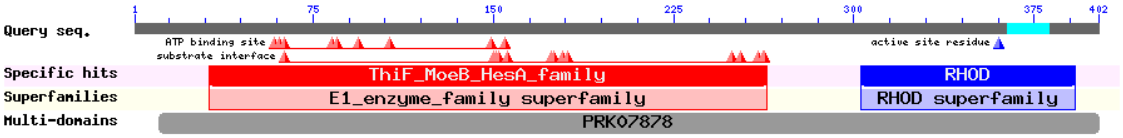


Figure 4-22. Putative conserved domains found in MoeZ.

The conserved domains are predicted by BLAST analysis. The gene number of *moeZ* is *orf* 02110.

Genes Related to Cysteine Desulfurases

Cysteine desulfurases are another important class of sulfur transfer proteins that utilize pyridoxal-5'-phosphate (PLP) and L-cysteine to generate a reactive protein persulfide. The BLAST analyses of the *A. orientalis* genomic DNA led to the identification of five cysteine desulfurase homologues (CD1−5, Table 4-8). This class of enzymes was considered for their potential to participate directly in donating sulfur to the BexX-D-glucose 6-phosphate ketone intermediate (**4-10**, Figure 4-5) or indirectly by serving a source of sulfur for the sulfur carrier proteins mentioned above. In fact, cysteine desulfurases have been, in many cases, proposed as the *in vivo* sulfur source for producing the thiocarboxylate group on many sulfur carrier proteins.[66,75,78-80,86]

Table 4-9. BLAST analysis of cysteine desulfurase homologues found in *A. orientalis.*

gene number	cysteine desulfurase	protein homologue and origin	identity / similarity (%)	protein accession number
10706	CD1	cysteine desulfurase/selenocysteine lyase [*Amycolatopsis mediterranei* U32]	97 / 98	YP_003765029
11099	CD2	cysteine desulfurase [*Amycolatopsis mediterranei* U32]	87 / 91	YP_003765163
14916	CD3	cysteine desulfurase [*Amycolatopsis mediterranei* U32]	88 / 93	YP_003766645
04763	CD4	cysteine desulfurase [*Amycolatopsis mediterranei* U32]	92 / 96	YP_003763873
09299	CD5	cysteine desulfurase [*Amycolatopsis mediterranei* U32]	97 / 99	YP_003762467

4.3.8 Investigation of ThiS as the Sulfur Donor.

MoeZ-Catalyzed ThiS Activation.

A. orientalis ThiS and MoeZ carring an *N*-terminal His_6-tag were heterologously expressed in *E. coli* (see also Figure 4-25), and the MoeZ-catalyzed ThiS-activating reaction (Figure 4-23A) was analyzed by ESI-MS. The purified *N*-His_6-ThiS from Ni affinity chromatography exhibited two mass signals corresponding to the calculated molecular weights of the recombinant enzyme (**4-56**) and its *N*-gluconoylated derivative (**4-60**; Figure 4-23B). Spontaneous modification of the *N*-terminal His_6-tag by D-glucono-1,5-lactone (**4-59**) is a previously reported phenomenon for proteins expressed in *E. coli.*[180] This modification is not expected to affect the enzyme activity of ThiS, because the predicted active site for ThiS is located at the *C*-terminus.

When *N*-His_6-ThiS was incubated with *N*-His_6-MoeZ and ATP in Tris (100 mM) buffer containing 6% glycerol, a signal consistent with the adenylated *N*-His_6-ThiS (**4-57**) was found along with other signals likely derived from nonenzymatic modification of the adenylated enzyme (**4-61** and **4-62**; Figure 4-23C) by the buffer components. In the presence of bisulfide (**4-25**), complete conversion of the enzyme to the corresponding thiocarboxylate (**4-58**; Figure 4-23D) was observed. Control experiments using bisulfide

in the absence of *N*-His_6-MoeZ displays no alteration of the original ThiS signals, consistent with the proposal that thiocarboxylate formation proceeds in a MoeZ-catalyzed adenylation-dependent manner (Figure 4-23E). Together, these data support our prediction that MoeZ is the *in vivo* activator for ThiS in *A. orientalis*.

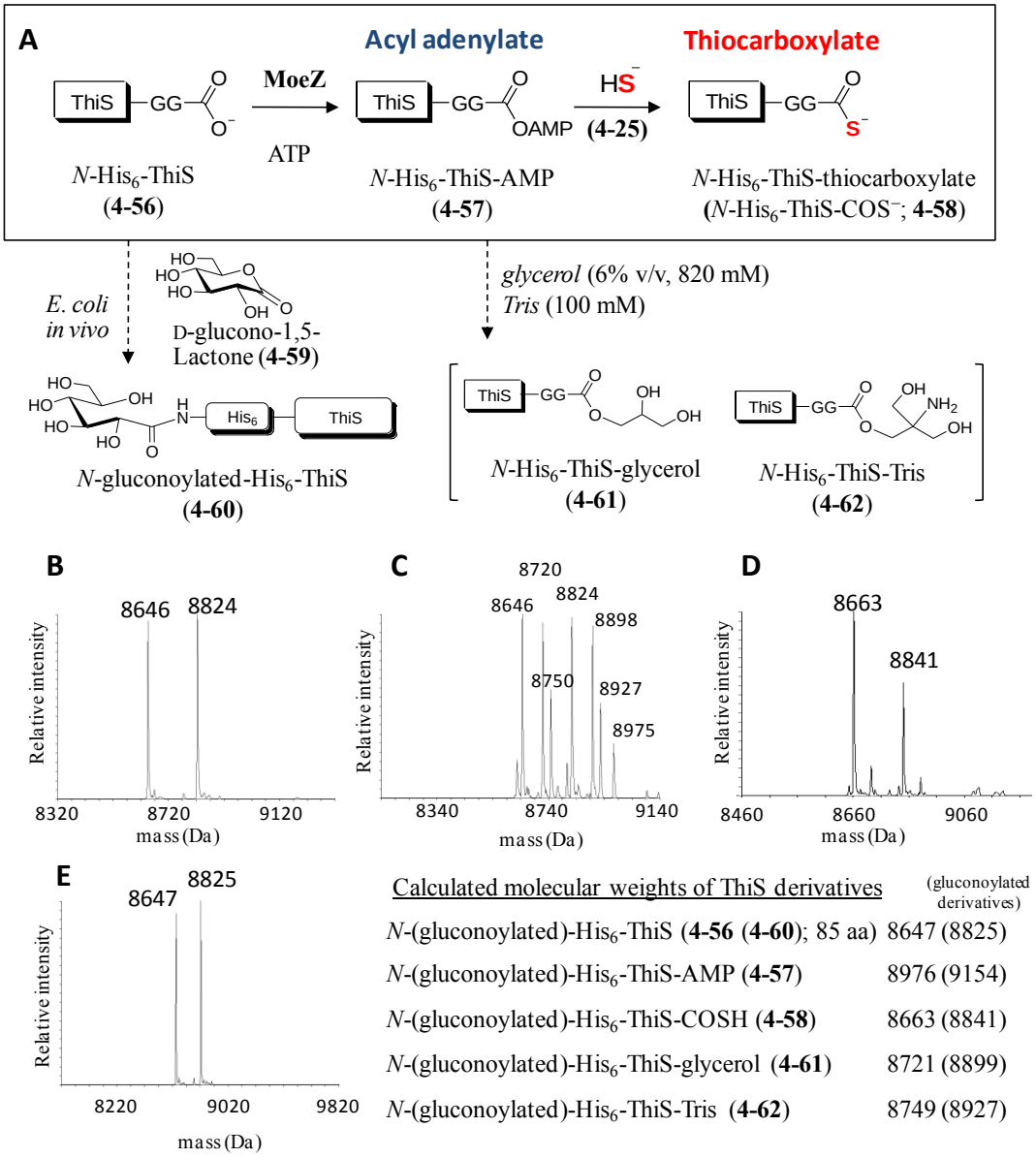


Figure 4-23. EIS-MS analyses of the MoeZ-catalyzed activation of ThiS

(A) Reaction scheme of the MoeZ-catalyzed activation of ThiS. (B) Deconvoluted ESI-MS of as isolated ThiS, (C) ThiS in the presence of MoeZ and ATP, (D) ThiS in the presence of MoeZ, ATP, and bisulfide (**4-25**), and (E) ThiS in the presence of bisulfide (control). The calculated molecular weights are shown as the neutral form in the right bottom.

Consideration of ThiS as the Sulfur Donor.

The *N*-His_6-ThiS-COS^- (**4-58**) obtained from the previous step was incubated with *C*-His_6-BexX, which forms a stable complex with D-glucose 6-phosphate (**4-2**) as described above. If sulfur transfer occurs and the 2-thiosugar product is released from the enzyme, then a mass signal corresponding to the *C*-His_6-BexX-D-glucose 6-phosphate should be converted to that of the free enzyme (Figure 4-24A). The protein *C*-His_6-BexX was freshly prepared, and most protein in this batch was found to be the enzyme−substrate complex (Figure 4-24B), thus, additional D-glucose 6-phosphate (**4-2**) was not added in this experiment. To our disappointment, we did not observe an increase of the free *C*-His_6-BexX in the presence of *N*-His_6-ThiS-COS^- (**4-58**; Figure 4-24 C and D). Instead, the *C*-His_6-BexX-G6P signal was stable even in the presence of ThiS, MoeZ, ATP and high concentrations of bisulfide (**4-25,** 10 mM). Although ThiS was originally considered to be the most likely candidate as the sulfur donor because of the protein sequence similarity between BexX and ThiG, our current result clearly ruled out this possibility and prompted us to investigate other possible sulfur donors.

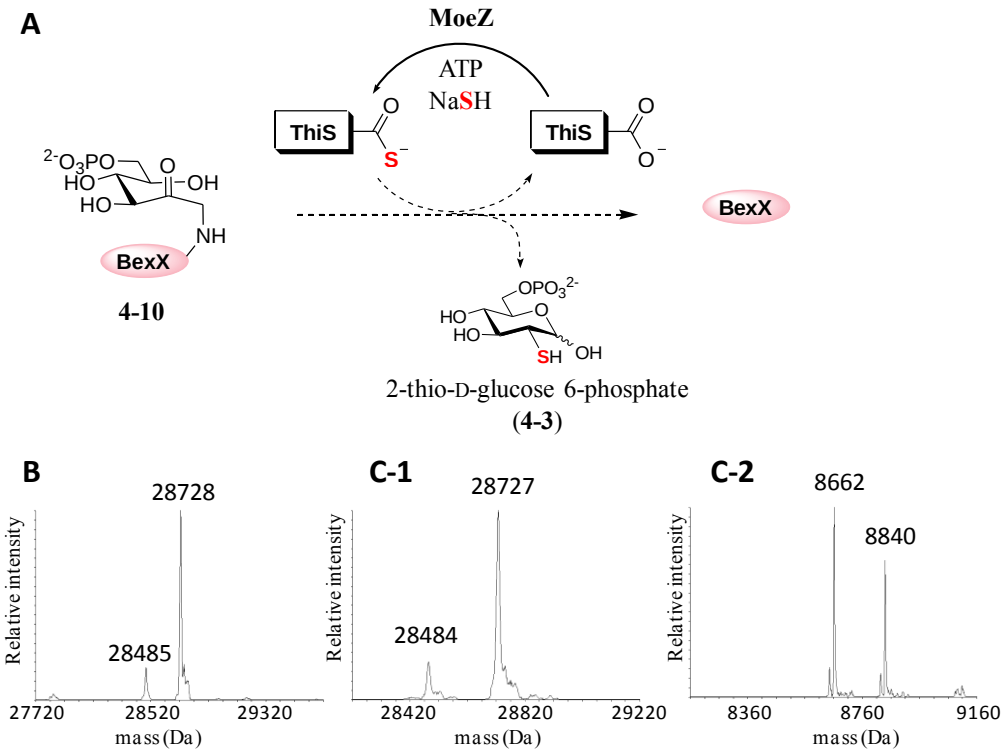


Figure 4-24. Investigation of potential sulfur transfer from ThiS-thiocarboxylate to BexX−D-glucose-6-phosphate complex (**4-10**).

(A**)** Reaction scheme of the expected sulfur transfer reaction**.** (B) Deconvoluted ESI-MS of the as-isolated *C*-His_6-BexX shows that the BexX−D-glucose-6-phosphate complex (**4-10**) is the main species. The calculated molecular weights of *C*-His_6-BexX (258 aa) and its glucose 6-phosphate complex (**4-10**) are 28488 and 29730, respectively. (C) **1**, Deconvoluted ESI-MS of the reaction mixture showing the mass range for BexX. Observed mass signals are consistent with *C*-His_6-BexX [28488 (calculated), 28484 (observed)] and *C*-His_6-BexX−D-glucose-6-phosphate [29730 (calculated), 29727 (observed)]. Conversion of the **4-10** to free BexX was not observed under these conditions, suggesting the proposed sulfur transfer does not occur from ThiS-thiocarboxylate. **2**, Deconvoluted ESI-MS of the same reaction mixture showing the mass range for ThiS. Observed mass signals are consistent with *N*-His_6-ThiS-COSH [8663 (calculated), 8662 (observed)] and *N*-gluconoyl-His_6-ThiS-COSH [8841 (calculated), 8840 (observed)].

4.3.9 Investigation of Other Sulfur Carrier Proteins.

MoeZ-Catalyzed Activation of Sulfur Carrier Proteins

As discussed above, three other sulfur carrier proteins, MoaD, CysO, and MoaD2, were found in the *A. orientalis* genome. While MoaD and CysO are likely involved in molybdopterin[70] and cysteine[84] biosyntheses, respectively (Figure 4-19 and 4-20), the physiological function of MoaD2 (MoaD homologue) is unknown. Although their native partner proteins, CysM and MoaE, do not show sequence similarities with BexX, one cannot predict the structure of the sulfur carrier protein-binding sites based simply on sequence analysis. Thus, it remains possible that MoaD, CysO, or MoaD2 may be able to interact with BexX and transfer its sulfur atom to the BexX–D-glucose 6-phosphate complex (**4-10**).

To test this hypothesis, we expressed and purified recombinant *N*-His_6-MoaD, *N*-His_6-CysO and *N*-His_6-MoaD2. While the latter two proteins were purified in good yields and quality (Figure 4-25, lane lane 2 and 3), *N*-His_6-MoaD was not expressed well in *E. coli*, and the latter protein was obtained in much less quantity and lower purity as compared to the former proteins (Figure 4-25, lane 7). ESI-MS analyses showed mass signals consistent with those expected for the corresponding enzymes and their *N*-gluconoylated derivatives as were seen in the case of *N*-His_6-ThiS (Figure 4-26B and C), though it was possible to obtain *N*-His_6-MoaD in the unmodified form (Figure 4-26A). Since a *thiF*-like gene does not exist near *moaD, cysO* or *moaD*2 in the *A. orientalis* genome, MoeZ was again proposed to be the activating enzyme, instead of ThiF or its equivalents, for these sulfur carrier proteins (Figure 4-19−21).

Thiocarboxylate formation on MoaD, CysO and MoaD2 in the presence of MoeZ, ATP and NaSH was confirmed in each case by MS (Figure 4-26D−F). These results suggest that MoeZ may act as a universal sulfur carrier protein-activating enzyme in *A.*

orientalis (Figure 4-27). Similar multifunctionality of a single ThiF-like protein in prokaryotes was recently predicted based on genome analyses.[181] Notably, TtuB, a ThiF-like enzyme from *Thermus thermophilus,* has been experimentally verified to contribute to both the thiamin and molybdopterin biosynthetic pathways as well as the tRNA thiouridine modification pathway.[86] Taken together, a common sulfur carrier protein activating-enzyme contributing to several independent biosynthetic pathways may be a common phenomenon in prokaryotes.

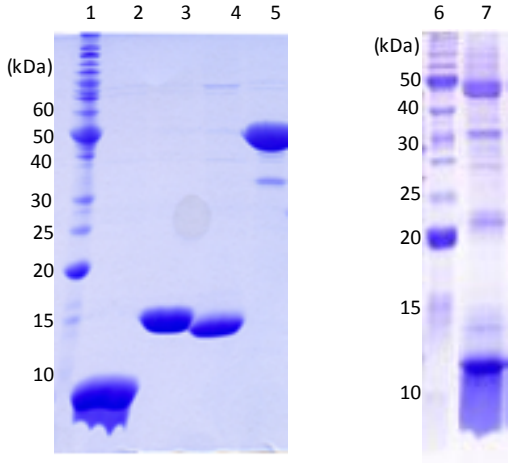


Figure 4-25. SDS-PAGE gel of purified sulfur carrier proteins and MoeZ.

N-His_6-ThiS (85 aa, 8.7 kDa, lane 2), *N*-His_6-CysO (109 aa, 11.7 kDa, lane 3), *N*-His_6-MoaD2 (115 aa, 12.5 kDa, lane 4), *N*-His_6- MoeZ (421 aa, 45.0 kDa, lane 5), and *N*-His_6-MoaD (105 aa, 11.0 kDa, lane 7). The molecular weight marks are 220, 160, 120, 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, and 10 kDa (top to bottom, lane 1 and 6). The protein MoaD was not expressed well, and the partially purified protein solution contained significant amounts of endogenous proteins from the *E. coli* host.

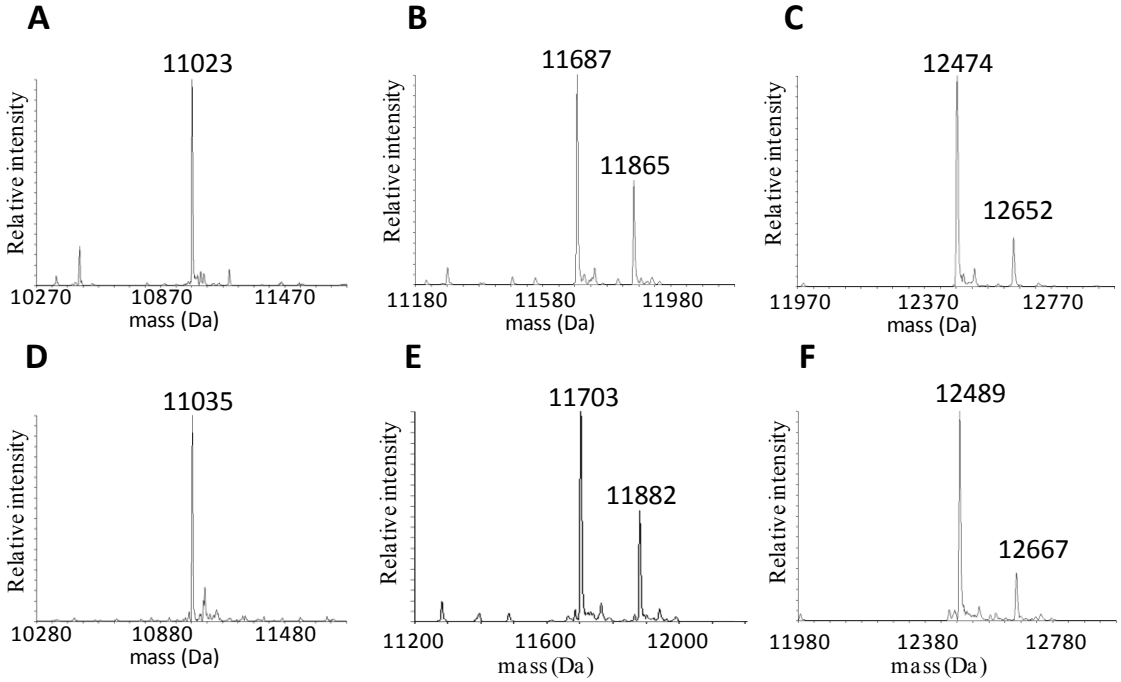


Figure 4-26. ESI-MS analyses of the MoeZ-catalyzed activation of sulfur carrier proteins

(A) Deconvoluted ESI-MS of as-isolated MoaD (the calculated molecular weight of *N*-His_6-MoaD (105 aa) is 11022 Da), (B) as-isolated CysO **(**the calculated molecular weight of *N*-His_6-CysO (109 aa) and its *N*-gluconoylated derivative are 11688 and 11866, respectively), (C) as-isolated MoaD2 **(**the calculated molecular weights of *N*-His_6-MoaD2 (115 aa) and its *N*-gluconoylated derivative are 12473 and 12651, respectively), (D**)** MoaD incubated with MoeZ, ATP, and NaSH **(**the calculated molecular weight of *N*-His_6-MoaD-COSH is 11038 Da), (E) CysO incubated with MoeZ, ATP, and NaSH (the calculated molecular weights of *N*-His_6-CysO-COSH and its *N*-gluconoylated derivative are 11704 and 11882, respectively), (F) MoaD2 incubated with MoeZ, ATP, and NaSH (the calculated molecular weights of *N*-His_6-MoaD-COSH and its *N*-gluconoylated derivative are 12489 and 12667, respectively).

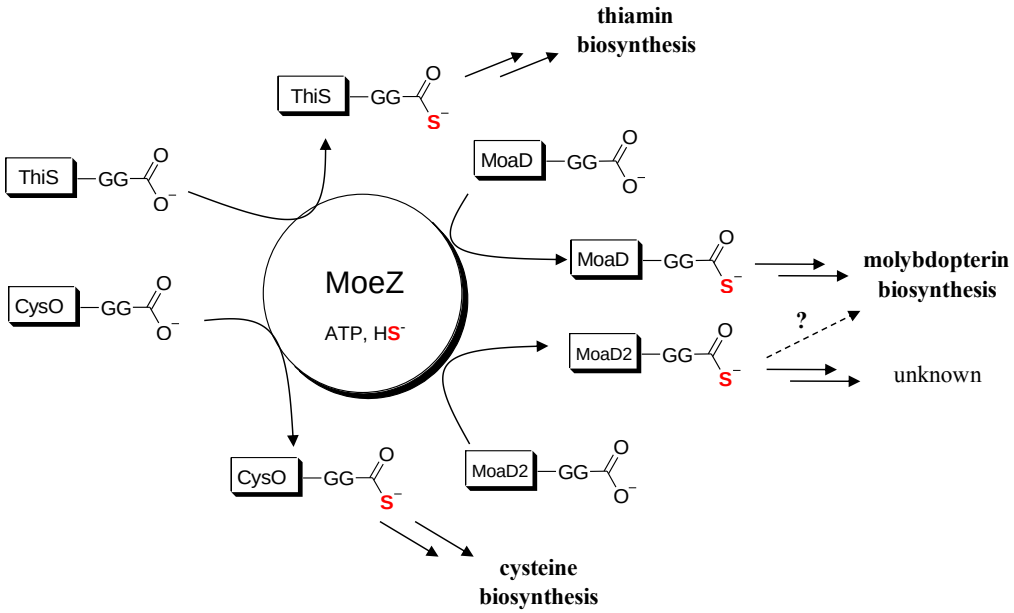


Figure 4-27. Multifunctionality of MoeZ observed in *A. orientalis.*

MoeZ appears to be a universal activator of sulfur carrier proteins in *A. orientalis*. These sulfur carrier proteins are involved in the biosynthetic pathways of thiamin, cysteine, molybdopterin, and possibly other unidentified sulfur-containing biomolecules.

Investigation of Sulfur Carrier Protein-Mediated Sulfur Incorporation

With activated CysO and MoaD2 in hand, we next examined the proposed sulfur transfer reaction from these enzymes to the BexX–D-glucose 6-phosphate complex (**4-10**). MoaD was not used for this assay because of its low concentration and purities (Figure 4-25, lane 7). Relative intensities of mass signals corresponding to **4-10** and the free enzyme were monitored before and after addition of the activated sulfur carrier proteins. In contrast to the analogous experiment using *N*-His_6-ThiS-COS^- (Figure 4-24), we observed the free *C*-His_6-BexX after addition of these sulfur carrier proteins (Figure 4-28). This indicates that the anticipated sulfur transfer followed by the thiosugar product release occurred from both *N*-His_6-CysO-COS^- and *N*-His_6-MoaD2-COS^-.

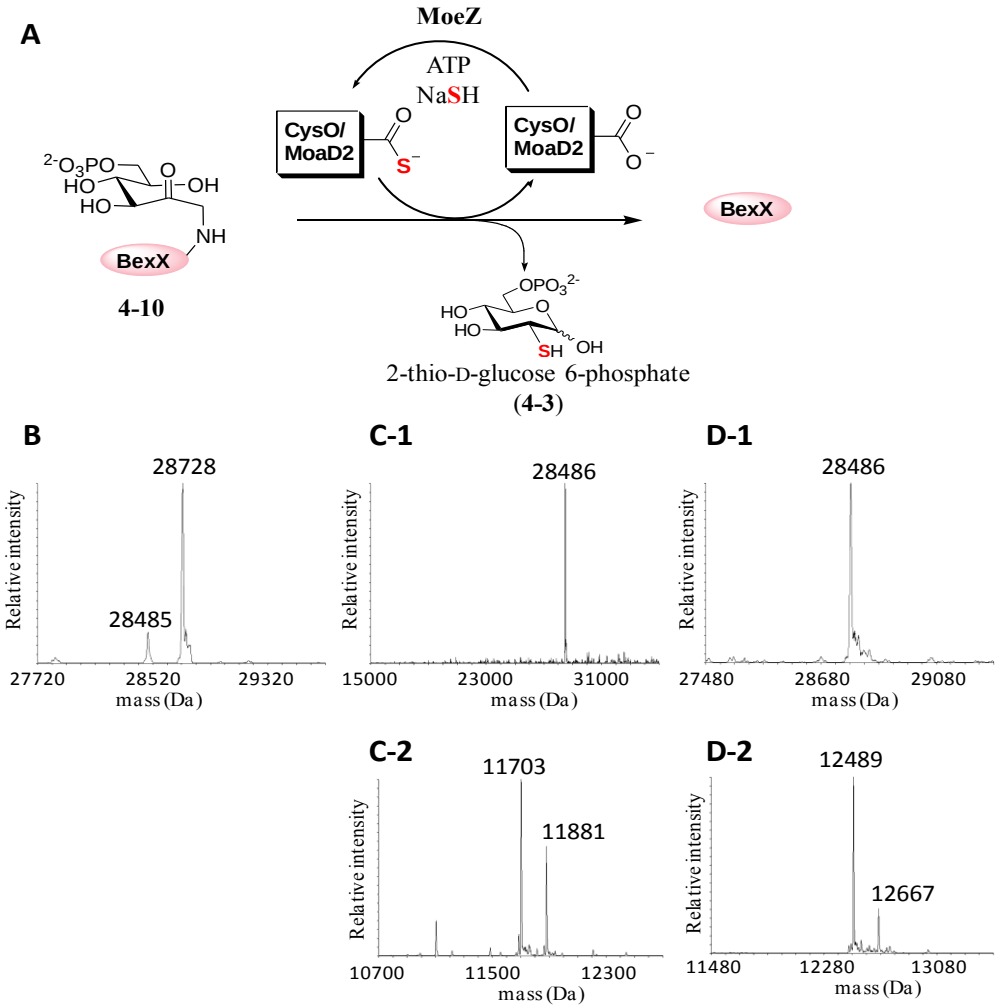


Figure 4-28. Sulfur transfer from CysO- or MoaD2-COS$^-$ to BexX-D-glucose-6-phosphate complex (**4-10**).

(A) Reaction scheme of the expected sulfur transfer reaction. (B) Deconvoluted ESI-MS of the as-isolated *C*-His$_6$-BexX shows that **4-10** is the main species. The calculated molecular weights of *C*-His$_6$-BexX (258 aa) and **4-10** are 28488 and 29730, respectively. (C) Deconvoluted ESI-MS of the reaction mixture using CysO. **1**, showing only the free form of *C*-His$_6$-BexX, suggesting the expected sulfur transfer and **2**, showing fully activated CysO-COS$^-$ and its *N*-gluconoylated derivative. (D) Deconvoluted ESI-MS of the reaction mixture using MoaD2 also shows **1**, only the free form of *C*-His$_6$-BexX, and **2**, fully activated MoaD2-COS$^-$.

Spectrophotometric Analysis of the Sulfur Transfer

MoeZ-catalyzed sulfur carrier-protein activation and the following sulfur transfer to the BexX–D-glucose-6-phosphate complex (**4-10**) were also analyzed spectrophotometrically using adenylate kinase (AK), pyruvate kinase (PK), and lactate dehydratase (LDH) to assay for the formation of AMP. As shown in Figure 4-29, one molecule of AMP is released upon formation of one molecule of sulfur carrier protein-thiocarboxylate. AK converts the produced AMP and another molecule of ATP to two molecules of ADP, which leads to the consumption of two molecules of NADH *via* the coupled enzymatic reactions. The consumption of NADH can be conveniently monitored based on its absorbance at 340 nm.

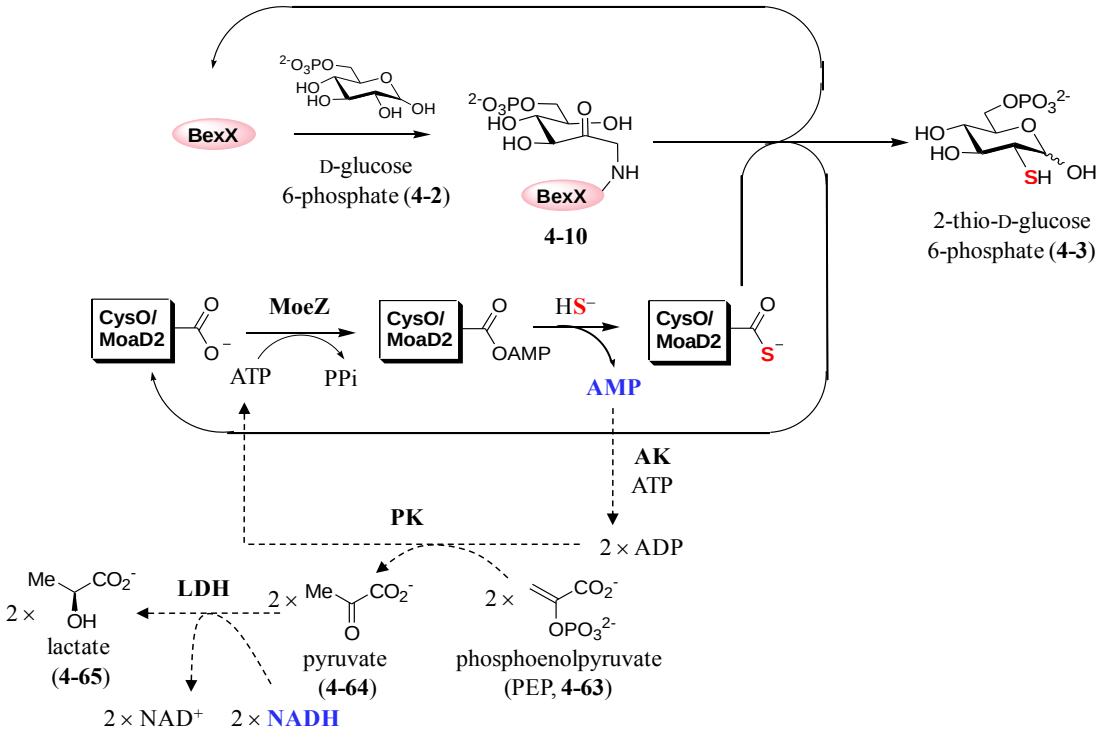
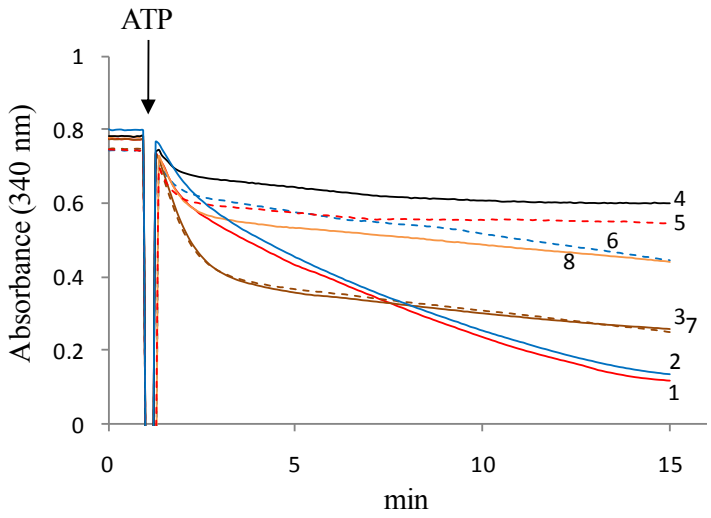


Figure 4-29. Spectrophotometric analysis of AMP production during the course of sulfur carrier protein activation by MoeZ and its sulfur transfer to BexX-D-glucose-6-phosphate complex (**4-10**).

When *N*-His_6-CysO or *N*-His_6-MoaD2 was used, consumption of NADH was faster in the presence of *C*-His_6-BexX (Figure 4-30, trace 1 or 2) compared to the control reaction without added BexX (Figure 4-30, trace 5 or 6). In contrast, addition of BexX did not affect the rate of NADH consumption when *N*-His_6-ThiS was used (Figure 4-30, trace 3 and 7). These observations are consistent with the ESI-MS analysis described above that *N*-His_6-CysO and *N*-His_6-MoaD2 produced the free form of *C*-His_6-BexX (Figure 4-28) while *N*-His_6-ThiS-COS^- did not (Figure 4-24). The observed decrease of absorbance at 340 nm in the first 2–3 min was mainly caused by a small amount of AMP contamination found in the ATP solution used for this experiment (Figure 4-30, trace 4). The single turnover of the MoeZ-catalyzed sulfur carrier protein-activation reaction was deduced by observing the greater decrease in absorbance in the first 2–3 min when sulfur carrier proteins were added (Figure 4-30, trace 5–7).

Although the purified *N*-His_6-MoaD contained impurities, a similar experiment was performed for this enzyme. The preliminary data suggested that *N*-His_6-MoaD cannot transfer its sulfur atom to the BexX-D-glucose-6-phosphate complex (**4-10**) as seen for *N*-His_6-ThiS.



	1	2	3	4	5	6	7	8
***C*-His$_6$-BexX**	+	+	+	+	–	–	–	+
***N*-His$_6$-MoaD2**	+	–	–	–	+	–	–	–
***N*-His$_6$-CysO**	–	+	–	–	–	+	–	–
***N*-His$_6$-ThiS**	–	–	++	–	–	–	++	+

Figure 4-30. Spectrophotometric analysis of AMP production using sulfur carrier proteins, MoeZ, and BexX with the AK, PK, and LDH-coupled reaction system.

Reaction conditions of used enzymes are shown in the table. *C*-His$_6$-BexX (+): 10 μM, *N*-His$_6$-MoaD2 (+): 9 μM, *N*-His$_6$-CysO (+): 6 μM, *N*-His$_6$-ThiS (++): 20 μM, *N*-His$_6$-ThiS (+): 10 μM. The reaction mixture also contained MoeZ (3 μM), ATP (1 mM), D-glucose 6-phosphate (5 mM), NaSH (5 mM), PEP (0.3 mM), NADH (0.1 mM), AK, PK, LDH, and $MgCl_2$ (5 mM) in 50 mM Tris·HCl buffer (pH 8.0). ATP was added to initiate the catalytic cycles at time = 1 min.

