STRUCTURAL ANALYSIS OF

AMINOGLYCOSIDE MODIFYING ENZYMES

STRUCTURAL ANALYSIS OF AMINOGLYCOSIDE MODIFYING ENZYMES: TOWARDS RATIONAL DRUG DESIGN

By

JEFF SCHWARTZENHAUER, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (1999)
(Biochemistry)

McMaster University
Hamilton, Ontario

TITLE: Structural Analysis of Aminoglycoside Modifying Enzymes: Towards Rational Drug Design.

AUTHOR: Jeff Schwartzenhauer, B.Sc. (McMaster University)

SUPERVISOR: Dr. Albert Berghuis

NUMBER OF PAGES: xii, 115

Abstract

Bacterial resistance to the aminoglycoside antibiotics is a major health concern because of the elimination of a therapeutic option for the treatment of nosocomial infections. Clinical resistance is commonly caused by the acquisition of genes that encode an aminoglycoside modifying enzyme. These enzymes offer a potential therapeutic target in the fight against aminoglycoside resistance. By gaining a structural understanding of these enzymes the potential is created for rational drug design. The research presented here deals with structural studies on two aminoglycoside resistance enzymes. First the initial stages of structural determination for the bifunctional Aminoglycoside 6'-N-Acetyltransferase Aminoglycoside 2"-O-Phosphotransferase (AAC(6')-APH(2")) including the optimization of the purification procedure for this enzyme. The second enzyme is the Aminoglycoside 3'-O-Phosphotransferase (APH(3')-IIIa). Computational studies on this enzyme have been carried out in order to determine models for aminoglycoside binding and also to search for potential enzyme inhibitors. The molecular docking studies for both the aminoglycoside binding and inhibitor search involved the development of a number of novel methods to improve the chance of obtaining a correct model, and to aid in the analysis of the data from the docking studies. These methods have the potential to be applied in future structure based drug design

Acknowledgements

I would like to thank my supervisor Dr. Albert Berghuis for all his assistance and helpful suggestions over my tenure in his lab. I truly enjoyed working in his lab and appreciated his door always being open when I needed help brainstorming new ideas and directions for my projects. I would also like to thank the members of my supervisory committee Dr. Daniel Yang and Dr. Eric Brown for numerous useful suggestions and criticisms of my work. In addition Dr. Gerry Wright and members of his lab Paul Thompson and Denis Daigle were a great assistance throughout my work in the lab. Their collaborative efforts and helpful suggestions on my projects were greatly appreciated. Thanks also to the members of Dr. Berghuis' lab who created an atmosphere that was both fun and exciting to work in. Thanks to Laura Rossi, Desiree Fong and David Burk for editing my thesis and for other assistance, and thanks to Byron DeLaBarre and Leanne Wybenga-Groot for helpful suggestions over the years. Additionally I would like to thank my parents George and Pat Schwartzenhauer for their continued support over the years and Jennifer Prince for keeping me motivated and supporting me during the writing of this thesis.

Table of Contents

List of Abbreviations

AAC – Aminoglycoside acetyltransferase
ADP – Adenosine 5'-diphosphate
AMBER – Assisted Model Building with Energy Refinement
AMPPNP – Adenylylimido-diphosphate
ANT – Aminoglycoside nucleotidyltransferase
APH – Aminoglyocoside phosphotransferase
ATP – Adenosine 5'-triphosphate
BSA – Bovine serum albumin
CASP2 – Critical Assessment of Methods of Protein Structure Prediction: Round II
CoA – Coenzyme A
DLS – Dynamic light scattering
DNA – 2' deoxyribonucleic acid
DTDP – 4,4'-dithiodipyridine
DTT – Dithiothreitol
EDTA – Ethylenediamine tetraacetic acid
FE – Free Energy
GNAT – GCN5 – related *N*-Acetyltransferase
GRASP – Graphical Respresentation and Analysis of Surface Properties
HB Score – Hydrogen Bond Score
HEPES – (N-[2-hydroxyethyl] piperazine-N'-[2-ethane sulfonic acid])
IPTG – Isopropyl β-D-thiogalactopyranoside
LB – Luria-Bertani medium
MES – (2-[N-morpholino] ethane sulfonic acid
MOPS – (3-[N-morpholino] propane sulfonic acid
MMFF – Merck Molecular Force Field
mRNA – Messenger ribonucleic acid
MPD – 2-methyl-2,4-pentanediol
NADH - β-Nicotinamide adenine dinucleotide, reduced form
NCI – National Cancer Institute
NMP – Nucleotide monophosphate
NMR – Nuclear Magnetic Resonance
NPFE – Non Polar Free Energy
O.D. – Optical density
PDB – Protein Data Bank
PEG – Polyethylene glycol
PEP – Phosphoenol pyruvate
PK/LDH – Pyruvate kinase / lactate dehydrogenase
PMSF – Phenylmethylsulfonyl fluoride

