The Mechanism of UDP-*N*-Acetylglucosamine 2-Epimerase

by

PAUL M. MORGAN

B.Sc., The University of Western Ontario, 1992

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

Department of Chemistry

We accept this thesis as conforming
To the required standard:

THE UNIVERSITY OF BRITISH COLUMBIA

May, 1998

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Chemistry

The University of British Columbia
Vancouver, Canada

Date June 30/98

DE-6 (2/88)

Abstract

The bacterial enzyme UDP-*N*-acetylglucosamine 2-epimerase (UDP-GlcNAc 2-epimerase) catalyzes the interconversion of UDP-*N*-acetylglucosamine (UDP-GlcNAc) and UDP-*N*-acetylmannosamine (UDP-ManNAc), by inverting the stereochemistry at C-2" of the nucleotide sugars. The mechanism of this enzyme is interesting, because the proton at the C-2" position is not acidic, and therefore the reaction must proceed via a pathway that is more complicated than a simple deprotonation followed by reprotonation in the inverted stereochemical sense.

A coupled enzyme assay for UDP-GlcNAc 2-epimerase activity has been developed using UDP-ManNAc dehydrogenase, an enzyme that catalyzes the NAD^+-dependent two-fold oxidation of UDP-ManNAc to form UDP-*N*-acetylmannosaminuronic acid. In order to obtain the substantially large amounts of the dehydrogenase required for the coupled assay, it was necessary to clone and overexpress the enzyme from *Escherichia coli*. The recombinant dehydrogenase was purified to homogeneity, and determined to have a k_{cat} of 1.2 ± 0.2 s^{-1} and an apparent K_m of 1.2 ± 0.4 mM under the conditions relevant to the coupled assay (at pH 8.8).

The coupled enzyme assay permitted the kinetic characterization of recombinant *E. coli* UDP-GlcNAc 2-epimerase. The epimerase was purified to homogeneity and determined to have a k_{cat} of 4.8 ± 0.2 s^{-1}, and an apparent K_m of 0.73 ± 0.09 mM. The epimerase also displays positive cooperativity, with a Hill coefficient of 2.3 ± 0.2. These values agree with those reported previously (Kawamura *et al.*, 1979).

The enzymatic epimerization in D_2O proceeds with the incorporation of deuterium into the C-2" position, supporting a mechanism that ultimately involves proton transfer at this position. The epimerization with 2"-^{2}H-UDP-GlcNAc as the substrate is slowed by a primary kinetic isotope effect (k_H/k_D = 1.8 ± 0.1) indicating that the C-H bond at C-2" is cleaved during a rate determining step of the reaction. The enzyme does not require the addition of any NAD^+ cofactor for activity, and experiments failed to detect any cofactor tightly bound within the enzyme, suggesting that NAD^+ is not involved in the epimerization mechanism. The enzyme is also observed to slowly release UDP and 2-acetamidoglucal when a large quantity of epimerase is incubated with substrate over extended periods. These observations are consistent with an unprecedented enzymatic epimerization mechanism that proceeds via cleavage of the anomeric C-O bond. The simplest reasonable description of the enzyme mechanism involves the *anti* elimination of UDP from UDP-GlcNAc, to form 2-acetamidoglucal and UDP as reaction intermediates, followed by the *syn* addition of UDP to the glycal double bond. Similarly, the reverse reaction requires the *syn* elimination of UDP from UDP-ManNAc coupled with the *anti* addition of UDP to the 2-acetamidoglucal intermediate.

An experiment to identify the general base or bases involved in the deprotonation of the C-2" proton required the synthesis of uridine 5'-diphosphate 2,3-epoxypropanol (UDP-glycidol), an affinity label specific for the UDP-GlcNAc 2-epimerase active site. However, under the experimental conditions, the covalent-labeling compound failed to inactivate the epimerase in an irreversible fashion. UDP-Glycidol was found, however, to irreversibly inactivate UDP-ManNAc dehydrogenase, and a specific cysteine residue is implicated as the site of covalent modification of the enzyme.

Table of Contents

List of Figures

List of Tables

Abbreviations and Symbols

δ	chemical shift
ε	extinction coefficient
A	(deoxy)adenosine (in nucleic acids)
Abs	absorbance
Ac	acetate
ADP	adenosine 5'-diphosphate
Amp^r	the phenotype of the gene which confers ampicillin resistance
ATP	adenosine 5'-triphosphate
B. cereus	*Bacillus cereus*
*Bam*HI	a restriction endonuclease
bp	base pairs
C	(deoxy)cytosine (in nucleic acids)
ca.	circa
CH_3CN	acetonitrile
Ci	Curie
CoA	Coenzyme A
ColE1 ori	origin of replication in pET11a
CPS	capsular polysaccharide
d	doublet (NMR)
D_2O	deuterium oxide
Da	dalton(s)
DEAE	diethylaminoethyl
DNA	deoxyribonucleic acid
dNTP	deoxyribonucletotides
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylene diamine tetraacetate (disodium salt)

Enz	enzyme
ESMS	electrospray ionization mass spectrometry
G	(deoxy)guanosine (in nucleic acids)
GalE	protein encoded by *galE* (UDP-galactose 4-epimerase)
GDP	guanosine 5'-diphosphate
h	hour(s)
*Hin*DIII	a restriction endonuclease
HPLC	high pressure/performance liquid chromatography
Hz	hertz
IPTG	isopropyl-1-thio-β-D-galactopyranoside
J	coupling constant (in NMR)
k_{cat}	catalytic rate constant (turnover number)
kDa	kilodalton(s)
K_{eq}	equilibrium constant
K_I	dissociation constant for an enzyme-inhibitor complex
K_m	Michaelis constant
lacA	gene encoding β-galactoside transacetylase
lacI	gene encoding the *lac* repressor protein
lacY	gene encoding galactoside permease
lacZ	gene encoding β-galactosidase
LB	Luria Bertani medium
LPS	lipopolysaccharide
LSIMS	Liquid Soft Ionization Mass Spectrometry
min	minute(s)
NAD^+	nicotinamide adenine dinucleotide
NADH	nicotinamide adenine dinucleotide, reduced form
$NADP^+$	nicotinamide adenine dinucleotide phosphate
$NADPH^+$	nicotinamide adenine dinucleotide phosphate, reduced form
NMR	nuclear magnetic resonance
NMWC	nominal molecular weight cut-off
m	multiplet (NMR)

mRNA	messenger RNA
*Nde*I	a restriction endonuclease
NfrC	protein encoded by the *nfrC* gene (UDP-GlcNAc 2-epimerase)
nfrC	gene identified as encoding UDP-GlcNAc 2-epimerase
ORF	open-reading frame
PAGE	polyacrylamide gel electrophoresis
pBR322	commercially available cloning vector
pCA62	plasmid containing a 4.9 kbp fragment of genomic *E. coli* DNA cloned into pBR322
PCR	Polymerase Chain Reaction
pET11a	commercially available expression vector
pKI86	plasmid containing *nfrC* (*rffE*) cloned into pET11a
PLP	pyridoxal phosphate
ppm	parts per million
psi	pounds per square inch
pUS01	plasmid containing *rffD* subcloned into pET11a
PVDF	polyvinyldifluoride
rfbE	gene encoding CDP-paratose 2-epimerase
rffD	gene encoding UDP-ManNAc dehydrogenase
RffD	protein encoded by *rffD* (UDP-ManNAc dehydrogenase)
rffE	gene encoding UDP-GlcNAc 2-epimerase
RffE	protein encoded by *rffE* (UDP-GlcNAc 2-epimerase)
RNA	ribonucleic acid
rpm	revolutions per minute
s	singlet (NMR); second(s)
S. enterica	*Salmonella enterica*
S. pneumonaie	*Streptococcus pneumoniae*
S. typhimurium	*Salmonella typhimurium*
*Sal*I	a restriction endonuclease
SDS	sodium dodecyl sulphate
t	triplet (NMR)

T	(deoxy)thymidine (in nucleic acids)
T4	a bacteriophage
T7	a bacteriophage
TBAHS	tetrabutylammonium hydrogen sulphate
TDP	thymidine 5'-diphosphate
Trien	triethanolamine
Tris	tris(hydroxy)amino methane
tt	triplet of triplets (NMR)
U	unit of enzyme activity
UDC	uridine 5'-diphosphate chloroacetol
UDP	uridine diphosphate
UDP-GlcNAc	uridine 5'-dihphosphate *N*-acetylglucosamine
UDP-glycidol	uridine 5'-diphosphate 2,3-epoxypropanol
UDP-ManNAc	uridine 5'-diphosphate *N*-acetylmannosamine
UDP-ManNAcUA	uridine 5'-diphosphate *N*-acetylmannosaminuronic acid
UMP	uridine 5'-monophosphate
UV	ultraviolet
Vis	visible
V_{max}	maximal reaction rate
V_o	initial reaction rate

Standard Abbreviations for Amino Acids:

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine

I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

Acknowledgements

I am indebted to the many who helped with the various stages of this thesis over the past few years. Specifically, I am grateful to all who contributed thoughts, time, equipment and/or materials. Some of these people provided specific assistance, and are acknowledged at the pertinent locations within these pages. However, I must extend an all-encompassing and hardy thanks to the many unsung others who have assisted in a general and kindly fashion. In particular, the lion-share of gratitude belongs to my supervisor, Martin Tanner, without whom absolutely none of this would have been possible. Only slightly less important were the ensemble cast of Tanner's minions, especially the high-priced editorial committee, who often seemed to be able to figure out what it was I was trying to say better than I could myself. The award for best contributor in a supporting role naturally goes out to my darling Jennifer, whose unbridled enthusiasm and encouragement has been irrepressible. Finally, I could not close without thanking my parents, who continue to have the uncanny knack of coming through when needed the most.

For

my grandparents

Chapter One:

Epimerases and Racemases

1.1 Introduction

Living organisms often exist in a hostile environment, competing for the limited available nutrients, while maintaining a defensive strategy to combat aggressors. On the molecular level, these two features are incontrovertibly linked to the three-dimensional arrangement of bio-molecules. Often, the available nutrients are present within the environment in unusual chemical forms, which first must be converted to readily metabolized molecules. In terms of survival, an organism capable of utilizing an unusual energy source would prove to be a far superior scavenger than one that couldn't. Similarly, organisms have adapted to mount defences against aggressive competitors. Their environment is full of offensive weapons evolved by other organisms, ranging from degradative enzymes, to antibiotics, to complex immune responses, each of which recognizes and targets highly specific three-dimensional cell-surface features. Defence against these types of attacks often lies in stealth and invisibility, since the offensive weapons will only perform their destructive tasks after recognizing their specified target.

Both survival modes are rooted on the same principle: stereochemical diversity. Organisms have adapted methods of deriving nutrients from unusual chemical compounds, while at the same time developing their own unique building blocks containing geometry unrecognizable to competitors, from which they build their own protective features. This chemical diversity comes from the ability of many of these organisms to invert chiral centres in order to create or metabolize molecules of unusual stereochemistry.

A class of enzymes has evolved to catalyze the inversion of stereocentres. In cases where the stereocentre acted upon is the sole chiral atom in the substrate molecule, these enzymes are known as racemases; in cases where multiple stereocentres are present in the

substrate, they are called epimerases. A large number of epimerases and racemases have been studied in some detail, and the mechanisms employed by these enzymes are as diverse as the types of compounds they act upon. As a result, these enzymes may be classified in terms of their substrates, and further subdivided by their mechanism type. Three comprehensive reviews have been published detailing the research that has been accumulated concerning each type (Glaser, 1972; Adams, 1976; Tanner and Kenyon, 1998).

Epimerases and racemases that have been identified to date act on amino acids, peptides, α-hydroxyacids, acyl-Coenzyme A, and most pertinent to this thesis – carbohydrates. Despite their variety of substrates, with only a few rare exceptions they all share a common general reaction scheme. During the reaction a carbon-hydrogen bond is first broken, then reformed in the inverted stereochemical sense (Figure 1.1). No enzymatic inversion is known to occur at an asymmetric carbon that is not hydrogen-substituted (Adams, 1976). The conversion is accomplished via a variety of methods, including proton transfer and hydride transfer mechanisms, which may involve the assistance of an enzyme cofactor. For instance, the α-hydroxyacid inverting enzyme, mandelate racemase, requires a divalent metal ion for the interconversion of the (*R*)- and (*S*)- forms of mandelate (Fee *et al.*, 1974). Among the examples of amino acid racemases, alanine racemase employs a pyridoxal phosphate (PLP) cofactor (Wood and Gunsalas, 1951), while another, glutamate racemase requires no cofactor at all (Nakajima *et al.*, 1986; Gallo and Knowles, 1993).

Like the amino acid racemases, the carbohydrate epimerases fall into two categories: those requiring cofactors, and those that do not. Generally, the cofactor used in these mechanisms is nicotinamide adenine dinucleotide (NAD^+), which serves as a hydride

acceptor during the catalytic reaction. The mechanisms of the carbohydrate epimerases will be explored in detail in the following section.

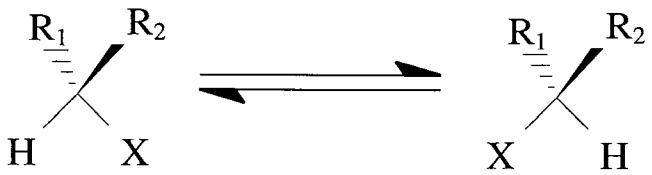


Figure 1.1 The general reaction scheme for epimerases and racemases. X can be OH as in the carbohydrate and the α-hydroxyacid epimerases, NH_2 as in the amino acid epimerases, or CH_3 for acyl-CoA epimerases.

1.2 Carbohydrate Epimerases

For the most part, the carbohydrate epimerases act by removing a proton (as opposed to a hydride) from the chiral centre of interest, and reattach it in the opposite stereochemical sense, resulting in the inversion of the stereocentre (Adams, 1976). This proton involved in the reaction must therefore be acidic enough to be removed by a base in the enzyme's active site. The simplest reasonable mechanism is typified by the enzyme D-ribulose 5-phosphate 3-epimerase, which catalyzes the removal of a proton directly adjacent to a carbonyl functionality, generating an enediol intermediate (Figure 1.2). Reprotonation of the opposite face of the planar ene-diol yields the inverted stereochemical product, D-xylulose 5-phosphate (Adams, 1976; Glaser, 1972; Tanner and Kenyon, 1998).

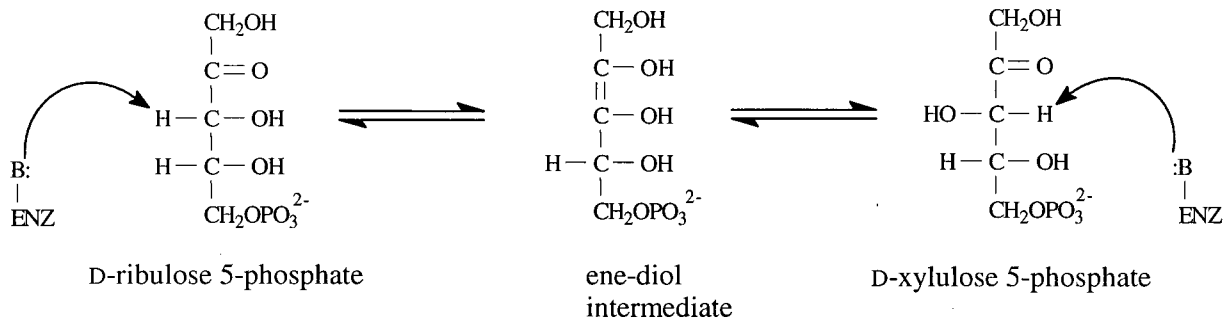


Figure 1.2 The mechanism of D-ribulose-5-phosphate 3-epimerase.

A similar mechanism is observed in the reaction catalyzed by dTDP-L-rhamnose synthetase (Glaser, 1972; Tanner and Kenyon, 1998). Although this enzyme cannot be accurately classified as an epimerase, since it doesn't simply invert the stereochemistry of its substrate, it does however have epimerization steps along its proposed mechanistic pathway (Figure 1.3). The synthetase consists of two separate polypeptide chains, each of which is responsible for one of the two reactions carried out by this enzyme. During the course of the first reaction, both C-3" and C-5", the two centres flanking the carbonyl at C-4", are concomitantly inverted. As in the previous example, both labile protons adjacent to the carbonyl are ultimately removed, and replaced in the opposite stereochemical sense.

A third example is the enzyme *N*-acetylglucosamine 2-epimerase. The mechanism of this enzyme has not been studied in any great detail, but it is reasonable to surmise it follows the previous examples by reacting through an opened-chain intermediate. The open chain bears an aldehyde adjacent to the reaction site (Figure 1.4), which effectively lowers the pKa of the proton at the reaction centre, facilitating deprotonation (Tanner and Kenyon, 1998).

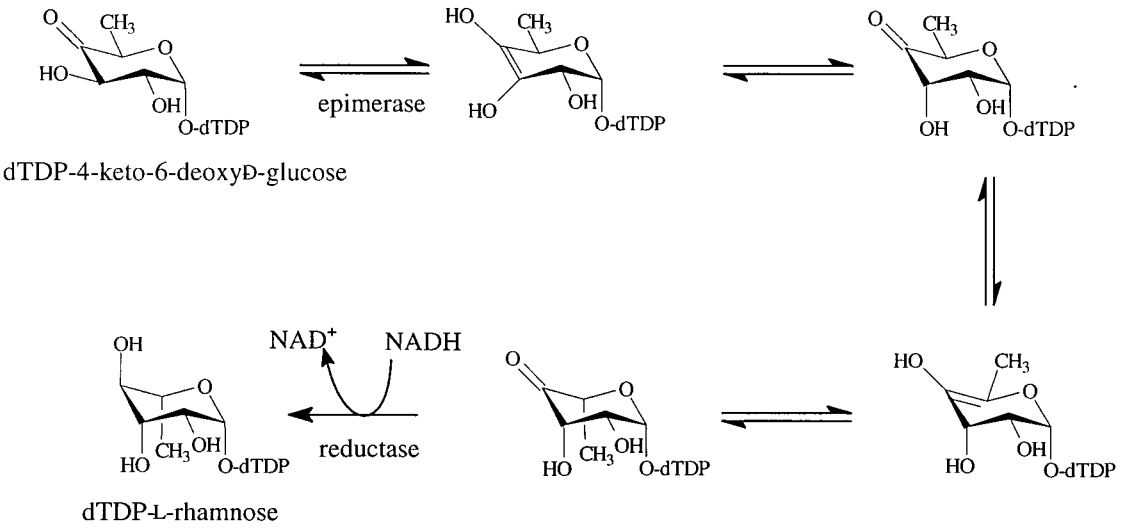


Figure 1.3 The proposed intermediates formed during the reaction catalyzed by dTDP-L-rhamnose synthetase.

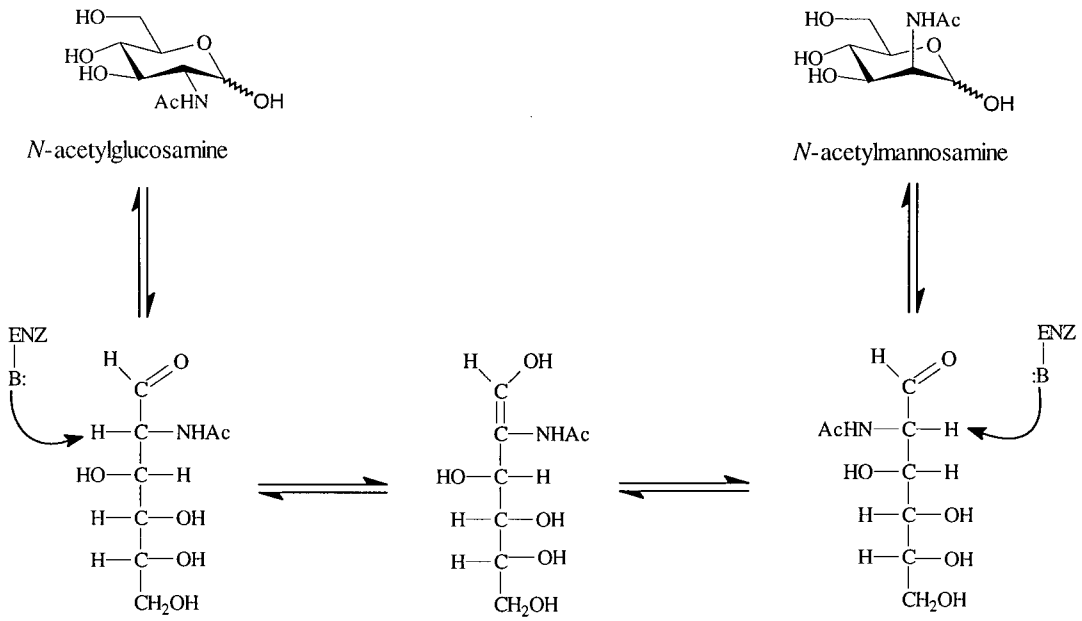


Figure 1.4 The proposed mechanism of *N*-acetylglucosamine 2-epimerase.

These first three examples together illustrate the case where the proton to be ultimately removed during the reaction is adjacent to an electron withdrawing group, and thus is acidic enough, or "activated" for deprotonation.

Alternatively, cases exist where the stereochemical inversion occurs at an "unactivated" or non-acidic proton. Enzymes of this type commonly employ a cofactor as a strategy to assist the reaction. This is exemplified by an enzyme fraction from the green algae *Chlorella pyrenoidosa*, which reportedly catalyzes the inversion at two stereocentres, in order to interconvert GDP-D-mannose to GDP-L-galactose (Barber, 1979; Barber and Hebda, 1982). This reaction, showing similarities to the epimerization half of the reaction catalyzed by TDP-L-rhamnose synthetase noted above (Figure 1.4), has NAD^+ first abstract a hydride from the carbon adjacent to the two protons destined for removal (Figure 1.5). The generated

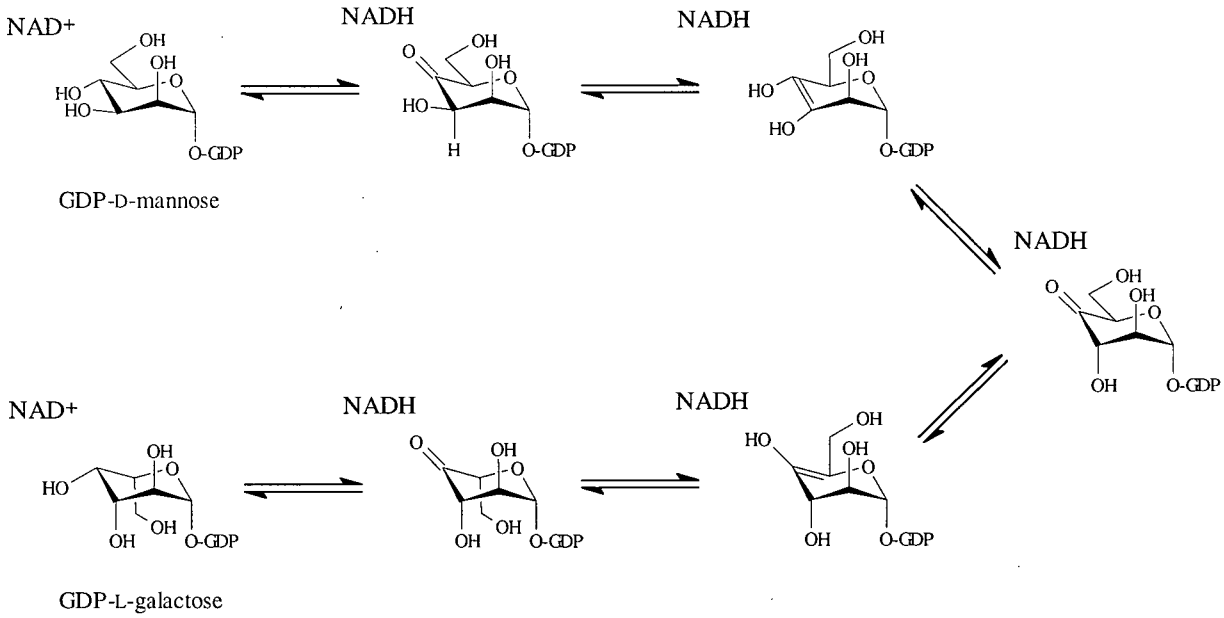


Figure 1.5 The proposed intermediates formed during the enzymatic interconversion of GDP-D-mannose and GDP-L-galactose.

ketone can then serve to alternately 'activate' both the protons at C-3" and C-5" for the reaction. Inversion of the stereochemistry at both of these carbons is achieved through the reprotonation of the opposite faces of the respective enol-intermediates. Subsequent reduction of the ketone affords the product, GDP-L-galactose.

The one enzyme that stands alone among the carbohydrate epimerases is UDP-galactose 4-epimerase. This enzyme is the best understood of all carbohydrate epimerases. While the previously mentioned enzymes proceed through a deprotonation/reprotonation mechanism, this enzyme operates instead through a direct hydride transfer (Frey, 1987).

UDP-galactose 4-epimerase contains a NAD^+ cofactor tightly bound within the enzyme's active site. This cofactor remains bound even after extensive dialysis. Furthermore, the successful removal of the NAD^+ from the active site is accompanied by denaturation of the enzyme. Rather than employing the cofactor to facilitate deprotonation, the cofactor in this case oxidizes the centre of inversion directly, converting the reactive carbon's functionality from a hydroxyl to a ketone (Figure 1.6). Studies have shown that substrate labeled with a deuterium isotope at C-4" is enzymatically converted without any loss of label to the bulk solvent, indicating that the hydrogen atom removed is the same that is returned (Maxwell, 1957; Frey, 1987). Further experiments have used either NaB^3H_4 or UDP-[4-^{3}H]-fucose in order to generate an inactive enzyme-NAD^3H abortive-complex with the tritium label exclusively located on the β-face of the cofactor (Nelsestuen and Kirkwood, 1971). Therefore, during the course of the reaction only one face of the NAD^+ cofactor is actually presented to the substrate.

Conversely, to bring about an inversion in the manner depicted, the keto-intermediate must be capable of exposing either face to the cofactor. Further insight into this requirement

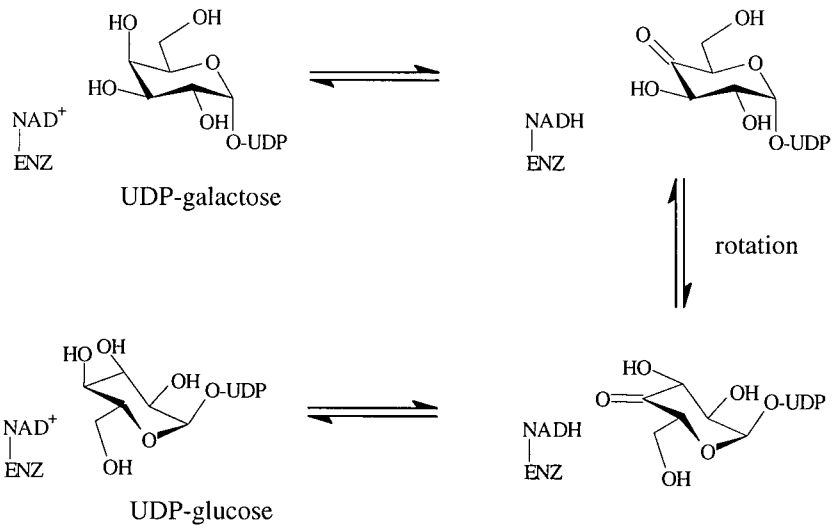


Figure 1.6 The mechanism of UDP-galactose 4-epimerase.

was gained by the salient observation that a variety of aldoses (such as glucose and galactose), when incubated with UMP and the epimerase, inactivated the enzyme (Kang *et al.*, 1975; Ketley and Schellenberg, 1973; Seyema and Kalckar, 1972). Under these conditions, the bound sugar is oxidized by the enzyme's cofactor, and released from the active site, leaving behind the enzyme in an inactive, reduced NADH form. In the example of inactivation by glucose, the aldose is oxidized at C-1, rather than at the normal reaction centre (C-4). The fact that the enzyme is relatively indiscriminate towards the identity and the orientation of the sugar has led to the suggestion that the sugar moiety of the normal substrate is held only loosely by the enzyme. On the other hand, the anionic diphosphate present in the nucleotide moiety provides a tenable "handle" for substrate binding. The overall picture that emerges then involves the binding of either substrate to the enzyme's active site, and the subsequent oxidation to the UDP-4-ketopyranose intermediate. During

the lifetime of the intermediate, the keto-hexose moiety of the intermediate rotates within the active site, independently of the nucleotide moiety, exposing one face or the other to the cofactor for subsequent reduction.

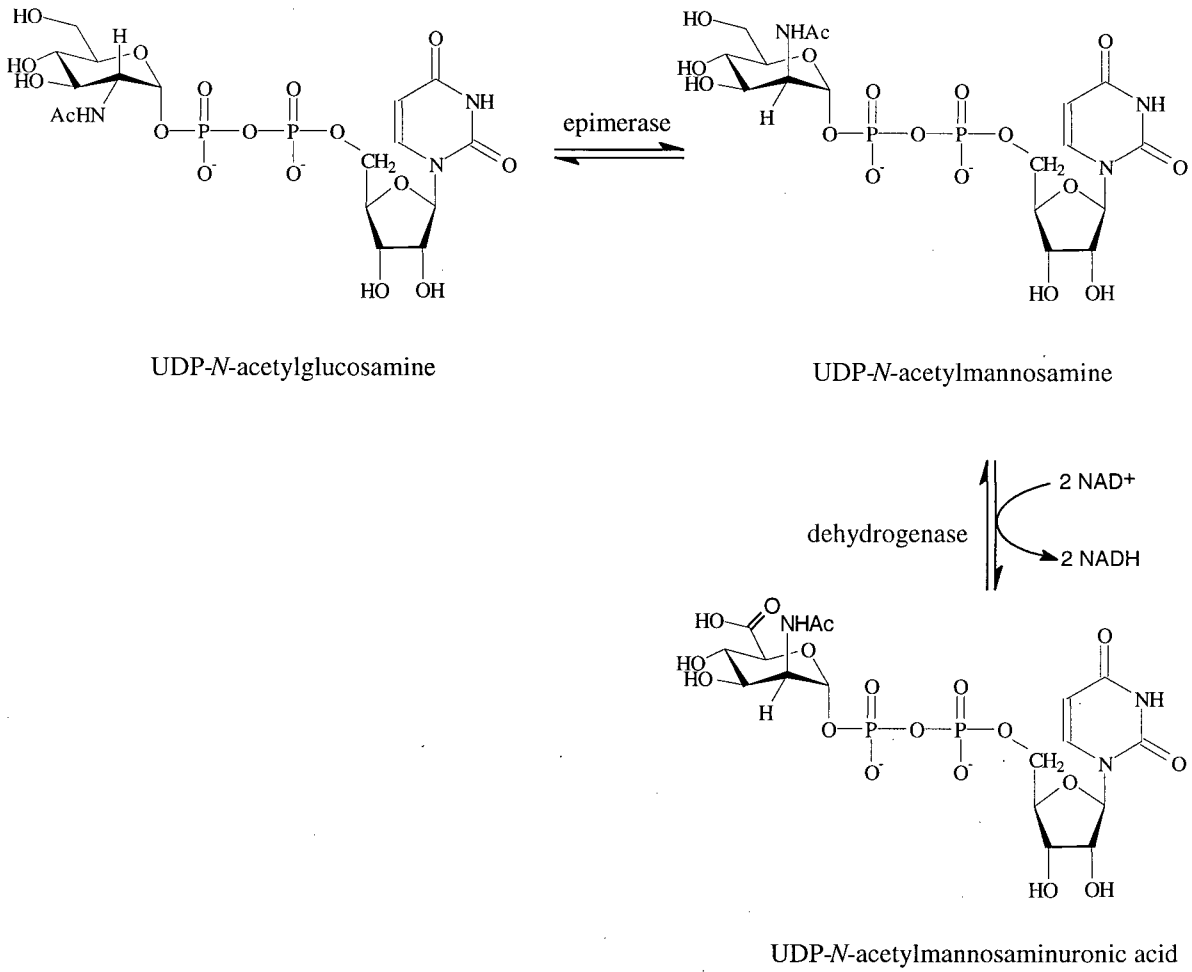
Interestingly, it has also been observed that the keto-intermediate, UDP-4-ketoglucose is occasionally released from the enzyme's active site into free solution, leaving behind an inactivated NADH-form of the enzyme, incapable of any further turnovers (Wee *et al.*, 1972; Wee and Frey, 1973; Liu *et al.*, 1996). This inactive form can be regenerated in two ways. First, if the enzyme were denatured, it can refold spontaneously in the presence of fresh NAD^+. Alternatively, the enzyme can be exposed to an abundance of synthetic UDP-4-ketoglucose, which can accept the hydride from the NADH in the normal mechanistic fashion, and consequently return the cofactor to its active NAD^+ form.

It is uncommon however, for an enzyme to favour release of its intermediates into free solution, as that practice would be catalytically inefficient, or as exemplified by UDP-galactose 4-epimerase, catalytically suicidal. Furthermore, in many cases enzyme reaction intermediates are unstable outside the enzyme's active site pocket.

In summary, carbohydrate epimerases employ a variety of strategies to bring about their inversions. In the cases that involve a deprotonation step along the mechanistic route the proton to be removed must be relatively acidic, or it must somehow be activated towards deprotonation. One strategy for altering the pKa at an "unactivated" centre is to employ a cofactor to generate a ketone adjacent to the reaction site. Alternatively, in the case of UDP-galactose 4-epimerase, the enzyme employs a direct hydride-transfer rather than a proton-transfer mechanism. Any other mechanism would be truly unique amidst this class of enzyme.

1.3 UDP-*N*-Acetylglucosamine 2-Epimerase and Its Biosynthetic Role

The bacterial enzyme UDP-*N*-acetylglucosamine 2-epimerase (UDP-GlcNAc 2-epimerase) catalyzes the interconversion of UDP-*N*-acetylglucosamine (UDP-GlcNAc) and UDP-*N*-acetylmannosamine (UDP-ManNAc) by inverting the stereochemistry at the inactivated C-2" stereocentre of the two substrates (Figure 1.7). The mechanistic possibilities of this enzyme will be explored later in this chapter.



UDP-*N*-acetylglucosamine

UDP-*N*-acetylmannosamine

UDP-*N*-acetylmannosaminuronic acid

Figure 1.7 The reactions catalyzed by UDP-*N*-acetylglucosamine 2-epimerase and UDP-*N*-acetylmannosamine dehydrogenase.

Biosynthetically, UDP-ManNAc can either be used directly in subsequent steps, or first converted to UDP-*N*-acetylmannosaminuronic acid (UDP-ManNAcUA) through the two-fold oxidation at C-6". This reaction is catalyzed by the enzyme UDP-ManNAc dehydrogenase. These two UDP-sugars are the activated forms of *N*-acetylmannosamine (ManNAc) and ManNAcUA respectively, both of which are incorporated into a wide variety of bacterial cell surface polysaccharides.

Traditionally, bacteria have been divided into two different classes based on their empirical susceptibility to staining by crystal violet and iodine. This chemical combination, which is known as the Gram stain, stains cells, as the crystal violet penetrates the cell wall and forms an insoluble complex with iodine. Using this stain, bacteria can be classified as either gram-positive or gram-negative depending on the characteristics of their cell-wall structures.

The first layer that bounds the cell is a phospholipid bilayer membrane common to both types of bacteria, known as the inner membrane (Figure 1.8). Beyond this membrane, gram-positive cells have a thick polymeric coating of sugar and peptide that wraps around the cell, giving it shape and rigidity, and protects it from lysing under osmotic pressure. This outermost layer of the gram-positive bacteria is commonly referred to as peptidoglycan. In contrast, gram-negative cells have only a thin layer of peptidoglycan along with an additional coating of lipopolysaccharide and lipoprotein coating the cell-surface, called the outer membrane.

During the Gram-staining procedure, following the initial application of crystal violet and iodine, ethanol is added to the cells. The added ethanol serves to dehydrate the cell wall. In the case of the gram-positive bacteria, the thick peptidoglycan is dehydrated, and acts as a

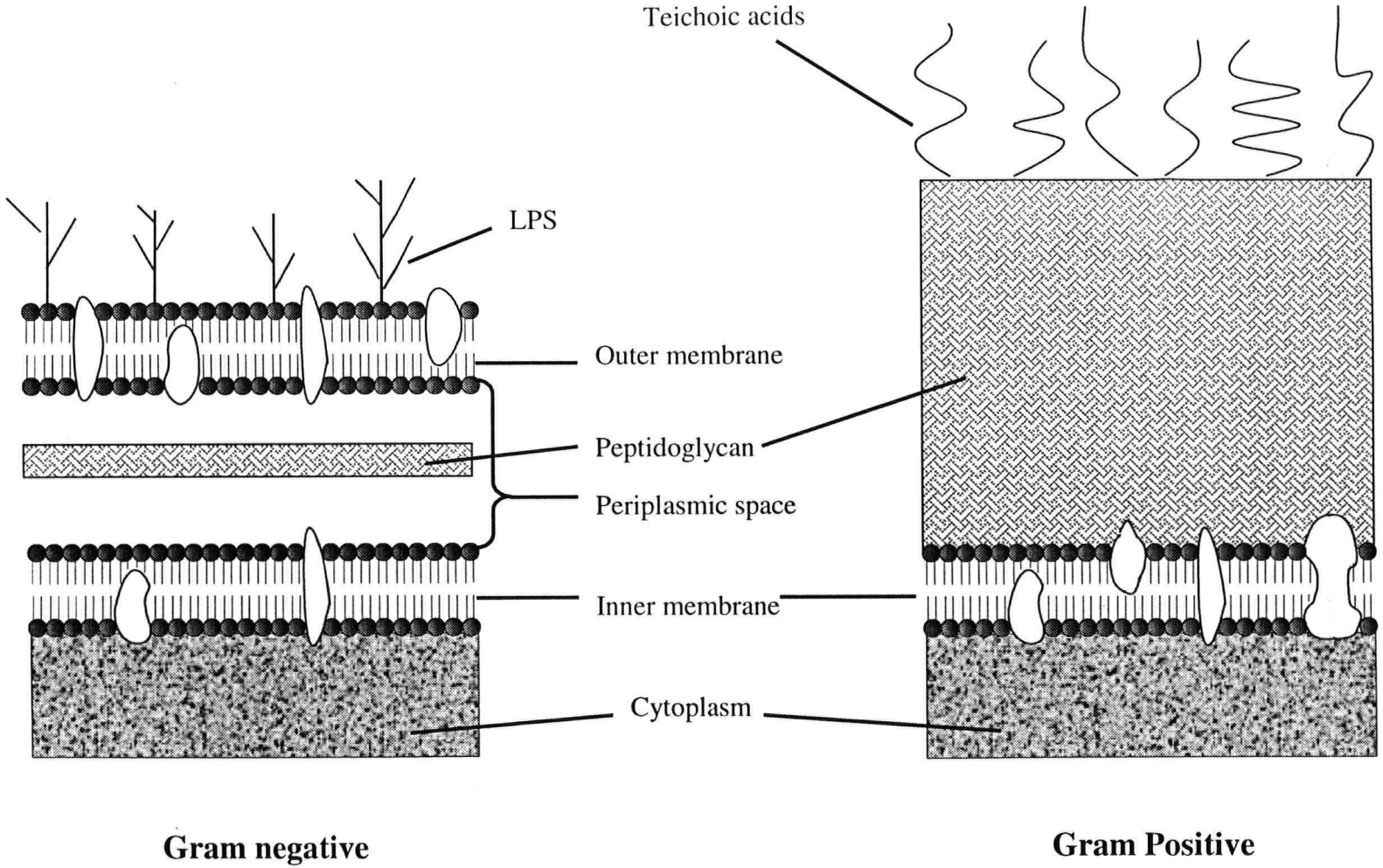


Figure 1.8 The cell wall features of gram-negative and gram-positive bacteria

barrier, which traps the stain within the cell. Conversely, gram-negative cells have only a thin peptidoglycan layer that does not serve as an adequate barrier to retain the stain inside the cell following treatment with ethanol.

Beyond their outermost layers, both types of bacteria present complex polysaccharides and lipopolysaccharides that protrude from the cell surface. In gram-positive bacteria, these polymers are the teichoic acids, which are covalently bound to the periphery of the thick peptidoglycan layer. In gram-negative bacteria, an entirely different type of polymeric surface antigen, the lipopolysaccharide (LPS) is presented on the outermost bacterial envelope.

1.3.1 Cell Wall Polysaccharides in Gram-Positive Bacteria

Gram-positive bacteria, whose many members include, for example, *Streptococcus pneumoniae*, *Staphylococcus aureus*, and *Bacillus subtilis*, all contain polymeric structures called teichoic acids that are found covalently attached to the peptidoglycan, accounting for up to 60% of the total cell-wall mass (Heptinstall *et al.*, 1970; Baddiley, 1989). Teichoic acids are polymers of either ribitol or glycerol, linked through phosphodiester bonds (Figure 1.9). The polymers are covalently modified with glycosyl or alanyl residues to provide structural diversity, and linked directly to the peptidoglycan. Although the chemical (and geometrical) structure of the wall-bound teichoic acids can differ significantly, in the cases studied to date they are all attached to the peptidoglycan by the same common linking unit, consisting of glycerophosphate, GlcNAc and ManNAc (Figure 1.10).

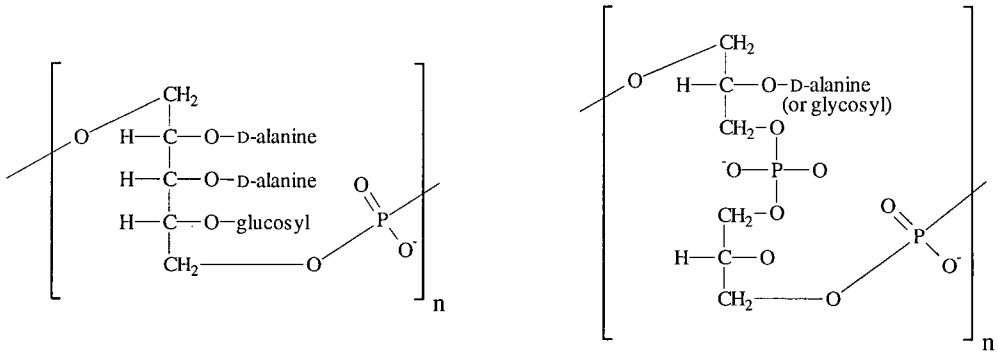


Figure 1. 9 The two types of repeating units of teichoic acids: the ribitol type (left) and the glycerol type (right).