4.3.10 Accomplishment of the Enzymatic 2-Thiosugar Formation.

Encouraged by the results described above, direct detection of the product thiosugar was then pursued. The purified sulfur carrier protein, *N*-His$_6$-CysO, *N*-His$_6$-MoaD2 or *N*-His$_6$-ThiS was incubated with *C*-His$_6$-BexX and *N*-His$_6$-MoeZ in the

presence of D-glucose 6-phospahte (**4-2**), ATP and NaSH. If the expected sulfur transfer occurs, the single catalytic cycle should produce one molecule of AMP and one molecule of 2-thio-D-glucose-6-phosphate (**4-3**) (Figure 4-29, solid line).

The reaction products were analyzed by HPLC after derivatizing the predicted thiosugar with the thiol reactive chromophore, monobromobimane[182,183] (mBBr, **4-66**; Figure 4-31A). When *N*-His_6-ThiS was used, the extent of AMP was similar to that observed with the control having no sulfur carrier protein, suggesting the detected AMP is likely derived from self degradation of ATP during the incubation (Figure 4-31B, traces 5 and 6). On the other hand, when *N*-His_6-CysO or *N*-His_6-MoaD2 was used, a clear increase in AMP formation was observed along with the appearance of a new peak (product peak) (Figure 4-31B, trace 3 or 4). In addition, control experiments showed that formation of AMP and the product peak were BexX- and MoeZ-dependent (Figure 4-31B, trace 1 and 2). The product peak was isolated and analyzed by ESI-MS. The observed signal (m/z 465.0741) matched the calculated value for 2-thio-D-glucose 6-phosphate-bimane (**4-67**, $C_{16}H_{22}N_2O_{10}PS^-$ [M – H^+] 465.0738; Figure 4-31C).

To further evaluate the assignment of the new product peak, each reaction was also treated with alkaline phosphatase prior to the addition of mBBr (**4-66**) and similarly analyzed by HPLC (Figure 4-32A). The retention time of the product peak was identical to that of the synthetic standard of the 2-thio-D-glucose-bimane derivative (**4-69**, Figure 4-32B). The product peak was also isolated and analyzed by ESI-MS. The observed signal (m/z 409.1038) indeed matches the calculated value (m/z 409.1040; Figure 4-32C). Taken together, we can conclude that BexX-catalyzed 2-thiosugar formation proceeds in the presence of either CysO-COS^- or MoaD2-COS^- but not ThiS-COS^-. These results likewise demonstrated the first *in vitro* enzymatic synthesis of a 2-thiosugar.

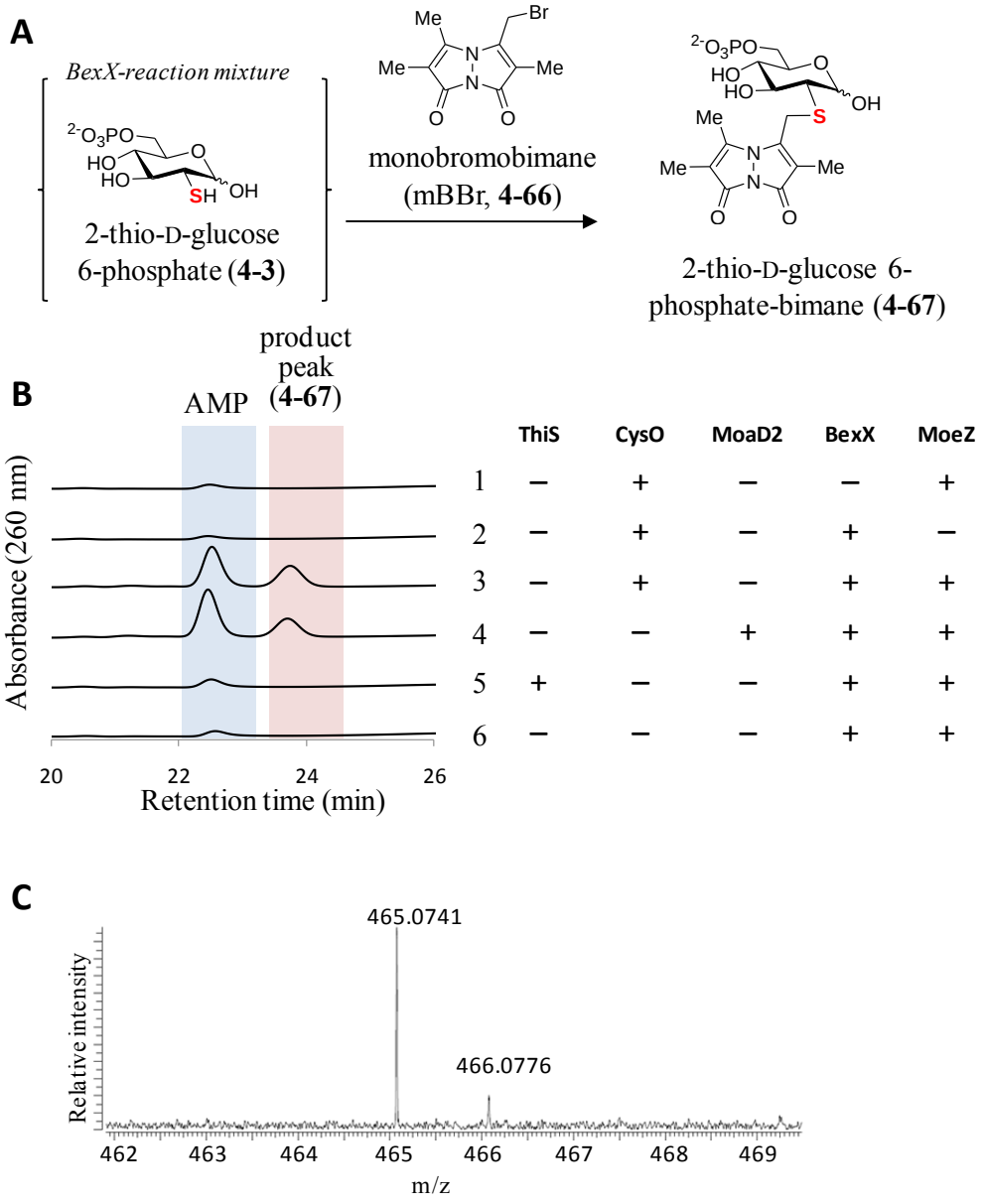


Figure 4-31. BexX-catalyzed 2-thio-D-glucose 6-phosphate formation with mBBr derivatization.

(A) Reaction scheme to make the expected bimane derivative (**4-67**). (B) HPLC traces of the *C*-His_6-BexX-catalyzed reactions utilizing *N*-His_6-ThiS, *N*-His_6-CysO or *N*-His_6-MoaD2 and the control reactions. The thiosugar product was derivatized with mBBr (**4-66**). (C) ESI-MS of the isolated peak (**4-67**). HRMS (ESI^-) calculated for $C_{16}H_{22}N_2O_{10}PS^-$ [$M - H^+$] 465.0738, observed 465.0741.

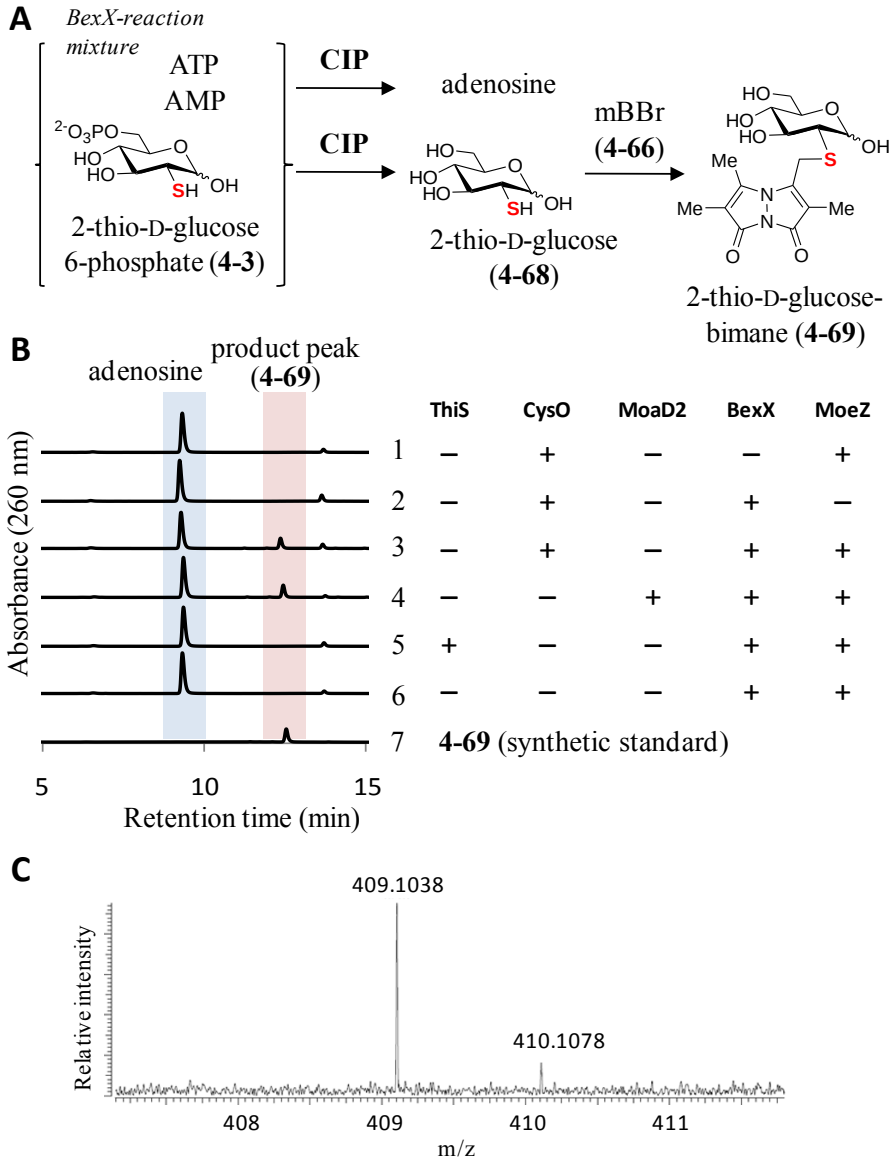


Figure 4-32. BexX-catalyzed 2-thio-D-glucose 6-phosphate formation with alkaline phosphatase (CIP) and mBBr.

(A) Reaction scheme to make the expected bimane derivative (**4-69**). (B) HPLC traces of the *C*-His_6-BexX-catalyzed reactions utilizing *N*-His_6-ThiS, *N*-His_6-CysO or *N*-His_6-MoaD2 and the control reactions. The thiosugar product was derivatized with mBBr (**4-66**). The synthetic standard of **4-69** is shown on the bottom trace (trace 7). (C) ESI-MS of the isolated peak (**4-69**). HRMS (ESI^+) calculated for $C_{16}H_{22}N_2NaO_7S^+$ $[M + Na]^+$ 409.1040, observed 409.1038.

4.3.11 Sequence Analysis of Sulfur Carrier Proteins.

It is surprising that BexX can utilize either CysO or MoaD2 but not ThiS (and MoaD) as its partner despite the sequence similarity between BexX and ThiG (58% identity and 75% similarity to ThiG from *Stigmatella aurantiaca* DW4/3-1 (ZP_01459245),[141] 38% identity and 59% similarity to ThiG from *Bacillus subtilis* subsp. *subtilis* str 168[142]). Confirmation that BexX is the desired 2-thioglucose 6-phosphate synthase, prompted us to consider that the relatively high sequence similarity found between BexX and ThiG from *S. aurantiaca* (ZP_01459245) may be an indication of the latter enzyme as a possible thiosugar synthase rather than a thioazole synthase. Furthermore, analysis of the sequence surrounding the putative *thiG* gene in *S. aurantiaca* failed to locate any other genes related to thiamin biosynthesis. Instead, a putative sulfur carrier protein annotated as PdtH homologue (ZP_01459250) can be located nearby. This leads to the proposal that the PdtH homologue (ZP_01459250) may be responsible for providing a sulfur atom to a complex involving the ThiG homologue (ZP_01459245) and its unknown substrate.

To gain more insights into the specific interaction between BexX and the sulfur carrier proteins, the protein sequences of the four sulfur carrier proteins from *A. orientalis* were compared with those of other previously reported sulfur carrier proteins and the related proteins as well as the PdtH homologue from *S. aurantiaca* (Figure 4-33). In addition, the secondary structures of the sulfur carrier proteins from *A. orientalis* and the PdtH homologue from *S. aurantiaca* were also analyzed (Figure 4-34). Although the former analysis did not provide significant information to explain the observed interaction between BexX and CysO or MoaD2, the latter analysis revealed a similar organization of the predicted secondary structures in CysO, MoaD2, and PdtH from *S. aurantiaca.* This suggests that such a structural organization may be necessary for the

interaction with BexX (or with the ThiG homologue from *S. aurantiaca*) and sulfur transfer. Further experiments are clearly required to address these hypotheses, and the structural analyses of these enzymes are in progress.

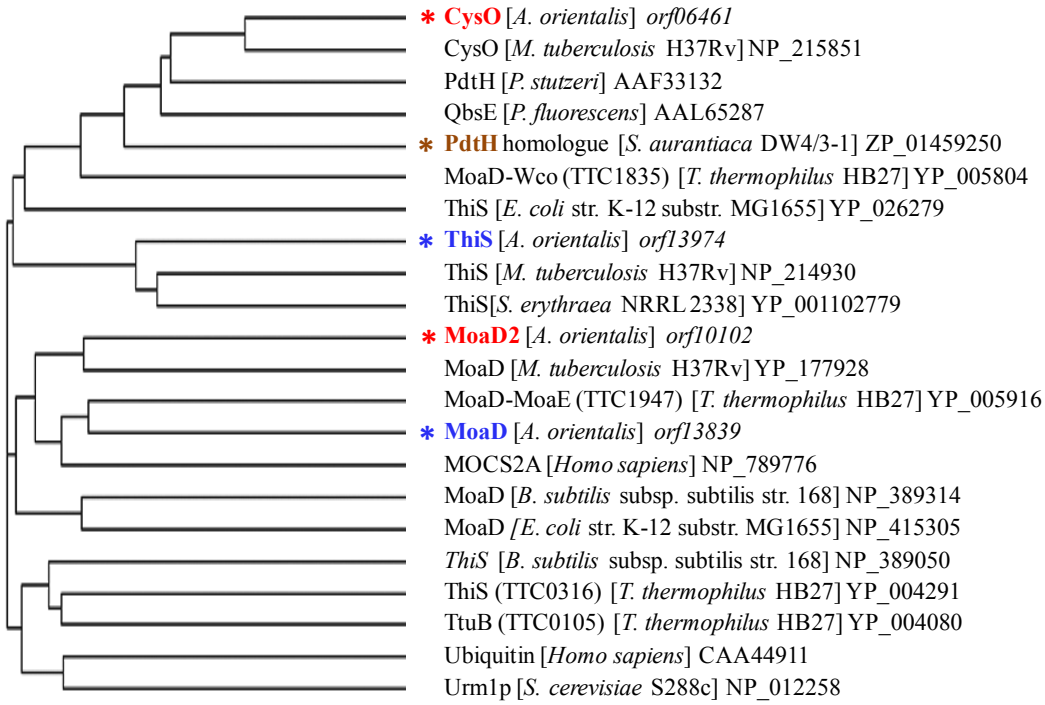


Figure 4-33. Protein sequence similarity analysis of sulfur carrier proteins, ubiquitin, and ubiquitin-like proteins.

The sulfur carrier proteins from *A. orientalis* and the PdtH homologue from *S. aurantiaca* are labeled with asterisks.

Figure 4-34. Predicted secondary structures of sulfur carrier proteins from *A. orientalis* and the PdtH homologue from *S. aurantiaca*.

(A) ThiS from *A. orentalis* (*orf13974*). (B) MoaD from *A. orentalis* (*orf13839*). (C) CysO from *A. orentalis* (*orf06461*). (D) MoaD2 from *A. orentalis* (*orf10102*). (E) PdtH homologue from *S. orientalis* (ZP_01459250). Prediction contains three states: alpha helix (H), beta strand (S) and coil (C), with confidence scores for each residue, in which higher score indicates a prediction with higher confidence.

4.3.12 *In vivo* Sulfur Source of the Sulfur Carrier Proteins.

Although CysO and MoaD2 can act as the sulfur donor for the 2-thiosugar biosynthetic pathway, the *in vivo* sulfur donor required to activate these enzymes remains unknown. Since the sulfur carrier protein-acyl adenylate intermediate formed by MoeZ (Figure 4-23A) was shown to be labile in the solution (Figure 4-23C) it must be orderly attacked by bisulfide or its equivalent as soon as it forms. In many cases, cysteine desulfurases has been implicated as the sulfur donors for the sulfur carrier proteins (see Chapter 1). Thus, the sulfur carrier proteins found in *A. orientalis* may interact with the endogenous cysteine desulfurases *in vivo*. Another possible candidate for the sulfur

source is a rhodanese-type enzyme. Interestingly, the *C*-terminal domain of MoeZ display good characteristics of a rhodanese-type enzyme (Figure 4-22). If this domain exhibits rhodanse activity, then the formed protein persulfide may be transferred to the adenylated sulfur carrier proteins, which may still be bound to the *N*-terminal adenylation domain of the same enzyme. In this last section, a preliminary investigation of the BexX-catalyzed 2-thiosugar formation utilizing MoeZ and thiosulfate, or cysteine desulfurases and L-cysteine instead of exogenous bisulfide is described.

Rhodanese Activity of MoeZ

The sulfur carrier protein, *N*-His_6-CysO was incubated with *C*-His_6-BexX and *N*-His_6-MoeZ in the presence of D-glucose 6-phospahte (**4-2**), ATP, thiosulfate, and DTT. The reaction was analyzed by HPLC after derivatization with mBBr as described above. The product peaks corresponding to 2-thio-D-glucose 6-phosphate-bimane (**4-67**) and 2-thio-D-glucose-bimane (**4-69**; with CIP treatment) were similarly observed as the reaction using sodium bisulfide (Figure 4-31B and 4-32B). This suggests that MoeZ does possess the expected rhodanese activity (Figure 4-35). At this point, however, whether the protein persulfide formed in MoeZ can directly attack the acyl adenylate group of the sulfur carrier protein is not known because the generated persulfide can be easily reduced by DTT to produce bisulfide in the reaction solution.

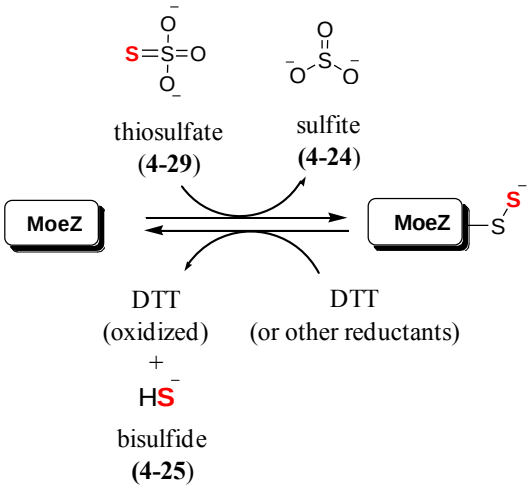


Figure 4-35. Rhodanse activity of MoeZ.

Cysteine Desulfrases in A. Oreintalis

Whole genome analysis of *A. orientalis* led to the identification of five cysteine desulfurase homologues (Table 4-8). These enzymes with *N*-terminal His_6-tag were purified except CD5, which was expressed as inclusion bodies (Figure 4-36). The sulfur carrier protein, *N*-His_6-MoaD2 was then incubated with one of the purified cysteine desulfurases (CD1–4), *C*-His_6-BexX and *N*-His_6-MoeZ in the presence of D-glucose 6-phospahte (**4-2**), ATP, L-cysteine (**4-26**), PLP, and DTT. The reaction was analyzed by HPLC after derivatization with mBBr as described above.

The product peak corresponding to 2-thio-D-glucose 6-phosphate-bimane (**4-67**) was observed, as in the control reaction, using exogenous bisulfide and *N*-His_6-CD2 and *N*-His_6-CD4. The reaction mixture containing *N*-His_6-CD1 did not produce a detectable amount of **4-67,** and the *N*-His_6-CD3 reaction gave a significantly lower yield of **4-67** than those of *N*-His_6-CD2 and *N*-His_6-CD4 reactions. When DTT was omitted, the yields of **4-67** from *N*-His_6-CD2 and *N*-His_6-CD4 reaction mixtures were dramatically decreased. This is consistent with the expected cysteine desulfurase activities of the

enzymes, which require a reducing agent for multi-turnovers (Figure 4-37A). Thus, the results of these experiments fully established the cysteine desulfurase activities of CD2 and CD4.

In contrast, the reaction mixture containing *N*-His$_6$-CD3 in the absence of DTT produced similar amount of **4-67** as the control reaction using exogenous bisulfide. This suggests that CD3 is not a cysteine desulfurase but may be a cystine lyase, which utilizes PLP to produce cysteine persulfide and pyruvate (**4-64**) from cystine (**4-70**; Figure 4-37B).[184,185] Indeed, production of **4-64** was confirmed spectrophotometrically using lactate dehydrogenase by monitoring the decrease of NADH (absorbance at 340 nm) along with the reduction of **4-64** to lactate (**4-65**).

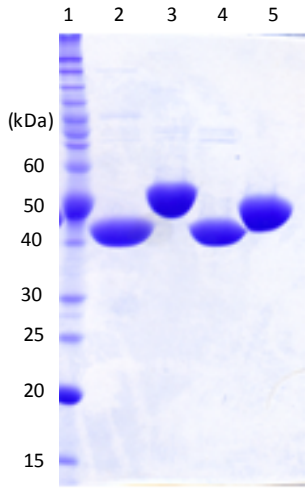


Figure 4-36. SDS-PAGE gel of purified cysteine desulfurases.

N-His$_6$-CD1 (415 aa, 44.1 kDa, lane 2), *N*-His$_6$-CD2 (485 aa, 51.4 kDa, lane 3), *N*-His$_6$-CD3 (432 aa, 44.8 kDa, lane 4), and *N*-His$_6$-CD4 (417 aa, 43.3 kDa, lane 5). The molecular weight marks are 220, 160, 120, 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, and 10 kDa (top to bottom, lane 1).

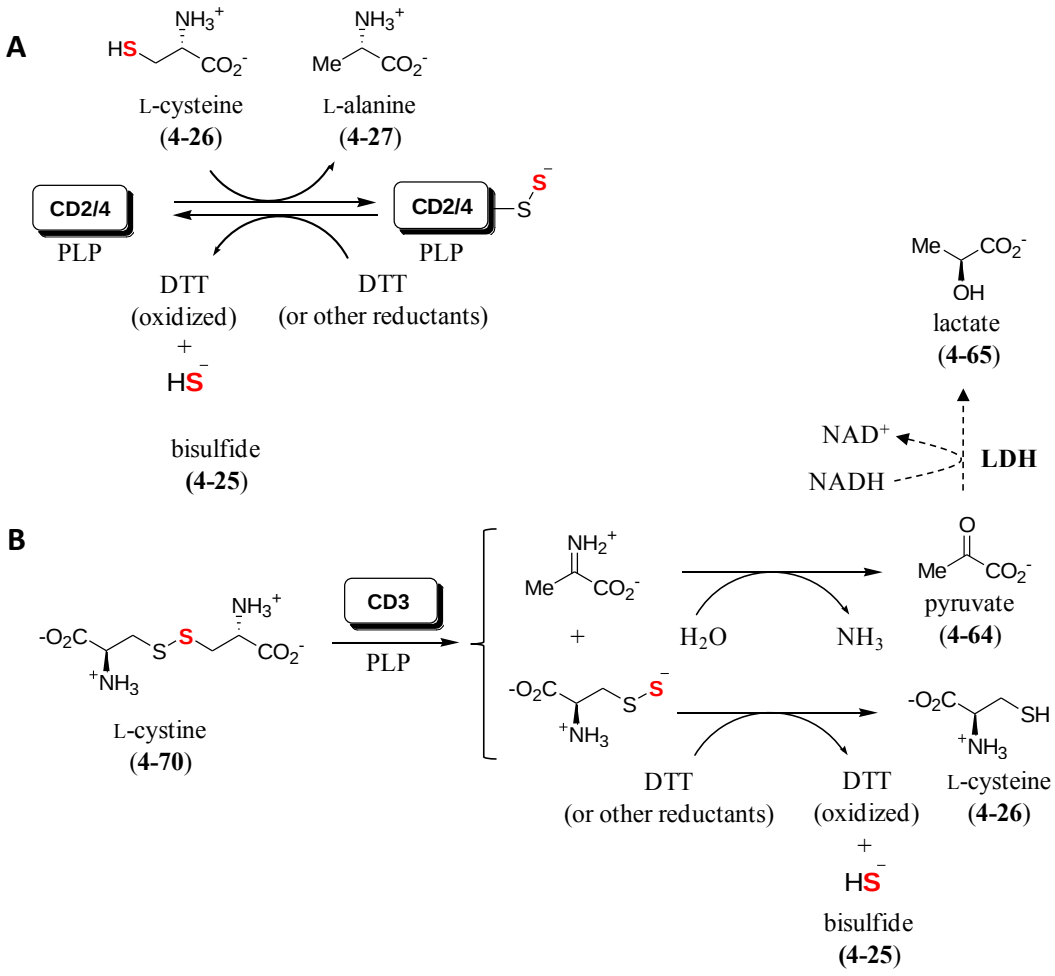


Figure 4-37. Cysteine desulfurase acitivity of CD2 and CD4

(A), and the proposed cystine lyase activity of CD3 (B). Detection of pyruvate using LDH and NADH is shown in dashed lines.

In summary, rhodanese activity of MoeZ, cysteine desulfurase activity of CD2 and CD4 and cystine lyase-like activity of CD3 were confirmed. They are potentially responsible for acting as sulfur donors for the sulfur carrier proteins in *A. orientalis*. However, whether a mechanism of direct sulfur transfer mediated by protein–protein interactions is involved or not is not fully understood. The resulting protein persulfides may also act as an indirect sulfur source because they can easily liberate bisulfide upon

reduction by endogenous reductants. Furthermore, major intracellular thiols in *Streptomyces* species have been determined using the mBBr derivatization method.[182,183] The identified species include cysteine, coenzyme A, sulfide, thiosulfate, and mycothiol.[182,186,187] Thus, cysteine, thiosulfate, and sulfide considered above are all potential *in vivo* sulfur sources for the sulfur carrier proteins.

4.4 CONCLUSIONS

In this chapter, the biosynthetic pathway of the 2-thiosuar moiety of BE-7585A were investigated. BexX was characterized as the 2-thio-D-glucose 6-phoshate (**4-3**) synthase, which produces **4-3** from D-glucose 6-phoshate (**4-2**) through a mechanism reminiscent to that of thiazole synthase (ThiG). The BexX-catalyzed sulfur incorporation reaction requires a sulfur delivery protein, whose encoded gene is not in the BE-7585A biosynthetic gene cluster. Therefore, it must be recruited from the primary metabolite biosynthetic pathways. The desired enzyme was identified through a genome mining approach followed by *in vitro* characterization. A sulfur carrier protein found in the thiamin biosynthetic gene cluster (ThiS) is not the sulfur donor despite the high sequence similarity between BexX and ThiG. Instead, a sulfur carrier protein found in the cysteine biosynthetic gene cluster (CysO) and a sulfur carrier protein (MoaD2) similar to MoaD in molybdopterin biosynthesis were confirmed as productive sulfur donors for the BexX reaction. During the course of these studies, the non-pathway specific sulfur carrier protein-activating enzyme, MoeZ, was also discovered.

The works presented here are significant for several reasons. First, this is the first report of the reconstitution of 2-thiosugar biosynthesis *in vitro*. Second, recruitment of a sulfur carrier protein from primary metabolism to be used in the biosynthesis of

secondary metabolites is an intriguing finding. Indeed, it is the first example of this type of enzymes to pair with a non-pathway protein which is encoded by a gene out of the biosynthetic gene cluster. Third, BexX is the first enzyme having ability to accept sulfur atom from multiple sulfur donor proteins. These results demonstrate that the collaboration interaction between the sulfur carrier protein and its acceptor protein is not always very specific. This implies that there may be more unidentified pairs of sulfur carrier and acceptor proteins. Namely, Nature may have evolved a combinational approach to match the sulfur carrier and the acceptor protein in different applications. For example, a pathway-independent sulfur carrier protein may play similar roles in the biosyntheses of other types of thio-containing compounds. Finally, MoeZ was identified as the universal sulfur carrier protein activator in *A. orientalis*. Based on the reported genome sequences of several Actinobacteria, this may be a common feature for this class of organisms. Moreover, the *C*-terminal rhodanese activity of the enzyme indicates that the *in vivo* sulfur source of sulfur carrier proteins may be thiosulfate although the possibility for utilizing bisulfide or an equivalent produced by other enzymes cannot be excluded.

In terms of the mechanism, a stable C-2 keto form of the BexX–D-glucose-6-phoshphate complex was identified as the key intermediate in this reaction. While NH_2OH could specifically trap this carbonyl functional group without denaturing the enzyme, bisulfide alone cannot appear to facilitate sulfur incorporation at C-2. A sulfur carrier protein-thiocarboxylate terminal is essential for delivers sulfur to BexX–2-ketosugar complex to complete the catalytic cycle. Site specific covalent attachment of the sugar substrate to Lys110 of BexX was confirmed by mutagenesis studies and LC-MS/MS analysis using trypsin digestion. Based on previous studies of ThiG, Asp112 and Glu194 were also predicted to be important for catalysis (Figure 4-9).[143] The proposed

mechanism of BexX-catalyzed 2-thio-D-glucose 6-phoshate (**4-3**) formation is shown in Figure 4-38. The initial steps are likely several isomerization reactions to generate the C-2 keto intermediate (**4-8** → **4-9** → **4-10**), priming the sugar substrate to accept the thiol group from a nucleophilic sulfur donor. Predicted active site residues, Asp112 and Glu194, likely play important roles in this process. A sulfur carrier protein CysO or MoaD2 is then recruited from the primary metabolite biosynthetic pathway to attack the C-2 keto intermediate (**4-10**) and introduce the thiofunctional group. Dehydration at C-2 through hydrolysis of the thioester bond formed with sulfur carrier protein then occurs to form of the C-2 thioketone intermediate (**4-12**). Subsequent isomerization of the thiol intermediate gives the enamine (**4-72**) and finally the imine intermediate (**4-13**). This would then be hydrolyzed to yield the 2-thio-D-glucose 6-phosphate (**4-3**) product. Mutagenesis studies targeting Asp112 and Glu194 as well as BexX protein structural analysis to further investigate the proposed mechanism are in progress.

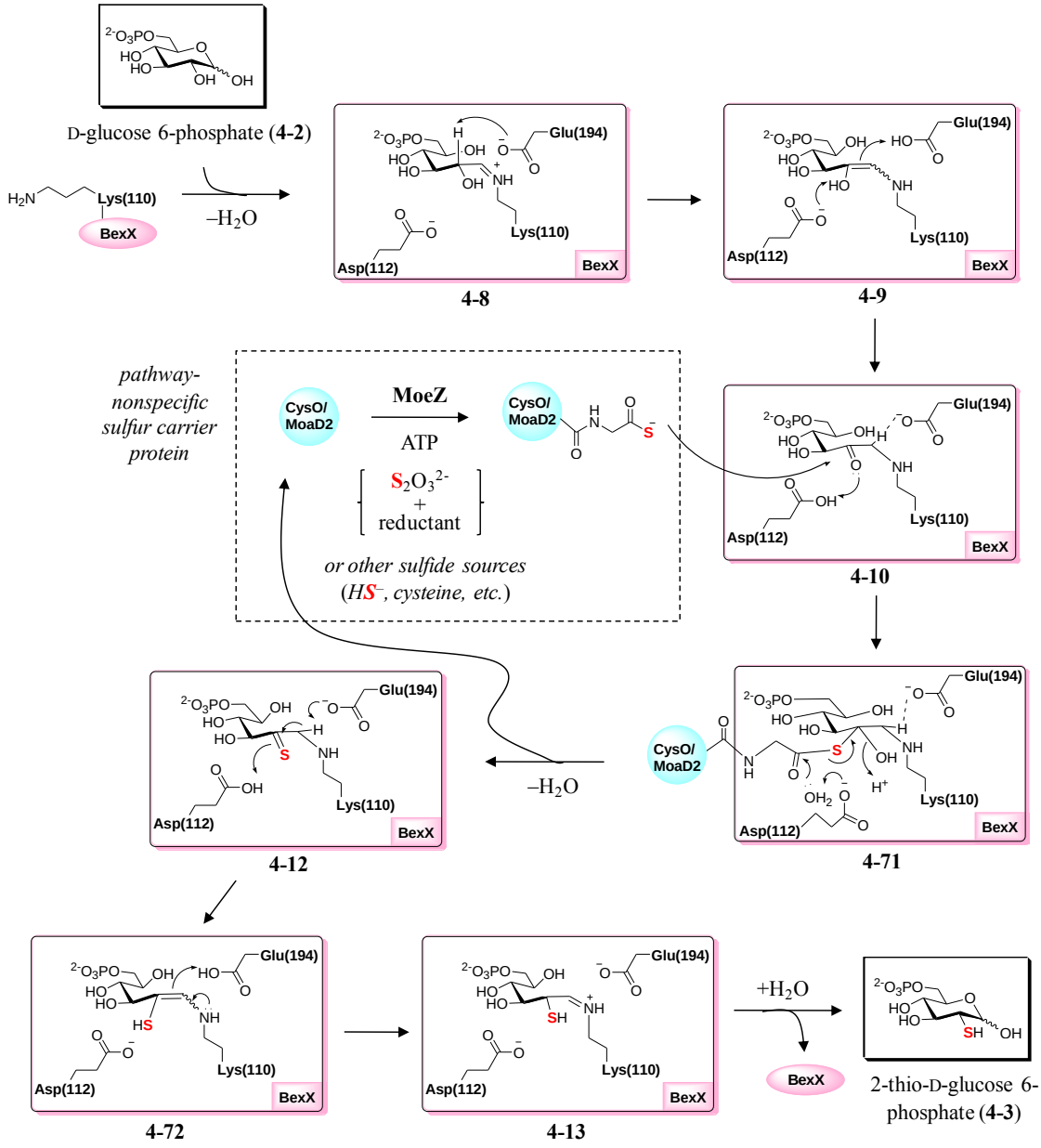


Figure 4-38. Proposed mechanism of BexX-catalyzed 2-thio-D-glucose 6-phosphate synthesis.

