

CLONING AND CHARACTERIZATION OF NEOCARZINOSTATIN GENE CLUSTER FROM *STREPTOMYCES CARZINOSTATICUS*

Ju Yun Bae

Under the supervision of Professor Ben Shen

at the University of Wisconsin-Madison

ABSTRACT

1. Overexpression of *svp* and purification of Svp from *Escherichia coli*

All carrier proteins must be converted from the inactive apo-forms to the functional holo-forms by the covalent attachment of the phosphopantetheine (P-pant) group. This posttranslational modification is catalyzed by a superfamily of enzymes known as phosphopantetheinyl transferase (PPTase). For structural characterization a PPTase from *Streptomyces verticillus*, Svp, was overproduced and purified in *E. coli* and sent to a collaborator for crystallography.

2. Cloning and characterization of neocarzinostatin (NCS) gene cluster from *Streptomyces carzinostaticus*

The NCS biosynthetic gene cluster was identified by using *ncsA* as a probe. After screening the cosmid library, the DNA fragments contain the *ncsA* gene were cloned. Among several recombinant clones, two were confirmed by

Southern blotting. A 7.5 kb fragment (37II-2) was cloned from one cosmid and was digested with *Bam*HI for subcloning and sequencing. Among 5 fragments, the 0.8 kb *Bam*HI fragment from 37II-2 revealed an incomplete open reading frame (ORF) whose deduced product resemble a dNDP-glucose 4,6-dehydratase. To clone the dNDP-glucose 4,6-dehydratase gene from 37II-2, PCR was performed with primers designed according to the sequence data. The amplified PCR product was sequenced, confirming the identity as a putative dNDP-glucose 4,6-dehydratase, and ligated to pOJ260 and pKC1139 for gene disruption with intent at determining if the dNDP-glucose 4,6-dehydratase was regarded for NCS biosynthesis.

Chromosomal walking from the 2 positive cosmid was also undertaken for cloning the whole NCS gene cluster by using two ends covering 64 kb DNA fragment. After extension to 140 kb, the DNA was sent to Ecopia Biosciences for sequencing.

3. Functional analysis of the NCS biosynthetic gene cluster

To identify the function of putative deoxyaminosugar biosynthetic genes, both *in vivo* gene replacement and *in vitro* biochemical experiments were pursued. GDP-mannose-4,6-dehydratase (*ncsC1*), GDP-mannose-2-dehydratase (*ncsC2*), and phenylalanine ammonia lyase (*ncsC3*) were targeted for gene replacement. For *in vitro* assay, six genes putatively involved in aminosugar biosynthesis were recombinantly produced in *Escherichia coli*: GDP-mannose synthase (NcsC), GDP-mannose-4,6-dehydratas (NcsC1), GDP-mannose-2-dehydratase

(NcsC2), phenylalanine ammonia lyase (NcsC3), GDP-hexose 4-ketoreductase (NcsC4) and N-methyltransferase (NcsC5).

Gene replacement of *ncsC1*, *ncsC2*, and *ncsC3* was not successful largely due to the presence of three copies of NCS gene cluster in NCS producer. For heterologous expression in *E. coli*, NcsC1, NcsC2, and NcsC3 were partially soluble and isolated as His-tagged construct. NcsC1 was assayed with possible substrates, TDP-glucose and GDP-mannose, but no activity was observed based on HPLC analysis. Expression of *ncsC* in *Streptomyces lividans* TK64 was also attempted. GDP-mannose synthase activity was detected with the crude cell free extract and GTP and mannose-1-phosphate as substrates, on the basis of a HPLC assay. However, extensive control experiments showed that this activity resulted from the *S. lividans* host background, failing to establish NcsC as a GDP-mannose synthase.

ACKNOWLEDGMENTS

First of all, I want to thank Prof. Ben Shen for his support and patience on my graduate study and research. His advices and suggestions were always challenge to me.

Many thanks are extended to my lab colleagues and my fellow students in the Pharmaceutical Science division who gave me much help and enjoyable time.

Thanks to my thesis committee members, Prof. Jon Thorson and Charles Lauhon for evaluation of the draft and comments about my research.

Finally, I appreciate the faith and love of my family and friends. Without them, I cannot stand here. I hope I improve myself to prove my love for them.

TABLE OF CONTENTS

ABSTRACT.....	i
ACKNOWLEDGEMENTS.....	iv
TABLE OF CONTENTS.....	v
LIST OF SYMBOLS AND ABBREVIATIONS.....	viii

CHAPTER 1: Overexpression of *svp* and purification of Svp from *E. coli*.....1

1. 1 Introduction.....	2
1.2. Materials and Methods.....	5
1.2.1. Bacterial strains, plasmids, and chemicals.....	5
1.2.2. General procedures.....	5
1.2.3. Protein overproduction and purification.....	5
1.3. Results.....	6
1.3.1. Overexpression in <i>E. coli</i> and purification of the Svp protein.....	6
1.4. References.....	7

Chapter 2: Cloning and characterization of neocarzinostatin (NCS)

gene cluster from *Streptomyces carzinostaticus*.....10

2.1. Introduction.....	11
2.2. Materials and Methods.....	16

2.2.1. DNA isolation and manipulation.....	16
2.2.2. DNA Sequencing.....	16
2.2.3. PCR.....	17
2.3. Results and Discussion.....	18
2.3.1. Cloning of NCS biosynthetic gene cluster.....	18
2.3.2. Subcloning and sequencing of Ncs cluster.....	21
2.3.3. Gene disruption of dNDP-glucose 4,6-dehydratase.....	23
2.4. References.....	25
Chapter 3: Functional analysis of NCS biosynthetic gene cluster.....	28
3.1. Introduction.....	29
3.1.1. Studies on NCS deoxyaminosugar biosynthetic pathway.....	29
3.2. Materials and Methods.....	37
3.2.1. Bacterial strains, DNA isolation, and manipulation.....	37
3.2.2. DNA Sequencing.....	37
3.2.3. PCR for <i>in vivo</i> assay.....	38
3.2.4. Conjugation and transformation.....	39
3.2.5. PCR for protein production in <i>E. coli</i>	40
3.2.6. Overproduction and purification of proteins.....	42
3.2.7. <i>In vitro</i> assay.....	45
3.2.8. HPLC methods.....	46
3.3. Results and Discussions.....	47

3.3.1. Gene disruption of dNDP-glucose 4,6-dehydratase.....	47
3.3.2. Construction of recombinant plasmids pJY-II38, 39, and 40 for gene replacement.....	48
3.3.3. Overproduction and purification of deoxyaminosugar biosynthetic gene cluster in <i>E. coli</i>	50
3.3.4. Enzyme assay.....	54
3.3.5. Expression of <i>ncsC</i> in <i>Streptomyces</i>	55
3.3.6. Enzyme assay with crude extract from <i>Streptomyces</i>	56
3.4. References.....	59

LIST OF SYMBOLS AND ABBREVIATIONS

ACP	Acyl carrier proteins
Apr	Apramycin
ArCP	Aryl carrier proteins
BLM	Bleomycin
CoA	Coenzyme A
DH	NDP-hexose-4,6-dehydratase
DNA	Deoxyribonucleic acid
EPS	Exopolysaccharide
FAS	Fatty acid synthases
GDP	Guanosine diphosphate
GTP	Guanosine triphosphate
HPLC	High-performance liquid chromatography
Kb	Kilobase
KDa	Kilodalton
LB	Luria-Bertani media
LPS	Lipopolysaccharides
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NCS	Neocarzinostatin
NDP	Nucleotide diphosphate
Ni-NTA	Nickel-nitrilotriacetic acid resin

NRPS	Nonribosomal peptide synthetases
PCP	Peptidyl carrier proteins
PCR	Polymerase chain reaction
PKS	Polyketide synthases
P-pant	4'-phosphopantetheine
PPTase	Phosphopantetheinyl transferase
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TDP	Thymidine diphosphate
TTP	Thymidine triphosphate

CHAPTER 1: Overexpression of *svp* and purification of Svp from *Escherichia coli*

1. 1 Introduction

Fatty acid synthases (FASs), polyketide synthases (PKSs), and nonribosomal peptide synthetases (NRPSs) typically contain one or more small proteins, either as integrated domains (type I) or as discrete subunits (type II), that function as carrier proteins for the growing intermediate in the biosynthesis of fatty acids, polyketides, and nonribosomal peptides, respectively. Depending on the nature of the growing intermediate, these carrier proteins are called acyl carrier proteins (ACPs), aryl carrier proteins (ArCPs), or peptidyl carrier proteins (PCPs). A characteristic feature of carrier proteins is the covalently bound 4'-phosphopantetheine (P-pant) prosthetic group. The P-pant group plays a dual role in these multienzyme systems – acting as a nucleophile to activate acyl groups by a thioester linkage to the terminal thiol and providing a flexible arm to channel the growing intermediates between the neighboring active sites [1,2]. In order to be functional, all carrier proteins must be converted from the inactive apo-forms to the functional holo-forms by the covalent attachment of the P-pant group to a conserved serine residue of the carrier proteins. This posttranslational modification is catalyzed by a superfamily of enzymes known as phosphopantetheinyl transferase (PPTase) that derives the P-pant group from coenzyme A (CoA) (Fig. 1) [3,4].

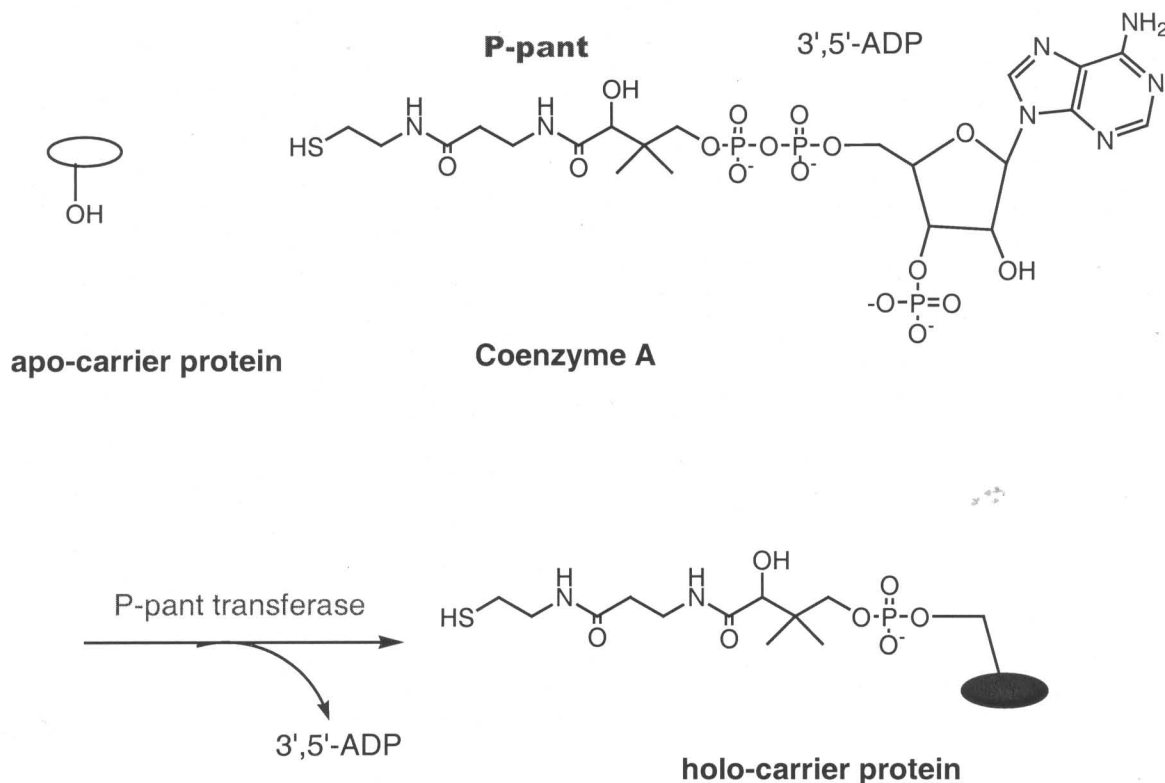


Figure 1. Activation of polyketide and polypeptide synthases by the covalent attachment of 4'-phosphopantetheine (P-pant).

PPTases have recently been re-classified, on a basis of amino acid sequence, into two subfamilies – ACPS-type and Sfp-type. The former is generally involved in fatty acid biosynthesis and, hence, is essential for cell survival. The latter is generally involved in secondary metabolite biosynthesis and hence is non-essential, and as a result may have evolved relaxed substrate specificity [5,6].

The *Bacillus subtilis* Sfp, the most studied PPTases to date, activates the peptidyl carrier protein (PCP) domains of surfactin synthetase by transferring the 4'-phosphopantetheinyl moiety of CoA to a conserved serine residue. It has also

been shown that Sfp activates several carrier proteins both ACPs and PCPs [7]. The wide spectrum of PCP substrates renders Sfp a biotechnologically valuable enzyme for use in combinatorial non-ribosomal peptide synthesis. The structure of the Sfp-CoA complex was determined at 1.8 Å resolution to reveal a novel α/β -fold containing an unexpected intramolecular 2-fold pseudosymmetry [8]. As proposed, this new fold and dimer structure may explain differences within the two PPTase families. The active site of Sfp accommodates a magnesium ion, which is complexed by the CoA pyrophosphate, the side chains of three acidic amino acids and one water molecule. The structure reveals regions likely to be involved in the interaction with the PCP substrate.

The biosynthesis of bleomycin (BLM) in *Streptomyces verticillus* ATCC15003 has been studied as a model system to investigate intermodular communication between NRPS and PKS modules [9-13]. BLM is an antitumor drug belonging to a group of natural products of hybrid peptide-polyketide origin, synthesized by a hybrid NRPS-PKS, the Blm NRPS/PKS/NRPS megasynthetase [11]. The *Streptomyces verticillus* PPTase gene (*svp*) was cloned, overexpressed, and the recombinant protein was purified and characterized biochemically as a PPTase [14]. Based on sequence analysis and activity characterization, Svp is an Sfp-like PPTase and produces a second example of a PPTase with broad substrate use.

Structural determination of new Sfp-like PPTases is imperative to understand the interactions of PPTases and the PCP substrate, and to elucidate a catalytic mechanism thereby explaining substrate specificity. Therefore, a large quantity of Svp was overproduced, purified, and sent to our collaborator for structural

determination by X-ray crystallography to produce insights into Svp's promiscuous nature.

1.2. Materials and Methods

1.2.1. Bacterial strains, plasmids, and chemicals

Bacterial strains used were *Escherichia coli* DH5 α , *E. coli* M15 (pREP) (Qiagen, Santa Clarita, CA), and *Streptomyces verticillus* ATCC15003 (American Type Culture Collection, Rockville, MD). The expression vector pQE70 was from Qiagen. Restriction enzymes were purchased from New England Biolabs (Beverly, MA) and Ni-NTA resin was from Qiagen.

1.2.2. General procedures

Plasmid preparation was carried out by using a commercial kit (Qiagen). Restriction digests, ligations, and transformations were done by standard methods [15].

1.2.3. Protein overproduction and purification

Expression of *svp* and overproduction and purification of Svp in *E. coli* were carried out according to published procedures [14]. The purification was monitored by SDS-PAGE with 15% acrylamide. The final pure Svp protein was desalted by dialysis in 50 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 1 mM EDTA, 2 mM DTT, 10% Glycerol, and the pure protein was flash frozen in liquid nitrogen and stored at -80 °C.

1.3. Results

1.3.1. Overproduction in *E. coli* and purification of Svp protein

The *svp* gene was previously cloned, and the recombinant Svp protein was successfully purified as a His₆-tag protein according to published procedures [14]. The purity of the isolated Svp protein was confirmed by SDS-PAGE analysis (Fig.2). A sample of 160 mg Svp was sent to our collaborator for structural determination by X-ray crystallography and results are pending.