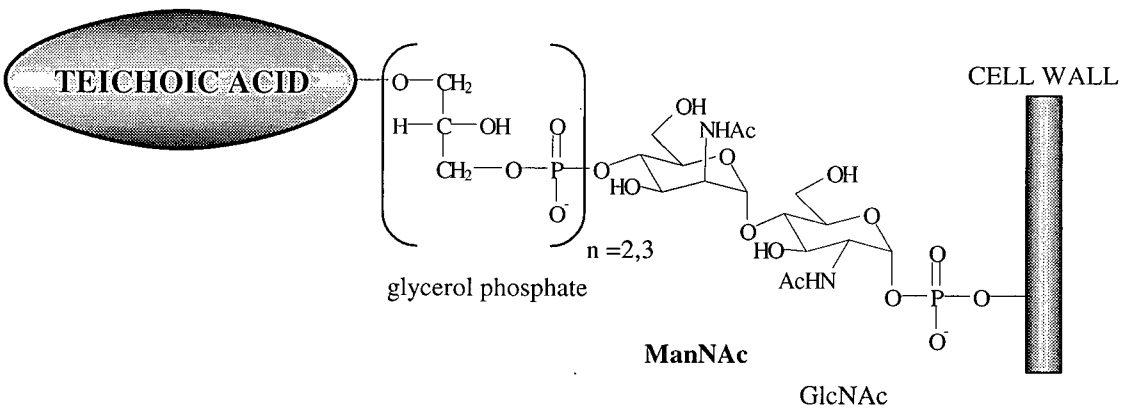


Figure 1.10 The common linking unit that joins wall-bound teichoic acids to the bacterial cell wall.

While the function of teichoic acids remains poorly understood, the fact they are so widespread among most gram-positive bacteria hints at the magnitude of their biological importance. Furthermore, bacteria that lack teichoic acids are observed to present chemically similar replacement polysaccharides on their surface, which presumably function in the same

manner (Heptinstall *et al.*, 1970). In one unique case, in lieu of teichoic acids, the gram-positive bacterium *Micrococcus lysodeikticus* expresses a polysaccharide whose repeating unit is known to contain ManNAcUA (Biely and Jeanloz, 1969; Page and Anderson, 1972). It has been suggested that teichoic acids are important in maintaining the cation concentration at the cell surface, in order to ensure a proper ionic environment for bacterial enzymes that operate near the cell surface (Baddiley, 1970; Hepinstall *et al.*, 1970; Baddiley, 1989). The cation concentration also has implications in preventing cell aggregation, since the like-charged surfaces would repel one another. Also, teichoic acids might play a role in cell division, as cells lacking these surface polysaccharides aren't capable of separating during cell division, resulting in elongated strings of joined cells (Tomasz, 1981). Finally, they may also have some immunological properties, arising from interactions between cells and the immune system of a host. Empirically, teichoic acids have been implicated in inflammation during pneumococcal infections (Tuomanen *et al.*, 1985; Carlsen *et al.*, 1992; Fischer *et al.*, 1993), and provoke antibody responses in rabbit and human models (Knox and Wicken, 1973).

1.3.2 Outer Membrane Polysaccharides in Gram-Negative Bacteria

UDP-GlcNAc 2-epimerase is also found in gram-negative bacteria. Its most notable function lies in the biosynthesis of the Enterobacterial Common Antigen (ECA). The ECA is a cell-surface polysaccharide present in all members of the enterobacteria (gut bacteria) family, which includes *Salmonella*, *Escherichia* and *Morganella*. ECA is exclusive to this family, and has an identical structure in all species. Mannosaminuronic acid is found in the trisaccharide repeating-unit of this polysaccharide (Figure 1.11, Kuhn *et al.*, 1988).

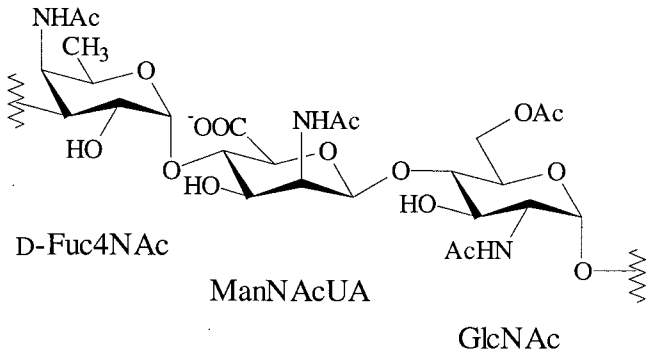


Figure 1.11 The *N*-acetylmannosaminuronic acid containing repeating unit of the Enterobacterial Common Antigen (ECA).

Similar to teichoic acids, the role of these acidic sugar molecules are thought to be involved in divalent cation binding, and the maintenance of the desired Ca^{2+} and Mg^{2+}concentration near the cell surface (Kuhn *et al.*, 1988). In addition, ECA may play a pathogenic role, as mutants of *Salmonella typhimurium* lacking ECA are 10-fold less virulent than the normal bacteria (Valtonen *et al.*, 1976). While the true function of ECA remains speculative, the fact that evolution has retained this invariable structure in all enterobacteria suggests an important role.

While ECA is found exclusively among the enterobacteria, other gram-negative bacteria also express their own LPS. The most common of these is known as the O-antigen. This lipopolysaccharide is composed of three regions: Lipid A, which anchors the structure to the outer cell membrane, the core polysaccharide, and the O-antigen region. In all natural occurrences, the Lipid A and the core region are both highly conserved. Only the peripheral O-antigen region is variable. In a single rare case, *Salmonella enterica* has been found to contain a disaccharide repeating unit consisting of only ManNAc sugars, alternately joined

by β(1→4) and β(1→3) linkages (Keenleyside *et al.*, 1994). The O-antigen is considered to play an important pathogenic role in bacteria, perhaps participating in the recognition of the host cells (Voet and Voet, 1990a).

1.3.3 Capsular Polysaccharides

In addition to the above-mentioned structural features, both types of bacteria are often coated with a cellular capsule, sometimes descriptively called the "slime layer". This capsule consists of a polysaccharide (Capsular Polysaccharide, CPS), that in several cases contains either ManNAc or ManNAcUA in the polymeric repeating unit. The capsule is not essential for bacterial growth, as it is observed under certain environmental conditions and absent in others, but it has been determined to be the major virulence factor in many strains of pathogenic bacteria (Lee *et al.*, 1991). The capsule itself does not cause toxicity, but it does prevent immunological response against bacterial infection, presumably by allowing the bacteria to evade detection by hiding recognizable cell-surface features – the features that would normally induce a response by the host immune system. Of the eighty-three known types of *Streptococcus pneumoniae*, ManNac appears in the capsular polysaccharide of six, including types 19A and 19F (Figure 1.12), which are the most frequent source of pneumococcal disease in infants and children (Lee *et al.*, 1991).

Certain strains of gram-negative *E. coli* have a capsule, which is commonly referred to as the K antigen (K for Kapsule). One particular type, known as the K7 antigen contains ManNAcUA in the disaccharide repeating-unit (Mayer, 1969; Ichihara *et al.*, 1974). In a fashion akin to the pneumococcal capsule, the *E. coli* K7 capsule might play a role in

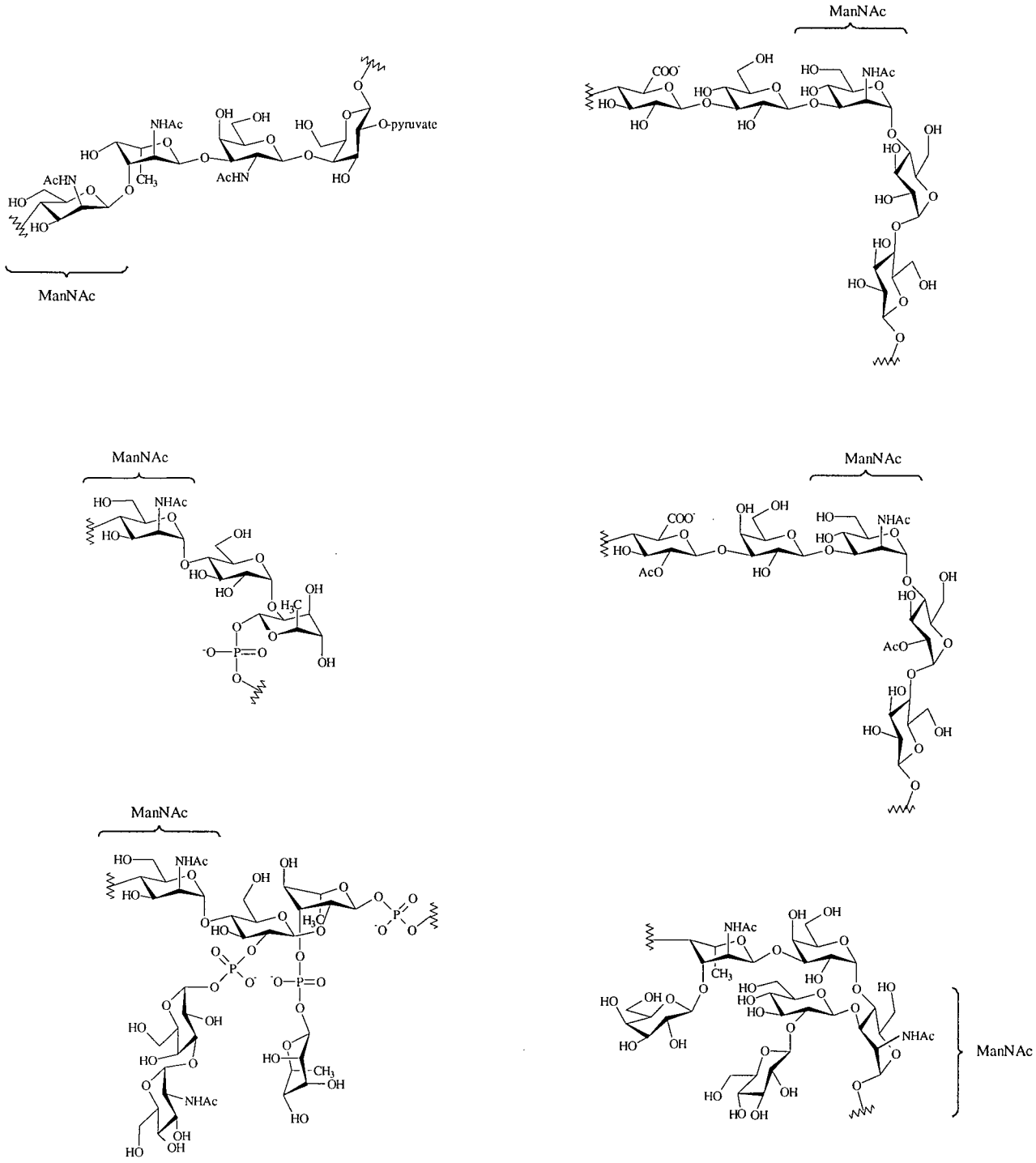


Figure 1.12 ManNAc appears in several *S. pneumoniae* CPS repeating units. Clockwise from the top left are types 4, 9N, 9V, 12N, 19A and 19F.

assisting the bacterium evade host immune response, and therefore act as a major virulence factor (Tsui *et al.*, 1982).

1.4 Previous Studies on Mammalian UDP-*N*-Acetylglucosamine 2-Epimerase

In 1958, Comb and Roseman noted that a crude enzyme extract obtained from rat liver was capable of converting UDP-GlcNAc to UDP and free *N*-acetylmannosamine (ManNAc). This was followed by the subsequent purification of a single "bifunctional" rat liver enzyme that could perform both the epimerization and hydrolysis steps apparently required for this reaction (Spivak and Roseman, 1966). At some point along the way, this mammalian enzyme was somewhat inappropriately named UDP-*N*-acetylglucosamine 2-epimerase, even though it is not in fact a true epimerase, as the enzyme's product is not an epimer of the substrate. Therefore, as the isolation of the true, bacterial epimerase was still a decade into the future, the preponderance of the early reports in the literature concentrate on the misnamed mammalian enzyme.

The size of the 2-epimerase isolated from rat liver was determined in two separate investigations to lie between 300-400 kDa and 400-600 kDa respectively (Sommar and Ellis, 1972a; Kikuchi and Tsuiki, 1973). Detailed studies on this enzyme however proved to be difficult, as the enzyme was found to be notoriously unstable, often losing all of its activity in only a few hours (Glaser, 1960; Spivak and Roseman, 1966; Sommar and Ellis, 1972a). Despite the problems handling the enzyme a few experiments bearing mechanistic relevance were successful.

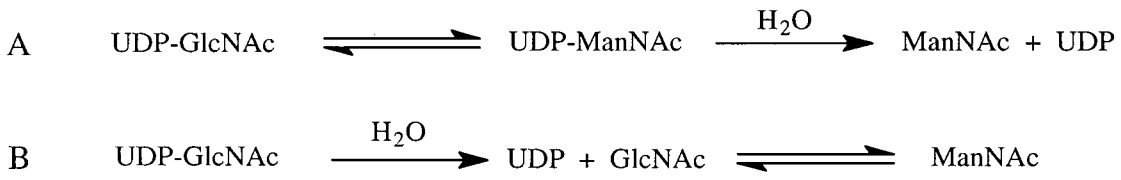


Figure 1.13 Two possible sequences for the reactions catalyzed by mammalian UDP-GlcNAc 2-epimerase.

If the mammalian epimerase does in fact catalyze two discrete steps, two obvious possibilities exist: either the epimerization is catalyzed first, followed by hydrolysis of the UDP nucleotide, or vice-versa. In the first case, the reaction would necessarily have UDP-ManNAc as a reaction intermediate (Figure 1.13, path A), while the other pathway would instead have an intermediary *N*-acetylglucosamine (GlcNAc) epimerized by the enzyme in the second catalytic step (Figure 1.13, path B). The preliminary efforts focussing on the mammalian epimerase concentrated on distinguishing between these two possibilities.

In order to test the viability of the first proposed pathway, enzyme and substrate were incubated in tritium-enriched water. Under these experimental conditions, a solvent derived tritium label was observed to be incorporated into ManNAc at C-2, indicating that the proton at the site of inversion is removed during the reaction, and replaced by another derived from the bulk solvent. When the reaction was quenched before all the UDP-GlcNAc had been converted into ManNAc, however, no isotopic label was observed in the recovered starting substrate (Glaser, 1960). This observation suggests that the epimerization could not be the first step (i.e. path A in Figure 1.13), as this reversible step would allow trititium-substituted

UDP-ManNAc to equilibrate with UDP-GlcNAc, resulting in isotope incorporation back into the UDP-GlcNAc pool. Furthermore, no UDP-ManNAc could be detected at all in the substrate pool.

In a later experiment, synthetic UDP-ManNAc was similarly incubated with the mammalian epimerase in tritium-enriched water (Salo and Fletcher, 1970). Again, if the reversible epimerization were to occur first in the reaction sequence, UDP-ManNAc would equilibrate with UDP-GlcNAc, simultaneously introducing an isotopic label into the latter sugar nucleotide. However, no UDP-GlcNAc at all was detected, despite the fact that the enzyme did handle UDP-ManNAc in the expected fashion (albeit slower than the natural substrate), hydrolyzing the sugar nucleotide to form free ManNAc. Furthermore, none of the UDP-ManNAc recovered before complete turnover possessed a tritium label. Both these observations are consistent with UDP-ManNAc acting as an alternative substrate in an irreversible reaction, rather than as an intermediate.

With the first pathway collapsing under scrutiny, focus turned to the second viable path. But this path was quickly dismissed when it was determined that the enzyme did not epimerize *N*-acetylglucosamine at all, which would be expected if GlcNAc were an intermediate (Spivak and Roseman, 1966). If neither GlcNAc nor UDP-ManNAc serve as intermediates then the enzymatic mechanism must be more complicated than simply following two discrete reaction steps.

In an effort to explain all the experimental results, Salo and Fletcher proposed the mechanism depicted in Figure 1.14 (Salo and Fletcher, 1970). In the first proposed step, the substrate reacts with the enzyme forming a glycosyl-enzyme intermediate, with the concomitant loss of UDP. This first step is effectively irreversible, preventing the

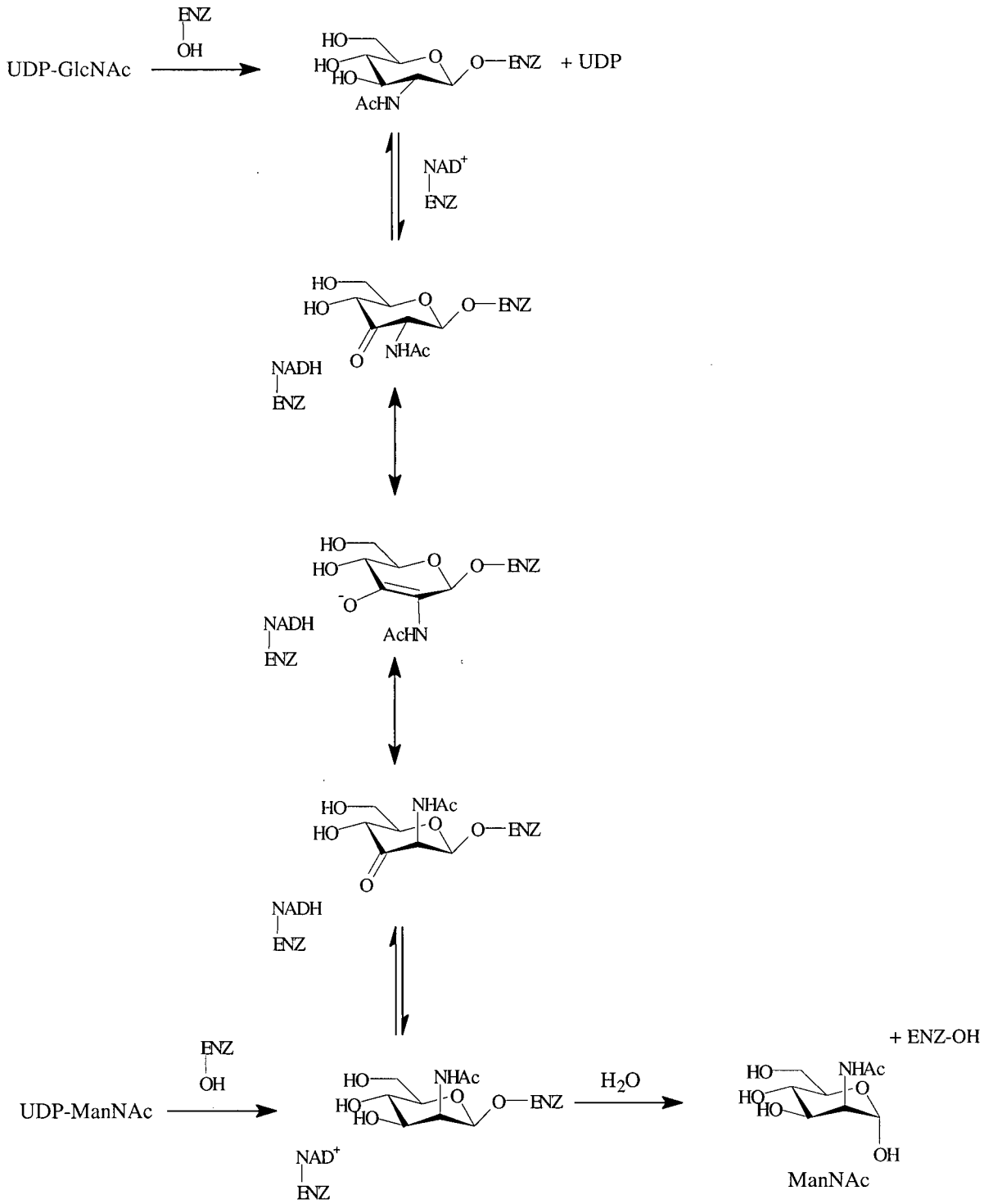


Figure 1.14 Salo's proposed mechanism (Adapted from Salo and Fletcher, 1970).

transformation of UDP-ManNAc to UDP-GlcNAc. During the second step of the proposed reaction, applying a strategy similar to the enzyme fraction that converts GDP-D-mannose to GDP-D-galactose (described above), the C-3 hydroxyl is oxidized to form the enzyme-bound 3-keto intermediate, which is readily converted to the other epimeric form of the sugar by deprotonation and subsequent reprotonation at C-2. The reaction can then completed by the regeneration of a hydroxyl at C-3, and the hydrolysis of ManNAc from the enzyme's active site.

While this proposed mechanism does conform to the experimental evidence concerning the mammalian enzyme, it is however based on two foundationless assumptions (Adams, 1976). First, the enzyme shows no rate enhancement upon the addition of extraneous NAD^+ cofactor, although the enzyme might very well contain such a cofactor tightly bound in the enzyme's active site, which was indeed the case for UDP-galactose 4-epimerase (see above). Second, the proposed mechanism requires the stereospecific addition of a solvent derived proton across the central enol-intermediate. Otherwise, hydrolysis of the enzyme bound substrate would yield both ManNAc and GlcNAc, which is not observed. Furthermore, no prior evidence exists supporting a covalently bound enzyme-intermediate, which is obligatory in the model to obviate the presence of GlcNAc in the reaction pathway, since the enzyme's inability to handle that substrate had previously been determined.

Sommar and Ellis later proposed an alternative mechanism involving the elimination of UDP from the UDP-GlcNAc substrate, to produce 2-acetamidoglucal as an intermediate, which is accompanied by the release of UDP from the active site, and into free solution (Figure 1.15; Sommar and Ellis, 1972b). Product inhibition studies presented in the same

report support the notion that UDP is released in an ordered mechanism followed by the irreversible formation of ManNAc.

The strongest support for the Sommar and Ellis mechanism however is derived from the experimental observation that synthetic 2-acetamidoglucal was converted exclusively to ManNAc by the enzyme. Conversely, non-enzymatic hydration of 2-acetamidoglucal favours the thermodynamically more stable GlcNAc.

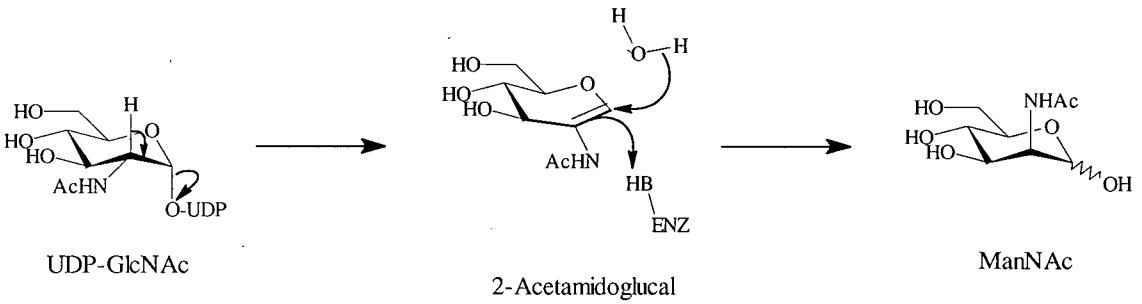


Figure 1.15 The mechanism proposed by Sommar and Ellis (adapted from Sommar and Ellis, 1972b).

1.5 Bacterial UDP-*N*-Acetylglucosamine 2-Epimerase

At the onset of this study very little was known concerning the bacterial epimerase. This enzyme was first reported partially purified by Kawamura *et al.* in 1975, and has since been characterized. In terms of prior mechanistic studies however, only a single paper exists in the literature. Within this article, the bacterial enzyme is reported to have been reacted in tritium-enriched water, resulting in the observed incorporation of a solvent-derived tritium label into the C-2" position of both nucleotide sugars in the equilibrated substrate pool (Salo, 1976). Based on this solitary experimental result, Salo proposed a similar NAD^+ requiring

mechanism for the bacterial epimerase as he did for the mammalian. He again invokes an NAD^+-dependent transient oxidation at C-3", in order to make the proton at C-2" more labile (Figure 1.16). Contrary to the mammalian enzyme though, the bacterial epimerase reaction is non-hydrolytic, so that this new proposal differs in the lack of requirement for an enzyme-bound intermediate. As in the case with the mammalian enzyme, the bacterial enzyme does not require any exogenous cofactor added to the assay buffer, but the presence of an NAD^+ tightly bound in the active site can not yet be ruled out. At any rate, this mechanism stands only as a single candidate of several, which for the purposes of this thesis, will be referred to as Path A.

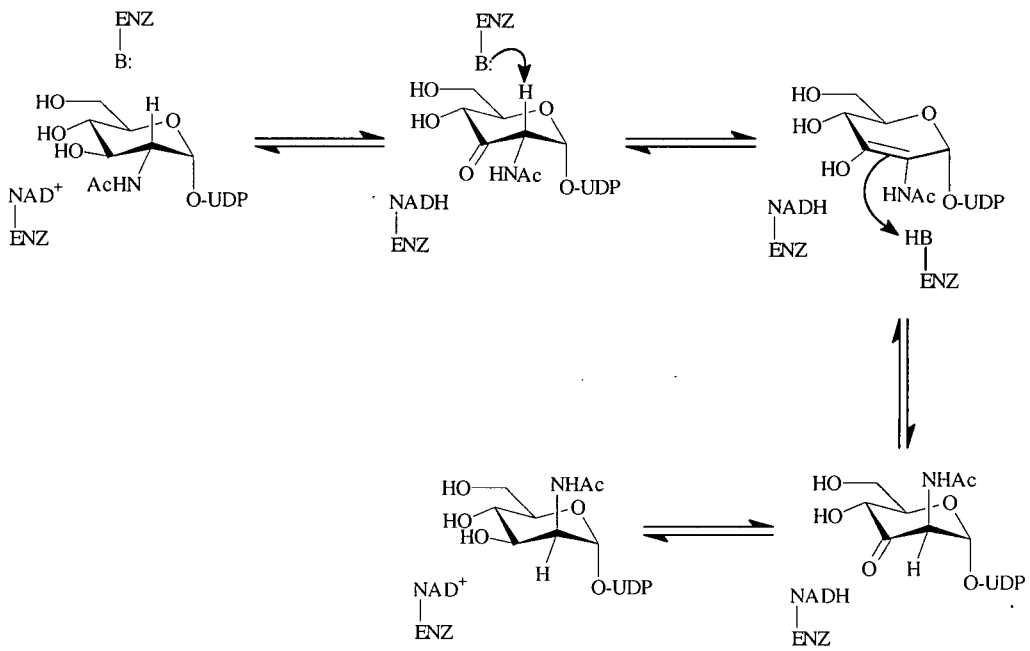


Figure 1.16 Path A: Transient oxidation at C-3" by a tightly-bound active site NAD^+ cofactor.

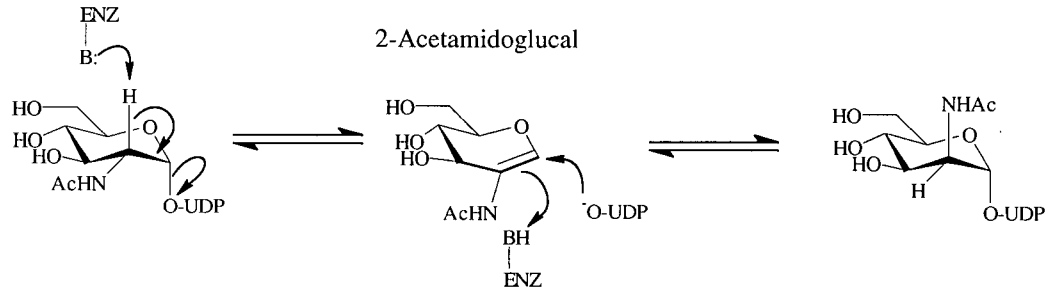


Figure 1.17 Path B: Elimination/addition of UDP.

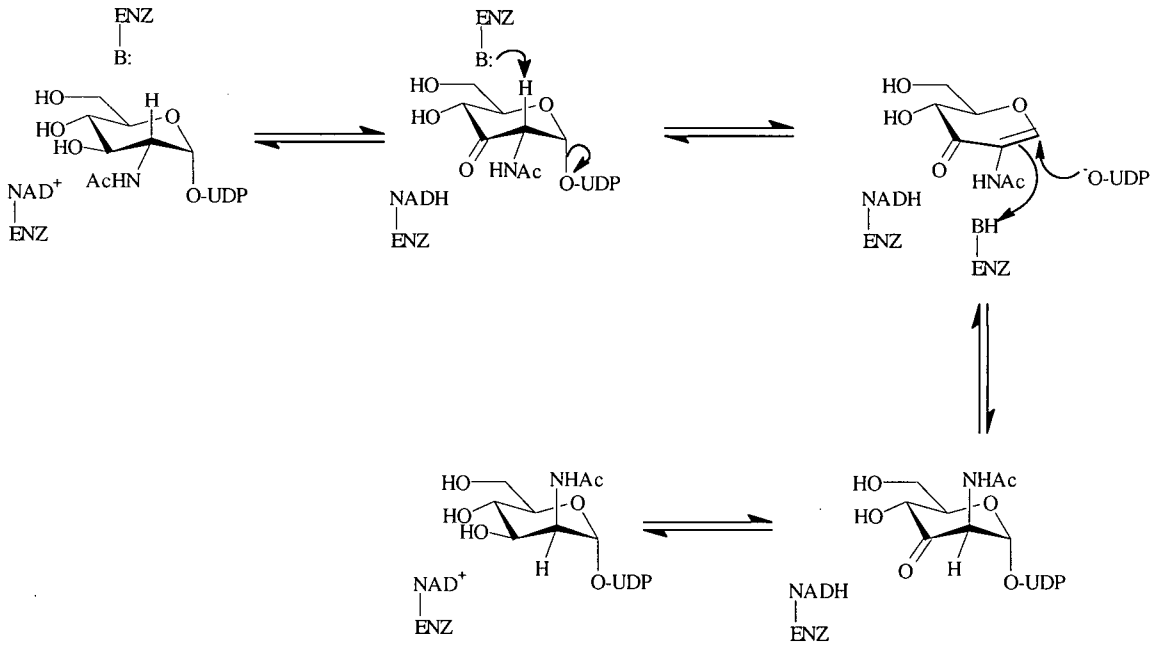


Figure 1.18 Path C: Transient oxidation at C-3", accompanied by elimination and congugate addition of UDP.

A second pathway, which is based on the mechanism proposed for the mammalian enzyme by Sommar and Ellis (Figure 1.15), is presented as Path B (Figure 1.17). This proposed two-step route involves the elimination of UDP from the substrate to form a 2-acetamidoglucal intermediate. Readdition of UDP is accompanied with reprotonation on either face of the intermediate glycal, respectively generating either epimer. In the forward direction of this unusual mechanism, the enzyme would catalyze the *trans*-elimination of UDP in the first step, followed by the *syn*-addition of the nucleotide in the second. In the reverse direction, from UDP-ManNAc to UDP-GlcNAc, a *syn*-elimination would precede the *trans*-addition of UDP.

Path C shows a mechanism that has never been previously considered, arising from a combination of the first two possibilities (Figure 1.18). This hybrid mechanism also has the transient oxidation at C-3" by an active site NAD^+ cofactor, but this might be followed by the subsequent elimination of UDP from the substrate, allowing conjugate readdition of the eliminated UDP.

A fourth potential mechanism might have a nucleophile in the enzyme's active site attacking the β-phosphorus instead, displacing the sugar moiety of UDP-GlcNAc as the open-chain form (Figure 1.19). The C-2 proton of the GlcNAc intermediate is adjacent to an aldehyde functional group, so that deprotonation may now occur to form the enolate-intermediate. Reprotonation on the opposite face, closure of the ring, and the readdition to UDP completes the inversion.

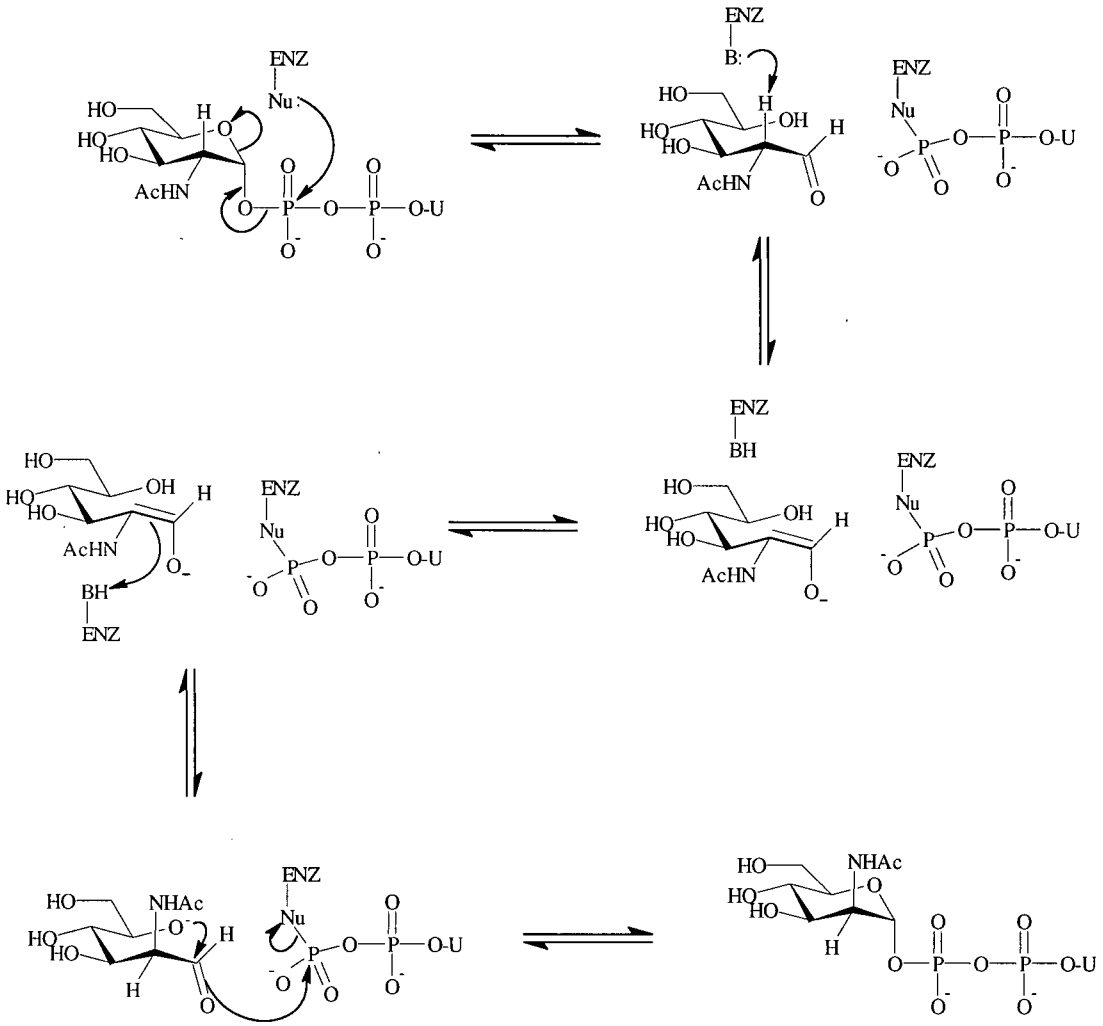


Figure 1.19 Path D: Attack at the β-phosphorus by an active site nucleophile.

One other pathway might be proposed, following the mode of action of UDP-galactose 4-epimerase. It is possible that an active site NAD^+ might oxidize the site of inversion directly. In such a mechanism, the C-2" position of the substrate would be oxidized to an enamine intermediate. Similar to the UDP-galactose 4-epimerase, the sugar moiety might rotate within the active site pocket, in order to present the opposite face of the

intermediate to the cofactor for the subsequent reduction step. However, this mechanism can be discounted at the outset, since prior studies have shown the proton at C-2" to be exchangeable with a solvent proton during the reaction. This observation indicates that a deprotonation/reprotonation mechanism is operable in the 2-epimerase, and not a hydride transfer, since water is not a good hydride source.

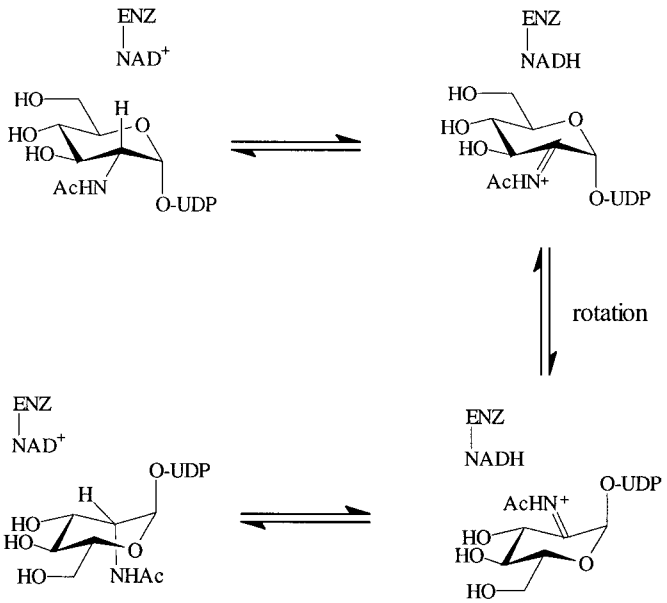


Figure 1.20 A possible direct hydride transfer mechanism.

1.6 Aims of This Thesis

Since UDP-GlcNAc 2-epimerase plays a role in the biosynthesis of a wide variety of bacterial cell surface polysaccharides, this enzyme provides a unique starting point for studies into the function of these recurring structures. One possible approach would be the

design of mechanism-based inhibitors, which may lead to the development of a novel class of antibiotic drugs. Such antibiotics might find application against *Streptococcus pneumoniae* for instance, since UDP-GlcNAc 2-epimerase is implicated in the biosythesis of the pneumococcal cell capsules present in several of the most virulent forms of this deadly bacterium. A first step in this process, however, would require the elucidation of the basic mechanistic mode of action of the enzyme in question.

Prior to the current work, very little was known about UDP-GlcNAc 2-epimerase, especially in terms of its catalytic mechanism. However, due to the nature of the reaction it catalyzes, its mechanism must be unique in comparison to the other carbohydrate epimerases. This enzyme therefore provides an interesting and challenging target for study. This thesis describes experiments designed to explore UDP-GlcNAc 2-epimerase from *E. coli*, and probe it in mechanistic detail.

Chapter Two describes the unambiguous identification of the epimerase gene within the *E. coli* chromosome, the purification of recombinant *E. coli* epimerase, and its subsequent characterization. In order to carry out this latter task, it was necessary to develop a direct continuous spectrophotometric assay for epimerase activity, which required the cloning of a separate enzyme – UDP-ManNAc dehydrogenase. The cloning, overexpression, and purification of this enzyme is discussed in Chapter Three.

The main focus of this thesis is finally broached in Chapter Four in which a series of experiments are detailed, which were carried out in order to distinguish between the four proposed mechanisms outlined earlier in this introduction chapter. Based on the results of these experiments, it was possible to suggest the simplest likely reaction mechanism utilized

by UDP-GlcNAc 2-epimerase. One further experiment, probing the enzyme for a deuterium kinetic isotope effect, and its mechanistic implications is also discussed in that chapter.

Finally, Chapter Five will discuss the synthesis of a proposed mechanism-based inactivator of the enzyme, and the attempts to use the inactivator to uncover the active site base or bases responsible for deprotonation of the substrates during the course of the catalytic reaction.

Chapter Two:

The Purification and Characterization of Recombinant UDP-*N*-Acetylglucosamine 2-Epimerase from *Escherichia coli*

2.1 Introduction

The first true UDP-GlcNAc 2-epimerase was identified in *Escherichia coli*, in 1975, when a partially purified cell extract was determined to catalyze the reversible interconversion of UDP-GlcNAc and UDP-ManNAc (Kawamura *et al.*, 1975). Epimerase activity was also reported in strains of *Bacillus cereus, Bacillus subtilis*, *Bacillus megaterium*, and *Micrococcus lysodeikticus*, each of which was known to have mannosamine–containing polysaccharides as components of their cell walls (Kawamura *et al.*, 1978). Following the initial study, the same group succeeded in characterizing both the partially purified enzyme from *Bacillus cereus* (Kawamura *et al.*, 1978), and the rigorously purified enzyme from *E. coli* (Kawamura *et al.*, 1979; Kawamura *et al.*, 1982). The characteristics and kinetic data for both the *E. coli* and the *B. cereus* epimerases were found to be very similar.

The *E. coli* epimerase was determined to be a dimer with a subunit molecular weight of 38 kDa. The enzyme did indeed catalyze the reversible interconversion of UDP-GlcNAc and UDP-ManNAc, affording an equilibrium ratio of 9:1 for the two epimers respectively. The epimerase has a broad pH optimum, ranging from pH 7 to 9. No catalytic rate enhancement was observed by the addition of NAD^+, NADH, $NADP^+$ or NADPH, indicating that any such cofactor potentially involved in the enzymatic mechanism must be bound tight enough in the active site to be retained through the purification procedure. Also, the enzyme was specific for UDP-GlcNAc and UDP-ManNAc, and was not capable of recognizing any of the other substrates tested, which included *N*-acetylglucosamine, *N*-acetylglucosamine-α-1-phosphate, *N*-acetylglucosamine-6-phosphate, UDP-*N*-acetyl-D-galactosamine, and UDP-D-glucose. In fact, the enzyme is so stringent towards UDP-GlcNAc, UDP-ManNAc itself isn't handled by the enzyme unless UDP-GlcNAc is present as well.

Indeed, one of the most striking features of the bacterial epimerase, is the marked dependence on the presence of UDP-GlcNAc for activity. Extended incubations of relatively large amounts of epimerase with UDP-ManNAc demonstrated that the reaction does not proceed appreciably in the UDP-ManNAc to UDP-GlcNAc direction until a small quantity of UDP-GlcNAc is added to activate the enzyme (Kawamura *et al.*, 1979). Furthermore, activation was not observed with any other substrate, including GlcNAc-α-1-phosphate. A plot of the rate of UDP-GlcNAc epimerization as a function of substrate concentration yielded an apparent K_m of 0.63 mM for UDP-GlcNAc, and produced a sigmoidal curve, which is typical in cases involving allosteric activation of an enzyme by its substrate. A Hill Plot of this data resulted in a straight line, with a Hill coefficient of 2.0, a number whose magnitude reflects the degree of allosteric activation (Fersht, 1985b). Together, these observations were interpreted to mean that the enzyme contains a modulator binding site, distinct from the catalytic site, which is strictly specific for UDP-GlcNAc (Kawamura *et al.*, 1979). One cannot rule out, however, a co-operative model in which the active site of one subunit of the dimeric enzyme must be occupied by the substrate in order for the other subunit to function.