Chapter 5: Biosynthetic Studies of Lincomycin A: Functional Characterization of the Transaldolase, LmbR, and Isomerase, LmbN, to Construct the Octose 8-Phosphate Intermediate Towards the Biosynthesis of Methylthiolincosamide

5.1 INTRODUCTION

Lincomycin A (**5-1**) is a lincosamide antimicrobial agent isolated from *Streptomyces lincolnensis* var. *lincolnensis* in 1962.[188] The lincosamide antibiotics, which bind to the 23S rRNA of the 50S subunit, act as the structural mimics of the 3'-ends of L-Pro-Met-tRNA and deacylated-tRNA in the initial phase of the peptide elongation cycle.[189] Thus, lincomycin and its derivatives exhibit biological activity against Gram-positive bacteria by blocking microbial protein synthesis.[190] For example, the semi-synthetic chlorinated lincomycin derivative, clindamycin (7-epichlorolincomycin, **5-3**), possesses improved biological activitiy and is a clinically useful antibiotic.[191-193] The structure of lincomycin A (**5-1**) is composed of an *N*-methyl-4-propyl-L-proline moiety and an unusual eight-carbon sugar, which contains a unique methylthio group at C-1, called methylthiolincosamide (MTL, **5-2**; Figure 5-1).[28] Despite the long-term use of this class of antibiotics and the identification of the biosynthetic gene cluster from the industrial overproduction strain *S. lincolnensis* 78-11,[46] the biosynthetic pathway of lincomycin A (**5-1**), especially the MTL (**5-2**) moiety, remains poorly understood.[194] Thus far, only three enzymes have been biochemically characterized *in vitro*. These include LmbB1,[195,196] LmbB2,[195] and LmbJ,[197] which are involved in the biosynthesis of the *N*-methyl-4-propyl-L-proline moiety (Figure 5-2). No intermediate has been experimentally identified in the MTL biosynthetic pathway.

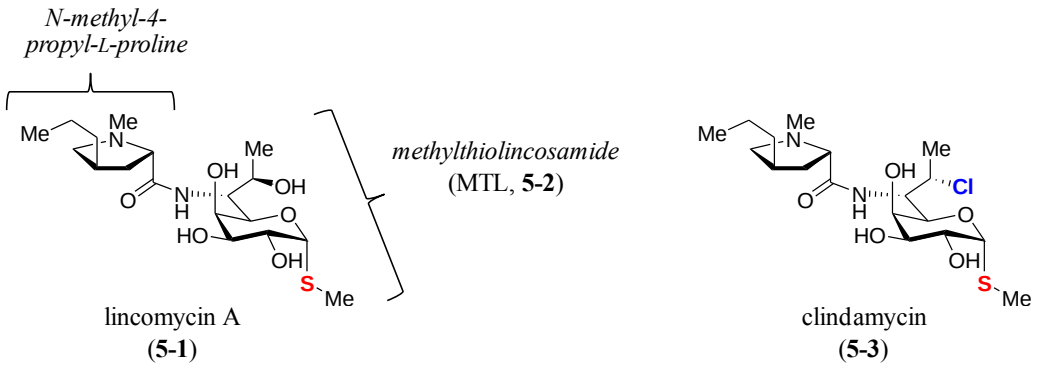


Figure 5-1. Structures of lincomycin A and its semi-synthetic derivative, clindamycin.

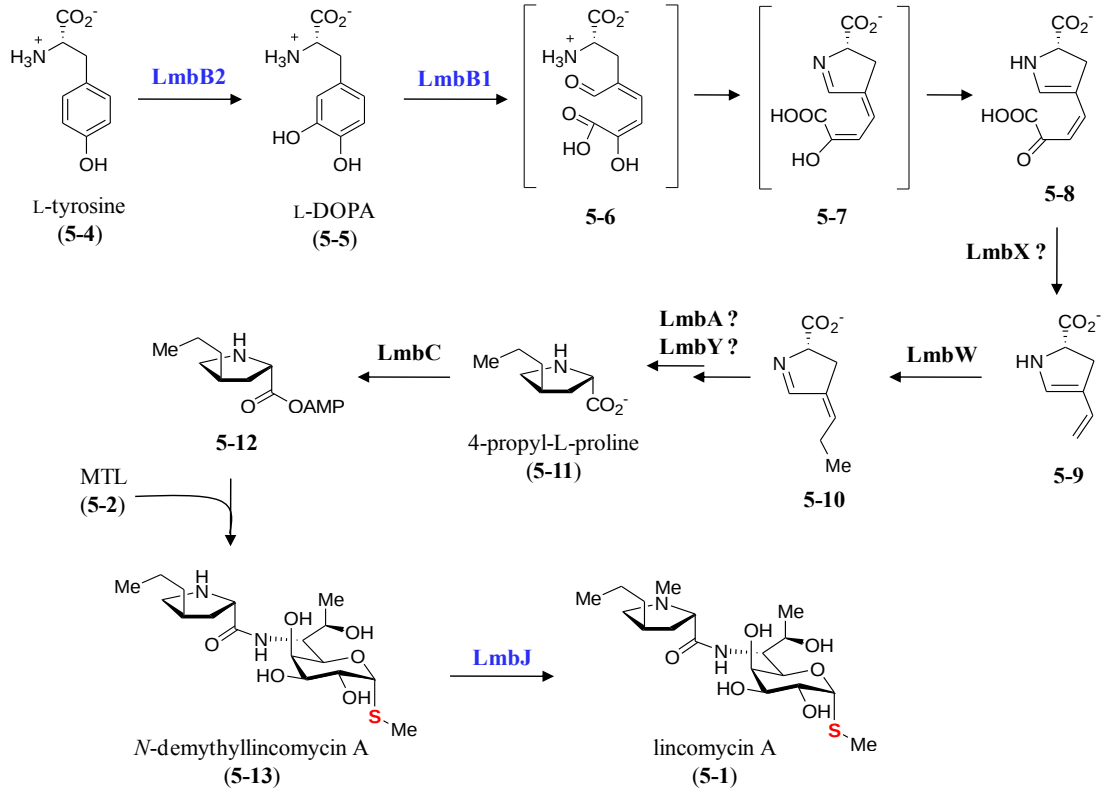


Figure 5-2. Proposed biosynthetic pathway of the 4-propyl-L-proline part of lincomycin A and its coupling with MTL to form lincomycin A.

The functions of LmbB1, LmbB2 and LmbJ were reported experimentally. The functions of other enzymes were predicted based on the sequence similarity analyses using the anthramycin, sibiromycin, tomaymycin and lincomycin gene clusters.

In 1984, before the biosynthetic gene cluster of lincomycin A (**5-1**) was discovered, feeding experiments using D-[$^{13}C_6$]glucose were performed to determine the origin of the C_8 scaffold of MTL (**5-2**). The data suggested that **5-2** is assembled through condensation of a C_5 and a C_3 units.[198] It was proposed that the C_5 unit might be a pentose 5-phosphate derived from the pentose phosphate pathway (Figure 5-3). On the other hand, the C_3 unit was proposed to be transferred from sedoheptulose 7-phosphate (**5-20**) using a transaldolase reaction analogous to the reaction seen in the pentose phosphate pathway (**5-20** + **5-21** ↔ **5-23** + **5-22**). The resulting octulose 8-phosphate was proposed to be further modified to form MTL (**5-2**) through an intermediate **5-24** (Figure 5-4, pathway A).

Meanwhile, the biosynthetic gene cluster of lincomycin A (**5-1**) was sequenced in 1995.[46] A few genes, such as *lmbO*, *lmbM*, and *lmbS*, were found to share sequence similarities with those involved in well-studied NDP-6-deoxyhexose pathways. Thus, another route for the MTL biosynthesis starting from TDP-glucose (C_6 unit) through a TDP-deoxyaminosugar intermediate (**5-28**), which might be then transferred to a C_2 extender unit *via* a *C*-glycosylation reaction, was proposed (Figure 5-4, pathway B).

An alternative pathway to reconcile the two contradictory pathways mentioned above was suggested after finding a gene encoding a putative transaldolase, *lmbR*, in the biosynthetic gene cluster. As proposed in pathway A, the C_8 carbon unit may be first formed from a C_5 unit and a C_3 unit *via* a transaldolase reaction. The resulting octulose 8-phosphate would then be converted to an NDP-activated pyranosidic octose intermediate and further modified by analogous reactions found in NDP-deoxyhexose biosynthetic pathways (pathway C).[194,199-201]

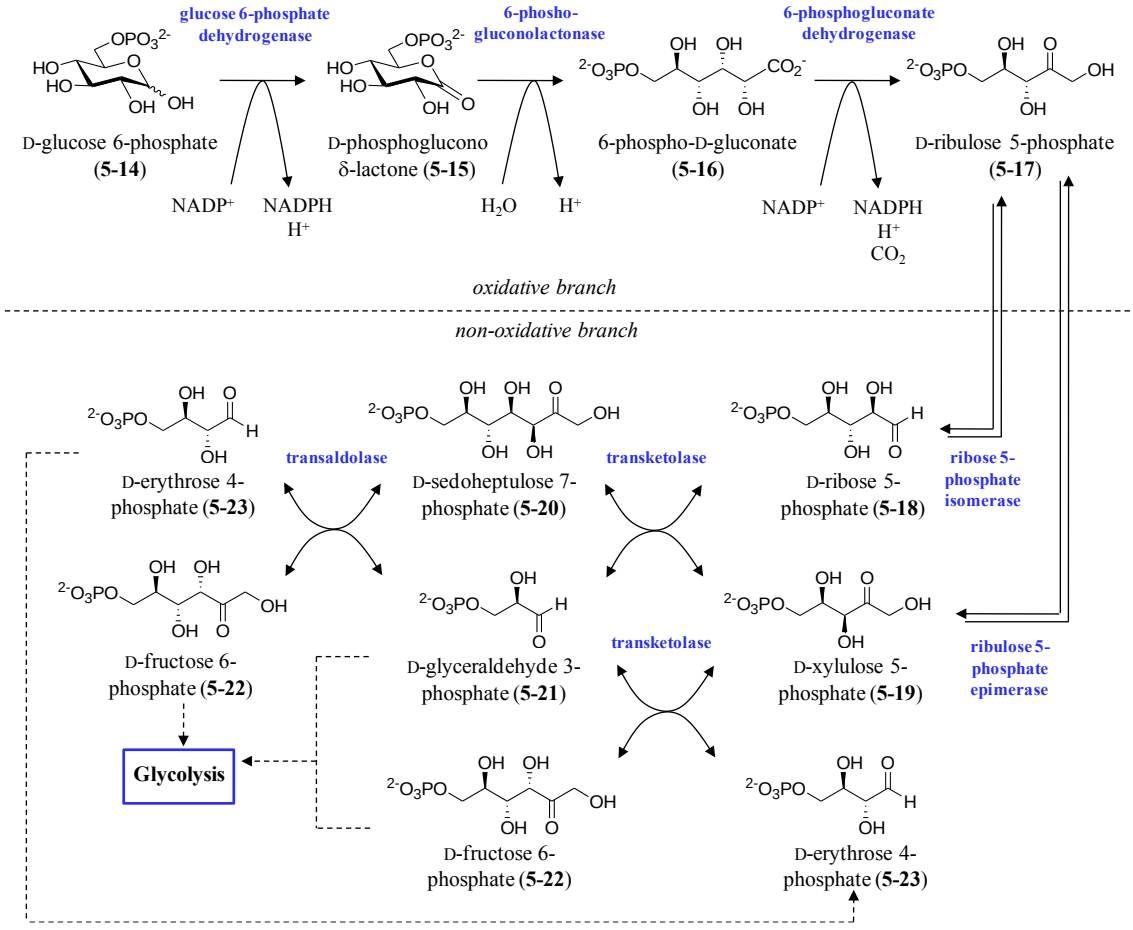


Figure 5-3. Pentose phosphate pathway.

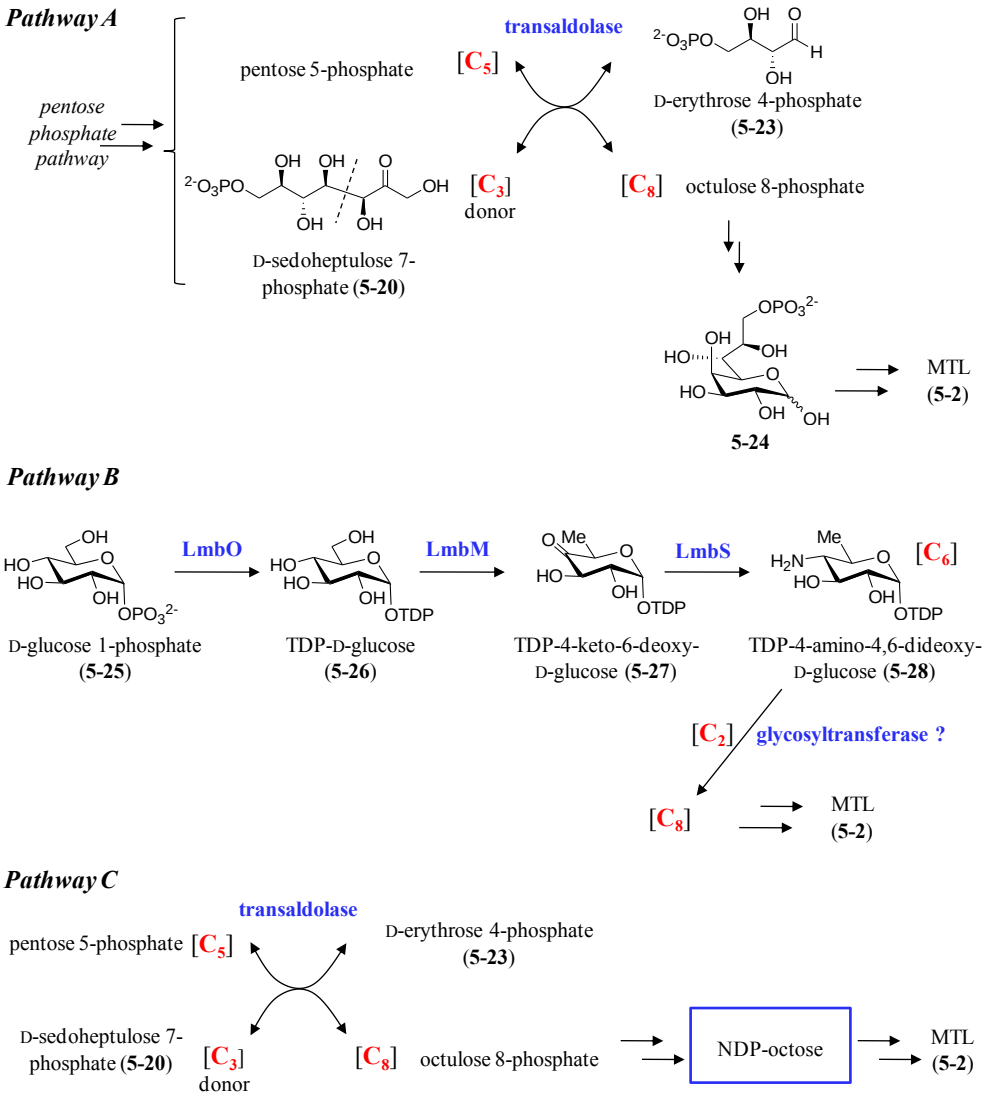


Figure 5-4. Proposed biosynthetic pathways for methylthiolincosamide (MTL, **5-2**).

(A) Pathway based on the D-[$^{13}C_6$]glucose feeding studies. (B) Pathway based on the putative NDP-hexose modification genes found in the lincomycin A biosynthetic gene cluster. (C) Pathway reconciling A and B based on the putative transaldolase gene identified in the gene cluster.

In this chapter, sequence alignment of the genes in the lincomycin A biosynthetic gene cluster using the updated database are first described. The biosynthetic pathway of

MTL (**5-2**) is proposed and discussed based on the sequence similarity results and recent reports of structurally similar NDP-heptose biosynthetic pathways.

The overall pathway of our proposal is consistent with pathway C (Figure 5-4) but is described in more detail including possible functional assignments of each enzyme. Briefly, the C_8 backbone was proposed to be constructed from a pentose 5-phosphate and a C_3 unit derived either sedoheptulose 7-phosphate or fructose 6-phosphate using the transaldolase homologue, LmbR. The resulting octulose 8-phosphate may be isomerized by a sedoheptulose 7-phosphate isomerase homologue, LmbN, to produce octose 8-phosphate. This must be further modified and activated by enzymes encoded by the genes similar to those involved in NDP-heptose biosynthetic pathways to yield NDP-octose. The subsequent modification steps for NDP-octose may be performed by enzymes homologous to well-studied NDP-hexose modification enzymes to generate the basic structure of MTL (**5-2**). Finally, C-1 sulfur incorporation is likely achieved *via* an *S*-glycosyl transfer reaction analogous to the reaction observed in the glucosinolates biosyntheses (Chapter 1, Figure 1-12).

To provide proof for the pathway proposed above, the LmbR-catalyzed transaldolase reaction is biochemically characterized since this is the key step to construct the unique C_8 backbone and, thus, its validation could help establish the entire pathway. The characterization of the LmbR reaction along with with the following LmbN-catalyzed isomerization reaction is presented. The products of enzymatic incubation were identified by comparing to the synthetic standards. In addition, enzymes involved in the later steps of MTL (**5-2**) biosynthesis were expressed, purified, and utilized for several model reactions to test their activities.

5.2 EXPERIMENTAL PROCEDURES

5.2.1 General

Materials

All chemicals and reagents were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA), and were used without further purification unless otherwise specified. Adenylate kinase (AK), pyruvate kinase (PK), lactate dehydrogenase (LDH), transketolase (TK), transaldolase (TA), glyceraldehyde 3-phosphate dehydrogenase (GAPDH) were purchased from Sigma-Aldrich Chemical Co. Calf intestinal alkaline phosphatase (CIP) was a product of New England Biolabs (Ipswich, MA). Materials used for molecular cloning were identical to those described in Chapter 2 except for the DNeasy Tissue Kit from Qiagen (Valencia, CA) used for genomic DNA extraction. Materials used for protein expression and purification were identical to those described in Chapter 3. The CarboPac PA1 high-performance liquid chromatography (HPLC) columns were obtained from Dionex (Sunnyvale, CA). Analytical C_{18} HPLC columns were purchased from Varian (Palo Alto, CA). Oligonucleotide primers were prepared by Invitrogen or Integrated DNA Technologies (Coralville, IA).

Bacterial Strains and Plasmids

Streptomyces lincolnensis NRRL ISP-5355 was obtained from the Agricultural Research Service (ARS) Culture Collection of the National Center for Agricultural Utilization Research (Peoria IL). *Escherichia coli* DH5α, acquired from Bethesda Research Laboratories (Gaithersburg, MD), was used for routine cloning experiments. The protein overexpression host *E. coli* BL21 star (DE3) was obtained from Invitrogen. Vectors pET24b(+) and pET28b(+) for protein over-expression were purchased from

Novagen (Madison, WI). Chaperone protein expression plasmids, pG-KJE8, pGro7, and pKJE7 were products of Takara Bio (Shiga, Japan). Vector pWHM3 for protein expression in *S. lividans* was kindly provided by Dr. C. R. Hutchinson, and the ermE* promoter retrieved as a 279-bp *Kpn*I–BamHI fragment was ligated into the same restriction sites of pWHM3 to generate pWHM3-ermEp.[202] Standard genetic manipulations of *E. coli* and *S. lividans* were performed as described by Sambrook et al[107] and Kiser et al,[203] respectively.

Instrumentation

UV–vis spectra were recorded using a Beckman DU 650 spectrophotometer. Evaporation of solution samples in 1.5 mL microcentfiguge-tubes was performed using a Savant Speedvac SC100 equipped with a Savant refrigerated condensation trap RT100 and Gel Pump GP100. NMR spectra were acquired on a Varian Unity 500 MHz spectrometer, and chemical shifts (in ppm) are reported relative to that of the solvent peak (δ_H = 7.24 for deuterated chloroform, δ_H = 4.65 for deuterated water, δ_H = 3.30 and δ_C = 49.0 for deuterated methanol). HPLC was performed on a Beckman Coulter System Gold equipped with a UV detector or a Corona CAD (charged aerosol detector). DNA concentrations were measured using a NanoDrop ND-1000 UV−vis instrument from Thermo Fisher Scientific. DNA sequencing was performed by the core facility of the Institute of Cellular and Molecular Biology at the University of Texas, Austin. Vector NTI Advance 10.1.1 from Invitrogen was used for sequence alignments. Mass spectroscopy was performed at the Mass Spectrometry core facility in the Department of Chemistry and Biochemistry and in the College of Pharmacy at the University of Texas, Austin.

5.2.2 Protein Sequence Analyses and Functional Predictions.

The protein sequences encoded in the lincomycin biosynthetic gene cluster from *Streptomyces lincolnensis* ATCC 25466 (accession number EU124663) were analyzed using the National Center for Biotechnology Information (NCBI) database by the basic local alignment search tool (BLAST).

5.2.3 *S. Lincolnensis* Genomic DNA Extraction.

Bacterial cells of *S. lincolnensis* NRRL ISP-5355 (identical to the ATCC 25466 strain) in a dormant state were inoculated into 7 mL of International *Streptomyces* Project (ISP) medium 2 and grown in a rotary incubator at 30 °C and 250 rpm for 16 h. The resultant seed culture (1 mL) was transferred to 50 mL of ISP-2 medium and grown under the same conditions for 2 days. The mycelia were harvested by centrifugation at 5000 × *g* for 10 min. The cells were resuspended in 10 mL of TES buffer (25 mM Tris, 25 mM EDTA, and 10.3% sucrose, pH8) and treated with lysozyme (15 mg) at 37 ºC and 250 rpm for 1 h. To the resulting mixture was added SDS (final 0.3%) and proteinase K (1.5 mg) and the combined solution was incubated under the same conditions for an additional 1 h. The released genomic DNA and RNA mixture was isolated by phenol-chloroform extraction followed by sodium acetate-isopropyl alcohol precipitation. The resulting solution was treated with ribonuclease (RNase), and the genomic DNA was isolated by another round of phenol-chloroform extraction and sodium acetate-isopropyl alcohol precipitation. The final concentration of the genomic DNA obtained was 160 ng/μL in 200 μL of 10 mM Tris·HCl buffer (pH 8.5).

5.2.4 Cloning of the Methylthiolincosamide Biosynthetic Genes.

The *lmbF*, *lmbG*, *lmbK*, *lmbL*, *lmbM*, *lmbN*, *lmbZ*, *lmbP*, *lmbO*, *lmbS*, *lmbR*, *lmbT, and lmbV* genes was PCR-amplified from *S. lincolnensis* genomic DNA using

primers with engineered *Nde*I and *Hind*III restriction sites. The sequences of the primers are shown in Table 5-1. The PCR-amplified gene fragments were purified, digested with *Nde*I and *Hind*III, and ligated into both pET24b(+) and pET28b (+) vectors, which had also been digested with the same restriction enzymes. The resulting plasmids were sequenced using the T7 or T7 terminal universal primer and used to transform the *E. coli* BL21 star (DE3) strain for protein overexpression.

The plasmids, *lmbP*/pET28b(+), *lmbO*/pET28b(+), *lmbS*/pET28b(+), *lmbT*/pET28b(+), and *lmbV*/pET28b(+), were digested with *Xba*I and *Hind*III restriction sites and ligated into pWHM3-ermEp vector that was also digested with the same enzymes. The resulting plasmids were used to transform *S. lividans* using a protoplast method for protein overexpression.[203]

Table 5-1. Primers used for the construction of the plasmids containing methylthiolincosamide biosynthetic genes.

Primer name	Sequence
lmbF for pET24/28b(+)-forward	5'-TACTAAC**CATATG**ACCGCCACGGCGAGC-3'
lmbF for pET24b(+)-reverse	5'-AGAT**AAGCTT**CCGGTACCGCCACTCGG-3'
lmbF for pET28b(+)-reverse	5'-ATAT**AAGCT*T***CA*CCGGTACCGCCACTCG-3'
lmbG for pET24/28b(+)-forward	5'-TACTAAC**CATATG**CGGGACTACCGTCTCTGGC-3'
lmbG for pET24b(+)-reverse	5'-GGAT**AAGCTT**TCGGGTCGCCCCCGCCG-3'
lmbG for pET28b(+)-reverse	5'-ATAT**AAGCTT**GGATCAGCCAGCCGTCGTCG-3'
lmbK for pET24/28b(+)-forward	5'-ATGTGAA**CATATG**GGACGCGAGGGACAGTCG-3'
lmbK for pET24b(+)-reverse	5'-AGAT**AAGCTT**GCGGCCGCCGTGCCCG-3'
lmbK for pET28b(+)-reverse	5'-AGAT**AAGCTT**GGTGACCGGTGCGCATGC-3'
lmbL for pET24/28b(+)-forward	5'-AGTTAAG**CATATG**ACCGACGCGACGCACC-3'
lmbL for pET24b(+)-reverse	5'-ATAT**AAGCTT**CGGGGCGGCATGCG-3'
lmbL for pET28b(+)-reverse	5'-GTAC**AAGCT*T***CA*CGGGGCGGCATGC-3'
lmbM for pET24/28b(+)-forward	5'-AGTTAAG**CATATG**AGCGGGCGGTACTGC-3'
lmbM for pET24b(+)-reverse	5'-ATAT**AAGCTT**CGGCCGCTGCCACC-3'
lmbM for pET28b(+)-reverse	5'-GTAC**AAGCT*T***CA*CGGCCGCTGCC-3'
lmbN for pET24/28b(+)-forward	5'-ATCACAA**CATATG**AGCACTCTGGACGAGGTCC-3'
lmbN for pET24b(+)-reverse	5'-ACAT**AAGCTT**TGCCGGCTCCCCCGC-3'
lmbN for pET28b(+)-reverse	5'-AGAT**AAGCTT***TCA*TGCCGGCTCCCCC-3'
lmbZ for pET24/28b(+)-forward	5'-ATCACAA**CATATG**ACCCACAGGTGCGGC-3'
lmbZ for pET24b(+)-reverse	5'-TGAT**AAGCTT**CGGTTTCTCCCAGGTGAGG-3'
lmbZ for pET28b(+)-reverse	5'-ACAT**AAGCT*T***CA*CGGTTTCTCCCAGGTGAGG-3'
lmbP for pET24/28b(+)-forward	5'-ATCACTA**CATATG**ATCGACGTCACGGCGCC-3'
lmbP for pET24b(+)-reverse	5'-TCAT**AAGCTT**TGCCCGCCCCTCTCCG-3'
lmbP for pET28b(+)-reverse	5'-TCAT**AAGCTT***TCA*TGCCCGCCCCTCTCC-3'
lmbO for pET24/28b(+)-forward	5'-ATCACTA**CATATG**GTCGCGTCGACCGAACC-3'
lmbO for pET24b(+)-reverse	5'-ATCT**AAGCTT**TTTCTCACCTGTCCGCAGATAGC-3'
lmbO for pET28b(+)-reverse	5'-TCAT**AAGCTT**CGA*TCA*TTTCTCACCTGTCCGC-3'
lmbS for pET24/28b(+)-forward	5'-GGCTAAG**CATATG**AGCGACTACATCCCCTTCG-3'
lmbS for pET24b(+)-reverse	5'-ATAA**AAGCTT**CCGTCCCGCCTCGTCG-3'
lmbS for pET28b(+)-reverse	5'-ATAA**AAGCT*T***CA*CCGTCCCGCCTCG-3'
lmbR for pET24/28b(+)-forward	5'-AACCGCC**CATATG**AAGATCTTTCTGGATACGCC-3'
lmbR for pET24b(+)-reverse	5'-AACT**AAGCTT**CTCGTACTTCACCCCCGC-3'
lmbR for pET28b(+)-reverse	5'-ATTC**AAGCT*T***CA*CTCGTACTTCACCCCCG-3'
lmbT for pET24/28b(+)-forward	5'-AGTTAAG**CATATG**ACGGCGCGGACGG-3'
lmbT for pET24b(+)-reverse	5'-GGAT**AAGCTT**TGACACCTCCGCCAGC-3'
lmbT for pET28b(+)-reverse	5'-GTCG**AAGCT*T***CA*TGACACCTCCGCCAGC-3'
lmbV for pET24/28b(+)-forward	5'-AGTTAAG**CATATG**CAGCGCAAGGGACTGG-3'
lmbV for pET24b(+)-reverse	5'-ATAT**AAGCTT**CGGCCCCACCAGCACC-3'
lmbV for pET28b(+)-reverse	5'-GTAA**AAGCT*T***CA*CGGCCCCACCAGC-3'

[a] The engineered restriction sites are shown in bold, the start codon is shown in bold and also underlined, and the stop codon is italic.

5.2.5 Expression and Purification of the Methylthiolincosamide Biosynthetic Proteins.

Protein Overexpression in E. Coli

A typical procedure for the overexpression and protein purification of the *N*-His_6-tagged or *C*-His_6-tagged MTL (**5-2**) biosynthetic proteins are shown below. An overnight culture of an *E. coli* BL21 star (DE3) transformant, grown in the LB medium (10 mL) containing 50 μg/mL of kanamycin at 37 °C, was used to inoculate 1 L of the same growth medium. The culture was incubated at 37 °C with shaking (230 rpm) until the OD_{600} reached ~0.5. Protein expression was then induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM, and the cells were allowed to grow at 18 °C and 125 rpm for an additional 24 h. The cells were harvested by centrifugation at 4500 × *g* for 15 min and stored at –80 °C until lysis. All purification steps were carried out at 4 °C using Ni-NTA resin according to the manufacturer's protocol with minor modifications. Specifically, the thawed cells (~5 g) were resuspended in the lysis buffer (20 mL) containing 10% (v/v) glycerol and 10 mM imidazole. After incubation with lysozyme (20 mg) for 30 min, the cells were disrupted by sonication using 10 × 10-s pulses with a 30-s cooling pause between each pulse. The resulting lysate was centrifuged at 20,000 × *g* for 20 min, and the supernatant was subjected to Ni-NTA chromatography. Bound protein was eluted using 250 mM imidazole buffer containing 10% glycerol. The collected protein solution was dialyzed against 3 × 1-L of 50 mM Tris·HCl buffer (pH 8) containing 300 mM NaCl and 15% glycerol. The protein solution was then flash-frozen in liquid nitrogen and stored at −80 °C until use. Protein concentration was determined by the Bradford assay using bovine serum albumin as the standard.[149] The proteins purified according to this method and their yields from 1 L cultures are *N*-His_6-LmbF (~10 mg), *N*-His_6-LmbG (~20 mg),

N-His_6-LmbK (~10 mg), *C*-His_6-LmbL (~15 mg), *C*-His_6-LmbM (~15 mg), and *C*-His_6-LmbN (~40 mg). The molecular mass and purity of the proteins were determined by SDS-PAGE analysis.

Protein Coexpression with Chaperone Proteins in E. Coli

Since protein overexpression of LmbZ, LmbP, LmbO, LmbS, LmbR, LmbT, and LmbV led to the formation of inclusion bodies, they were not further purified. Instead, coexpression with the chaperone proteins was tried for these proteins. Plasmids pG-KJE8 (encoding chaperones, DnaK, DnaJ, GrpE, GroES and GroEL), pGro7 (encoding chaperones, GroES and GroEL), or pKJE7 (encoding chaperones, DnaK, DnaJ, and GrpE) were used to cotransform *E. coli* BL21 star (DE3) strain with the target protein overexpression plasmids. The overexpression of chaperone proteins was induced by L-arabinose and tetracycline according to the manufacture's instructions. *C*-His_6-LmbR was obtained in soluble form by this method and was purified using Ni-NTA resin as described above. However, the other proteins remained practically insoluble under the conditions. The yield of *C*-His_6-LmbR was ~40 mg from 1 L culture. The molecular mass and purity of the protein were determined by SDS-PAGE analysis.

Protein Overexpression in S. Lividans

LmbP, LmbO, LmbS, LmbT, and LmbV proteins were also heterologously expressed with an *N*-terminal His_6 tag in *S lividans* using a plasmid constructed from pWHM3-ermEp as described above. The *S. lividans* transformants selected by thiostrepton on the R1R2 agar plates were used to inoculate YEME medium (10 mL) containing 10 μg/mL of thiostrepton, and the culture was incubated at 30 °C and 230 rpm for 3 days. The resulting seed culture was transferred to 1 L of YEME medium containing 5 μg/mL of thiostrepton, and incubated under the same conditions for 3 days.

The cells were harvested by centrifugation at 6000 × *g* for 20 min and stored at –80 °C until lysis. All purification steps were similarly carried out at 4 °C using Ni-NTA resin as described for the protein purification from *E. coli*. While *N*-His_6-LmbS and *N*-His_6-LmbT were obtained in good purity, the other proteins were either not expressed well or were difficult to purify. The yields of *N*-His_6-LmbS and *N*-His_6-LmbT were ~10 mg and ~6 mg from 1 L cultures, respectively. The molecular mass and purity of the protein were determined by SDS-PAGE analysis.

5.2.6 ESI-MS Analysis of LmbR.

The purified *C*-His_6-LmbR protein stock solution was added to 50 mM Tris·HCl buffer (pH 8.0) (final concentration of protein was 40 μM) and incubated with D-fructose 6-phosphate (**5-22**) at 30 °C for 15 min. The resulting solution (200 μL) was cooled on ice. To this solution was added 20 μL of cold $NaBH_4$ solution in methanol (530 mM; final concentration of $NaBH_4$ was 50 mM). The mixture was incubated on ice for 1 h. The excess salts and glycerol were removed at 4 °C using an Amicon ultrafiltration unit through a YM-10 membrane. The resulting protein solution was subjected to ESI-MS analysis. As a control, the purified *C*-His_6-LmbR protein (40 μM) in 50 mM Tris·HCl buffer (pH 8.0) in the absence of **5-22** was subjected to ESI-MS analysis without $NaBH_4$ treatment.