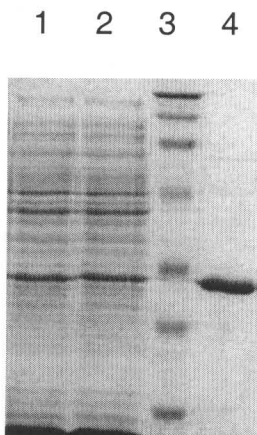


Figure 2. SDS-PAGE analysis of Svp overproduction and purification. Lane 1, total protein; 2, soluble protein; 3, MW markers; 4, purified Svp (Ni-eluent)

1.4. References

1. Hopwood, D.A. & Sherman, D.H. (1990). Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu. Rev. Genet.* **24**, 37-66
2. Kleinkauf, H. & von Dohren, H. (1996). A nonribosomal system of peptide biosynthesis. *Eur. J. biochem.* **236**, 335-351
3. Lambalot, R.H., et al., & Walsh, C.T. (1996). A new enzyme superfamily – the phosphopantetheinyl transferases. *Chem. Biol.* **3**, 923-936
4. Walsh, C.T., Gehring, A.M., Weinreb, P.H., Quadri, L.E.N. & Flugel, R.S. (1997). Posttranslational modification of polyketide and nonribosomal peptide synthases. *Curr. Opin. Chem. Biol.* **1**, 309-315
5. Coderre, P.E. & Earhart, C.F. (1989). The *entD* gene of the *Escherichia coli* K12 enterobactin gene cluster. *J. Gen. Microbiol.* **135**, 3043-3055
6. Nakano, M.M., Marahiel, M.A. & Zuber, P (1988). Identification of a genetic locus required for biosynthesis of the lipopeptide antibiotic surfactin in *Bacillus subtilis*. *J. Bacteriol.* **170**, 5662-5668
7. Quadri, L.E.N., Weinreb, P.W., Lei, M., Nakano, M.M., Zuber, P. & Walsh, C.T. (1998). Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry*, **37**, 1585-1595
8. Reuter, K., Mofid, M.R., Marahiel, M.A. & Ficner, R. (1999). Crystal structure of the surfactin synthetase-activating enzyme Sfp: a prototype of the 4'-phosphopantetheinyl transferase superfamily. *EMBO J.*, **18**, 6823-6831

9. Shen, B., Du, L., Sanchez, C., Chen, M. & Edwards, D.J. (1999). Bleomycin biosynthesis in *Streptomyces certicillus* ATCC15003: a model of hybrid peptide and polyketide biosynthesis. *Bioorg. Chem.* **27**, 155-171
10. Du, L. & Shen, B. (1999). Identification of a type II peptidyl carrier protein from the bleomycin producer *Streptomyces certicillus* ATCC15003. *Chem. Biol.* **6**, 507-517
11. Du, L., Sanchez, C., Chen, M., Edwards, D.J. & Shen, B. (2000) The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces certicillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. *Chem. Biol.* **7**, 623-642
12. Du, L., Chen, M., Sanchez, C. & Shen, B. (2000). An oxidation domain in the BImIII nonribosomal peptide synthetase probably catalyzing thiazole formation in the biosynthesis of the antitumor drug bleomycin *Streptomyces certicillus*. *FEMS Microbiol. Lett.* **289**, 171-175
13. Shen, B. Du, L., Sanchez, C., Edwards, D.J. Chen, M. & Murrell J. (2002). Cloning and characterization of the bleomycin biosynthetic gene cluster from *Streptomyces verticillus* ATCC15003. *J. Nat. Prod.* **65**, 422-431
14. Sanchez, C., Du, L., Edwards, D.J., Toney, M.D. & Shen, B. (2001) Cloning and characterization of a phosphopantetheinyl transferase from *Streptomyces verticillus* ATCC15003, the producer of the hybrid peptide-polyketide antitumor drug bleomycin. *Chem. Biol.* **8**, 725-738
15. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular cloning: a*

laboratory manual, (2nd ed), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA

Chapter 2: Cloning and characterization of neocarzinostatin (NCS) gene cluster
from *Streptomyces carzinostaticus*

2.1. Introduction

As a family, the enediyne antibiotics are the most potent, highly active antitumor agents ever discovered. Some members are 1,000 times more potent than adriamycin, one of the most effective, clinically used antitumor antibiotics [1]. The enediyne antibiotics are the focus of intense research activity in the fields of chemistry, biology, and medical sciences because of their unique molecular architectures, biological activities, and modes of action [2-3]. Thus far, there have been two basic groups within the enediyne antibiotic family: (i) the calicheamicin-esperamicin type, which includes the calicheamicins, the esperamicins, and namenicin; (ii) and the chromoprotein type, consisting of an apoprotein and an unstable enediyne chromophore. The former consists of a 10-membered enediyne ring system that is relatively stable. The latter group, which includes neocarzinostatin (NCS), kedarcidin, C-1027, and maduropeptin, possesses a novel 9-membered ring chromophore core structure and require a specific binding protein for chromophore stabilization (Fig. 1 and 2) [3].

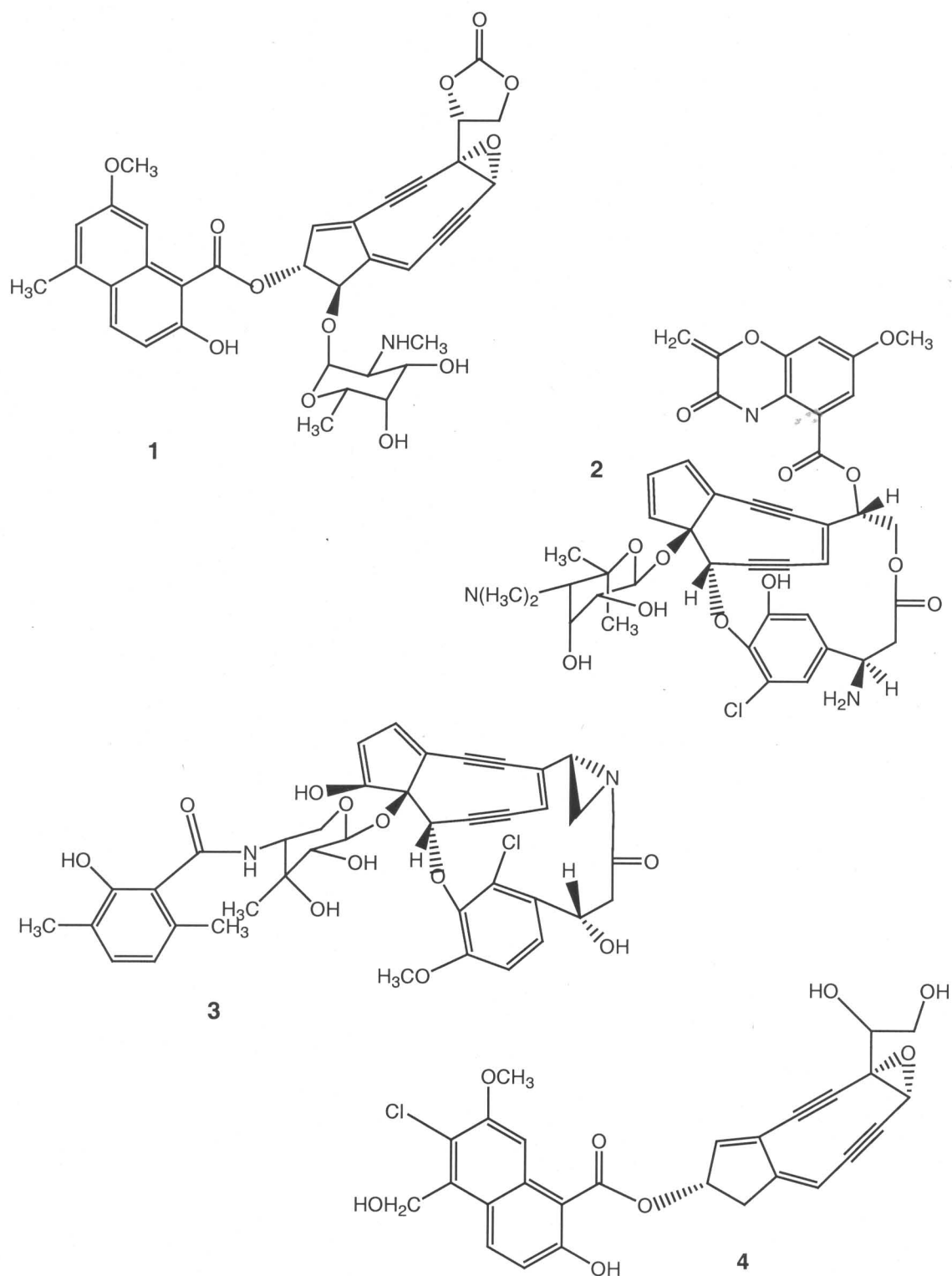


Figure 1. Bicyclo enediyne chromophore for neocarzinostatin (1), C-1027 (2), maduropeptin (3), and N19999A2 (4).

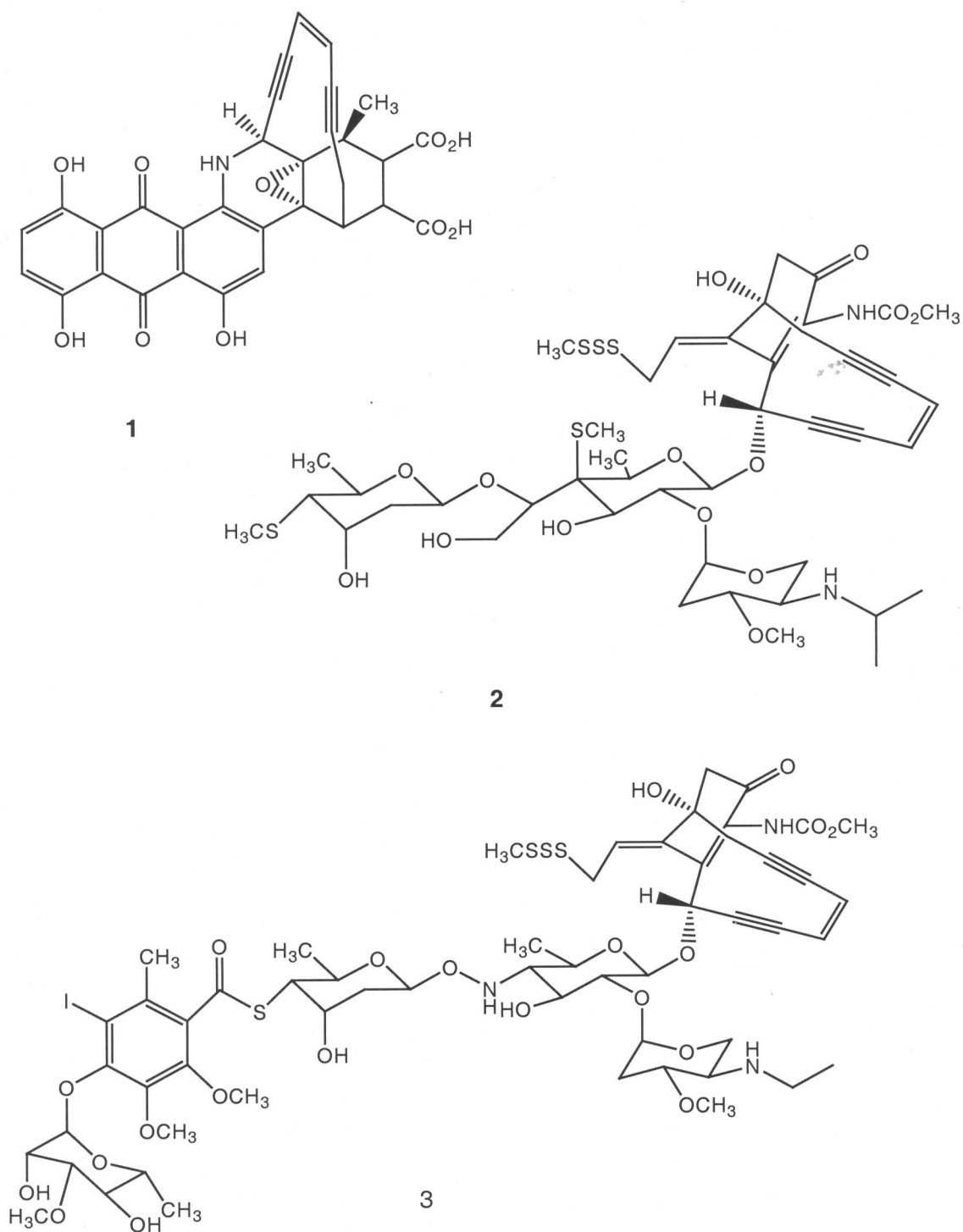


Figure 2. The structure of the 10-membered ring enediynes: dynemicin (1), namenamicin (2), and calicheamicin γ (3).

All members of enediyne family contain a core or “warhead” consisting of two acetylenic groups conjugated to a double bond or incipient double bond within a 9- or 10-membered ring, i.e., the enediyne core as exemplified by NCS in Fig. 3. As the consequence of this structural feature, these compounds share a common mechanism of action: the enediyne core undergoes an electronic rearrangement to form a transient benzenoid diradical, which is positioned in the minor groove of DNA so as to damage DNA by abstracting hydrogen atoms from deoxyriboses on both strands. This novel mechanism of DNA damage has important implications for their application as potent cancer chemotherapeutic agents [4-6].

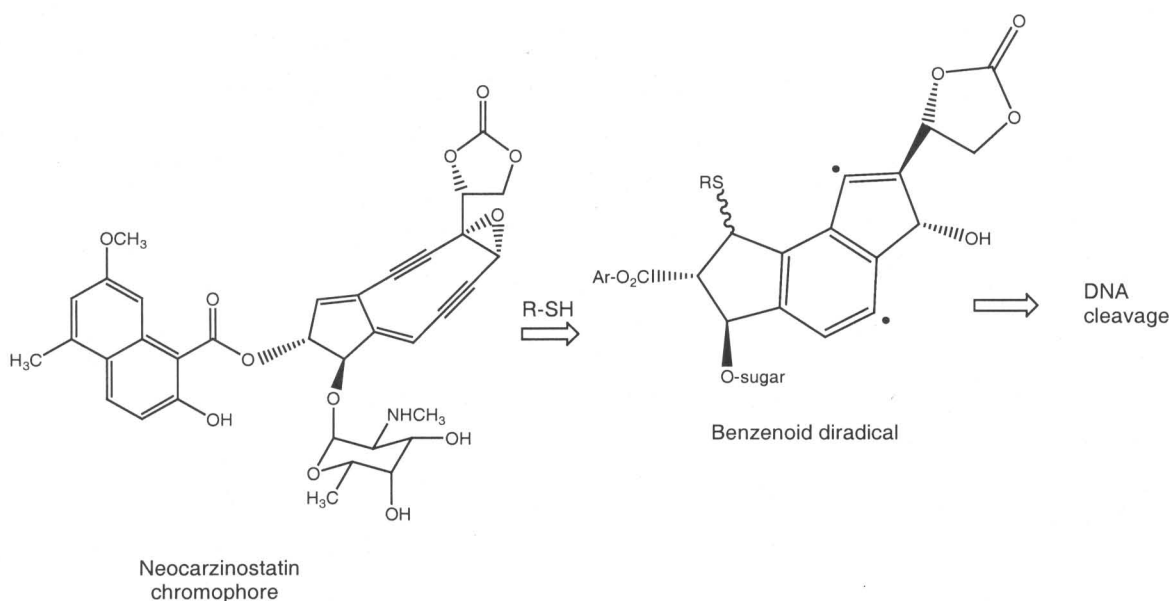


Figure 3. Structure of NCS chromophore and the benzenoid diradical intermediate proposed to initiate DNA cleavage.

Neocarzinostatin (NCS) is a natural chromoprotein antibiotic produced by *Streptomyces carzinostaticus* and the first chromoprotein antitumor antibiotic to structurally elucidated [7-8]. NCS is composed of an aminoglycoside chromophore component (NCS chromophore) bound noncovalently to a 113-amino acid carrier protein (NcsA) [9-11]. It was recognized that essentially all the biological activity of NCS resides in the chromophore and that the protein component plays an important role in transporting and stabilizing the complex [12-13].

The natural enediynes that have so far been introduced into the clinic have been limited due to unpredictable toxicity [2, 14]. NCS is the only molecule of the 9-membered enediyne family that has received approval for the clinical treatment of cancer in Japan. However, various polymer-based delivery systems or enediyne-antibody conjugates have shown great clinical promise in anticancer chemotherapy demonstrating that the enediynes can be developed into powerful drugs when their extremely potent cytotoxicity is harnessed and delivered directly onto the targeted tumor cells. Although a total synthetic approach for the 9-membered enediynes has been reported, the schemes are laborious and yields are low. Genetic manipulation of genes governing secondary metabolism offers a promising alternative for structurally making complex microbial metabolites. Therefore, biosynthetic and genetic studies of NCS will help us understand the general mechanisms for enediyne biosynthesis and is critical for developing ways to make novel, clinically useful enediyne analogs. Significant progress has been made in recent years in genetic engineering of antibiotic biosynthesis in actinomycetes.