As part of these initial efforts to characterize the bacterial epimerase, the wild type enzyme was purified over 10 000 fold from *E. coli* (Kawamura *et al.*, 1979). The reported multiple step purification scheme (Table 2.1), which included an ammonium sulfate precipitation, and two different affinity chromatography resins, began with 70 g of wet bacterial cells harvested from 25 L of cell culture, to yield a total of 0.24 mg of the purified UDP-GlcNAc 2-epimerase.

Initially, our own study began with attempts to reproduce the purification of the genomic epimerase from natural sources. This proved to be difficult due to the small protein yields, and

the lack of an adequately sensitive assay for the enzyme. In addition, our intended experiments with the epimerase required larger amounts of the enzyme than had been previously isolated. As a result, it was decided that the best course of action would be to obtain larger quantities of the epimerase by cloning and overexpressing its gene.

The first step towards cloning the gene was to uncover the available information in the literature concerning the epimerase. In particular, we had hopes that the epimerase gene had been previously identified, or better still, its DNA sequence had been previously established. Prior knowledge of the DNA sequence encoding the epimerase would greatly facilitate further cloning steps.

Table 2.1. The purification steps for UDP-GlcNAc 2-epimerase from *E. coli* reported by Kawamura *et al.* (1979).

Step	Fraction	Protein (mg)	Total Activity (units)	Specific Activity (units/mg)
1	Crude extract	5050	3.32	0.00065
2	ADP-agarose	4590	2.97	0.00064
3	Protamine/ammonium sulfate	1670	3.50	0.00209
4	1st UDP-agarose	4.42	2.21	0.50
5	DEAE-Sepharose CL-6B	0.96	2.20	2.29
6	2nd UDP-agarose	0.24	1.70	7.08

2.2 Identification of the Epimerase Gene

In *E. coli*, UDP-GlcNAc 2-epimerase is found in the biosynthetic route of ECA (see Chapter 1). There are three clusters of genes in the chromosome of enterobacteria implicated in the biosynthesis of ECA (Makela *et al.*, 1976). These gene clusters are known as *rfb*, *rfe* and *rff*. Mutant bacteria unable to produce ECA were found to have mutations located in the *rff* gene cluster, which abolished epimerase activity (Lew *et al.*, 1978). Subsequently, the *rff* gene cluster was mapped to a position adjacent to another gene cluster (known as *ilv*) whose exact location on the *E. coli* chromosomal map was known (Meier and Mayer, 1985). This result located the ECA biosynthetic genes at a position near 85 minutes on the circular 100-minute *E. coli* chromosomal map. Eventually however, the observation that *E. coli* with specific genetic mutations located between 84 and 85 minutes lacked epimerase activity further narrowed the location of the enzyme's gene (Meier-Dieter *et al.*, 1990).

One minute on the chromosomal map corresponds to a section of the bacterial DNA over 40 000 base pairs in length, which is still a relatively large segment of the chromosome. In a subsequent study, a part of the region between 84 and 85 minutes was further subdivided into smaller fragments, and each fragment was cloned and used to test for the activity of enzymes related to ECA biosynthesis. One particular DNA fragment from this region, 4 900 base pairs in length, was discovered to possess the ability to restore both UDP-GlcNAc 2-epimerase and UDP-ManNAc dehydrogenase activity to a strain of *E. coli* that did not have those enzymes. Therefore, the epimerase gene had been pinpointed within that short stretch of the *E. coli* chromosome (shown as pCA62 in Figure 2.1; Meier-Dieter *et al.*, 1992).

In 1992, preliminary results were reported from an ambitious effort to sequence the entire *E. coli* K12 genome (Daniels *et al.*, 1992). (The sequence of the entire genome was

only recently reported by Blattner *et al.*, 1997.) The report of the initial findings was most fortuitous since it happened to begin in the region most pertinent to our own study. It is entirely coincidental that this first report revealed the sequences of every gene between 84.5 and 85.5 minutes of the *E. coli* chromosome. This region had a particularly relevant series of six consecutive open-reading frames (ORF's – potential protein encoding genes that contain discrete start and stop sequences), only one of which had been previously identified (Figure 2.1). That one identified gene, known as *rfe*, has been postulated to encode UDP-GlcNAc:undecaprenylphosphate GlcNAc-1-phosphate transferase, and is required for ECA biosynthesis (Meier-Dieter *et al.*, 1992). The other five genes were each assigned a simple name based on the number of amino acids that would be hypothetically transcribed into the putative protein expressed from the determined DNA sequence. In the order of transcription,

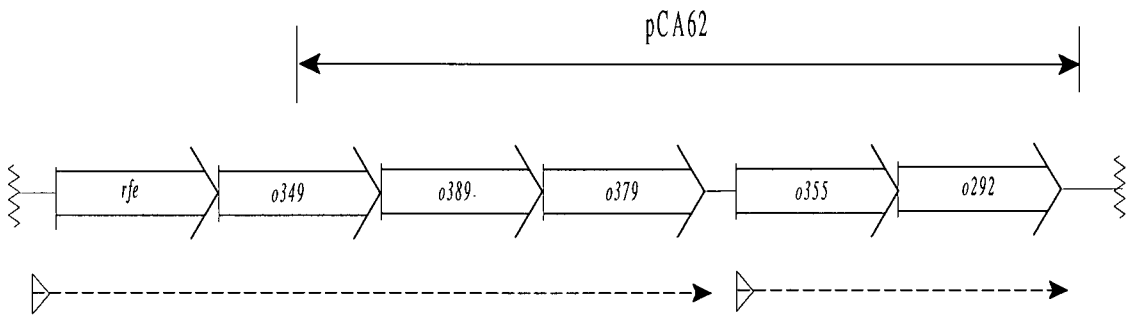


Figure 2.1. The short stretch of the *E. coli* genome encoding five proteins of unknown function, as reported by Daniels *et al.* (Figure adapted from Daniels *et al.*, 1992). The open arrows represent each of the six open reading frames, the dotted arrows represent the mRNA transcriptional units, and the solid double-headed arrow shows the fragment of the genomic *E. coli* DNA cloned into the pCA62 plasmid.

these ORF's are o349, o389, o379, o355 and o292. Since the latter four genes all lie completely within the 4 900 base pair fragment that was known to contain the gene for UDP-GlcNAc 2-epimerase, it was apparent that one of those ORF's encoded the epimerase.

The primary goal of the *E. coli* genome-sequencing project was to determine the sequence and locations of as many of the organism's genes as possible. In order to discover a likely function of each protein encoded by an ORF, the determined DNA sequences for each ORF were compared to all previously known sequences stored in the GenBank and EMBL DNA sequence databases. Based on the results of the sequence comparisons, the genes for UDP-GlcNAc 2-epimerase and UDP-ManNAc dehydrogenase were each tentatively matched to one of the four ORF's in this region. First, o379, reportedly displayed 25.2% sequence similarity with GDP-mannose dehydrogenase from *Pseudomonas aeruginosa*, so it was deemed likely that o379 was in fact the DNA sequence encoding UDP-ManNAc dehydrogenase, whose gene is known as *rffD*. Second, o355, had a few notable similarities with UDP-glucose 4-epimerase from *Streptomyces lividans*, and CDP-paratose 2-epimerase from *Salmonella typhimurium*. As a result, this gene was suggested to be *rffE,* the gene encoding UDP-GlcNAc 2-epimerase. And thus, the sequence for the epimerase was apparently known.

However, this latter assignment was found to be inconsistent with observations by several different groups. Notably, the o355 gene product was actually found to have 71% sequence identity with the *E. coli rfbB* gene product, dTTP-glucose-4,6-dehydratase (Macpherson *et al.*, 1993; Stevenson *et al.*, 1994). It appears that *E. coli* contains two genes at different locations in its genome, *rfbB* and o355, that both encode the dehydratase. Furthermore, another study showed that both o389 and o379 together were required for UDP-

GlcNAc 2-epimerase and UDP-ManNAc dehydrogenase activity (Robertson *et al.*, 1994). This information led us and others to suggest that the gene for the epimerase was actually coded within o389 (Marolda and Valvano, 1995; Sala *et al.*, 1996; Morgan *et al.*, 1997).

In retrospect, this new assignment seems more plausible considering the roles of the epimerase and dehydrogenase, which are both common to ECA biosynthesis. Figure 2.1 shows the order of the ORF's, and below that, the dotted arrows show the messenger RNA (mRNA) transcriptional units that are subsequently used as the template for protein biosynthesis. On the molecular level, biosynthesis is often regulated by controlling the expression of the biosynthetic enzymes. This can be easily accomplished if the cell transcribes all the enzymes required for a common pathway simultaneously, therefore expressing or suppressing them at the same time. The ECA biosynthetic enzymes are very likely to all be present on a single mRNA template after transcription. If o355 were the epimerase gene, as first suggested, the dehydrogenase (o379) and the epimerase would be transcribed onto different mRNA templates. However, o389 and o379 are part of the same transcriptional unit, allowing the organism to mutually express both enzymes during the times in its lifecycle when ECA biosynthesis is required.

Remarkably, in the midst of this confusion, another group was pursuing an unrelated project, studying the genes required for the adsorption of Bacteriophage N4 onto *E. coli.* Their research revealed four genes that were required for Bacteriophage N4 adsorption, which they named *nfrA*, *nfrB*, *nfrC* and *nfrD*. The location for each gene was mapped out on the *E. coli* genome, placing *nfrA* and *nfrB* side-by-side at 12 minutes on the chromosomal map, *nfrD* between 44 and 58 minutes, and *nfrC* at 85 minutes (Kiino and Rothman-Denes, 1989). The *nfrC* gene was then sequenced and unambiguously found to be o389 (Kiino *et al.*, 1993). The

gene was subsequently cloned into a pET11a expression vector and overexpressed. The protein expressed from the *nfrC* gene was reported to be a soluble cytosolic protein, however, due to the misassignment by Daniels *et al.*, no function was suggested for NfrC.

In light of this new information, it was evident that the NfrC protein was not only the epimerase, but it had also already been cloned and overexpressed, which had been our own intention. Upon contacting Dr. L. B. Rothman-Denes, whose laboratory had performed the cloning, she graciously agreed to forward a sample of the pKI86 plasmid containing the cloned epimerase gene.

2.3 Expression and Purification of UDP-*N*-Acetylglucosamine 2-Epimerase

The pKI86 plasmid consists of the *nfrC*-coding region cloned into the cloning/expression region of the pET11a expression vector. The pET11a vector is a small fragment of circular DNA (around 5000 base pairs, as compared to 4 million base pairs for the entire *E. coli* chromosome) designed to provide maximum expression of the protein-encoding gene inserted into a highly specified position of the vector.

The pKI86 plasmid was transformed into *E. coli* JM109(DE3), grown in LB media in the presence of 100 μg/ml ampicillin, and induced with 0.4 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG). As described elsewhere (Kiino *et al.*, 1993), this resulted in high levels of UDP-GlcNAc 2-epimerase production.

The previous purification of genomic wild-type epimerase, reported by Kawamura *et al.* (1979, 1982) included glycerol and dithiothreitol (DTT) in all purification buffers for protein stability. This practice was maintained throughout all purification steps. Furthermore,

all protein manipulations were performed at 4°C, whenever possible. The purification procedure for the overexpressed recombinant epimerase differed significantly from the original reported purification. Due to the high protein production, many of the steps that were necessary for the purification of the small quantity of epimerase present in natural sources were no longer required.

The crude cell extract was prepared by passing the cells twice through a chilled French pressure cell, and centrifuging the resultant extract to remove the ruptured cell wall debris. After centrifugation, the epimerase remained exclusively in the soluble fraction. The crude lysate was partially purified by elution from a DE52 weak anion-exchange column with a linear gradient of 0 to 0.2 M NaCl in Trien-HCl buffer at pH 8.5, containing 10% glycerol and 2mM DTT. This step effectively removed the strong-binding negatively charged impurities from the crude lysate. The partially purified protein was subsequently concentrated, desalted, and purified by anion-exchange HPLC, using a Waters AP-1 Protein-Pak Q 8HR HPLC column, and eluting with a linear gradient of 0 to 0.4 mM NaCl. Amidst the smaller protein peaks, a single large peak eluted from the ion-exchange column at around 0.15 M NaCl. All of the protein fractions were assayed for epimerase activity, and the large peak was determined to solely contain all of the epimerase activity.

Initially, the eluted column fractions were assayed for the epimerase in one of three ways. First, all protein fractions were examined by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Fractions containing the major protein band from crude cell lysate could be readily identified by SDS-PAGE (Figure 2.2). Second, it had been previously established that the enzymatic reaction in D_2O resulted in the solvent-derived incorporation of deuterium at C-2" (Salo, 1976). As discussed in Chapter 4, this observation was reconfirmed with our own

preparation of purified recombinant UDP-GlcNAc 2-epimerase. Analogous to the experiment outlined in section 4.2, it was possible to monitor the enzymatic reaction by ^{1}H NMR spectroscopy by incubating UDP-GlcNAc dissolved in phosphate buffer prepared with D_2O with an aliquot from an HPLC fraction. Only the active epimerase-containing fractions resulted in the incorporation of deuterium at C-2". Third, the ion-pair reversed-phase technique described by Meynial *et al.* (1995) was useful in determining whether or not a particular column fraction was capable of catalyzing the epimerization. This analytical technique, also described more thoroughly in Chapter 4 (section 4.5), is capable of separating the two nucleotide-sugar epimers, UDP-GlcNAc and UDP-ManNAc, on a reversed-phase HPLC column. Eluent fractions were incubated with UDP-GlcNAc, and the resultant sample was injected onto a reversed-phase HPLC column, while the eluent was monitored at 262 nm. Only epimerase-containing fractions produced a second peak in the HPLC trace, one-tenth the size of the first, UDP-GlcNAc peak. This smaller second peak, which elutes just prior to the UDP-GlcNAc, is attributed to an equilibrium concentration of UDP-ManNAc.

Due to the time-scale of the assay techniques, as compared to the rate of the catalytic reaction, these techniques are not practical for the quantitative evaluation of epimerase activity. The development of a coupled enzyme assay provided a more convenient manner for qualitative, and more importantly, quantitative detection of enzyme activity. The details of the coupled assay are described in section 2.5.

The purified protein appeared to be homogeneous when analyzed by SDS-PAGE (Figure 2.2). Electrospray mass spectroscopy showed the subunit mass was consistent with that expected from the gene sequence: calculated, 42 246 Da; found 42 254 $\pm$ 4 Da. The accuracy of this technique provides a far superior measurement for the molecular weight of the

epimerase, when compared to the mass of 38 kDa previously determined using SDS-PAGE (Kawamura *et al.*, 1979).

Using the coupled enzyme assay, the specific activity of the pure epimerase preparations was determined to be 6.8 ± 0.5 units/mg. This value agrees closely with that previously reported for the epimerase purified directly from natural sources (7.08 units/mg; Kawamura *et al.*, 1979, 1982). Therefore, in two steps, the overexpressed recombinant epimerase was purified to yield the same specific activity as that previously reported.

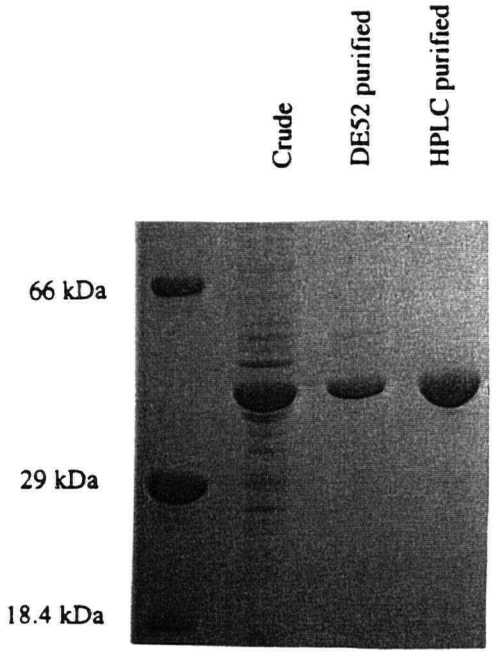


Figure 2.2 SDS-polyacrylamide gel of crude recombinant UDP-GlcNAc 2-epimerase, epimerase purified by DE52 anion exchange chromatography, and epimerase purified by a Waters AP-1 Protein-Pak Q 8HR HPLC column. The protein components of the crude and purified samples were separated by 12% SDS-PAGE. The molecular weight standards are: 66 kDa, bovine serum albumin; 29 kDa, carbonic anhydrase; and 18.4 kDa, β-lactoglobulin.

2.4 Identification of the Epimerase Reaction Products

It was necessary to confirm that the expressed protein encoded by o389 (NfrC) was in fact the epimerase (RffE), and could catalyze the formation of UDP-ManNAc from UDP-GlcNAc. Since UDP-ManNAc is not commercially available, GlcNAc and ManNAc standards were used instead to identify the reaction products of the purified UDP-GlcNAc 2-epimerase. An enzymatically equilibrated mixture of UDP-GlcNAc and UDP-ManNAc was hydrolysed under acidic conditions that were known to cleave the glycosidic bond, but not cause further epimerization of the hydrolyzed sugars (Kawamura *et al.*, 1978). The resulting hydrolysate was applied to an HPLC column capable of separating GlcNAc and ManNAc (Lai and Withers, 1994). Two peaks were eluted from the column in the approximate equilibrium ratio of 10:1 as expected for the two sugar epimers. The two peaks had retention times identical to authentic GlcNAc and ManNAc standards, respectively. This experiment confirms that *nfrC*, o389 and *rffE* are all the same gene. Furthermore, the overexpressed, purified protein from this gene does indeed catalyze the reversible conversion of UDP-GlcNAc and UDP-ManNAc.

2.5 Kinetic Characterization of UDP-*N*-Acetylglucosamine 2-Epimerase

In order to perform the kinetic characterization of the epimerase, it was necessary to develop a continuous spectrophotometric enzyme assay. Since many enzyme activities cannot be directly detected spectrophotometrically and quantified, a common technique is to couple the reaction of the enzyme of interest with a second, auxiliary, enzyme. In the coupled assay,

the auxiliary enzyme consumes the product from the primary enzyme, and quantitatively generates a spectrophotometrically measurable compound (P in Figure 2.3).

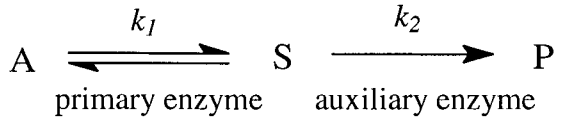


Figure 2.3 The general scheme for a coupled enzyme assay.

This scheme is feasible as long as the two enzymes operate under compatible conditions. If, for instance, the pH optima for the two enzymes differ significantly, or if one of the co-substrates required for the auxiliary enzyme inhibits the primary enzyme, then a continuous assay is not possible. In that case, it would be necessary to run the assay in two steps. However, if none of the conditions for either enzyme are detrimental to the other, then both stages of the assay may be carried out simultaneously.

It is possible to measure the rate of the primary enzyme indirectly by monitoring the rate of formation of product P. To do so, two criteria are necessary to ensure that the overall rate is independent of the auxiliary enzyme concentration (McClure, 1969; Segel, 1974). First, the primary reaction must be zero-order with respect to [A]. This condition is met if only a small fraction of the initial amount of substrate is catalytically turned-over during the assay period. Second, the velocity of the auxiliary enzyme must be first-order with respect to [S], and must catalyze an effectively irreversible reaction. This condition is achieved if the steady-state concentration of S is significantly lower than the K_m of S with respect to the auxiliary

enzyme. That is to say, sufficient amounts of the auxiliary enzyme must be present to turn over S rapidly enough to maintain an appropriately low steady-state concentration of S.

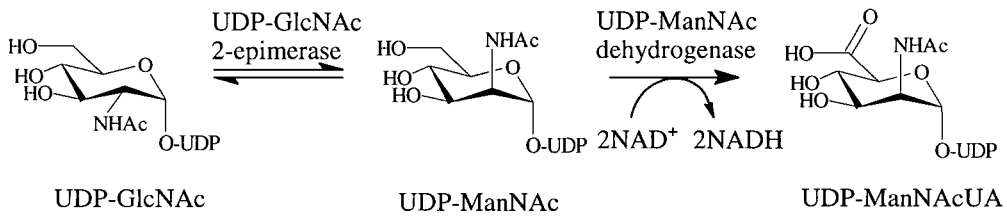


Figure 2.4 A coupled spectrophotometric assay for UDP-GlcNAc 2-epimerase.

The auxiliary enzyme used in our assay is the enzyme that follows UDP-GlcNAc 2-epimerase in the biosynthesis of ECA, UDP-ManNAc dehydrogenase. UDP-ManNAc dehydrogenase catalyzes the irreversible two-fold oxidation of UDP-ManNAc to UDP-ManNAcUA with the production of two equivalents of NADH (Figure 2.4). This enzyme would therefore allow us to follow the reaction continuously while monitoring the increasing concentration of NADH, at 340 nm. This enzyme has previously been characterized, and its kinetic constants (measured at pH 10) have been reported (Kawamura *et al.*, 1979). In that study, the dehydrogenase was determined to have an apparent K_m of 0.38 mM for UDP-ManNAc, and an apparent K_m for NAD^+ (measured in the presence of 0.6 mM UDP-ManNAc) of 0.21 mM. Both products of the enzyme were discovered to be competitive inhibitors, displaying a K_i of 0.23 mM for NADH, and a K_i of 0.75 mM for UDP-ManNAcUA. In contrast to observations with the epimerase, UDP was surprisingly not an inhibitor of UDP-

GlcNAc dehydrogenase (at a concentration of 1 mM). The optimum pH for UDP-ManNAc dehydrogenase was determined to lie between 9.5 and 10. However, their results showed that the dehydrogenase still retains approximately 75% of its activity at pH 8.8. This latter pH value is near the upper limit of the optimum pH range for UDP-GlcNAc 2-epimerase, and was the pH selected for the assay used to monitor the epimerase activity in the earlier reported study (Kawamura *et al.*, 1978, 1979). A pH of 8.8 was also employed for our own studies with UDP-GlcNAc 2-epimerase.

The reported product inhibition of UDP-ManNAc dehydrogenase poses a potential problem in the design of a coupled assay system. If the two products of the dehydrogenase reaction, UDP-ManNAcUA and NADH, were allowed to accumulate during the time of the assay, the auxiliary enzyme would be increasingly inhibited over time, potentially invalidating the required criteria for the coupled assay. This can be overcome by measuring only the initial velocity kinetics, before significant concentrations of the products are generated.

The next question that arises concerns the quantity of each enzyme required. Adequate amounts of UDP-GlcNAc 2-epimerase are required to generate a measurable rate for the assay. Consequently a proportional excess of UDP-ManNAc dehydrogenase is necessary to maintain a condition where the overall reaction rate depends only on the concentration of the epimerase. For the epimerase, it was empirically determined that approximately 8×10^{-3} units/mL of enzyme gave a reasonably measurable rate. The relative amount of dehydrogenase needed for the assay is less trivial to determine, and more crucial to the validity of the assay.

In the valid coupled assay under the conditions defined above, the velocity of the auxiliary enzyme, v_2, is given by the equation:

$$v_2 = k_2[S]_{ss}[E_{aux}]$$

where k_2 is the first-order rate constant of the reaction catalyzed by the auxiliary enzyme, $[S]_{ss}$ is the steady state concentration of the substrate S, and $[E_{aux}]$ is the concentration of the auxiliary enzyme (McClure, 1969; Segel, 1974). Under valid assay conditions, doubling the amount of the primary enzyme, will double the steady-state concentration of S, and therefore double v_2. Conversely, doubling $[E_{aux}]$ would subsequently halve the $[S]_{ss}$, and the net result would leave v_2 unchanged. Using this relationship, it was determined through trial and error that 8.3 mg/mL of dehydrogenase was required to ensure that the kinetics of the coupled assay were valid.

This large amount of dehydrogenase required to ensure a properly coupled assay can be explained by the kinetic constants measured for the dehydrogenase at pH 8.8. (For the characterization of UDP-ManNAc dehydrogenase refer to Chapter Three.) The kinetics of UDP-ManNAc dehydrogenase, previously determined at pH 10, showed that the enzyme at its optimum pH has a relatively high apparent K_m for UDP-ManNAc (0.38 mM), and a low k_{cat} of 5.3 s^{-1} (Kawamura *et al.*, 1979). At pH 8.8, we determined the apparent K_m to be 1.2 ± 0.4 mM and the k_{cat} to be 1.2 ± 0.2 s^{-1}. The difference between the two sets of reported values for the dehydrogenase can be easily attributed to the change in pH, which would be expected to alter the binding and catalytic properties of the enzyme. Overall, the dehydrogenase, which is relatively inefficient under its own optimal conditions, is evidently even less efficient under the assay conditions that are optimal for epimerase activity. In order to maintain the stipulation that the steady-state concentration of UDP-ManNAc must be much less than the K_m value of the dehydrogenase for that substrate, a surprisingly large amount of dehydrogenase is

required. Thus, it was absolutely necessary to obtain the large quantities of dehydrogenase inevitably required for use in the coupled enzyme assay in an efficient and inexpensive manner. We were therefore required to clone and overexpress the gene for this enzyme. The details of the cloning, overexpression, purification, and characterization of the dehydrogenase are discussed in Chapter Three.

Once the dehydrogenase was successfully cloned, we had a steady source of UDP-ManNAc dehydrogenase, and it was possible to determine the kinetic constants for UDP-GlcNAc 2-epimerase. The initial velocity kinetics of UDP-GlcNAc 2-epimerase were determined by plotting the rate of epimerization as a function of the initial UDP-GlcNAc concentration. In a typical experiment, each cuvette contained 5.8×10^{-3} units of epimerase, 5 mg of dehydrogenase (6.0 units at pH 8.8), 4 mM NAD^+ and variable concentrations of UDP-GlcNAc (0.07 mM to 3.1 mM; 600 μL total volume). The reactions were allowed to proceed until 5% of the initial UDP-GlcNAc was consumed. For the larger initial UDP-GlcNAc concentrations, this amounted to the eventual production of 0.12 mM NADH, which is half of the K_i of NADH with respect to the dehydrogenase. At the lower initial UDP-GlcNAc concentrations, this was significantly less.

Coupled enzyme assays typically display a brief lag phase, lasting a few seconds, before the steady-state concentration of S is reached (MacClure, 1969; Segel, 1974; Rudolph *et al.*, 1979). After the non-linear lag phase, the observed rate is linear, and proportional to the rate of the primary enzyme. The lag is independent of the primary enzyme present in the assay, but dependent on the amount of auxiliary enzyme and the K_m of S for the auxiliary enzyme. As a result of the high K_m value of the dehydrogenase for UDP-ManNAc, a lag of several minutes could be observed during the assay when lower dehydrogenase concentrations

were employed. Initial velocity conditions are typically assumed to persist during the period of time in which the substrate concentration is within 10% of its initial value (Allison and Purich, 1979). Therefore, in order to obtain a valid rate measurement with a coupled enzyme assay, it is necessary to ensure that the lag is minimized, otherwise, if the lag phase is too long, the steady-state concentration of S might be only reached after the initial velocity conditions are no longer valid. The lag phase can be eliminated by adding sufficient auxiliary enzyme, but in the case of the inefficient UDP-ManNAc dehydrogenase, which has a relatively poor specific activity (1.6 ± 0.2 units/mg) at pH 8.8, the amount of auxiliary enzyme required was exceedingly high. Under the optimal coupled assay conditions used to measure epimerase activity however, the high concentrations of dehydrogenase successfully eliminated the lag.

Consequently, the high concentrations of coupling enzyme required resulted in a noticeable background rate in NAD^+ reduction caused by the impurities in the dehydrogenase preparation. This background, however, was determined to be independent of substrate concentration. As a result, the background, which typically amounted to 15% of the V_{max}, was subtracted from the final rate.

A direct plot of epimerase rate versus substrate concentration gave a sigmoidal curve with a Hill coefficient of 2.3 ± 0.2, and an apparent K_m value of 0.73 ± 0.09 mM (Figure 2.5). These results are similar to those previously obtained with the enzyme purified from natural sources (Kawamura *et al.*, 1979). A k_{cat} value of 4.8 ± 0.2 s^{-1} was also determined, based on the assumption that each subunit of the dimeric epimerase contributes one active site, as opposed to a single active site formed at the dimer interface between the two converging subunits. The error reported with the data is indicative of the deviation of the data from the curve-of-best-fit. However, it was found that the values determined from four independent

data sets, each performed with different enzyme preparations gave results that differed by as much as 20%. Therefore, the error on the values should be considered slightly larger than reported.

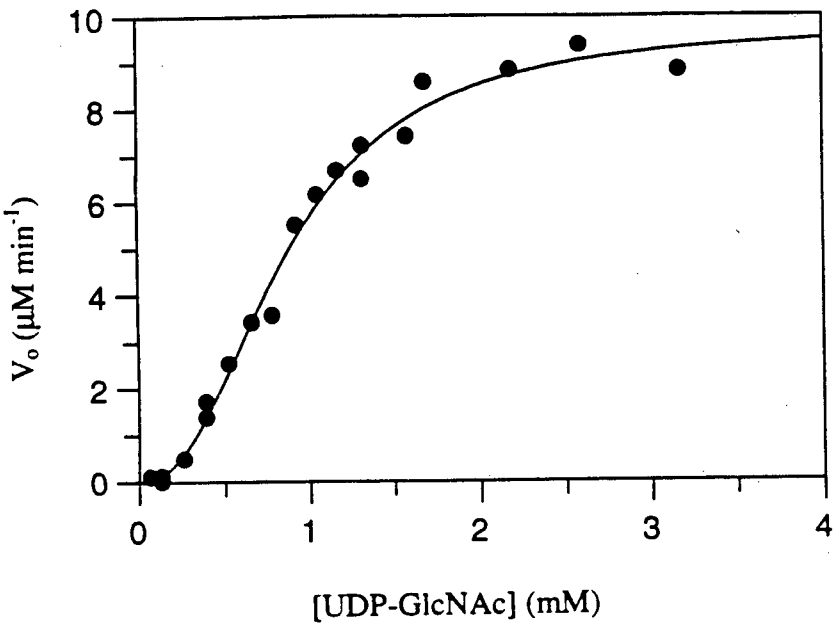


Figure 2.5 A plot of the initial velocity of UDP-GlcNAc epimerization as a function of UDP-GlcNAc concentration.

2.6 Conclusions

A literature search aiming to uncover the identity of the gene for UDP-GlcNAc 2-epimerase revealed that it had been previously misidentified. It was also discovered that the new gene suspected of encoding the epimerase had previously been cloned into a pET11a

expression vector, resulting in high levels of expression of the resultant protein. We have purified the protein to homogeneity, and confirmed it to catalyze the reversible interconversion of UDP-GlcNAc and UDP-ManNAc. Therefore, the gene encoding UDP-GlcNAc 2-epimerase has now been positively identified.

The initial velocity kinetics of the reaction catalyzed by the recombinant UDP-GlcNAc 2-epimerase were measured by employing a coupled enzyme assay that uses UDP-ManNAc dehydrogenase as the auxiliary enzyme. The data gave a sigmoidal curve with a Hill coefficient of 2.3 ± 0.2, an apparent K_m of 0.73 ± 0.09 mM and a k_{cat} of 4.8 ± 0.2 s^{-1}. These results were in agreement with the similar kinetic constants and positive cooperativity observed previously with epimerase isolated from natural sources. However, due to the large amounts of recombinant epimerase available, as compared to the quantities of enzyme present naturally within *E. coli*, it was possible to purify the recombinant epimerase to the same specific activity as the previous study, but in only two steps. The large quantities of epimerase available make it possible to perform detailed studies in order to investigate the unique mechanism of this enzyme.

2.7 Experimental Methods

2.7.1 Chemicals and Enzyme Substrates

UDP-GlcNAc and NAD^+ used in the determination of UDP-GlcNAc 2-epimerase kinetic constants were obtained from Sigma Chemical Company. All other buffers and chemicals were obtained from Sigma Chemical Company, Boehringer Mannheim

Biochemicals, or Fisher Scientific Company, Inc., unless otherwise stated. UDP-ManNAc dehydrogenase was prepared as described in Chapter 3.

2.7.2 Strains, Media and Plasmids

The pKI86 plasmid, which consists of the *nfrC/rffE* gene cloned into a pET11a expression vector, was a gift from Drs. Lucia B. Rothman-Denes and Diane R. Kiino, Department of Molecular Genetics and Cell Biology, University of Chicago. *E. coli* JM109(DE3) (Promega) was used as the host organism for pKI86 during the overexpression of UDP-GlcNAc 2-epimerase. Cells were grown in LB media (Difco) supplemented with ampicillin (100 μg/mL). Unless otherwise stated, general microbiology techniques are as those described by Ausubel *et al.* (1992).

2.7.3 Protein Determination

Protein concentrations were determined by the method of Bradford (1976), using bovine serum albumin (Bio-Rad) as standard.

2.7.4 Definition of a Unit

One unit of UDP-GlcNAc 2-epimerase is defined as the amount of epimerase that produces 1 μmol of UDP-ManNAc per minute under the standard coupled assay conditions (37°C, pH 8.8), and in the presence of 4 mM UDP-GlcNAc. Under these conditions, the specific activity of UDP-GlcNAc 2-epimerase was determined to be 6.8 ± 0.5 units/mg.

For the purposes of this study, one unit of UDP-ManNAc dehydrogenase is defined as the amount of dehydrogenase that produces 1 μmol of UDP-ManNAcUA per minute (or 2

μmoles of NADH) under the same standard conditions (pH 8.8). These conditions are the relevant conditions for this study, but are not the optimal conditions for UDP-ManNAc dehydrogenase. Previous studies (Kawamura *et al.*, 1979) have determined the value of one unit at the optimal pH of 10. The optimally determined specific activity (pH 10) was reported as 7.58 units/mg, whereas our determination at pH 8.8 was 1.6 ± 0.2 units/mg.

2.7.5 Purification of UDP-*N*-Acetylglucosamine 2-Epimerase

The pKI86 plasmid was used to transform $CaCl_2$-competent *E. coli* JM109(DE3). A colony from the transformation was used to inoculate 2 L of Luria Bertani (LB) media supplemented with 100 μg/mL ampicillin. LB media contains 10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl. The cell culture was grown at 37°C to an optical density (600 nm) between 0.8 and 1.2, at which point high-level overexpression of UDP-GlcNAc 2-epimerase was induced by the addition of IPTG to a final concentration of 0.4 mM. After three hours of further growth, the cells were harvested by centrifugation, and the cell pellet (2 g) was stored at –78°C. The pellet was later thawed rapidly with warm water, and the cells were resuspended in 50 mL of cold 50 mM Tris-HOAc buffer, pH 7.9 containing 2 mM DTT, 1 mM EDTA (disodium salt), 1 mM phenylmethylsulfonyl fluoride, 1 μg/mL pepstatin, and 5 μg/ml aprotonin (buffer A). The resuspended cells were passed twice through a chilled French pressure cell (SLM Aminco), rupturing the cells at 20 000 psi.

The fresh crude cell extract was centrifuged for 20 minutes at 6000 rpm (Sorvall GSA rotor). The supernatant was diluted with buffer B (50mM Trien-HCl, pH 8.5, 2mM DTT, 10% glycerol) until the conductivity of the solution was approximately the same as that of the stock buffer B, as determined with a VWR expanded-range digital conductivity meter. Generally,

this required a two-fold dilution of the cell lysate. The diluted crude cell extract was applied to a 50 ml column of DE52 pre-equilibrated in buffer B, and the column was washed with 100 ml of Buffer B, then eluted with a linear gradient of 0 to 0.4 M NaCl (400 mL over 180 minutes at 2.2 mL/min) in buffer B. The eluted protein was monitored at 280 nm, using a Spectrum Spectra/Chrom Flow Thru UV monitor. The fractions were assayed for protein content by SDS-PAGE or assayed for epimerase activity. The epimerase fractions were pooled, concentrated using Amicon centriprep-10 and centricon-10 concentrators, and desalted by repeated dilution and concentration with buffer B.

The concentrated, DE52-purified protein was applied to a Waters AP-1 Protein-Pak Q 8HR HPLC column (10 X 100 mm) that had been pre-equilibrated with degassed buffer B (at 22°C). The HPLC system used consisted of a Waters 625 LC system, monitored at 280 nm using a Waters 486 tunable absorbance detector. The column was washed with 20 mL of buffer B and eluted with a linear gradient of NaCl (0 to 0.4 M NaCl over 40 min at 1.0 mL/min) in buffer B. The epimerase eluted at a salt concentration of 0.15 M. The epimerase fractions, which comprised 75% of the injected protein, were pooled, concentrated and desalted, then frozen in liquid N_2 for storage at –75°C.

2.7.6 Purity Assessment of UDP-*N*-Acetylglucosamine 2-Epimerase

The purity of the epimerase was assessed by 12% SDS-PAGE, using a Bio-Rad Mini-PROTEAN II electrophoresis system. Protein bands were visualized by staining with Coomassie Brilliant Blue. The epimerase band was compared to β-lactoglobulin (19 kDa) carbonic anhydrase (29 kDa) and BSA (66 kDa) as molecular weight standards.

2.7.7 Molecular Weight Determination of UDP-*N*-Acetylglucosamine 2-Epimerase

The molecular weight of the epimerase was determined by electrospray ionization mass spectrometry, performed by Shouming He in the lab of Dr. Stephen Withers, Department of Chemistry, UBC. Electrospray ionization experiments used an HPLC-ESMS setup consisting of a microbore HPLC (Michrom UMA) connected in-line to a PE-SCIEX API 300 MS as described by Hess *et al.* (1993). Intact protein samples (10 μL, ≥1μg/μL) were injected into a microbore PLRP-S reversed-phase column (1 X 50 mm) and eluted with a linear gradient of 16 to 80% acetonitrile in water (containing 0.05% trifluoroacetic acid) over three minutes, and maintained at 80% acetonitrile for an additional seven minutes. The eluent was introduced into the spectrometer, which was operated in the single quadrupole mode. Protein molecular weights were determined from the mass spectroscopic data using deconvolution software supplied by SCIEX.

2.7.8 Identification of the Epimerase Reaction Products

The enzymatic products (3 mM total concentration of the combined epimers) were hydrolyzed in 0.06 N HCl at 100°C for 15 minutes, according to the conditions described in Kawamura *et al.* (1978). The resulting free sugars were analyzed using a Bio-Rad Aminex HPX-87H column (300 X 7.8 mm), eluting with 13 mM H_2SO_4 at a flow rate of 0.5 mL/min, while monitoring the eluent at 190 nm (Lai and Withers, 1994). The retention times of the two eluent peaks were compared to authentic GlcNAc and ManNAc standards. Control samples confirmed that the conditions of the analysis did not cause non-enzymatic epimerization.

2.7.9 Coupled Enzyme Assay

The kinetic parameters of the purified recombinant epimerase were determined in the UDP-GlcNAc to UDP-ManNAc direction with a continuous coupled spectrophotometric assay that employs UDP-ManNAc dehydrogenase as the auxiliary enzyme. Initial rates were measured at 37°C by following the increase in absorbance at 340 nm (NADH formation, ε_{340} = 6220 $M^{-1}cm^{-1}$) using a Varian Cary 3E spectrophotometer. The assay mixtures contained 50 mM Tris-HCl (pH 8.8), 2 mM DTT, $5.8x10^{-3}$ units of epimerase ($8.5x10^{-4}$ mg), 8.5 units of UDP-ManNAc dehydrogenase (5 mg of protein in each cuvette, 1.6 units/mg determined under the assay conditions at pH 8.8), 3.8 mM NAD^+, and a variable concentration of UDP-GlcNAc (0.07 mM to 3.1 mM). Epimerization was initiated by the addition of UDP-GlcNAc 2-epimerase. Prior to the addition of epimerase, a background rate was observed, and attributed to minor impurities in the dehydrogenase preparation, amplified by the requirement for such large quantities of auxiliary enzyme. This background rate was determined to be independent of UDP-GlcNAc concentration. The kinetic rates were therefore determined before and after the addition of UDP-GlcNAc 2-epimerase, and the background rates, which typically amounted to 15% of V_{max}, were subtracted from the final rate. The adjusted rate data were subsequently divided by a factor of two, to account for the stoichiometry of the dehydrogenase reaction. The resulting calculated enzyme rate was plotted as a function of substrate concentration, and the kinetic parameters were determined by a direct fit of the data to a Hill equation:

$$\frac{v}{V_{max}} = \frac{[S]^n}{K' + [S]^n}$$

using the computer program Grafit (Erithacus Software, 1994). The program performs a non-linear regression analysis on the data following the method of Marquart (1963), and reports an error for the data based on the deviation of the data from the calculated curve-of-best-fit. This is the value reported with the data.

Chapter Three:

Subcloning, Overexpression, Purification and Characterization of Recombinant UDP-*N*-Acetylmannosamine Dehydrogenase from *Escherichia coli*

3.1 Introduction

DNA is the molecule of heredity. It alone contains the entire information required for all the functions of a cell's life. As a result, a great deal of an organism's metabolic machinery and energy is focused on the accurate reproduction of DNA and the expression of the information stored within. These processes are known in considerable detail, and many of their elegant subtleties have been mimicked and exploited, to serve as the tools of the modern molecular biologist. One organism in particular, the bacterium *Escherichia coli*, has been thoroughly studied. At the time cloning techniques were beginning to be developed, it was the most understood of all organisms in terms of the biochemistry of its life processes. As a result, this bacterium has found a prominent role in molecular biology.

The *E. coli* genome is a single circular piece of double-stranded DNA, around 4 million base pairs in length. The long fibrous DNA molecule is subdivided into genes, each of which encodes a single component necessary for a specific cellular function. During cellular reproduction, *E. coli* uses a DNA polymerase to make an exact copy of the DNA for its progeny. This enzyme begins at a highly specific site on the DNA strand, called the origin of replication, and copies one strand of the DNA, base pair by base pair, as it travels around the circumference of the circular DNA. In order to catalyze the formation of a new DNA copy, DNA polymerase absolutely requires (1) a DNA template to copy, (2) all four nucleotide base precursors (dATP, dCTP, dGTP, and dTTP) to incorporate into the new DNA chain, (3) a divalent Mg^{2+} for activity, and (4) a short pre-existing DNA chain in order to initiate the DNA chain elongation reaction (Stryer, 1988). This short oligonucleotide chain is known as a primer.