5.2.7 Synthesis of D-Xylose 5-Phospahte.

As shown in Figure 5-5, D-xylose 5-phospahte (**5-29**) was synthesized following previously reported procedures. The ^{1}H NMR spectral data of **5-29** and all the synthetic intermediates are consistent with those reported in the literatures.[204-207]

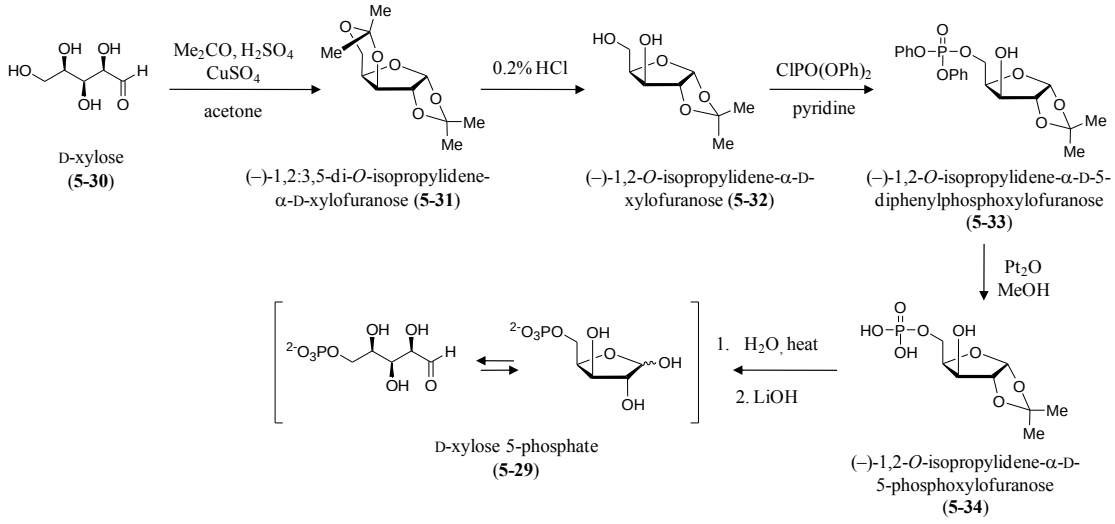


Figure 5-5. Synthetic scheme of D-xylose 5-phosphate.

5.2.8 LmbR and LmbN Reaction Assays.

LmbR Reaction Using D-Ribiose 5-Phospahte and D-Fructose 6-Phosphate

The purified *C*-His$_6$-LmbR protein (40 μM final concentration) was incubated with D-ribose 5-phosphate (**5-18**, 10 mM), D-fructose 6-phosphate (**5-22**, 10 mM), and $MgCl_2$ (1 mM) in 50 mM glycylglycine buffer (pH 8.0) at 30 °C for 6 h. A series of control reactions omitting the enzyme, **5-18**, or **5-22** was similarly prepared. The reaction mixture (50 μL each) was filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the protein. HPLC analysis was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm). The sample (6 μL each) was eluted with 100 mM NH_4OAc at 1 mL/min and monitored by a Corona CAD. The HPLC fraction containing the product peak (retention time ~10 min) was collected (6 μL × 5 injections). The collected fraction was lyophilized and redissolved in deionized water (250 μL) prior to ESI-MS analysis (negative ion detection mode).

LmbN Reaction Using D-Ribiose 5-Phospahte and D-Fructose 6-Phosphate

The LmbR reaction prepared as described above was filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the protein. To the resulting mixture was added purified *C*-His_6-LmbN (50 μM final concentration), and this reaction was incubated at 30 °C for 3 h. The reaction mixture was filtered through a YM-10 membrane to remove the protein prior to HPLC analysis. The HPLC conditions were the same as described for the LmbR reaction. The LmbN product peak (retention time ~8 min) was similarly collected and subjected to ESI-MS analysis (negative ion detection mode).

LmbR and LmbN Reaction Using D-Xylose 5-Phospahte and D-Fructose 6-Phosphate

The purified *C*-His_6-LmbR protein (40 μM final concentration) was incubated with D-xylose 5-phosphate (**5-29**, 10 mM), D-fructose 6-phosphate (**5-22**, 10 mM), and $MgCl_2$ (2 mM) in 50 mM glycylglycine buffer (pH 8.0) at 30 °C for 9 h. Similarly, both *C*-His_6-LmbR (40 μM) and *C*-His_6-LmbN (50 μM) were incubated together with D-xylose 5-phosphate (**5-29**, 10 mM), D-fructose 6-phosphate (**5-22**, 10 mM) and $MgCl_2$ under the same conditions. As a control, **5-29** was also replaced by D-ribose 5-phosphate (**5-18**, 10 mM). Each reaction mixture was filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the protein. The samples were subjected to HPLC analyses under the same conditions as described above. The LmbR product peak (retention time ~7.5 min) was collected and subjected to ESI-MS analysis (negative ion detection mode).

LmbR Reaction Using Other Substrates

Dihydroxyacetone (**5-35**), D-fructose (**5-36**), and D-sedoheptulose 7-phosphate (**5-20**) were investigated as alternative C_3 sugar donors for the LmbR reaction. The former

two compounds (10 mM or 50 mM each) simply replaced D-fructose 6-phosphate (**5-22**, 10 mM) in the standard LmbR reaction described above. In the latter case, D-sedoheptulose 7-phosphate (**5-20**) was prepared from D-fructose 6-phosphate (**5-22**, 10 mM) and β-hydroxypyruvate (**5-37**, 10 mM) using transketolase from *E. coli* (0.04 units) in 50 mM glycylglycine buffer (pH 8.0) containing $MgCl_2$ (1 mM) and thiamine pyrophosphate (0.5 mM). The reaction mixture (50 μL) was incubated at 30 ºC for 8 h and filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the protein. To this solution containing the transketolase product, D-sedoheptulose 7-phosphate (**5-20**, final 2.5 mM), was added *C*-His_6-LmbR (40 μM) and D-ribose 5-phosphate (**5-18**, 6.6 mM), and this reaction was incubated at 30 ºC for 3 h. HPLC analyses were similarly performed as described for the standard LmbR reaction.

LmbN Reaction Using D-Sedoheptulose 7-Phosphate

The LmbN reaction was also investigated using D-sedoheptulose 7-phosphate (**5-20**) as a potential substrate. Compound **5-20** was produced *in situ* from D-fructose 6-phosphate (**5-22**, 10 mM) and β-hydroxypyruvate (**5-37**, 20 mM) using transketolase from *E. coli* (0.04 units) and thiamine pyrophosphate (0.5 mM) in 100 mM glycylglycine buffer (pH 8.0) containing $MgCl_2$ (1 mM) and LmbN (50 μM). The reaction mixture (50 μL) was incubated at 30 ºC for 2 h. and filtered through a YM-10 membrane to remove the proteins prior to HPLC analysis. The HPLC conditions were the same as described above.

5.2.9 Acetyl derivatization of the LmbR and LmbN Products.

The purified enzymes *C*-His_6-LmbR and *C*-His_6-LmbN were dialyzed against 100 mM $NH_4 \cdot HCO_3$ buffer (pH 8.0) containing 10 mM $MgCl_2$ at 4 ºC to remove glycerol and salts present in the enzyme stock solution. To this enzyme solution were added D-ribose

5-phosphate (**5-18**, 10 mM) and D-fructose 6-phosphate (**5-22**, 50 mM). The final concentration of *C*-His_6-LmbR and *C*-His_6-LmbN were 40 μM and 50 μM, respectively. As a control, the reaction containing only *C*-His_6-LmbR and the sugar substrates was similarly prepared. Each reaction mixture (400 μL) was incubated at 30 °C for 4.5 h and an additional 3.5 h at 37 °C. To the mixture was then added calf intestinal alkaline phosphatase (CIP) (4 μL, 40 units), and the reaction was incubated at 37 °C for 12 h. To the resulting solution was added MeOH (400 μL) and CH_3CN (800 μL), and then it was mixed well. The resultant precipitates were removed by centrifugation at 16,000 × *g* for 5 min. The solution (400 μL) was evaporated *in vacuo* using a Speedvac SC100 (Savant). The resultant residue was completely dried and then treated with anhydrous pyridine (20 μL) and acetic anhydride (20 μL). This mixture was incubated at room temperature for 12 h, and quenched by the addition of deionized water (150 μL). As a control, D-fructose (**5-36**, 1 mg, ~5 μmol) was similarly treated with pyridine and acetic anhydride. The acetylated products were extracted with ethyl acetate (150 μL) and subjected to HPLC analysis using an analytical C_{18} column (4 × 250 mm). The sample (5 μL) was eluted with a gradient of water (solvent A) and 80% acetonitrile (solvent B). The gradient was run from 40 to 60% B over 15 min, 60−100% B over 5 min with a 3 min wash at 100% B, 100−40% B over 2 min, followed by re-equilibration at 40% B for 5 min. The flow rate was 1 mL/min, and a corona CAD was used for the detection. Authentic standards of heptaacetyl octose derivatives were synthesized by another graduate student, Chia-I Lin, and the procedure will be published elsewhere. The peaks corresponding to the enzymatic reaction products were isolated and subjected to ESI-MS analyses.

5.2.10 Spectrophotometric Analysis of the LmbR Reaction.

LmbR or an authentic transaldolase (TA) from yeast was used for the assay. The reaction solution (100 μL) contained a sugar acceptor (D-ribose 5-phosphate (**5-18**) or D-erythrose 4-phosphate (**5-23**), 150 μM), D-fructose 6-phosphate (**5-20**, 150 μM), nicotinamide adenine dinucleotide (oxidized form, NAD^+, 500 μM), and glyceraldehyde 3-phosphate dehydrogenase (GAPDH, 1 unit) in 50 mM sodium phosphate buffer (pH 7.8) comtaining 10 mM $MgCl_2$. The reaction was initiated by the addition of LmbR (1 μL, 5 μM final) or TA (1 μL, 0.05 units) into the reaction mixture. Absorbance at 340 nm was monitored by a UV-vis spectrophotometer for 5 min.

5.2.11 LmbM Reaction Assay.

LmbM reaction

The typical LmbM reaction mixture (100 μL) contained purified *C*-His_6-LmbM (final 20 μM), UDP-glucuronic acid (**5-38**, 5 mM), and nicotinamide adenine dinucleotide (oxidized form, NAD^+, 12 mM) in 100 mM Tris·HCl (pH 7.5). The mixture was incubated at room temperature for 1h. A half of the mixture (50 μL) was filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the protein. The remaining half was incubated at room temperature for additional 11 h and then filtered through a YM-10 membrane. HPLC analysis was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm). The samples (5 μL each) were eluted with a gradient of water (solvent A) and 1 M NH_4OAc (solvent B). The gradient was run from 0 to 100% B over 15 min and 100−0% B over 5 min, followed by re-equilibration at 0% B for 5 min. The flow rate was 1 mL/min, and the detector was set at 260 nm.

Isolation of the LmbM Product

A large scale reaction (3 mL) was prepared as described above. HPLC analysis was performed using a Dionex CarboPac PA1 semiprep column (9 × 250 mm). The sample (300 μL × 8 injections) was eluted with a gradient of water (solvent A) and 1 M $NH_4 \cdot HCO_3$ (solvent B). The gradient program was as follows: 20% B for 10 min, 20−100% B over 10 min, and 100−20% B over 5 min, followed by re-equilibration at 20% B for 5 min. The flow rate was 5 mL/min, and the detector was set at 260 nm. The HPLC fraction containing the product peak (retention time ~8 min) was collected. The collected fraction was lyophilized, redissolved in D_2O (650 μL), and subjected to ^{1}H-NMR and ESI-MS (negative mode) analyses.

Hydroxylamine Derivatization

The LmbR reaction prepared above (12 h incubation) was filtered through a YM-10 membrane to remove the protein and treated with hydroxylamine (300 mM final concentration). The resulting mixture was incubated at room temperature for 1 h and subjected to HPLC anlaysis. The HPLC conditions were the same as described above.

5.3 RESULTS AND DISCUSSION

5.3.1 Proposal of the Biosynthetic Pathway.

Lincomycin Biosynthetic Gene Cluster and BLAST Analyses

The biosynthetic gene cluster of lincomycin A (**5-1**) was first identified from the genomic DNA of the industrial overproduction strain *S. lincolnensis* 78-11 in 1995.[46] The gene cluster was also found in another *S. lincolnensis* strain ATCC 25466, reported in 2008 (Figure 5-6A).[208] Overall, the organization of the two reported gene clusters is very similar although several hundred point mutations were found. Integration of the cosmid

bearing the latter gene cluster into the chromosomes of *S. coelicolor* CH 999 and *S. coelicolor* M 145 led to the heterologous production of lincomycin A (**5-1**) in these non-producing strains, suggesting that the genes in the cloned cluster should be sufficient to enable lincomycin A production. Moreover, hybridization analysis of the biosynthetic gene cluster of the structurally similar celesticetin (**5-39**) from *S. caelestis* strain ATCC 1584 was successfully performed using probes designed according to the lincomycin biosynthetic gene sequences.[47] Recently, the complete sequence of the celesticetin gene cluster became available in the NCBI database (Figure 5-6B). Since 1-thiolincosamide is a common structural component shared by lincomycin A (**5-1**) and celeticetin (**5-39**), the genes responsible for the construction of this moiety likely exist in both biosynthetic gene clusters (shared genes are shown in blue in Figure 5-6, sequence identities and similarities are shown in Table 5-2).

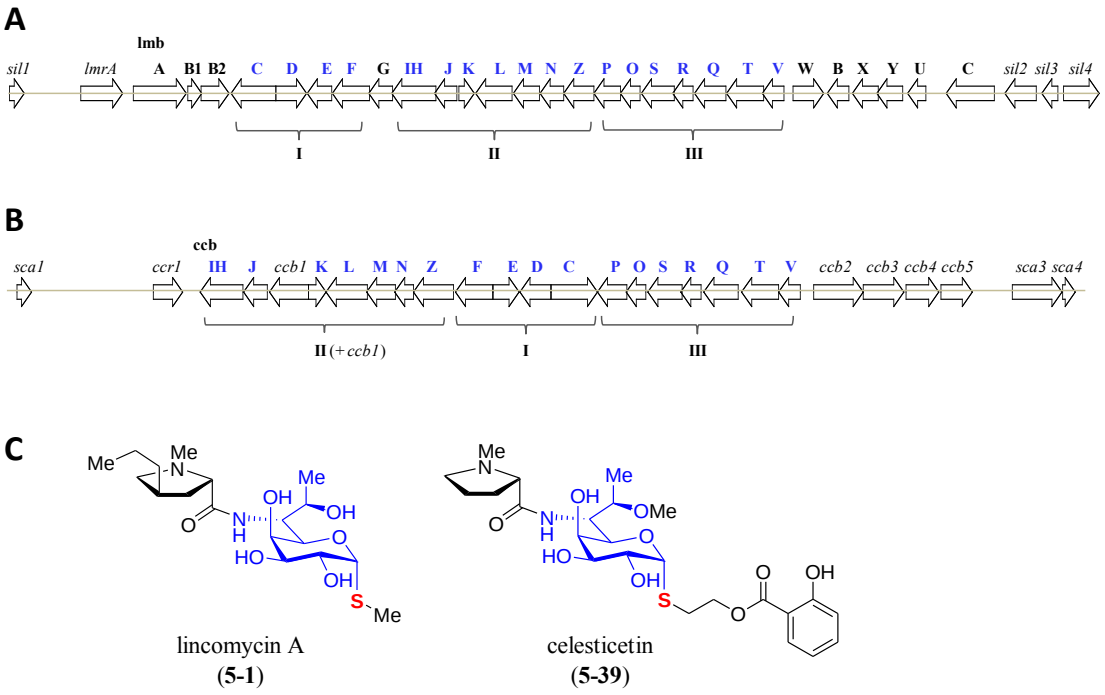


Figure 5-6. Organization of the lincomycin and celesticetin biosynthetic gene clusters.

(A) The lincomycin biosynthetic gene cluster from *S. lincolnensis* ATCC 25466 (EU124663, 38217 bp). (B) The celesticetin biosynthetic gene cluster from *S. caelestis* strain ATCC 1584 (GQ844764, 36543 bp). Homologous genes found in both clusters are shown in blue. (C) Structures of lincomycin A and celesticetin. The 1-thiolincosamide structure is shown in blue.

Table 5-2. Seqeuce similarity analysis of putative 1-thiolincosamide biosynthetic proteins in lincomycin A and celesticetin pathways

Lincomycin A biosynthetic protein[a]	protein accession number	Celesticetin biosynthetic protein[b]	protein accession number	identity / similarity (%)
LmbC	ABX00600	CcbC	ADB92568	57 / 67
LmbD	ABX00601	CcbD	ADB92567	56 / 65
LmbE	ABX00602	CcbE	ADB92566	60 / 71
LmbF	ABX00603	CcbF	ADB92565	40 / 55
LmbIH	ABX00605	CcbIH	ADB92568	63 / 72
LmbJ	ABX00606	CcbJ	ADB92568	60 / 71
LmbK	ABX00607	CcbK	ADB92557	64 / 72
LmbL	ABX00608	CcbL	ADB92558	63 / 75
LmbM	ABX00609	CcbM	ADB92560	74 / 83
LmbN	ABX00610	CcbN	ADB92561	72 / 80
LmbZ	ABX00611	CcbZ	ADB92562	64 / 75
LmbP	ABX00612	CcbP	ADB92563	59 / 70
LmbO	ABX00613	CcbO	ADB92564	65 / 75
LmbS	ABX00614	CcbS	ADB92571	74 / 81
LmbR	ABX00615	CcbR	ADB92572	69 / 80
LmbQ	ABX00616	CcbQ	ADB92573	44 / 54
LmbT	ABX00617	CcbT	ADB92574	65 / 76
LmbV	ABX00618	CcbV	ADB92575	57 / 69

[a] Putative lincomycin A biosynthetic proteins found in *S. lincolensis* ATCC 25466 strain. [b] Putative celesticetin biosynthetic proteins found in *S. caelestis* ATCC1584.

Next, we repeated BLAST analyses of the genes found in the lincomycin A biosynthetic gene cluster using the updated database. Table 5-2 shows similar genes found from the NCBI database excluding strains of *S. lincolensis* and *S. caelestis* to predict the functions of the deduced gene products. The proposed biosynthetic pathways of the *N*-methyl-4-propyl-L-proline moiety and the MTL moiety are shown in Table 5-3, and Figure 5-7, respectively.

Table 5-3. BLAST analyses of the lincomycin biosynthetic gene cluster.

protein[b]	protein accession number	protein homologue and origin[a]	identity / similarity (%)	protein accession number
LmrA	ABX00596	major facilitator superfamily permease [*Streptomyces cf. griseus* XylebKG-1]	54 / 67	ZP_08226252
LmbA	ABX00597	gamma-glutamyltranspeptidase [*Streptomyces griseoflavus* Tu4000]	65 / 75	ZP_07314001
LmbB1	ABX00598	putative L-DOPA-2,3-dioxygenase [*Streptomyces achromogenes*]	59 / 66	ACN39021
LmbB2	ABX00599	putative tyrosine hydroxylase [*Streptomyces achromogenes*]	42 / 53	ACN39022
LmbC	ABX00600	amino acid adenyltransferase [*Streptomyces clavuligerus* ATCC 27064]	41 / 50	ZP_05004791
LmbD	ABX00601	No significant similarity found		
LmbE	ABX00602	mycothiol conjugate amidase Mca [*Saccharomonospora viridis* DSM 43017]	39 / 53	YP_003134979
LmbF	ABX00603	class I and II aminotransferase [*Stackebrandtia nassauensis* DSM 44728]	28 / 43	YP_003511287
LmbG	ABX00604	methyltransferase [*Chlorobium ferrooxidans* DSM 13031]	34 / 47	ZP_01385426
LmbIH	ABX00605	peptidase U62 modulator of DNA gyrase [*Dictyoglomus turgidum* DSM 6724]	37 / 53	YP_002352129
LmbJ	ABX00606	Methyltransferase [*Micromonospora aurantiaca* ATCC 27029]	41 / 56	YP_003836969
LmbK	ABX00607	D-alpha,beta-D-heptose 1,7-bisphosphate phosphatase [*Candidatus Solibacter usitatus* Ellin6076]	52 / 62	YP_821716
LmbL	ABX00608	UDP-glucose 6-dehydrogenase [*Lawsonia intracellularis* PHE/MN1-00]	35 / 50	YP_595630
LmbM	ABX00609	NAD-dependent epimerase/dehydratase [*Aciduliprofundum boonei* T469]	42 / 62	YP_003482522
LmbN	ABX00610	sedoheptulose 7-phosphate isomerase [*Sorangium cellulosum* 'So ce 56']	42 / 57	YP_001614387
LmbZ	ABX00611	oxidoreductase domain protein [*Chthoniobacter flavus* Ellin428]	35 / 51	ZP_03129380
LmbP	ABX00612	GHMP kinase [*Candidatus Solibacter usitatus* Ellin6076]	36 / 51	YP_821713
LmbO	ABX00613	putative nucleotidyl transferase [*Leptospirillum rubarum*]	35 / 50	EAY57000
LmbS	ABX00614	glutamine–scyllo-inositol transaminase [*Caldicellulosiruptor kristjanssonii* 177R1B]	42 / 61	YP_004027350
LmbR	ABX00615	Transaldolase [*Cellulomonas flavigena* DSM 20109]	45 / 62	YP_003638673
LmbQ	ABX00616	peptidase U62 modulator of DNA gyrase [*Thermotoga lettingae* TMO]	29 / 44	YP_001471158
LmbT	ABX00617	glycosyltransferase [*Yersinia aldovae* ATCC 35236]	24 / 44	ZP_04619711
LmbV	ABX00618	hypothetical protein Tbis_3508 [*Thermobispora bispora* DSM 43833]	44 / 53	YP_003654089
LmbW	ABX00619	methyltransferase [*Streptomyces refuineus* subsp. *thermotolerans*]	78 / 86	ABW71836
LmrB	ABX00620	rRNA methyltransferase PikR2 [*Streptomyces venezuelae*]	61 / 74	AAC69327
LmbX	ABX00621	phenazine biosynthesis protein PhzF family [*Micromonospora aurantiaca* ATCC 27029]	46 / 58	YP_003833811
LmbY	ABX00622	putative F-420 dependent reductase [*Streptomyces achromogenes*]	55 / 68	ACN39023
LmbU	ABX00623	hypothetical protein FRAAL4627 [*Frankia alni* ACN14a]	56 / 71	YP_714812
LmrC	ABX00624	putative ABC transporter ATP-binding protein [*Streptomyces ambofaciens* ATCC 23877]	63 / 77	CAJ88384
Sli2	ABX00625	pentapeptide repeat-containing protein [*Streptomyces viridochromogenes* DSM 40736]	31 / 45	ZP_07308783
Sli3	ABX00626	hypothetical protein SGR_277 [*Streptomyces griseus subsp. Griseus* NBRC 13350]	80 / 87	YP_001821789
Sli4	ABX00627	amino acid permease [*Streptomyces roseosporus* NRRL 15998]	83 / 90	ZP_06586134

[a.] The homologous proteins were searched in organisms excluding strains of *S. lincolensis* and *S. caelestis*. [b.] The gene products shared with the lincomycin A and the celesticetin biosynthetic gene clusters are shown in blue. The gene products shared with lincomycin A, anthramycin, and sibiromycin are shown in orange.

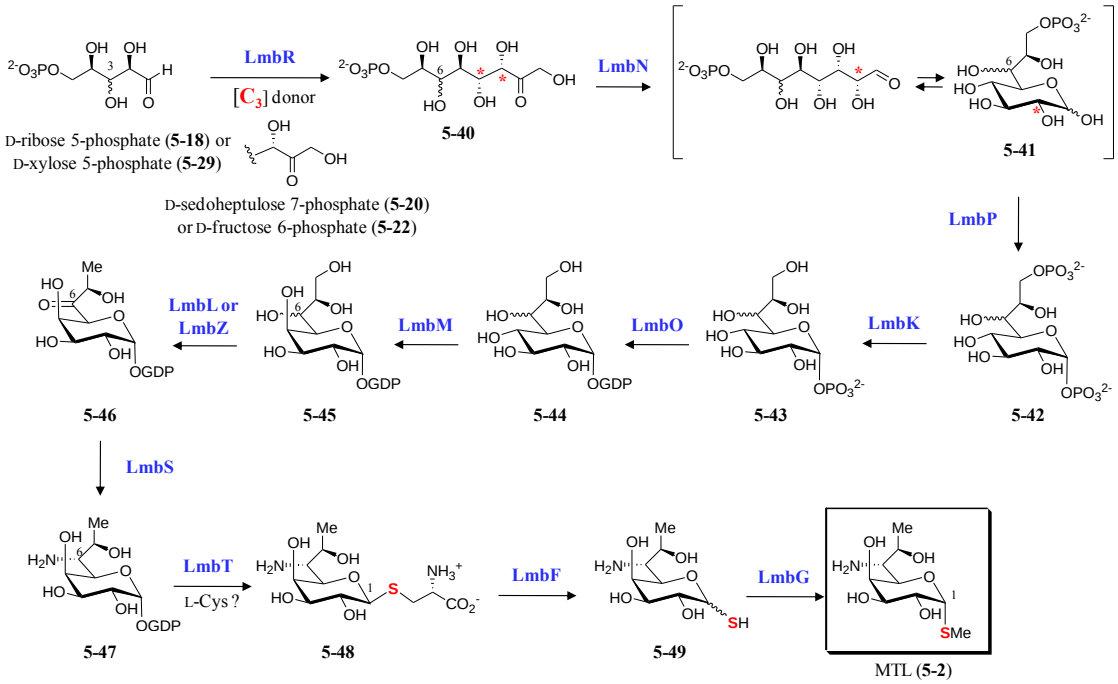


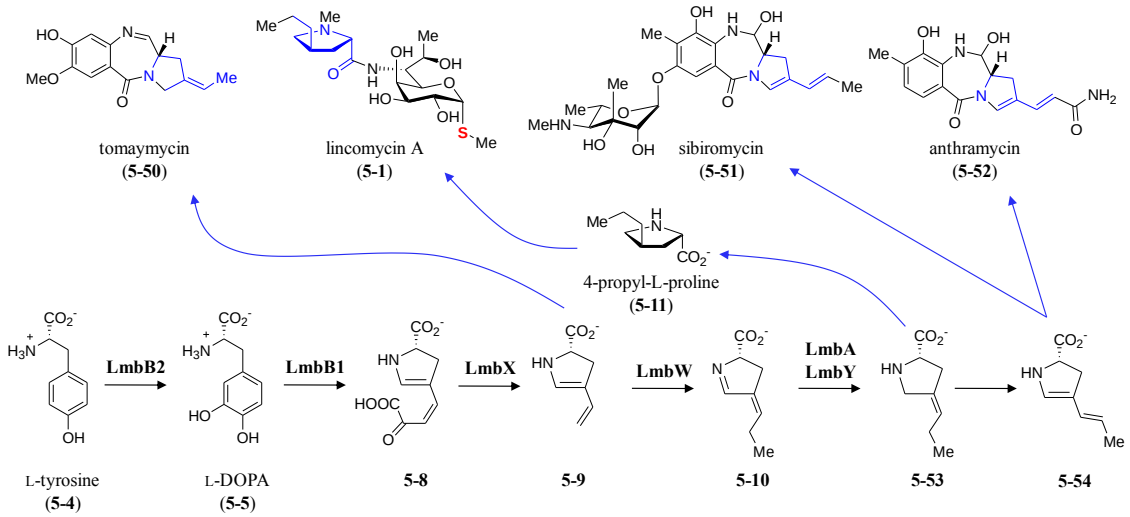
Figure 5-7. Proposed biosynthetic pathway of methylthiolincosamide (MTL).

* Predicted stereochemistry.

Biosynthetic Pathway of the 4-Propyl-L-Proline Moiety

Previous feeding studies using ^{14}C-labeled tyrosine reveled that besides the *N*-methyl and terminal aliphatic side chain, the carbon skeleton of this moiety is derived from tyrosine.[209] The first two steps of the pathway catalyzed by LmbB2 (L-tyrosine 3-hydroxylase) and LmbB1 (L-DOPA 2,3-dioxygenase) were biochemically verified (Figure 5-2; **5-4** → **5-5** → **5-6**).[195,196] Although *in vitro* characterization of the corresponding steps in the lincomycin A (**5-1**) pathway remains to be investigated, a similar route has been proposed as key reactions in the biosyntheses of the pyrrolo[1,4]benzodiazepine antibiotics anthramycin (**5-50**), sibiromycin (**5-51**) and tomaymycin (**5-52**).[210,211] Notably, the biosynthetic gene clusters of **5-50**–**5-52** were recently sequenced.[212-214] Sequence comparison of these gene clusters led to the

functional assignments of LmbX (C–C hydrolase), LmbW (methyltransferase), LmbA (γ-glutamyltransferase, likely involved in the synthesis of F-420), and LmbY (F-420-dependent reductase), which are involved in the construction of 4-propyl-L-proline (**5-11**; Figure 5-8).



	Homologous proteins found in the gene clusters					
lincomycin A [*Streptomyces lincolnensis*]	**LmbB2**	**LmbB1**	**LmbX**	**LmbW**	**LmbA**	**LmbY**
anthramycin [*Streptomyces refuineus* subsp. *thermotolerans*]	ORF13	ORF12	ORF15	ORF5	ORF6	ORF14
sibiromycin [*Streptosporangium sibiricum*]	SibU	SibV	SibS	SibZ	SibY	SibT
tomaymycin [*Streptomyces achromogenes*]	TomI	TomH	TomK	-	TomL	TomJ

Figure 5-8. Shared pathways to produce 4-propyl-L-proline derivatives in the biosyntheses of lincomycin A and the pyrrolo[1,4]benzodiazepine antibiotics.

Construction of the C_8 Backbone

As mentioned above, a transaldolase homologue, LmbR, was proposed to construct the C_8 backbone of MTL (**5-2**) from a C_5 and a C_3 units according to the early feeding experiments.[198] A candidate for the C_5 unit is D-ribose 5-phosphate (**5-18**) from

the pentose phosphate pathway. The C_3 unit donor may be D-sedoheptulose 7-phosphate (**5-20**) or D-fructose 6-phosphate (**5-22**) as observed in the physiological transaldolase reaction in the pentose phosphate pathway (Figure 5-3). The expected octulose 8-phosphate product (**5-40a**) likely inherits the stereochemistry at the C-3 and C-4 positions from the corresponding positions of the C_3 unit donor, a well-trait of the aldolase-type reactions.[215] Although aldolase-catalyzed reactions giving inversed stereochemistry in the products are also known, they are very rare.[216] The expected product possessing the retained stereochemistry would be converted to a pyranosidic octose intermediate in the later step having an equatorial hydroxyl groups at both C-3 with the (*S*)-configuration and C-4 with the (*R*)-configuration, such as seen in **5-42** (Figure 5-9). Although the final MTL (**5-2**) structure has an axial hydroxyl group with the (*S*)-configuration at C-4, a C-4 epimerization reaction catalyzed by LmbM could account for this mismatch (**5-44** $\rightarrow$ **5-45**). On the other hand, the C-6 hydroxyl group of **5-42** is expected to be replaced by an amino group *via* the transamination reaction by LmbT *via* the corresponding ketone (**5-46**). Since the stereochemical information at the C-6 position would be lost during this process, D-xylose 5-phosphate (**5-29**), which is a C-3 epimer of **5-18**, is another possible C_5 unit for the LmbR reaction yielding **5-40b** (C-6 epimer of **5-40a**; Figure 5-9). Although D-xylose 5-phosphate (**5-29**) is not a common chemical entity produced in primary metabolism, it may be derived from D-xylulose 5-phosphate (**5-19**) in the pentose phosphate pathway by an isomerase.

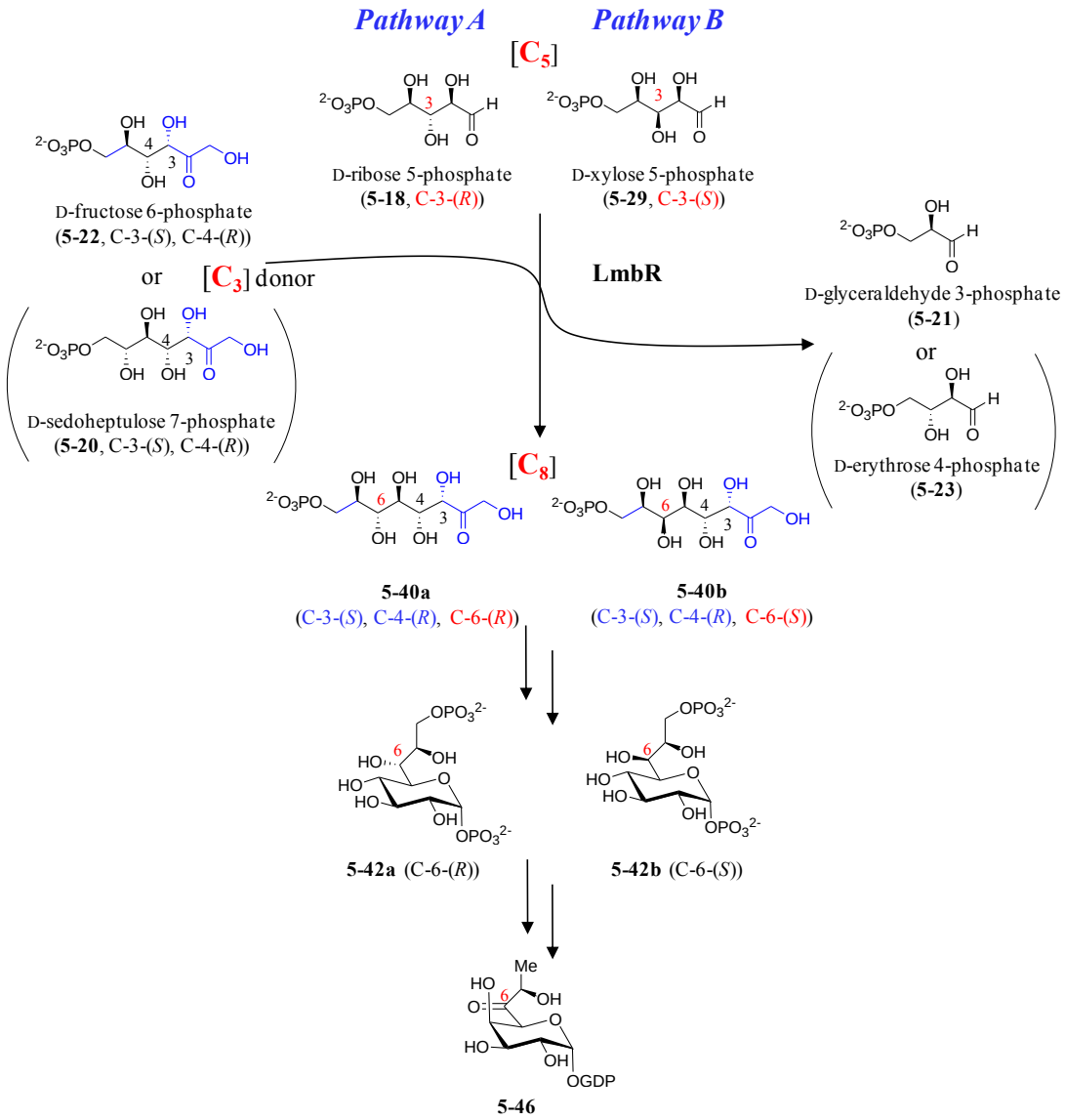


Figure 5-9. Proposed LmbR-catalyzed reaction and its stereochemical consideration.

Pathway A: using D-robiose 5-phosphate as the starting C_5 unit. *Pathway B*: using D-xylose 5-phosphate as the starting C_5 unit.