A cosmid library of *Streptomyces carzinostaticus* ATCC15944 was previously constructed, and a genetic system was developed [15]. To clone the NCS biosynthetic gene cluster from the NCS-producing strain *S. carzinostaticus*, we took advantage of previous work regarding NCS and knowledge of the common machinery governing deoxyaminosugar biosynthesis. Two probes corresponding to the *ncsA* gene, the sequence of which is known, and a NDP-hexose-4,6-dehydratase gene (DH), were generated by PCR and used to screen the *S. carzinostaticus* cosmid library. Using the *ncsA* probe, a 7.5 kb DNA was identified and shown to contain a putative dNDP-glucose 4,6-dehydratase. In collaboration with Dr. Liping Nie, sequencing of the entire cluster was pursued.

2.2. Materials and Methods

2.2.1. DNA isolation and manipulation

Plasmid isolation and DNA gel extraction was carried out using commercial kits (Qiagen, Santa clarita, CA). Restriction of DNA and ligation were carried out using appropriate enzymes (Invitrogen Corporation, Carlsbad, CA) and standard procedures [16]. *E. coli* XL1-Blue and pGEM-3zf were used as a general cloning host and general cloning vector, respectively. Plasmids were transformed into *E. coli* by heat shock as described [16].

2.2.2. DNA Sequencing

A typical sequencing reaction consisted of 1000 ng of DNA, 25 pmol of each primer, 5% of DMSO, 4 μ L 2.5X buffer, 2 μ L of enzyme reaction mixture in a final

volume of 20 μ L. Sequencing reaction cycle was as follows: an initial melting at 95 °C for 3 min followed by 50 cycles of 10 min at 96 °C, 4 min at 58 °C, and 7 min at 72 °C. Sequencing reaction mixtures were cleaned by spin column AutoSeqG-50 (Amersham pharmacia Biotech Inc., Piscataway, NJ). The dried pellets were resuspended in loading buffer, and analyzed by agarose gel electrophoresis. Automated DNA sequencing was carried out on an ABI automated sequencer. Sequencing was provided by UW-Biotech Center Sequence Facility.

2.2.3. PCR

Primers were synthesized at the UW-Biotech Center. PCR reactions were carried out with a Gene Amp PCR system 2400 (Perkin-Elmer/ABI) and either *Taq* (Invitrogen Corporation, Carlsbad, CA) or *Vent* (New England Biolabs) polymerase with their respective commercial buffers. A typical reaction mixture consisted of 100 pmol of each primer, 0.2 mM deoxynucleoside triphosphates, 10% dimethyl sulfoxide, 2 U of selected polymerase, and buffer in a final volume of 100 μ L. A typical programmed cycling were as follows: initial denaturation at 95 °C for 5 min followed by 30 cycles of 30 sec at 95 °C, 30 sec at 62 °C, and 2.5 min at 72 °C, and an additional 7 min at 72 °C.

For *NcsA*, the following pair of primers were used: 5'-GCTTCGGG CGCTTDTCTCAG-3' (forward), and 5'-GCGGAGCGGATCCTCCGATC-3' (reverse). A single distinctive product of the predicted 0.6 kb was amplified in good yield from 37II-2.

For gene disruption, primers were designed based on the sequence of dNDP-glucose 4,6-dehydratase and 37II-2: 5'-CCGGGTGCGGGCC ATTGGCGAATAC-3' (forward) and 5'-TCCCTGATCGAC GCGTCGGCATCC-3' (reverse) were used. Because restriction sites were not designed into the primers, it was necessary to use *Taq* polymerase for TA-cloning (Promega). A distinctive product with the predicted size of 0.7 kb was amplified as judged by agarose electrophoresis.

2.3. Results and Discussion

2.3.1. Cloning of NCS biosynthetic gene cluster

Fourteen cosmids (cosmid 1, 5, 6, 7, 11, 12, 13, 15, 17, 23, 27, 28, 36, and 37) were isolated by using the dioxigenin-labeled 600 bp *ncsA* fragment as a probe. These cosmids were digested with several enzymes and the patterns of fragments were analyzed to generate a contig map (Fig. 4).

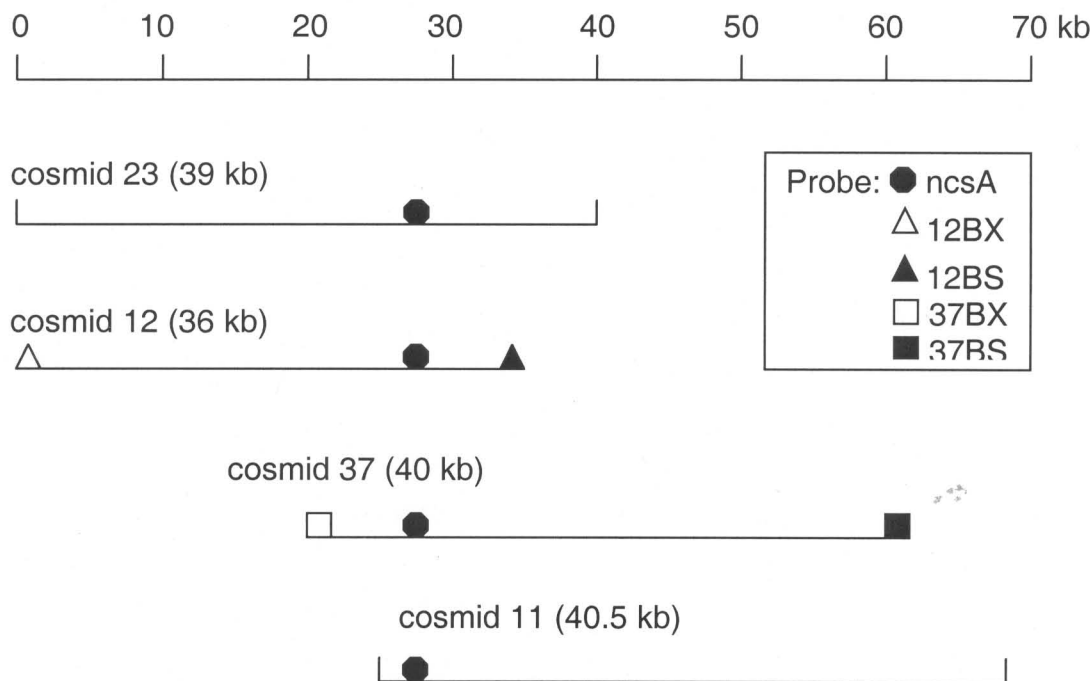


Figure 4. Relative position of the overlapping cosmids that were identified with the *ncsA* probe as exemplified by 4 cosmids.

In order to clone the gene cluster that is responsible for the biosynthesis of NCS, the DNA fragments with positive signals for *ncsA* were cloned. Cosmid 22, 24, 12, 23, 37, 11 and 28 were digested with *Bam*HI, *Bgl*II, *Nco*I, and *Pst*I. Four of the cosmids, cosmid 12, 23, 37, and 11, gave the expected 2.1 kb signals; For *Bgl*II digestion, three of the cosmids, cosmid 12, 23, and 37, gave the 7.5 kb signals (Fig. 5). These two DNA fragments were cloned into *Nco*I site of pGEM-5zf and *Bam*HI site of pGEM-3zf from cosmid37, respectively. In each group, several recombinant clones were obtained, and two of them were confirmed with Southern blotting (Fig. 6).

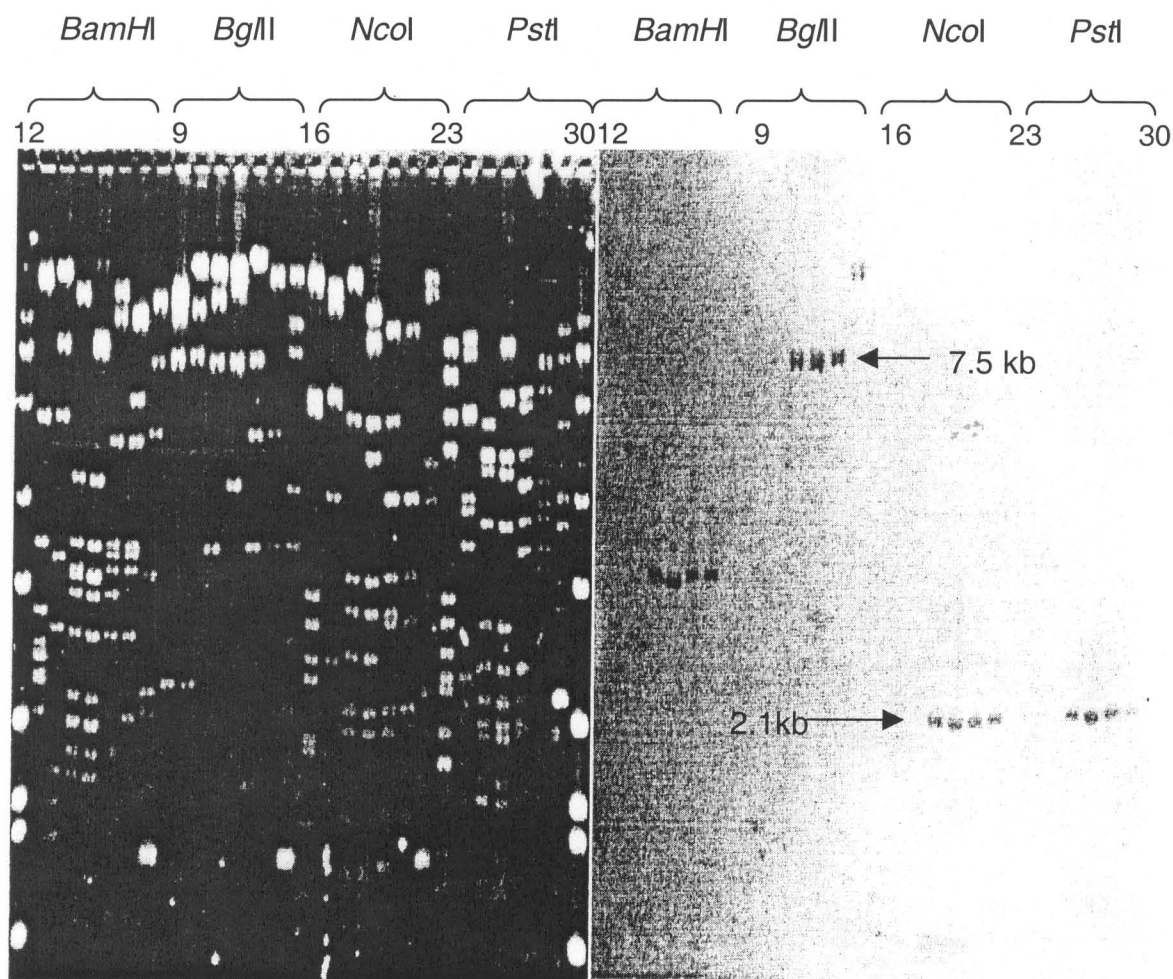


Figure. 5. Southern blot with *ncsA* probe. Lane 1, 30; DNA marker, lane 2-8, *Bam*HI digestion; lane 9-15, *Bgl*II digestion; lane 16-22, *Nco*I digestion; lane 23-29; *Pst*I digestion. In each group, the sequence of cosmids is 22, 14, 12, 23, 37, 11, and 28

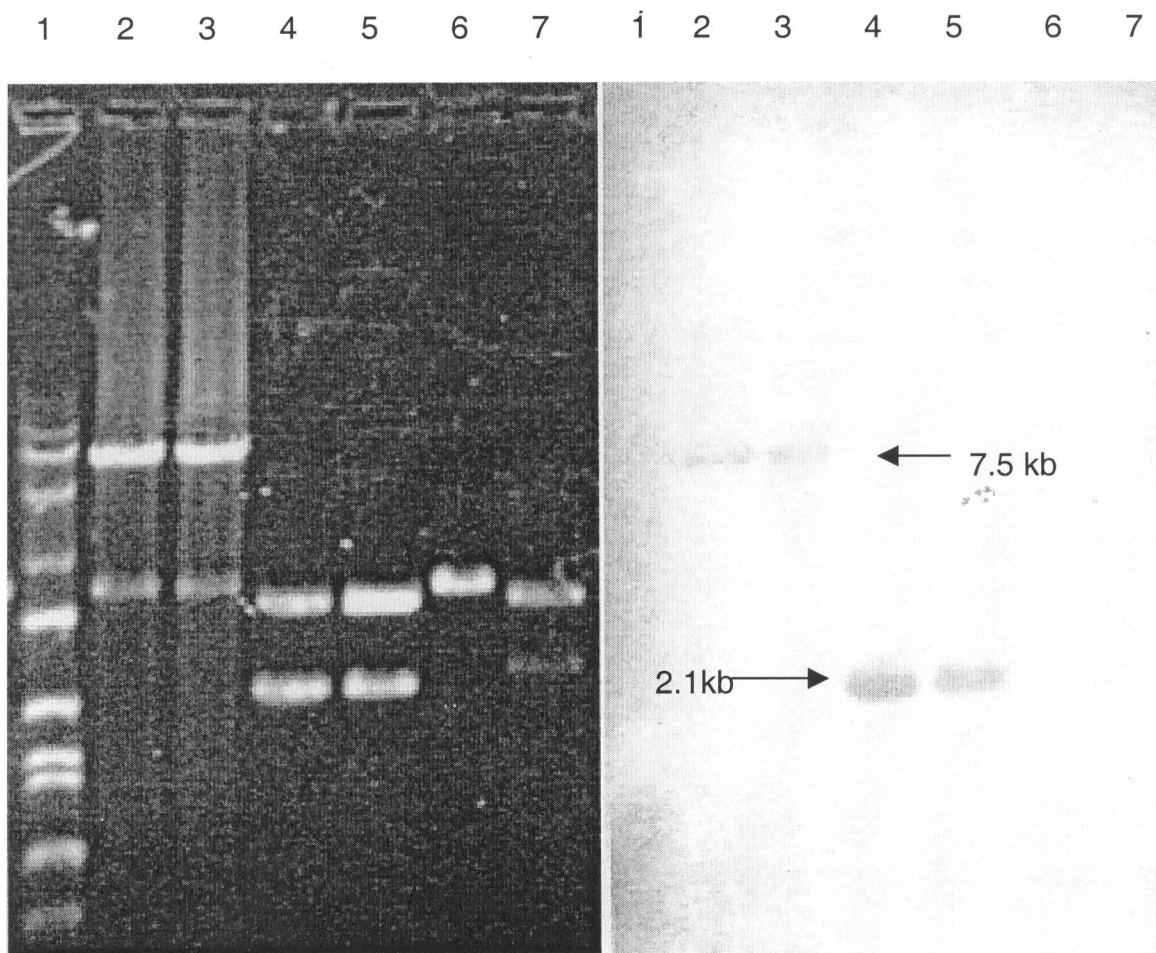


Figure 6. Southern blot of recombinant plasmids. Lane 1, DNA marker; lane 2, 37II-1/*Xba*I and *Eco*RI; lane 3, 37II-2/*Xba*I and *Eco*RI; lane 4, 37c-1/*Nco*I; lane 5, 37c-2/*Nco*I; lane 6, pGEM3zf/*Bam*HI; lane 7, pGEM5zf/*Nco*I.

2.3.2. Subcloning and sequencing of NCS cluster

A single cosmid 37II-2 was digested with *Bam*HI for subcloning and sequencing to give five fragments (Fig. 7). 2.1 kb fragment was found when it was digested with *Nco*I (data not shown), which means this 7.5 kb *Bgl*II fragment cover 2.1 kb *Nco*I fragment.