List of Figures

Page Number

List of Tables

Page Number

Chapter 4

Chapter 1: Introduction

1.1 Antimicrobial Discovery, Mechanisms, and Resistance

The discovery of antimicrobial compounds over the past 70 years has been a key part of medical history. These compounds have played major roles in the fight against numerous bacterial infections. The antimicrobial compounds are mostly natural products isolated from bacteria or fungi, or derivatives of these compounds. The majority of antibiotic compounds were discovered in the relatively short period of time between the 1930's and the 1960's. The confidence in these compounds has been so great that in 1969 the U.S. Surgeon General stated that it was time to "close the book on infectious disease" (Bloom and Murray, 1992). However, despite this confidence, infectious diseases are still among the leading causes of death worldwide (Berkelman and Hughes, 1993).

Fleming's accidental discovery that *Penicillin notatum* was able to inhibit the growth of staphylococci in 1928 (Fleming, 1929) was the beginning of the antibiotic era. This initial discovery led to the work of Howard Florey and colleagues, a decade later, who isolated and purified the compound penicillin, the first β-lactam antibiotic, and eventually put it into clinical practice in 1941 (Kunin, 1993). However, even before

the first use of penicillin a different class of antimicrobial compound had been discovered and put into clinical practice. These compounds were the sulfonamides discovered in 1935 by Dogmagk. The first sulfonamide was Prontosil used for the prevention and cure of many bacterial infections (Mandell and Petri, 1996a). After the accidental discovery of penicillin many researchers purposely used Fleming's methodology to actively search for new compounds that could inhibit the growth of bacteria. One such researcher was René Jules Dubos. In the late 1930's Dubos worked with *Bacillus brevis* and was able to isolate the compound tyrothricin, which was later separated into gramcidin and tyrocidin, both of which were effective as topical antibiotics for the treatment of external infections caused by cocci or bacilli (Böttcher, 1959). Another researcher who actively pursued the possibility of fungi or bacteria producing compounds able to inhibit the growth of bacteria was Selman Waksman. In 1944 his work focusing on the streptomycetes was rewarded with the discovery of the first aminoglycoside antibiotic streptomycin (Schatz et al., 1944). Later he discovered the first 2-deoxystreptamine aminoglycoside antibiotic neomycin (Waksman and Lechevalier, 1949). The forties also saw the discovery of other classes of antibiotics, including the tetracyclines, discovered by Duggar from streptomycetes in 1945 (Böttcher, 1959). In addition, the cephalosporins belonging to the β-lactam family, were discovered by Brotzu in 1948 (Mandell and Petri, 1996b), and Bartz discovered chloramphenicol in 1948 (Bartz, 1948). The fifties continued on the successes of the forties with the discovery of erythromycin, from streptomycetes, in 1952 by McGuire (Kapusnik-Uner et al., 1996), this being the most used macrolide antibiotic. And then the 1956 discovery of

vancomycin, from streptomycetes, the first glycopeptide (McCormick et al., 1956). The sixties saw the introduction of the quinolones (Gootz, 1990) which were improved upon in the seventies with the introduction of the fluoro-quinolones (Mandell and Petri, 1996a). However since this time, the discovery or development of new antibiotic compounds has been virtually non-existent.

Table 1.1.1: Discovery of Antibiotics: Major classes, dates and researcher responsible for discovery.