Formation of NDP-Octose

NDP-octose has never been identified as a precursor/intermediate in a natural product biosynthetic pathway. In contrast, NDP-heptose has been established as an

intermediate in the biosynthesis of the inner core of lipopolysaccharide (LPS), which is a major component of the outer membrane of Gram-negative bacteria.[217] Two pathways from D-sedoheptulose 7-phosphate (**5-20**) to either GDP-D-*glycero*-α-D-*mannno*-heptose (**5-58**) or ADP-D-*glycero*-β-L-*mannno*-heptose (**5-61**) have been characterized for the biosyntheses of NDP-heptoses (Figure 5-10).[218,219] Sequence analyses using the lincomycin A biosynthetic gene cluster from *S. lincolnensis* ATCC 25466 (accession number EU124663), S-layer (cell surface layer) glycan glycosylation gene cluster from *Aneurinibacillus thermoaerophilus* strain DSM 10155/G+ (accession number AF324836), and the corresponding genes from *E. coli* str. K-12 substr. MG1655 (accession number NC_000913) led to the proposed NDP-activation pathway starting from octulose 8-phosphate (**5-40**), as a key part of lincomycin A (**5-1**) pathway.

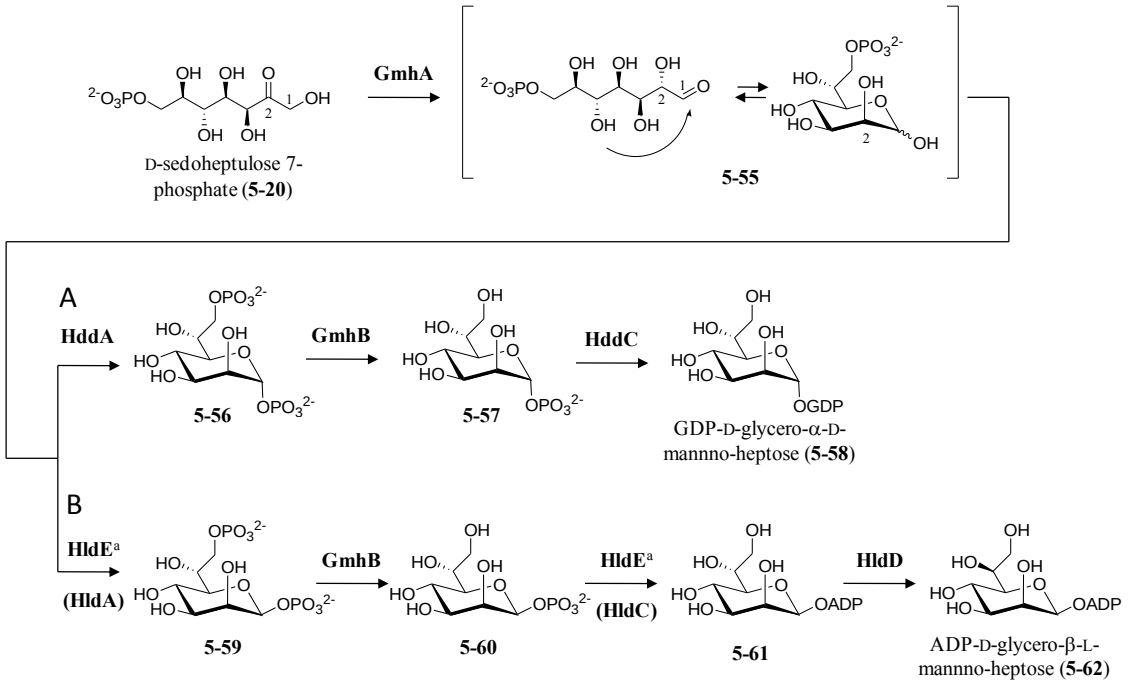


Figure 5-10. NDP-heptose biosynthetic pathways

(A) GDP-D-*glycero*-α-D-*mannno*-heptose biosynthetic pathway from *Aneurinibacillus thermoaerophilus*. (B) ADP-D-*glycero*-β-L-*mannno*-heptose biosynthetic pathway from *E. coli*. [a] A bifunctional enzyme, HldE, from *E. coli* is replaced by HldA and HldC in some other organisms.

As shown in Figure 5-7, LmbN was proposed to catalyze the C-1–C-2 isomerization of **5-40** to produce the corresponding octose 8-phosphate (**5-41**) in a reaction analogous to the GmhA-catalyzed isomerization (**5-20** → **5-55**). Sequence identities and similarities of LmbN to GmhA from *A. thermoaerophilus* and *E. coli* are 26% / 43% and 31% / 47%, respectively (Table 5-4). GmhA is known to produce the product **5-55** with a (*S*)-configuration at C-2.[218,220,221] In contrast, based on the stereochemistry at C-2 in MTL (**5-2**), the LmbN-catalyzed reaction is proposed to produce a product with an (*R*)-configuration at the C-2 position (**5-41**; Figure 5-7).

The following steps in the NDP-heptose biosynthetic pathway are phosphorylation, dephosphorylation, and nucleotidyl transfer reactions (Figure 5-10). The NDP-octose biosynthetic pathway may follow the same reaction sequence using LmbP, LmbK and LmbO, which are the putative kinase, phosphatase, and nucleotidyl transferase, respectively (Table 5-3). While LmbK displays moderate sequence similarities to GmhB in the NDP-heptose biosynthetic pathways, LmbP and LmbO showed similarities to the corresponding proteins only in the GDP-D-*glycero*-α-D-*mannno*-heptose (**5-58**) biosynthetic pathway (Table 5-4). Thus, phosphorylation and nucleotidyl transfer reactions in the MTL (**5-2**) biosynthetic pathway resemble those of the **5-58** biosynthetic pathway (Figure 5-7), and the product is predicted to be GDP-α-octose (**5-44**), not the ADP-α-derivative. However, formation of other phosphodinucleotide derivatives cannot be ruled out.

Table 5-4. Sequence similarity analysis of lincomycin A biosynthetic proteins and the corresponding proteins in NDP-heptose biosynthetic pathways.

	lincomycin A [*Streptomyces lincolnensis*] (identity / similarity %)			
	LmbN	**LmbP**	**LmbK**	**LmbO**
GDP-D-glycero-α-D-manno-heptose [*Aneurinibacillus thermoaerophilus*]	GmhA (26 / 43)	HddA (23 / 41)	GmhB (25 / 42)	HddC (27 / 47)
ADP-D-glycero-β-L-manno-heptose [*Escherichia coli*]	GmhA (31 / 47)	HldE (no significant similarity found)	GmhB (32 / 47)	HldE (no significant similarity found)

* HldE from *E. coli* is a bifunctional enzyme.

NDP-Octose Modification Pathway

The proposed GDP-α-octose (**5-44**) intermediate would be then modified through C-4 epimerization, C-6 amination, and C-8 dehydration as well as C-1 sulfur incorporation, to produce MTL (**5-2**) (see Figure 5-7). An epimerase homologue, LmbM, is proposed to be responsible for the C-4 epimerization (**5-44** → **5-45**). The C-6 oxidation and C-8 dehydration (**5-45** → **5-46**) are possibly catalyzed by a nucleotide sugar dehydrogenase homologue, LmbL, or an oxidoreductase homologue, LmbZ. The proposed reaction mechanism is analogous to that of the reaction catalyzed by TDP-α-D-glucose 4,6-dehydratase, which converts TDP-α-D-glucose (**5-63**) to TDP-4-keto-6-deoxy-α-D-glucose (**5-64**; Figure 5-11).[14,222] Finally, an aminotransferase homologue, LmbS, is proposed to install the amino group at C-6 (**5-46** → **5-47**; Figure 5-7).

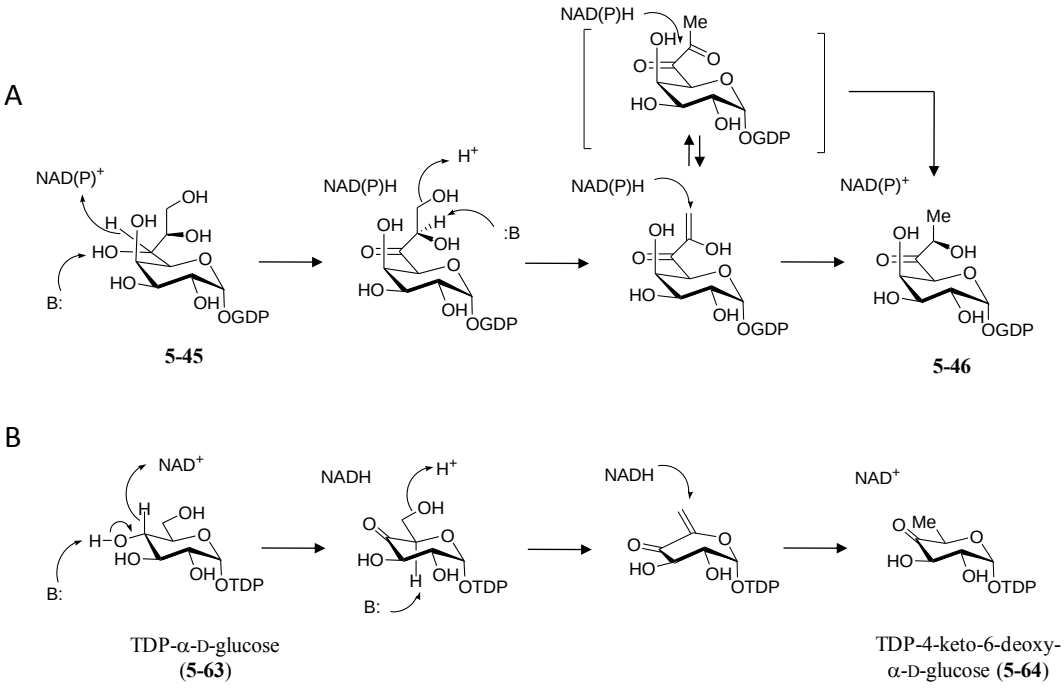


Figure 5-11. (A) Proposed mechanism of GDP-α-D-octose 6,8-dehydratase reaction and (B) the analogous TDP-α-D-glucose 4,6-dehydratase reaction.

Proposed Pathway of the C-1 Sulfur Incorporation

A glycosyltransferase homologue, LmbT, is proposed as a key enzyme for the sulfur incorporation at C-1 in the MTL (**5-2**) biosynthetic pathway. The enzyme may act as an *S*-glycosyltransferase to form an *S*-conujugate at C-1. A candidate for the glycosyl acceptor is L-cysteine (**5-47**→ **5-48**; Figure 5-7). Notably, similar C-1 thioglycoside is known for glucosinolates and sublancin (see Chapter 1, Figure 1-12, 1-13). To expose a free thiol group at C-1, an aminotransferase homologue, LmbF, is proposed to act as a C–S lyase to produce **5-49** using pyridoxal-5'-phosphate (PLP). A methyl transferase homologue, LmbG, would then transfer a methyl group to the C-1 thiol group to complete the MTL biosynthetic pathway (**5-49**→ **5-2**; Figure 5-7). Alternatively, the *S*-glycosyl acceptor of the LmbT reaction may be mycothiol (**5-65**) rather than L-cysteine, because the deduced gene product of *lmbE* displays moderate sequence similarity (33% identity and 48% similarity) to a mycothiol *S*-conjugate amidase (Mca) from *Mycobacterium tuberculosis* H37Rv (Figure 5-12).[223] Additionally, a conserved domain search revealed that the protein sequence of LmbV contains a putative mycothiol maleylpyruvate isomerase *N*-terminal domain. Although the function of LmbV is not clear at this point, it may also play a role in the proposed mycothiol-related C-1 sulfur incorporation process.

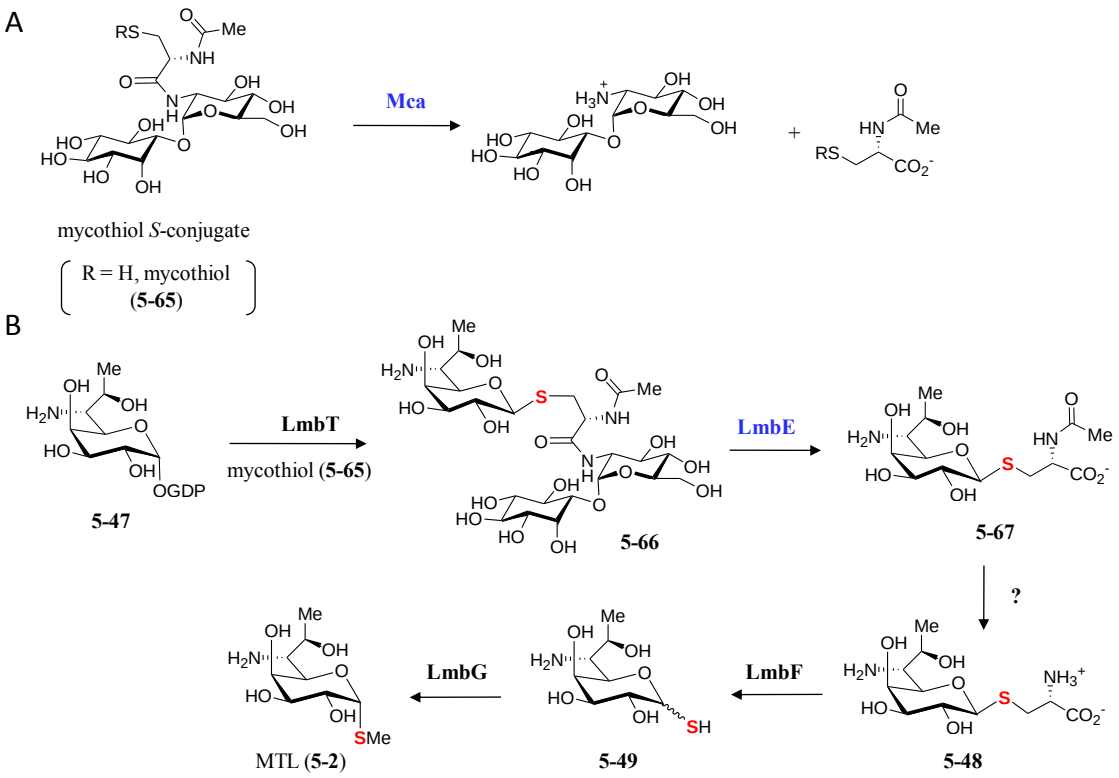


Figure 5-12. (A) Reaction scheme of mycothiol *S*-conjugate amidase and (B) the possible analogous reaction catalyzed by LmbE in the MTL biosynthetic pathway.

5.3.2 Functional Characterizations of LmbR.

To verify the proposed C_8 backbone construction catalyzed by LmbR (Figure 5-9), the recombinant LmbR with a *C*-terminal His_6-tag was overexpressed in *E. coli* and partially purified (Figure 5-13).

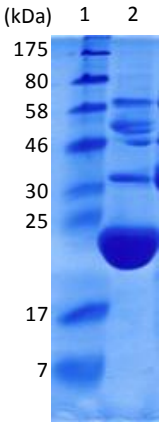


Figure 5-13. SDS-PAGE gel of the partially purified LmbR.

C-His_6 tagged LmbR (233 aa, 25.0 kDa) is shown in lane 2. Minor impurities are likely coexpressed chaperone proteins (DnaK ~70 kDa, GroEL ~60 kDa and DnaJ ~40 kDa). The molecular weight marks are 175, 80, 58, 46, 30, 25, 17 and 7 kDa (top to bottom, lane 1).

Mechanism of the transaldolase reaction catalyzed by LmbR was proposed to involve iminium bond formation between a lysine residue of the enzyme and C-2 ketone of the substrate sugar (**5-68**; Figure 5-14). To address this hypothesis, the purified *C*-His_6-LmbR was incubated with a putative substrate, D-fructose 6-phosphate (**5-22**), and treated with sodium borohydride prior to the MS analysis. Under this condition, the molecular masses observed are consistent with those of the reduced form of *C*-His_6-LmbR–D-glyceraldehyde 3-phosphate conjugate (**5-69**, caluculated 25024 Da, observed 25024 Da), the reduced *C*-His_6-LmbR–D-fructose 6-phosphate conjugate (**5-68**, calculated 25194 Da, observed 25193 Da), and the unmodified enzyme (caluculated 24952 Da, observed 24952 Da; Figure 5-15). These results suggest that D-fructose 6-phosphate (**5-22**) is indeed a possible substrate of LmbR and a proposed mechanism through formation of the iminium intermediate (**5-68**) followed by C–C bond cleavage *via* retro aldol reaction (**5-68** → **5-69**) could occur.

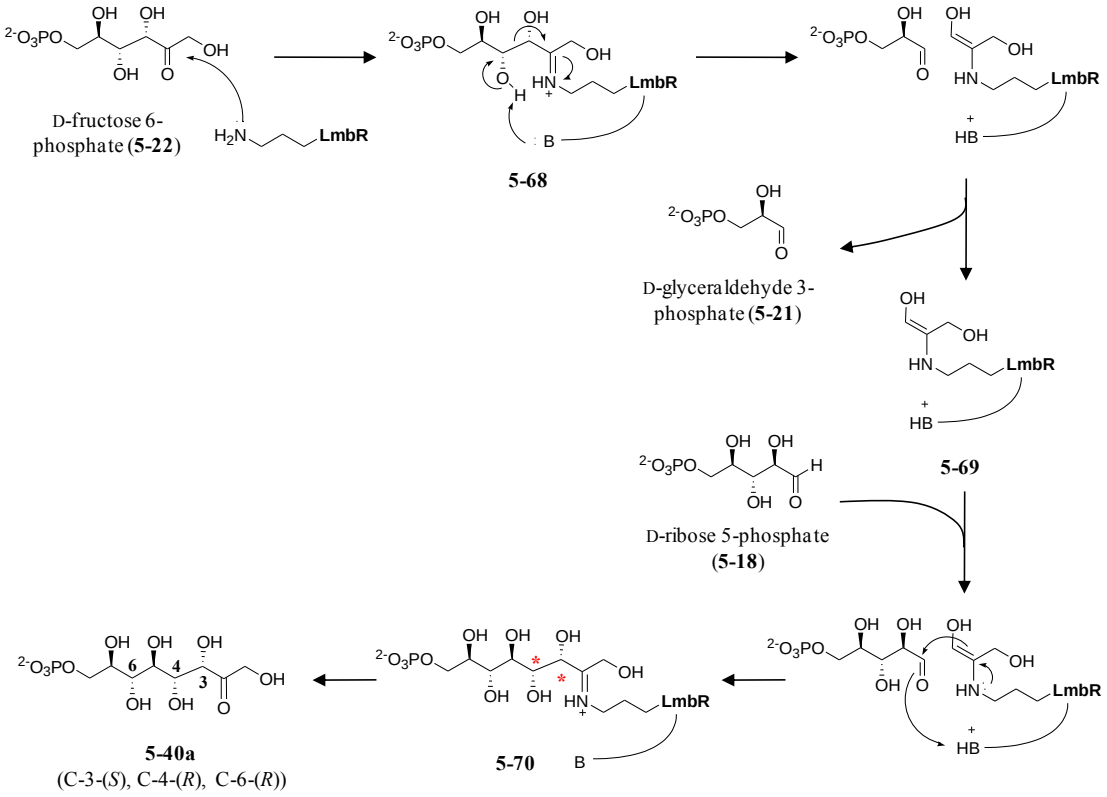


Figure 5-14. Proposed mechanism of the LmbR-catalyzed tranaldolase reaction using D-fructose 6-phoshate and D-ribose 5-phosphate as the substrates.

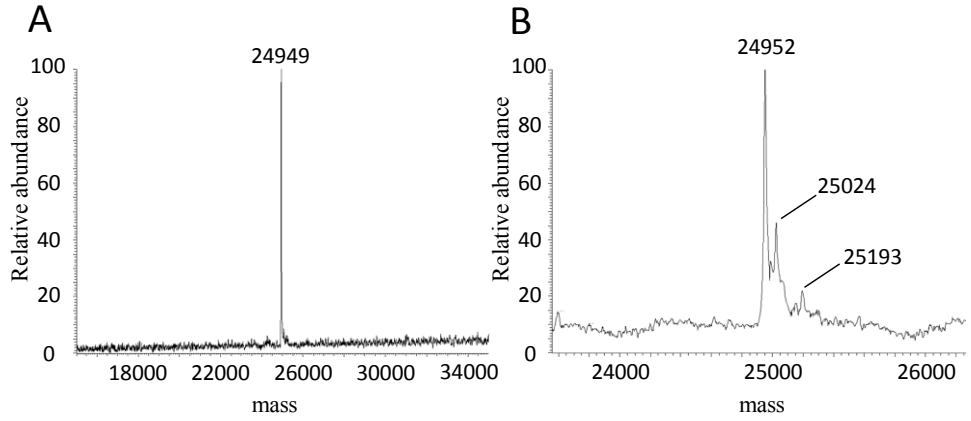


Figure 5-15. ESI-MS analyses of the LmbR reaction.

Deconvoluted ESI-MS of the *C*-His_6-LmbR. (A) *C*-His_6-LmbR in the absence of substrate. (B) *C*-His_6-LmbR in the presence of D-fructose 6-phosphate treated with sodium borohydride. The calculated molecular weight of *C*-His_6-LmbR (233 aa) is 24952 Da. The calculated molecular weights of *C*-His_6-LmbR–D-glyceraldehyde 3-phosphate conjugate (reduced) and *C*-His_6-LmbR–D-fructose 6-phosphate conjugate are 25024, and 25194 Da, respectively.

The proposed transaldolase activity of LmbR was further investigated using D-fructose 6-phosphate (**5-22**) and D-ribose 5-phosphate (**5-18**) as the C_3 unit donor and acceptor, respectively. HPLC analysis of the reaction mixture showed the decrease of the substrates and appearance of two new peaks (Figure 5-16, trace a). Control reactions omitting the enzyme or one of each substrate gave no product peaks (trace b–d). The reaction is clearly LmbR-dependent, and the results support the proposed transaldolase reaction. The new peak with a shorter retention time was isolated and subjected to ESI-MS analysis. The molecular weight of the observed mass signal is consistent with the proposed octulose 8-phosphate product (**5-40a**; ESI^- calculated for $C_8H_{16}O_{11}P^-$ $[M - H^+]$ 319.0436, observed 319.0431).

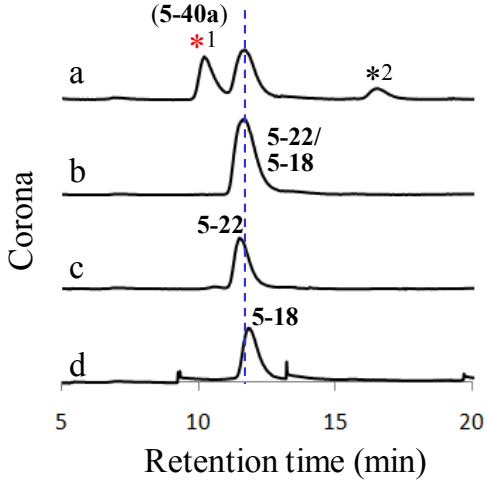


Figure 5-16. HPLC analysis of the LmbR reaction.

(a) Reaction containing *C*-His_6-LmbR 40 μM), D-fructose 6-phosphate (**5-22**, 10 mM) and D-ribose 5-phosphate (**5-18**, 10 mM). *[1] Product peak corresponding to octulose 8-phosphate (**5-40a**) based on the MS analysis). *[2] Product peak predicted as glyceraldehyde 3-phosphate (**5-21**). (b) No enzyme control. (c) No D-ribose 5-phosphate control. (d) No D-fructose 6-phosphate control. Signals at 9, 13, 20 min are irregular noise of the detection system.

5.3.3 Functional Characterizations of LmbN.

To verify the proposed isomerization activity of LmbN (**5-40** → **5-41**; Figure 5-8), the recombinant enzyme with a *C*-terminal His_6-tag was overexpressed in *E. coli* and purified (Figure 5-17). The purified *C*-His_6-LmbN protein was added to the LmbR reaction containing D-fructose 6-phosphate (**5-22**) and D-ribose 5-phosphate (**5-18**). The HPLC trace of the LmbN reaction shows the appearance of a product peak which has a shorter retention time than that of the LmbR product (Figure 5-18, trace d). A control reaction without LmbR did not yield this product peak (trace c). These results suggest that the LmbN product is indeed derived from the LmbR product, whose structure is proposed as **5-40a**, and not from D-fructose 6-phosphate (**5-22**) or D-ribose 5-phosphate (**5-18**) present in the reaction solution. In the absence of the LmbR product, minor peaks

likely derived from D-fructose 6-phosphate (**5-22**) and D-ribose 5-phosphate (**5-18**) were observed (trace c). These peaks might be arised from the C1–C2 isomerized products of **5-22** and **5-18**, namely, D-glucose 6-phosphate (**5-14**) and D-ribulose 5-phosphate (**5-17**), respectively, although they were not characterized. The LmbN product peak was isolated and subjected to ESI-MS analysis. The molecular weight of the observed mass signal is consistent with the proposed octose 8-phosphate structure (**5-41a**; ESI$^-$ calculated for $C_8H_{16}O_{11}P^-$ [M – H$^+$] 319.0436, observed 319.0431).

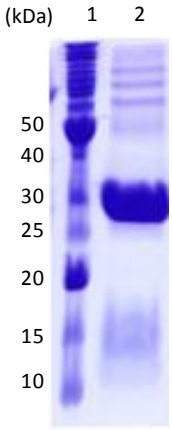


Figure 5-17. SDS-PAGE gel of the purified LmbN.

C-His$_6$ tagged LmbN (304 aa, 31.9 kDa) is shown in lane 2. The molecular weight marks are 220, 160, 120, 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, and 10 kDa (top to bottom, lane 1).

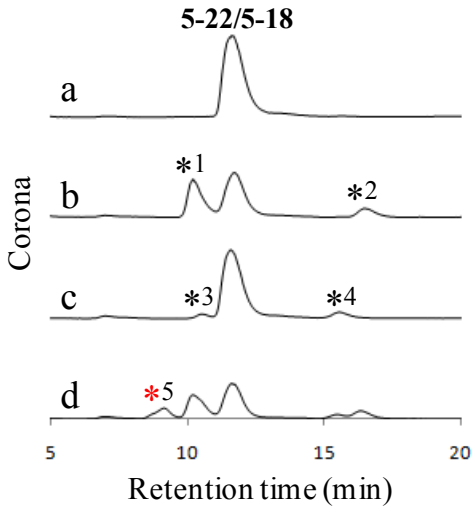


Figure 5-18. HPLC analysis of the LmbN reaction.

(a) Mixture of D-fructose 6-phosphate (**5-22**, 10 mM) and D-ribose 5-phosphate (**5-18**, 10 mM) substrates. (b) LmbR reaction using **5-22** and **5-18**. *[1] LmbR product peak corresponding to octulose 8-phosphate (**5-40a**). *[2] LmbR product peak predicted as glyceraldehyde 3-phosphate (**5-21**). (c) LmbN (50 μM) was added to the reaction a. *[3] Uncharacterized minor product possibly D-glucose 6-phosphate (**5-14**). (The retention time is consistent with the authentic standard of **5-14**.) *[4] Uncharacterized minor product possibly D-ribulose 5-phosphate (**5-17**). (d) LmbN (50 μM) was added to the reaction b after removing LmbR. *[5] LmbN product peak corresponding to octose 8-phosphate (**5-41a**) based on the MS analysis).

5.3.4 LmbR and LmbN Reactions Using D-Xylose 5-Phosphate.

Initial investigation of LmbR and LmbN *in vitro* assays showed that D-ribose 5-phosphate (**5-18**) and D-fructose 6-phosphate (**5-22**) are the potential substrates in the early stage of the MTL biosynthetic pathway. The proposed structures of these enzyme reaction products are **5-40a** and **5-41a** (Figure 5-9). However, D-xylose 5-phosphate (**5-29**) is another possible C_5 unit for the LmbR reaction, and the predicted product is **5-40b**, namely, a C-6 epimer of **5-40a**, in this case.

To test this alternative pathway, D-xylose 5-phosphate (**5-29**) was chemically synthesized according to a reported method. The purified *C*-His_6-LmbR protein was then incubated with **5-29** and D-fructose 6-phosphate (**5-22**), and the reaction mixture was analyzed by HPLC. A new peak having a shorter retention time than that of **5-41a** was observed (Figure 5-19, trace b). This peak was isolated and subjected to ESI-MS analysis. The molecular weight of the observed mass signal is consistent with the proposed octulose 8-phosphate (**5-40b**; ESI^- calculated for $C_8H_{16}O_{11}P^-$ [M – H^+] 319.0, observed 319.1). These results indicate that LmbR can accept both D-ribose 5-phosphate (**5-18**) and D-xylose 5-phosphate (**5-29**) as the C_3 sugar acceptor. The products from each substrate are clearly different based on the distinct retention times on the HPLC traces but possess the same molecular weight, consistent with the proposed octulose 8-phosphate C-6 epimers, **5-40a** and **5-40b**.

When *C*-His_6-LmbN was incubated together with the LmbR reaction mixture containing D-xylose 5-phosphate (**5-29**) and D-fructose 6-phosphate (**5-22**), no change was observed of the LmbR product peak (predicted as **5-41b**; Figure 5-19, trace b,c). This result suggests that LmbN cannot isomerize the LmbR product derived from **5-29**, although it can isomerize the product resulting from D-ribose 5-phosphate (**5-18**) as described above. It is important to note that the proposed C1-C2 isomerization is a prerequisite for the formation of the pyranosidic octose intermediate in the proposed MTL biosynthetic pathway (Figure 5-7). Hence, the substrate specificity observed in the LmbN reaction suggests that the starting substrate of MTL biosynthetic pathway is likely D-ribose 5-phosphate (**5-18**) and not D-xylose 5-phosphate (**5-29**), even though both substrates can be processed by LmbR *in vitro*.

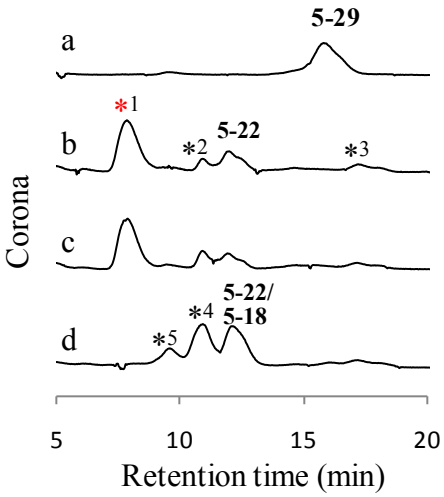


Figure 5-19. HPLC analysis of the LmbR and LmbN reaction using D-xylose 5-phosphate

(a) Synthetic standard of D-xylose 5-phosphate (**5-29**). (b) LmbR (40 μM) reaction using **5-29** (10 mM) and D-fructose 6-phosphate (**5-22**, 10 mM) as the substrates. *[1] LmbR product peak corresponding to octulose 8-phosphate (**5-40b**) based on the MS analysis. *[2] Uncharacterized degradation peak. *[3] LmbR product peak predicted as glyceraldehyde 3-phosphate (**5-21**). (c) Both LmbR (40 μM) and LmbN (50 μM) were incubated with **5-29** and **5-22**. (d) Control reaction using D-ribose 5-phosphate (**5-18**) instead of D-xylose 5-phosphate (**5-29**). *[4] LmbR product peak corresponding to octulose 8-phosphate (**5-40a**). *[5] LmbN product peak corresponding to octose 8-phosphate (**5-41a**).

5.3.5 Characterization of the LmbR and LmbN Products.

To further characterize the LmbR and LmbN reaction products, heptaacetyl octose derivatives **5-71a** and **5-71b** were chemically synthesized as the authentic standards. The compounds **5-71a** and **5-71b** possess the same stereochemistry at the C-2–C-7 carbons as the proposed LmbR/LmbN reaction products derived from D-ribose 5-phosphate (**5-18**) and D-xylose 5-phosphate (**5-29**), respectively (Figure 5-20). The enzymatic reaction mixture containing the LmbR product (**5-40a**, proposed structure) and the LmbN product (**5-41a**, proposed structure) from D-ribose 5-phosphate (**5-18**) and D-

fructose 6-phosphate (**5-22**) was treated with alkaline phosphatase. The resulting dephosphorylated sugar compounds were then reacted with acetic anhydride and pyridine to peracetylate the free alcohol groups. The derivatized enzymatic product mixture and the authentic standards were analyzed by HPLC.

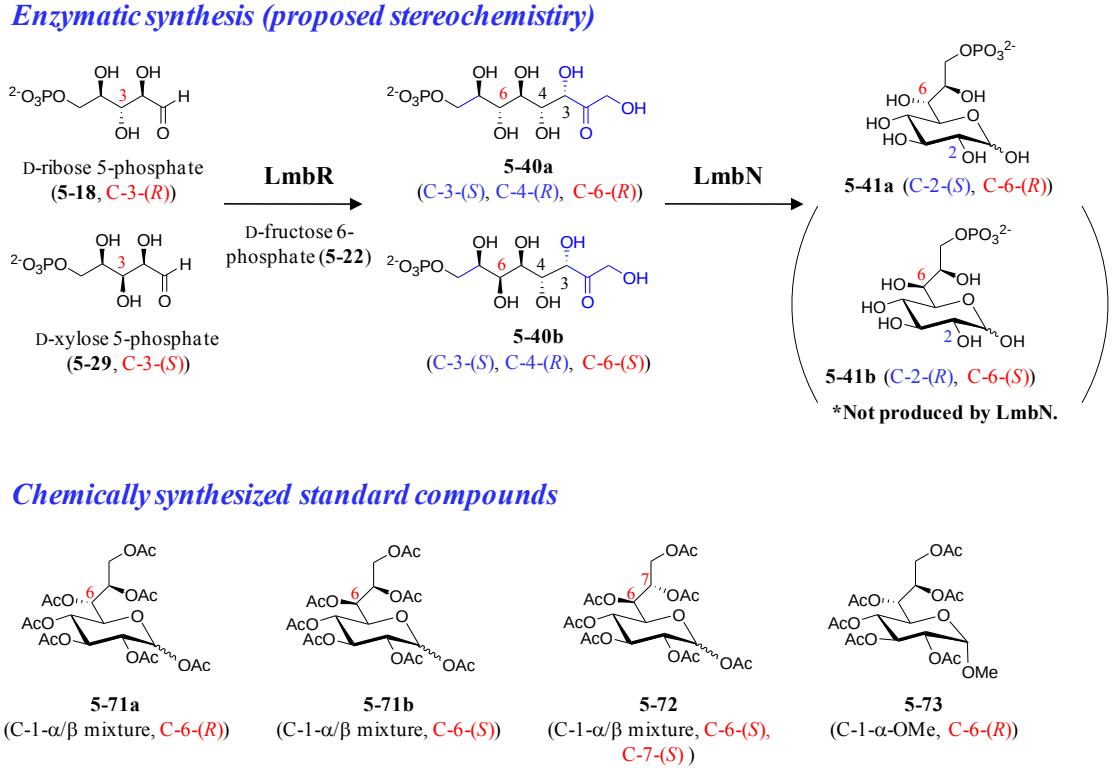


Figure 5-20. Proposed structures of LmbR and LmbN products and structures of chemically prepared authentic standards.