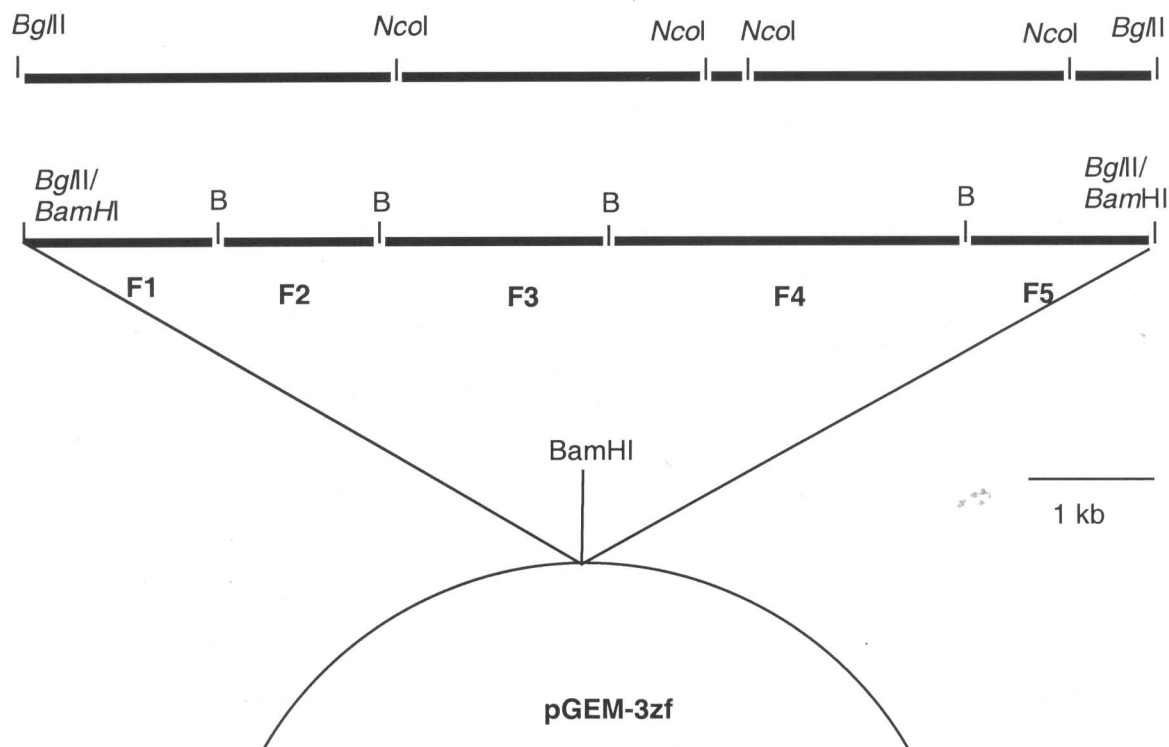


Figure 7. Partial genetic map of 7.5 kb *Bgl*II fragment and its construction into pGEM-3zf *Bam*HI site.

Three fragments (F2, F3, and F4) were inserted into pGEM-3zf. F1 + F5 fragment was self-ligated and then digested with several enzymes to find suitable enzymes for subcloning to pGEM-3zf. No restriction digest activity was observed with *Sst*I and *Xba*I, therefore F1 and F5 were subcloned into the *Bam*HI/*Sst*I sites of pGEM-3zf and *Bam*HI/*Xba*I sites of pGEM-3zf, respectively. These subclones were confirmed with proper enzyme digestions (data not shown). Both ends of the five cloned DNA fragments were sequenced and putative ORFs were

determined by using GCG. BLAST analysis of the gene products resulted in a map as shown in Figure 8.

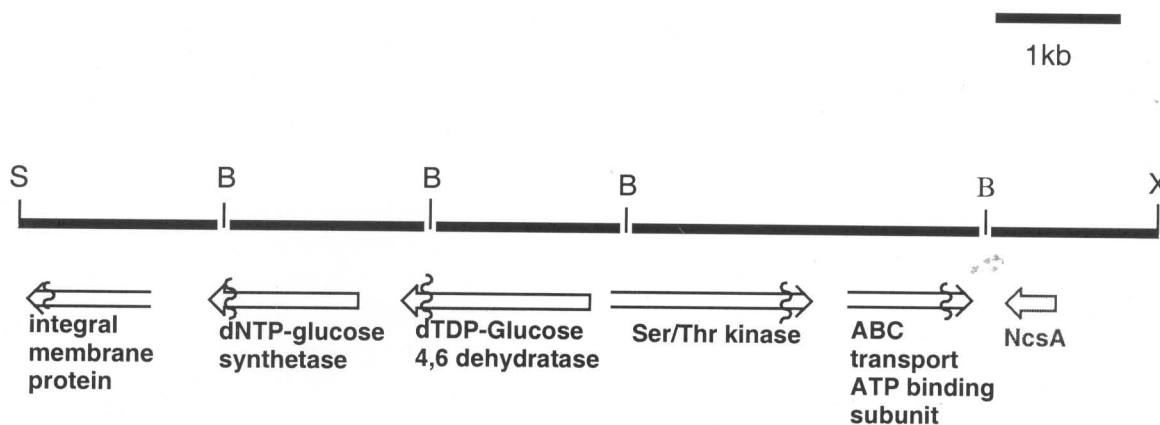


Figure 8. A putative gene map of a region of the NCS gene cluster

2.3.3. Gene disruption of dNDP-glucose 4,6-dehydratase

The sequence data showed that 0.8 kb *Bam*HI fragment from 37II-2 contained the incomplete dNDP-glucose 4,6-dehydratase (Fig. 8).

Cloning of the gene cluster for C-1027 biosynthesis in *S. globisporus* revealed a dNDP-glucose 4,6-dehydratase gene (*sgcA*) residing around 14kb downstream of the *cagA* gene, encoding the CagA apoprotein. The involvement of *sgcA* in C-1027 biosynthesis was demonstrated by disrupting *sgcA* to generate *S. globisporus* C-1027-nonproducing mutants and by complementing the *sgcA* mutants *in vivo* to restore C-1027 production [17].

Moreover, we found the putative NDP-glucose 4,6-dehydratase (DH) locus by using 0.5 kb DH probe. But, the putative DH gene-disrupted mutant produced

NCS at the level of wild-type strain, excluding the involvement of this gene in NCS biosynthesis. The cosmids which were screened with both DH and *ncsA* probe do not overlap, either [18].

Therefore, we thought the cloned DH gene from 37II-2 is likely involved in NCS biosynthesis. To clone the DH gene from 37II-2, PCR was performed with primers designed according to the sequence data. The distinctive product with the predicted size of 0.7 kb was amplified (data not shown). This product was cloned to pGEM-easy vector for restriction analysis and sequencing. Although a genetic system for *S. carzinostaticus* has been developed, attempts to inactivate this DH gene to confirm its involvement in NCS biosynthesis all failed likely due to our recent finding of 3 separate NCS biosynthetic gene clusters are present with the producer.

For cloning the whole NCS gene cluster, chromosomal walking from the two positive cosmids was undertaken by using two ends covering 64 kb DNA fragment (Fig. 4). After extension to 140 kb, the DNA was sent to Ecopia Bioscience for sequencing.

2.4. References

1. Zhen, Y.S., Ming, X.Y., Yu, B., Otani, T., Saito, H., & Yamada, Y. (1989). A new macromolecular antitumor antibiotic, C-1027.III. Antitumor activity. *J. Antibiot.* **42**, 1294-1298
2. Doyle, T.W. & Borders, D. B. (ed.). 1995. Eneidyne antibiotics as antitumor agents. Larcel Dekker, Inc., New York, N.Y.
3. Thorson, J.S., Shen, B., Whitwam, R.E., Liu, W., Li, Y. & Ahlert, J. (1999). Eneidyne biosynthesis and self-resistance: a process report. *Bioorg. Chem.* **27**, 172-188
4. Myers, A.G., Kort, M.E., & Hammond, M. (1997). A comparison of DNA cleavage by neocarzinostatin chromophore and its aglycon: evaluating the role of the carbohydrate residue. *J. Am. Chem. Soc.* **119**, 2965-2972
5. Stassinopoulos, A., Ji, J., Gao, X., & Goldberg, I.H. (1996). Solution structure of a two-base DNA bulge complexed with an enediyne cleaving analog. *Science*. **272**, 1943-1946
6. Ikemoton, N., Kumar, R.A., Ling, T.T., Ellestad, G.A., & Danishefky, S.J. (1995). Calicheamicin-DNA complexes – war head alignment and saccharide recognition of the minor groove. *Proc. Natl. Acad. Sci. USA*. **92**, 10506-10510
7. Ishida, N., Miyazaki, K., & Rikimaru, M. (1965). Neocarzinostatin, an antitumor antibiotic of high molecular weight. *J. Antibiot.* **18**, 68-76
8. Kim, K.-H., Kwon, B.-M., Myers, A.G., & Rees, D.C. (1993). Crystal Structure of Neocarzinostatin, an Antitumor Protein-Chromophore Complex. *Science*.

262, 1042-1046

9. Napier, M.A., Holmquist, B., Strydom, D.J., & Goldberg, I.H. (1979). Neocarzinostatin: spectral characterization and separation of a non-protein chromophore. *Biochem Biophys Res Commun.* **89**, 635-642
10. Koide, N., Ishii, F., Hasuda, K., Koyama, Y., Edo, K., Kitamine, S., & Ishida, N. (1980). Isolation of a non-protein component and a protein component from neocarzinostatin (NCS) and their biological activities. *J. Antibiot.* **33**, 342-346
11. Heyd, B., Leart, G., Adjadj, E., Minard, P., & Desmadril, M. (2000) Reinvestigation of the Proteolytic Activity of Neocarzinostatin. *J. Bacteriol.* **182**, 1812-1818
12. Napier, M.A., Kappen, L.S., & Goldberg, I.H. (1980) Effect of nonprotein chromophore removal on neocarzinostatin action. *Biochemistry.* **19**, 1767-1773
13. Kappen, L.S., Napier, M.A., & Goldberg, I.H. (1980) Roles of chromophore and apo-protein in neocarzinostatin action. *Proc. Natl. Acad. Sci. USA.* **77**, 1970-1974
14. Maeda, H., Edo, K., Ishida, N. (ed.). 1997. Neocarzinostatin: the past, present, and future of an anticancer drug. Springer-verlag, Inc., New York, N.Y.
15. Dr. Liping's research report (2001)
16. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, (2nd ed), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA
17. Liu, W. & Shen, B. (2000) Genes for production of the Eneidyne Antitumor

Antibiotic C-1027 in *Streptomyces globisporus* Are Clustered with the *cagA* Gene That Encodes the C-1027 Apoprotein. *Antimicrobiol. Agen. Chemoth.* **44**, 382-392

18. Liu, W., Nonaka, K., Nie, L., Zhang, J., Christenson, S.D., Bae, J., Lanen, S.G., Zazopoulos, E., Farnet, C.M., Yang, C.F., & Shen, B. (2004) Cloning and sequencing of the neocarzinostatin biosynthetic gene cluster from *Streptomyces carzinostaticus* AT 15944: a convergent model for biosynthesis involving two iterative type I polyketide synthases. *Proc. Natl. Acad. Sci. USA*. (in press)

Chapter 3: Functional analysis of neocarzinostatin biosynthetic gene cluster

3.1. Introduction

3.1.1. Studies on NCS deoxyaminosugar biosynthetic pathway

Several variations of unusual sugars are found in natural products that have hydroxyl groups replaced by hydrogen, an amino group, or an alkyl side chain. These sugars have been shown to play crucial roles in biological activity of the natural products in bacterial lipopolysaccharides (LPSs) or secondary metabolite antibiotics [1]. Whereas the 3,6-dideoxysugars are limited to the LPS of gram-negative bacteria, 2,6-dideoxysugars are abundant in secondary metabolites, such as in the macrolide and aminoglycoside antibiotics. With the exception of N-acetyl-D-glucosamine, aminosugars are only found in secondary metabolites [1-2].

Common deoxysugars found in LPS and antibiotics are shown in Figure 1 [2]. L-rhamnose 1, L-fucose 2, 6-deoxy-D-gulose 3, and D-perosamine 4 are well-known examples of 6-deoxysugars. L-ascarylose 5, D-abequose 6, D-tyvelose 7, and D-colitose 8 are frequently found 3,6-dideoxyhexoses that are frequently found in secondary metabolites. D-olivose 9 is an example of a prevalent 2,6-dideoxysugars, while L-daunosamine 10, N-acetyl-L-glucosamine 11, and N-acetyl-D-glucosamine 12 are representative amino sugars found in natural products.

A few deoxysugar biosynthetic pathways have been well studied, including rhamnose and some other common LPS deoxysugars (Fig. 2) [3], but to date several pathways have not been investigated.

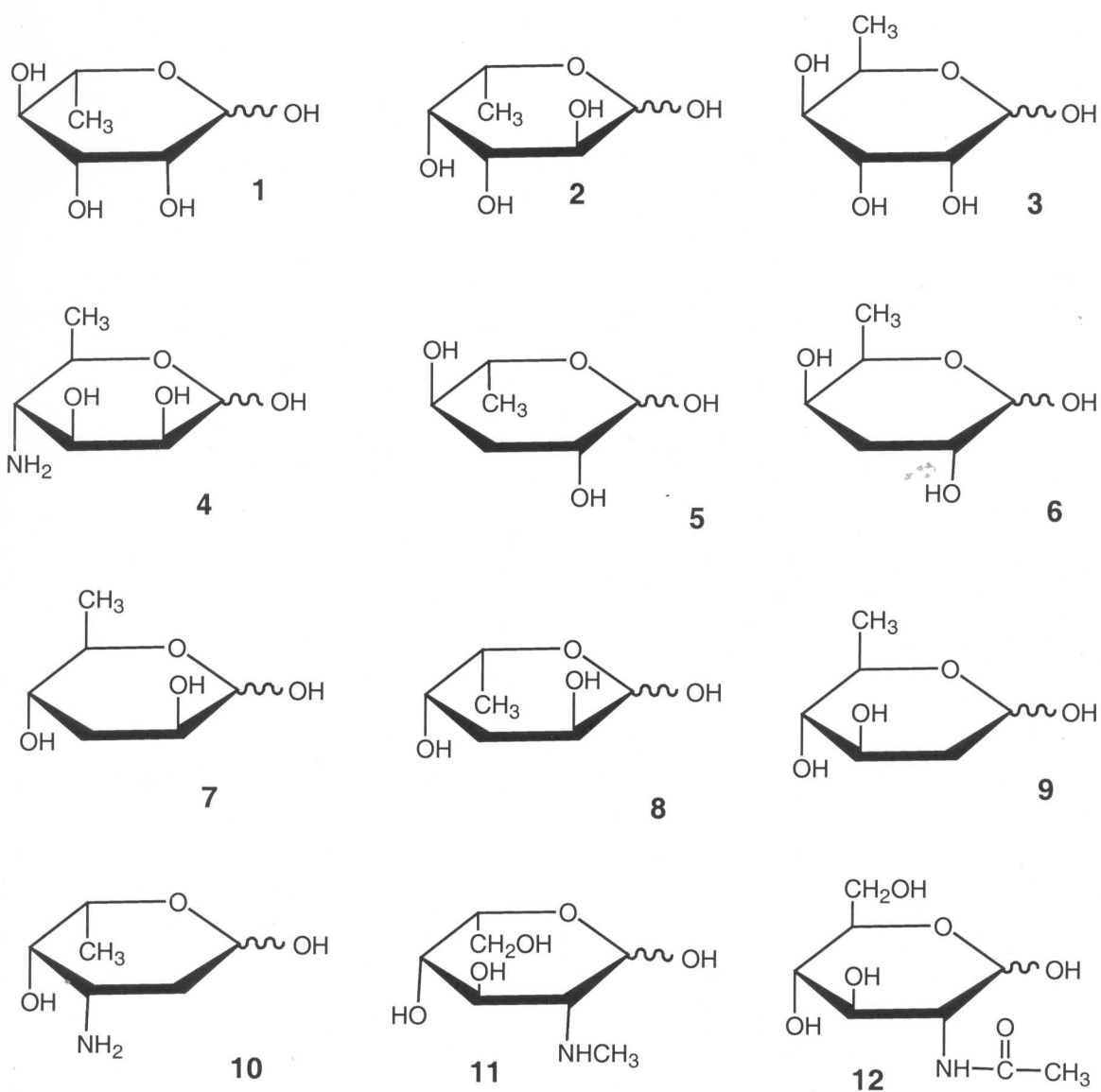


Figure 1. Common deoxysugars found in LPS and antibiotics

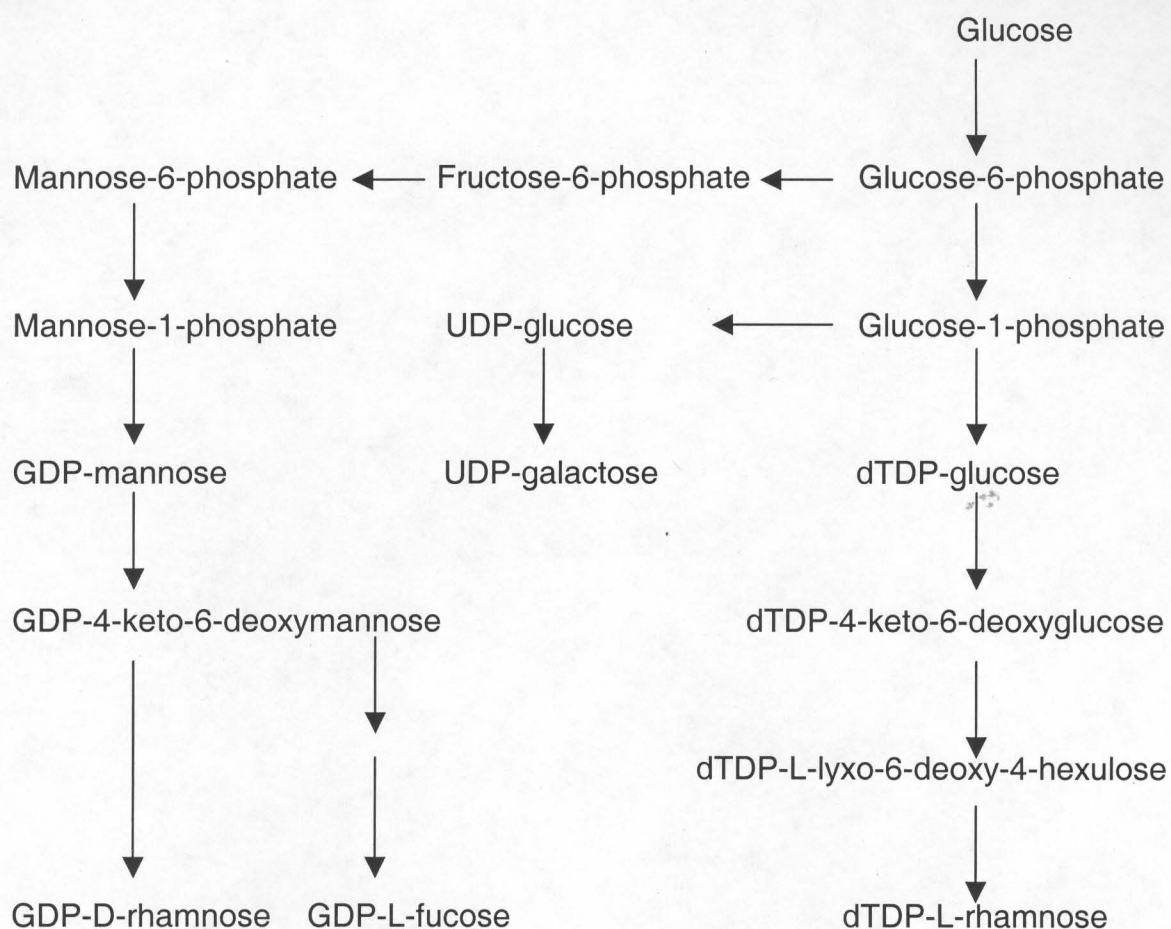


Figure 2. The rhamnose pathway, and some of its interactions with other deoxysugar biosynthetic pathway.