After reproduction has been completed, the cell must begin a new lifecycle, which will require the expression of different genes at various stages. The process of gene expression is divided in two parts: transcription and translation. Even though the sequence of DNA bases in a gene directly codes the amino acid sequence of a protein, the DNA strand itself does not act as the actual template during protein biosynthesis. During the process of transcription, RNA polymerase first forms an intermediary messenger RNA (mRNA) template. The mRNA molecules are copies of short regions of the DNA strand, but with ribose replacing the normal deoxyribose found in the DNA backbone and uracil replacing thymine as a nucleotide base. Each mRNA usually contains the sequence for only a single gene, or a small cluster of several related genes. The protein synthesizing ribosomes then use the mRNA molecule as the direct template for a new protein. During this second, translation step, the information encoded by the sequence of the four nucleotide bases is converted into the larger expanded vocabulary of the twenty natural amino acids that are the building blocks used to construct the proteins.

Often, protein expression is controlled by suppressing the transcription of individual genes, and preventing the formation of new mRNA templates. To accomplish this task, the gene, or a set of biosynthetically related genes are often part of a co-ordinated genetic unit known as an operon. An operon consists of the structural protein genes and all of the additional genes required for the regulatory system that controls the transcription of the structural genes. One of the most studied operons is the *lac* operon, whose three structural genes are required for the transport and breakdown of lactose (Figure 3.1). During transcription, RNA polymerase recognizes and binds to a highly specific DNA sequence known as the *lac* promoter. The promoter indicates to the polymerase the starting point from which to begin mRNA transcription. From the initiation point, a new mRNA copy is dictated

from the DNA strand, and includes the three *lac*Z, *lac*Y and *lac*A structural genes. The mRNA copy that is produced is then used as the template for the translation of *lacZ*, *lacY,* and *lacA* into their respective proteins. However, in lactose-deficient times, it is energetically wasteful to express the unneeded lactose metabolizing proteins. As a result, the *lac* operon also contains the gene for another protein known as the *lac* repressor (la*cI* in Figure 3.1).

When the *lac* repressor protein binds to the operator sequence in the operon, it prevents RNA polymerase from binding to the promoter, and therefore simultaneously blocks the transcription of all three *lac* proteins. However, when lactose is present in the cell, the *lac* repressor will bind a lactose molecule instead of the operator sequence, thereby freeing the blocked promoter, allowing mRNA transcription to occur, and therefore permitting the subsequent expression of the lactose recognizing enzymes. When the enzymes have once again depleted the cell of lactose, the repressor again binds to the operator sequence, and expression of the three *lac* genes is once again halted. Therefore, the production of the lactose handling genes is only triggered when those enzymes are required.

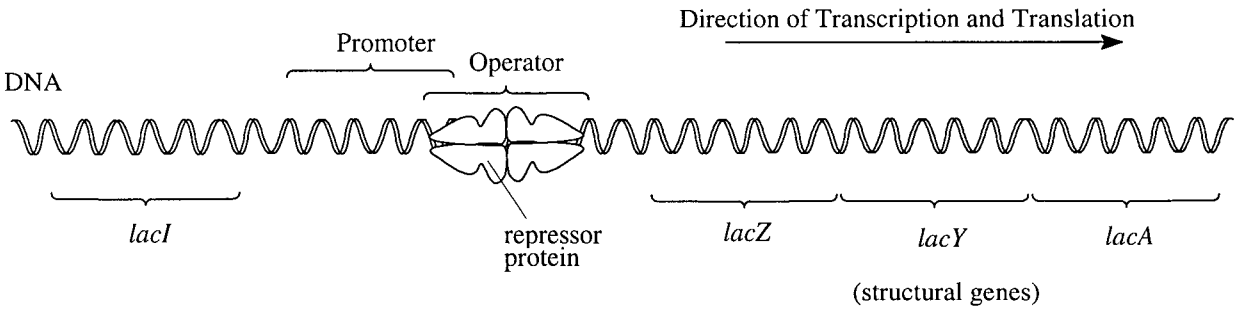


Figure 3.1 Schematic representation of the *lac* operon during times the cell is deficient in lactose.

These naturally occurring features of DNA replication and the regulation of protein expression have been incorporated into the cloning vehicles that are used to overproduce proteins in host organisms.

3.2 The pET11a Vector

Our goal in cloning the UDP-ManNAc dehydrogenase gene (which is called *rffD* and had been previously identified as *o379* in Figure 2.1) was to obtain an inexpensive, convenient and abundant source of the auxiliary enzyme required for the coupled spectrophotometric assay that was used to monitor UDP-GlcNAc 2-epimerase activity (described in Chapter Two). Previously, UDP-ManNAc dehydrogenase had been purified from natural sources, with 25 L of *E. coli* cell culture yielding 1.9 mg of dehydrogenase after a four step purification (Kawamura *et al.*, 1979). Due to the amount of dehydrogenase required for each run with the coupled assay, this recovered amount of enzyme was insufficient for even a single experiment. As a result, it was desired to clone *rffD* and overproduce UDP-ManNAc dehydrogenase in *E. coli*.

It is possible to introduce DNA encoding a protein into a suitable bacterial host, such as *E. coli*, to cause the protein in question to be expressed by the host at higher levels than a typical endogenous protein. Theoretically, through such manipulations, it is possible for a bacterial cell to produce over 50% of its total manufactured protein as a single overexpressed protein (Novagen, 1997 catalogue).

A variety of cloning vehicles, or vectors, have been developed for this task, and are commercially available. These commercial vectors are small circular double-stranded DNA

molecules that have been highly engineered to possess the genetic features necessary for maximal protein expression. While the *E. coli* genome is over four million nucleotide base pairs in length, vectors are only several thousand base pairs in total, and as a result, they are easier to purify and manipulate than the larger, unwieldy genomic DNA. Their small size also allows them, under appropriate conditions, to penetrate the walls of living cells. This latter feature makes it possible to infect a host cell with suitably prepared vector DNA.

Not all vectors are used for protein overexpression. Regardless of their intended use however, commercial vectors all have three common features necessary for their utility. First, each has its own origin of replication. This is required to permit a host organism that has been infected with vector DNA, to use its DNA polymerase to copy the foreign vector as if it were its own native DNA. This ensures that the host cell is capable of providing copies of the invasive vector to its progeny. Therefore, as the cells within a bacterial culture multiply in number, so too will the number of copies of the vector. This provides an inexpensive method of amplifying the vector DNA.

Second, vectors confer resistance to antibiotics. The gene for an enzyme responsible for degrading a particular antibiotic is included in the vector sequence, which allows cells containing the vector to survive in media containing that antibiotic. This provides a facile screening method for cells containing the vector, since all other cells – those that don't have the vector, and therefore the gene for resistance – cannot survive in the presence of the antibiotic.

Third, vectors are highly engineered with respect to the sequences recognized by restriction endonucleases. Restriction endonucleases are enzymes that recognize DNA at highly specific palindromic nucleotide sequences, and cleave both strands of the double-

stranded DNA at that site (Figure 3.2). These enzymes have been discovered in many prokaryotes, and it is thought that they likely serve a defensive role in their natural context. They are probably responsible for degrading DNA alien to the organism. (Standard bacterial strains used in the laboratory for molecular biological manipulations have had their own native restriction endonucleases disabled.) Nowadays, a wide library of these enzymes is commercially available, and these are employed as the highly specific cutting tools used for genetic engineering. As a result, the commercially available vectors have been altered so that they are cut by several common restriction endonucleases only once, and at a single desired position. This was done by inserting new restriction sites, and removing redundant restriction sites from the commercial vector DNA.

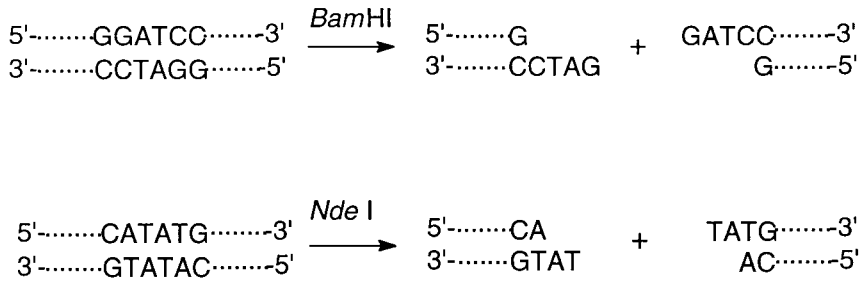


Figure 3.2 The DNA cleavage reactions catalyzed by the restriction endonucleases, *Nde*I and *Bam*HI.

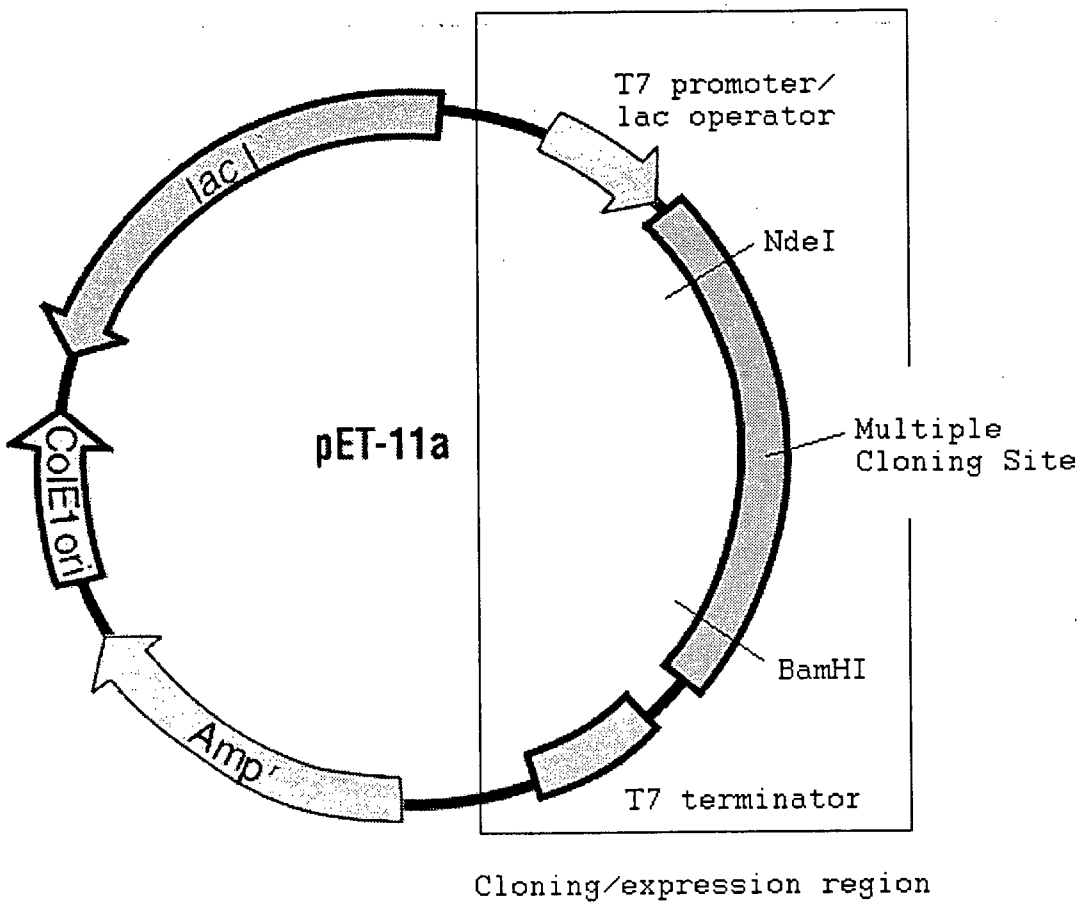


Figure 3.3 The pET11a expression vector. (Not shown to scale.)

Some vectors are intended only for the cloning of DNA fragments. This is the situation when the DNA itself is of interest. However, in the cases where it is desired to study the protein that is encoded within the cloned DNA, expression vectors are used instead. The expression vectors have the same basic features as the cloning vectors, but they also contain the features necessary to promote protein expression.

The pET11a expression vector, which is 5677 base pairs in size, contains each of the three features mentioned above (Figure 3.3). It has its own origin of replication (ColE1 ori in Figure 3.3), a gene to confer resistance to the antibiotic ampicillin (Amp^r in Figure 3.3), and several unique restriction endonuclease cleaving positions at unique locations in the DNA sequence. These cutting sites appear only in a highly specialized region of the vector known as the cloning/expression region. This region is another highly engineered feature of the vector, designed to maximally express the protein encoded by the fragment of DNA inserted into this section of the vector. Optimum expression is accomplished through the action of an assembly of genes that have been incorporated into the vector for this very purpose.

As shown in Figure 3.3, the pET11a cloning/expression region contains the regulatory machinery of the *lac* operon. Instead of the three normal structural genes whose expression is controlled by the operon (i.e. *lacZ*, *lacY* and *lacA* in Figure 3.1), the vector contains a multiple cloning site designed to accept the insertion of any new structural gene instead. This multiple cloning site contains several restriction endonuclease cleavage sites, such as *Nde*I and *Bam*HI, which are not located anywhere else in the vector. It is possible to insert a DNA fragment containing the complementary restriction sites directionally into the cloning/expression region between these two restriction sites. The resultant recombinant plasmid, which consists of the vector and the foreign DNA insert, would have the new inserted gene positioned such that its

expression is rigorously controlled by the *lac* operator. Protein expression can be initiated by displacing the *lac* repressor (whose gene, *lacI*, is also found in the pET11a vector) from its position associated with the *lac* operator. *In vivo*, this is accomplished by the presence of lactose within the cell. However, inducing protein expression with the non-metabolizable lactose-analogue isopropyl-1-thio-β-D-galactopyranoside (IPTG) is more useful for the purposes of *in vitro* protein overexpression.

The other change that has been engineered into the pET11 cloning/expression region is the promoter (Rosenberg *et al.*, 1987). The stronger T7 promoter has been inserted, to replace the normally present *lac* promoter that is recognized by the native *E. coli* RNA polymerase as the position to begin transcription. Similarly, the corresponding T7 terminator sequence is also included to mark the endpoint of mRNA transcription. The T7 promoter and terminator are recognized by T7 RNA polymerase, which is not normally present in *E. coli*. For this reason, a special strain of *E. coli*, such as *E. coli* JM109(DE3), that has been engineered to express this RNA polymerase must be used. *E. coli* JM109(DE3) has the gene for T7 RNA polymerase from bacteriophage λ DE3 integrated into the *E. coli* chromosome. Transcription of T7 RNA polymerase is under the control of the *lac* promoter, and therefore the T7 RNA polymerase gene is not expressed until it is induced by the addition of IPTG. This allows tight regulation of protein expression (Rosenberg *et al.*, 1987; Makrides, 1996).

3.3 Subcloning of the *rffD* Gene

The overall cloning strategy is shown in Figure 3.4. The scheme can be divided into two stages: (1) the preparation of the vector DNA, and (2) the preparation of the *rffD* gene for

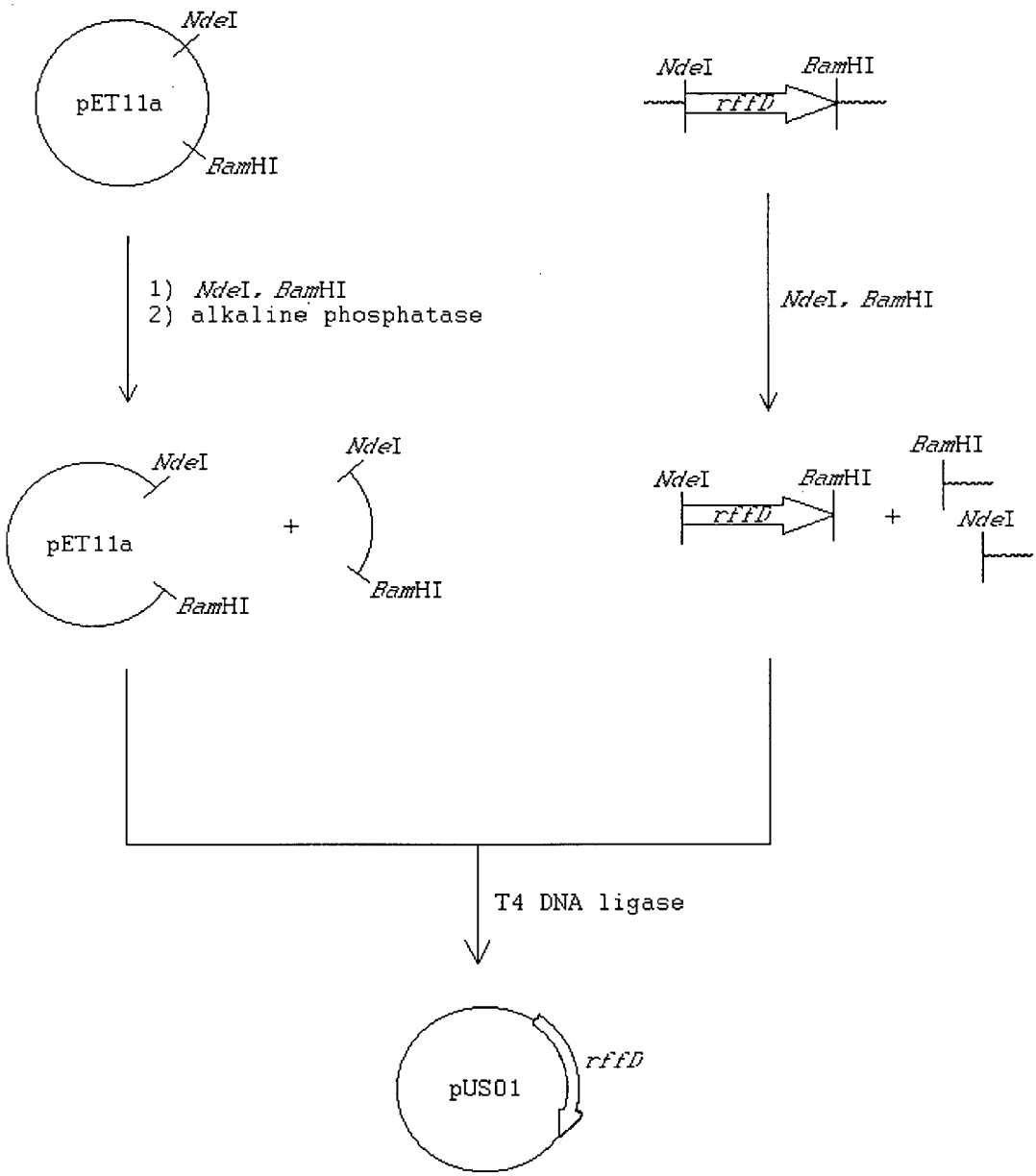


Figure 3.4 The general strategy for cloning *rffD* into the pET11a expression vector.

insertion. Once the two fragments are generated, they are subsequently ligated using T4 DNA ligase, resulting in the formation of a new chimeric recombinant DNA molecule consisting of the pET11a expression vector with *rffD* inserted between the *Nde*I and *Bam*HI restriction endonuclease sites of the cloning/expression region.

3.3.1 Preparation of the pET11a Vector DNA

The pET11a vector was digested with both *Bam*HI and *Nde*I restriction endonucleases simultaneously. Both enzymes operate efficiently under the same conditions, allowing simultaneous digestion in a single step. Following the digestion, the resultant DNA was treated with calf intestine alkaline phosphatase, an enzyme that removes the 5'-phosphate from the ends of DNA chains. Since the 5'-phosphate is necessary for the final ligation reaction to proceed, this treatment ensures that any vector that had been inefficiently cut by only one of the two restriction endonucleases could not recyclize during the ligation step. Otherwise, the recyclization would result in trace amounts of contaminating "empty" pET11a, which would be laborious to differentiate from the desired successfully ligated plasmid.

Following digestion and the alkaline phosphate treatment, the resultant vector DNA was purified by agarose gel electrophoresis, and visualized with ethidium bromide. The band containing the vector DNA was sliced from the gel, the DNA was recovered, further purified by ethanol precipitation, and stored at –20°C until the insert had been prepared, and the ligation of the two fragments could be performed.

3.3.2. Preparation of the *rffD* Gene Insert

The previous step had the pET11a vector prepared to allow a DNA fragment to be specifically and directionally inserted between the *Nde*I and *Bam*HI sites of the expression region. The complimentary insert is prepared by engineering a length of DNA to contain the *rffD* gene flanked by the *Nde*I-specific sequence at the start of the gene, and the *Bam*HI-recognized sequence at some position following.

The engineering is performed using a technique known as the polymerase chain reaction (PCR, Mullis and Faloona, 1987). PCR is used to amplify small quantities of double-stranded DNA, by incubating a DNA polymerase with a double-stranded DNA template and two site-specific primers. The incubated mixture is subjected to repeated cycles, during which the two strands of the DNA duplex are separated by heat denaturation, the two primers are annealed to their respective complementary sequences within the denatured DNA template, and the DNA polymerase extends the region of DNA between the two primers. The result is the exclusive exponential amplification of the DNA flanked by the primers. Generally, the primers are sequence specific, but several mismatches can be tolerated. The synthetic oligonucleotide primer may be designed to partly complement the sequence, and to simultaneously incorporate specifically engineered changes, such as a new restriction endonuclease-recognized sequence (Figure 3.5). Therefore, primer design requires prior knowledge of the nucleotide sequence at the periphery of the DNA region to be amplified.

In the literature, two different sequences for the *E. coli rffD* gene have been reported. The first came from the *E. coli* genome sequencing project, and reported the UDP-ManNAc dehydrogenase gene as encoding 379 amino acids (Daniels *et al.*, 1992). In the second published sequence, however, the same gene encodes a protein that is 420 amino acids long

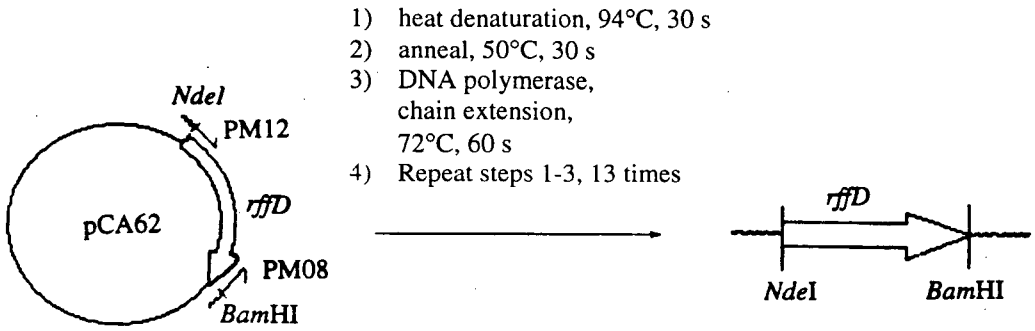


Figure 3.5 The PCR amplification of *rffD* between primers PM08 and PM12 results in the production of copies of *rffD* containing the *Nde*I restriction site at the start of the gene, and *Bam*HI after.

Table 3.1 A comparison of the two different reported sequences of *rffD* showing the conflicting residues after amino acid 17. The upper sequence, from Daniels *et al.* (1992), differs in the six nucleotides shown in bold. The lower sequence has nine nucleotides in that same location (Kiino *et al.*, 1994). The numbers above the sequence indicate the nucleotide, while those below indicate the corresponding amino acid position in the predicted protein sequence.

49	66
5'- CCA **ACC CGA** GCG TTT GCC -3'	
pro **thr arg** ala phe ala	
17	22
49	69
5'- CCA **ACG GCA GCA** GCG TTT GCC -3'	
pro **thr ala ala ala phe** ala	
17	23

(Kiino *et al.,* 1994). Closer inspection reveals that the two reported sequences differ in three positions.

The first discrepancy between the two reported sequences appears after nucleotide 52 of *rffD*, which corresponds to amino acid 17 of UDP-ManNAc dehydrogenase (Table 3.1). The six nucleotides, from 52 to 57, in the sequence reported by Daniels *et al.*, code for threonine followed by arginine. At the same position in the sequence reported by Kiino *et al.*, there are nine nucleotides instead, which code for an amino acid sequence of: threonine-alanine-alanine. The consequence of this discrepancy results in different amino acids predicted for amino acid 19 in the two resulting proteins, with the hypothetical protein reported by Kiino *et al.,* containing one more amino acid than the other. However, the three additional nucleotides present in the sequence reported by Kiino *et al.* do not result in a comparative shift in the nucleotide reading-frame for the remainder of the sequence.

The second conflict between the two published sequences appears to be a jumbled order reported for nucleotides 229, 230 and 231 in the Daniels *et al.* sequence, as compared to the corresponding three nucleotides in the Kiino *et al.* sequence (which are nucleotides 232, 233 and 234 in the Kiino *et al.* sequence due to the extra three nucleotides reported earlier in the sequence). The first case has a reported sequence of C-G-C, the triplet codon that encodes an arginine residue to be incorporated into the protein, while the other has G-C-C, which instead encodes alanine.

The third discrepancy results in a far more drastic difference between the resultant proteins predicted by the two published *rffD* sequences. The sequence reported by Daniels *et al.* has an additional guanine base at position 1124 that was not reported in the nucleotide sequence reported by Kiino *et al.* In the Daniels *et al.* sequence, this would result in a

Table 3.2 A comparison of the two different reported terminal sequences of *rffD*. The upper sequence, from Daniels *et al.* (1992), reports an additional G after nucleotide 1123 (in bold underline). The lower sequence has only a single G in that same location (Kiino *et al.*, 1994). The numbers above the sequence indicate the nucleotide position, while those below indicate the corresponding amino acid position in the predicted protein sequence.

1110										1142
5'-CTG	ACC	GGG	CTT	TG**G**	TAC	TCT	GGC	GCA	GCT	TGA -3'
leu	thr	gly	leu	trp	tyr	ser	gly	ala	ala	STOP
370										379

1110																			1158
5'-CTG	ACC	GGG	CTT	TGT	ACT	CTG	GCG	CAG	CTT	GAC	GAG	GCG	CTG	GCA	ACG-3'				
leu	thr	gly	leu	cys	thr	leu	ala	gln	leu	asp	glu	ala	leu	ala	thr				
371										380									
1159																			1206
5'-GCA	GAC	GTG	CTG	GTG	ATG	CTG	GTC	GAT	CAT	AGT	CAG	TTC	AAA	GTT	ATC-3'				
ala	asp	val	leu	val	met	leu	val	asp	his	ser	gln	phe	lys	val	ile				
			390										400						
1207																			1254
5'-AAT	GGC	GAC	AAT	GTC	CAT	CAG	CAG	TAT	GTC	GTC	GAT	GCC	AAA	GGA	GTC-3'				
asn	gly	asp	asn	val	his	gln	gln	tyr	val	val	asp	ala	lys	gly	val				
								410											

1255		1263
5'-TGG	CGC	TGA-3'
trp	arg	STOP
	420	

tryptophan residue incorporated at amino acid 375 in the expressed dehydrogenase (Table 3.2). In that same sequence, the tryptophan residue would be followed by an additional five amino acids after which a stop-signal is encountered on the DNA, signalling the end of the putative protein. Overall, the protein encoded by the nucleotide sequence reported by Daniels *et al.* is predicted to contain a total of 379 amino acids. Conversely, the sequence reported by Kiino *et al.* lacks the additional guanine base, resulting in a shift in the triplet reading-frame of that sequence relative to the reading frame in the shorter, 379 amino-acid encoding, sequence. The result is that a stop codon is not encountered in the reading-frame of the Kiino *et al.* sequence, until nucleotide 1261. As a result, the protein encoded by the nucleotide sequence reported by Kiino *et al.* is predicted to be longer than the other reported sequence, with a total of 420 amino acids. Therefore, a single error present in one of the two sequences, results in the termination signal occurring at two different positions, raising uncertainties regarding the actual number of amino acid residues in UDP-ManNAc dehydrogenase.

In terms of the oligonucleotide primers necessary for the PCR reaction, the first and second errors are inconsequential, and have no bearing on the expected start or terminal sequences of *rffD*. The third error also has no effect on the reported sequence for the nucleotides at the start of the protein. As a result, it was possible to unambiguously design primer PM12 to be identical to the first eighteen nucleotides of the sense strand of *rffD*, and incorporate the *Nde*I restriction sequence into the start of the intended PCR product (Table 3.3).

The second error does however dramatically change the expected terminal sequence of the UDP-ManNAc dehydrogenase gene. At the time when the primers were initially designed, the sequence reported by Kiino *et al.* had not yet been known. The primer, PM04, was

designed to complement the sense strand at the termination of *rffD*, based on the sequence from Daniels *et al.*, and incorporate a *Bam*HI endonuclease site into the PCR product. Later however, as we became aware of the second sequence, a second primer, PM08, was designed to complement the sense strand of the gene at the termination of the longer (420 amino acid-encoding) sequence, and to incorporate the *Bam*HI restriction site (Table 3.3).

Table 3.3 The primers used to subclone *rffD*.

PM04[a]	5' – TAG GGA TCC TCA AGC TGC GCC AGA GTA CCA – 3'	
	*Bam*HI	
PM08[b]	5' – CGG GCC GGA TCC TCA GCG CCA GAC TCC TTT – 3'	
	*Bam*HI	
PM12[c]	5' – GG CGG GCC CAT ATG AGT TTT GCG ACC ATT TCT – 3'	
	*Nde*I	

[a] PM04 is complementary to the sense strand at the termination of the *rffD* sequence reported by Daniels *et al.* (1992) which encodes a protein 379 amino acids in length.
[b] PM08 is complementary to the sense strand at the termination of the *rffD* sequence reported by Kiino *et al.* (1993) which encodes a protein 420 amino acids in length.
[c] PM12 is identical to the sense strand at the start of *rffD* (consistent with both reported sequences).

With the oligonucleotide PCR primers designed, our attention turned towards acquiring the template for the PCR reaction. It is possible to use genomic DNA directly as the template, although working with smaller, manageable vector DNA is preferable. It turns out that a previous study (Meier-Deiter *et al.*, 1992) had cloned a 4.9 kbp segment of *E. coli* genomic DNA into a cloning vector (as opposed to an expression vector, which was our goal). That

study had subsequently shown that the resultant plasmid, which was named pCA62, was able to provide both UDP-GlcNAc 2-epimerase and UDP-ManNAc dehydrogenase activity to a strain of *E. coli* that lacks both enzymes. Evidently, the *E. coli* cells could use the cloning-vector DNA as a template to express a functional quantity of both enzymes. This indicated that the section of DNA cloned into pCA62 (shown in Figure 2.1) contains the genes for both the epimerase (*rffE*) and the dehydrogenase (*rffD*). Since the dehydrogenase had already been cloned, it was more convenient to use the pCA62 plasmid as the template for our own cloning efforts.

The *rffD* gene was amplified by PCR from pCA62 using the PM08 and PM12 primers. As judged by agarose gel electrophoresis, a single PCR product was generated, whose electrophoretic mobility confirmed it was the expected size for the *rffD* gene (ca. 1300 bp). The PCR product was digested with *Nde*I and *Bam*HI simultaneously in order to remove both of the ends, and free the restriction sites for the ligation step. Using agarose gel electrophoresis, the resultant DNA was separated from the clipped oligonucleotide ends. The insert was recovered from the agarose gel, and desalted by ethanol precipitation.

3.3.3 The pUS01 Plasmid

The prepared insert and pET11a vector were ligated with T4 DNA ligase in the presence of ATP, resulting in the pUS01 plasmid. The ligated mixture was transformed directly into $CaCl_2$-competent *E. coli* JM109(DE3), and plated onto LB agar media containing ampicillin. Several of the resulting colonies were randomly selected and grown as a small culture. These cultures were treated with 0.4 mM IPTG in order to induce protein synthesis. Crude cell extracts were tested for high expression of a protein the same size as that expected

for the dehydrogenase by analyzing samples by SDS-PAGE. All colonies tested showed the intended overexpression.

The crude cell extract from one colony was tested, and showed significant dehydrogenase activity when incubated with UDP-GlcNAc 2-epimerase, NAD^+ and UDP-GlcNAc. The protein was also analyzed by electrospray mass spectrometry, and was determined to have a subunit mass similar to that expected for the dehydrogenase, based on the mass calculated for the 420 amino acid sequence: found, 45 718 ± 4 Da; calculated, 45 838 Da. The difference between the expected mass and the observed mass can be attributed to the post-translational loss of the N-terminal methionine from the enzyme, which is commonly observed with proteins expressed in prokaryotes (Voet and Voet, 1990b). This loss was confirmed by performing an N-terminal sequence on the enzyme. The first six amino acid residues were unambiguously determined to be Ser-Phe-Ala-Thr-Ile-Ser, which matched the expected sequence, without the N-terminal methionine (Figure 3.6).

The cloning/overexpression region of the pUS01 plasmid was sequenced using the commmercial T7 and T7 terminator primers, and a series of internal sequence specific primers complementary to positions within the *rffD* gene (Figure 3.6). Most of the obtained sequences possessed overlapping regions, allowing the sequence of many of the nucleotides in *rffD* to be confirmed from different independent sequence determinations. The result was an unambiguous sequence determination with no errors within *rffD*. It was clear that the sequence of *rffD* subcloned into pET11a agreed with the second published sequence for that gene (from Kiino *et al.*, 1993), which corresponds to an expected protein 420 amino acids in length. Therefore, *rffD* has been successfully ligated into pET11a to yield the pUS01 plasmid,

1 T7 promoter *lac* operator
CGC GAA ATT AAT ACG ACT CAC TAT AGG GGA ATT GTG AGC GGA TAA CAA TTC CCC TCT
T7 promoter primer→

58 ribosome binding site *Nde*I **PM07 primer→**
AGA AAT AAT TTT GTT TAA CTT TAA GAA GGA GAT ATA CAT ATG AGT TTT GCG ACC ATT
M S F A T I

115
TCT GTT ATC GGA CTG GGT TAT ATC GGG CTG CCA ACG GCA GCA GCG TTT GCC TCA CGG
S V I G L G Y I G L P T A A A F A S R

172
CAA AAA CAG GTA ATT GGT GTC GAT ATC AAC CAA CAT GCG GTT GAT ACC ATC AAT CGT
Q K Q V I G V D I N Q H A V D T I N R

229
GGC GAA ATC CAT ATC GTC GAA CCT GAT TTG GCG AGT GTA GTA AAA ACT GCC GTA GAA
G E I H I V E P D L A S V V K T A V E

286
GGC GGT TTT TTA CGA GCG AGC ACG ACG CCA GTT GAA GCG GAT GCC TGG CTG ATT GCT
G G F L R A S T T P V E A D A W L I A

343
GTA CCC ACG CCG TTT AAG GGC GAT CAT GAG CCA GAT ATG ACC TAC GTT GAA TCG GCT
V P T P F K G D H E P D M T Y V E S A

400
GCT CGC TCC ATT GCG CCA GTG CTG AAA AAA GGC GCG CTG GTG ATC CTT GAA TCC ACC
A R S I A P V L K K G A L V I L E S T

457 **PM09 primer→**
TCG CCG GTG GGG TCA ACC GAG AAG ATG GCA GAA TGG TTA GCA GAG ATG CGT CCG GAT
S P V G S T E K M A E W L A E M R P D

514
CTC ACT TTC CCG CAG CAG GTG GGC GAG CAG GCG GAC GTC AAC ATT GCT TAC TGC CCG
L T F P Q Q V G E Q A D V N I A Y C P

571
GAA CGC GTG TTA CCA GGA CAG GTA ATG GTC GAG CTG ATT AAA AAC GAT CGC GTG ATT
E R V L P G Q V M V E L I K N D R V I

628
GGT GGT ATG ACG CCG GTT TGT TCG GCC CGC GCC AGC GAA CTG TAC AAA ATT TTC CTC GAA
G G M T P V C S A R A S E L Y K I F L E

688
GGT GAG TGT GTC GTC ACT AAC TCG CGG ACG GCG GAA ATG TGT AAG CTC ACC GAA AAC
G E C V V T N S R T A E M C K L T E N

745 **←PM11 primer**
AGC TTC CGC GAT GTG AAT ATC GCT TTT GCT AAT GAA TTG TCG CTG ATT TGT GCC GAT
S F R D V N I A F A N E L S L I C A D

Figure 3.6 The sequence of the cloning/expression region of pUS01 showing the *rffD* gene inserted between *Nde*I and *Bam*HI. Primers used for sequencing are shown in bold. Other sequences of interest are indicated by underline.

802
CAG GGG ATT AAC GTC TGG GAA CTG ATT CGC CTG GCG AAT CGT CAC CCT CGC GTT AAT ATT
Q G I N V W E L I R L A N R H P R V N I

862
CTT CAG CCT GGC CCT GGC GTG GGC GGT CAC TGC ATT GCT GTT GAT CCG TGG TTT ATC GTG
L Q P G P G V G G H C I A V D P W F I V

922 **PM10 primer→**
GCA CAG AAC CCC CAG CAG GCG CGG CTT ATC CGT ACC GCG **CGC GAA GTG AAC GAT CAC**
A Q N P Q Q A R L I R T A R E V N D H

979
AAA CCG TTC TGG GTT ATC GAT CAG GTG AAA GCG GCG GTG GCT GAT TGC CTG GCG GCT
K P F W V I D Q V K A A V A D C L A A

1036
ACC GAT AAA CGC GCC AGT GAA CTG AAA ATC GCC TGC TTT GGT CTG GCG TTT AAA CCG
T D K R A S E L K I A C F G L A F K P

1093
AAT ATT GAT GAC CTG CGC GAA AGC CCG GCG ATG GAA ATC GCT GAA CTG ATC GCC CAG
N I D D L R E S P A M E I A E L I A Q

1150
TGG CAT AGC GGC GAA ACT CTG GTT GTT GAG CCT AAC ATC CAC CAG TTG CCG AAA AAA
W H S G E T L V V E P N I H Q L P K K

1207
CTG ACC GGG CTT TGT ACT CTG GCG CAG CTT GAC GAG GCG CTG GCA ACG GCA GAC GTG
L T G L C T L A Q L D E A L A T A D V

1264
CTG GTG ATG CTG GTC GAT CAT AGT CAG TTC AAA GTT ATC AAT GGC GAC AAT GTC CAT CAG
L V M L V D H S Q F K V I N G D N V H Q
1324 *Bam*HI
CAG TAT GTC GTC GAT GCC AAA GGA GTC TGG CGC TGA GGA TCC GGC TGC TAA CAA AGC
Q Y V V D A K G V W R STOP

1381 **←T7 terminator primer**
CCG AAA GGA AGC TGA GTT GGC TGC TGC CAC **CGC TGA GCA ATA ACT AGC** ATA ACC CCT

1438 T7 terminator
TGG GGC CTC TAA ACG GGT CTT GAG GGG TTT TTT G

Figure 3.6 (continued) The sequence of the pUS01 cloning/expression region.

which can be induced for the high level expression of an active recombinant UDP-ManNAc dehydrogenase identical in sequence to the enzyme from natural sources.

A similar attempt was made to subclone *rffD* using the PM04 primer instead of PM08. This primer matches the terminal region of the shorter 379 amino acid encoding sequence. The PCR product from the amplification of the segment of *rffD* between was digested with *Bam*HI and *Nde*I, and ligated into the pET11a vector. The resulting plasmid could be induced to express a protein of comparable size to that expected for the dehydrogenase. However, the expressed protein remained insoluble in the crude cell lysate. It is likely that this protein was the truncated version of the dehydrogenase, which was unable to fold into the native conformation due to the absence of forty C-terminal amino acids.

3.4 Expression and Purification of Recombinant UDP-*N*-Acetylmannosamine Dehydrogenase

The pUS01 plasmid was transformed into competent *E. coli* JM109(DE3), and grown at 37°C in LB media in the presence of 100 μg/ml ampicillin. Protein expression was induced with 0.4 mM IPTG. After 60 minutes, an aliquot of the cell culture showed the presence of a new overexpressed protein (Figure 3.7).

The purification scheme for the recombinant UDP-ManNAc dehydrogenase was similar to the purification of UDP-GlcNAc 2-epimerase. The difference, however, was the dehydrogenase was required in larger quantities, which necessitated a larger scale purification. In order to maximize protein yields, the cell cultures were allowed to grow for five hours following induction, before the cells were harvested.

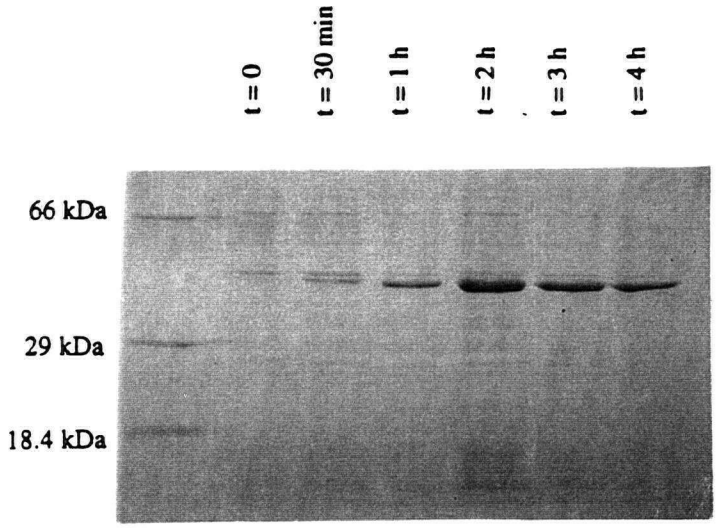


Figure 3.7 SDS-polyacrylamide gel showing the expression of UDP-ManNAc dehydrogenase over time, following induction with 0.4 mM IPTG. The protein components of the samples from each time point were separated by 12% SDS-PAGE. The molecular weight standards are: 66 kDa, bovine serum albumin; 29 kDa, carbonic anhydrase; and 18.4 kDa, β-lactoglobulin.