As shown in Figure 5-21, four major signals were observed from the LmbR reactions (retention time ~13.5 min and ~17.5 min, trace b). The earlier two peaks matched the signals from the control reaction yielding pentaacetyl fructose (trace a). The MS data of the isolated peaks (mixture of the two peaks) were consistent with this

assignment (ESI^+, calculated for $C_{14}H_{19}O_9^+$ [M - AcO^-] 331.1, found 331.2), corresponding to penta-*O*-acetyl-α-D-fructose (**5-74α**) and penta-*O*-acetyl-β-D-fructose (**5-74β**, Figure 5-22). Similarly, the latter two peaks may correspond to C-1 anomers of furanosidic octulose heptaacetate, **5-75α** and **5-75β** (Figure 5-22). The MS data of the isolated peaks were consistent with this hypothesis (ESI^+, calculated for $C_{20}H_{27}O_{13}^+$ [M - AcO^-] 475.1, found 475.2).

When both LmbR and LmbN were used, a new peak emerged (Figure 5-21, trace c, retention time ~19 min). Retention time of the authentic standard **5-71a** matches this peak (trace d), and they were coeluted when coinjected (trace h). In contrast, the other stereoisomers of pyranosidic octose heptaacetate **5-71b** and **5-72** show distinct retention times (trace f, g). The product peak (*3 in trace c) derived from the lmbN reaction was isolated and subjected to HRMS analysis. The observed mass signal is indeed consistent with the proposed heptaacetyl octose 8-phospahte product (ESI^+, calculated for $C_{22}H_{30}O_{15}Na^+$ $[M + Na]^+$ 557.1477, found 557.1478).

The results shown above support the proposed structures of the LmbR and LmbN products as **5-40a** and **5-41a**, respectively. The data are consistent with the proposed transaldolase reaction catalyzed by LmbR with retention of the configuration at the C-3 and C-4 positions as well as the proposed C-1–C-2 isomerization catalyzed by LmbN yielding the (*S*)-configuration at the C-2 position. Importantly, the products **5-40a** and **5-41a** are the first two examples of experimentally verified intermediates in the MTL biosynthetic pathway.

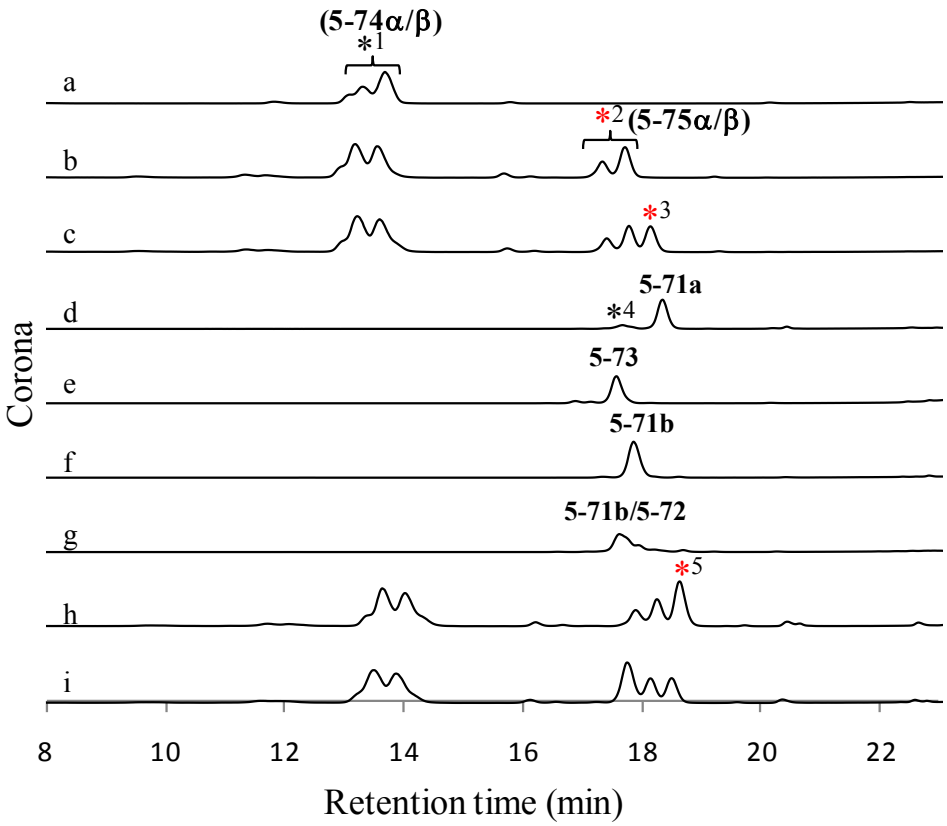


Figure 5-21. HPLC analysis of the acetyl derivatized LmbR and LmbN products.

(a) Control reaction using D-fructose (**5-36**). *[1] The two peaks at ~13.5 min are expected to be α/β mixture of pentaacetyl D-fructose (**5-74α/β**). (b) LmbR reaction followed by dephosphorylation and acetylation. *[2] The two peaks at ~17.5 min are expected to be α/β mixture of heptaactyl furanosidic octulose (**5-75α/β**). (c) LmbR and LmbN reactions followed by dephosphorylation and acetylation. *[3] The peak at ~19 min is expected to be heptaacetyl pyranosidic octose **5-71a**. (d) Authentic standard of **5-71a**. *[4] The peak at ~17 min is a synthetic intermediate **5-73** (detected by NMR, see Figure 5-20). (e) Authentic standard of **5-73**. (f) Authentic standard of **5-71b**. (g) Mixture of the authentic standards **5-71b** and **5-72**. (h) Coinjection of the sample derived from LmbR and LmbN reaction, and the authentic standard **5-71a**. *[5] LmbN reaction derived peak and the authentic standard of **5-71a** were coeluted. (i) Coinjection of the sample derived from LmbR and LmbN reaction, and the authentic standard **5-71b**.

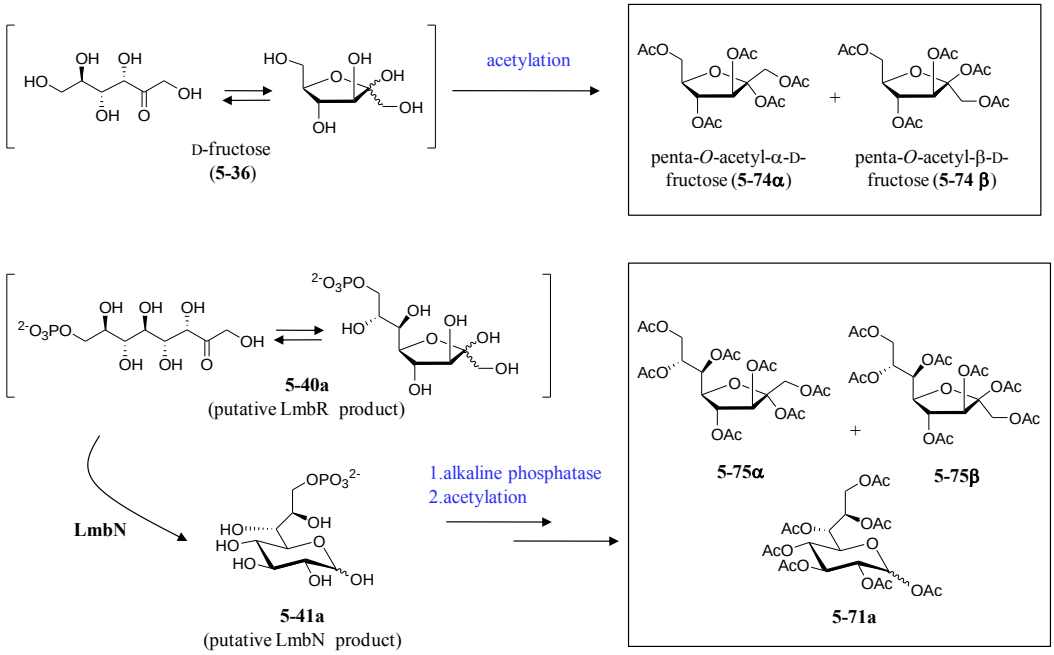


Figure 5-22. Proposed structures of the acetylated sugar derivatives.

5.3.6 Substrate Specificity of the LmbR and LmbN Reactions.

As described above, while LmbR could utilize both D-ribose 5-phosphate (**5-18**) and D-xylose 5-phosphate (**5-29**) as the C_5 unit, only the product formed in the former reaction was accepted by LmbN. The observed substrate specificity of LmbN suggested that D-ribose 5-phosphate (**5-18**) is likely the physiological C_5 unit substrate for the LmbR reaction. It was also confirmed that D-fructose 6-phosphate (**5-22**) could serve as the C_3 unit donor for this reaction.

As mentioned earlier, D-sedoheptulose 7-phosphate (**5-20**) from the pentose phosphate pathway (Figure 5-3) is also a possible C_3 unit donor for the LmbR reaction (Figure 5-9). In addition, D-fructose (**5-36**) and dihydroxyacetone (**5-35**) are other possible C_3 unit donors based on the proposed mechanism of the LmbR reaction (Figure 5-14, 5-23). To test whether these C_3 unit substrate candidates can be utilized by LmbR

in vitro, LmbR reaction assays followed by HPLC analyses were performed using these potential substrates. Moreover, LmbR-catalyzed transaldolase reaction to make a C_8 sugar was compared with the usual transaldolase reaction forming a C_7 sugar. Finally, D-sedoheptulose 7-phosphate (**5-20**) was also tested as a LmbN reaction substrate because it is the substrate of the homologous enzyme, GmhA, in the NDP-heptose biosynthetic pathways (Figure 5-10).

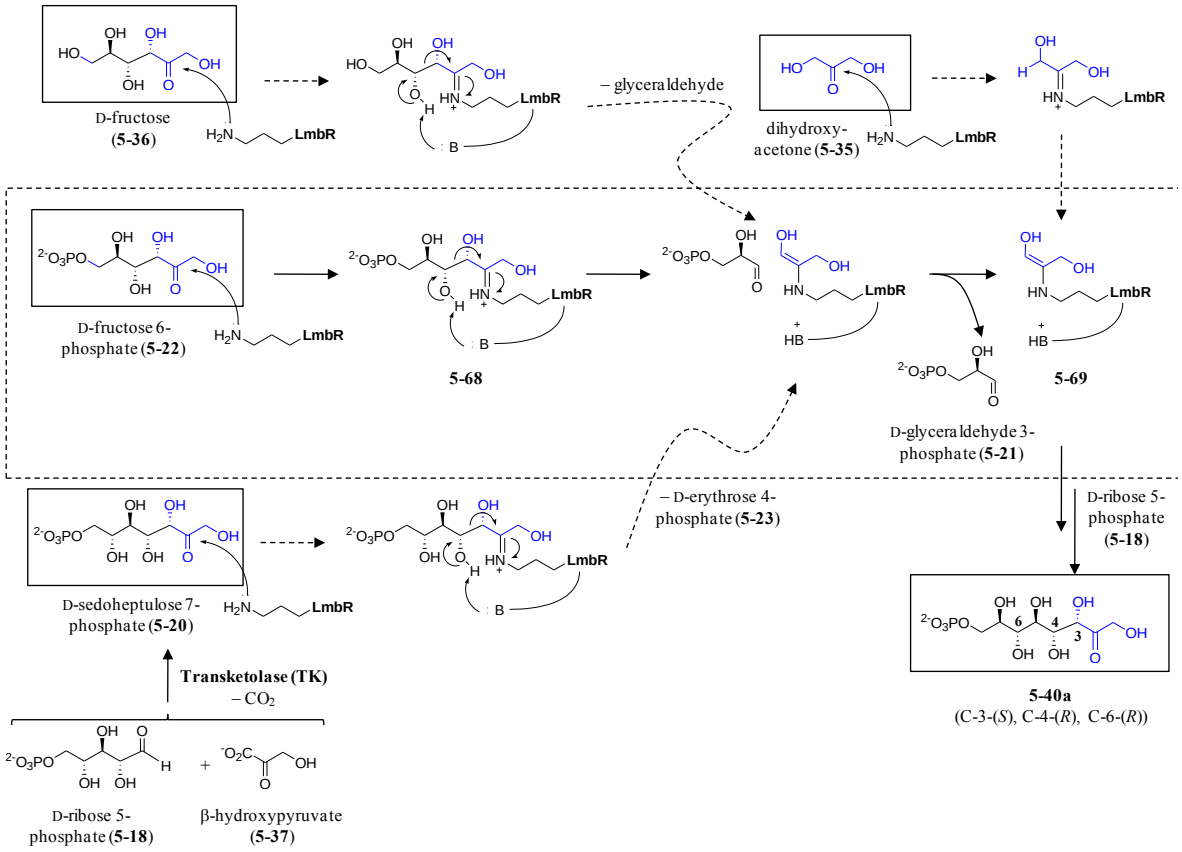


Figure 5-23. Potential C_3 unit donors of the LmbR reaction.

LmbR Reaction Using D-Sedoheptulose 7-Phosphate

To study the LmbR reaction with D-sedoheptulose 7-phosphate (**5-20**), compound **5-20** was prepared from D-ribose 5-phosphate (**5-18**) and β-hydroxypyruvate (**5-37**) using transketolase (TK) from *E. coli* (Figure 5-23).[215] The peak corresponding to D-sedoheptulose 7-phosphate (**5-20**) was detected by HPLC (Figure 5-24, trace a, retention time ~11 min). After removing the TK, LmbR and D-ribose 5-phosphate (**5-18**) were added to the reaction mixture. A new peak having a similar elution profile as the LmbR product (**5-40a**) generated in the reaction using D-fructose 6-phosphate (**5-22**) as the C_3 unit donor was observed (trace b, c). However, the retention time of the new peak is too close to the substrate (**5-20**) peak, and it is difficult to verify its identity using this method.

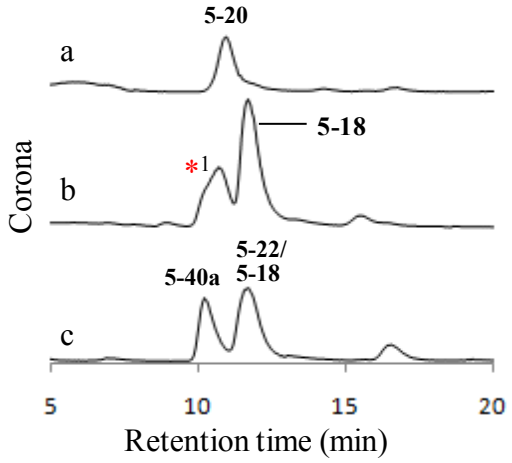


Figure 5-24. HPLC analysis of the LmbR reaction using D-sedoheptulose 7-phosphate (**5-20**)

(a) Transketolase (TK) reaction showing the formation of **5-20** as the product. (b) LmbR and D-ribose 5-phosphate (**5-18)** were added to the TK reaction. *[1] LmbR product (**5-40a**) seems to be formed although it overlaps with **5-20**. (c) Control LmbR reaction using D-ribose 5-phosphate (**5-18)** and D-fructose 6-phosphate (**5-22)** to yield **5-40a** and glyceraldehydes 3-phosphate (**5-21**).

LmbR Reaction Using D-Fructose or Dihydroxyacetone

Similarly, D-fructose (**5-36**) and dihydroxyacetone (**5-35**) were tested as potential C_3 unit donors for the LmbR reaction (Figure 5-23). When D-fructose 6-phosphate (**5-22**) was substituted with D-fructose (**5-36**), a similar product peak was observed but the yield was much lower than that of the control reaction using **5-22** (Figure 5-25, trace a, d). The yield was slightly improved if a higher concentration (50 mM) of **5-36** was used (trace b). Similarly, the product peak was confirmed but formed in a lower yield when dihydroxyacetone (**5-35**, 50 mM) was utilized as the C_3 unit donor (trace c). These results suggest that both D-fructose (**5-36**) and dihydroxyacetone (**5-35**) can serve as the substrate for LmbR *in vitro*, however, D-fructose 6-phosphate (**5-22**) is clearly the preferred C_3 unit donor for the reaction.

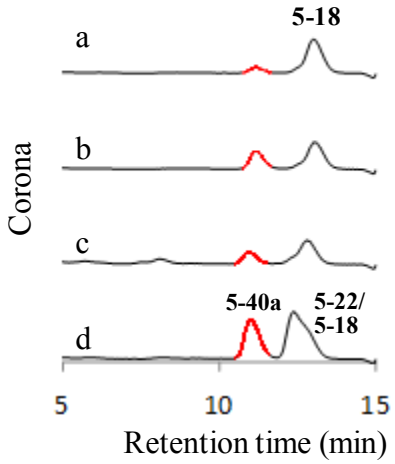


Figure 5-25. HPLC analysis of the LmbR reaction using D-fructose or dihydroxyacetone

(a) D-Fructose (**5-36**, 10 mM); (b) **5-36** (50 mM); (c) dihydroxyacetone (**5-35**, 50 mM); (d) D-Fructose 6-phospahte (**5-22**, 10 mM). D-Ribose 5-phospahte (**5-18**, 10 mM) was used as the C_5 acceptor in all reactions.

Comparison of LmbR with a Regular Transaldolase

Transaldolase (TA) is an essential enzyme catalyzing the reversible reaction from D-sedoheptulose 7-phosphate (**5-20**, C_7) and D-glyceraldehyde 3-phosphate (**5-21**, C_3) to D-fructose 6-phosphate (**5-22**, C_6) and D-erythrose 4-phosphate (**5-23**, C_4) in the pentose phosphate pathway (Figure 5-3). LmbR is now confirmed to catalyze the reaction utilizing D-fructose 6-phosphate (**5-22**, C_6) and D-ribose 5-phosphate (**5-18**, C_5) to yield D-octulose 8-phosphate (**5-40a**, C_8) and D-glyceraldehyde 3-phosphate (**5-21**, C_3) (Figure 5-9). Since LmbR showed relatively relaxed substrate specificity for the C_3 sugar donor (D-fructose 6-phosphate (**5-22**), D-fructose (**5-36**), and dihydroxyacetone (**5-35**)) and also for the C_3 sugar acceptor (D-ribose 5-phosphate (**5-18**) and D-xylose 5-phosphate (**5-29**)), it is possible that LmbR could catalyze the regular transaldolase reaction as well. In fact, there is an old report about a TA (from a special strain of Torula yeast) which is capable of catalyzing the same reaction to produce octulose 8-phosphate from D-ribose 5-phosphate (**5-18**) and D-fructose 6-phosphate (**5-22**, C_6) as LmbR.[224] However, whether such observed activity is common for a general TA is not clear because detailed studies of the substrate specificity of TA have not been explored thus far.[215]

To address these questions, a series of reactions using LmbR or commercially available TA from yeast were investigated. The reactions were spectrophotometrically monitored by following the producton of D-glyceraldehyde 3-phosphate (**5-21**) using glyceraldehydes 3-phosphate dehydrogenase (GAPDH) (Figure 5-26). As expected, the increase in absorbance at 340 nm was observed for both the usual TA reaction and standard LmbR reactions (Figure 5-27, trace a, b). LmbR could also catalyze the physiological transaldolase reaction utilizing D-erythrose 4-phosphate (**5-23**) as the sugar acceptor (trace d). Surprisingly, TA (from yeast) could not catalyze the LmbR reaction utilizing D-ribose 5-phosphate (**5-18**) as the sugar acceptor (trace c). The increase in

absorbance at 340 nm for the control reaction using LmbR with only D-fructose 6-phosphate (**5-22**) is likely caused by the retro-aldol reaction. This is consistent with the previous MS data showing the release of the LmbR reaction intermediate, a C_3 unit (**5-69**), in the absence of the sugar acceptor (Figure 5-14 and 5-15).

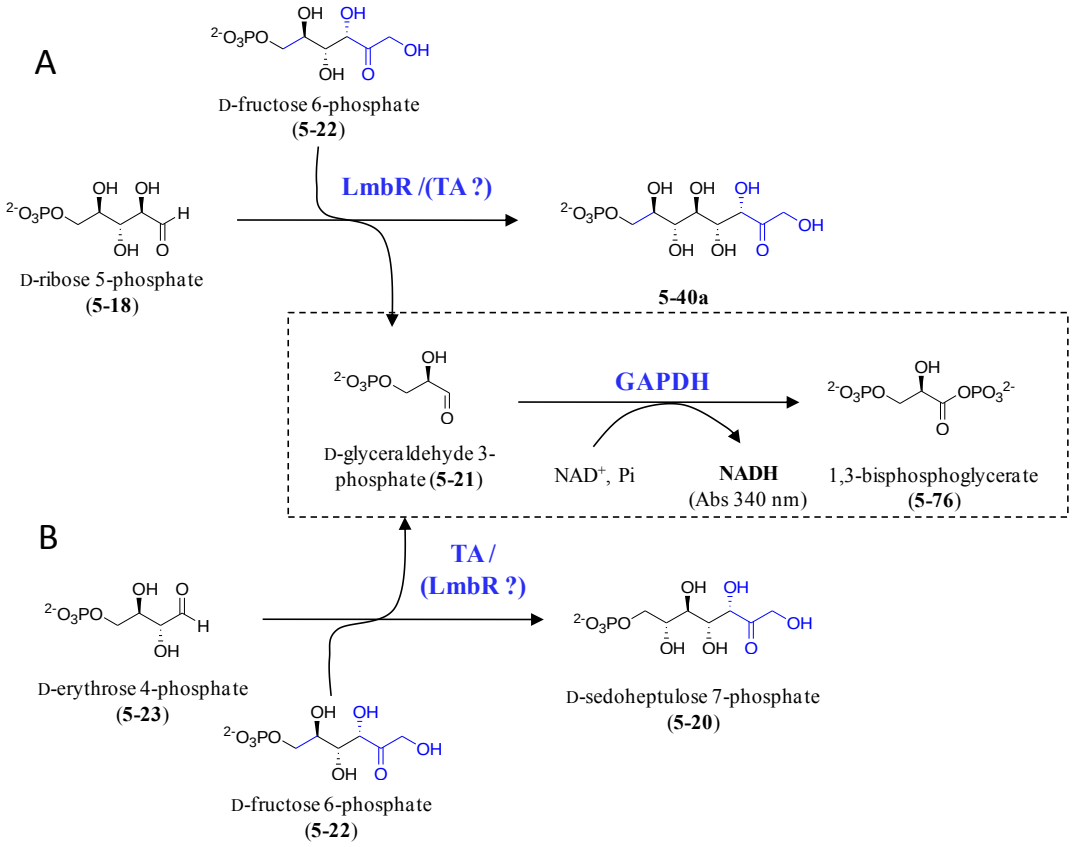


Figure 5-26. Comparison of the LmbR reaction and the regular transaldolase reaction.

(A) LmbR reaction. (B) Transaldolase (TA) reaction. The production of D-glyceraldehyde 3-phosphate (**5-21**) was detected by the coupled GAPDH-catalyzed assay shown in the dashed square.

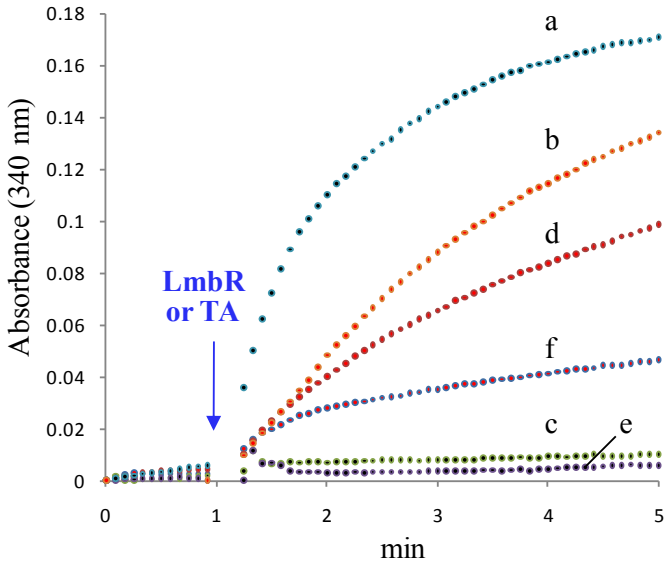


Figure 5-27. Investigation of the LmbR and TA reactions.

(a) Usual TA reaction using D-erythrose 4-phosphate (**5-23**) and D-fructose 6-phosphate (**5-22**). (b) Standard LmbR reaction using D-ribose 5-phosphate (**5-18**) and **5-22**. (c) TA with **5-18** and **5-22**. (d) LmbR with **5-23** and **5-22**. (e) Control reaction using TA and **5-22** only. (f) Control reaction using LmbR and **5-22** only.

LmbN Reaction Using D-Sedoheptulose 7-Phosphate

As described above, LmbN displays moderate sequence similarity to GmhA in the NDP-heptose biosynthetic pathways (Table 5-4), however it utilizes an octulose 8-phosphate instead of a heptulose 7-phosphate as the substrate. Interestingly, the product of the GmhA reaction has an (*R*)-configuration of its C-2 hydroxyl group (Figure 5-10). In contrast, C-2 hydroxyl group of the LmbN reaction product (**5-41a**) has an (*S*)-configuration based on HPLC comparison with an authentic standard (Figure 5-21). In order to learn more about the substrate specificity of LmbN reaction, D-sedoheptulose 7-phosphate (**5-20**), which is the substrate of GmhA reaction, was tested as a possible substrate for the LmbN reaction.

Compound **5-20** was generated from D-ribose 5-phosphate (**5-18**) and β-hydroxypyruvate (**5-37**) using transketolase (TK) from *E. coli* as described above (Figure 5-23). The HPLC trace did not show any change in this peak in the presence of LmbN (Figure 5-28). This result suggests that LmbN cannot take **5-18** as the substrate although, the possibility that the isomerized product might have an identical retention time as the substrate cannot be excluded. Together with the previous result that LmbN cannot take the LmbR product derived from D-xylose 5-phosphate as the substrate, it is clear that LmbN has more strict substrate specificity than LmbR.

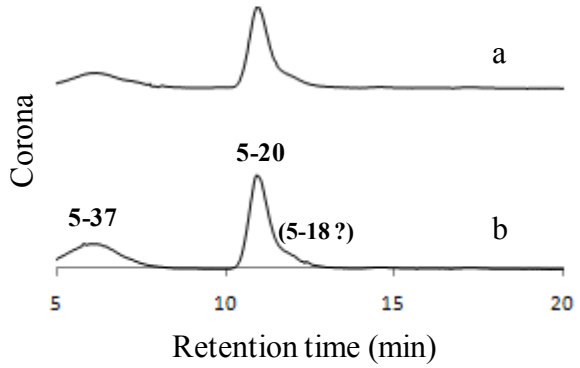


Figure 5-28. HPLC analysis of the LmbN reaction using D-sedoheptulose 7-phosphate.

(a) LmbN and D-sedoheptulose 7-phosphate (**5-20**). The compound **5-20** was produced by transketolase (TK) using D-ribose 5-phosphate (**5-18**, 10 mM**)** and β-hydroxypyruvate (**5-37**). (b) No LmbN control reaction.

5.3.7 Preparation of Other Enzymes in the MTL Biosynthetic Pathway.

As shown in Figure 5-7, it is proposed that the LmbR and LmbN reactions are followed by at least ten more enzymatic reactions in the MTL (**5-2**) biosynthetic pathway. To verify the proposed functions, each enzyme was expressed with a His_6-tag in *E. coli*

or *S. lividans*. Some enzymes were expressed as inclusion bodies, and, thus, could not be purified. The proteins purified from *E. coli* are *N*-His_6-LmbF, *N*-His_6-LmbG, *N*-His_6-LmbK, *C*-His_6-LmbL and *C*-His_6-LmbM, along with *C*-His_6-LmbR and *C*-His_6-LmbN described above. The proteins purified from *S. lividans* are *N*-His_6-LmbS and *N*-His_6-LmbT. SDS-PAGE images of the isolated proteins are shown in Figure 5-29. *In vitro* functional characterizations of each enzyme using the proposed substrate or substrate analogue are in progress.

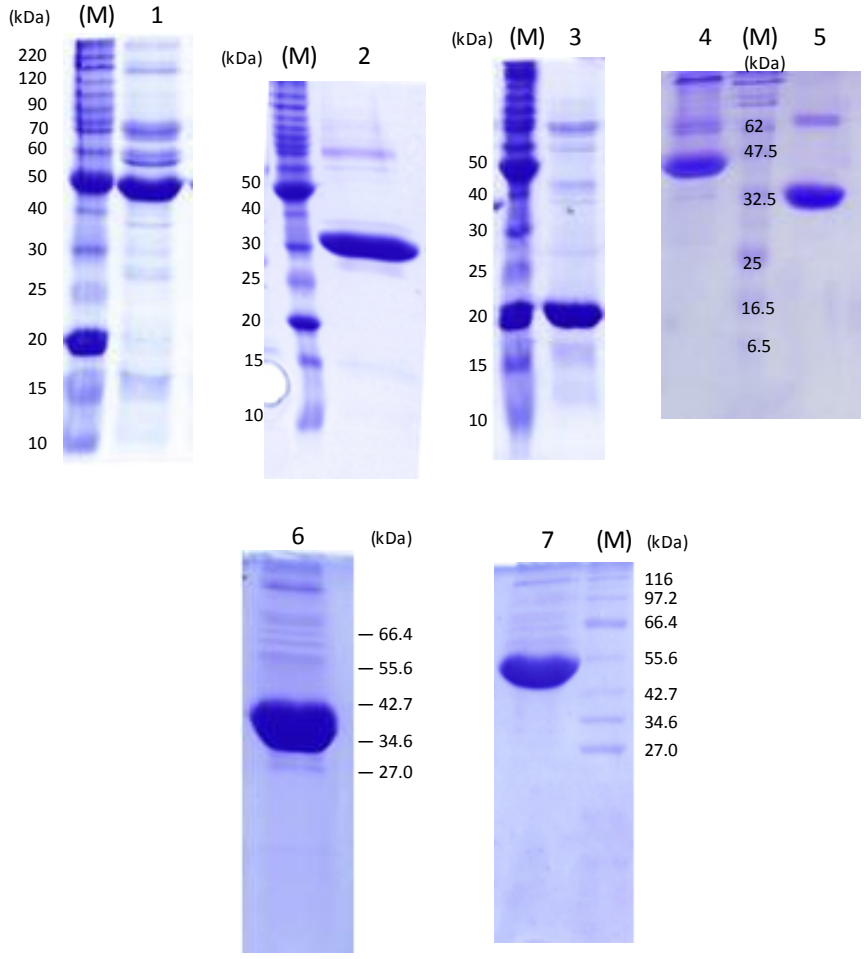


Figure 5-29. SDS-PAGE gel images of the purified enzymes in the MTL biosynthetic pathway.

Calculated molecular weights of each protein are (1) *N*-His_6-LmbF (446 aa, 48.8 kDa), (2) *N*-His_6-LmbG (285 aa, 31.3 kDa), (3) *N*-His_6-LmbK (210 aa, 22.5 kDa), (4) *C*-His_6-LmbL (449 aa, 47.4 kDa), (5) *C*-His_6-LmbM (336 aa, 36.8 kDa), (6) *N*-His_6-LmbS (401 aa, 43.0 kDa), and (7) *N*-His_6-LmbT (455 aa, 49.1 kDa). (M) stands for molecular weight markers.

5.3.8 Functional Investigation of LmbM.

LmbM is proposed as the C-4 epimerase in the MTL biosynthetic pathway (**5-44a** → **5-45a**; Figure 5-30A, route A). Alternatively, the epimerization step may take place

after the 6,8-dehydration process, which is proposed to be catalyzed by LmbL or LmbZ (route B). LmbM belongs to the short chain dehydrogenase/reductase superfamily, which includes UDP-galactose epimerase,[225] TDP-glucose 4,6-dehydratase,[226] and UDP-glucuronic acid decarboxylase.[227] Since the proposed substrates, **5-44a** and **5-77**, are not readily available through either chemical or enzymatic synthesis, model reactions using commercially available NDP-sugars were tested. Interestingly, we found that LmbM could take UDP-D-glucuronic acid (**5-38**) as the substrate and produce UDP-4-keto-D-xylose (**5-78**; Figure 5-30B). In this last section, the demonstrated UDP-4-keto-D-xylose synthase activity of LmbM is presented.

Purified *C*-His_6-LmbM was incubated with UDP-D-glucuronic acid (**5-38**) and NAD^+, and the reaction was analyzed using HPLC. Two new peaks (retention time ~11 min and ~19 min) appeared along with the consumption of **5-38** and NAD^+ (Figure 5-31A, trace a, b). One was confirmed to have an identical retention time (~19 min) to NADH using an authentic standard. The other peak was, thus, suspected to be an oxidized product resulting from **5-38**. This peak was isolated and subjected to MS and ^{1}H-NMR analyses. The observed mass signals were consistent with the UDP-4-keto-D-xylose (**5-78**; ESI^- calculated for $C_{14}H_{19}N_2O_{16}P_2$ [M – H^+] 533.0, observed 533.2) and its hydrated form (**5-79**; ESI^- calculated for $C_{14}H_{21}N_2O_{17}P_2$ [M – H^+] 551.0, observed 551.1). The ^{1}H-NMR data are also consistent with the previously reported values of UDP-4-keto-D-xylose (hydrate form) produced by ArnA in the biosynthetic pathway of 4-amino-4-deoxy-L-arabinose.[228] Compounds **5-78** and **5-79** are likely in equilibrium in aqueous solution.

To further confirm the formation of **5-78**, a carbonyl directing reagent, hydroxylamine, was added to the LmbM reaction mixture. If the keto form, **5-78**, exists in the solution, it should be converted to the corresponding oxime, **5-82**.[229,230] In fact,

upon addition of hydroxylamine, almost complete disappearance of the **5-78**/**5-79** peak and the appearance of a new peak eluted with a longer retention time (~14 min) was observed (Figure 5-31, trace c).