NDP-glucose has been shown to be the precursor for most deoxyaminosugars, and typically TDP-D-glucose is the precursor for these unusual sugars produced by *Streptomyces* spp. [1]. In contrast, GDP-mannose has been shown to be the precursor for fucose, rhamnose, and perosamine, components of capsular polysaccharides and LPS [4-5].

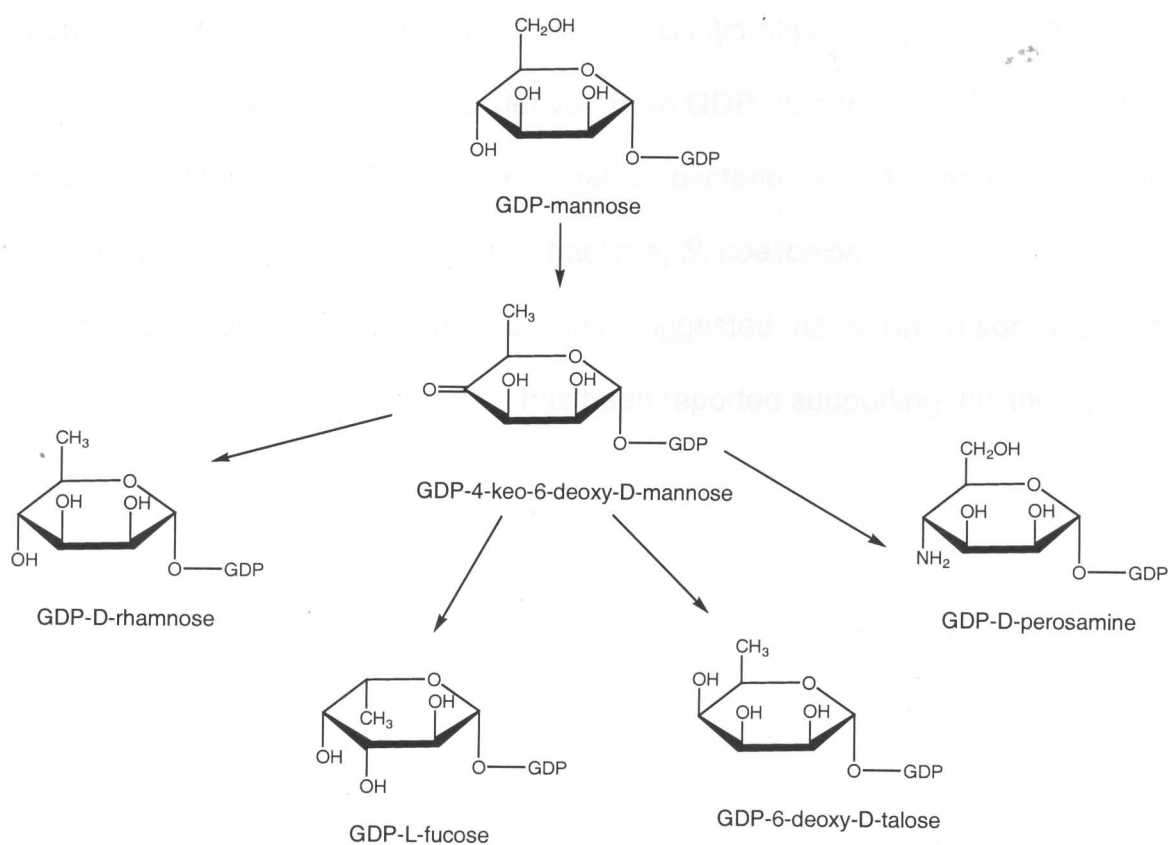


Figure 3. The biosynthetic pathway of several different deoxyhexoses originating from the same 4-keto-6-deoxy intermediate.

GDP-mannose is known to be a typical precursor in primary metabolism, exemplified by alginate, which is a major component of exopolysaccharide (EPS)

[6]. However, there are few examples of GDP-mannose as a component of secondary metabolites produced by bacteria. Mycosamine is a deoxyaminosugar moiety of the polyene macrolide antibiotics amphotericin B, candicidin, nystatin, and pimaricin, produced by *Streptomyces* spp., and it has been suggested that D-mycosamine (3-amino-3,6-dideoxy-D-mannose) is synthesized via GDP-mannose, since each of the four polyene macrolide gene clusters were found to contain a homolog of GDP-mannose 4,6-dehydratase (Fig. 4) [7]. A second example of a deoxyaminosugar derived from GDP-mannose is D-perosamine, a main constituent of LPS in gram-negative bacteria and a component of the antibiotic perimycin in gram-positive bacteria, *S. coelicolor*.

In both cases, GDP-mannose was suggested as a precursor based on sequence alignments, but no data has been reported supporting this theory.

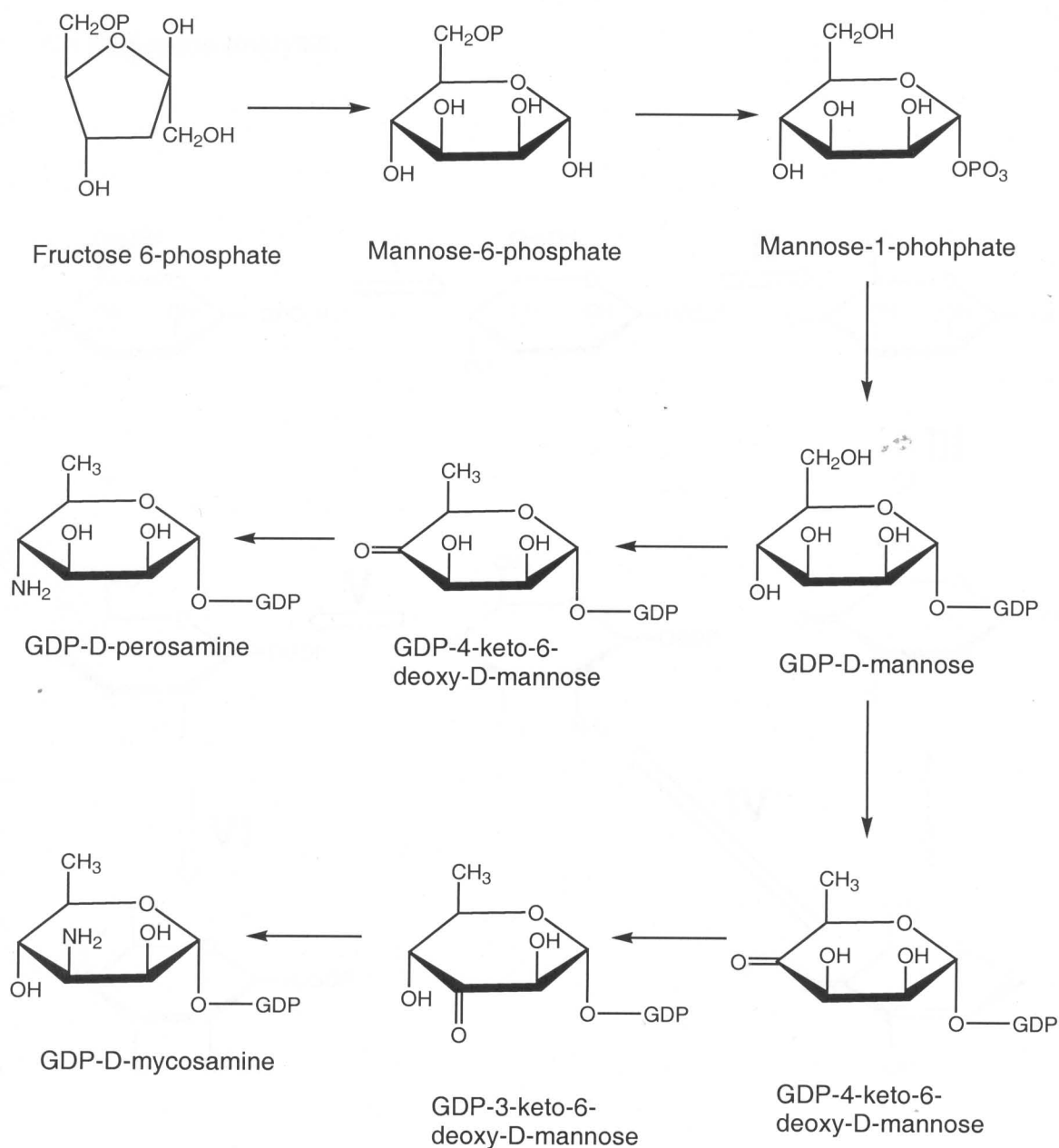


Figure 4. Proposed pathways for D-mycosamine and D-perosamine biosynthesis.

DNA sequencing of the NCS biosynthetic gene cluster revealed six genes (*ncsC*, *C1*, *C2*, *C3*, *C4*, and *C5*) whose deduced gene products support their role in the biosynthesis of the deoxyaminosugar moiety of NCS. Figure 5 shows the proposed biosynthetic pathway for the deoxyaminosugar moiety on the basis of

the sequence analysis.

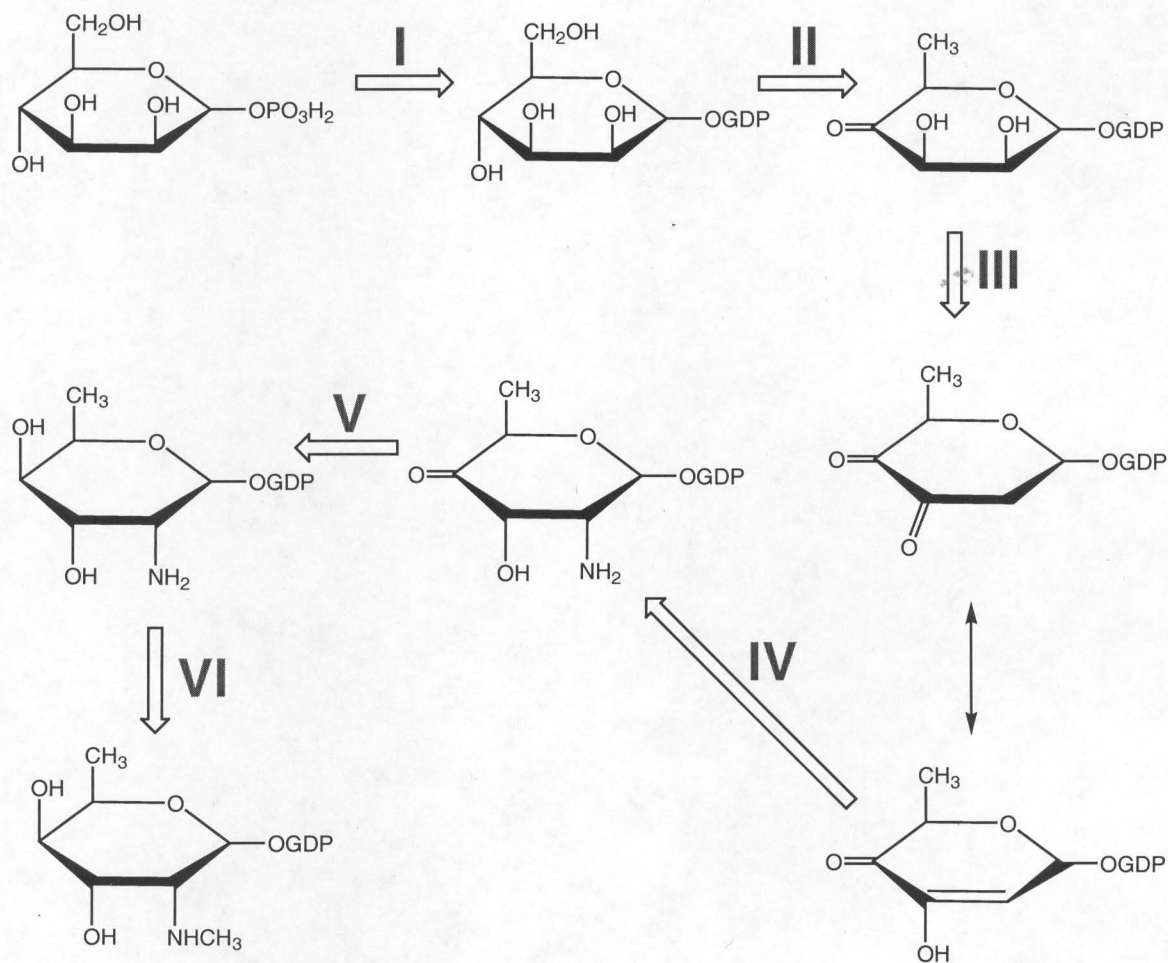


Figure 5. Proposed biosynthesis of deoxyaminosugar moiety of NCS. I. GDP-mannose synthase (*ncsC*), II. GDP-mannose-4,6-dehydratase (*ncsC1*), III. GDP-mannose-2-dehydratase (*ncsC2*), IV. Phenylalanine ammonia lyase (*ncsC3*), V. NDP-hexose 4-ketoreduatase (*ncsC4*), VI. N-methyltransferase (*ncsC5*)

NcsC showed 57% of identity with GDP-mannose synthase, suggesting that similar to mycosamine and perosamine, the deoxyaminosugar is derived from GDP-mannose. However, NcsC1 showed no significant homology with GDP-mannose 4,6-dehydratase, but instead has 60% identity with NDP-sugar dehydratase. Therefore, the origin of the NCS deoxyaminosugar remained unclear based on sequence analysis.

To identify the function of putative deoxyaminosugar biosynthetic genes, both *in vivo* gene replacement assay and *in vitro* assays were pursued. GDP-mannose-4,6-dehydratase (*ncsC1*), GDP-mannose-2-dehydratase (*ncsC2*), and phenylalanine ammonia lyase (*ncsC3*) were targeted for gene replacement and GDP-mannose-4,6-dehydratase gene was also tried with gene disruption. For *in vitro* assay, all six genes proposed to be involved in aminosugar biosynthesis were cloned, and the gene products expressed in *E. coli*: GDP-mannose synthase (*ncsC*), GDP-mannose-4,6-dehydratase (*ncsC1*), GDP-mannose-2-dehydratase (*ncsC2*), phenylalanine ammonia lyase (*ncsC3*), GDP-hexose 4-ketoreductase (*ncsC4*) and N-methyltransferase (*ncsC5*) (Fig. 5).