Antibiotic Class	Date first discovered	Researcher
Sulfonamides	1935	Dogmagk
β-lactams	1929	Fleming
Aminoglycoside	1944	Waksman
Tetracyclines	1948	Duggar
Chloramphenicol	1948	Bartz
Macrolides	1952	McGuire
Glycopeptides	1956	McCormick

Antimicrobial compounds act on a variety of targets in the bacterial cell that are essential either for maintaining life or for reproduction of the bacteria. Antibiotics that are able to kill the bacteria are termed bactericidal and those that only inhibit its growth and reproduction are termed bacteriostatic. The mechanisms of the various classes of antimicrobials are reviewed briefly by Richard Wise (Wise, 1999), and are summarized in figure 1.1.1 adapted from Neu (1992). The β-lactam antibiotics (penicillins and cephalosporins) act by binding to peptidoglycan transpeptidases, these enzymes are involved in the synthesis of the bacterial cell wall (peptidoglycan layer) in bacteria.

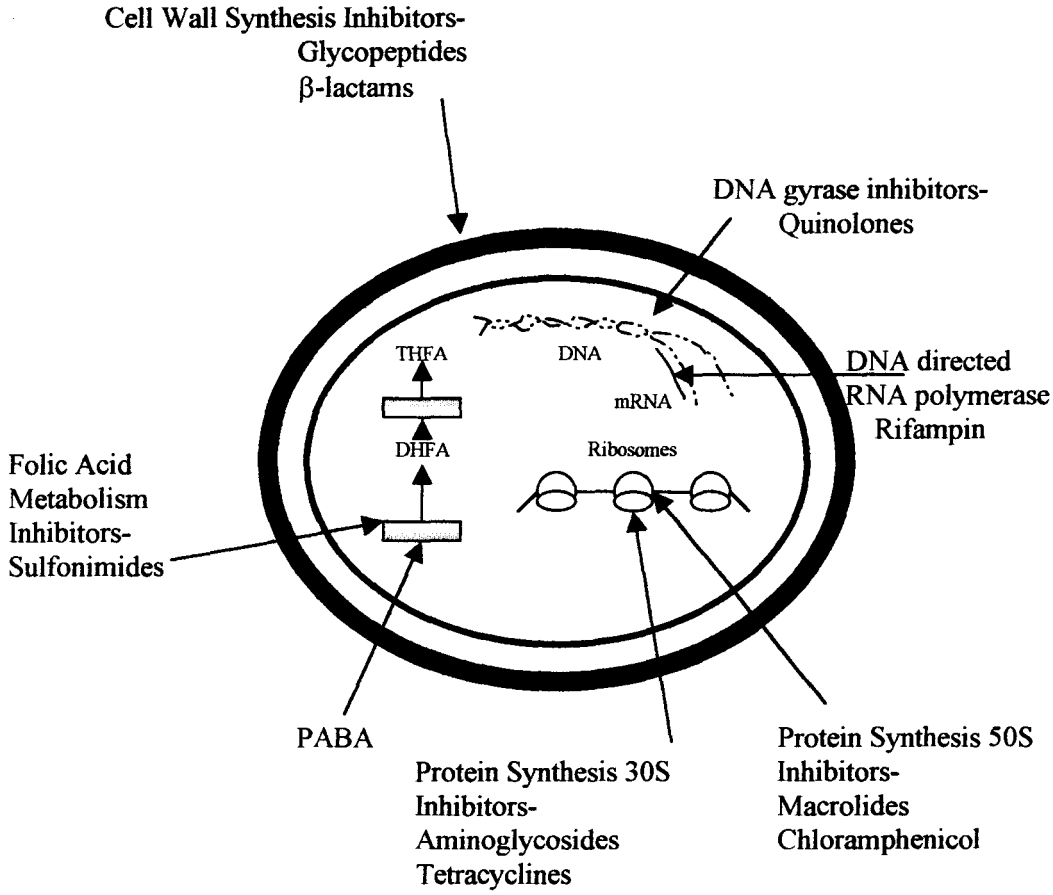


Figure 1.1.1: Sites of action of various antibiotics. Adapted from Neu (1992).

Binding to these enzymes prevents cross-linking in the peptidoglycan leading to cell lysis and therefore cell death. The glycopeptides also work by preventing the formation of the peptidoglycan layer in gram-positive bacteria. These drugs bind to the D-alanyl-D-alanine preventing cross-linking from occurring during cell wall synthesis leading to cell lysis and death of the bacteria. The quinolones bind to the alpha subunit of DNA gyrase (topoisomerase) thereby inhibiting DNA supercoiling and leading to bacterial cell death.