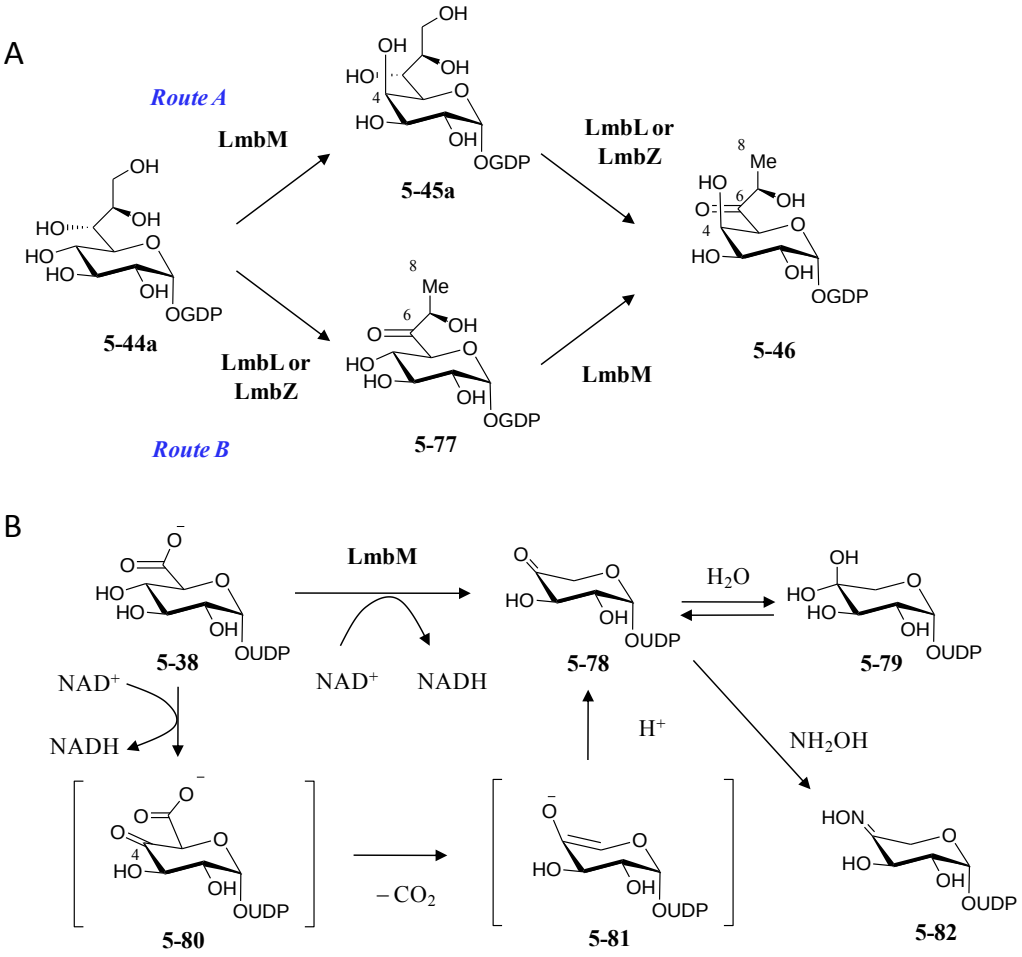


Figure 5-30. LmbM reaction

(A) Proposed LmbM reaction in the MTL biosynthetic pathway. (B) UDP-4-keto-D-xylose producing reaction catalyzed by LmbM.

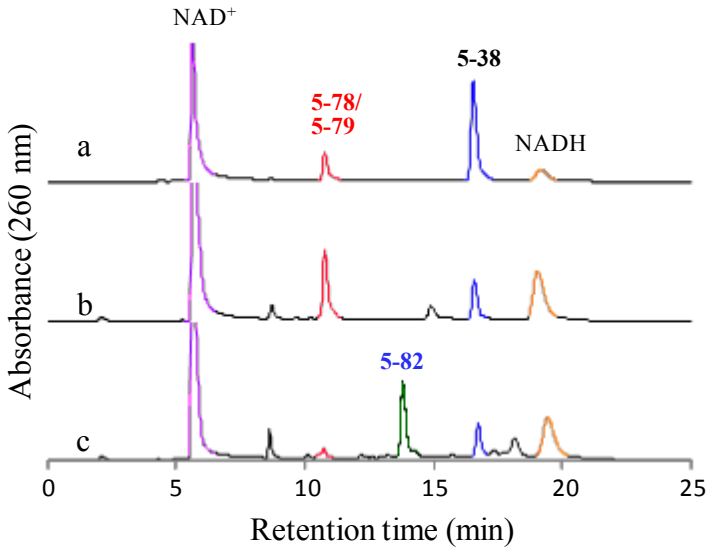


Figure 5-31. HPLC analysis of the LmbM-catalyzed model reaction.

(a) LmbR reaction mixture using UDP-D-glucuronic acid (**5-38**) and NAD^+ (1-h incubation). New peaks corresponding to UDP-4-keto-D-xylose (**5-78**) or its hydrated form (**5-79**) and NADH were observed. (b) LmbR reaction mixture (12-h incubation). (b) The reaction was treated with NH_4OH (**5-78** → **5-82**).

Although the LmbM reaction presented here is likely a side reaction, it is clear that the enzyme has the potential to oxidize the C-4 position of an NDP-sugar. In the MTL biosynthetic pathway, LmbM may oxidize the C-4 position of **5-44a** or **5-77** and reduce it back to epimerize the position rather than perform the decarboxylation observed with UDP-D-glucuronic acid (**5-38**).

5.4 CONCLUSIONS

In this chapter, the biosynthetic pathway of lincomycin A (**5-1**), which possesses a unique C-1 thiosugar, methylthiolincosamide (MTL, **5-2**), was studied. Previous feeding experiments and genetic information led us to propose the entire biosynthetic pathway of

MTL (**5-2**). As a key reaction, the C_8 backbone construction catalyzed by LmbR has now been characterized *in vitro*. Our data show that LmbR has relatively relaxed substrate specificity for both the C_3 sugar donors and acceptors. In contrast, LmbN shows rather strict substrate specificity, since only the octulose 8-phosphate resulting from the LmbR reaction with D-ribose 5-phosphate (**5-18**) can be recognized and processed by LmbN. These results are consistent with the proposed biosynthetic pathway involving octulose 8-phosphate **5-40a** and octose 8-phospahte **5-41a** as the early intermediates (Figure 5-7 and 5-20). The MS data and the HPLC analyses using the chemically synthesized standards fully support these assignments.

This work represents the first biochemical verification of **5-40a** and **5-41a** as key intermediates in the MTL (**5-2**) biosynthetic pathway. Identification of the unique C_8 back bone construction mechanism provides important information shedding light on the latter stage of the MTL biosynthesis. The next key intermediate proposed is a GDP-activated octose (**5-44**). Although similar NDP-heptoses are known in the biosynthesis of the inner core of lipopolysaccharide (LPS), the involvement of NDP-octose in a biosynthetic pathway is unique and has never been reported. As far as we know, the aminoglycoside antibiotic apramycin is the only other type of a natural product that may potentially be biosynthesized through an NDP-octose intermediate (gene cluster accession number: AJ875019 and AJ629123). Finally, C-1 thio incorporation mechanism is the most intriguing question in the biosynthesis of lincomycin A (**5-1**). Gene analysis led us to propose a pathway involving a putative *S*-glycosyltransferase, *lmbT*, and a C–S lyase, *lmbF*. Sulfur donor may be cysteine or mycothiol (**5-65**) as discussed above (Figure 5-12). The recombinant proteins, *N*-His_6-LmbT and *N*-His_6-LmbS, as well as several other proteins in the MTL biosynthetic pathway have been purified. Substrate preparation for these enzymes is in progress to elucidate the C-1 thiosugar biosynthesis.

Appendix: List of Abbreviations

ACP	Acyl carrier protein
ADP	Adenosine diphosphate
AK	Adenylate kinase
AMP	Adenosine monophosphate
A. orientalis	*Amycolatopsis orientalis*
ARS	Agricultural Research Service
Asp	Aspartic acid
A. thermoaerophilus	*Aneurinibacillus thermoaerophilus*
ATP	Adenosine triphosphate
bex cluster	BE-7585A biosynthetic gene cluster
BLAST	Basic local alignment search tool
B. subtilis	*Bacillus subtilis*
CAD	Charged aerosol detector
CD	Cysteine desulfurase
CDP	Cytidine diphosphate
CIP	Calf intestinal alkaline phosphatase
CoA	Coenzyme A
COSY	Correlated spectroscopy
Cys	Cysteine
DEPT	Distortionless enhancement by polarization transfer
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid

DNPH	2,4-Dinitrophenylhydrazine
L-DOPA	L-3,4-Dihydroxyphenylalanine
DTT	Dithiothreitol
DXP	1-Deoxy-D-xylulose 5-phosphate
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionization
FDA	Food and drug administration
FMN	Flavin mononucleotide
GAPDH	Glyceraldehydes 3-phosphate dehydrogenase
GDP	Guanosine diphosphate
Glu	Glutamic acid
Gly	Glycine
GSH	Glutathione
HMBC	Heteronuclear multiple bond correlation
HPLC	High performance liquid chromatography
HR	High resolution
HSQC	Heteronuclear single quantum coherence
INADEQUATE	Incredible natural abundance double quantum transfer experiment
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ISP	International *Streptomyces* Project
KS_α	β-ketoacyl synthase α subunit
KS_β	β-ketoacyl synthase β subunit
LB	Luria-Bertani broth
LC	Liquid chromatography

LDH	Lactate dehydrogenase
LPS	Lipopolysaccharide
Lys	Lysine
mBBr	Monobromobimane
mcm^5s^2U	5-Methoxycarbonylmethyl-2-thiouridine
Met	Methionine
MS	Mass spectrometry
MTL	Methylthiolincosamide
M. tuberculosis	*Mycobacterium tuberculosis*
NAD^+	nicotinamide adenine dinucleotide (oxidized form)
NADH	nicotinamide adenine dinucleotide (reduced form)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
NCBI	National center for biotechnology information
NDP	Nucleoside diphosphate
Ni-NTA	Nickel-nitrilotriacetic acid
NMP	Nucleoside monophosphate
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser enhancement spectroscopy
OASS	*O*-acetylserine sulfhydrylase
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDA	Photodiode array
Pdtc	Prydine dithiocarboxylic acid
PEP	Phosphoenolpyruvate

PK	Pyruvate kinase
PKS	Polyketide synthase
PLP	Pyridoxal 5'-phosphate
Pro	Proline
PVDF	Polyvinylidene fluoride
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
SAM	*S*-adenosyl-L-methionine
S. aurantiaca	*Stigmatella aurantiaca*
S. caelestis	*Streptomyces caelestis*
S. coelicolor	*Streptomyces coelicolor*
S. fradiae	*Streptomyces fradiae*
S. lincolnensis	*Streptomyces lincolnensis*
S. lividans	*Streptomyces lividans*
SDS	Sodium dodecyl sulphate
TA	Transaldolase
TDP	Thymidine diphosphate
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TK	Transketolase
TPP	Trehalose 6-phosphate phosphatase
TPS	Trehalose 6-phosphate synthases
tRNA	Transfer ribonucleic acid
TSB	Tryptone soya broth
T. thermophilus	*Thermus thermophilus*

UDP	Uridine diphosphate
UV	Ultraviolet
UV−vis	Ultraviolet−visible spectroscopy
YEME	Yeast extract/malt extract

References

1. Wu, S. Q., and Chappell, J. (2008) Metabolic engineering of natural products in plants; tools of the trade and challenges for the future, *Current Opinion in Biotechnology 19*, 145-152.

2. Fox, E. M., and Howlett, B. J. (2008) Secondary metabolism: regulation and role in fungal biology, *Curr Opin Microbiol 11*, 481-487.

3. Challis, G. L., and Hopwood, D. A. (2003) Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by Streptomyces species, *Proc. Natl. Acad.Sci. U.S.A. 100*, 14555-14561.

4. Price-Whelan, A., Dietrich, L. E. P., and Newman, D. K. (2006) Rethinking 'secondary' metabolism: physiological roles for phenazine antibiotics, *Nat. Chem. Biol. 2*, 71-78.

5. Dickschat, J. S. (2010) Quorum sensing and bacterial biofilms, *Nat. Prod. Rep. 27*, 343-369.

6. Newman, D. J., Cragg, G. M., and Snader, K. M. (2000) The influence of natural products upon drug discovery, *Nat. Prod. Rep. 17*, 215-234.

7. Demain, A. L. (2009) Antibiotics: Natural Products Essential to Human Health, *Med. Res. Rev. 29*, 821-842.

8. Li, J. W., and Vederas, J. C. (2009) Drug discovery and natural products: end of an era or an endless frontier?, *Science 325*, 161-165.

9. Butler, M. S. (2008) Natural products to drugs: natural product-derived compounds in clinical trials, *Nat. Prod. Rep. 25*, 475-516.

10. Walsh, C. T., and Fischbach, M. A. (2010) Natural products version 2.0: connecting genes to molecules., *J. Am. Chem. Soc. 132*, 2469-2493.

11. Szpilman, A. M., and Carreira, E. M. (2010) Probing the Biology of Natural Products: Molecular Editing by Diverted Total Synthesis, *Angew. Chem., Int. Ed. 49*, 9592-9628.

12. Varki, A. (1999) *Essentials of glycobiology*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

13. Varki, A. (2009) *Essentials of glycobiology*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

14. Thibodeaux, C. J., Melancon, C. E., and Liu, H. W. (2008) Natural-Product Sugar Biosynthesis and Enzymatic Glycodiversification, *Angew. Chem., Int. Ed. 47*, 9814-9859.

15. Schlunzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A., and Franceschi, F. (2001) Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria, *Nature 413*, 814-821.

16. Thibodeaux, C. J., Melancon, C. E., and Liu, H. W. (2007) Unusual sugar biosynthesis and natural product glycodiversification, *Nature 446*, 1008-1016.

17. Chen, X. (2011) Fermenting next generation glycosylated therapeutics, *ACS Chem. Biol. 6*, 14-17.

18. Fontecave, M., Ollagnier-de-Choudens, S., and Mulliez, E. (2003) Biological radical sulfur insertion reactions, *Chem. Rev. 103*, 2149-2166.

19. Mueller, E. G. (2006) Trafficking in persulfides: delivering sulfur in biosynthetic pathways, *Nat. Chem. Biol. 2*, 185-194.

20. Tang, G. L., Cheng, Y. Q., and Shen, B. (2004) Leinamycin biosynthesis revealing unprecedented architectural complexity for a hybrid polyketide synthase and nonribosomal peptide synthetase, *Chem. Biol. 11*, 33-45.

21. Rohr, J. (1989) Biosynthetic Formation of the S-Methyl Group of the Angucycline Antibiotic Urdamycin-E, *J. Chem. Soc.-Chem. Commun.*, 492-493.

22. Palumbo, A., d'Ischia, M., Misuraca, G., and Prota, G. (1982) Isolation and structure of a new sulphur-containing aminoacid from sea urchin eggs, *Tetrahedron Lett. 23*, 3207-3208.

23. Anderson, B., Crowfoot.D, and Viswamit.Ma. (1970) The structure of thiostrepton, *Nature 225*, 233-235.

24. Sasaki, H., Oishi, H., Hayashi, T., Matsuura, I., Ando, K., and Sawada, M. (1982) Thiolactomycin, a new antibiotic. II. Structure elucidation, *J. Antibiot. 35*, 396-400.

25. Taori, K., Paul, V. J., and Luesch, H. (2008) Structure and activity of largazole, a potent antiproliferative agent from the Floridian marine cyanobacterium Symploca sp., *J. Am. Chem. Soc. 130*, 1806-1807.

26. Okabe, T., Suda, H., Sato, F., and Okanishi, M. (1990) 抗腫瘍性物質 BE-7585A、その製法及びその用途, In *Jpn. Kokai Tokkyo Koho*, pp JP 02-16894 A, Japan.

27. Etoh, H., Iguchi, M., Nagasawa, T., Tani, Y., Yamada, H., and Fukami, H. (1987) Structures of Rhodonocardins Produced by a Nocardia Sp, *Agr. Biol. Chem. 51*, 1819-1824.

28. Hoeksema, H., Bannister, B., Birkenmeyer, R. D., Kagan, F., Magerlein, B. J., Mackellar, F. A., Schroeder, W., Slomp, G., and Herr, R. R. (1964) Chemical Studies on Lincomycin. I. The Structure of Lincomycin, *J. Am. Chem. Soc. 86*, 4223-4224.

29. Hoeksema, H. (1964) Celesticetin. IV. The Structure of Celesticetin, *J. Am. Chem. Soc. 86*, 4224-4225.

30. Argoudelis, A. D., and Brodasky, T. F. (1972) Studies with Streptomyces caelestis. I. New celesticetins, *J Antibiot 25*, 194-196.

31. Golik, J., Clardy, J., Dubay, G., Groenewold, G., Kawaguchi, H., Konishi, M., Krishnan, B., Ohkuma, H., Saitoh, K., and Doyle, T. W. (1987) Esperamicins, a Novel Class of Potent Antitumor Antibiotics .2. Structure of Esperamicin-X, *J. Am. Chem. Soc. 109*, 3461-3462.

32. Lee, M. D., Dunne, T. S., Siegel, M. M., Chang, C. C., Morton, G. O., and Borders, D. B. (1987) Calichemicins, a Novel Family of Antitumor Antibiotics .1. Chemistry and Partial Structure of Calichemicin-Gamma-1, *J. Am. Chem. Soc. 109*, 3464-3466.

33. McDonald, L. A., Capson, T. L., Krishnamurthy, G., Ding, W. D., Ellestad, G. A., Bernan, V. S., MAiese, W. M., Lassota, P., Discafani, C., Kramer, R. A., and Ireland, C. M. (1996) Namenamicin, a new enediyne antitumor antibiotic from the marine ascidian Polysyncraton lithostrotum, *J. Am. Chem. Soc. 118*, 10898-10899.

34. Oku, N., Matsunaga, S., and Fusetani, N. (2003) Shishijimicins A-C, novel enediyne antitumor antibiotics from the ascidian Didemnum proliferum, *J. Am. Chem. Soc. 125*, 2044-2045.

35. Stefanska, A. L., Fulston, M., Houge-Frydrych, C. S. V., Jones, J. J., and Warr, S. R. (2000) A potent seryl tRNA synthetase inhibitor SB-217452 isolated from a Streptomyces species, *J. Antibiot. 53*, 1346-1353.

36. Benz, G., Schroder, T., Kurz, J., Wunsche, C., Karl, W., Steffens, G., Pfitzner, J., and Schmidt, D. (1982) Constitution of the Deferriform of the Albomycins Delta-1, Delta-2, and Epsilon, *Angew Chem Int Edit 21*, 527-528.

37. Pramanik, A., Stroeher, U. H., Krejci, J., Standish, A. J., Bohn, E., Paton, J. C., Autenrieth, I. B., and Braun, V. (2007) Albomycin is an effective antibiotic, as exemplified with Yersinia enterocolitica and Streptococcus pneumoniae., *Int. J. Med. Microbiol. 297*, 459-469.

38. Yoshikawa, M., Morikawa, T., Matsuda, H., Tanabe, G., and Muraoka, O. (2002) Absolute stereostructure of potent alpha-glucosidase inhibitor, salacinol, with unique thiosugar sulfonium sulfate inner salt structure from Salacia reticulata, *Bioorg. Med. Chem. 10*, 1547-1554.

39. Yoshikawa, M., Murakami, T., Yashiro, K., and Matsuda, H. (1998) Kotalanol, a potent alpha-glucosidase inhibitor with thiosugar sulfonium sulfate structure, from antidiabetic ayurvedic medicine Salacia reticulata, *Chem. Pharm. Bull. 46*, 1339-1340.

40. Sim, L., Jayakanthan, K., Mohan, S., Nasi, R., Johnston, B. D., Pinto, B. M., and Rose, D. R. (2010) New glucosidase inhibitors from an ayurvedic herbal treatment for type 2 diabetes: structures and inhibition of human intestinal maltase-glucoamylase with compounds from Salacia reticulata, *Biochemistry 49*, 443-451.

41. Yoshikawa, M., Xu, F. M., Nakamura, S., Wang, T., Matsuda, H., Tanabe, G., and Muraoka, O. (2008) Salaprinol and ponkoranol with thiosugar sulfonium sulfate structure from Salacia prinoides and alpha-glucosidase inhibitory activity of ponkoranol and kotalanol desulfate, *Heterocycles 75*, 1397-1405.

42. Halkier, B. A., and Gershenzon, J. (2006) Biology and biochemistry of glucosinolates, *Annu. Rev. Plant Biol. 57*, 303-333.

43. Capon, R. J., and Macleod, J. K. (1987) 5-Thio-D-Mannose from the Marine Sponge Clathria-Pyramida (Lendenfeld) - the 1st Example of a Naturally-Occurring 5-Thiosugar, *J. Chem. Soc.-Chem. Commun.*, 1200-1201.

44. Yuasa, H., Izumi, M., and Hashimoto, H. (2009) Thiasugars: potential glycosidase inhibitors, *Curr. Top. Med. Chem. 9*, 76-86.

45. Ahlert, J., Shepard, E., Lomovskaya, N., Zazopoulos, E., Staffa, A., Bachmann, B. O., Huang, K. X., Fonstein, L., Czisny, A., Whitwam, R. E., Farnet, C. M., and Thorson, J. S. (2002) The calicheamicin gene cluster and its iterative type I enediyne PKS, *Science 297*, 1173-1176.

46. Peschke, U., Schmidt, H., Zhang, H. Z., and Piepersberg, W. (1995) Molecular Characterization of the Lincomycin-Production Gene-Cluster of Streptomyces Lincolnensis-78-11, *Mol. Microbiol. 16*, 1137-1156.

47. Cermak, L., Novotna, J., Sagova-Mareckova, M., Kopecky, J., Najmanova, L., and Janata, J. (2007) Hybridization analysis and mapping of the celesticetin gene cluster revealed genes shared with lincomycin biosynthesis, *Folia Microbiol. 52*, 457-462.

48. Zeng, Y., Roy, H., Patil, P. B., Ibba, M., and Chen, S. (2009) Characterization of Two Seryl-tRNA Synthetases in Albomycin-Producing Streptomyces sp Strain ATCC 700974, *Antimicrob. Agents Chemother. 53*, 4619-4627.

49. Geu-Flores, F., Olsen, C. E., and Halkier, B. A. (2009) Towards engineering glucosinolates into non-cruciferous plants., *Planta 229*, 261-270.

50. Geu-Flores, F., Nielsen, M. T., Nafisi, M., Møldrup, M. E., Olsen, C. E., Motawia, M. S., and Halkier, B. A. (2009) Glucosinolate engineering identifies a gamma-glutamyl peptidase., *Nat. Chem. Biol. 5*, 575-577.

51. Mikkelsen, M. D., Naur, P., and Halkier, B. A. (2004) Arabidopsis mutants in the C-S lyase of glucosinolate biosynthesis establish a critical role for indole-3-acetaldoxime in auxin homeostasis, *Plant J. 37*, 770-777.

52. Grubb, C. D., Zipp, B. J., Ludwig-Muller, J., Masuno, M. N., Molinski, T. F., and Abel, S. (2004) Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis, *Plant J. 40*, 893-908.

53. Oman, T. J., Boettcher, J. M., Wang, H., Okalibe, X. N., and van der Donk, W. a. (2011) Sublancin is not a lantibiotic but an S-linked glycopeptide., *Nat. Chem. Biol.*, 1-3.

54. Booker, S. J., Cicchillo, R. M., and Grove, T. L. (2007) Self-sacrifice in radical S-adenosylmethionine proteins, *Curr. Opin. Chem. Biol. 11*, 543-552.

55. Kessler, D. (2006) Enzymatic activation of sulfur for incorporation into biomolecules in prokaryotes, *Fems Microbiol. Rev. 30*, 825-840.

56. Frey, P. A., and Magnusson, O. T. (2003) S-Adenosylmethionine: A wolf in sheep's clothing, or a rich man's adenosylcobalamin?, *Chem. Rev. 103*, 2129-2148.

57. Marquet, A. (2001) Enzymology of carbon-sulfur bond formation, *Curr. Opin. Chem. Biol. 5*, 541-549.

58. Begley, T. P., Xi, J., Kinsland, C., Taylor, S., and McLafferty, F. (1999) The enzymology of sulfur activation during thiamin and biotin biosynthesis, *Curr. Opin. Chem. Biol. 3*, 623-629.

59. Rabeh, W. M., and Cook, P. F. (2004) Structure and mechanism of O-acetylserine sulfhydrylase, *J. Biol. Chem. 279*, 26803-26806.

60. Ikeuchi, Y., Shigi, N., Kato, J., Nishimura, A., and Suzuki, T. (2006) Mechanistic insights into multiple sulfur mediators sulfur relay by involved in thiouridine biosynthesis at tRNA wobble positions, *Mol. Cell. 21*, 97-108.

61. Xi, J., Ge, Y., Kinsland, C., McLafferty, F. W., and Begley, T. P. (2001) Biosynthesis of the thiazole moiety of thiamin in Escherichia coli: Identification of an acyldisulfide-linked protein-protein conjugate that is functionally analogous to the ubiquitin/E1 complex, *Proc. Natl. Acad.Sci. U.S.A. 98*, 8513-8518.

62. Mueller, E. G., Palenchar, P. M., and Buck, C. J. (2001) The role of the cysteine residues of ThiI in the generation of 4-thiouridine in tRNA, *J. Biol. Chem. 276*, 33588-33595.

63. Lauhon, C. T., Erwin, W. M., and Ton, G. N. (2004) Substrate specificity for 4-thiouridine modification in Escherichia coli, *J. Biol. Chem. 279*, 23022-23029.

64. Wright, C. M., Christman, G. D., Snellinger, A. M., Johnston, M. V., and Mueller, E. G. (2006) Direct evidence for enzyme persulfide and disulfide intermediates during 4-thiouridine biosynthesis, *Chem. Commun.*, 3104-3106.

65. Begley, T. P. (2006) Cofactor biosynthesis: an organic chemist's treasure trove, *Nat. Prod. Rep. 23*, 15-25.

66. Dorrestein, P. C., Zhai, H. L., McLafferty, F. W., and Begley, T. P. (2004) The biosynthesis of the thiazole phosphate moiety of thiamin: The sulfur transfer mediated by the sulfur carrier protein ThiS, *Chem. Biol. 11*, 1373-1381.

67. Hazra, A., Chatterjee, A., and Begley, T. P. (2009) Biosynthesis of the Thiamin Thiazole in Bacillus subtilis: Identification of the Product of the Thiazole Synthase-Catalyzed Reaction, *J. Am. Chem. Soc. 131*, 3225-3229.

68. Jurgenson, C. T., Begley, T. P., and Ealick, S. E. (2009) The Structural and Biochemical Foundations of Thiamin Biosynthesis, *Annu. Rev. Biochem. 78*, 569-603.

69. Wang, C. Y., Xi, J., Begley, T. P., and Nicholson, L. K. (2001) Solution structure of ThiS and implications for the evolutionary roots of ubiquitin, *Nat. Struct. Biol. 8*, 47-51.

70. Rudolph, M. J., Wuebbens, M. M., Rajagopalan, K. V., and Schindelin, H. (2001) Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation, *Nat. Struct. Biol. 8*, 42-46.

71. Dye, B. T., and Schulman, B. A. (2007) Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins, *Annu. Rev. Bioph. Biom. 36*, 131-150.

72. Hochstrasser, M. (2000) Evolution and function of ubiquitin-like protein-conjugation systems, *Nat. Cell Biol. 2*, E153-E157.

73. Iyer, L. M., Burroughs, A. M., and Aravind, L. (2006) The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like beta-grasp domains, *Genome Biol. 7*, R60.

74. Hochstrasser, M. (2009) Origin and function of ubiquitin-like proteins, *Nature 458*, 422-429.

75. Leidel, S., Pedrioli, P. G. A., Bucher, T., Brost, R., Costanzo, M., Schmidt, A., Aebersold, R., Boone, C., Hofmann, K., and Peter, M. (2009) Ubiquitin-related modifier Urm1 acts as a sulphur carrier in thiolation of eukaryotic transfer RNA, *Nature 458*, 228-U229.

76. Van der Veen, A. G., Schorpp, K., Schlieker, C., Buti, L., Damon, J. R., Spooner, E., Ploegh, H. L., and Jentsch, S. (2011) Role of the ubiquitin-like protein Urm1 as a noncanonical lysine-directed protein modifier, *Proc. Natl. Acad.Sci. U.S.A. 108*, 1763-1770.

77. Miranda, H. V., Nembhard, N., Su, D., Hepowit, N., Krause, D. J., Pritz, J. R., Phillips, C., Soll, D., and Maupin-Furlow, J. A. (2011) E1-and ubiquitin-like proteins provide a direct link between protein conjugation and sulfur transfer in archaea, *Proc. Natl. Acad.Sci. U.S.A. 108*, 4417-4422.

78. Leimkuhler, S., Wuebbens, M. M., and Rajagopalan, K. V. (2001) Characterization of Escherichia coli MoeB and its involvement in the activation

of molybdopterin synthase for the biosynthesis of the molybdenum cofactor, *J. Biol. Chem. 276*, 34695-34701.

79. Leimkuhler, S., Freuer, A., Araujo, J. A. S., Rajagopalan, K. V., and Mendel, R. R. (2003) Mechanistic studies of human molybdopterin synthase reaction and characterization of mutants identified in group B patients of molybdenum cofactor deficiency, *J. Biol. Chem. 278*, 26127-26134.

80. Marelja, Z., Stocklein, W., Nimtz, M., and Leimkuhler, S. (2008) A novel role for human Nfs1 in the cytoplasm - Nfs1 acts as a sulfur donor for MOCS3, a protein involved in molybdenum cofactor biosynthesis, *J. Biol. Chem. 283*, 25178-25185.

81. Lewis, T. A., Cortese, M. S., Sebat, J. L., Green, T. L., Lee, C. H., and Crawford, R. L. (2000) A Pseudomonas stutzeri gene cluster encoding the biosynthesis of the CCl4-dechlorination agent pyridine-2,6-bis(thiocarboxylic acid), *Environ. Microbiol. 2*, 407-416.

82. Cortese, M. S., Caplan, A. B., and Crawford, R. L. (2002) Structural, functional, and evolutionary analysis of moeZ, a gene encoding an enzyme required for the synthesis of the Pseudomonas metabolite, pyridine-2,6-bis(thiocarboxylic acid), *BMC Evol. Biol. 2*, 8.

83. Godert, A. M., Jin, M., McLafferty, F. W., and Begley, T. P. (2007) Biosynthesis of the thioquinolobactin siderophore: An interesting variation on sulfur transfer, *J. Bacteriol. 189*, 2941-2944.

84. Burns, K. E., Baumgart, S., Dorrestein, P. C., Zhai, H. L., McLafferty, F. W., and Begley, T. P. (2005) Reconstitution of a new cysteine biosynthetic pathway in Mycobacterium tuberculosis, *J. Am. Chem. Soc. 127*, 11602-11603.

85. Krishnamoorthy, K., and Begley, T. P. (2011) Protein Thiocarboxylate-Dependent Methionine Biosynthesis in Wolinella succinogenes, *J. Am. Chem. Soc. 133*, 379-386.

86. Shigi, N., Sakaguchi, Y., Asai, S., Suzuki, T., and Watanabe, K. (2008) Common thiolation mechanism in the biosynthesis of tRNA thiouridine and sulphur-containing cofactors, *Embo J. 27*, 3267-3278.

87. Furukawa, K., Mizushima, N., Noda, T., and Ohsumi, Y. (2000) A protein conjugation system in yeast with homology to biosynthetic enzyme reaction of prokaryotes, *J. Biol. Chem. 275*, 7462-7465.

88. Kerscher, O., Felberbaum, R., and Hochstrasser, M. (2006) Modification of proteins by ubiquitin and ubiquitin-like proteins, *Annu. Rev. Cell Dev. Biol. 22*, 159-180.

89. Jarrett, J. T. (2005) The novel structure and chemistry of iron-sulfur clusters in the adenosylmethionine-dependent radical enzyme biotin synthase, *Arch. Biochem. Biophys. 433*, 312-321.

90. Berkovitch, F., Nicolet, Y., Wan, J. T., Jarrett, J. T., and Drennan, C. L. (2004) Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme, *Science 303*, 76-79.

91. Ugulava, N. B., Sacanell, C. J., and Jarrett, J. T. (2001) Spectroscopic changes during a single turnover of biotin synthase: Destruction of a [2Fe-2S] cluster accompanies sulfur insertion, *Biochemistry 40*, 8352-8358.

92. Rohr, J., and Thiericke, R. (1992) Angucycline group antibiotics, *Nat. Prod. Rep. 9*, 103-137.

93. Hertweck, C., Luzhetskyy, A., Rebets, Y., and Bechthold, A. (2007) Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork, *Nat. Prod. Rep. 24*, 162-190.

94. Imamura, N., Kakinuma, K., Ikekawa, N., Tanaka, H., and Omura, S. (1982) Biosynthesis of vineomycin-A1 and vineomycin-B2, *J. Antibiot. 35*, 602-608.

95. Rohr, J., Beale, J. M., and Floss, H. G. (1989) Urdamycins, new angucycline antibiotics from Streptomyces fradiae. IV. Biosynthetic studies of urdamycins A-D, *J. Antibiot. 42*, 1151-1157.

96. Gould, S. J., Cheng, X. C., and Halley, K. A. (1992) Biosynthesis of dehydrorabelomycin and PD 116740: prearomatic deoxygenation as evidence for different polyketide synthases in the formation of benz[a]anthraquinones, *J. Am. Chem. Soc. 114*, 10066-10068.

97. Seaton, P. J., and Gould, S. J. (1987) Kinamycin biosynthesis. Derivation by excision of an acetate unit from a single-chain decaketide intermediate, *J. Am. Chem. Soc. 109*, 5282-5284.

98. Seaton, P. J., and Gould, S. J. (1988) Origin of the cyanamide carbon of the kinamycin antibiotics, *J. Am. Chem. Soc. 110*, 5912-5914.

99. Takahashi, K., and Tomita, F. (1983) Gilvocarcins, new anti-tumor antibiotics. 5. Biosynthesis of gilvocarcins: incorporation of C13-labeled compounds into gilvocarcin aglycones, *J. Antibiot. 36*, 1531-1535.

100. Carter, G. T., Fantini, A. A., James, J. C., Borders, D. B., and White, R. J. (1985) Biosynthesis of chrysomycin A and B. Origin of the chromophore, *J. Antibiot. 38*, 242-248.

101. Gould, S. J., and Halley, K. a. (1991) Biosynthesis of the benz[a]anthraquinone antibiotic PD 116198: evidence for a rearranged skeleton, *J. Am. Chem. Soc. 113*, 5092-5093.

102. Raty, K., Kunnari, T., Hakala, J., Mantsala, P., and Ylihonko, K. (2000) A gene cluster from Streptomyces galilaeus involved in glycosylation of aclarubicin, *Mol. Gen. Genet. 264*, 164-172.

103. Hoffmeister, D., Ichinose, K., Domann, S., Faust, B., Trefzer, A., Drager, G., Kirschning, A., Fischer, C., Kunzel, E., Bearden, D. W., Rohr, J., and Bechthold, A. (2000) The NDP-sugar co-substrate concentration and the enzyme expression level influence the substrate specificity of glycosyltransferases: cloning and characterization of deoxysugar biosynthetic genes of the urdamycin biosynthetic gene cluster, *Chem. Biol. 7*, 821-831.