Gene disruption of *ncsC1* and gene replacement of *ncsC1*, *ncsC2* and *ncsC3* yielded inconclusive results. Among the six recombinant proteins produced in *E. coli*, *ncsC1*, *ncsC2* and *ncsC3* were partially soluble and isolated as a His₆-Tag constructs. NcsC1 was assayed with possible substrates: TDP-glucose and GDP-mannose, but no activity was observed based on HPLC analysis. In an attempt to obtain soluble and active NcsC, the gene was in *S. lividans* TK64 strain but the crude cell free extract showed activity with GTP and mannose-1-

phosphate not with TTP and glucose-1-phosphate based on HPLC analysis. To confirm this activity the putative product was purified by HPLC, but spectroscopic analysis was inconclusive. Furthermore, extensive control experiments showed that this activity resulted from the host background.

3.2. Materials and Methods

3.2.1. Bacterial strains, DNA isolation and manipulation

Plasmid isolation and DNA gel extraction was carried out using a commercial kit (Qiagen, Santa Clarita, CA). Restriction digests of DNA and ligation were carried out using appropriate enzymes (Invitrogen Corporation, Carlsbad, CA) and standard procedures [8]. *E. coli* XL1-Blue and pGEM-3zf, 5zf, and 7zf were used as a general cloning host and general cloning vector, respectively. Plasmids were transformed into *E. coli* by heat shock as described [8].

3.2.2. DNA Sequencing

A typical sequencing reaction consisted of 1000 ng of DNA, 25 pmol of each primer, 5% of DMSO, 4 μ L 2.5X buffer, 2 μ L of enzyme reaction mixture in a final volume of 20 μ L. Sequencing reaction cycle was as follows: an initial melting at 95 °C for 3 min followed by 50 cycles of 10 min at 96 °C, 4 min at 58 °C, and 7 min at 72 °C. Sequencing reaction mixtures were cleaned by spin column AutoSeqG-50 (Amersham pharmacia Biotech Inc., Piscataway, NJ). The dried pellets were resuspended in loading buffer, and analyzed by agarose gel electrophoresis. Automated DNA sequencing was carried out on an ABI

automated sequencer. Sequencing was provided by UW-Biotech Center Sequence Facility

3.2.3. PCR for *in vivo* assay

Primers were synthesized at the UW-Biotech Center. PCR reactions were carried out with a GeneAmp PCR system 2400 (Perkin-Elmer/ABI) and either *Taq* (Invitrogen Corporation, Carlsbad, CA) or *Vent* (New England Biolabs) polymerase with their respective commercial buffers. A typical reaction mixture consisted of 100 pmol of forward and reverse primer, 0.2 mM deoxynucleoside triphosphates, 10% dimethyl sulfoxide, 2 U of selected polymerase, and buffer in a final volume of 100 μ L. The typical programmed cycling temperatures were as follows: initial denaturation at 95 °C for 5 min followed by 30 cycles of 30 sec at 95 °C, 30 sec at 62 °C, and 2.5 min at 72 °C, and a final heat treatment 72 °C for 7 min.

For gene disruption, primers were designed based on the sequence of dNDP-glucose 4,6-dehydratase from 37II-2: 5'-CCGGGTGCGGGCCATTGGCGAATAC-3' (forward) and 5'-TCCCTGATCGACGCGTCGGCATCC-3' (reverse). Restriction sites were not designed into the primers, since the PCR fragment was cloned using TA-cloning *Taq* polymerase (Promega). A distinctive product with the predicted size of 0.7 kb was amplified (data not shown).

For gene replacement of *ncsC1*, the following pair of primers was used for left arm (Deh-1): 5' – CCAGGAGATCTGGTCACCGACCTCGGCAGAG - 3' (forward, *Bgl*II site is underlined), and 5'-GCTCTAGACCTTCGTCTCGGT CGGCACCG

CTCCGG-3' (reverse, *Xba*I site was inserted). A single distinctive product of the predicted 2.0 kb size was amplified in good yield from cosmid37 and 1.9 kb of right arm was produced from *Nco*I and *Pst*I digestion of 37II-2 (data not shown).

For *ncsC2*, the following pairs of primers were used for left arm (GD-1): 5' – GGAATCGAAATTCGAACAGGCTTCGGGCGG - 3' (forward, *Eco*RI site is underlined), and 5' – CCTGGTACCGGTCGCTCATG TCTGCCCCG - 3' (reverse, *Kpn*I site is underlined). For right arm (GD-2), 5' – GCTCTAGAGGC TCCGGTGATGACCACGCG - 3' (forward, *Xba*I site was inserted) and 5'-CCCAAGCTTGGCGAGTCTCCCGCCTTGAAGACCG-3' (reverse, *Hind*III site is underlined) were used. A distinctive product of the expected 1.8 kb and 1.7 kb were observed by using cosmid 15 (data not shown).

For *ncsC3*, the following pairs of primers were used for left arm (PAL-1): 5' – CACCAGCATGCCGCCGGAGCGGCG - 3' (forward, *Sph*I site is underlined), and 5' – CGCCACATCGATGCC TGGCGGGCTCTCACGG - 3' (reverse, *Cla*I site is underlined). For right arm (PAL-2), 5'-GCTCTAGAGCAGAACAA TGATCGGATCATTGCCATTGG-3' (forward, *Xba*I site was inserted) and 5'-GAGGGATCCAGCGTGCCG CGCCGTTACCG-3' (reverse, *Bam*HI site is underlined) were used. A distinctive product of the expected 2.0 kb and 1.7 kb were observed by using cosmid 12 as template (data not shown).

3.2.4. Conjugation and transformation

A 0.7 kb PCR fragment from 37II-2 was cloned into pOJ260 and pKC1139 to yield pJY-II-35 and pJY-II-36, respectively. Introduction of pJY-II-35 and pJY-II-36 into *Streptomyces* was done using conjugation. For conjugation, *E. coli* S17-1

(pJY-II-35 or pJY-II-36) was grown to an OD₆₀₀ of 0.3 to 0.4. Cells from a 20 mL culture were pelleted by centrifugation, washed in LB, and resuspended in 2 mL of LB as the *E. coli* donor. *S. carzinostaticus* (10^3 to 10^9) were washed, resuspended in TSB, and incubated at 50 °C for 10 minutes to activate germination. After additional incubation at 37 °C for 2-5 h, the spores were pelleted and resuspended in LB as the *S. carzinostaticus* recipient. The donor cells (100 µL) and recipient cells (100 µL) were mixed and spread equally onto two modified ISP-4 (IWL-4) plates, supplemented with freshly prepared 10 mM MgCl₂. The plates were incubated at 28 °C for 16-22 h. After removal of most of the *E. coli* S17-1 donors by washing the surface with sterile water, the plates were overlaid with 3 mL of soft LB (0.7% agar) containing nalidixic acid (final concentration, 50 µg/mL) and apramycin (final concentration, 100 µg/mL) and were incubated at 28 °C until exconjugants appeared.

Alternatively, transformation to the *Streptomyces* strain was performed using standard methods [9].

3.2.5. PCR for protein production in *E. coli*

Primers were synthesized at the UW-Biotech Center. PCR reactions were carried out using a GeneAmp PCR system 2400 (Perkin-Elmer/ABI) and either *Taq* (Invitrogen Corporation, Carlsbad, CA) or *Vent* (New England Biolabs) polymerase with their respective commercial buffers. A typical reaction mixture consisted of 100 pmol of each primer, 0.2 mM deoxynucleoside triphosphates, 10% dimethyl sulfoxide, 2 U of selected polymerase, and buffer in a final volume of 100 µL. A typical programming cycle is as follows: initial denaturation at 95 °C

for 5 min and 30 cycles of 30 sec at 95 °C, 30 sec at 62 °C, and 2.5 min at 72 °C, followed by an additional 7 min at 72 °C.

For protein production, the following pairs of primers were used:

For *ncsC* (705 bp), cosmid 12 was used as template with the primers 5' – CCGGAATTCTGTGATCCTGGCGGGCGGCAAGGGG - 3' (forward, *EcoRI* site is underlined), and 5'-CCGAAGCTTCATGCCCCCTTGAGCAGGGTCCC-3' (reverse, *HindIII* site is underlined).

For *ncsC1* (996 bp), cosmid 37 was used as template with the primers 5' – CCGGAATTCATGGCCGGA CTCTAGCGGTC - 3' (forward, *EcoRI* site is underlined), and 5'-CCGAAGCTTCAGCGGTTGTAGTGGGTCTG-3' (reverse, *HindIII* site is underlined).

For *ncsC2* (980 bp), cosmid 37BS15 was used as template with the primers 5' – CCGGAATTCATGAGCCCTCCACTGCGCGTGGTC - 3' (forward, *EcoRI* site is underlined), and 5'-CCGCTCGAGTCA GGGCGTCACCGTCCGGG C-3' (reverse, *XhoI* site is underlined).

For *ncsC3* (1.58 kb), cosmid 12BX10 was used as template with the primers 5' – CCGGAATTCGTGACCACAGCTGAGACCGCC - 3' (forward, *EcoRI* site is underlined), and 5'-CCGAAGCTTCAGCGCAGAACGATCCC CTCCTG-3' (reverse, *HindIII* site is underlined).

For *ncsC4* (797 bp), cosmid 12 was used as template with the primers 5' – GGAATTCATATGGAATTCGTGCCCGTCGACCTG - 3' (forward, *NdeI* and *EcoRI* sites are underlined), and 5'-CCGAAGCTTCACGCCGGACGGTCGCCT TCCAG-3' (reverse, *HindIII* site is underlined).

For *ncsC5* (719 bp), cosmid 37 was used as template with the following primers 5' – GGAATTCCCATATGGAATTCATGTCCGCCTGGCTC - 3' (forward, *NdeI* and *EcoRI* sites are underlined), and 5'-CCGAAGCTTCATCGGCTC GGTCGCACGGACGACGC-3' (reverse, *HindIII* site is underlined).

3.2.6. Overproduction and purification of proteins

The PCR product was digested and ligated to the expression vector pET28a (Table 3.1.), and the presence of the correct gene confirmed by DNA sequencing.

Table 3.1. An overview of constructs used in N-terminal His₆-tag protein production

Plasmid	Description	Preparation
pJY-III-3	<i>ncsC</i> expression construct with pET28a	PCR product was digested with <i>EcoRI/HindIII</i>
pJY-III-4	<i>ncsC1</i> expression construct with pET28a	PCR product was digested with <i>EcoRI/HindIII</i>
pJY-III-5	<i>ncsC2</i> expression construct with pET28a	PCR product was digested with <i>EcoRI/HindIII</i>
pJY-III-6	<i>ncsC3</i> expression construct with pET28a	PCR product was digested with <i>EcoRI/HindIII</i>
pJY-III-21	<i>ncsC4</i> expression construct in pET28a	PCR product was digested with <i>NdeI/HindIII</i>
pJY-III-22	<i>ncsC5</i> expression construct with pET28a	PCR product was digested with <i>NdeI/HindIII</i>

After sequence conformation, plasmids were transformed to BL21 (DE3) for protein production. Purification of His₆-tagged proteins by affinity chromatography on Ni-NTA resin was carried out under the standard conditions recommended by Qiagen. Protein production was examined with valuable incubation temperature

(16 °C and 30 °C), inducer concentration (1 mM and 0.1 mM), and incubation time (24 h and 48 h). Among six proteins, NcsC1 and NcsC3 showed good solubility and purity. To improve solubility of other four enzymes, PCR was performed again to introduce the gene into a different vector for expression as a C-terminal His₆-tag.

For *ncsC*, the primer 5'-CCGAAAGCTTTG CCCCCTTGAGCAGGGTCCC-3' (reverse, *Hind*III site is underlined) was used.

For *ncsC2*, the primer 5'-CCGCTCGAGTCAG GGCGTCACCGTCCGGGC-3' (reverse, *Xho*I site is underlined) was used.

For *ncsC4*, the primer 5'-CCGAAAGCTTCGC CGGACGGTCGCCTTCCAG-3' (reverse, *Hind*III site is underlined) was used.

For *ncsC5*, the primer 5'-CCGAAAGCTTTCGGCTCGGTCGCACGGACGACG C-3' (reverse, *Hind*III site is underlined) was used.

The PCR product was digested and ligated to pET29a and pET37b for C-terminal His₆-tag protein production. Cloning was confirmed by DNA sequencing.

Table 3.2. An overview of constructs used in C-terminal His₆-tag protein production

Plasmid	Description	Preparation
pJY-III-49	<i>nscC</i> expression construct with pET29a	PCR product was digested with <i>EcoRI/HindIII</i>
pJY-III-50	<i>nscC2</i> expression construct with pET29a	PCR product was digested with <i>EcoRI/XhoI</i>
pJY-III-51	<i>nscC4</i> expression construct with pET29a	PCR product was digested with <i>NdeI/HindIII</i>
pJY-III-52	<i>nscC5</i> expression construct with pET29a	PCR product was digested with <i>NdeI/HindIII</i>

Plasmids were transformed into BL21 (DE3) for production of C-terminal His₆-tag protein. In the end, there was no big difference of their solubility. The purification of each enzyme was monitored by SDS-PAGE on 15% acrylamide gel. In preparation for activity tests of NcsC1, NcsC2, and NcsC3, the collected Ni-NTA column eluents were desalted by dialysis in 50 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 1 mM EDTA, 2 mM DTT, 10% Glycerol but it caused precipitation. After dialysis at 6 °C overnight, a precipitate was formed, therefore future experiments were done without desalting.

Expression of NcsC in *Streptomyces* was attempted following the Raynal method [10]. NcsC was cloned to the vector pPWW50 that has N-terminal His₆-tag and the plasmid transformed into the *Streptomyces* protoplasts. NcsC was expressed both in *S. lividans* and NCS producer strain (*S. carzinostaticus*).

3.2.7. *In vitro* assay

NcsC1, NcsC2, and NcsC3 were used for enzyme assays either alone or in various combinations. For NcsC1 activity tests, a reaction mixture (300 μ L) containing 50 mM Tris-HCl, pH 8.0, 2.5 mM substrate (GDP-mannose or TDP-glucose), and 30 μ L of Ni-NTA column eluent (enzyme) was incubated at 30 °C for 2 h, and then the reaction was terminated by heat denaturation of the protein by boiling for 3 min followed by centrifugation. For NcsC2 activity assays, a typical reaction mixture of 300 μ L contained 10 μ L NcsC1, 50 μ L of NcsC2, 50 mM Tris-HCl, pH 8.0, and 1 mM substrate (GDP-mannose or glucose). After 4 h incubation at 30 °C, it was extracted twice with 200 μ L of ethyl acetate and the organic solvent was removed by speedvac. For NcsC3, a reaction mixture (500 μ L) containing 50 mM Tris-HCl, pH 8.0, 2.5 mM substrate (TDP-glucose or GDP-mannose), 50 mM L-Phenylalanine, 10 μ L of NcsC1, 50 μ L of NcsC2, and 80 μ L of NcsC3 was incubated at 30 °C for 6 h. The reaction was terminated by heat treatment and subjected to HPLC analysis after removal of protein.

It has been reported that TDP-glucose 4,6-dehydratase typically uses cofactor NAD and GDP-mannose 4,6-dehydratase prefers NADP as a cofactor [11-24]. Therefore a variety of conditions were tested for NcsC1. The reaction mixture (50 μ L) contained 50 mM KH_2PO_4 , pH 7.5, 2.5 mM substrate (TDP-glucose or GDP-mannose), 4 mM MgCl_2 , 1 mM cofactor (NAD or NADP) and 10 μ L of enzyme (Ni-NTA column eluent). Both the crude extract and purified enzyme were tested for activity in combination with null enzyme controls. NcsC2 and NcsC3 were also tested with positive controls, TDP-glucose 4,6-dehydratase and TDP-

glucose 2,3-dehydratase that were kindly given from Dr. Changsheng Zhang in Thorson lab. But it didn't show any activity.

The reaction mixture (50 μ L) for NcsC from *S. lividans* contained 50 mM KH_2PO_4 , pH 7.5, 2.5 mM substrate (TTP or GTP), 5 mM substrate (glucose-1-phosphate or mannose-1-phosphate), 5.5 mM MgCl_2 , and 10 μ L of crude extract. The reaction was terminated by heat treatment after 1 h at 37 °C and subjected to HPLC analysis.

3.2.8. HPLC methods

Reaction products were analyzed by HPLC on Alltech Altima C18 column (5 μ m, 250 x 4.6 mm, Deerfield, IL). The chromatogram was developed with a linear gradient of 0 to 50% over 30 min and 50 to 100% over 2 min at 0.7 ml/min, and the A_{254} of the effluent was measured. As a mobile phase, a phosphate buffer (50 mM KH_2PO_4 , pH 6; 5 mM tetrabutylammonium hydrogensulfate, 2% acetonitrile) and acetonitrile was used [25].