Sulfonamides are analogues of para-aminobenzoic acid, the precursor of folic acid, thus they act in a bacteriostatic way by inhibiting folic acid synthesis. The tetracyclines and macrolides each act in a bacteriostatic way by binding to bacterial ribosomes and interfering with the synthesis of proteins. Macrolides inhibit translocation, whereas tetracyclines prevent the binding of tRNA to the ribosome preventing elongation of the polypeptide chains. Like the tetracyclines and macrolides the aminoglycosides act on the bacterial ribosome in such a way as to cause misreading of mRNA. This creates downstream effects that eventually lead to cell death, this will be discussed in greater detail in section 1.2.

With the variety of antimicrobial mechanisms, bacteria have adopted a number of different strategies to defend themselves against these antibiotics. Shortly after the initial use of the antibiotic penicillin, resistance to this drug in staphylococci was noticed in hospitals (Neu, 1992; Kunin, 1993). Streptomycin was a similar story, the initial excitement surrounding the discovery of this drug led to its overuse, and resistance was beginning to appear quickly (Kunin, 1993). Recently resistance to other antibiotics has appeared including resistance in Enterococci and Staphylococci to Vancomycin (Perl, 1999), which was once termed the antibiotic of last resort for the treatment of infections caused by pathogenic gram-positive bacteria. The development of resistance to antibiotics has led to the reemergence of "old" diseases that were once treatable including tuberculosis (Berkelman and Hughes, 1993). As this happens the number of strains resistant to the various antibiotics also grow.

Four general strategies exist for resistance (Neu, 1992). The first strategy involves the alteration of the antibiotic's target. The second is the inactivation of the antibiotic itself by some sort of covalent alteration. The third mechanism is alteration of the permeability of the bacteria to the antibiotic. The final mechanism of antibiotic resistance is the active efflux of the antibiotic, to remove the drug from the cell. The resistance mechanisms used against the various classes of antimicrobial compounds are presented in Table 1.1.2.

Table 1.1.2: Method of bacteria resistance to various antibiotic classes

Antibiotic Class	Mechanisms of Resistance
β–lactams	Alteration of target (Penicillin binding proteins) Reduction in permeability Detoxifying enzymes (β-lactamases)
Quinolones	Alteration of target (DNA gyrase) Reduction in permeability
Chloramphenicol	Detoxifying enzymes (Chloramphenicol Acetyltransferases)
Tetracycline	Alteration of target (Ribosomes) Efflux of Drug Enzymatic detoxification
Macrolides	Enzymatic alteration of target (rRNA)
Sulfonamides	Reduced permeability Alteration of target
Glycopeptides	Alteration of target (D-alanyl-D-alanine)
Aminoglycosides	Alteration of target (rRNA) Reduction in uptake of drug Enzymatic detoxification (Aminoglycoside modifying enzymes)

Information regarding various mechanisms of resistance was adapted from Neu (1992), and Davies (1994).

Generally, antibiotic resistance elements are genetically encoded either on the bacterial chromosome, a plasmid or a transposon. In order for the resistance mechanism to exist on the chromosome either the organism in question has intrinsic resistance or a mutation occurred due to selective pressure resulting in the resistance. In cases where a particular strain of bacteria are first sensitive then resistant to a particular antibiotic, there must be some mutational events occurring, this is often involved in the alteration of the antibiotic target or the permeability of the cell to the drug (Davies, 1994). The mechanism of alteration of the drug itself is more likely encoded on a plasmid and is therefore able to be transferred from one bacteria to another and from one species to another (Chen and Williams, 1985).

1.2 Aminoglycoside Antibiotics

The aminoglycosides antibiotics have been an important class of antibiotics for the treatment of infections caused by both Gram-positive and Gram-negative organisms (Edson and Terrell, 1991). The first two aminoglycosides, streptomycin (Schatz et al., 1944) and neomycin (Waksman and Lechevalier, 1949), were discovered in the forties. Streptomycin was put into use for the treatment of tuberculosis (Hinshaw and Feldman, 1945) and it continues to be a major part of the anti-mycobacterial therapy (Musser, 1995). The utility and effectiveness of these early compounds led to the identification and characterization of many more aminoglycoside antibiotics over the three decades following their initial discovery (reviewed in Wright et al., 1999).