104. Niemi, J., and Mäntsälä, P. (1995) Nucleotide sequences and expression of genes from Streptomyces purpurascens that cause the production of new anthracyclines in Streptomyces galilaeus., *J. Bacteriol. 177*, 2942-2945.

105. Westrich, L., Domann, S., Faust, B., Bedford, D., Hopwood, D. A., and Bechthold, A. (1999) Cloning and characterization of a gene cluster from Streptomyces cyanogenus S136 probably involved in landomycin biosynthesis, *FEMS Microbiol. Lett. 170*, 381-387.

106. Ichinose, K., Bedford, D. J., Bibb, M. J., Revill, W. P., and Hopwood, D. A. (1998) The granaticin biosynthetic gene cluster of Streptomyces violaceoruber Tu22: sequence analysis and expression in a heterologous host, *Chem. Biol. 5*, 647-659.

107. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Mannual, 3rd ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

108. Bierman, M., Logan, R., Obrien, K., Seno, E. T., Rao, R. N., and Schoner, B. E. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp, *Gene 116*, 43-49.

109. Han, L., Yang, K. Q., Ramalingam, E., Mosher, R. H., and Vining, L. C. (1994) Cloning and characterization of polyketide synthase genes for jadomycin B biosynthesis in *Streptomyces venezuelae* ISP5230, *Microbiology 140*, 3379-3389.

110. Malpartida, F., and Hopwood, D. A. (1984) Molecular cloning of the whole biosynthetic pathway of a Streptomyces antibiotic and its expression in a heterologous host, *Nature 309*, 462-464.

111. Malpartida, F., Hallam, S. E., Kieser, H. M., Motamedi, H., Hutchinson, C. R., Butler, M. J., Sugden, D. A., Warren, M., McKillop, C., Bailey, C. R., Humphreys, G. O., and Hopwood, D. A. (1987) Homology between streptomyces genes-coding for synthesis of different polyketides used to clone antibiotic biosynthetic genes, *Nature 325*, 818-821.

112. Chen, H. W., Agnihotri, G., Guo, Z. H., Que, N. L. S., Chen, X. M. H., and Liu, H. W. (1999) Biosynthesis of mycarose: Isolation and characterization of enzymes involved in the C-2 deoxygenation, *J. Am. Chem. Soc. 121*, 8124-8125.

113. Draeger, G., Park, S. H., and Floss, H. G. (1999) Mechanism of the 2-deoxygenation step in the biosynthesis of the deoxyhexose moieties of the antibiotics granaticin and oleandomycin, *J. Am. Chem. Soc. 121*, 2611-2612.

114. Zhang, H., White-Phillip, J. A., Melancon, C. E., Kwon, H. J., Yu, W. L., and Liu, H. W. (2007) Elucidation of the kijanimicin gene cluster: Insights into the biosynthesis of spirotetronate antibiotics and nitrosugars, *J. Am. Chem. Soc. 129*, 14670-14683.

115. Hong, L., Zhao, Z. B., and Liu, H. W. (2006) Characterization of spnQ from the spinosyn biosynthetic pathway of Saccharopolyspora spinosa: Mechanistic and evolutionary implications for C-3 deoxygenation in deoxysugar biosynthesis, *J. Am. Chem. Soc. 128*, 14262-14263.

116. Melancon, C. E., and Liu, H. W. (2007) Engineered biosynthesis of macrolide derivatives bearing the non-natural deoxysugars 4-epi-D-mycaminose and 3-N-monomethylamino-3-deoxy-D-fucose, *J. Am. Chem. Soc. 129*, 4896-+.

117. Ishikawa, J., and Hotta, K. (1999) FramePlot: a new implementation of the Frame analysis for predicting protein-coding regions in bacterial DNA with a high G plus C content, *FEMS Microbiol. Lett. 174*, 251-253.

118. Gould, S. J., Cheng, X. C., and Halley, K. A. (1992) Biosynthesis of dehydrorabelomycin and PD 116740: prearomatic deoxygenation as evidence for

different polyketide synthases in the formation of benz[*a*]anthraquinones, *J. Am. Chem. Soc. 114*, 10066-10068.

119. Sato, Y., and Gould, S. J. (1986) Biosynthesis of the kinamycin antibiotics by Streptomyces murayamaensis. Determination of the origin of carbon, hydrogen, and oxygen atoms by carbon-13 NMR spectroscopy, *J. Am. Chem. Soc. 108*, 4625-4631.

120. Sato, Y., and Gould, S. J. (1985) Biosynthesis of kinamycin D. incorporation of [1,2-^{13}C] acetate and of [2-$^{2}H_3$,1-^{13}C]acetate *Tetrahedron Lett. 26*, 4023-4026.

121. Liu, T., Fischer, C., Beninga, C., and Rohr, J. (2004) Oxidative rearrangement processes in the biosynthesis of gilvocarcin V, *J. Am. Chem. Soc. 126*, 12262-12263.

122. Rix, U., Wang, C. C., Chen, Y. H., Lipata, F. M., Rix, L. L. R., Greenwell, L. M., Vining, L. C., Yang, K. Q., and Rohr, J. (2005) The oxidative ring cleavage in jadomycin biosynthesis: A multistep oxygenation cascade in a biosynthetic black box, *Chembiochem 6*, 838-+.

123. Gibson, M., Nur-e-alam, M., Lipata, F., Oliveira, M. A., and Rohr, J. (2005) Characterization of kinetics and products of the Baeyer-Villiger oxygenase MtmOIV, the key enzyme of the biosynthetic pathway toward the natural product anticancer drug mithramycin from Streptomyces argillaceus, *J. Am. Chem. Soc. 127*, 17594-17595.

124. Beam, M. P., Bosserman, M. A., Noinaj, N., Wehenkel, M., and Rohr, J. (2009) Crystal Structure of Baeyer-Villiger Monooxygenase MtmOIV, the Key Enzyme of the Mithramycin Biosynthetic Pathway, *Biochemistry 48*, 4476-4487.

125. Drautz, H., Zähner, H., Rohr, J., and Zeeck, A. (1986) Metabolic products of microorganisms. 234. Urdamycins, new angucycline antibiotics from Streptomyces fradiae. I. Isolation, characterization and biological properties., *J. Antibiot. 39*, 1657-1669.

126. Decker, H., and Haag, S. (1995) Cloning and characterization of a polyketide synthase gene from Streptomyces fradiae Tü2717, which carries the genes for biosynthesis of the angucycline antibiotic urdamycin A and a gene probably involved in its oxygenation, *J. Bacteriol. 177*, 6126-6136.

127. Faust, B., Hoffmeister, D., Weitnauer, G., Westrich, L., Haag, S., Schneider, P., Decker, H., Kunzel, E., Rohr, J., and Bechthold, A. (2000) Two new tailoring enzymes, a glycosyltransferase and an oxygenase, involved in biosynthesis of the

angucycline antibiotic urdamycin A in Streptomyces fradiae Tu2717, *Microbiology-(UK) 146*, 147-154.

128. Fischer, C., Lipata, F., and Rohr, J. (2003) The complete gene cluster of the antitumor agent gilvocarcin V and its implication for the biosynthesis of the gilvocarcins, *J. Am. Chem. Soc. 125*, 7818-7819.

129. Kharel, M. K., Zhu, L. L., Liu, T., and Rohr, J. (2007) Multi-oxygenase complexes of the gilvocarcin and jadomycin biosyntheses, *J. Am. Chem. Soc. 129*, 3780-+.

130. Gould, S. J., and Cheng, X. C. (1994) New benz a anthraquinone secondary metabolites from *Streptomyces phaeochromogenes*, *J. Org. Chem. 59*, 400-405.

131. Wang, L. R., McVey, J., and Vining, L. C. (2001) Cloning and functional analysis of a phosphopantetheinyl transferase superfamily gene associated with jadomycin biosynthesis in Streptomyces venezuelae ISP5230, *Microbiology-(UK) 147*, 1535-1545.

132. Olano, C., Gomez, C., Perez, M., Palomino, M., Pineda-Lucena, A., Carbajo, R. J., Brana, A. F., Mendez, C., and Salas, J. A. (2009) Deciphering Biosynthesis of the RNA Polymerase Inhibitor Streptolydigin and Generation of Glycosylated Derivatives, *Chem. Biol. 16*, 1031-1044.

133. Thorson, J. S., and Liu, H. W. (1993) Characterization of the first PMP-dependent iron-sulfur-containing enzyme which is essential for the biosynthesis of 3,6-dideoxyhexoses, *J. Am. Chem. Soc. 115*, 7539-7540.

134. Hallis, T. M., and Liu, H. W. (1999) Learning nature's strategies for making deoxy sugars: Pathways, mechanisms, and combinatorial applications, *Accounts Chem. Res. 32*, 579-588.

135. He, X. M., Agnihotri, G., and Liu, H. W. (2000) Novel enzymatic mechanisms in carbohydrate metabolism, *Chem. Rev. 100*, 4615-+.

136. Hong, L., Zhao, Z. B., Melancon, C. E., Zhang, H., and Liu, H. W. (2008) In vitro characterization of the enzymes involved in TDP-D-forosamine biosynthesis in the spinosyn pathway of Saccharopolyspora spinosa, *J. Am. Chem. Soc. 130*, 4954-4967.

137. Chang, S., Duerr, B., and Serif, G. (1988) An epimerase-reductase in L-fucose synthesis, *J. Biol. Chem. 263*, 1693-1697.

138. Alam, J., Beyer, N., and Liu, H. W. (2004) Biosynthesis of colitose: Expression, purification, and mechanistic characterization of GDP-4-keto-6-deoxy-D-mannose-3-Dehydrase (ColD) and GDP-L-Colitose synthase (ColC), *Biochemistry 43*, 16450-16460.

139. Trefzer, A., Hoffmeister, D., Kunzel, E., Stockert, S., Weitnauer, G., Westrich, L., Rix, U., Fuchser, J., Bindseil, K. U., Rohr, J., and Bechthold, A. (2000) Function of glycosyltransferase genes involved in urdamycin A biosynthesis, *Chem. Biol. 7*, 133-142.

140. Schuman, B., Alfaro, J. A., and Evans, S. V. (2007) Glycosyltransferase Structure and Function, *Top. Curr. Chem. 272*, 217-257.

141. Ronning, C. M., and Nierman, W. C. (2008) *Myxobacteria: Multicellularity and Differentiation*, ASM Press, Washington, DC.

142. Park, J. H., Dorrestein, P. C., Zhai, H., Kinsland, C., McLafferty, F. W., and Begley, T. P. (2003) Biosynthesis of the thiazole moiety of thiamin pyrophosphate (Vitamin B1), *Biochemistry 42*, 12430-12438.

143. Settembre, E. C., Dorrestein, P. C., Zhai, H. L., Chatterjee, A., McLafferty, F. W., Begley, T. P., and Ealick, S. E. (2004) Thiamin biosynthesis in Bacillus subtilis: Structure of the thiazole synthase/sulfur carrier protein complex, *Biochemistry 43*, 11647-11657.

144. Lehmann, C., Begley, T. P., and Ealick, S. E. (2006) Structure of the Escherichia coli ThiS-ThIF complex, a key component of the sulfur transfer system in thiamin biosynthesis, *Biochemistry 45*, 11-19.

145. Avonce, N., Mendoza-Vargas, A., Morett, E., and Iturriaga, G. (2006) Insights on the evolution of trehalose biosynthesis, *BMC Evol. Biol. 6*.

146. Smith, D. R., DoucetteStamm, L. A., Deloughery, C., Lee, H. M., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K., Harrison, D., Hoang, L., Keagle, P., Lumm, W., Pothier, B., Qiu, D. Y., Spadafora, R., Vicaire, R., Wang, Y., Wierzbowski, J., Gibson, R., Jiwani, N., Caruso, A., Bush, D., Safer, H., Patwell, D., Prabhakar, S., McDougall, S., Shimer, G., Goyal, A., Pietrokovski, S., Church, G. M., Daniels, C. J., Mao, J. I., Rice, P., Nolling, J., and Reeve, J. N. (1997) Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: Functional analysis and comparative genomics, *J. Bacteriol. 179*, 7135-7155.

147. Walsh, C. (2000) Molecular mechanisms that confer antibacterial drug resistance, *Nature 406*, 775-781.

148. Rohr, J. (1989) Biosynthetic formation of the S-methyl group of the angucycline antibiotic urdamycin E, *J. Chem. Soc.-Chem. Commun.*, 492-493.

149. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding., *Anal. Biochem. 72*, 248-254.

150. Deferrari, J. O., Gros, E. G., and Mastronardi, I. O. (1967) Methylation of carbohydrates bearing base-labile substituents, with diazomethane-boron trifludride etherate II. A new synthesis of 2-*O*-methyl-D-mannose, *Carbohydr. Res. 4*, 432-434.

151. Kovac, P. (1986) A short synthesis of 2-deoxy-2-fluoro-D-glucose, *Carbohydr. Res. 153*, 168-170.

152. Hamacher, K. (1984) Phase-transfer catalyzed synthesis of 4-S-beta-D-glucopyranosyl-4-thio-D-glucopyranose (thiocellobiose) and 2-S-beta-D-glucopyranosyl-2-thio-D-glucopyranose (thiosophorose), *Carbohydr. Res. 128*, 291-295.

153. Pavliak, V., and Kovac, P. (1991) Synthesis of ligands related to the *O*-specific antigen of shigella-dysenteriae type .1. a short synthesis of 1,3,4,6-tetra-*O*-acetyl-2-azido-2-deoxy-beta-D-glucopyranose and the corresponding alpha-glucosyl chloride from D-mannose, *Carbohydr. Res. 210*, 333-337.

154. Knapp, S., and Kirk, B. A. (2003) Glycosylation with 2 '-thio-S-acetyl participation, *Tetrahedron Lett. 44*, 7601-7605.

155. Knapp, S., Naughton, A. B. J., Jaramillo, C., and Pipik, B. (1992) Reactions of some pyranoside diol monotriflates with nucleophiles and bases, *J. Org. Chem. 57*, 7328-7334.

156. Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M., and Omura, S. (2003) Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis, *Nat. Biotechnol. 21*, 526-531.

157. Oliynyk, M., Samborskyy, M., Lester, J. B., Mironenko, T., Scott, N., Dickens, S., Haydock, S. F., and Leadlay, P. F. (2007) Complete genome sequence of the

erythromycin-producing bacterium Saccharopolyspora erythraea NRRL23338, *Nat. Biotechnol. 25*, 447-453.

158. Qu, Q. H., Lee, S. J., and Boos, W. (2004) TreT, a novel trehalose glycosyltransferring synthase of the hyperthermophilic Archaeon Thermococcus litoralis, *J. Biol. Chem. 279*, 47890-47897.

159. Ryu, S. I., Park, C. S., Cha, J., Woo, E. J., and Lee, S. B. (2005) A novel trehalose-synthesizing glycosyltransferase from Pyrococcus horikoshii: Molecular cloning and characterization, *Biochem. Biophys. Res. Commun. 329*, 429-436.

160. Kouril, T., Zaparty, M., Marrero, J., Brinkmann, H., and Siebers, B. (2008) A novel trehalose synthesizing pathway in the hyperthermophilic Crenarchaeon Thermoproteus tenax: the unidirectional TreT pathway, *Arch. Microbiol. 190*, 355-369.

161. Nobre, A., Alarico, S., Fernandes, C., Empadinhas, N., and da Costa, M. S. (2008) A Unique Combination of Genetic Systems for the Synthesis of Trehalose in Rubrobacter xylanophilus: Properties of a Rare Actinobacterial TreT, *J. Bacteriol. 190*, 7939-7946.

162. Silva, Z., Alarico, S., and da Costa, M. (2005) Trehalose biosynthesis in Thermus thermophilus RQ-1: biochemical properties of the trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatase., *Extremophiles 9*, 29-36.

163. Pan, Y. T., Carroll, J. D., and Elbein, a. D. (2002) Trehalose-phosphate synthase of Mycobacterium tuberculosis, *Eur. J. Biochem. 269*, 6091-6100.

164. Gibson, R. P., Lloyd, R. M., Charnock, S. J., and Davies, G. J. (2002) Characterization of Escherichia coli OtsA, a trehalose-6-phosphate synthase from glycosyltransferase family 20, *Acta Crystallogr. D. 58*, 349-351.

165. Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R., and Davies, G. J. (2002) Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA., *Chem. Biol. 9*, 1337-1346.

166. Gibson, R. P., Tarling, C. a., Roberts, S., Withers, S. G., and Davies, G. J. (2004) The donor subsite of trehalose-6-phosphate synthase: binary complexes with UDP-glucose and UDP-2-deoxy-2-fluoro-glucose at 2 A resolution., *J. Biol. Chem. 279*, 1950-1955.

167. Errey, J. C., Lee, S. S., Gibson, R. P., Fleites, C. M., Barry, C. S., Jung, P. M. J., O'Sullivan, A. C., Davis, B. G., and Davies, G. J. (2010) Mechanistic Insight into

Enzymatic Glycosyl Transfer with Retention of Configuration through Analysis of Glycomimetic Inhibitors, *Angew. Chem., Int. Ed. 49*, 1234-1237.

168. Dorrestein, P. C., Zhai, H. L., Taylor, S. V., McLafferty, F. W., and Begley, T. P. (2004) The biosynthesis of the thiazole phosphate moiety of thiamin (Vitamin B-1): The early steps catalyzed by thiazole synthase, *J. Am. Chem. Soc. 126*, 3091-3096.

169. Mouilleron, S., Badet-Denisot, M. A., and Golinelli-Pimpaneau, B. (2006) Glutamine binding opens the ammonia channel and activates glucosamine-6P synthase, *J. Biol. Chem. 281*, 4404-4412.

170. Leriche, C., BadetDenisot, M. A., and Badet, B. (1996) Characterization of a phosphoglucose isomerase-like activity associated with the carboxy-terminal domain of Escherichia coli glucosamine-6-phosphate synthase, *J. Am. Chem. Soc. 118*, 1797-1798.

171. Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S. C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N. M., Choi, S. K., Codani, J. J., Connerton, I. F., Cummings, N. J., Daniel, R. A., Denizot, F., Devine, K. M., Dusterhoft, A., Ehrlich, S. D., Emmerson, P. T., Entian, K. D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S. Y., Glaser, P., Goffeau, A., Golightly, E. J., Grandi, G., Guiseppi, G., Guy, B. J., Haga, K., Haiech, J., Harwood, C. R., Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M. F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., KlaerrBlanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S. M., Levine, A., Liu, H., Masuda, S., Mauel, C., Medigue, C., Medina, N., Mellado, R. P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., Oreilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S. H., Parro, V., Pohl, T. M., Portetelle, D., Porwollik, S., Prescott, A. M., Presecan, E., Pujic, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S. J., Serror, P., Shin, B. S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A., Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weitzenegger, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H. F., Zumstein, E., Yoshikawa, H., and Danchin, A. (1997) The

complete genome sequence of the Gram-positive bacterium Bacillus subtilis, *Nature 390*, 249-256.

172. Watanabe, C. M. H., and Townsend, C. A. (1998) The in vitro conversion of norsolorinic acid to aflatoxin B-1. An improved method of cell-free enzyme preparation and stabilization, *J. Am. Chem. Soc. 120*, 6231-6239.

173. Delcher, A. L., Bratke, K. A., Powers, E. C., and Salzberg, S. L. (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer, *Bioinformatics 23*, 673-679.

174. Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction, *BMC Bioinformatics 9*, 40.

175. Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction, *Nat. Protoc. 5*, 725-738.

176. Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J., and Hopwood, D. A. (2002) Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2), *Nature 417*, 141-147.

177. Ohnishi, Y., Ishikawa, J., Hara, H., Suzuki, H., Ikenoya, M., Ikeda, H., Yamashita, A., Hattori, M., and Horinouchi, S. (2008) Genome sequence of the streptomycin-producing microorganism Streptomyces griseus IFO 13350, *J. Bacteriol. 190*, 4050-4060.

178. Schwarz, G., and Mendel, R. R. (2006) Molybdenum cofactor biosynthesis and molybdenum enzymes, *Annu. Rev. Plant Biol. 57*, 623-647.

179. Schwarz, G., Mendel, R. R., and Ribbe, M. W. (2009) Molybdenum cofactors, enzymes and pathways, *Nature 460*, 839-847.

180. Geoghegan, K. F., Dixon, H. B. F., 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., and Stroh, J. G. (1999) Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: The cause of extra mass of 258 or 178 Da in fusion proteins, *Anal. Biochem. 267*, 169-184.

181. Burroughs, A. M., Iyer, L. M., and Aravind, L. (2009) Natural history of the E1-like superfamily: Implication for adenylation, sulfur transfer, and ubiquitin conjugation, *Proteins 75*, 895-910.

182. Newton, G. L., Fahey, R. C., Cohen, G., and Aharonowitz, Y. (1993) Low-molecular-weight thiols in streptomycetes and their potential role as antioxidants, *J. Bacteriol. 175*, 2734-2742.

183. Newton, G. L., and Fahey, R. C. (1995) Determination of Biothiols by Bromobimane Labeling and High-Performance Liquid-Chromatography, *Methods Enzymol.251*, 148-166.

184. Kaiser, J. T., Bruno, S., Clausen, T., Huber, R., Schiaretti, F., Mozzarelli, A., and Kessler, D. (2003) Snapshots of the cystine lyase C-DES during catalysis - Studies in solution and in the crystalline state, *J. Biol. Chem. 278*, 357-365.

185. Campanini, B., Schiaretti, F., Abbruzzetti, S., Kessler, D., and Mozzarelli, A. (2006) Sulfur mobilization in cyanobacteria: The catalytic mecahnism of L-cystine C-S lyase (C-DES) from synechocystis, *J. Biol. Chem. 281*, 38769-38780.

186. Newton, G. L., Bewley, C. A., Dwyer, T. J., Horn, R., Aharonowitz, Y., Cohen, G., Davies, J., Faulkner, D. J., and Fahey, R. C. (1995) The structure of U17 isolated from Streptomyces clavuligerus and its properties as an antioxidant thiol, *Eur. J. Biochem. 230*, 821-825.

187. Newton, G. L., Arnold, K., Price, M. S., Sherrill, C., Delcardayre, S. B., Aharonowitz, Y., Cohen, G., Davies, J., Fahey, R. C., and Davis, C. (1996) Distribution of thiols in microorganisms: Mycothiol is a major thiol in most actinomycetes, *J. Bacteriol. 178*, 1990-1995.

188. Herr, R. R., and Bergy, M. E. (1962) Lincomycin, a new antibiotic. II. Isolation and characterization., *Antimicrob. Agents Chemother.*, 560-564.

189. Fitzhugh, A. L. (1998) Antibiotic inhibitors of the peptidyl transferase center. 1. Clindamycin as a composite analogue of the transfer RNA fragments L-Pro-Met and the D-ribosyl ring of adenosine, *Bioorg. Med. Chem. Lett. 8*, 87-92.

190. Lewis, C., Clapp, H. W., and Grady, J. E. (1963) In vitro and in vivo evaluation of lincomycin, a new antibiotic, *Antimicrob. Agents Chemother.*, 570-582.

191. Magerlei.Bj, and Kagan, F. (1969) Lincomycin. 8. 4'-Alkyl-1'-demethyl-4'-depropylclindamycins, potent antibacterial and antimalarial agents, *J. Med. Chem. 12*, 780-784.

192. Birkenme.Rd, and Kagan, F. (1970) Lincomycin. XI. Synthesis and structure of clindamycin, a potent antibacterial agent, *J. Med. Chem. 13*, 616-619.

193. Douthwaite, S. (1992) Interaction of the antibiotics clindamycin and lincomycin with *Escheichia coli* 23S ribosomal RNA, *Nucleic Acids Res. 20*, 4717-4720.

194. Spízek, J., and Rezanka, T. (2004) Lincomycin, cultivation of producing strains and biosynthesis, *Appl. Microbiol. Biotechnol. 63*, 510-519.

195. Neusser, D., Schmidt, H., Spizek, J., Novotna, J., Peschke, U., Kaschabeck, S., Tichy, P., and Piepersberg, W. (1998) The genes lmbB1 and lmbB2 of Streptomyces lincolnensis encode enzymes involved in the conversion of L-tyrosine to propylproline during the biosynthesis of the antibiotic lincomycin A, *Arch. Microbiol. 169*, 322-332.

196. Novotna, J., Honzatko, A., Bednar, P., Kopecky, J., Janata, J., and Spizek, J. (2004) L-3,4-Dihydroxyphenyl alanine-extradiol cleavage is followed by intramolecular cyclization in lincomycin biosynthesis, *Eur. J. Biochem. 271*, 3678-3683.

197. Novotna, G., and Janata, J. (2006) A new evolutionary variant of the streptogramin A resistance protein, vga(A)(LC), from Staphylococcus haemolyticus with shifted substrate specificity towards lincosamides, *Antimicrob. Agents Chemother. 50*, 4070-4076.

198. Brahme, N. M., Gonzalez, J. E., Mizsak, S., Rolls, J. R., Hessler, E. J., and Hurley, L. H. (1984) Biosynthesis of the Lincomycins .2. Studies Using Stable Isotopes on the Biosynthesis of Methylthiolincosaminide Moiety of Lincomycin-A, *J. Am. Chem. Soc. 106*, 7878-7883.

199. Piepersberg, W., and Distler, J. (1997) Aminoglycosides and sugar components in other secondary metabolites, In *Biotechnology* (Rehm, H. J., and Reed, G., Eds.) 2nd ed., pp 397-488, Wiley-VCH, Weinheim.

200. Piepersberg, W. (2001) Glycosylation of antibiotics and other agents from actinomycetes, In *Novel frontiers in the production of compounds for biomedical use* (van Broekhoven, A., Shapiro, F., and Anne, J., Eds.), pp 161-167, Kluwer Academic, Dordrecht.

201. Spizek, J., Novotna, J., and Rezanka, T. (2004) Lincosamides: Chemical structure, biosynthesis, mechanism of action, resistance, and applications, *Adv. Appl. Microbiol. 56*, 121-154.

202. Vara, J., Lewandowskaskarbek, M., Wang, Y. G., Donadio, S., and Hutchinson, C. R. (1989) Cloning of genes governing the deoxysugar portion of the erythromycin biosynthesis pathway in Saccharopolyspora erythraea (Streptomyces erythreus), *J. Bacteriol. 171*, 5872-5881.

203. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000) *Practical streptomyces genetics*, The John Innes Foundation, Norwich, England.

204. Ahn, M., Cochrane, F. C., Patchett, M. L., and Parker, E. J. (2008) Arabinose 5-phosphate analogues as mechanistic probes for Neisseria meningitidis 3-deoxy-D-manno-octulosonate 8-phosphate synthase, *Bioorg. Med. Chem. 16*, 9830-9836.

205. Wang, Z. X., Wiebe, L. I., Balzarini, J., De Clercq, E., and Knaus, E. E. (2000) Chiral synthesis of 4-[1-(2-deoxy-beta-L-ribofuranosyl)] derivatives of 2-substituted 5-fluoroaniline: "Cytosine replacement" analogues of deoxy-beta-L-cytidine, *J. Org. Chem. 65*, 9214-9219.

206. White, J. D., and Jeffrey, S. C. (1996) Synthesis of the tricarbonyl subunit (C-8-C-19) of rapamycin via tandem chan rearrangement-oxidation, *J. Org. Chem. 61*, 2600-2601.

207. Lee, S., Kirschning, A., Muller, M., Way, C., and Floss, H. G. (1999) Enzymatic synthesis of [7-C-14, 7-H-3]- and [1-C-13]sedoheptulose 7-phosphate and [1-C-13]ido-heptulose 7-phosphate, *J. Mol. Catal. B-Enzym. 6*, 369-377.

208. Koberska, M., Kopecky, J., Olsovska, J., Jelinkova, M., Ulanova, D., Man, P., Flieger, M., and Janata, J. (2008) Sequence Analysis and Heterologous Expression of the Lincomycin Biosynthetic Cluster of the Type Strain Streptomyces lincolnensis ATCC 25466, *Folia Microbiol. 53*, 395-401.

209. Witz, D. F., Hessler, E. J., and Miller, T. L. (1971) Bioconversion of tyrosine into the propylhygric acid moiety of lincomycin, *Biochemistry 10*, 1128-1133.

210. Hurley, L. H. (1980) Elucidation and Formulation of Novel Biosynthetic Pathways Leading to the Pyrrolo[1,4]Benzodiazepine Antibiotics Anthramycin, Tomaymycin, and Sibiromycin, *Accounts Chem. Res. 13*, 263-269.

211. Hurley, L. H., Lasswell, W. L., Ostrander, J. M., and Parry, R. (1979) Pyrrolo[1,4]Benzodiazepine Antibiotics - Biosynthetic Conversion of Tyrosine to the C2-Proline and C3-Proline Moieties of Anthramycin, Tomaymycin, and Sibiromycin, *Biochemistry 18*, 4230-4237.

212. Hu, Y., Phelan, V., Ntai, I., Farnet, C. M., Zazopoulos, E., and Bachmann, B. O. (2007) Benzodiazepine biosynthesis in Streptomyces refuineus, *Chem. Biol. 14*, 691-701.

213. Gerratana, B., Li, W., Khullar, A., Chou, S., and Sacramo, A. (2009) Biosynthesis of Sibiromycin, a Potent Antitumor Antibiotic, *Appl. Environ. Microbiol. 75*, 2869-2878.

214. Gerratana, B., Li, W., Chou, S. C., and Khullar, A. (2009) Cloning and Characterization of the Biosynthetic Gene Cluster for Tomaymycin, an SJG-136 Monomeric Analog, *Appl. Environ. Microbiol. 75*, 2958-2963.

215. Takayama, S., McGarvey, G. J., and Wong, C. H. (1997) Microbial aldolases and transketolases: New biocatalytic approaches to simple and complex sugars, *Annu. Rev. Microbiol. 51*, 285-310.

216. Fitz, W., Schwark, J. R., and Wong, C. H. (1995) Aldotetroses and C(3)-Modified Aldohexoses as Substrates for N-Acetylneuraminic Acid Aldolase - a Model for the Explanation of the Normal and the Inversed Stereoselectivity, *J. Org. Chem. 60*, 3663-3670.

217. Valvano, M. A., Messner, P., and Kosma, P. (2002) Novel pathways for biosynthesis of nucleotide-activated glycero-manno-heptose precursors of bacterial glycoproteins and cell surface polysaccharides, *Microbiology-(UK) 148*, 1979-1989.

218. Kneidinger, B., Graninger, M., Puchberger, M., Kosma, P., and Messner, P. (2001) Biosynthesis of nucleotide-activated D-glycero-D-manno-heptose, *J. Biol. Chem. 276*, 20935-20944.

219. Kneidinger, B., Marolda, C., Graninger, M., Zamyatina, A., McArthur, F., Kosma, P., Valvano, M. A., and Messner, P. (2002) Biosynthesis pathway of ADP-L-glycero-beta-D-manno-heptose in Escherichia coli, *J. Bacteriol. 184*, 363-369.

220. Brooke, J. S., and Valvano, M. A. (1996) Biosynthesis of inner core lipopolysaccharide in enteric bacteria identification and characterization of a conserved phosphoheptose isomerase, *J. Biol. Chem. 271*, 3608-3614.

221. Brooke, J. S., and Valvano, M. A. (1996) Molecular cloning of the Haemophilus influenzae gmhA (lpcA) gene encoding a phosphoheptose isomerase required for lipooligosaccharide biosynthesis, *J. Bacteriol. 178*, 3339-3341.

222. Beis, K., Allard, S. T. M., Hegeman, A. D., Murshudov, G., Philp, D., and Naismith, J. H. (2003) The structure of NADH in the enzyme dTDP-D-glucose dehydratase (RmIB), *J. Am. Chem. Soc. 125*, 11872-11878.

223. Steffek, M., Newton, G. L., Av-Gay, Y., and Fahey, R. C. (2003) Characterization of Mycobacterium tuberculosis mycothiol S-conjugate amidase, *Biochemistry 42*, 12067-12076.

224. Racker, E., and Schroeder, E. (1957) Formation and utilization of octulose-8-phosphate by transaldolase and transketolase, *Arch. Biochem. Biophys. 66*, 241-243.

225. Frey, P. A. (1996) The Leloir pathway: A mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose, *Faseb J. 10*, 461-470.

226. Hegeman, A. D., Gross, J. W., and Frey, F. A. (2001) Probing catalysis by Escherichia coli dTDP-glucose-4,6-dehydratase: Identification and preliminary characterization of functional amino acid residues at the active site, *Biochemistry 40*, 6598-6610.

227. Hofmann, C., Boll, R., Heitmann, B., Hauser, G., Durr, C., Frerich, A., Weitnauer, G., Glaser, S. J., and Bechthold, A. (2005) Genes encoding enzymes responsible for biosynthesis of L-lyxose and attachment of eurekanate during avilamycin biosynthesis, *Chem. Biol. 12*, 1137-1143.

228. Breazeale, S. D., Ribeiro, A. A., and Raetz, C. R. H. (2002) Oxidative decarboxylation of UDP-glucuronic acid in extracts of polymyxin-resistant Escherichia coli - Origin of lipid a species modified with 4-amino-4-deoxy-L-arabinose, *J. Biol. Chem. 277*, 2886-2896.

229. Kalia, J., and Raines, R. T. (2008) Hydrolytic stability of hydrazones and oximes, *Angew. Chem., Int. Ed. 47*, 7523-7526.

230. Hang, H. C., and Bertozzi, C. R. (2001) Chemoselective approaches to glycoprotein assembly, *Accounts Chem. Res. 34*, 727-736.

Vita

Eita Sasaki was born in August, 1980, in Matsuyama, Japan. He grew up in Kawasaki and Yokohama cities with his parents, an older sister, a younger sister, and a younger brother. Eita graduated Shonan High School in March, 1999 and attended the University of Tokyo (Natural Sciences I at the College of Arts and Sciences). After completing the junior division in two years, he was admitted to the Faculty of Pharmaceutical Sicences for his senior education. He joined Prof. Tetsuo Nagano's research group and started his research project, the development of near-infrared fluorescent probes for nitric oxide, under the supervision of Dr. Hirotatsu Kojima. Eita passed the National Pharmacist Qualifying Examination in March, 2003 and continued his research at the Graduate School of Pharmaceutical Sciences at the University of Tokyo. His three-year research experience in Prof. Nagano's lab resulted in two research articles (first- and third-author) and one international patent. After obtaining the M.S. degree in 2005, Eita decided to pursue his doctoral studies overseas. He was admitted to the Department of Chemistry and Biochemistry at the University of Texas at Austin and joined Dr. Hung-wen (Ben) Liu's group in May, 2005. He received the BASF fellowship in his first year and was a teaching assistant for Organic Chemistry Lab and Organic Chemistry II for two years. His research on the biosynthetic studies of the 2-thiosugar-containing natural product, BE-7585A, led to two first-author publications by the time of submission of this dissertation, and several other research articles are in preparation.

Permanent email: e.sasaki@utexas.edu

This dissertation was typed by Eita Sasaki.