3.3. Results and Discussion

3.3.1. Gene disruption of dNDP-glucose 4,6-dehydratase

0.7 kb PCR product from 37II-2 was cloned to pGEM-easy vector (pJY-II-32) for restriction analysis and sequencing (Fig. 7).

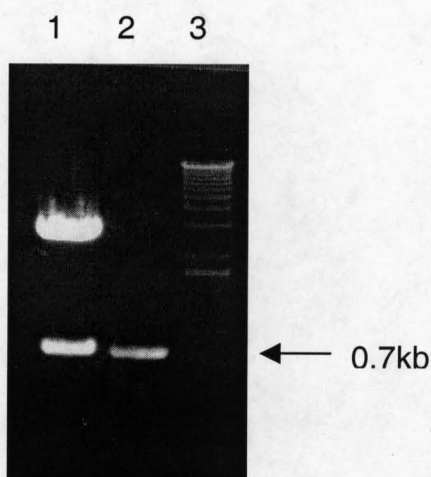


Figure 7. 0.7kb ligated to pGEM-Easy vector (pJY-II-32). Lane 1, pJY-II-32; lane 2, 0.7kb PCR product; lane 3, DNA marker.

The sequence data of pJY-II-32 showed identity with original genomic sequence (data not shown). Then this 0.7 kb fragment was cloned both to suicide vector pOJ260 and to temperature-sensitive vector pKC1139. 0.7 kb PCR fragment was cloned to *EcoRI* site of pOJ260 (pJY-II-35) and *EcoRI* and *XbaI* site of pKC1139 (pJY-II-36).

These constructs were used for conjugation but it was not successful.

3.3.2. Construction of recombinant plasmids pJY-II-38, 39, and 40 for gene replacement

For GDP-mannose-4,6-dehydratase (*ncsC*) gene replacement, pJY1-II-38 was constructed. 1.9 kb of *NcoI*-*PstI* fragment (Deh-2) from 37II-2 (Fig. 8), 2.0 kb *BglII*-*XbaI* fragment (Deh-1) from PCR, and apramycin cassette for selection marker from pUO9090 were ligated to pSET151 to yield pJY-II-38 (Fig. 9). Deh-1 was sequenced for confirmation (data not shown).

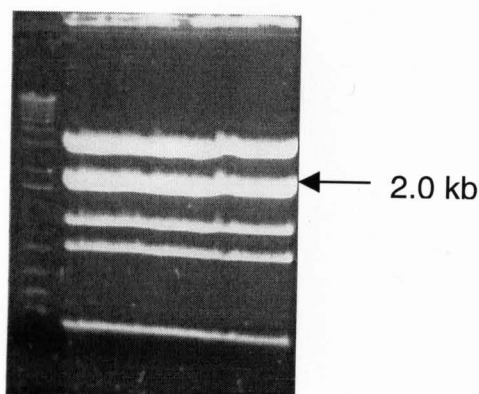


Figure 8. 2.0kb fragment from 37II-2 digestion with *NcoI*-*PstI*

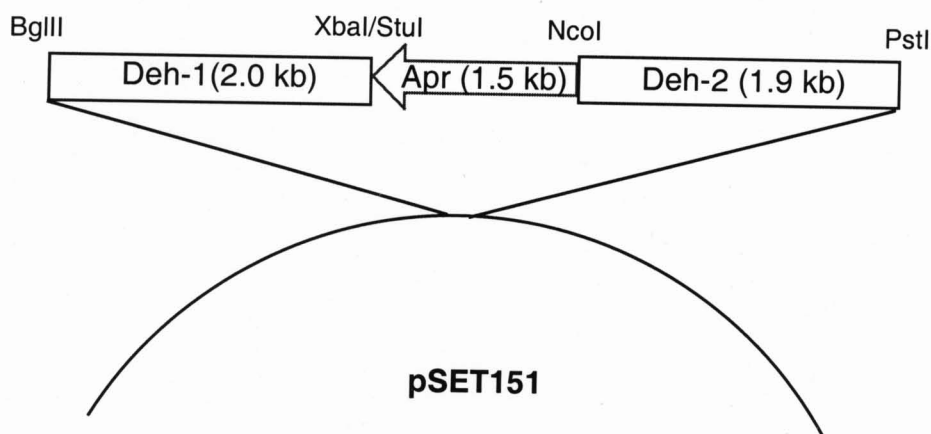


Figure 9. Final construct of pJY-II-38 for gene replacement

For phenylalanine ammonia lyase (*ncsC3*), pJY-II-39 was constructed. 2.0 kb of *SphI*-*Clal* fragment (PAL-1) and 1.7 kb of *XbaI*-*BamHI* fragment were isolated from PCR product (data not shown). Apramycin cassette was used and these three fragments were ligated to pSET151 to yield pJY-II-39 (Fig. 10). PAL-1 & 2 were sequenced for confirmation (data not shown).

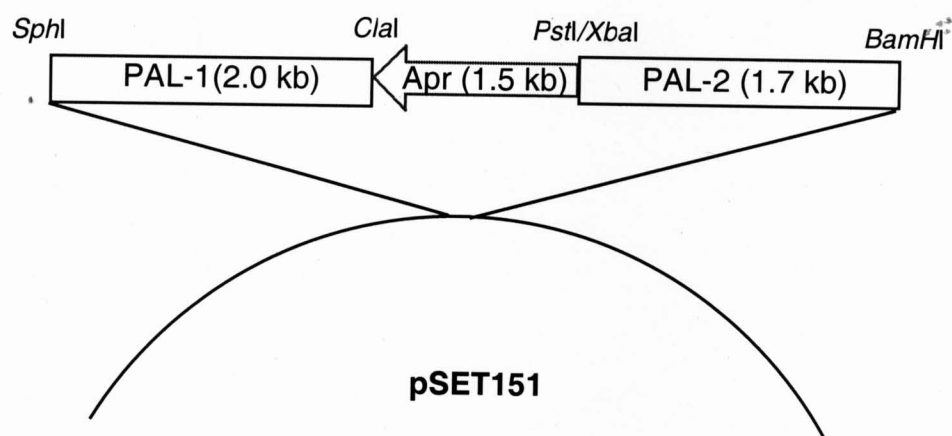


Figure 10. Final construct of pJY-II-39 for gene replacement

For GDP-mannose-2-dehydratase (*ncsC2*), pJY-II-40 was constructed. 1.8 kb of *EcoRI*-*KpnI* fragment (GD-1) and 1.7 kb of *XbaI*-*HindIII* fragment were isolated from PCR product (data not shown). Apramycin cassette was also used and these three fragments were ligated to pSET151 to yield pJY-II-40 (Fig. 11). GD-1 was confirmed with several enzyme digestions (data not shown) and GD-2 was sequenced for confirmation (data not shown).

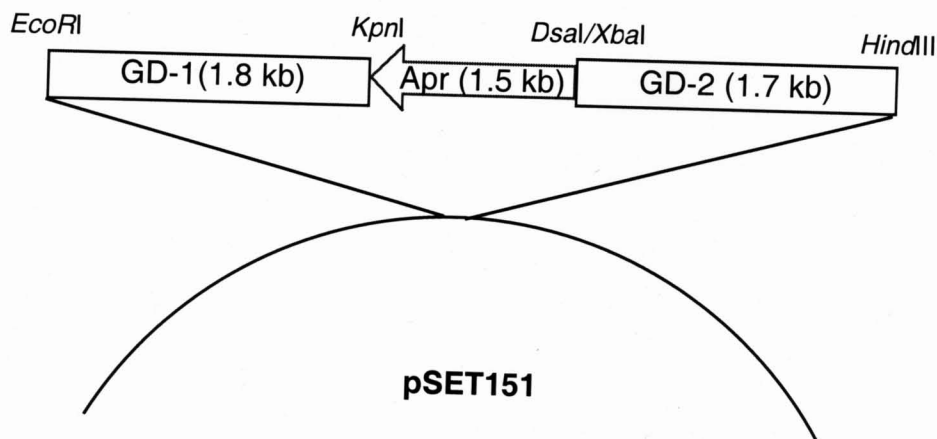


Figure 11. Final construct of pJY-II-40 for gene replacement

These three plasmids (pJY-II-38, 39, and 40) were constructed for gene replacement via conjugation and transformation.

Gene disruption of *ncsC1* and gene replacement of *ncsC1*, *ncsC2* and *ncsC3* yielded inconclusive results. There were several factors that can affect to the transformation frequencies. NCS producer was not efficient strain for transformation compared to *S. lividans*. It also has three copies of genome so it is supposed to be hard to get gene inactivation mutant either by gene disruption or gene replacement.

3.3.3. Overproduction and purification of deoxyaminosugar biosynthetic gene cluster in *E. coli*.

Every protein overproduction was better at 16 °C and 0.1 mM IPTG induction than 30 °C and 1 mM induction (data not shown). Among 6 proteins, NcsC1, NcsC2 and NcsC3 showed good solubility and purity (Fig. 12-17). To improve

solubility of other three enzymes, C-terminal His₆-tagged proteins were tried but there was no big difference of their solubility (data not shown).

The purification of each enzyme was monitored by SDS-PAGE on 15% gel.

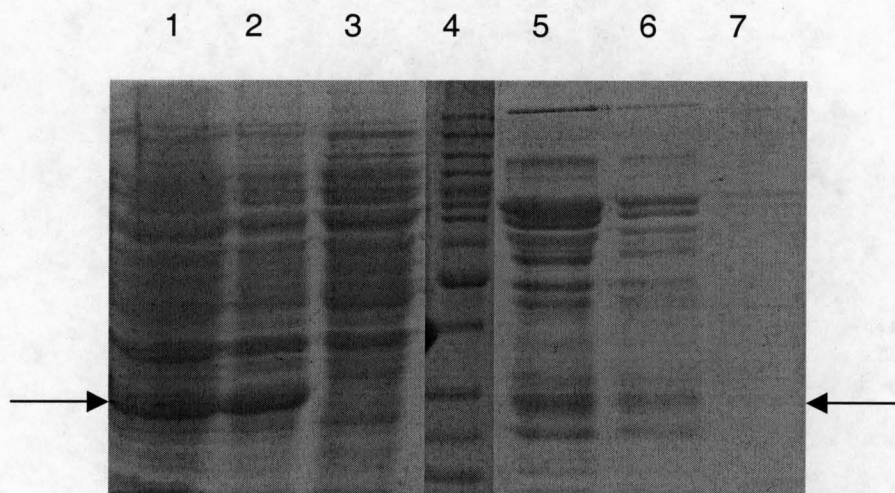


Figure 12. SDS-PAGE analysis of NcsC (26 kDa) overproduction and purification in *E. coli*. Lane 1, total protein; 2, soluble protein; 3, insoluble protein; 4, MW markers; 5-7, Ni-NTA eluents

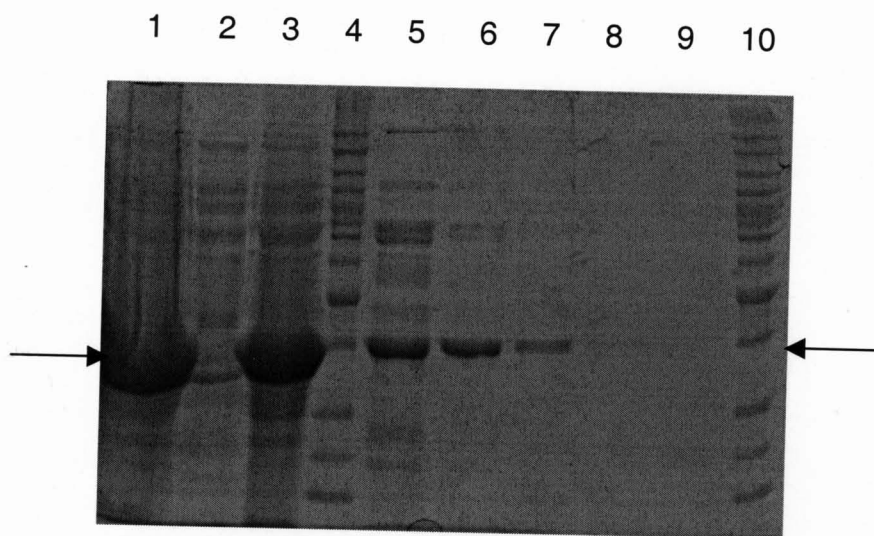


Figure 13. SDS-PAGE analysis of NcsC1 (37 kDa) overproduction and purification in *E. coli*. Lane 1, total protein; 2, soluble protein; 3, insoluble protein; 4, MW markers; 5-8, Ni-NTA eluents

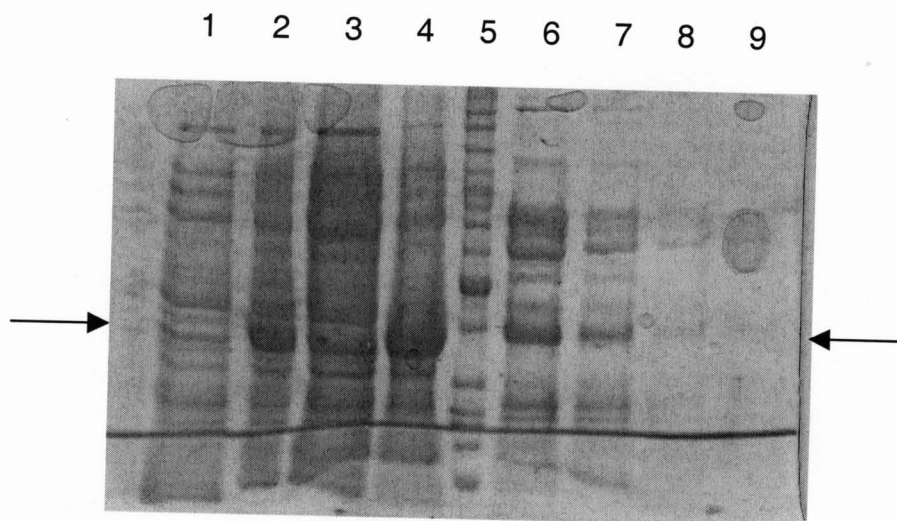


Figure 14. SDS-PAGE analysis of NcsC2 (36 kDa) overproduction and purification in *E. coli*. Lane 1, uninduced total protein; 2, induced total protein; 3, soluble protein; 4, insoluble protein; 5, MW markers; 6-9, Ni-NTA eluents.