The aminoglycoside antibiotics are all water soluble, cationic molecules that possess a six-membered aminocyclitol ring. Aminoglycosides are divided into two structural classes, those with a 2-deoxystreptamine ring and those without a 2-deoxystreptamine ring. The 2-deoxystreptamine aminoglycosides can be further subdivided into two classes, those that are substituted at the 4 and 5 positions of the 2-deoxystreptamine and those that are substituted at the 4 and 6 positions of the 2-deoxystreptamine ring. The aminohexose at the 4 position is conventionally labeled the (') ring and the pentose or hexose at the 5 or 6 position is labeled the (") ring (see Figure 1.2.1). The amino sugars at the 4, 5 and 6 positions are linked to the central aminocyclitol ring through glycosidic linkages. Aminoglycosides are widely produced by bacteria of the actinomycete group as well as other bacteria and can also be derived semi-synthetically from these natural products (reviewed in Wright et al., 1999).

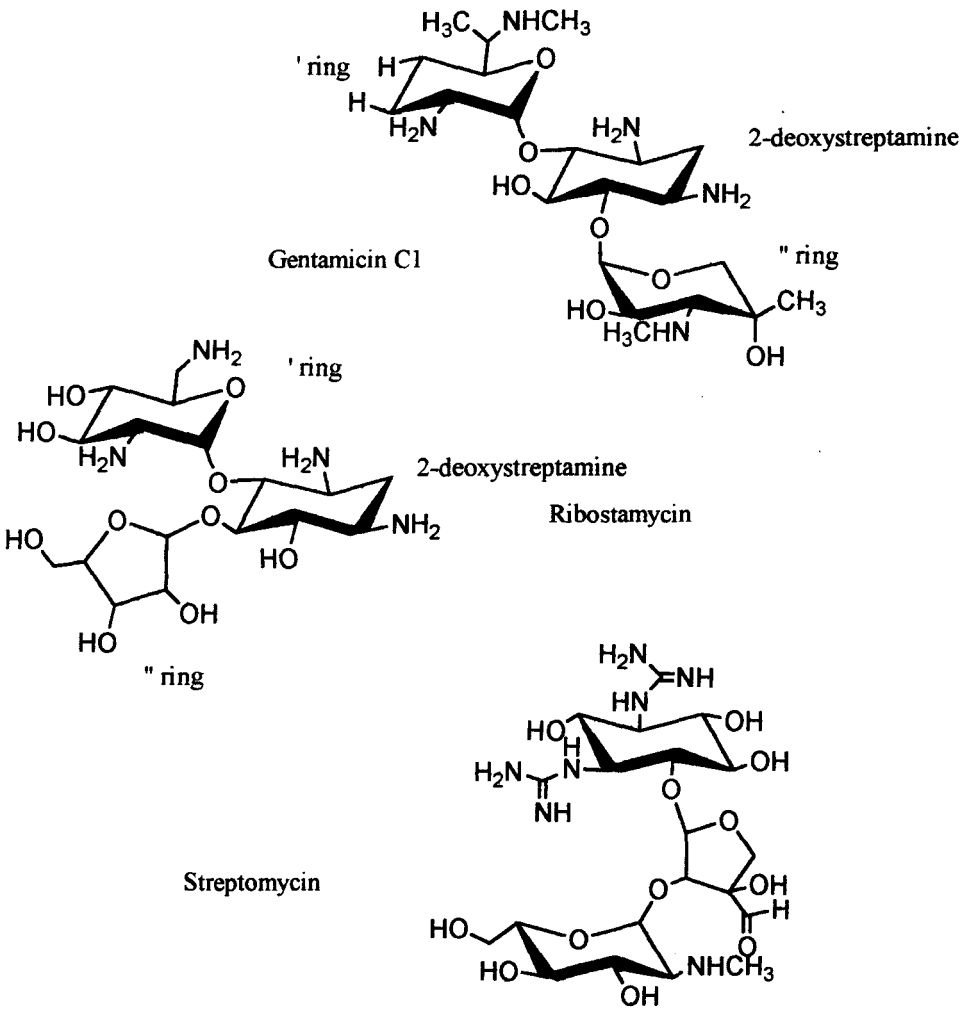


Figure 1.2.1: Structures of various aminoglycoside antibiotics. 4,6-disubstituted 2-deoxystreptamine gentamicin C1, 4,5-disubstituted 2-deoxystreptamine ribostamycin, and non 2-deoxystreptamine aminoglycoside Streptomycin. The ' and " rings on gentamicin C1 and ribostamycin are indicated.