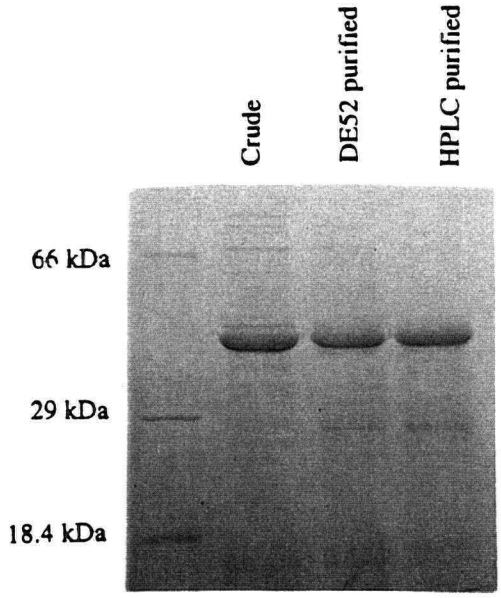


Figure 3.8 SDS-polyacrylamide gel of crude recombinant UDP-ManNAc dehydrogenase, deydrogenase purified by DE52 anion exchange chromatography, and dehydrogenase purified by a Waters AP-1 Protein-Pak Q 8HR HPLC column. The protein components of the crude and purified samples were separated by 12% SDS-PAGE. The molecular weight standards are: 66 kDa, bovine serum albumin; 29 kDa, carbonic anhydrase; and 18.4 kDa, β-lactoglobulin.

The crude cell extract was prepared by passing the resuspended cells twice through a chilled French pressure cell, and centrifuging the ruptured cells to separate out the cell wall debris and unbroken cells. In contrast to the epimerase, which was partially purified by eluting the crude protein from a DE52 anion-exchange column with a linear salt gradient, the crude dehydrogenase was loaded onto a larger column of the same resin, and eluted with 0.3 M NaCl buffer. This successfully removed the negatively charged impurities that could bind strongly to the anion-exchange HPLC column used in the following purification step.

The partially purified protein from the DE52 plug was desalted, concentrated and further purified by anion-exchange HPLC, using a large volume Waters AP-2 Protein-Pak Q 40HR column, and eluting with a linear gradient of 0 to 0.4 M NaCl. Fractions were tested for activity by incubating an aliquot of an eluent fraction at 37°C with UDP-GlcNAc, NAD^+, and UDP-GlcNAc 2-epimerase, while monitoring the sample at 340 nm. The combination of both the epimerase and the dehydrogenase, in fractions containing the active dehydrogenase, would show an increase at that wavelength, corresponding to the generation of NADH from the UDP-ManNAc produced by the epimerase. A single large peak that eluted from the HPLC around 0.2 M NaCl was determined to contain all the dehydrogenase activity. The resulting purified protein appeared greater than 85% homogeneous when analyzed by SDS-PAGE (Figure 3.8). The dehydrogenase purified in this manner was used directly in the coupled enzyme assay, described in Chapter Two, for the characterization of UDP-GlcNAc 2-epimerase.

3.5 Characterization of Recombinant UDP-*N*-Acetylmannosamine Dehydrogenase

The kinetic constants for UDP-ManNAc dehydrogenase had been previously measured for the enzyme at pH 10. At this optimal pH for the dehydrogenase, it was reported that the dehydrogenase has an apparent K_m for UDP-ManNAc of 0.38 mM and a k_{cat} of 5.3 s^{-1} (Kawamura *et al.*, 1979).

Since the purified dehydrogenase was to be used in the coupled enzyme assay for UDP-GlcNAc 2-epimerase activity, which operated at a pH of 8.8, we were interested in determining the dehydrogenase kinetic constants at this lower pH. The results of the pH/rate profile for the dehydrogenase from the previous study suggested that the enzyme retains 75% of its activity at pH 8.8 (Kawamura *et al.*, 1979).

Unfortunately the substrate, UDP-ManNAc, is not commercially available. The axial 2"-acetamido group renders the nucleotide sugar somewhat unstable, and therefore difficult to prepare. A small sample of synthetic UDP-ManNAc was generously provided by Mr. Rafael Sala of our group. However the low yields associated with the multi-step synthesis of this substrate limited the amount available. As a result, it was not possible to perform a more detailed kinetic characterization of the enzyme or to confirm the kinetic constants of the dehydrogenase at pH 10.

The rate of NADH formation was measured at pH 8.8, with various initial concentrations of UDP-ManNAc, ranging from 0.23 mM to 5.7 mM. A plot of the enzyme velocity versus the substrate concentration was fit to an equation describing Michaelis-Menten kinetics, and gave an apparent K_m of 1.2 ± 0.4 mM and a k_{cat} of 1.2 ± 0.2 s^{-1} (Figure 3.9). The error reported is a reflection of the conformity of the acquired data from one experiment to the determined curve-of-best-fit. Insufficient quantities of UDP-ManNAc limited our

determination to two independent data sets. A second experiment gave similar results, with the apparent K_m of 1.0 ± 0.6 mM and a k_{cat} of 1.4 ± 0.4 s^{-1} (data not shown). Therefore it was found that the two different determinations, which were performed at different times, differed by approximately 20%.

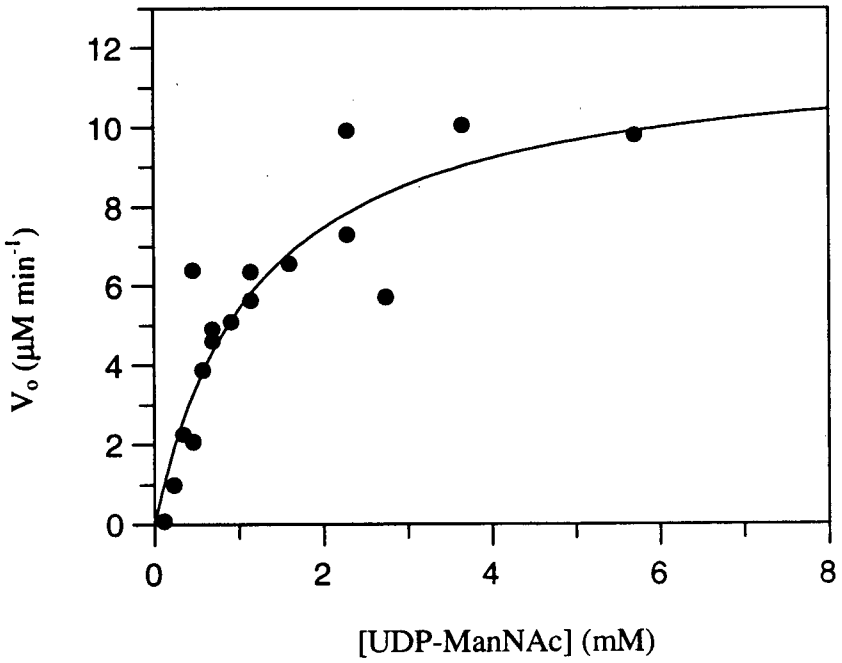


Figure 3.9 Plot of the initial velocity of UDP-ManNAc dehydrogenase as a function of UDP-ManNAc concentration, measured at pH 8.8.

The differences between the results reported here for a pH of 8.8, and those previously reported for pH 10 can be attributed to the change in pH between the two studies. However, the k_{cat} and K_m for the dehydrogenase measured at pH 8.8 are within the range expected for the enzyme operating slightly below its optimal pH. While the previous results suggested that the

dehydrogenase retained around 75% of its activity at pH 8.8, the determination in this study found that the dehydrogenase was only 20% active compared to the fully active protein at pH 10. It is possible that our own UDP-ManNAc dehydrogenase preparation was less active at pH 10 than that previously reported for the enzyme purified from natural sources in the previous study, but without any additional synthetic UDP-ManNAc available it was not possible to test this further.

3.6 Conclusions

The *rffD* gene, which encodes the enzyme UDP-ManNAc dehydrogenase, has been subcloned into the cloning/expression region of the pET11a expression vector, between the *Nde*I and *Bam*HI restriction sites, generating the pUS01 plasmid. Transformation of pUS01 into *E. coli* and induction with IPTG yielded high levels of expression of the dehydrogenase. The dehydrogenase was purified to homogeneity and kinetically characterized under the conditions similar to the assay for which its cloning had been intended. It was demonstrated that the original assignment for the *rffD* gene, made by Daniels *et al.* (1992) was correct, but the resulting protein actually contains 419 amino acids, following the post-translational loss of the N-terminal methinonine residue.

With UDP-ManNAc dehydrogenase abundantly available, it was possible to employ the enzyme in the direct spectrophotometric coupled enzyme assay for UDP-GlcNAc 2-epimerase activity that was described in Chapter Two.

3.7 Experimental Methods

3.7.1 Chemicals and Enzyme Substrates

All routinely used chemicals were obtained as described in Chapter Two, unless noted otherwise.

UDP-ManNAc was synthesized from ManNAc-α-1-phosphate, which was prepared according to the procedure outlined by Salo and Fletcher (1970). ManNAc-α-1-phosphate was coupled to uridine 5'-monophosphomorpholidate using the procedure of Wittman and Wong (1997). The UDP-ManNAc used in this study was synthesized and generously provided by Mr. Rafael Sala, a co-worker in laboratory of Dr. Martin Tanner, Department of Chemistry, University of British Columbia.

3.7.2 Oligonucleotides, Plasmids and Strains

Some of the primers for cloning and sequencing were prepared by Tracy Evans, Patrick Haering and Debbie Neufeld at the Nucleic Acid – Protein Service Unit, Biotechnology Laboratory, University of British Columbia, using a Perkin-Elmer Applied Biosystems Model 380B DNA synthesizer or a Model 394 DNA/RNA synthesizer. Oligonucleotides prepared on these instruments were deprotected overnight at 55°C, and subsequently ammonia-butanol purified. The remaining primers used in this study were prepared by Dr. Ivan J. Sadowski, in the Department of Biochemistry, University of British Columbia, using a Perkin-Elmer Applied Biosystems Model 391 PCR-MATE DNA synthesizer. Oligonucleotides prepared on this instrument were deprotected in a similar fashion.

The pCA62 plasmid, used as the template for subcloning *rffD*, consists of a 4.9 kbp fragment of genomic *E. coli* DNA inserted between the *Sal*I and *Hin*DIII restriction sites of a

pBR322 cloning vector. That plasmid was generously donated by Dr. Paul D. Rick, Department of Microbiology, Uniformed Services University of the Health Sciences, Bethesda, Maryland. The pET11a expression vector used to subclone *rffD*, and overexpress the gene product, was obtained from Novagen.

E. coli DH5α, was used for general plasmid manipulations, unless otherwise noted. *E. coli* JM109(DE3) was used as the host for the overexpression of UDP-ManNAc dehydrogenase.

3.7.3 Preparation of Plasmid DNA

LB broth (30 mL, containing 100 μg/mL ampicillin) was inoculated with *E. coli* DH5α that had been previously transformed with pET11a. The cells were incubated in a 37°C shaker (New Brunswick Scientific Innova 3000 digital incubator shaker) overnight. Cells were harvested from the growth medium by centrifugation at 13 000 rpm in a Fisher Scientific Allied Model 235C micro-centrifuge. The plasmid DNA was recovered from the cells using the WIZARD Mini-Preps Plasmid DNA purification system kit from Promega.

The recovered pET11a vector DNA (200 μL) was treated directly with both *Nde*I (20 U) and *Bam*HI (20 U) restriction endonucleases (from Promega), following addition of 22 μL of Promega "10X Buffer D" (60 mM Tris-HCl, pH 7.9, 1500 mM NaCl, 60 mM $MgCl_2$ and 10 mM DTT), the buffer that is included with the purchase of *Nde*I. The vector DNA was left to digest in a 37°C bath for 12 hours, before an additional 20 U of each enzyme was added, and the reaction was allowed to proceed overnight. The restriction endonucleases were heat inactivated at 70°C for an hour, and cooled to 37°C, at which time 20 U of calf intestine

alkaline phosphatase (Promega) was added. After incubation at 37°C for four hours, the alkaline phosphatase was heat-killed at 70°C for an hour.

The digested DNA was purified by 1.2 % agarose gel electrophoresis, using *Hin*dIII and *Eco*RI dually digested λ-phage DNA (Promega) as molecular weight markers, and visualizing with ethidium bromide/UV. The DNA band in the correct size range (ca. 5000 nt) was excised from the gel, and the digested DNA was recovered from the gel slice using the Promega WIZARD PCR Preps DNA Purification System kit. The recovered vector DNA was subsequently desalted by ethanol precipitation, dried *in vacuo*, resuspended in H_2O, and stored at –20 °C.

3.7.4 Preparation of *rffD* Insert DNA

The PM08 oligonucleotide primer was designed to complement the C-terminal sequence of *rffD* and introduce the *Bam*HI restriction site thereafter. The PM12 primer was identical to the start of *rffD*, and was designed to introduce the *Nde*I restriction site, with the fourth, fifth and sixth nucleotides of the *Nde*I sequence also encoding the start codon for RffD translation.

The *rffD* gene was amplified by PCR from plasmid pCA62, using PM08 and PM12 as the site-specific primers for DNA chain replication, using a M J Research MiniCycler. Each PCR reaction mixture (100 μL total volume) contained: 0.1 mM each dNTP, 0.5 μM each primer, 0.1 μg pCA62, 1.5 mM $MgCl_2$, 50 mM KCl, 20 mM Tris-Cl (pH 8.4), and 2.5 U *Taq* DNA polymerase (Gibco/BRL). 100 μL of mineral oil was added to prevent evaporation during thermocycling. Following an intial melt at 94°C for 1 min, the reaction mixture underwent 14 cycles involving a 94°C melt for 30 s, 50°C annealing for 30 s, and 72°C chain

extension for 1 min. The final chain extension, following the last cycle, proceeded for an additional 7 min at 72°C.

The amplified DNA samples from six PCR reactions were recovered using the Promega WIZARD PCR Preps PCR Product DNA purification System kit, and the recovered DNA was digested directly by the addition of 20 U each of *Nde*I and *Bam*HI, and Promega "Buffer D" (final concentrations: 6 mM Tris-HCl, pH7.9, 150 mM NaCl, 6 mM $MgCl_2$, and 1 mM DTT). The PCR product was digested at 37°C overnight, at which point an additional 20 U of each restriction endonuclease was added, and allowed to react for another 12 h.

The digested PCR product was purified by 1.2% agarose gel electrophoresis, and the solitary DNA band, which was consistent in size with the *rffD* gene (ca. 1200 nt) was excised. The PCR product was recovered from the gel slice using the Promega WIZARD PCR Preps kit. The recovered PCR product was further purified by ethanol precipitation, dried *in vacuo*, and resuspended in H_2O.

3.7.5 Preparation of pUS01

The purified digested PCR product was ligated into the digested pET11a vector using the manufacturer's recommended procedure for the ligation of cohesive-end DNA that accompanied the purchased T4 DNA ligase (Promega). The ligation reaction mixture contained buffer (50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, and 5% (w/v) polyethylene glycol-8000), 1U T4 DNA ligase, and an estimated 3:1 ratio of insert to vector in a total of 20 μL.

After ligating for 24 hours at 20°C, the entire ligation mixture was transformed into $CaCl_2$-competent *E. coli* JM109(DE3), and plated onto LB agar medium, containing 100

μg/mL ampicillin. Colonies were selected randomly from the plate, and tested for protein overexpression by inoculating 5 mL LB cultures (with 100 μg/mL ampicillin) and inducing protein expression by the addition of 0.4 mM IPTG when the optical density (600 nm) had reached 0.8-1.2. Whole cells were boiled in SDS sample buffer (stock SDS sample buffer contains 0.12 mM Tris-HCl, pH 6.8, 40 mg/mL SDS, 20% glycerol, 31 mg/mL DTT, and 0.01 mg/mL bromophenol blue) and analyzed for protein content by 12% SDS-PAGE. All colonies tested showed a single highly expressed protein band amidst the other proteins expressed at their normal levels. The molecular weight of the overexpressed protein had an electrophoretic mobility corresponding to *ca.* 40 kDa, as expected for UDP-ManNAc dehydrogenase.

3.7.6 DNA Sequencing of the Cloning/Expression Region of pUS01

DNA sequences of the cloning/expression region of the pUS01 plasmid were determined by Tracy Evans and Debbie Neufeld at the Nucleic Acid – Protein Service Unit, in the Biotechnology Laboratory, University of British Columbia. The sequencing reactions were performed using the Perkin-Elmer Applied Biosystems PRISM Dye Terminator Cycle Sequencing Ready Reaction kit, each of the primers listed in Table 3.4, and pUS01 plasmid recovered from cell cultures using the WIZARD Mini-Preps Plasmid DNA Purification kit.

With the exception of the T7 promoter and T7 terminator primers, which are specific for sequences on the pET11a vector DNA outside of the multiple cloning region, all other primers were specific internal primers designed for sequencing the *rffD* gene. The location in the *rffD* gene that each primer was designed to match is indicated in Figure 3.6. The sequence reactions were analyzed using either a Perkin-Elmer Applied Biosystems Model 373 or Model 377 DNA automated sequencer.

Table 3.4 The oligonucleotide primers used to sequence the cloning/expression region of the pUS01 plasmid.

PM09[a]	5' –CGG TGG GGT CAA CCG AGA – 3'
PM10[b]	5' – CGC GAA GTG AAC GAT CAC – 3'
PM11[c]	5' – CGC GGA AGC TGT TTT CGG – 3'
T7[d]	5' – TAA TAC GAC TCA CTA TAG G – 3'
T7 Terminator[e]	5' – GCT AGT TAT TGC TCA GCG G – 3'

[a] PM09 is identical to the sense strand of *rffD* from nucleotides 365 to 383.
[b] PM10 is identical to the sense strand of *rffD* from nucleotides 865 to 883.
[c] PM11 is complementary to the sense strand of *rffD* from nucleotides 658 to 640.
[d] The T7 primer is identical to a sequence in the sense strand of pET11a that encompasses the T7 promoter.
[e] The T7 terminator primer is complementary to a sequence in the sense strand of pET11a that overlaps the T7 terminator.

3.7.7 Expression and Purification of UDP-*N*-Acetylmannosamine Dehydrogenase

A single colony of *E. coli* JM109(DE3), transformed with pUS01, was used to inoculate 2 L of LB supplemented with 100 μg/mL ampicillin. The cell culture was grown to an optical density (600 nm) between 0.8 and 1.2, at which point protein expression was induced with the addition of 0.4 mM IPTG. After five hours of further growth, the cells were harvested by centrifugation (Sorvall GSA rotor, 5 min, 5000 rpm), and the cell pellet (4 g) was

stored at –78°C. The pellet was later thawed rapidly with warm water, and the cells were resuspended in 50 mL of cold 50 mM Tris-HOAc buffer, pH 7.9, containing 2 mM DTT, 1 mM EDTA, 1 mM PMSF, 1 μg/mL pepstatin, and 5 μg/mL aprotonin (buffer A). the resuspended cells were passed twice through a chilled French pressure cell, rupturing at 20 000 psi.

The fresh crude cell extract was centrifuged for 20 min at 6000 rpm (Sorvall GSA rotor). The supernatant was diluted with buffer B (50 mM Trien-HCl, pH 8.5, 2 mM DTT, 10% glycerol) until the conductivity of the solution was approximately the same as that of the stock buffer B. The diluted cell extract was applied to a 100 mL column of DE52 pre-equilibrated in buffer B, and the column was washed with 200 mL of buffer B. The crude protein was eluted from the weak anion-exchange resin with 0.3 M NaCl in buffer B, and the eluent was monitored at 280 nm, using a SPECTRUM Spectra/Chrom Flow Thru UV detector. The entire protein fraction was pooled, concentrated using Millipore UltraFree-15 Centrifugal Filter Device (10 kDa NMWC) concentrators, and desalted by repeated dilution and concentration with buffer B.

The concentrated, DE52 purified protein was purified using a Waters AP-2 Protein-Pak Q 40HR HPLC column (20 mm X 300 mm) pre-equilibrated with degassed buffer B (at 22°C), and eluting with a linear gradient from 0 to 0.4 M of NaCl in buffer B (4 mL/min over 200 min). A single large peak eluted at salt concentration of 0.2 M. The presence of the dehydrogenase within this large fraction, which comprised 90% of the total injected protein, was confirmed by SDS-PAGE. The purified dehydrogenase was concentrated, desalted, then frozen in liquid N_2, for storage at –78°C.

3.7.8 N-Terminal Sequence Determination of UDP-*N*-Acetylmannosamine Dehydrogenase

A sample of the purified UDP-ManNAc dehydrogenase (20 μg of protein) was run on a 12% SDS-PAGE gel, using a Bio-Rad MINI-PROTEAN II gel electrophoresis apparatus, with glycine/tris buffer (2.7g glycine, 5.8g Tris base, and 10% methanol in 1 L total volume) as the electrophoresis running buffer. The protein was subsequently immobilized onto a 0.45μm PROBLOTT PEABI polyvinyldifluoride (PVDF) membrane, by electrophoretically blotting for 5 h at 60 volts (at 4°C). The PVDF was stained with 0.025% Coomasie brilliant blue in 40% methanol, to confirm the success of the protein blot.

N-terminal protein sequencing was performed on the PVDF-immobilized protein by Suzanne Perry of the Nucleic Acid-Protein Service Unit, Biotechnology Lab, University of British Columbia. The band containing the immobilized UDP-ManNAc dehydrogenase was cut from the PVDF membrane, and analyzed on a Perkin-Elmer Applied Biosystems Model 476A automated protein sequencer, using standard gas phase Edman Chemistry. The 476A is equipped with an on-line reverse phase HPLC and 610A data analysis system. Separation and analysis of the amino acid sequence occurs on the basis of the amino acid affinity for the stationary phase of the PTH C-18 column packing material.

3.7.9 Physical Characterization of UDP-*N*-Acetylmannosamine Dehydrogenase

The purity of the dehydrogenase was assessed by 12% SDS-PAGE, in a fashion identical to the epimerase, described in Chapter Two. The molecular weight of the dehydrogenase was determined by electrospray ionization mass spectrometry, also under identical conditions to those described for the epimerase. The electrospray mass specrometry

was performed by Shouming He in the lab of Dr. Stephen Withers, Department of Chemistry, University of British Columbia.

3.7.10 Kinetic Characterization of UDP-*N*-Acetylmannosamine Dehydrogenase

The kinetic parameters of the purified recombinant dehydrogenase were determined at pH 8.8, under conditions relevant to the coupled enzyme assay for UDP-GlcNAc 2-epimerase activity. Initial rates were measured at 37°C with various concentrations of UDP-ManNAc, by following the generation of NADH at 340 nm (ε_{340} = 6220 $M^{-1}cm^{-1}$), using a Varian Cary 3E spectrophotometer. The assay mixtures contained 50 mM Tris-HCl (pH 8.8), 2 mM DTT, 4.5 mM NAD^+, $2.6x10^{-3}$ mg UDP-ManNAc dehydrogenase, and UDP-ManNAc (0.23 mM to 5.7 mM) in a total volume of 335 μL. The concentration of the UDP-ManNAc stock solution was determined using A_{262} (ε_{262} = 8770 M^{-1} cm^{-1}). Reactions were initiated by the addition of UDP-ManNAc dehydrogenase. The observed rates were divided by a factor of two, in order to account for the stoichiometric production of two NADH molecules from each molecule of substrate. A plot of the rate data as a function of the substrate concentration showed typical Michaelis-Menten kinetics. The kinetic parameters were determined by a direct fit of the data to the equation:

$$v = \frac{V_{max}\,[S]}{K_m + [S]}$$

using the computer program Grafit. The program performs a non-linear regression analysis of the data following the method of Marquart (1963), and reports an error for the data based on

the deviation of the data from the calculated curve-of-best-fit. This value is reported with the data.

Chapter Four:

The Mechanism of UDP-*N*-Acetylglucosamine 2-Epimerase from *Escherichia coli*

4.1 Introduction

In Chapter One, four possible catalytic mechanisms for UDP-GlcNAc 2-epimerase were proposed, each employing a different strategy for removing the relatively non-acidic proton at C-2". The main overall goal of this study was to distinguish between these four possibilities, and by process of elimination determine the most reasonable mechanism for this enzyme that accounts for all observed experimental data. Several key differences between the four pathways pointed towards specific experiments that could be used to probe particular elements of the enzyme's mechanism. For instance, both path A and path C involve the transient oxidation of the C-3" hydroxyl. Both of these routes necessarily involve an enzyme-bound hydride acceptor/donor. A NAD^+ cofactor is the most probable candidate for the role. The presence of this cofactor had been speculated upon in the literature (Salo, 1976), but no effort to detect it had previously been made. Any experiment to determine the presence of NAD^+ requires a significant quantity of purified enzyme, which was not reasonably feasible until the enzyme was cloned and overexpressed. With the large amounts of purified recombinant protein now available, it was possible to perform two different experiments that clearly demonstrate that UDP-GlcNAc 2-epimerase contains no bound NAD^+ cofactor.

A second approach to distinguish the four possible mechanisms would be to probe the fate of the anomeric carbon-oxygen bond during the course of the reaction. Pathways B and C both involve cleavage of the anomeric carbon-oxygen bond, generating enzyme-associated intermediates. In the case of path B the proposed intermediates are 2-acetamidoglucal and UDP, while path C would afford 2-acetamido-3-ketoglucal and UDP. Conversely, path A and path D retain an intact anomeric carbon-oxygen bond throughout the reaction. The result of

the positional isotope exchange experiment, described in section 4.4, provides strong evidence that the anomeric bond is cleaved during the enzymatic epimerization.

These two experiments together support the elimination/readdition mechanism outlined in path B. This mechanism is unprecedented among the epimerase and racemase family of enzymes. To further support this proposed mechanism, an experiment was conducted to test the enzyme's ability to catalyze the second half of the reaction – the addition of UDP to 2-acetamidoglucal to form an equilibrated mixture of UDP-GlcNAc and UDP-ManNAc. The results of this attempt to demonstrate that UDP and 2-acetamidoglucal are reaction intermediates are described in section 4.5.

The experiments described in this chapter provide a detailed view into the workings of the unique mechanism employed by UDP-GlcNAc 2-epimerase. The accumulated evidence makes it possible to describe the minimal reasonable reaction pathway for the enzymatic epimerization. However, many questions remain concerning the intimate details of the enzyme's mode of action, and a few of the future directions of study with the enzyme are discussed at the close of the chapter.

4.2 Solvent-Derived Deuterium Incorporation at C-2"

In the only prior study of the mechanism employed by UDP-GlcNAc 2-epimerase, Salo reported the incorporation of a solvent derived isotopic label when the bacterial enzyme catalyzed the reaction in tritium-enriched water (Salo, 1976). In order to confirm this observation with the purified recombinant epimerase, UDP-GlcNAc was incubated with the

enzyme in phosphate buffer prepared in D_2O, and the reaction was followed by ^1H NMR (Figure 4.1).

In the ^1H NMR spectrum of UDP-GlcNAc, the signal corresponding to the C-2" proton is obscured by the signals arising from the protons of the ribosyl moiety of the nucleotide. As a result, it is more practical to consider the distinct multiplet signal centered around 5.48 ppm in the UDP-GlcNAc ^1H NMR spectrum corresponding to the anomeric proton at C-1". The anomeric proton is coupled to both the proton at C-2" and the β-phosphorus of UDP, and therefore appears in the initial spectrum as a doublet of doublets.

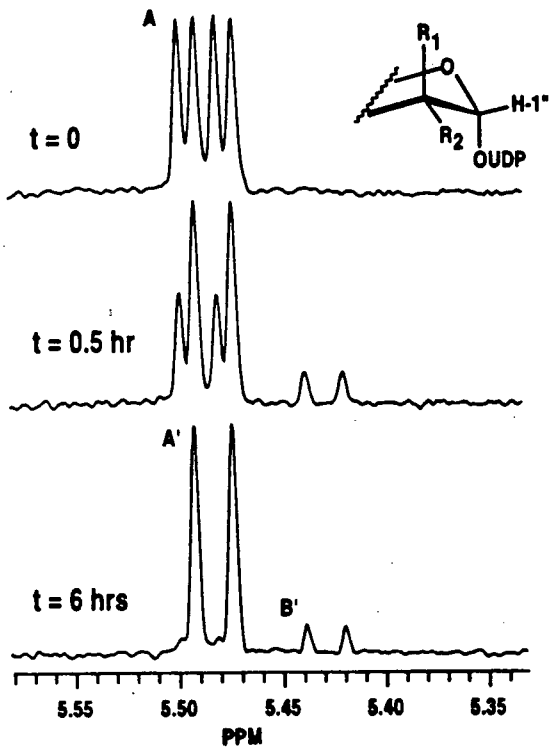


Figure 4.1 Expansion of the ^1H NMR spectrum of UDP-GlcNAc showing the anomeric proton signals during enzymatic epimerization in D_2O. Species A = UDP-GlcNAc (R_1 = H, R_2 = NHAc); species A' = deuterated UDP-GlcNAc (R_1 = D, R_2 = NHAc); species B' = deuterated UDP-ManNAc (R_1 = NHAc, R_2 = D).

As shown in Figure 4.1, over the course of the reaction in D_2O, the signal for the anomeric proton clearly collapses to a doublet, indicating that the proton at C-2" has been replaced by a solvent derived deuterium atom. The expansion of the ^1H NMR region around the anomeric proton signal also reveals the appearance of a new doublet, approximately one-tenth the size of the more prominent UDP-GlcNAc anomeric peak. This new signal is due to the anomeric proton of 2"-^2H-UDP-ManNAc also present in the equilibrated mixture.

The incorporation of a single deuterium into UDP-GlcNAc has also been demonstrated in the LSI mass spectrum taken of 2"-^2H-UDP-GlcNAc recovered after the enzymatic epimerization was run in D_2O, which clearly shows an increase of one mass unit.

These results confirm the previously reported observation that UDP-GlcNAc 2-epimerase catalyzes the solvent derived deuterium incorporation at C-2". This result supports a mechanism involving breakage of the C-2" carbon-hydrogen bond via deprotonation. All of the four proposed pathways follow this notion. However, as mentioned in Chapter One, a scheme involving direct hydride transfer can be immediately ruled out.

4.3 Test for Bound NAD^+

Two of the mechanistic pathways proposed for UDP-GlcNAc 2-epimerase require NAD^+ for the transient oxidation at C-3". To support or discount these two pathways, it was required to test for the presence of any such cofactor tightly bound in the enzyme's active site. This was accomplished through two different experiments. The first method analyzed the native enzyme directly for a chromophoric NADH transiently generated during the course of

the reaction. The second experiment was carried out by first proteolytically digesting the epimerase before assaying the resultant digest for free NAD^+.

What makes these experiments possible is the fact that NADH has a characteristic absorption at 340 nm, while NAD^+ does not. This difference in the chromophoric properties of the oxidized and reduced forms of the cofactor is commonly exploited to monitor enzymatic reactions, such as in the coupled assay described earlier in Chapter Two. It can also be used to detect a transiently generated NADH cofactor, but since only one cofactor is present per active site, a large amount of enzyme is necessary to ensure NADH is present at reasonably detectable levels.

The presence of an NAD^+ cofactor has been successfully demonstrated through the UV spectrum of the enzymes UDP-galactose 4-epimerase (Wilson and Hogness, 1964), and *S*-adenosylhomocysteine lyase (Palmer and Abeles, 1979). In the absence of substrate, these two enzymes both exist in the oxidized NAD^+ form. Addition of substrate causes an increase in the absorbance around 340 nm, which is attributed in both cases to the production of NADH in the active site during the catalytic reaction.

In a similar fashion, the UV absorption spectrum was taken of a large quantity of UDP-GlcNAc 2-epimerase. As expected for the resting enzyme, the spectrum showed no significant absorption above 310 nm that might be attributed to a chromophoric cofactor (Figure 4.2). After addition of saturating amounts of UDP-GlcNAc, the enzyme showed no significant change in the spectrum above 310 nm. If it were employing an NAD^+ cofactor, in the presence of substrate, the active enzyme should shuttle between oxidized and reduced forms at a maximal rate of 4.8 turnovers per second (k_{cat}). Therefore, if a mechanism involving an NAD^+ cofactor were employed here, it might have been possible to observe a

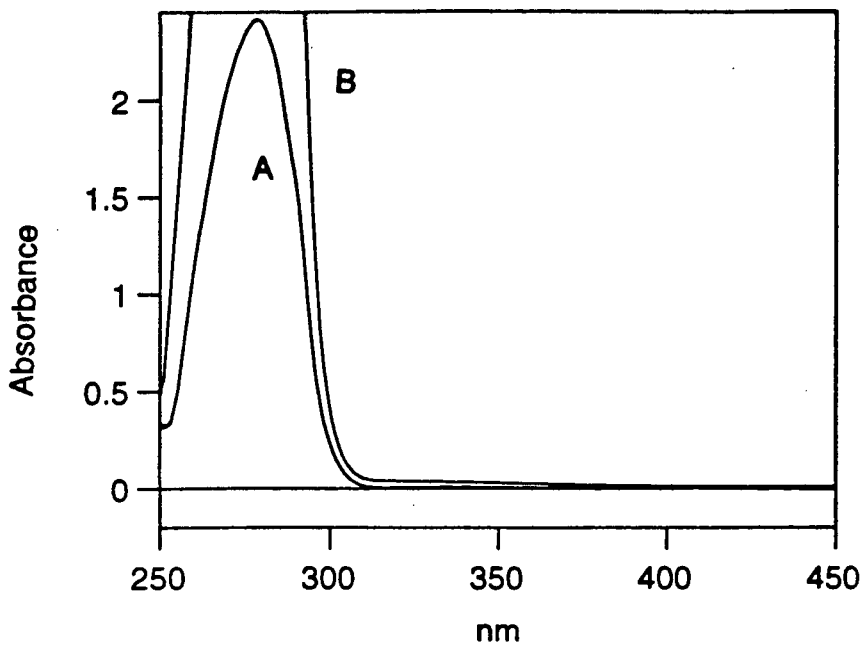


Figure 4.2 UV-visible spectra of UDP-GlcNAc 2-epimerase: (A) Enzyme alone, and (B) Enzyme in the presence of 12 mM UDP-GlcNAc.

change in the absorbance spectrum at 340 nm (Palmer and Abeles, 1979). This would depend on the internal thermodynamics controlling the relative amounts of the two enzyme forms.

In the second experiment, a large sample of the enzyme was proteolytically digested at 37°C overnight with both trypsin and chymotrypsin in order to release any bound NAD^+, according to the method of Palmer and Abeles (1976). The digested enzyme pool was assayed for the presence of NAD^+ using an enzymatic assay system that requires ethanol, semi-carbazide and horse liver alcohol dehydrogenase, to quantitatively generate NADH from NAD^+ according to the reaction illustrated in Figure 4.3. If NAD^+ were present, the UV

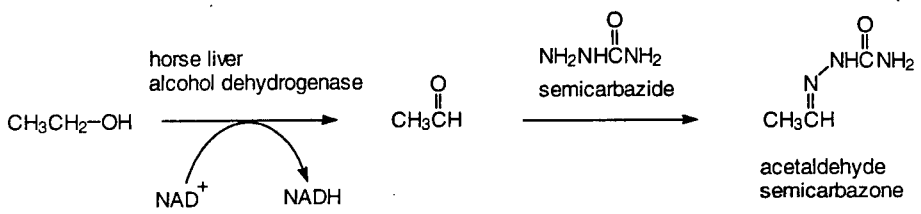


Figure 4.3 NADH is quantitatively generated from NAD^+ using an assay that requires ethanol, horse liver alcohol dehydrogenase and semicarbazide.

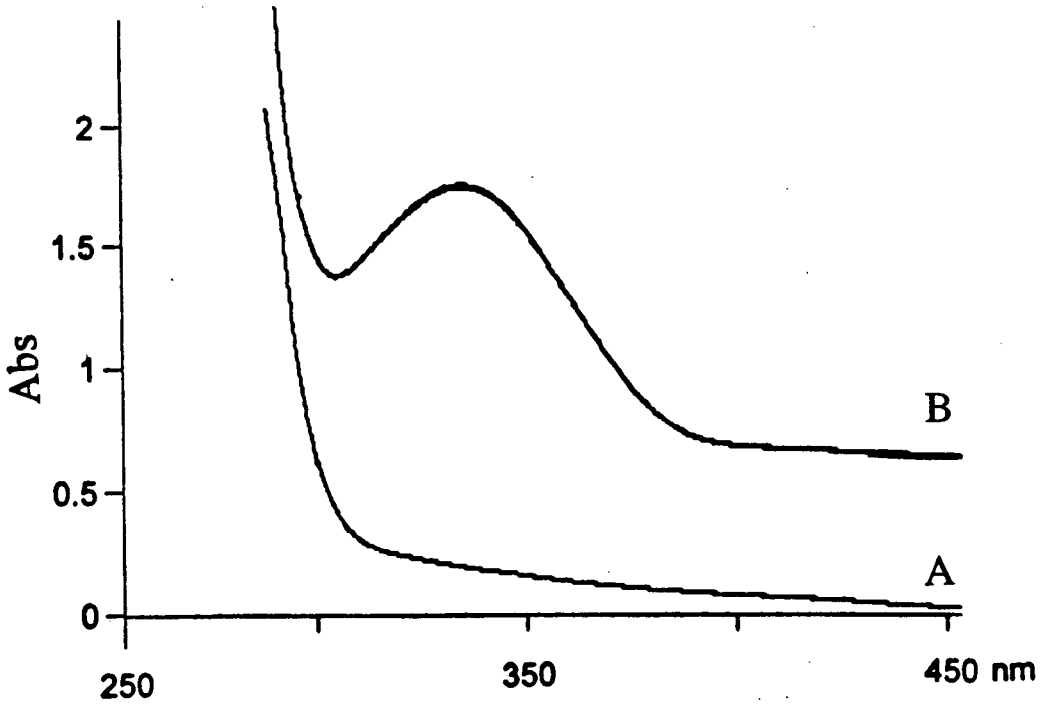


Figure 4.4 UV-visible spectra from the test for NAD^+ released from preoteolyzed UDP-GlcNAc 2-epimerase: (A) 10 mg of epimerase is digested and assayed for the presence of NAD^+ using the assay in Figure 4.3, and (B) 10 mg of bovine serum albumin and 1 equivalent NAD^+ (1 eq. with respect to 10 mg of epimerase) are treated identically as a control to show the expected spectrum. (Spectrum B is offset by 0.75 Absorbance units.)

spectrum would show a significant absorption at 340 nm, corresponding to the generated NADH. The expected spectrum is clearly demonstrated by the control sample, in which a comparable amount of bovine serum albumin was digested in lieu of UDP-GlcNAc 2-epimerase, and one equivalent of NAD^+ (one equivalent with respect to the amount of epimerase present in the actual sample) was added to the assay reaction (Figure 4.4, curve B). However, using this method, the presence of NAD^+ was not observed in the digested UDP-GlcNAc 2-epimerase samples (Figure 4.4, curve A).

The fact that neither experiment provided evidence to support an NAD^+ associated with the enzyme provides a convincing argument against any reaction mechanism requiring that cofactor. This includes both the pathway originally proposed by Salo (1976), described in this thesis as path A, and the conjugate-addition pathway, path C.

The failure to observe NAD^+ associated with the epimerase is not too surprising in terms of the epimerase amino acid sequence. A comparison of the X-ray crystal structures of NAD^+ utilizing enzymes has revealed a common compact fold in the three-dimensional structure at the N-terminal region of these proteins, consisting of two parallel β-sheets associated with a single α-helix (Figure 4.5). This βαβ-fold is responsible for binding the ADP portion of the NAD^+ cofactor (Eventoff *et al.*, 1977). In particular, eleven specific amino acid residues were determined to be required within the structure of the binding motif in order for the proper geometry to be obtained by the protein chain (Table 4.1; Wierenga *et al.*, 1985, 1986).

Although this "fingerprint" is a general guideline for the amino acids expected in the NAD^+ binding region, it is by no means rigorous. In fact, the total length of the binding motif

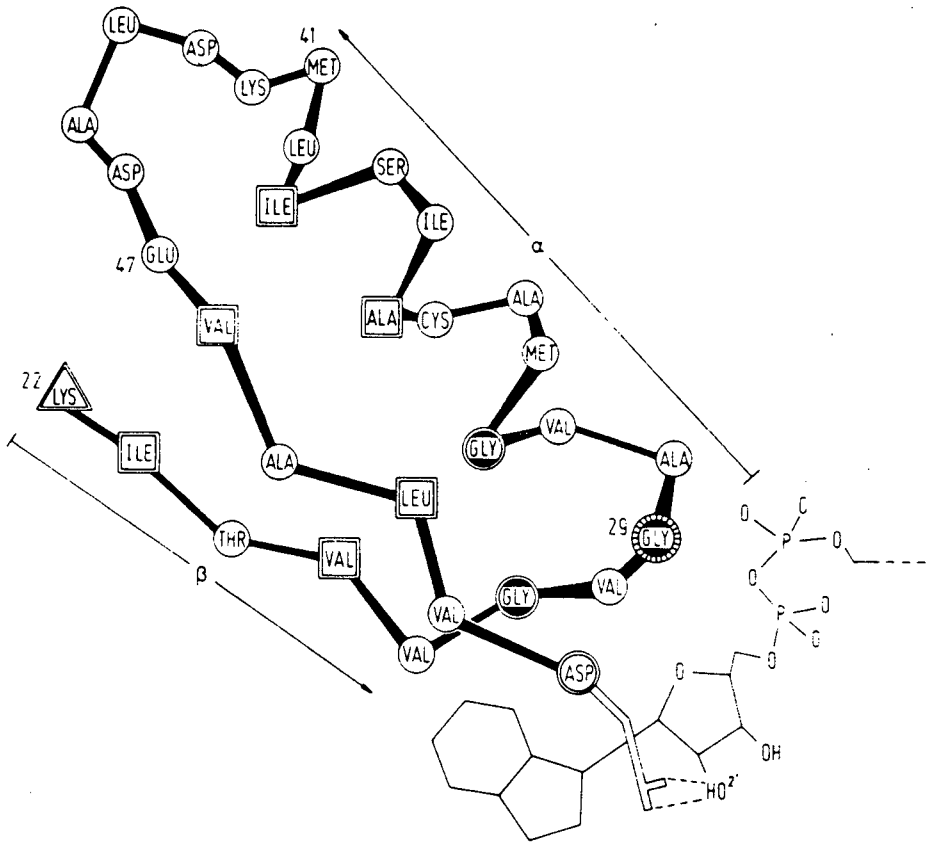


Figure 4.5 Schematic drawing of the βαβ-fold from spiny dogfish M-lactate dehydrogenase highlighting the eleven fingerprint residues (from Wierenga et al., 1986).

Table 4.1 The residues of the NAD^+ binding site fingerprint (from Wierenga *et al.*, 1986).