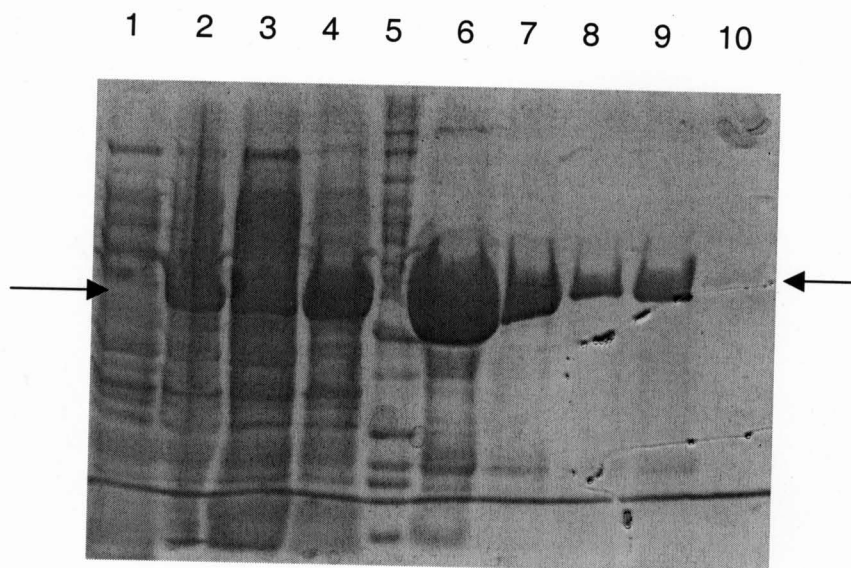


Figure 15. SDS-PAGE analysis of NcsC3 (58 kDa) overproduction and purification in *E. coli*. Lane 1, uninduced total protein; 2, induced total protein; 3, soluble protein; 4. insoluble protein; 5, MW markers; 6-10, Ni-NTA eluents

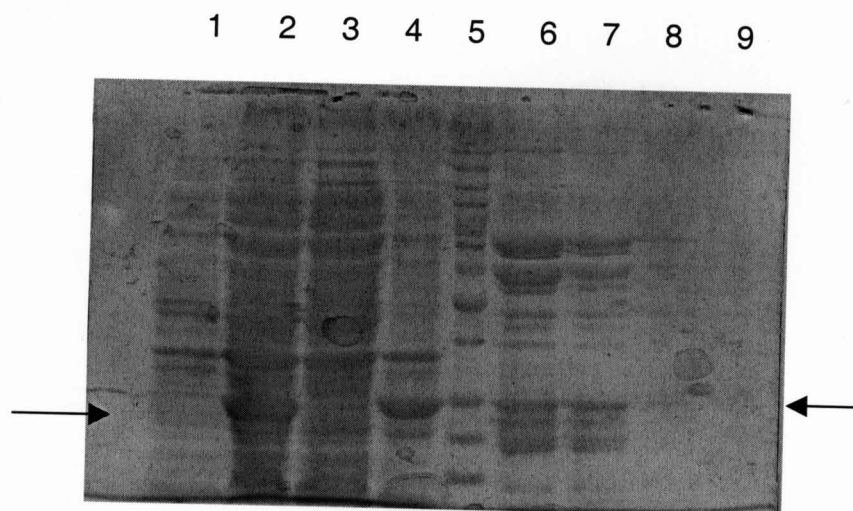


Figure 16. SDS-PAGE analysis of NcsC4 (30 kDa) overproduction and purification in *E. coli*. Lane 1, uninduced total protein; 2, induced total protein; 3, soluble protein; 4. insoluble protein; 5, MW markers; 6-9, Ni-NTA eluents

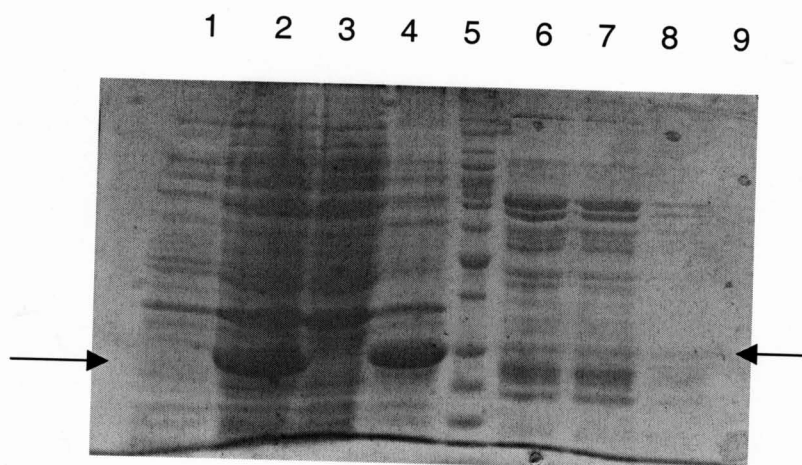


Figure 17. SDS-PAGE analysis of NcsC5 (27 kDa) overproduction and purification in *E. coli*. Lane 1, uninduced total protein; 2, induced total protein; 3, soluble protein; 4. insoluble protein; 5, MW markers; 6-9, Ni-NTA eluents

3.3.4. Enzyme assay

For functional analysis of each enzyme especially NcsC1, NcsC2, and NcsC3, the collected Ni-NTA column eluents were desalted by dialysis but it caused precipitation. So enzyme assay was done without desalting.

NcsC1 was tested with TDP-glucose and GDP-mannose to identify its substrate. But it didn't show any activity with TDP-glucose or GDP-mannose according to HPLC analysis (data not shown). Based on these results, NcsC1 seemed to be not active even though it was partially soluble. NcsC2 was tested with TDP-glucose 4,6-dehydratase and NcsC3 was tested with TDP-glucose 2-dehydratase. But there was no activity with TDP-glucose (data not shown).

Three proteins were at least partially soluble but they didn't show any activity. There could be several reasons. One possibility is I didn't prepare proper substrate for their reaction. For example, NcsC1 could use other NDP-sugar substrates and NcsC3 could work with other deoxysugar intermediate. Second possibility is our hypothesis is not right especially in terms of reaction order. That is, NCS deoxyaminosugar could be from totally new precursor. The other possible explanation is that enzyme itself is not active due to His-tag caused improper folding.

3.3.5. Expression of *ncsC* in *Streptomyces*

To get the active enzyme, NcsC was expressed in *Streptomyces*. The expression was better in *S. lividans* TK64 than NCS producer (*S. carzinostaticus*). However, expression was quite different from each transformants (Fig. 18). The crude extract from TK64 transformant was used for enzyme assay.

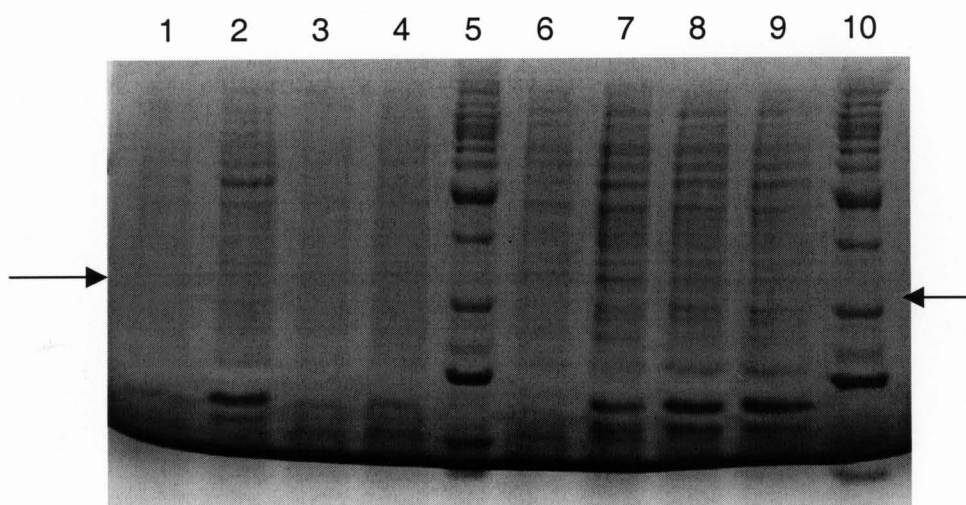


Figure 18. SDS-PAGE analysis of NcsC (26 kDa) overproduction in *Streptomyces*. Lane 1 and 6, plasmid control; Lane 2-4, total protein from each NCS producer transformant; Lane 5 and 10, MW markers; Lane 7-9, total protein from each TK64 transformant

3.3.5. Enzyme assay with crude extract from *Streptomyces*

The crude extract of TK64 transformant showed activity with GTP and mannose-1-phosphate not with TTP and glucose-1-phosphate. It showed increased product peak with time intervals (Fig. 19).

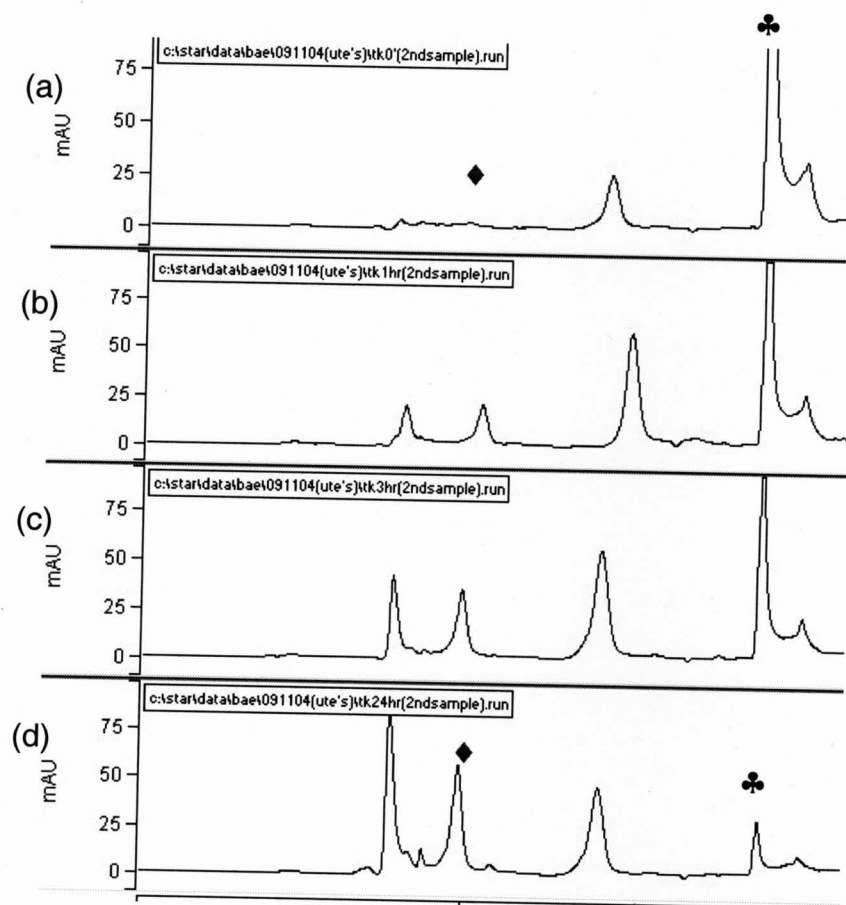


Figure 19. HPLC analysis of the *in vitro* NcsC reaction. Aliquots were taken at 15 seconds (a), 1 hr (b), 3 hr (c), and 18 hr (d). ♣ is GTP and ♦ is GDP-mannose.

Then there was a try to isolate NcsC from the crude extract with Ni-NTA resin. But it was not successful. Moreover, there were also GDP-mannose peak from the plasmid control (data not shown). Careful examination of these experiments concluded that the observed NcsC activity resulted from the *S. lividans* host not from the *ncsC* gene product.

Experiments so far, therefore, failed to establish NcsC as either TDP-glucose or GDP-mannose synthase. To identify the NCS deoxyaminosugar biosynthetic pathway, production of NcsC1, NcsC2, and NcsC3 in *Streptomyces* will give us chance to work with active enzymes even though NcsC didn't. Moreover, there is a need to test each enzyme's activity with several possible substrates. For *in vivo* assay, it is still worth to try targeted gene replacement even though it's not so easy to get gene-disrupted mutant and hard to interpret the data if their all three copies of gene cluster were replaced or not.

3.4. References

1. He, X.M. & Liu, H.W. Formation of unusual sugars; mechanistic studies and biosynthetic applications. *Annu. Rev. Biochem.* **2002**, 71:701-54
2. Trefzer, A., Salas, J.A., & Bechthold, A. Genes and enzymes involved in deoxysugar biosynthesis in bacteria. *Nat. Prod. Rep.* **1996**, 16:283-299
3. Giraud, M.-F. & Naismith, J.H. The rhamnose pathway. *Curr. Opin. Struct. Biol.* **2000**, 10:687-696
4. Menon, S., Stahl, M., Kumar, R., Xu, G.-Y., & Sullivan, F. Stereochemical course and steady state mechanism of the reaction catalyzed by the GDP-fucose synthetase from *E. coli*. *J. Biol. Chem.* **1999**, 274:26743-26750
5. Webb, N.A., Mulichak, A.M., Lam, J.S., Rocchetta, H.L., & Garavito, R.M. Crystal structure of a tetrameric GDP-D-mannose 4,6-dehydratase from a bacterial GDP-D-rhamnose biosynthetic pathway. *Protein Science* **2004**, 13:529-539
6. Snook, C.F., Tipton, S.A., & Beamer, L.J. Crystal Structure of GDP-mannose dehydrogenase: a key enzyme of alginate biosynthesis in *P. aeruginosa*. *Biochemistry* **2003**, 42:4658-4668
7. Nedal, A. & Zotchev, S.B. Biosynthesis of deoxyaminosugars in antibiotic-producing bacteria. *Appl. Microbiol. Biotechnol.* **2004**, 64:7-15
8. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, (2nd ed), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA
9. Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., & Hopwood, D.A. (2000)

Practical Streptomyces Genetics, (2nd ed), The John Innes Foundation, Norwich, England

10. Doumith, M., Weingarten, P., Wehmeier, U.F., Salah-Bey, K., Benhamou, B., Capdevila, C., Michael, J.-M., Piepersberg, W., & Raynal, M.-C. (2000). Analysis of genes involved in 6-deoxyhexose biosynthesis and transfer in *Saccharopolyspora erythraea*. *Mol. Gen. Genet.* **264**, 477-485
11. Bonn, C.P., Potter, I., Vanzin, G.F., & Reiter, W.-D. The MUR1 gene of *Arabidopsis thaliana* encodes an isoform of GDP-D-mannose-4,6-dehydratase, catalyzing the first step in the *de novo* synthesis of GDP-L-fucose. *Proc. Natl. Acad. Sci.* **1997**, 94:2085-2090
12. Tonetti, M.T., Zanardi, D., Gurnon, J.R., Fruscione, F.F., Armirotti, A., Damonte, G., Sturla, L., Flora, A.D., & Etten, J.L.V. *Paramecium bursaria* *Chlorella* virus 1 encodes two enzymes involved in the biosynthesis of GDP-L-fucose and GDP-D-rhamnose. *J. Biol. Chem.* **2003**, 278:21559-21565
13. Sturla, L., Bisso, A., Zanardi, D., Benatti, U., Flora, A.D., & Tonetti, M. Expression, purification and characterization of GDP-D-mannose 4,6-dehydratase from *E. coli*. *FEBS Lett.* **1997**, 412:126-130
14. Somoza, J.R., Menon, S., Schmidt, H., J.-McCarthy, D., Dessen, A., Stahl, M.L., Somers, W.S., & Sullivan, F.X. Structural and kinetic analysis of *E. coli* GDP-mannose 4,6 dehydratase provides insights into the enzyme's catalytic mechanism and regulation by GDP-fucose. *Structure* **2000**, 8:123-135
15. Nakayama, K., Maeda, Y., & Jigami, Y. Interaction of GDP-4-keto-6-deoxymannose-3,5-epimerase-4-reductase with GDP-mannose-4,6-

- dehydratase stabilizes the enzyme activity for formation of GDP-fucose from GDP-mannose. *Glycobiol.* **2003**, 13:673-680
16. Ohyama, C., Smith, P.L., Angata, K., Fukuda, M.N., Lowe, J.B., & Fukuda, M. Molecular cloning and expression of GDP-mannose-4,6-dehydratase, a key enzyme for fucose metabolism defective in Lec13 cells. *J. Biol. Chem.* **1998**, 273: 14582-14587
 17. Sullivan, F.X., Kumar, R., Kriz, R., Stahl, M., Xu, G.-Y., Rouse, J., Chang, X.-J., Boodhoo, A., Potvin, B., & Cumming, D.A. Molecular cloning of human GDP-mannose 4,6-dehydratase and reconstitution of GDP-fucose biosynthesis *in vitro*. *J. Biol. Chem.* **1998**, 273:8193-8202
 18. Albermann, C. and Piepersberg, W. Expression and identification of the RfbE protein from *V. cholerae* O1 and its use for the enzymatic synthesis of GDP-D-perosamine. *Glycobiol.* **2001**, 11:655-661
 19. Sturla, L., Bisso, A., Zanardi, D., Benetti, U., Flora, A.D., & Tonetti, M. Expression, purification and characterization of GDP-D-mannose 4,6-dehydratase from *E. coli*. *FEBS Lett.* **1997**, 412:126-130
 20. Yamamoto, K., Katayama, I., Onoda, Y., Inami, M., Kumagai, H., & Tochikura, T. Evidence that the enzyme catalyzing the conversion of GDP-D-mannose to a 4-keto sugar nucleotide intermediate requires NADP. *Archiv. Biochem. Biophys.* **1993**, 300:694-698
 21. Preiss, J. and Wood, E. Sugar nucleotide reactions in *Arthrobacte*. *J. Biol. Chem.* **1964**, 239:3119-3126

22. Oths, P., Mayer, R.M., & Floss, H.G. Stereochemistry and mechanism of the GDP-mannose dehydratase reaction. *Carbohydr. Resear.* **1990**, 198:91-100
23. Jarvinen, N., Maki, M., Rabina, J., Roos, C., Mattila, P., & Renkonen, R. Cloning and expression of *H. pylori* GDP-L-fucose synthesizing enzymes (GMD and GMER) in *S. cerevisiae*. *Eur. J. Biochem.* **2001**, 268:6458-6464
24. Hegeman, A., Gross, H.W., & Frey, P.A. Probing catalysis by *E. coli* dTDP-glucose-4,6-dehydratase: identification and preliminary characterization of functional amino acid residues at the active site. *Biochemistry* **2001**, 40:6598-6610
25. Payne, S.H., & Ames, B.N. A procedure for rapid extraction and high-pressure liquid chromatographic separation of the nucleotides and other small molecules from bacterial cells. *Anal. Biochem.* **1982**, 123:151-161