Position in amino acid sequence	Residues obeying the fingerprint	Comment
First β-sheet:		
1	K,R,H,S,T,Q,N	Basic or hydrophilic
2	A,I,L,V,M,C	Small and hydrophobic
4	A,I,L,V,M,C	Small and hydrophobic
α-helix:		
6	G	Glycine
8	G	Glycine
11	G	Glycine
15	A,I,L,V,M,C	Small and hydrophobic
18	A,I,L,V,M,C	Small and hydrophobic
Variable Loop Region		
Second β-sheet:		
28	A,I,L,V,M,C	Small and hydrophobic
30	A,I,L,V,M,C	Small and hydrophobic
32	D,E	Acid

is not entirely consistent from enzyme to enzyme, on account of the variable-length loop region that connects the α-helix with the second β-sheet. As a result, without prior knowledge of the exact length of the loop itself (which can only be determined from the X-ray crystal structure) it can be difficult to match the fingerprint residues that occur after the loop, in the second β-sheet. Furthermore, the eleven-residue fingerprint is not absolute, and can tolerate a few substitutions from the predicted pattern. However, at the bare minimum, the three glycine residues near the beginning of the α-helix strand are considered to be the essential core of the fingerprint, and are strictly required for the proper fold in the protein backbone. In addition to the three glycines, the final amino acid of the fingerprint must be an acidic residue, which as the crystal structures have shown, forms a hydrogen bond with the 2'-OH of the ADP-ribose in NAD^+.

The amino acid sequence of UDP-GlcNAc 2-epimerase does not contain any of the features predicted for the NAD^+ binding motif. In fact, the required core sequence of three glycines is entirely absent in the epimerase sequence (rffE in Table 4.2).

Table 4.2 The first forty N-terminal amino acid residues of UDP-ManNAc dehydrogenase (RffD), UDP-galactose 4-epimerase (GalE), and UDP-GlcNAc 2-epimerase (RffE). The eleven fingerprint positions of the NAD^+ binding motif are underlined. Residues matching the fingerprint are indicated in bold.

RffD	M S F A **T** **I** S **V** I **G** L **G** Y I **G** L P T **A** A A F A S R Q K Q **V** I G V **D** I N Q H A V D
GalE	M R V **L** V T G **G** S **G** Y I **G** S H T C V Q **L** L Q N G H D **V** I **I** L **D** N L C N S K R S V
RffE	M K V L T V F G T R P E A I K M A P L V H A L A K D P F F E A K V C V T A Q H R

For the sake of comparison, it is instructive to consider other proteins known to require NAD^+. The X-ray crystal structure of UDP-galactose 4-epimerase from *E. coli* has been solved and refined to a high resolution (Bauer *et al.*, 1992; Thoden *et al.*, 1996; Thoden *et al.*, 1997). Despite the fact that this bacterial enzyme has little amino acid sequence homology with other NAD^+ binding proteins, UDP-galactose 4-epimerase folds into a structure remarkably similar to the three-dimensional structure common to dehydrogenases. This recurring structure consists of two distinct domains, one involved with NAD^+ binding and the other in substrate binding. The sequence of UDP-galactose 4-epimerase does conserve the requisite three glycine core structure implicated in maintaining the proper fold in this NAD^+ binding domain (GalE in Table 4.2). In total, the 4-epimerase sequence matches with a total of nine out of the eleven predicted fingerprint residues. Furthermore, the presence of the requisite folding pattern can be confirmed by the known crystal structure of the enzyme, which clearly shows a hydrogen bond between the 2'-OH of the ADP ribose and aspartate-31, as predicted by the fingerprint.

Although its X-ray crystal structure is not known, it is likely that the other enzyme involved in this study, UDP-ManNAc dehydrogenase, also contains the common NAD^+-binding domain, since NAD^+ is one of its substrates. The sequence of the dehydrogenase matches the first seven positions of the eleven-residue fingerprint exactly, including the three core glycine residues (RffD in Table 4.2). Analogous to UDP-galactose 4-epimerase, it also seems reasonable that aspartate-33 contributes the obligatory hydrogen-bond to NAD^+. If, by crystallographic evidence this proves to be correct, the dehydrogenase has only two mismatches, found at position eight (phenylalanine in lieu of a smaller hydrophobic residue) and at position 30 of the fingerprint where a glycine is present. Both replacements for the

expected residues however do not seem completely unreasonable. In total, UDP-ManNAc dehydrogenase apparently matches nine out of the expected eleven, a score observed in the known crystal structures of several dehydrogenases (Wierenga *et al.*, 1986).

4.4 Positional Isotope Exchange Experiment

Another key distinction between several of the proposed mechanisms is the fate of the anomeric carbon-oxygen bond during the course of the reaction. An experiment that can provide evidence regarding the cleavage of the anomeric carbon-oxygen bond in UDP-GlcNAc is the positional isotope exchange (PIX) experiment (Raushel and Villafranca, 1988). This experiment was run in our lab by Rafael Sala at the same time the work described in this thesis was carried out, and is presented here for the sake of continuity, and to convey the chronological order associated with the experiments conducted.

In the proposed experiment, a single ^{18}O label is positioned in the labile carbon-oxygen bond prior to the reaction (i.e. the anomeric position of UDP-GlcNAc) where it would have a P-O bond order of 1 (Figure 4.6). If that C-O bond were broken (as in, for instance, our proposed pathway B), the intermediate phosphate would have three chemically equivalent oxygens with which it could reform the carbon-oxygen bond. Statistically, the C-^{18}O would reform one third of the time, while a new C-^{16}O bond would form the remaining two-thirds of the time. In the latter case, the ^{18}O-label will have scrambled out of the bridging position where the P-^{18}O bond had an order of 1, and into a position with a P-^{18}O bond order of 1.5. As long as the lifetime of the intermediate was sufficient to allow free rotation of the

phosphate, and rotation wasn't otherwise hindered within the active site, the statistical distribution of two non-bridging to one bridging ^{18}O-label would be observed.

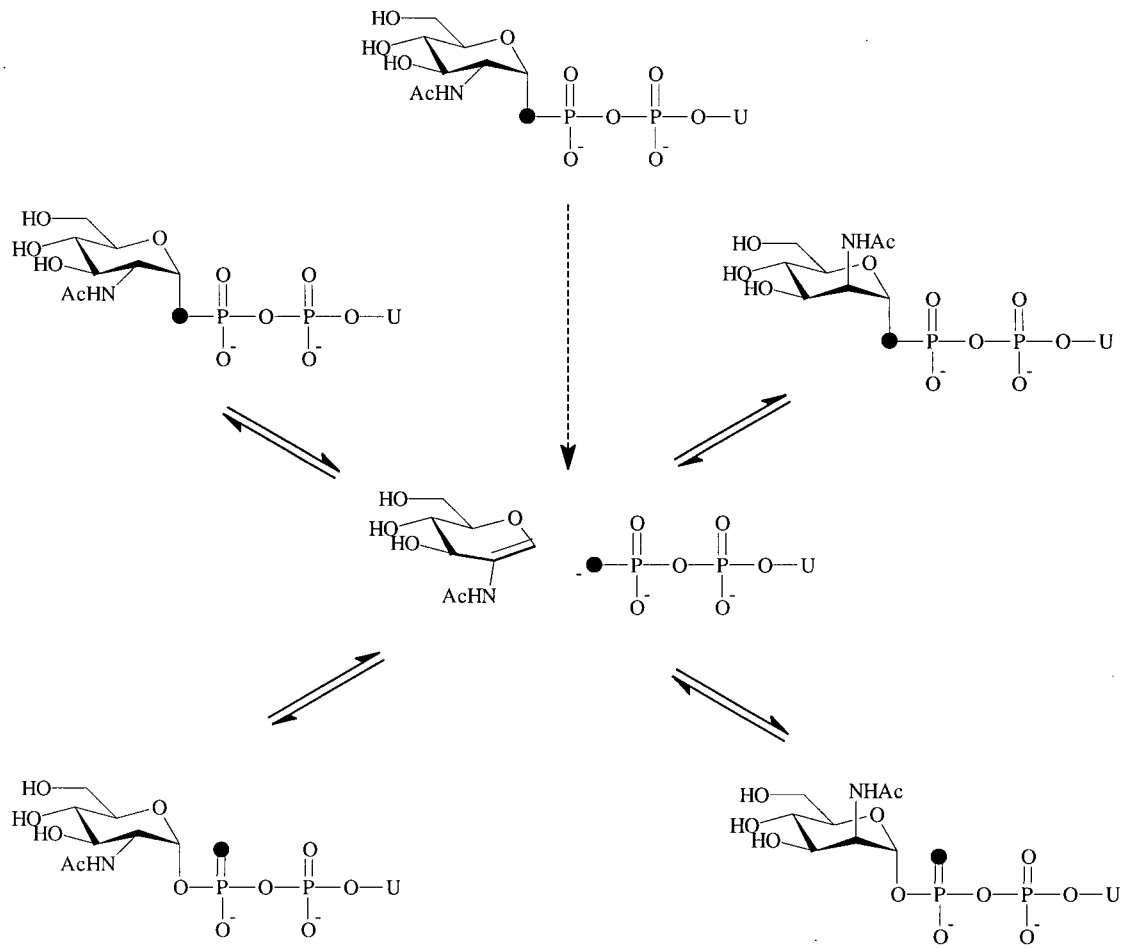


Figure 4.6 A positional isotope exchange experiment that can test for cleavage of the anomeric C-O bond during the epimerization reaction. Initial: ^{18}O-labeled UDP-GlcNAc (before dotted arrow). Final: the ^{18}O label is scrambled into a statistically equilibrated mixture of UDP-GlcNAc and UDP-ManNAc labeled in both bridging and non-bridging positions (after dotted arrow). U=uridine; the darkened atom indicates the ^{18}O label.

The isotopic substitution of ^{18}O for ^{16}O in phosphates causes a small upfield shift of the signals in the ^{31}P NMR spectrum, the magnitude of which depends on the P-O bond order (Hester and Raushel, 1987). An ^{18}O label singly bonded to a phosphorus atom results in a 0.013 ppm upfield shift with respect to an unlabeled P-^{16}O single bond. The scrambling of the ^{18}O label into a higher bond order position shifts the signal further upfield still.

A sample of UDP-GlcNAc containing ^{18}O isotope at the anomeric position was synthesized and incubated with UDP-GlcNAc 2-epimerase. The ^{31}P NMR spectrum of the equilibrated pool demonstrated that ^{18}O did indeed scramble from the bridging position, and into a new position with a higher P-O bond order (Morgan *et al.*, 1997). In the NMR spectrum, a new doublet appears, shifted 0.029 ppm upfield from the unlabeled material, and with an integrated area that shows the expected 2:1 ratio reflecting the statistical distribution of the label between the bridging and non-bridging positions.

The observation of the positional isotope exchange supports the notion that the anomeric C-O bond is cleaved during the enzymatic reaction. While paths B and C both follow suitable reaction schemes, of the two, only path B does not utilize an NAD^+ cofactor. Therefore, path B remains the only likely pathway that accounts for all the experimental observations.

4.5 Test for 2-Acetamidoglucal and UDP as Reaction Intermediates

To further test the validity of path B as the most reasonable enzymatic mechanism, UDP-GlcNAc 2-epimerase was probed for its ability to catalyze the addition of 2-acetamidoglucal and UDP, the two proposed reaction intermediates for this pathway. If these

two compounds are indeed intermediates, the enzyme would potentially be capable of accepting both as substrates for the second half of the reaction pathway.

As discussed earlier in Chapter One, mechanistic studies with the hydrolyzing mammalian enzyme suggested that that enzyme likely proceeded through the same intermediates. With the mammalian enzyme however, rather than completing the epimerization by adding UDP back to the glycal, it catalyzed the stereospecific hydration of 2-acetamidoglucal to form ManNAc (as opposed to GlcNAc).

With this in mind, an experiment was devised to test whether the bacterial UDP-GlcNAc 2-epimerase is capable of catalyzing the addition of UDP to 2-acetamidoglucal. Unfortunately, the approach to this experiment is not as straightforward as simply incubating the epimerase with UDP and 2-acetamidoglucal. Three considerations prevent this direct approach. First, since the epimerase has an absolute requirement for the presence of UDP-GlcNAc as an allosteric activator, the enzyme would remain inactive in its absence, unable to catalyze the second half of the reaction, as desired. As a result, this requisite presence of the natural substrate would have the normal epimerization reaction competing with the intended intermediate-coupling reaction. Second, it had previously been reported that UDP is a competitive inhibitor of the enzyme, with 1 mM of UDP reducing the enzyme efficiency to 50% of its maximal rate (Kawamura *et al.*, 1979). In the previously discussed example of UDP-galactose 4-epimerase, it was shown that the UDP-moiety of the substrate binds to the active site strongly, while the sugar-moiety binds only weakly (Kang *et al.*, 1975; Ketley and Schellenberg, 1973; Seyema and Kalckar, 1972; Frey, 1989). If this binding preference were extended to UDP-GlcNAc 2-epimerase, this would suggest that while UDP has a strong affinity for the 2-epimerase (since it is a competitive inhibitor), the non-polar 2-

acetamidoglucal does not. If UDP binds to the active site first, and in such a way as to exclude 2-acetamidoglucal from accessing the active site, it seems reasonable that UDP itself might inhibit the coupling reaction as well. Third, an experiment with the 4-epimerase has shown that UMP and a variety of ketohexoses can concomitantly bind to that enzyme, with the ketohexose binding in a variety of orientations (Kang *et al.*, 1975; Ketley and Schellenberg, 1973; Seyema and Kalckar, 1972). This might also be the case for UDP and 2-acetamidoglucal. Only one proper binding orientation of the glycal however, would actually result in the desired addition reaction.

Anticipating all these factors, the experiment was designed with the intention of reducing the competing events as much as possible. The incubated samples contained as little UDP-GlcNAc as possible in order to reduce the competing reaction, yet high enough quantities to maintain sufficient amounts of the epimerase in its active conformation. The concentration of UDP was selected such that it would inhibit the enzyme no more than 50% (with respect to the normal reaction). Significantly more 2-acetamidoglucal than either UDP or UDP-GlcNAc was added in order to make the binding of the non-polar sugar molecule more favorable, and drive the coupling reaction. Finally, long incubation times and large quantities of epimerase were required to ensure that a detectable amount of the addition reaction had occurred.

In the actual experiment, a synthetic sample of 2-acetamidoglucal (2 mM) was incubated with 0.3 units of UDP-GlcNAc 2-epimerase, 0.6 mM UDP-GlcNAc and 0.9 mM ^{14}C-labeled UDP. Control samples contained either heat-killed enzyme, or no glycal. After a four hour incubation at 37°C, the samples were each injected onto a reversed-phase HPLC column, and the components were separated using the ion-pair reversed-phase HPLC

technique described by Meynial *et al.* (1995). This sensitive technique uses tetrabutylammonium hydrogen sulfate (TBAHS) in the elution buffer in order to retain the charged species on the reversed-phase column, while eluting with a linear gradient of acetonitrile. This method allows the separation of the two epimers, UDP-GlcNAc and UDP-ManNAc, both of which have a retention time significantly distinct from UDP. The eluent is monitored at 262 nm, which is the wavelength of maximal absorbance for the uridine chromophore. Interestingly, at this wavelength a small peak corresponding to the eluting 2-acetamidoglucal is observed, at a retention time that indicates the non-polar molecule is only slightly retained by the reverse phase column. It appears that the enamine chromophore of the glycal has a detectable absorbance at 262 nm.

Thus, with the equilibrated components separated by HPLC, each separate fraction was counted for radioactivity. All detectable ^{14}C-label remained present in the UDP pool, with essentially no radioactivity detected above background in either the UDP-GlcNAc or the UDP-ManNAc fractions. This indicated that the coupling did not occur under these conditions to any significant extent. Interestingly, the control sample lacking 2-acetamidoglucal showed a very small peak at the retention time where the glycal normally eluted.

A second attempt was performed to confirm the results of the prior experiment. This time however, in light of the fact that the epimerase did not catalyze the glycal/UDP addition reaction to any appreciable extent within four hours, the samples of this second trial were allowed to react for 24 hours under identical incubation conditions. These samples also contained 10 mM of 2-acetamidoglucal to further drive the equilibrium of the coupling reaction. All other concentrations were kept the same. In addition to the repeated samples

and controls, one new sample was also prepared containing 5-fold the quantity of UDP-GlcNAc 2-epimerase previously used (for a total of 1.5 units of enzyme), and this sample was similarly incubated for the same extended period.

Like the first set of samples, the samples of this trial were analyzed by ion-pair reversed-phase HPLC, and the levels of radioactivity in each fraction were determined. Again, all detectable ^{14}C-label remained in the UDP pool, indicating that no appreciable UDP/2-acetamidoglucal addition had occurred. However, the intriguing observation persisted that the "no glycal" control sample had gained appreciable amounts of 2-acetamidoglucal. This time however, after the extended incubation, the glycal had a greater peak area than UDP-GlcNAc and UDP-ManNAc, despite a clearly lower extinction coefficient, implying that new glycal had formed from UDP-GlcNAc. More dramatically, in the sample containing the 5-fold quantity of epimerase, the UDP-GlcNAc peak had all but disappeared, and was clearly dwarfed by the peaks corresponding to 2-acetamidoglucal and UDP. Remarkably, under the conditions of these extended incubations, it appeared as though UDP-GlcNAc 2-epimerase was producing more of the two intermediates, rather than consuming them. HPLC traces following the time course of this conversion are shown in Figure 4.7.

To confirm the identity of the new peak, a larger amount of UDP-GlcNAc (38 mM) was incubated in phosphate buffer with 8.8 units of epimerase at 37°C for 12 hours. The resulting sample was applied to a short column of anion-exchange resin, eluted with water, and lyophilized to dryness. Under these conditions, all charged species present in the reaction sample, including UDP-GlcNAc, UDP-ManNAc, and UDP, were expected to be retained by the ion-exchange resin, while the non-charged glycal would wash through. The lyophilized

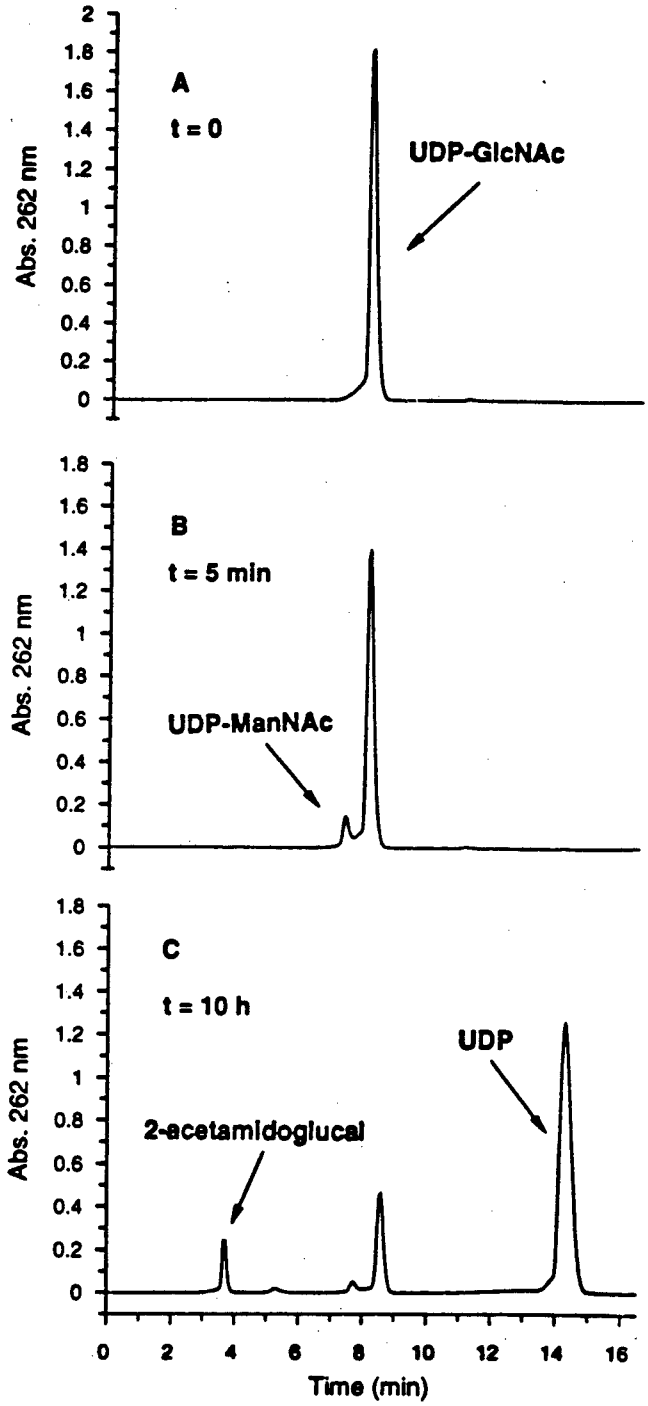


Figure 4.7 Ion-pair reversed-phase HPLC traces obtained during an extended incubation of UDP-GlcNAc with UDP-GlcNAc 2-epimerase, showing (A) substrate, (B) equilibrated UDP-GlcNAc and UDP-ManNAc, and (C) the appearance of UDP and 2-acetamidoglucal.

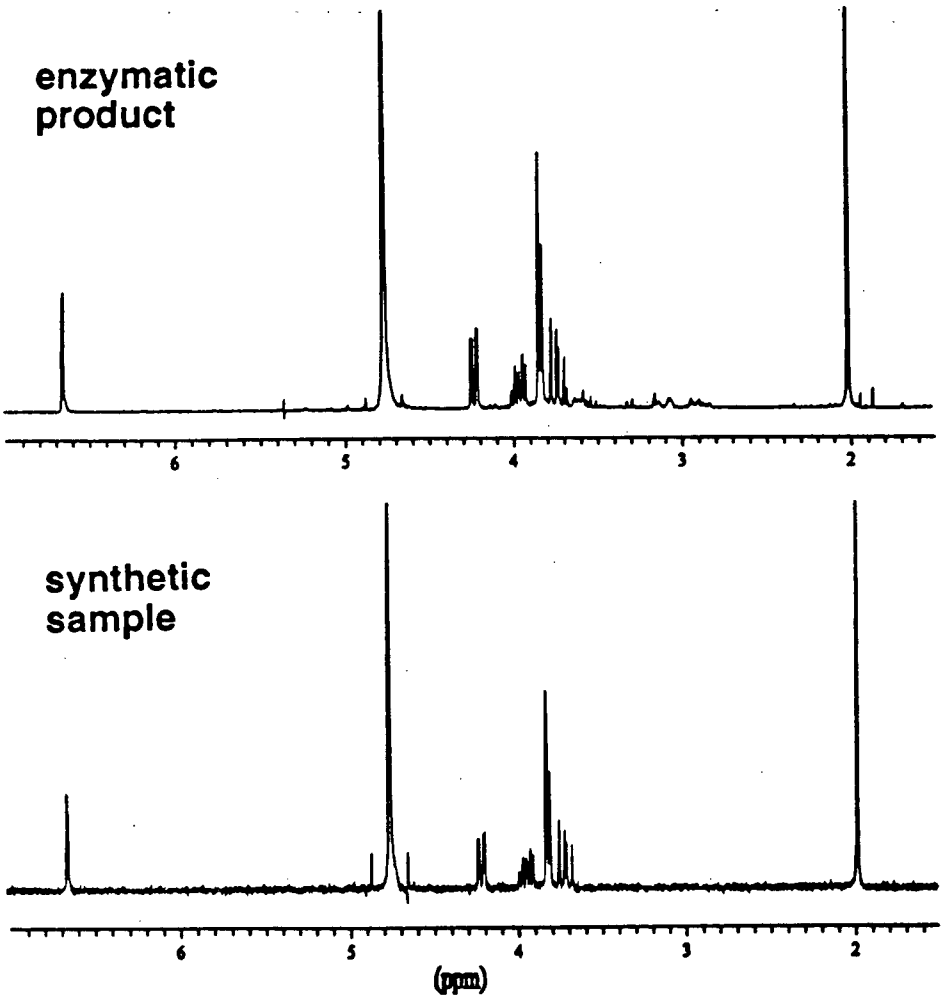


Figure 4.8 ^{1}H NMR sprectra of 2-acetamidoglucal: (A) obtained during an extended incubation of UDP-GlcNAc with UDP-GlcNAc 2-epimerase, and (B) prepared by independent synthesis.

product was redissolved in D_2O, and a ^1H NMR spectrum was acquired (Figure 4.8). This spectrum proved to be identical with the spectrum of an authentic synthetic sample of 2-acetamidoglucal. Furthermore, no peaks from free GlcNAc or ManNAc were observed, indicating that the glycal is not significantly hydrolysed under the experimental conditions.

The observation that UDP and 2-acetamidoglucal accumulate in solution has two implications. First, it strongly supports the notion that they are reaction intermediates in the UDP-GlcNAc 2-epimerase reaction, and they are occasionally released from the enzyme active site. A similar situation was observed with UDP-galactose 4-epimerase in which the 4-keto-hexose intermediate was released from the enzyme, resulting in the inactivation of that enzyme (Frey, 1989). Second, the accumulation of these intermediates implies that the free intermediates are thermodynamically more stable than the equilibrating mixture of UDP-sugar epimers. This thermodynamic preference explains why ^{14}C-UDP was not significantly incorporated into the UDP-sugar pool. The formation of the intermediates is evidently much faster than the reverse reaction.

An attempt to determine the external equilibrium constant was made by incubating a solution of UDP-GlcNAc with a large amount of epimerase. Aliquots were removed at timed intervals over a 48 hour period. Each aliquot was analyzed by ion-pair reversed-phase HPLC. The areas of the peaks attributed to each component were determined (with the concentration of 2-acetamidoglucal taken to be equal to UDP, and the concentration of UDP-ManNAc estimated to be one-tenth that of UDP-GlcNAc). The equilibrium constant was determined from the equation:

$$K_{eq} = \frac{[\text{UDP}]\ [\text{2-acetamidoglucal}]}{[\text{UDP-GlcNAc}]\ [\text{UDP-ManNAc}]}$$

using the ratio of the peaks due to the equilibrated species, and assuming that UDP and UDP-GlcNAc have the same ε_{262}, and that the 2-acetamidoglucal concentration is equal to the concentration of UDP. Unfortunately, the equilibrium lies heavily in favour of UDP and 2-acetamidoglucal, so that the peaks for UDP-GlcNAc and UDP-ManNAc become too small for accurate integration. Furthermore, the epimerase is undoubtedly inactivated during the course of the incubations, since the reduction in the UDP-GlcNAc concentration effectively removes the enzyme's allosteric activator, and a competitive inhibitor (UDP) is produced over time. Using this technique, the equilibrium constant was conservatively determined to exceed 25 000 in favour of the intermediates.

The rate at which the intermediates are released into solution was determined by following the course of the reaction by ion-pair reversed-phase HPLC. The HPLC analysis was calibrated by injecting samples of known UDP concentration, and subsequently comparing the amount injected with the integrated peak area. Actual samples were prepared by incubating saturating amounts of UDP-GlcNAc with enough enzyme to produce the intermediates at an appreciable rate – fast enough that the enzyme did not lose significant activity during the experiment, but slow enough that timed aliquots could be removed at reasonable intervals. The aliquots were removed over the course of 200 minutes and immediately frozen in liquid N_2. In order to approximate initial velocity conditions all measurements were taken prior to the consumption of 10% of the epimeric UDP-sugars. The samples were thawed immediately before injection onto the HPLC. The peak areas were compared to the linear calibration curve to determine the concentration of UDP present in the sample at each time point. The representative HPLC traces of the timed aliquots, which is shown in Figure 4.8, demonstrates the initial equilibration of the two substrate epimers

followed by the subsequent appearance of UDP and 2-acetamidoglucal, as the two intermediates slowly accumulate in solution. The plot of UDP concentration as a function of time was linear (Figure 4.9). From the plot it was determined that the enzyme had a specific activity of 0.017 μmol of UDP generated per min per mg of UDP-GlcNAc 2-epimerase under these conditions. By way of comparison, the specific activity of the enzyme towards the epimerization of UDP-GlcNAc is 6.8 μmol of UDP-ManNAc generated per min per mg of UDP-GlcNAc 2-epimerase.

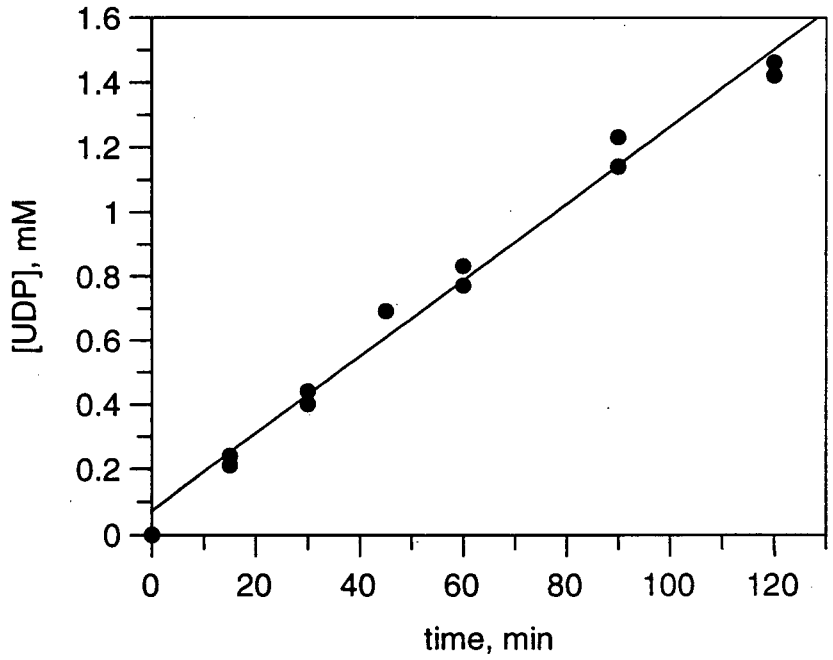


Figure 4.9 A plot of the accumulated concentration of UDP released into solution by UDP-GlcNAc 2-epimerase as a function of time.

4.6 The Measurement of a Kinetic Isotope Effect at C-2"

The incorporation of a solvent-derived deuterium atom at C-2" of UDP-GlcNAc suggests that the mechanism of UDP-GlcNAc 2-epimerase involves the cleavage and

reformation of the C2"-H bond at some point during the catalytic reaction. If this step of the reaction is rate determining, the substitution of a heavier deuterium atom for the proton at C-2" of the substrate can have a profound effect on the energetics of the reaction. Typically, the result of isotopic substitution of a deuterium for a proton is an overall two to fifteen fold slowing in the reaction rate, when the C-D bond is broken or formed during the rate-determining step (Richards, 1970; Walsh, 1979). In a simplified view, this phenomenon arises from the lower vibrational zero-point energy for a bond to deuterium as compared to a bond to hydrogen. The result is a greater amount of energy required for a C-D bond to reach the transition state in a bond cleaving reaction, as compared to a C-H bond. The resulting differences in reaction rates caused by heavy isotope substitution, called kinetic isotope effects, are often used to probe the rate-determining step of a reaction.

The previously described solvent derived isotope incorporation at C-2" allowed the preparation of 2"-^{2}H-UDP-GlcNAc. A sample of UDP-GlcNAc was enzymatically epimerized in D_2O buffer to generate an equilibrated mixture of C-2" deuterated UDP-sugar epimers. The smaller component, UDP-ManNAc, was oxidized to 2"-^{2}H-UDP-ManNAcUA by the addition of UDP-ManNAc dehydrogenase and NAD^+, leaving the 2"-^{2}H-UDP-GlcNAc as the sole remaining epimer. The labeled substrate was isolated by ion-exchange chromatography, by eluting it from a Dowex AG1 X8 column with a linear gradient of 0-2 M triethylammonium bicarbonate. Fractions containing the deuterated substrate were further purified by passage through a column of Amberlite IR-120(plus) resin, followed by a column of Biogel P-2. Both ^{1}H NMR spectroscopic and LSI-mass spectrometric analyses indicated that the substrate was greater than 97% enriched with deuterium at C-2".

The rates of enzymatic epimerization for both the labeled and unlabelled substrates were measured under saturating conditions using the coupled enzyme assay described earlier (Figure 4.10). Both the protonated and deuterated substrate concentrations analyzed were 2.5 mM as determined using A_{262} with ε = 8700 $M^{-1}cm^{-1}$. Under these conditions the V_{max} isotope effect k_H/k_D was determined to be 1.8 ± 0.1. There is clearly a primary isotope effect that slows the epimerization of the deuterated substrate, indicating that the C-2" carbon-hydrogen bond is broken during a rate-determining step of the reaction. The determined value of 1.8 for the kinetic isotope effect lies below the expected range (between 2 to 15) that would

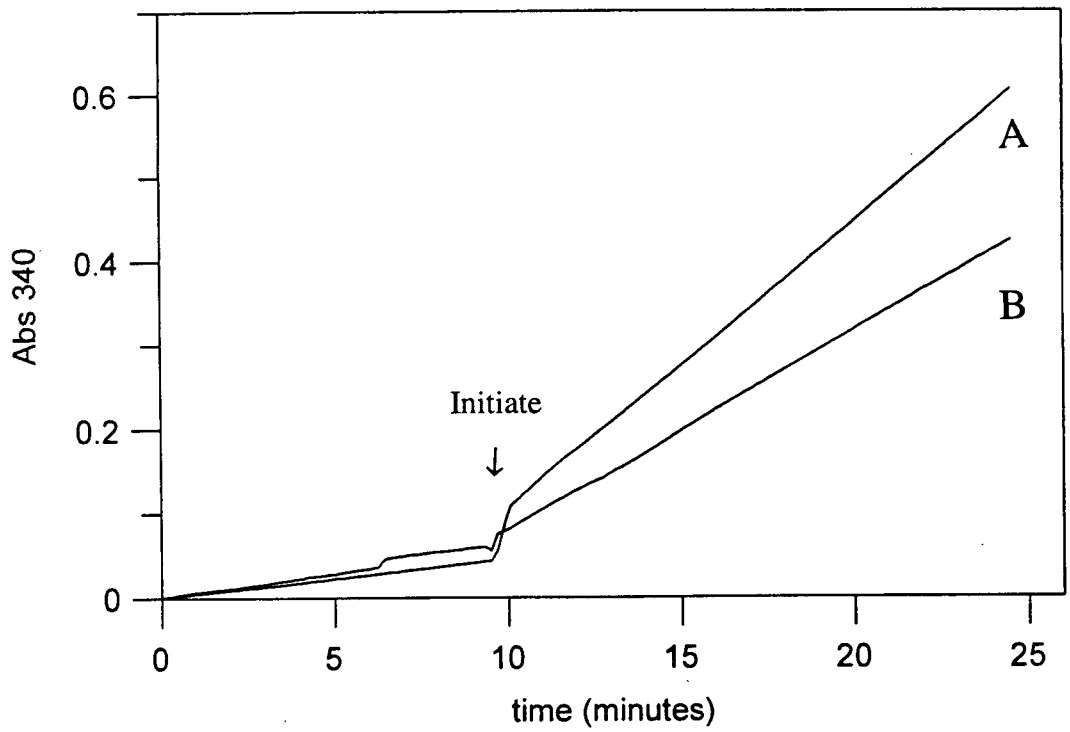


Figure 4.10 The UV traces showing the rate of enzymatic epimerization for UDP-GlcNAc 2-epimerase incubated with (A) 2.5 mM 2"-^{1}H-UDP-GlcNAc, and (B) 2.5 mM 2"-^{2}H-UDP-GlcNAc. Both reactions were initiated after 10 min by the addition of epimerase. Prior to initiation, a substrate-independent background rate is observed due to the impurities present in the dehydrogenase (auxiliary enzyme) preparation.

be expected for a primary isotope effect observed in a mechanism that has the cleavage of the labeled carbon-deuterium bond during the sole rate determining step. The somewhat lower-than-expected value therefore suggests that the C2"-H bond of UDP-GlcNAc is being cleaved during a step in the mechanism that is only partially rate-limiting.

This result is consistent with the β-elimination pathway indicated by the other evidence previously presented in this chapter. The mechanistic implications of this experiment will be discussed within the next section.

4.7 Mechanistic Implications and Future Directions

The absence of any cofactor requirement or a tightly-bound NAD^+ cofactor suggests C-3" oxidation is not involved in the epimerase-catalyzed reaction, and the observation of isotopic scrambling in the UDP-β-phosphate supports a mechanism involving anomeric C-O bond cleavage. The most reasonable minimal description of this enzyme-catalyzed epimerization is therefore path B, the elimination/readdition mechanism. This notion is further supported by the observed release of the proposed reaction intermediates, UDP and 2-acetamidoglucal. This pathway employs an *anti* elimination when UDP-GlcNAc is the substrate, followed by the *syn* readdition of UDP. Conversely, a *syn* elimination occurs with UDP-ManNAc as the substrate, followed by the *anti* readdition of UDP.

Several precedents are known for both *syn* and *anti* eliminations in enzyme mechanisms, and have been recently reviewed (Anderson, 1998). It has been generally recognized that less acidic protons require *anti* eliminations, whereas *syn* eliminations all have adjacent carbonyl groups enhancing the acidity of the abstracted proton. For the most part, the

anti eliminations predominate, as this type is observed throughout the enolase and fumarase superfamily of enzymes. Examples of *syn* eliminations at "activated" centres include muconate lactonizing enzyme, enoyl CoA hydratase, and a few polysaccharide lyases. These reactions are all reversible, the reverse reaction being formally a 1,2-conjugate addition.

Two phosphoenolpyruvate-transferring enzymes, UDP-GlcNAc enolpyruvyl transferase, and 5-enolpyruvylshikimate-3-phosphate synthase, catalyze the addition of phosphoenolpyruvate to an alcohol functional group in their substrates to form tetrahedral intermediates, which are subsequently broken down with the elimination of phosphate. These two enzymes likely proceed through an *anti* addition step, followed by a *syn* elimination (Grimshaw *et al.*, 1984; Lee *et al.*, 1984; Lees and Walsh, 1995; Skarzynski *et al.*, 1998). Since the proton removed is relatively non-acidic, these enzymes are precedents for *syn* elimination at an "unactivated" centre.

Other examples of *anti*-additions are observed in the hydration of glycals by retaining β–glucosidases (Legler, 1990), with the exception of a *syn* addition in the hydration of 2-acetamidoglucal by β–*N*-acetylhexosaminidase (Lai and Withers, 1994). In non-enzymatic acid-catalyzed additions of alcohols to glycals, *syn* products are found to predominate, although *trans* addition products are also observed (Kaila *et al.*, 1992).

The results reported here also parallel the previous observations with the mammalian hydrolysing enzyme, UDP-GlcNAc 2-epimerase, whose mechanism likely proceeds through 2-acetamidoglucal (as discussed in Chapter One). In support of that mechanism, the mammalian enzyme was observed to catalyze the *syn* addition of water to 2-acetamidoglucal. Recently, a bifunctional mammalian enzyme was reportedly cloned, functionally expressed and found to catalyze both the conversion of UDP-GlcNAc to ManNAc, as well as the

subsequent ATP-dependent phosphorylation of ManNAc to ManNAc-6-phosphate (Stasche *et al.*, 1997; Hinderlich *et al.*, 1997). It is likely that this bifunctional UDP-GlcNAc 2-epimerase/ManNAc kinase enzyme was the same enzyme that had been previously isolated and studied (Glaser, 1960; Spivak and Roseman, 1966; Sommar and Ellis, 1972a, 1972b). However, those initial reports did not involve testing the enzyme preparation for kinase activity, which would explain why only the first of the two reactions catalyzed by this enzyme had been noted. The new studies show that the bifunctional enzyme has a single polypeptide chain 75 kDa in size, which associates to form a partially active dimer (150 kDa) or a fully active hexamer (450 kDa). The dissociation of the hexamer could easily account for both the difficulties in prior molecular weight determinations, and the observed diminishment of enzyme activity over time. It is also interesting to note that the first half of the reported amino acid sequence of the bifunctional mammalian enzyme shows 26.8% sequence identity and 56.8% amino acid similarity with the bacterial UDP-GlcNAc 2-epimerase. This would suggest that the epimerization catalyzed by the mammalian 2-epimerase/kinase is most likely localized to a single domain found in the first half of the 75 kDa bifunctional enzyme, and is distinct from a second ManNAc kinase domain. In light of the apparent amino acid sequence homology, it is not surprising that the two enzymes share similar β-elimination mechanisms.

The mechanism of the bacterial UDP-GlcNAc 2-epimerase might also shed some light on another enzyme, CDP-paratose 2-epimerase. This enzyme, which has been identified in *Salmonella typhimurium* and *Yersinia pseudotuberculosis*, catalyzes the epimerization of a similar sugar nucleotide by inverting the stereochemistry at a carbon bearing a non-acidic proton (Figure 4.11; Verma *et al.*, 1989; Thorson *et al.*, 1993). However, CDP-paratose does not have a hydroxyl in a position adjacent to the proton at C-2", and therefore this enzyme

Table 4.3 A comparison of the first forty N-terminal amino acid residues of UDP-GlcNAc 2-epimerase (RffE), and CDP-paratose 2-epimerase (RfbE). The eleven fingerprint positions of the NAD^+ binding motif are underlined. Residues matching the fingerprint are indicated in bold.

RffE	MKVLTVFGTRPEAIKMAPLVHALAKDPFFEAKVCVTAQHR
RfbE	MK**L**_**L**_I_T_G**G**C**G**FL**G**SNL**A**SF**A**LSQGID**L**I_**V**_F_**D**_NLSRKGATD

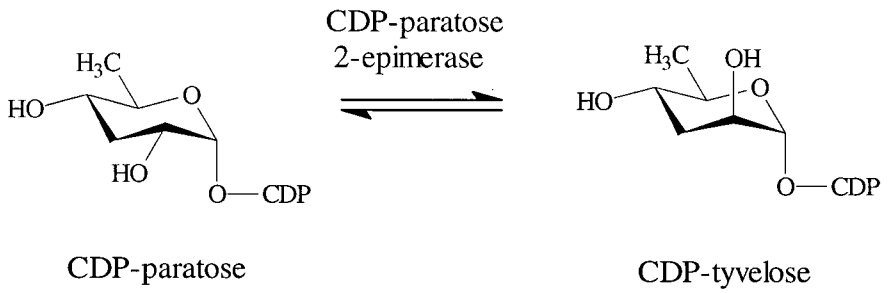
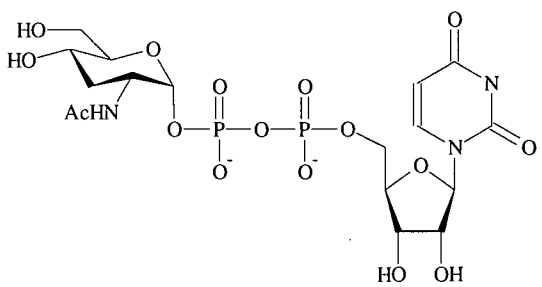


Figure 4.11 The reaction catalyzed by CDP-paratose 2-epimerase.

cannot utilize any mechanism involving transient oxidation at C-3", equivalent to paths A and C described here. It is quite possible that this enzyme proceeds through the β-elimination of CDP in a mechanism analogous to UDP-GlcNAc 2-epimerase. However, a glance at the N-terminal amino acid sequence clearly shows that the enzyme has a potential NAD^+ binding domain, as predicted by the eleven residue fingerprint (Table 4.3). It would be interesting if CDP-paratose 2-epimerase and UDP-GlcNAc 2-epimerase do share the same mechanism,

since these two enzymes do not share any obvious amino acid sequence homology, and have therefore evolved independently. Without any additional mechanistic information, however, it also seems plausible that CDP-paratose 2-epimerase might catalyze a direct hydride transfer mechanism akin to UDP-galactose 4-epimerase.

The substrate for CDP-paratose 2-epimerase did however suggest a substrate analogue for UDP-GlcNAc 2-epimerase, in the form of 3"-deoxy-UDP-GlcNAc, **1**. Unfortunately though, studies performed by Rafael Sala (in our lab) with this substrate demonstrated that despite the observation that UDP-GlcNAc 2-epimerase did bind the 3"-deoxy substrate-homologue (since that substrate did act as a competitive inhibitor of the epimerase), the 3"-deoxy substrate did not turnover, even in the presence of sufficient UDP-GlcNAc to maintain the enzyme in its active conformation (Morgan *et al.*, 1997). Based on the accumulated mechanistic evidence it would be expected that the enzyme could still epimerize the 3"-deoxy substrate through the usual β-elimination. Apparently the 3"-hydroxyl is required in the substrate for activity, which might indicate that this group is necessary for conformational, or electronic properties in the molecule itself, or it is a requisite for a crucial binding interaction between UDP-GlcNAc and the active site pocket. Further studies, including incubations of



1

the 3"-deoxy-UDP-GlcNAc with the epimerase and UDP-ManNAc also resulted in no detectable traces of equilibration of UDP-ManNAc to UDP-GlcNAc, indicating that the removal of the 3"-hydroxyl from UDP-GlcNAc also abolishes the ability of the sugar nucleotide to activate the epimerase.

Although the results within this thesis support the β-elimination pathway, the class of β-elimination reaction that is operable, E1cb, E1 or E2, has yet to be determined (Figure 4.12). The first of these mechanisms is the E1cb, in which deprotonation occurs first, affording a carbanion intermediate. There is ample precedent for both *syn* and *anti* E1cb mechanisms in

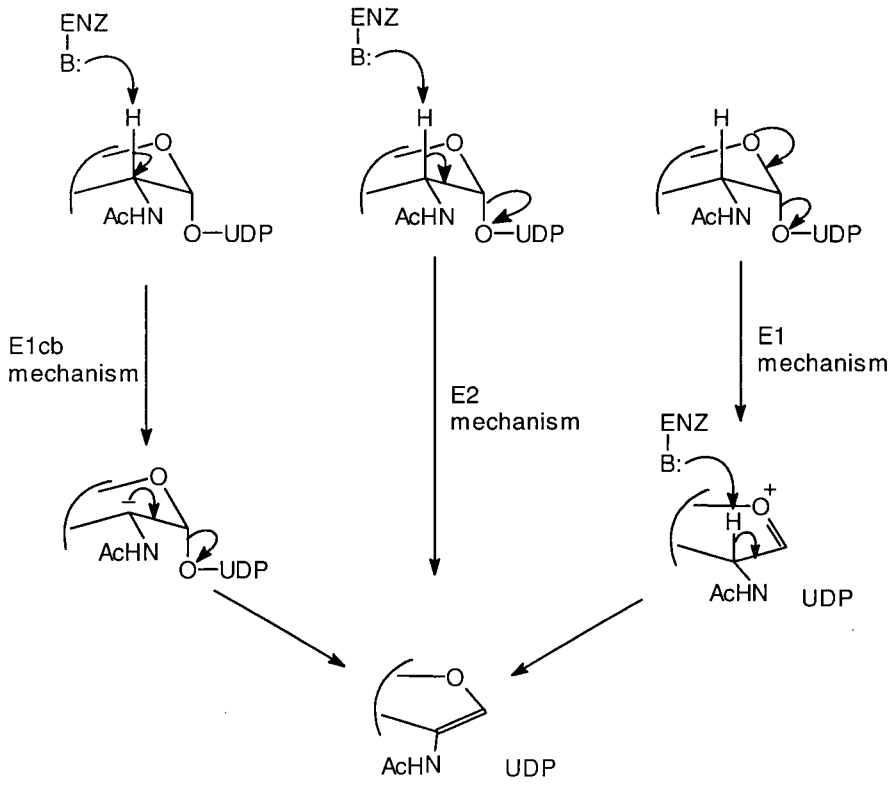
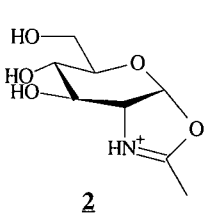


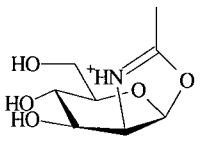
Figure 4.12 The three possible mechanisms for the β-elimination of UDP from UDP-GlcNAc.

enzymatic reactions; however, all cases have an electron withdrawing functionality present in the intermediate in order to stabilize the generated carbanion (Anderson, 1991; Creighton and Murthy, 1990, Anderson, 1998). Since UDP-GlcNAc has no such stabilization available, the pKa value of the C-2" proton is very high, and it seems unlikely that UDP-GlcNAc 2-epimerase proceeds in this fashion. The second possibility is for the enzyme to catalyze a concerted E2 elimination, proceeding in a single step without any intermediate. In enzymatic reactions, convincing evidence for an E2 elimination reaction is limited to the example of crotonase, which has been reported to catalyze a concerted *syn* elimination of water (Bahnson and Anderson, 1991). The most likely mechanism however, would proceed via an E1-like process. The combination of a good leaving group and a relatively stable oxocarbenium ion intermediate should permit this type of mechanism to occur. Precedence for an E1 elimination mechanism can be found in the reaction catalyzed by imidazoleglycerol phosphate dehydratase, where the cationic intermediate is thought to be stabilized by conjugation to the imadazole ring (Parker *et al.*, 1995). The observation of a primary kinetic isotope effect is consistent with either an E2 mechanism or an E1 mechanism in which deprotonation of the oxocarbenium ion intermediate is partially rate-determining.

It is further possible that the enzyme mechanism is more complicated than the four proposed pathways, potentially involving enzyme-bound intermediates or neighbouring group participation by the acetamido functionality. In particular, oxazolines **2** and **3** could be

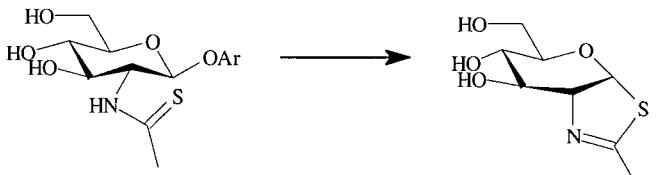


2



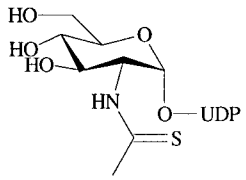
3

intermediates in the *anti* and *syn* eliminations respectively. Non-enzymatic oxazoline formation in both GlcNAc and ManNAc systems often occurs under conditions that promote oxocarbenium formation (Collins and Ferrier, 1995). Recently, a study has shown that oxazoline **2** is a potential intermediate in the reaction catalyzed by *N*-acetyl-β-hexosaminidase (Knapp *et al.*, 1996). In that study, the 2-thioacetamido glucoside **4** was incubated with jackbean *N*-acetyl-β-hexosaminidase, resulting in the enzyme-dependent conversion of **4** to **5**, which is a stable analogue of the oxazoline **2** (Figure 4.13). In a similar fashion, UDP-GlcNAc 2-epimerase might be probed with compound **6**. If evidence for neighbouring group participation were obtained, it might suggest why CDP-paratose 2-epimerase probably follows a different mechanism from UDP-GlcNAc 2-epimerase: while UDP-GlcNAc and UDP-ManNAc both have the potential for neighbouring group participation the same potential is not present with the 2"-hydroxyl of CDP-paratose and CDP-tyvelose.



4 **5**

Figure 4.13 N-Acetyl-β-hexosaminidase catalyzes the conversion of 2-thioacetamido glucoside, **4**, to thiazoline **5**. Ar = 4-methylumbelliferyl.



6

Finally, a great deal of information could be obtained from the X-ray crystal structure of the epimerase. Most useful would be the identification of the active site base or bases involved in the deprotonation during the catalytic reaction. The results of this study do not provide any information to distinguish between the possible involvement of a single general base or two separate active site bases in the catalytic mechanism (Figure 4.14; Rose, 1966). In the one-base mechanism, a single enzymic base abstracts a proton from the substrate, and then delivers it to the opposite face of the intermediate. The two-base mechanism involves different amino acid residues positioned on opposite sides of the substrate. One of the bases deprotonates the substrate, and the other (in its conjugate acid form) protonates the resulting intermediate to generate the inverted product.

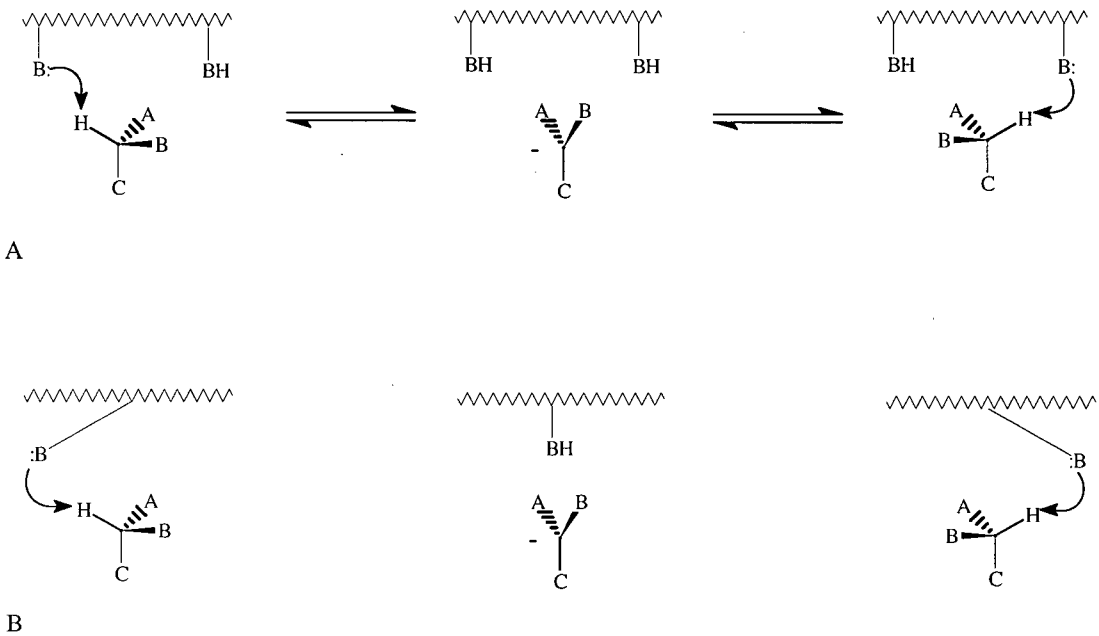


Figure 4.14 Schematic representations of (A): a two-base mechanism, and (B): a one-base mechanism.

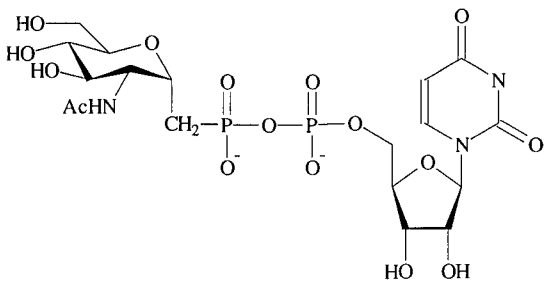
These two mechanisms can be distinguished by experiments with isotopically labeled substrates in order to trace the fate of the substrate or solvent protons. For instance the exchange of a C-2" deuterium for a proton in 2"-^{2}H-UDP-GlcNAc when the enzymatic reaction is run in H_2O under single-turnover conditions is indicative of a two-base mechanism. Conversely, return of the C-2" deuterium to the opposite face of the intermediate provides evidence for the involvement of a single active site base in the reaction. However, the presence of a polyprotic base, or the possibility that the lone base of a one-base mechanism is accessible and can exchange a proton with the solvent during the lifetime of the enzyme intermediate complicates the picture somewhat.

Another approach would involve the design and synthesis of an electrophilic reagent that could specifically target the active site bases (see Chapter Five). The catalytic bases could also act as nucleophiles, and become irreversibly labeled by a covalently-modifying reagent. However, such an affinity-label might not react at all, or react instead with a catalytically unimportant active site nucleophile. As a result, a more definitive answer might be derived from X-ray structural data.

Work towards this end has been initiated by Mr. Robert Campbell in collaboration with Dr. Natalie Strynadka in the Department of Biochemistry, The University of British Columbia. There is an important caveat however, that should be mentioned concerning the pursuit of crystallographic structural data. The epimerase has both a resting and an active conformation, which might differ significantly. Crystals obtained in the absence of a substrate analogue would show the inactive conformation, and would provide no information as to the position of the active site. It would be useful to design a non-reactive substrate analogue that could serve

to activate the epimerase and bind in the active site, acting to both induce the active conformation and mark the location of the active site in the crystallographic structure.

The 3"-deoxy substrate analogue did bind the enzyme, and it was not turned over by the epimerase, but it also failed to activate the enzyme as well. Another likely candidate for the role as a non-reactive activator is the nucleotide phosphonate 7. The synthesis of the precursor phosphonate was recently reported (Casero *et al.*, 1996). Only one additional synthetic step, to couple that precursor with UMP using standard morpholidate chemistry, would yield 7. The resultant UDP-GlcNAc phosphonate analogue would be incapable of undergoing the usual β-elimination reaction catalyzed by UDP-GlcNAc 2-epimerase. It is possible that this compound might bind to both the epimerase modulator site and the active site, serving in the capacity of both an allosteric enzyme activator and a substrate analogue.



7

One final interesting feature of the epimerase may also be determined from the X-ray crystal structure. Considering that enzyme active sites are notoriously stringent concerning the geometry of the substrate, epimerases and racemases are an unusual class of enzyme, since they must be capable of accommodating both stereoisomers of their substrate. In UDP-GlcNAc 2-epimerase this is especially unique since the two epimeric substrates, and the

intermediate glycal all differ by the position of a relatively bulky 2-acetamido group. Hopefully a detailed picture of the geometric construction of the active site pocket will shed some light on this intriguing issue.

4.8 Conclusions

The observed solvent derived deuterium incorporation at C-2" and primary deuterium kinetic isotope effect at C-2" supports a mechanism in which the C-2" proton is removed during a rate-determining step of the enzymatic epimerization reaction. The failure to observe any NAD^+ cofactor associated with the enzyme indicates that oxidation at C-3" is not involved in the reaction. The observation of isotopic scrambling in the β-phosphate of UDP-GlcNAc supports a mechanism involving cleavage of the anomeric C-O bond during the course of the reaction. As a result, only path B remains as a reasonable minimal depiction of the overall mechanism that accounts for all experimental observations.

This mechanism is supported by the further observation that 2-acetamidoglucal and UDP, the proposed intermediates for path B, accumulate free in solution upon extended incubations with a large excess of UDP-GlcNAc 2-epimerase. These two intermediates are produced through the *anti* elimination of UDP from UDP-GlcNAc, and the *syn* elimination of the UDP from UDP-ManNAc. The reverse reaction, readdition of UDP to 2-acetamidoglucal during the enzyme catalyzed epimerization can proceed in both manners, the *syn* addition resulting in UDP-ManNAc, and the *anti* addition generating UDP-GlcNAc. Although *syn* and *anti* eliminations and additions are known in enzyme reactions, it is unprecedented for an enzyme to employ both concomitantly in order to afford an overall stereochemical inversion.

Elucidation of the features involved in the enzyme mechanism suggests a variety of further experiments that might be performed, including the design of mechanism-based inhibitors with the intention of identifying the active site bases involved in the catalytic reaction.

4.9 Experimental Methods

4.9.1 General Chemicals

A sample of synthetic 2-acetamido glucal was generously provided by Dr. Ellen C. K. Lai from the lab of Dr. Steven Withers, Department of Chemistry, University of British Columbia. ^{14}C-UDP (10 μCi, 105 μL) was obtained from Sigma Radiochemicals. UDP-GlcNAc 2-epimerase was prepared as described in Chapter Two. All other chemicals were obtained as described in Chapter Two, unless otherwise noted.

4.9.2 Solvent Deuterium Incorporation

A sample of potassium phosphate buffer (50 mM, pH 8.8), containing 15.4 mM UDP-GlcNAc was lyophilized and resuspended two times in an equal volume of D_2O. UDP-GlcNAc 2-epimerase (0.27 units) was added to 450 μL of this deuterated buffer. The reaction was monitored by ^{1}H NMR at 25°C. The particular spectrum depicted in Figure 4.1 was acquired on a 400 MHz Bruker WH400 NMR spectrometer, by Mr. Rafael Sala, a co-worker in the lab of Dr. Martin Tanner.

4.9.3 Test for Bound NAD^+

The UV-visible spectra of the epimerase (8.5 mg/mL) in potassium phosphate buffer (100 mM, pH 8.8, containing 2 mM DTT) were taken at 37°C with a Varian Cary 3E Spectrophotometer. Spectra were run before and after the addition of 12 mM UDP-GlcNAc.

The epimerase was also tested for NAD^+ release by digesting 20 mg of the enzyme in 1 mL of potassium phosphate buffer (pH 7.0, containing 2 mM DTT, and 1 mM EDTA), according to the method outlined by Palmer and Abeles (1979). This sample was denatured at 100°C for 3 minutes, cooled to 37°C, and incubated at that temperature for 24 h with 0.5 mg each of trypsin and chymotrypsin. Half of the digested epimerase sample (0.5 mL) was added to 0.5 mL of 200 mM glycine buffer, pH 9.0, and 0.4 mg semicarbazide. UV-visible spectra from 500 to 250 nm were taken at 37°C of the sample, and again after the subsequent additions of 1 U horse liver alcohol dehydrogenase, followed by 15 μL of ethanol. No noticable change in the absorbance spectrum above 310 nm was noted until 1 equivalent (0.17 mg) of NAD^+ was added to the cuvette as well.

A control sample was prepared and digested with trypsin and chymotrypsin in an identical fashion. The control contained 20 mg bovine serum albumin in lieu of the epimerase, as well as 0.35 mg NAD^+ (1 equivalent with respect to the 20 mg amount of UDP-GlcNAc 2-epimerase used in the actual sample). The UV-Visible spectrum of the control sample was determined, and repeated following the subsequent addition of 1 U horse liver alcohol dehydrogenase, and 15 μL ethanol.

4.9.4 Test for Enzymatic Addition of UDP and 2-Acetamidoglucal

A sample (100 μL) containing 2 mM 2-acetamidoglucal, 0.6 mM UDP-GlcNAc, 0.9 mM ^{14}C-UDP (7.6 pCi ^{14}C) and 40 μg UDP-GlcNAc 2-epimerase (0.3 U), was prepared in 50 mM Tris-HOAc buffer (pH 8.8, containing 2 mM DTT). Two control samples containing heat killed enzyme, and no 2-acetamidoglucal were similarly prepared. Each sample was incubated for 4 h at 37°C. A 50 μL aliquot of each sample was injected onto a Waters Radial-Pak 8NVC18 reversed-phase HPLC column pre-equilibrated with 50 mM phosphate buffer (pH 7.0, containing 2.5 mM TBAHS), and eluted with a linear gradient of 0-10% CH_3CN in the same buffer, over the span of 20 min, according to the procedure described by Meynial *et al.* (1996). The eluent was monitored at 262 nm. The components were separated into 1 mL fractions, added to 5 mL of ScintiVerse II (general-purpose scintillation cocktail), and each fraction was counted for ^{14}C-isotope using a Beckman Instruments Model LS 6000 IC Scintillation counter.

A second trial involved samples prepared using identical conditions to those of the first, with the exception that 10 mM 2-acetamidoglucal was present in the final 100 μL sample. Again, the two controls contained heat-killed epimerase, and no-glycal respectively. One other sample was additionally prepared, containing 10 mM 2-acetamidoglucal, 0.6 mM UDP-GlcNAc, 1 mM UDP and 1.5 U (0.2 mg) UDP-GlcNAc 2-epimerase. All samples in this second trial were incubated for 24 h at 37°C, and again following incubation, 50 μL aliquots analyzed by ion-pair reversed-phase HPLC, with the eluent fractions were individually counted for ^{14}C-radioisotope.

4.9.5 Identification of the Enzyme-Released 2-Acetamidoglucal

Enzyme-produced 2-acetamidoglucal was identified by incubating a solution of UDP-GlcNAc (38 mM) in potassium phosphate buffer (50 mM, 1 mL, pH 8.8) containing 2 mM DTT and 8.8 units of UDP-GlcNAc 2-epimerase at 37°C for 12 h. The resulting sample was applied to a column of Dowex AG1 X8 (20 mL, formate form, 100-200 mesh), eluted with water (200 mL), and then lyophilized to dryness. The sample was redissolved in D_2O and a ^1H NMR spectrum was collected on a Bruker AC-200E spectrometer. This spectrum was compared to a similar spectrum obtained using an authentic sample of synthetic 2-acetamidoglucal in D_2O.

4.9.6 The Rate of Enzymatic Production of 2-Acetamidoglucal and UDP

The rate of UDP and 2-acetamidoglucal formation was measured by incubating solutions of UDP-GlcNAc and UDP-GlcNAc 2-epimerase, and monitoring timed aliquots by ion-pair reversed-phase HPLC. Samples (1 mL) contained 50 mM Tris-HCl buffer (pH 8.8), 2 mM DTT, 9.2 mM UDP-GlcNAc, and 0.88 units of epimerase at 37°C. Aliquots (50 μL) were removed at timed intervals over 120 min and frozen rapidly with liquid N_2. The samples were thawed immediately prior to injection onto the Waters Radial-Pak 8NVC18 reversed-phase HPLC column prequilibrated with 50 mM potassium phosphate buffer (containing 2.5 mM TBAHS). Components were seperated by elution with a linear gradient of 0 to 10% CH_3CN in the same buffer, over the span of 20 min. The eluent peaks, monitored at 262 nm, were integrated, and compared to a linear calibration curve previously obtained by the injection of known quantities of UDP. During the allotted time, 10% of the epimeric UDP-sugars had been converted to UDP and 2-acetamidoglucal. A plot of the UDP concentration

determined for the samples as a function of time yielded linear kinetics for UDP release. Error was determined based on the linear regression of the data. Control samples containing heat-killed epimerase were run to ensure that significant amounts of UDP/2-acetamidoglucal did not spontaneously form in the absence of active enzyme.

4.9.7 Determination of the External Equilibrium Constant for UDP and 2-Acetamidoglucal Release

An attempt to determine the external equilibrium constant was made by incubating a solution of UDP-GlcNAc (17 mM, 1 mL) in Tris-HCl buffer (50 mM, pH 8.8, containing 2 mM DTT) with 51 units of UDP-GlcNAc 2-epimerase at 37°C. Aliquots were removed at various timed intervals over the course of 48 h and analyzed by ion-pair reversed-phase HPLC, as described above. The concentration of 2-acetamidoglucal was taken to be equal to that of UDP, and for better accuracy with the small UDP-ManNAc peak, the UDP-ManNAc was taken to be one-tenth the concentration of UDP-GlcNAc (as expected for the equilibrated epimers).

4.9.8 Kinetic Isotope Effect Measurement

(a) Synthesis of 2"-^{2}H-UDP-*N*-Acetylglucosamine

A 20 mL sample of Tris-HCl buffer (50 mM, pH 8.8, containing 2 mM DTT) was lyophilized to dryness, and resuspended in an equal volume of D_2O two times. UDP-GlcNAc (100 mg, 7.7 mM) and UDP-GlcNAc 2-epimerase (400 μL, 3.5 mg/mL, 9.5 units, exchanged into the same "deuterated" buffer), was incubated at 37°C in "deuterated" buffer for 2 h. The reaction was quenched by heat denaturation of the epimerase at 50°C for 1 h, followed by the removal of the denatured protein using a centricon-10 concentrator (Amicon). The

equilibrated 2"- ^{2}H-UDP-ManNAc was removed by the addition of 15 mg NAD^{+} and UDP-ManNAc dehydrogenase (1 mg, 12 units). The solution was incubated at 37°C for 12 h, and analyzed by ion-pair reversed-phase HPLC to ensure all of the 2"- ^{2}H-UDP-ManNAc had been consumed. The product 2"-^{2}H-UDP-GlcNAc was purified by anion exchange chromatography using a column of Dowex AG1 X8 (20 mL, formate form, 100-200 mesh) and eluting with a linear gradient from 0 to 2 M triethylammonium bicarbonate (400 mL total volume). Fractions containing the product were lyophilized, resuspended in water, and lyophilized again. The product was passed through a column of Amberlite IR-120(plus) resin (20 mL, Na^{+} form, eluted with water) followed by a column of Bio-Gel P-2 (2.5 X 45 cm, eluted with water).

1H NMR spectroscopic and –LSIMS mass spectroscopic analyses indicated that the extent of deuterium incorporation was >97%: -LSMS (thioglycerol) m/z 629, (monosodium salt), ^{2}H-2", 100%).

(b) Kinetic Isotope Effect Determination

The reported kinetic isotope effect is k_H/k_D for the epimerase reaction under saturating conditions. The rates of epimerization were determined in triplicate for the labeled and the unlabeled 2"-H substrates. The reaction was monitored using the coupled enzyme assay described previously, in Chapter Two. The assay mixtures (325 μL) contained 50 mM Tris-HCl (pH 8.8), 2 mM DTT, 2 x 10^{-4} mg of UDP-GlcNAc 2-epimerase (1.4 x 10^{-3} U), 5 mg of UDP-ManNAc dehydrogenase (8.5 U at pH 8.8), 2.5 mM UDP-GlcNAc (either 2"-^{1}H or 2"-^{2}H), and 7 mM NAD^{+}. Reactions were initiated by the addition of UDP-GlcNAc 2-epimerase. A background rate due to impurities in the dehydrogenase preparation, and determined to be

independent of substrate, was subtracted from the observed reaction rate. The velocity of the epimerase with both protonated and deuterated substrate was calculated from the observed rate of NADH formation by dividing the rate by two, in order to account for the stoichiometry of the auxiliary enzymatic reaction. The error reported for the kinetic isotope effect is the standard deviation of the data points from the average determined rate of epimerization.

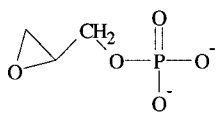
Chapter Five:

An Affinity Label to Probe the Catalytic Bases of UDP-*N*-Acetylglucosamine 2-Epimerase

5.1 Introduction

Several amino acid side-chains are capable of acting as nucleophiles, provided they are not buried within a folded protein. These nucleophiles can react with electrophilic reagents, resulting in the covalent modification of proteins. The majority of the commonly used reagents are non-specific. For example, the thiol-modifying reagent iodoacetate will label all the accessible cysteine side-chains, regardless of whether the side-chains are important for catalysis or not. Since it is often desirable to label only a single active site residue that is involved in catalysis, the non-specific reagents have limited use in revealing information regarding enzymatic mechanisms.

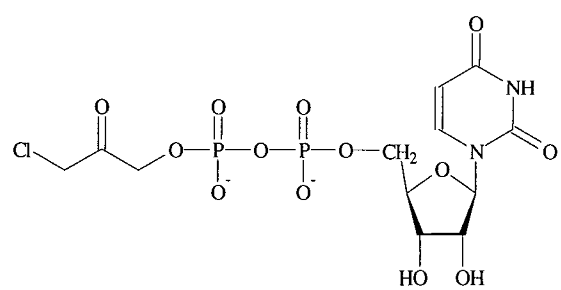
It is possible however, to target an enzyme active site for covalent modification by incorporating specific binding elements found in natural substrates into the design of electrophilic reagents. Many successful affinity-labeling reagents have been synthesized, and been found to irreversibly inhibit their target enzymes by covalently modifying active site nucleophiles (Glazer *et al.*, 1976; Fersht, 1987). One example of note is glycidol phosphate (2,3-epoxypropanol phosphate), **8**, which reacts with an active site glutamate of triosephosphate isomerase (Rose and O'Connell, 1969; Waley *et al.*, 1970), and also with an active site cysteine of glyceraldehyde-3-phosphate dehydrogenase (McCaul and Byers, 1976). In this compound, the dianionic phosphate provides similar binding properties to the natural



8

substrates, allowing the affinity label to bind in a similar orientation, and present the reactive epoxide moiety to the catalytically active amino acid residues.

Similarly, the uridine 5'-diphosphate moiety of UDP-sugars also provides a convenient handle that might be exploited in the design of highly specific affinity labels targeted for the active site of an enzyme that has an UDP-sugar as a substrate. For example, the compound uridine 5'-diphosphate chloroacetol (UDC), **9**, was designed to react with a possible general base required for catalysis in UDP-galactose 4-epimerase (Flentke and Frey, 1990). UDC did inactivate the 4-epimerase, but unexpectedly, instead of alkylating an active site residue, the compound formed a covalent attachment to the nicotinamide ring of the NAD^+ cofactor within the enzyme active site. However, since the resulting chromophoric adduct remained tightly bound in the active site, the enzyme was effectively inactivated.



9

A more conventional reaction with UDC was observed with UDP-glucose dehydrogenase, when an active site thiol required for catalysis was irreversibly alkylated by the affinity label (Campbell *et al.*, 1997). Electrospray mass spectrometric analysis showed that the inactivation was accompanied by an increase in the mass of the enzyme, and the size of the increase was consistent with the covalent attachment of one inhibitor molecule (less a chlorine atom).

UDC was also an attractive candidate for the irreversible inhibition of UDP-GlcNAc 2-epimerase since the reactive electrophilic position (C-3") of UDC lies three bonds away from the bridging oxygen of the β-phosphate. That number of bonds is analogous to the number of bonds between the C-2" proton and the anomeric oxygen of UDP-GlcNAc. Ideally, C-3" of UDC could sit in the active site in a position within reach of the active site base or bases involved the removal of the C-2" proton from UDP-GlcNAc and UDP-ManNAc during the catalytic mechanism. Studies probing the ability of UDC to react with UDP-GlcNAc 2-epimerase however showed that UDC did not inactivate the epimerase (Mr. Rafael Sala in the lab of Dr. Martin Tanner, Department of Chemistry, University of British Columbia; unpublished results).

5.2 Synthesis of UDP-Glycidol

The inability of UDC to inactivate the epimerase prompted us to design an alternative affinity label to probe the epimerase, with the goal of uncovering the base or bases involved in the catalytic mechanism. The success of glycidol phosphate, **8**, as an affinity label, suggested an alternative to UDC. Epoxides, like glycidol phosphate, are reportedly reactive with cysteines, lysines and carboxylate side chains, but remain relatively stable in solution (Fersht, 1987a). In comparison, the chloroacetol moiety of UDC reacts with thiol reagents such as DTT, a compound that is a commonly required buffer supplement necessary to maintain enzyme stability. Uridine 5'-diphosphate 2,3-epoxypropanol (UDP-glycidol), **10**, would have similar reactivity, but greater specificity for UDP-GlcNAc 2-epimerase than the monophosphate reagent, glycidol phosphate. Similar to UDC, UDP-glycidol would have an

electrophilic site at the C-3" position of the reagent, which would be expected to rest within the active site in an appropriate position to undergo nucleophilic attack by a catalytic base, possibly with the assistance of a general acid to increase reactivity.

Initial unsuccessful attempts were made to synthesize UDP-glycidol by directly coupling glycidol phosphate and UMP-morpholidate using standard UMP-morpholidate chemistry (Figure 5.1; Wittman and Wong, 1997). It is likely that the epoxide functionality in glycidol phosphate is not stable under the reaction conditions, which requires dry pyridine as the solvent. Instead, an alternate scheme was devised, in which the coupling reaction was carried out between allyl phosphate, **11**, and UMP-morpholidate under similar conditions (Figure 5.2). The resulting UDP-allyl alcohol, **12**, was recovered in 40% yield. Interestingly, attempts to perform the coupling in the absence of the tetrazole catalyst, did not yield any coupling product and only unreacted UMP-morpholidate was recovered.

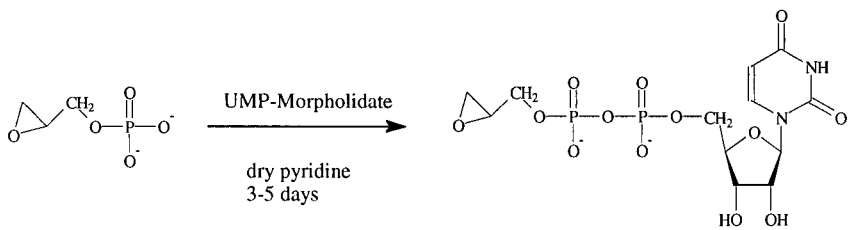


Figure 5.1 Initial scheme for the synthesis of UDP-glycidol.

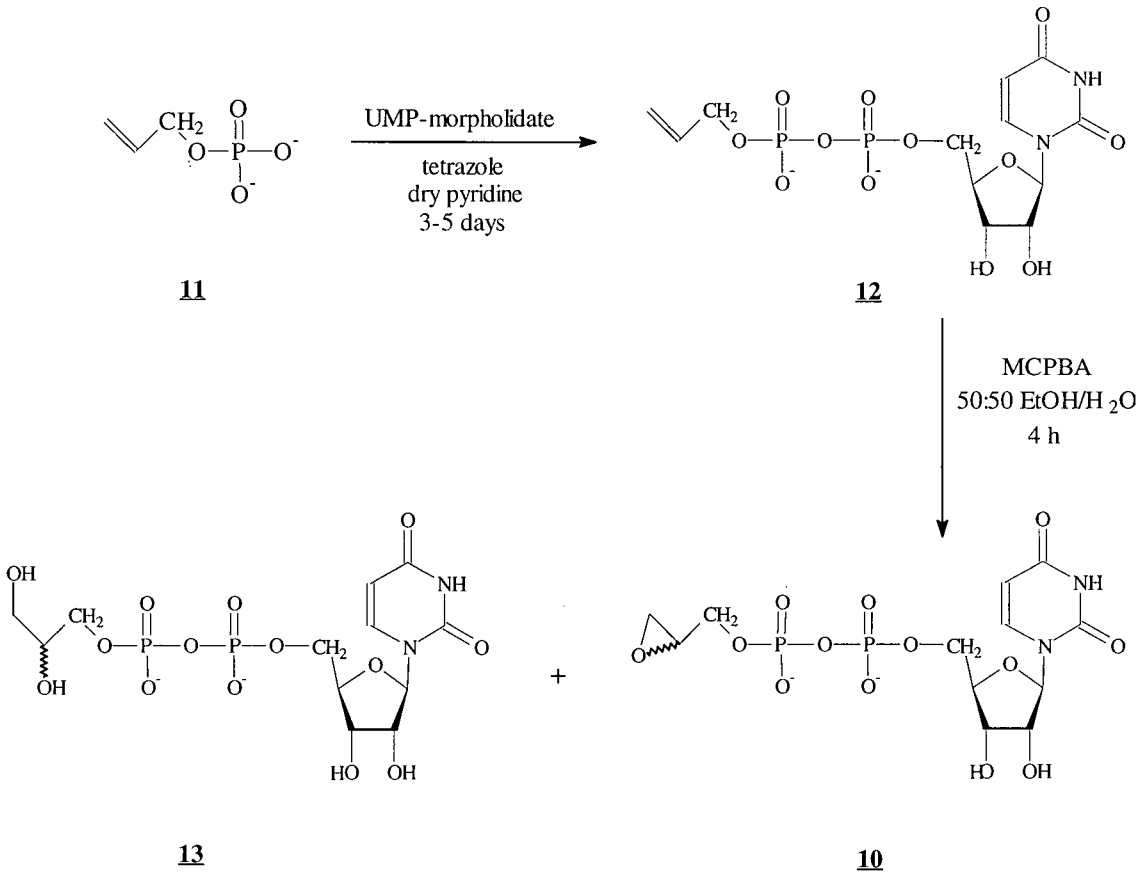


Figure 5.2 Modified synthetic scheme for UDP-glycidol.

The epoxidation of **12** was carried out with *m*-chloroperoxybenzoic acid (MCPBA) in a 50:50 ethanol/water solution. The ^{1}H NMR spectrum of the product was consistent with the complete consumption of the alkene starting material, as indicated by the disappearance of the signals in the ^{1}H NMR spectrum between 5.0 and 5.3 ppm and the multiplet between 5.7 and 5.8 ppm, that are attributed to the alkene protons of **12**. This was accompanied by the appearance of two multiplets at 2.85 and 3.35 ppm, indicative of the protons of a mono-

substituted epoxide ring. However it was evident from the mass spectrum (-LSIMS) that under the conditions employed during the reaction and subsequent work-up, a fair amount of the synthesized UDP-glycidol had been hydrolyzed (presumably at the epoxide to give **13**). From the mass spectrum, it was not possible to quantitatively determine the percentage of the compound that had hydrolyzed. The mass spectrum did however show that the $M+H^+$ peaks due to the epoxides (m/z 459) and hydrolysis products (m/z 477) were in a 1:1 ratio, while the corresponding $M+Na^+$ signals (m/z 481 and 499 for UDP-glycidol, and hydrolyzed UDP-glycidol, respectively) were in an approximate 3:5 ratio, indicating the preparation contained slightly more hydrolyzed compound than epoxide.

Since the remaining ^1H NMR signal from the C1"-proton of the epoxide moiety appears between 3.5 and 4.5 ppm, it could not be easily distinguished from the signals from the hydrolyzed compound present in the mixture, or from the signals arising from the hydroxyls of the ribosyl moiety of UDP. Therefore, it was not possible to thoroughly assign all peaks from this portion of the ^1H NMR spectrum.

However, based on the ^1H NMR and mass spectrometric analyses it was clear that the prepared sample was an inseparable mixture containing the diastereomeric epoxides and the corresponding hydrolysis products. As a result, due to the difficulties separating the components of the mixture, which are all of similar size and charge, the results reported here are considered preliminary.

5.3 Attempts to Inactivate UDP-*N*-Acetylglucosamine 2-Epimerase

The unique β-elimination mechanism of UDP-GlcNAc 2-epimerase makes UDP-glycidol an interesting prospect for an affinity label, particularly if a two-base mechanism is employed. The potential exists that the mechanism of enzyme inactivation by UDP-glycidol could be more complicated than a simple alkylation (Figure 5.3). The possibility exists that one of the bases could act as a nucleophile while the other assists as a general acid catalyst. Following the initial opening of the epoxide, it might then be possible for the second active site base to deprotonate C-2", resulting in the mechanism-based elimination of UDP from the covalent adduct.

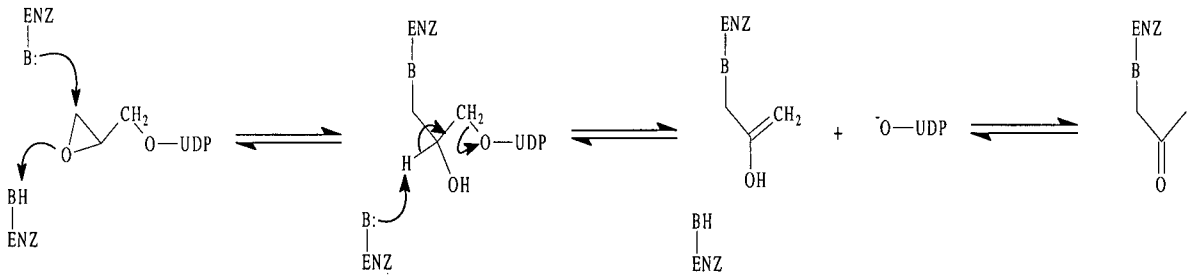


Figure 5.3 A possible mechanism for the proposed mechanism-based inhibition of UDP-GlcNAc 2-epimerase with UDP-glycidol.

To test this possiblilty, a sample of UDP-GlcNAc 2-epimerase was incubated overnight with 10 mM of the UDP-glycidol preparation and 1 mM UDP-GlcNAc. The latter substrate was included to ensure the epimerase was present in the active conformation. While the epimerase remained active following incubation, the electrospray mass spectrometric

analysis of the enzyme showed that only unmodified epimerase was present in the sample after incubation with the affinity label.

Since neither UDP-glycidol, **10**, nor UDC, **9**, is capable of inactivating the epimerase, it might be suggested that the enzyme base or bases involved in the deprotonation of the C-2" proton of UDP-GlcNAc and UDP-ManNAc during the normal catalytic reaction are not a cysteine thiol, a carboxylate residue, or the ε-amine from a lysine side-chain, all of which are known to react with epoxide and chloroacetol reagents. However, the possibility remains that neither affinity label is actually binding to the active site, or is binding in a non-productive conformation. Studies to test either UDC or UDP-glycidol as a competitive inhibitor have not yet been performed. The kinetic analysis of UDP-GlcNAc 2-epimerase, reported in Chapter Two, involved the use of a coupled enzyme assay, employing UDP-ManNAc dehydrogenase as the auxiliary enzyme. However, during experiments using the coupled assay, UDP-glycidol was observed to irreversibly inactivate UDP-ManNAc dehydrogenase.

5.4 Inactivation of UDP-*N*-Acetylmannosamine Dehydrogenase with UDP-Glycidol

As it became evident that the epimerase would not react with the UDP-glycidol, but UDP-ManNAc dehydrogenase would, we turned our focus on the nature of the inactivation of the dehydrogenase. Similar to the epimerase, UDP-ManNAc dehydrogenase was incubated with the UDP-glycidol preparation and analyzed by electrospray mass spectrometry for evidence of the successful covalent modification of the enzyme. The mass spectrum of the native dehydrogenase determined a mass of 45 723 ± 4 Da for the enzyme, agreeing with the expected mass calculated from the amino acid sequence (Figure 5.4, spectrum A). However,

following incubation with the UDP-glycidol preparation, the mass of the dehydrogenase increased to 46 181 ± 4 Da, which is consistent with the covalent attachment of one molecule of inactivator (molecular weight, 458 g/mol; Figure 5.4, spectrum B). Excess UDP-glycidol was removed from the covalently modified dehydrogenase by repeated dilution and concentration using fresh inactivator-free buffer. The process did not regenerate the active unmodified enzyme, suggesting that the inactivation is irreversible.

UDP-ManNAc dehydrogenase contains a total of nine cysteine residues. An experiment was performed to test the possibility that one of these residues was the position of covalent attachment, following the procedure of Campbell *et al.* (1997). A sample of the dehydrogenase was denatured in 5 M urea, and allowed to react with excess iodoacetate. The electrospray mass spectrum of the iodoacetate-treated protein gave a mass of 46 241 ± 4 Da, which corresponds to nine acetate units (58 daltons per acetate) each covalently linked to one of the nine cysteine residues (Figure 5.5, spectrum A). A sample of dehydrogenase first pre-treated with UDP-glycidol then treated with iodoacetate in an identical manner, gained the mass equivalent to only eight acetate units, in addition to the one molecule of inhibitor already covalently bound to the enzyme (Figure 5.5, spectrum B). This is consistent with UDP-glycidol reacting with a specific cysteine residue.

Like UDP-ManNAc dehydrogenase, the enzyme UDP-glucose dehydrogenase catalyzes the two-fold oxidation of an UDP-sugar at C-6". Mechanistic studies on UDP-glucose dehydrogenase have shown that that enzyme likely proceeds through the mechanism outlined in Figure 5.6. An experiment involving the incubation of UDP-glucose dehydrogenase with UDP-hexodialdose, **14**, has demonstrated that the aldehyde is kinetically

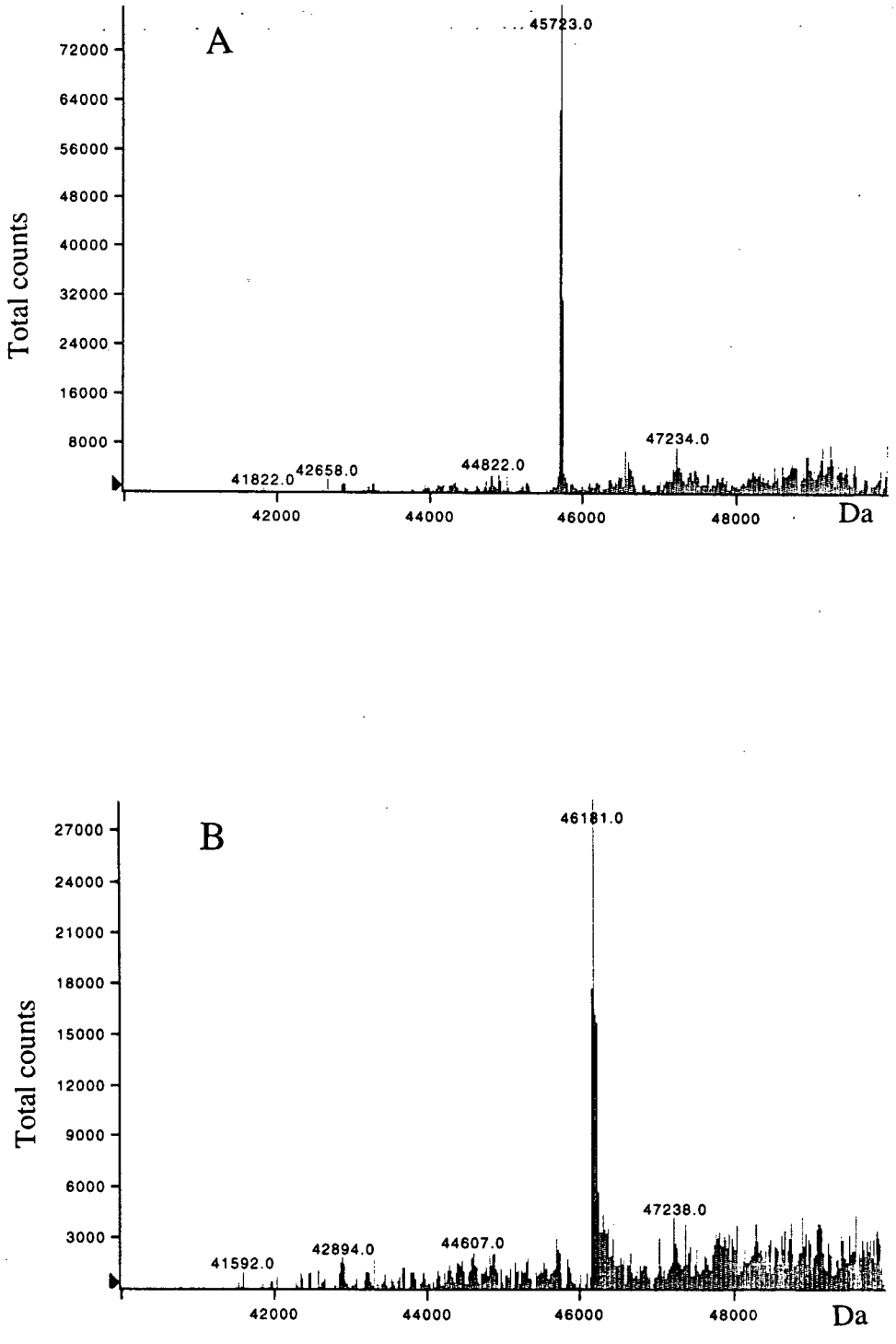


Figure 5.4 Deconvoluted electrospray ionization mass analysis of (A) native UDP-ManNAc dehydrogenase, and (B) UDP-ManNAc dehydrogenase following incubation with the UDP-glycidol preparation (10 mM).

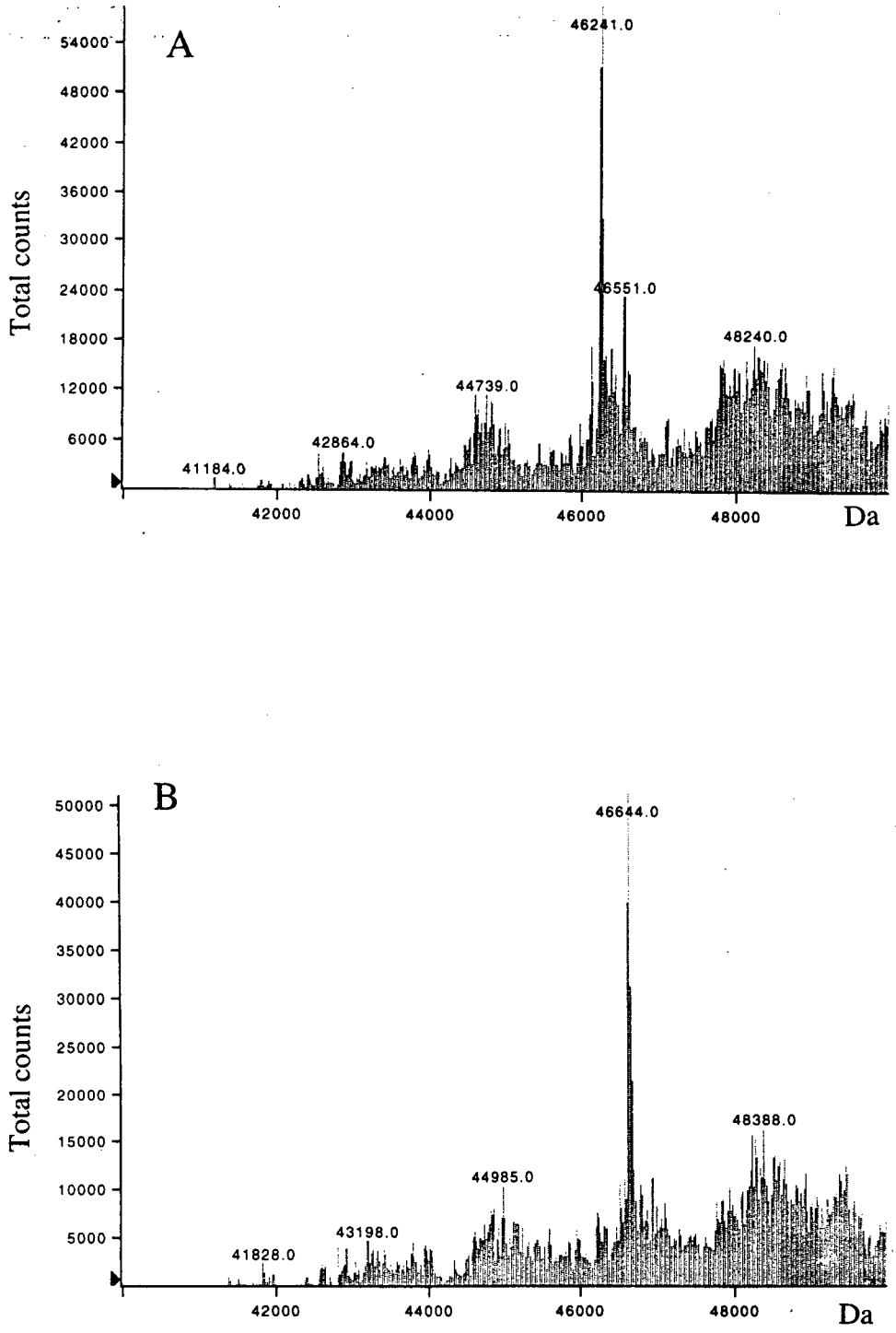


Figure 5.5 Deconvoluted electrospray ionization mass analysis of (A) UDP-ManNAc dehydrogenase, first denatured then treated with 2 mM iodoacetate, and (B) UDP-ManNAc dehydrogenase inhibited with UDP-glycidol, before denaturation and incubation with 2 mM iodoacetate.

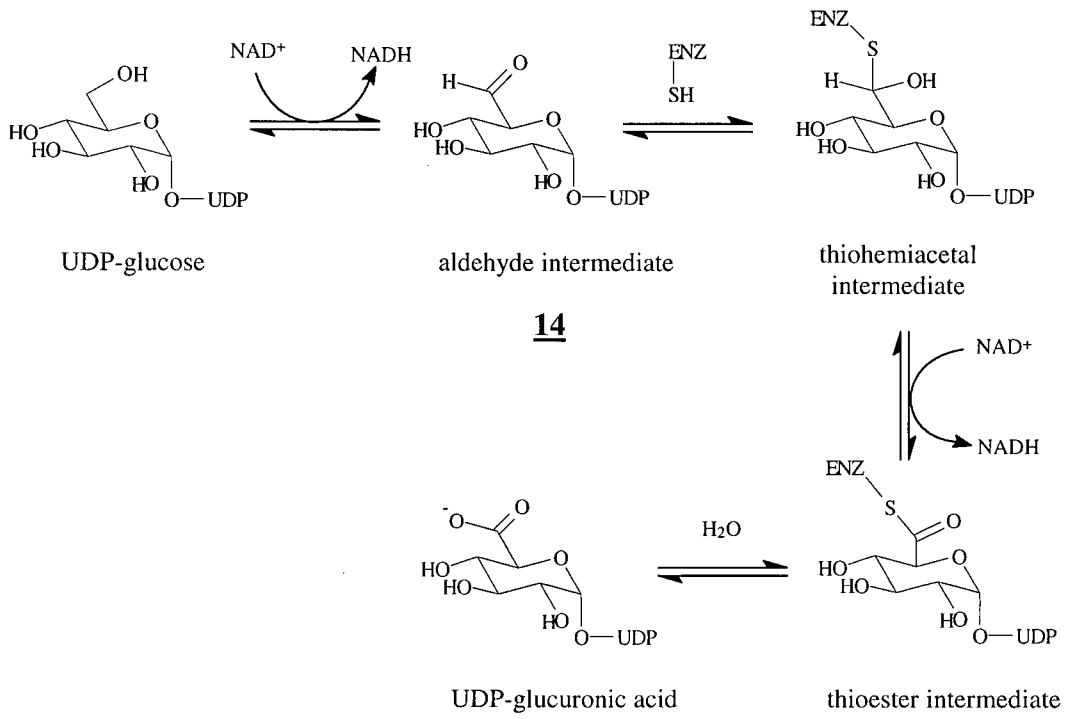


Figure 5.6 The proposed mechanism of UDP-glucose dehydrogenase.

competent, utilizing a single NAD^+ to generate UDP-glucuronic acid, suggesting the involvement of **14** as an intermediate in the normal reaction pathway (Campbell and Tanner, 1997). Covalent modification and site-directed mutagenesis studies have demonstrated the requirement for cysteine-260 of UDP-glucose dehydrogenase for catalytic activity (Campbell *et al.*, 1997). This cysteine, one of only two in the UDP-glucose dehydrogenase sequence, is also strictly conserved in the sequences of several other known nucleotide diphosphate sugar

dehydrogenases, including GDP-mannose dehydrogenase from *Pseudomonas aeruginosa*, and bovine UDP-glucose dehydrogenase (Dougherty and van de Rijn, 1993).

Sequence alignment of UDP-ManNAc dehydrogenase with each of these other dehydrogenases shows that cysteine-266 of UDP-ManNAc dehydrogenase matches the conserved catalytic cysteine found in the active site of all the other dehydrogenases (Table 5.1). It is likely that this cysteine is the active site nucleophile, and that same nucleophile is the site of covalent attachment of the inactivator, UDP-glycidol.

Table 5.1 An alignment of the putative active site residues for five nucleotide-diphosphate sugar dehydrogenases. The highly conserved catalytic cysteine residue is underlined.

HasB[a]	H Y N N P S F G Y G G Y C L P K D
AlgD[b]	Y Y M R P G F A F G G S C L P K D
RffD[c]	N I L Q P G P G V G G H C I A V D
Udg[d]	H Y N N P S F G Y G G Y C L P K D
bovine[e]	A S V G F G G S C F Z Z G

[a]UDP-glucose dehydrogenase from *Streptococcus pyogenes*
[b]GDP-mannose dehydrogenase from *Pseudomonas aeruginosa*
[c]UDP-ManNAc dehydrogenase from *Escherichia coli.*
[d]UDP-glucose dehydrogenase from *Salmonella typhimurium*
[e]Bovine UDP-glucose dehydrogenase

An effort was made to confirm the suspicion that cysteine-266 was the site of covalent modification in the labeling experiments. Samples of both UDP-glycidol-labeled and

unlabeled UDP-ManNAc dehydrogenase were proteolytically digested. Two different proteases were used in different trials (trypsin and pepsin) with both digestion methods yielding inconclusive results. In both cases there were obvious differences in the electrospray mass spectra between the labeled and unlabeled digestions, but it was not possible to identify any peptides in the labeled digestions that had gained a mass equivalent to a single inactivator molecule, or of a fragment of UDP-glycidol.

5.5 Conclusions

The preliminary results reported in this chapter demonstrate that the affinity label UDP-glycidol, whose design was based on another successful affinity reagent, reacts with UDP-ManNAc dehydrogenase in an irreversible manner. The evidence supports the covalent attachment of a single inactivator molecule to an active site cysteine, that likely acts as an active site nucleophile during the normal catalytic reaction. This observation is consistent with the mechanism that has been proposed for a number of similar dehydrogenases, all of which catalyze a similar two-fold oxidation at the C-6" position of a sugar nucleotide.

Unfortunately, however, UDP-glycidol did not react irreversibly with UDP-GlcNAc 2-epimerase. As a result, the enzymic bases used by the epimerase during the catalytic reaction remain elusive. It is hoped that X-ray crystallographic studies might reveal a more detailed picture concerning the active site residues involved in catalysis.

5.6 Experimental Methods

5.6.1 Chemicals

Chemicals were obtained as described in Chapter Two, unless noted otherwise.

Bis(cyclohexylammonium) allyl phosphate, **11**, was prepared according to the procedure detailed in Müller et al. (1990), by Anne Johnson, a co-worker in the lab of Dr. Martin Tanner, Department of Chemistry, University of British Columbia.

5.6.2 Synthesis of Uridine 5'-Diphosphate 2-Propenol, 12

A solution of bis(cyclohexylammonium) allyl phosphate (**11**, 250 mg, 0.73 mmol) in dry pyridine (20 mL) was evaporated to dryness *in vacuo*. The dried residue was redissolved in dry pyridine and again evaporated to dryness twice. Following the final evaporation, the residue was redissolved in 10 mL dry pyridine. Separately, a sample of 4-morpholine *N,N'*-dicyclohexylcarboxamidinium uridine 5'-monophosphomorpholidate (205 mg, 0.31 mmol) was treated similarly. After the second evaporation of pyridine from the 5'-monophosphomorpholidate, 40 mg of tetrazole was added, followed by 20 mL of dry pyridine. This solution was evaporated once to dryness, and redissolved in 10 mL dry pyridine. The two dry pyridine solutions were mixed, sealed, and stirred for three days. After three days, the pyridine was removed from the solution *in vacuo*, the resultant syrup was resuspended in water, and dried, in order to remove the excess pyridine. The syrup was redissolved in 50 mL H_2O, and applied to a column of Dowex AG1 X-8 (20 mL, formate form, 100-200 mesh) prequilibrated in water. The column was washed with water and eluted with a linear gradient of 0 to 1 M triethylammonium bicarbonate (800 mL total volume). The column eluent was monitored at 254 nm with a SPECTRUM Spectra/Chrom Flow-Thru UV

monitor. Two peaks were eluted from the column, the first attributed to unreacted 5'-monophosphomorpholidate, while the later peak (eluting at 0.6 M triethylammonium bicarbonate) had a ^{1}H NMR spectrum consistent with the expected product. The pooled fractions of uridine 5'-diphosphate 2-propenol, **12** were concentrated to 5 mL and applied to a column of Bio-Gel P-2 (2.5 X 90 cm) and eluted with water, again monitored at 254 nm. The appropriate fractions were pooled and lyophilized to yield 75 mg (40%) of the triethylammonium salt of **12**. ^{1}H NMR (200 MHz, D_2O): δ 7.92 (d, J=8.1 Hz, 1H, H-C6); δ 5.93 (d, J=4.9 Hz, 1H, H-C1'); δ 5.91 (d, J=8.1 Hz, 1H, H-C5); δ 5.84 (m, 1H, H-C2"); δ 5.35 (dd, J=17.2, 1.8 Hz, 1H, H_{trans}-C3"); δ 5.19 (dd, J=10.4, 1.5 Hz, 1H, H_{cis}-C3"); δ 4.34-4.25 (m, 2H, H-C2', H-C3'); δ 4.22-4.10 (m, 3H, H-C4', H-C5a', H-C5b'); δ 4.21 (tt, J=4.7, 1.4 Hz, 2H, H-C1"). ^{31}P NMR (51 MHz, D_2O): δ -11.01 (d,Pβ, J=20.3 Hz); δ -12.90 (d, Pα, J=20.3 Hz). LRMS (-LSIMS) Calcd for; $C_{12}H_{18}O_{12}P_2N_2$ ($M-H^+$): 443, Found: 443

5.6.3 Synthesis of Uridine 5'-Diphosphate 2,3-epoxypropanol (UDP-Glycidol), 10

A solution of 20 mg (0.03 mmol) uridine 5'-diphosphate 2-propenol, **12**, in 5 mL H_2O was mixed with a solution of 500 mg (2.9 mmol) *m*-chloroperoxybenzoic acid dissolved in 5 mL ethanol. The resultant solution was stirred overnight at 4°C. The solution and white precipitate was extracted twice with 20 mL portions of diethyl ether, concentrated *in vacuo* to 2 mL, and purified through a column of Bio-Gel P2 (2.5 cm X 90 cm) eluting with water. The UV absorbing fractions (254 nm) were lyophilized to give 12 mg of a mixture containing the triethylammonium bicarbonate salts of the diastereomeric epoxides, **10**, and the hydrolyzed diols, **13**. The 200 MHz ^{1}H NMR spectrum was taken of the mixture, but due to the overlapping of peaks from the contributing species present, it was not possible to fully

assign all the signals. However, it was possible to identify certain signals corresponding to the epoxide moiety of **10**. In particular, the doublet of quartets at 2.85 ppm is typical for the two C-3" protons of the mono-substituted epoxide ring, and the multiplet at 3.35 ppm can be attributed the other epoxide-ring proton (H-C2"). These signals are indicative that the epoxide had formed from **12**. It was evident that little **12** remained, on account of the absence of the alkene peaks between 5 and 6 ppm. The ^{1}H NMR spectrum also possessed many of the peaks expected from the uridine moeity of **10**, including the uridine ring protons, H-C6 and H-C5, at 7.92 and 5.91 ppm respectively. The ^{31}P NMR spectrum also confirmed the presence of an intact phosphodiester bond. ^{31}P NMR (51 MHz, D_2O): δ -11.01 (d,Pβ, J=20.3 Hz); δ -12.90 (d, Pα, J=20.3 Hz). Mass spectroscopic analysis clearly showed the presence of both **10** and **13**. LRMS (-LSIMS) Calcd for **10**, $C_{12}H_{18}O_{13}P_2N_2$ ($M-H^+$): 459, Found: 459; Calcd for **13**, $C_{12}H_{20}O_{14}P_2N_2$ ($M-H^+$): 477, Found: 477.

5.6.4 Incubations of UDP-Glycidol with UDP-GlcNAc 2-Epimerase

The UDP-glycidol mixture (10 mM; determined at A_{262}, based on uridine chromophore; ε_{262} = 8770 $M^{-1}cm^{-1}$) was incubated with 1 mg of the epimerase, and 1 mM UDP-GlcNAc in 400 µL of 50 mM Tris-HCl buffer (pH 8.8, containing 2 mM DTT). The sample was incubated overnight at 37°C, and submitted directly for electrospray mass spectrometric analysis.

5.6.5 Incubations of UDP-Glycidol with UDP-ManNAc Dehydrogenase

A 10 mg/mL sample of UDP-ManNAc dehydrogenase was incubated overnight at 37°C with 10 mM UDP-glycidol (determined at A_{262}), in 300 µL of 50 mM Tris-HCl (pH 8.8,

containing 2 mM DTT). The sample was subsequently submitted for mass spectrometric analysis.

5.6.6 Iodoacetate Labeling of UDP-ManNAc Dehydrogenase

A sample of dehydrogenase (2 mg/mL in 30 μL of 50 mM Tris-HCl, pH 8.8) was denatured by the addition of 30 μL of saturated urea in the same buffer. Following incubation at room temperature for 1 h, iodoacetate was added (final concentration 2 mM). The solution was incubated in the dark for a further 1 h, and quenched by the addition of 10 mM DTT. A sample of dehydrogenase (2 mg/mL) that had previously been covalently modified by UDP-glycidol was treated in an identical fashion. Both samples were submitted directly for electrospray mass spectrometric analysis.

5.6.7 Electrospray Ionization Mass Spectrometric Analysis of Proteins

The molecular weights of the labeled and unlabeled whole proteins were determined by electrospray ionization mass spectrometry, performed by Shouming He, in the lab of Dr. Stephen Withers, Department of Chemistry, University of British Columbia. The electrospray ionization experiments were performed in an identical fashion to those reported previously in Chapter Two. The intact protein samples were injected onto a microbore PLRP-S reversed-phase column (1 X 50 mm), and eluted with a linear gradient of 16 to 80% CH_3CN in water (containing 0.05% trifluoroacetic acid) over three min, and maintained at 80% CH_3CN for an additional seven min.

5.6.8 Proteolytic Digestions of Labeled and Unlabeled UDP-ManNAc Dehydrogenase

(a) Tryptic Digests

A sample containing 2.5 mg/mL UDP-ManNAc dehydrogenase was incubated overnight at 37°C with 0.025 mg/mL trypsin (1000 μL total volume, with 25 mM Tris-HCl, pH 8.0, 1 mM EGTA and 2 mM $CaCl_2$). The sample was inactivated by heat denaturation at 100°C for 5 min, and submitted directly for electrospray mass spectrometric analysis. A similar sample containing 2.5 mg/mL dehydrogenase previously covalently modified by UDP-glycidol was prepared and analyzed in an identical manner.

(b) Peptic Digests

A sample of UDP-ManNAc dehydrogenase (0.5 mg/mL) was proteolytically digested with 0.05 mg/mL of pepsin in 200 μL of 100 mM sodium phosphate buffer, pH 2. The reaction was rapidly quenched after a 4 h incubation at 37°C by immersion in liquid N_2, and the sample was submitted directly for electrospray mass spectrometric analysis. A similar sample containing 5 mg/mL of dehydrogenase previously covalently modified by UDP-glycidol was prepared and analyzed in an identical manner.

5.6.9 Electrospray Ionization Mass Spectrometric Analysis of Peptides

The molecular weights of the labeled and unlabeled protein proteolytic digests were determined by electrospray ionization mass spectrometry, performed by Shouming He, in the lab of Dr. Steven Withers, Department of Chemistry, University of British Columbia. Electrospray ionization mass spectrometry experiments used an HPLC-ESMS setup consisting of a microbore HPLC (Michrom UMA) connected in-line to a PE-SCIEX AP 300

MS. The digested protein samples were injected onto a microbore Reliasil C-18 reversed-phase column (1 X 50 mm), and eluted with a linear gradient of 0 to 60% CH_3CN in water (containing 0.05% trifluoroacetic acid) over 60 min. The eluent was introduced directly into the spectrometer, which was operated in the single quadrupole mode. The molecular weights of the peptides were determined from the mass spectrometric data using software supplied by SCIEX.

5.6.10 Database Searches and Protein Sequence Alignments

The protein sequences for the nucleotide sugar dehydrogenases listed in Table 5.1 were retrieved from the Swiss Protein Data Bank (Bairoch and Apweiler, 1997) or from the GenBank Sequence Database (Benson *et al.*, 1998). The protein sequence alignments were performed using the on-line program ALIGN, on the Genestream SSEARCH Network server at CRBM, Montpellier, France. ALIGN performs an optimal global alignment of two sequences with no short-cuts.

References

Adams, E. (1976) *Adv. Enzymol. Relat. Areas Mol. Biol.* **44**, 69-138

Allison, R.D. and Purich, D.L. (1979) *Methods Enzymol.* **63**, 3-22

Anderson, V.E. (1991) In *Enzyme Mechanisms from Isotope Effects*, Cook, P.F., Ed.; CRC Press: Cleveland. pp 389-417

Anderson, V.E. (1998) in *Comprehensive Biological Catalysis: A Mechanistic Reference, Vol. 2: Reactions of Nucleophilic/Carbanionoid Carbon*, Sinnott, M.L. Ed.; Academic Press: NewYork. pp. 115-133

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1992) in *Short Protocols in Molecular Biology (A Compendium of Protocols from Current Protocols in Molecular Biology)*; Green Publishing Associates and John Wiley and Sons: New York.

Baddiley, J. (1970) *Acc. Chem. Res.* **3**, 98-105

Baddiley, J. (1989) *Bio Essays* **10**, 207-210

Bahnson, B.J. and Anderson, V.E. (1991) *Biochemistry* **30**, 5894-5906

Bairoch, A. and Apweiler, R. (1997) *Nucleic Acid Res.* **25**, 31-36

Barber, G.A. (1979) *J. Biol. Chem.* **254**, 7600-7603

Barber, G.A. and Hebda, P.A. (1982) *Methods Enymol.* **83**, 522-525

Bauer, A.J., Rayment, I., Frey, P.A. and Holden, H.M. (1992) *Proteins* **12**, 372-381

Benson, D.A., Boguski, M.S., Lipman, D.J., Ostell, J. and Ouellette, B.F. (1998) *Nucleic Acids Res.* **26**, 1-7.

Biely, P. and Jeanloz, R.W. (1969) *J. Biol. Chem.* **244**, 4929-4935

Blattner, F.R., Plunkett III, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau,B. and Shao, Y. (1997) *Science* **277**, 1453-1474

Bradford, M.M. (1976) *Anal. Biochem.* **72**, 248-254

Campbell, R.E., Sala, R.F., van de Rijn, I. and Tanner, M.E. (1997) *J. Biol. Chem.* **272**, 3416-3422

Campbell, R.E. and Tanner, M.E. (1997) *Angew. Chem. Int. Ed. Engl.* **36**, 1520-1522

Carlsen, B.D., Kawana, M., Kawana, C., Tomasz, A. and Giebink, G.S. (1992) *Infection and Immunity* **60**, 2850-2854

Casero, F., Cipolla, L., Lay, L., Nicotra, F., Panza, L. and Russo, G. (1996) *J. Org. Chem.* **61**, 3428-3432

Collins, P. and Ferrier, R. (1995) In *Monosaccharides*; John Wiley and Sons: Chichester. pp 236-237

Comb, D.G. and Roseman, S. (1958) *Biochim. Biophys. Acta* **29**, 653-654

Creighton, D.J. and Murthy, N.S.R.K. (1990) In *The Enzymes, Vol. 19*, Sigman, D.S. and Boyer, P.D., Eds.; Academic Press: San Diego. pp 397-403

Daniels, D.D., Plunkett III, G., Burland, V. and Blattner, F.R. (1992) *Science* **257**, 771-778

Dougherty, B.A. and van de Rijn, I. (1993) *J. Biol. Chem.* **268**, 7118-7124

Eventoff, W., Rossman, M.G., Taylor, S.S., Torff, H.-J., Meyer, H., Keil, W. and Kiltz, H.-H. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 2677-2681

Fee, J.A., Hegeman, G.D. and Kenyon, G.L. (1974) *Biochemistry* **13**, 2528-2532

Fersht, A. (1985a) in *Enzyme Structure and Mechanism, 2nd ed.*; W.H. Freeman and Co.: New York. pp. 248-259.

Fersht, A. (1985b) in *Enzyme Structure and Mechanism, 2nd ed.*; W.H. Freeman and Co.: New York. pp. 272-273.

Fischer, W., Behr, T., Hartmann, R., Peter-Katalinic, J. and Egge, H. (1993) *Eur. J. Biochem.* **215**, 851-857

Flentke, G.R. and Frey, P.A. (1990) *Biochemistry*, **29**, 2430-2436

Frey, P. (1987) in *Pyridine Nucleotide Coenzymes*, Dolphin, D., Poulson, R., and Avramovic, Eds.; John Wiley and Sons: NewYork. pp. 461-511.

Gallo, and Knowles (1993) *Biochemistry* **32**, 3981-3990

Glaser, L. (1960) *Biochem. Biophys. Acta* **41**, 534-536

Glaser, L. (1972) in *The Enzymes, 3rd ed.*, Boyer, P.D., Ed.; Academic Press: New York. pp. 355-380.

Glazer, A.N., DeLange, R.J. and Sugman, D.S. (1976) in *Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 4, Part 1: Chemical Modification of Proteins*, Work, T.S. and Work, E., Eds.; American Elesevier Publishing Co.: New York. pp. 1-205

Grimshaw, C.E., Sogo, S.G., Copley, S.D. and Knowles, J.R. (1984) *J. Am. Chem Soc.* **106**, 2699-2700

Hepinstall, S., Archibald, A.R. and Baddiley, J. (1970) *Nature* **225**, 519-521

Hester, L.S. and Raushel, F.M. (1987) *Biochemistry* **26**, 6465-6471

Hess, D., Covey, T.C., Winz, R., Brownsey, R.W. and Aebersold, R. (1993) *Protein Sci.* **2**, 1342-1351

Hinderlich, S., Stäsche, R., Zeitler, R., and Reutter, W. (1997) *J. Biol. Chem.* **272**, 24313-24318

Ichihara, N., Ishimoto, N. and Ito, E. (1974) *FEBS Letters* **40**, 309-311

Kaila, N., Blumenstein, M., Bielawska, H. and Franck, R.W. (1992) *J. Org. Chem.* **57**, 4576-4578

Kang, U.G., Nolan, L.D. and Frey, P.A. (1975) *J. Biol. Chem.* **250**, 7099-7105

Kawamura,, T., Ichihara, N., Ishimoto, N. and Ito, E. (1975) *Biochem. Biophys. Res. Comm.* **66**, 1506-1512

Kawamura, T., Kimura, T., Yamamori, S. and Ito, E. (1978) *J. Biol. Chem.* **253**, 3595-3601

Kawamura, T., Ishimoto, N. and Ito, E. (1979) *J. Biol. Chem.* **254**, 8457-8465

Kawamura, T., Ishimoto, N. and Ito, E. (1982) *Methods Enzymol.* **83**, 515-519

Keenlyside, W.J., Perry, M., Maclean, L., Poppe, C. and Whitfield, C. (1994) *Mol. Microbiol.* **11**, 437-448

Ketley, J.N. and Schellenberg, K.A. (1973) *Biochemistry* **12**, 315-319

Kiino, D.R. and Rothman-Denes, L.B. (1989) *J. Bacteriol.* **171**, 4595-4602

Kiino, D.R., Licudine, R., Wilt, K., Yang, D.H.C. and Rothman-Denes, L.B. (1993) *J. Bacteriol.* **175**, 7074-7080

Kikuchi, K. and Tsuiki, S. (1973) *Biochim. Biophys. Acta* **327**, 192-206

Knapp, S., Vocadlo, D., Gao, Z., Kirk, B. Lou, J. and Withers, S.G. (1996) *J. Am. Chem. Soc.* **118**, 6804-6805

Knox, K.W. and Wicken, A.J. (1973) *Bacteriol. Rev.* **37**, 215-257

Kuhn, H.-M., Meier-Deiter, U. and Mayer, H. (1988) *FEMS Microbiology Reviews* **54**, 195-222

Lai, E.C.K. and Withers, S. (1994) *Biochemistry* **33**, 14743-14749

Lee, C.-J., Banks, S.D. and Li, J.P. (1991) *Crit. Rev. Microbiol.* **18**, 89-114

Lee, J.J., Asano, Y., Shieh, T.-L., Spreafico, F., Lee, K. and Floss, H.G. (1984) *J. Am. Chem. Soc.* **106**, 3367-3368

Lees, W.J. and Walsh, C.T. (1995) *J. Am. Chem. Soc.* **117**, 7329-7337

Legler, G. (1990) *Adv. Carbohydr. Chem. Biochem.* **48**, 349-357

Lew, H.C., Nikaido, H. and Mäkela, P.H. (1978) *J. Bacteriol.* **136**, 227-233

Liu, Y., Vanhooke, J.L., and Frey, P.A. (1996) *Biochemistry* **35**, 7615-7620

MacPherson, D.F., Manning, P.A. and Morona, R. (1993) *Mol. Microbiol.* **11**, 281-292

Makela, P.H., Schmidt, G., Mayer, H., Nikaido, H., Whang, H.Y. and Neter, E. (1976) *J. Bacteriol.* **127**, 1141-1149

Makrides, S. (1996) *Microbiological Reviews* **60**, 512-538

Marquart, D.W. (1963) *J. Soc. Ind. Appl. Math.* **11**, 431-441

Mayer, H. (1969) *Eur. J. Biochem.* **8**, 139-145

Maxwell, E.S. (1957) *J. Biol. Chem.* **229**, 139-151

McCaul, S. and Byers, L.D. (1976) *Biochem. Biophys. Res. Comm.* **72**, 1028-1034

McClure, W.R. (1969) *Biochemistry* **8**, 2782-2786.

Meier, U. and Mayer, H. (1985) *J. Bacteriol.* **163**, 756-762

Meier-Deiter, U., Starman, R., Barr, K., Mayer, H. and Rick, P.D. (1990) *J. Biol. Chem.* **265**, 13490-13497

Meier-Deiter, U., Barr, K., Starman, R., Hatch, L. and Rick, P.D. (1992) *J. Biol. Chem.* **267**, 746-753

Meynial, I., Paquet, V. and Combes, D. (1995) *Anal. Chem.* **67**, 1627-1631

Morgan, P.M., Sala, R.F. and Tanner, M.E. (1997) *J. Am. Chem. Soc.* **119**, 10269-10277

Morolda, C.L. and Volvano, M.A. (1995) *J. Bacteriol.* **177**, 5539-5546

Müller, D., Pitsch, S., Kittaka, A., Wagner, E., Wintner, C.E., and Eschenmoser, A. (1990) *Helv. Chim. Acta* **73**, 1410-1468

Mullis, K.B. and Faloona, F.A. (1987) *Methods Enzymol.* **155**, 335-350

Nakajima, N., Tanizawa, K., Tanaka, H. and Soda, K. (1986) *Agric. Biol. Chem.* **50**, 2823-2830

Nelsestuen, G.L. and Kirkwood, S.J. (1971) *J. Biol. Chem.* **246**, 7533-7543

Page, R.L. and Anderson, J.S. (1972) *J. Biol. Chem.* **247**, 2471-2479

Palmer, J.L. and Abeles, R.H. (1976) *J. Biol. Chem.* **251**, 5817-5819

Palmer, J.L. and Abeles, R.H. (1979) *J. Biol. Chem.* **254**, 1217-1226

Parker, A.R., Moore, J.A., Schwab, J.M. and Davisson, V.J. (1995) *J. Am. Chem. Soc.* **117**, 10605-10613

Raushel, F.M. and Villafranca, J.J. (1988) *CRC Crit. Rev. Biochem.* **23**, 1-26

Richards, J.H. (1970) in *The Enzymes, Vol. 2*, Boyer, P., Ed.; Academic Press: New York. pp 321-333

Robertson, B.D., Frosch, M. and van Putten, J.P.M. (1994) *J. Bacteriol.* **176**, 6915-6920

Rose, I.A. and O'Connell, E.L. (1969) *J. Biol. Chem.* **244**, 6548-6557

Rose, I.A. (1966) *Annu. Rev. Biochem.* **35**, 23-56

Rosenberg, B.H., Lade, B.N., Chui, D., Lin, S.-W., Dunn, J.J. and Studier, F.W. (1987) *Gene* **56**, 125-135.

Rudolph, F.B., Baugher, B.W. and Beissner, R.S. (1979) *Methods Enzymol.* **63**, 22-41

Sala, R.F., Morgan, P.M. and Tanner, M.E. (1996) *J. Am. Chem. Soc.* **118**, 3033-3034

Salo, W.L. and Fletcher, H.G. (1970) *Biochemistry* **9**, 882-885

Salo, W.L. (1976) *Biochim. Biophys. Acta* **452**, 625-628

Segel, I.H. (1975) in *Enzyme Kinetics: Behaviour and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*; John Wiley and Sons: New York. pp 83-89.

Seyema, Y. and Kalckar, H.M. (1972) *Biochemistry* **11**, 36-40

Skarzynski, T., Kim, D.H., Lees, W.J., Walsh, C.T. and Duncan, K. (1998) *Biochemistry* **37**, 2572-2577

Sommar, K.M. and Ellis, D.B. (1972) *Biochim. Biophys. Acta* **268**, 581-589

Sommar, K.M. and Ellis, D.B. (1972) *Biochim. Biophys. Acta* **268**, 590-595

Spivak, C.T. and Roseman, S. (1966) *Methods Enzymol.* **9**, 612-615

Stäsche, R., Hinderlich, S., Weise, C., Karin, E., Lucka, L., Moorman, P. and Reutter, W. (1997) *J. Biol. Chem.* **272**, 24319-24324

Stevenson, G., Neal, B., Liu, D., Hobbs, M., Packer, N.H., Batley, M., Redmond, J.W., Lindquist, L. and Reeves, P. (1994) *J. Bacteriol.* **176**, 4144-4156

Stryer, L. (1988) in *Biochemistry, 3rd ed.*; W.H. Freeman and Co.: New York. p. 84.

Tanner, M.E., and Kenyon, G.L. (1998) in *Comprehensive Biological Catalysis: A Mechanistic Reference, Vol. 2: Reactions of Nucleophilic/Carbanionoid Carbon*, Sinnott, M.L. Ed.; Academic Press: NewYork. pp. 7-41.

Tanner, M.E. (1998) in *Comprehensive Biological Catalysis: A Mechanistic Reference, Vol. 3: Radical Reactions and Oxidation/Reduction*, Sinnott, M.L. Ed.; Academic Press: New York. pp. 76-82.

Thoden, J.B., Frey, P.A. and Holden, H.M. (1996) *Biochemistry* **35**, 2557-2566

Thoden, J.B., Frey, P.A. and Holden, H.M. (1997) *Biochemistry* **36**, 1212-1222

Thorson, J.S., Lo, S.F. and Liu H.-W. (1993) *J. Am. Chem. Soc.* **115**, 5827-5828

Tomasz, A. (1981) *Reviews of Infectious Diseases* **3**, 190-211

Tsui, F.-P., Boykium, R.A. and Egan, W. (1982) *Carbohydr. Res.* **102**, 263-271

Tuomanen, E., Liu, H., Hengstler, B., Zak, O. and Tomasz, A. (1985) *Journal of Infectious Diseases* **151**, 859-867

Valtonen, M.V., Larinkari, U.M., Plosila, M., Valtonen, V.V. and Makela, P.H. (1976) *Infect. Immun.* **7**, 386-392

Verma, N. and Reeves, P. (1989) *J. Bacteriol.* **171**, 5694-5701

Voet, D. and Voet, J.G. (1990a) in *Biochemistry, 2nd ed.*; John Wiley and Sons: New York. p 264.

Voet, D. and Voet, J.G. (1990b) in *Biochemistry, 2nd ed.*; John Wiley and Sons: New York. p 935.

Walsh, C. (1979) in *Enzymatic Reaction Mechanisms*; W.H. Freeman and Co.: San Francisco. pp 109-123

Waley, S.G., Miller, J.C., Rose, I.A. and O'Connell, E.L. (1970) *Nature* **227**, 181

Wee, T.G., Davis, J. and Frey, P.A. (1972) *J. Biol. Chem.* **247**, 1339-1342

Wee, T.G. and Frey, P.A. (1973) *J. Biol. Chem.* **248**, 33-40

Wierenga, R.K., De Maeyer, M.C.H. and Hol, W.G.J. (1985) *Biochemistry* **24**, 1346-1357

Wierenga, R.K., Terpstra, P. and Hol, W.G.J. (1986) *J. Mol. Biol.* **187**, 101-107

Wilson, D.B. and Hogness, D.S. (1964) *J. Biol. Chem.* **239**, 2469-2481

Wittman, V. and Wong, C.-H. (1997) *J. Org. Chem.* **62**, 2144-2147

Wood, W.A. and Gunsalus, I.C. (1951) *J. Biol. Chem.* **190**, 403-416