Aminoglycosides have been the principal weapon for the treatment of seriously ill patients suffering from a variety of bacterial infections (Price, 1986). Aminoglycoside antibiotics have been extensively used against many clinically important gram-negative bacteria including *Klebsiella, Enterobacter, Serratia, Citrobacter, Pseudomonas*

aeroginosa, Hafnia alvei, and *Acinetobacter* species (reviewed by Lortholary, 1995). In addition to their effectiveness against the gram-negative bacteria, aminoglycosides are used against many gram-positive bacteria. Streptomycin was found to be effective in treating bacterial endocarditis that was caused by penicillin resistant streptococci (Hunter, 1947). This finding led to extensive research of streptomycin and subsequently discovered aminoglycosides for the treatment of gram-positive bacterial infections. Jawetz et al. (1950) found that the combination of stretomycin with penicillin was effective at killing enterococci *in vitro*. The mechanism for this synergism appears to be the increased uptake of the streptomycin in the presence of penicillin (Moellering et al., 1971). In addition to the synergism between penicillin and streptomycin, other aminoglycosides are also able to synergistically kill gram-positive bacteria when combined with penicillin. These include kanamycin (Standiford et al., 1970), tobramycin, and gentamicin (Moellering et al., 1973; Weinstein and Moellering, 1973).

One of the most significant factors that makes aminoglycosides useful compounds is that unlike other protein synthesis inhibiting antibiotics the aminoglycosides are bactericidal. Other antibiotics, such as the tetracycline, chloramphenicol and the macrolides that bind to the same target as the aminoglycosides (the bacterial ribosome) are only bacteriostatic. This means that the aminoglycosides must have another activity that leads to the death of the bacteria. The aminoglycosides are known to cause misreading during polypeptide synthesis by interfering with the ribosome-messenger RNA complex (Davies et al., 1964). Footprinting studies have shown that the

aminoglycosides interact specifically with the A-site of the 16S ribosomal RNA (Moazed and Noller, 1987; Woodcock et al., 1991). A recent NMR structure of paromomycin bound to a region of the 16S rRNA shows this aminoglycoside bound in the A-site of the ribosome (Fourmy et al., 1996). The actual mechanism by which the aminoglycosides are bactericidal is not yet fully understood. There seems to be a series of pleiotropic events that ultimately lead to death of the bacterial cell. Four important events include the blockade of the ribosome, misreading in translation, membrane damage, and irreversible uptake of drug (Davis, 1987). There also is some evidence that points towards a direct interference on DNA replication by the aminoglycosides (Tanaka et al., 1984; Matsanuga et al., 1986). Davis (1988) offers one mechanism for aminoglycoside action that could account for all the effects triggered by the antibiotic. First a small amount of antibiotic enters the cell, this may be through a multiphasic process involving protein mediators (reviewed in Wright et al., 1999) the aminoglycoside would then bind to the ribosome leading to the formation of misread proteins. The misread proteins could target and insert themselves into the membrane and create pores or holes in the membrane (Davis et al., 1986). With the holes in the membrane it is conceivable that a large amount of aminoglycoside could enter into the cell causing a variety of effects and eventually leading to cell death of the bacteria (reviewed in Davis, 1988; Wright et al., 1999; Mingeot-Leclerq et al., 1999). In addition to their ability to kill the bacterial cells, the aminoglycosides are also very specific for prokaryotic ribosomes. The antibiotics bind to a conserved region of the ribosome, however, a single nucleotide difference at

position 1408 (based on *E.coli* numbering) makes the prokaryotic ribosome sensitive to the aminoglycoside and the eukaryotic ribosome resistant (Recht, 1999).

1.3 Aminoglycoside Resistance

Shortly after the initial introduction of the aminoglycoside antibiotics into clinical practice, bacterial resistance to these compounds began to appear (Kunin, 1993). This resistance has grown and spread to a wide variety of bacteria. One of the most significant bacteria that have acquired resistance to the aminoglycosides are the gram-positive cocci such as *Streptococcus*. Enterococci, in particular those resistant to aminoglycosides, have become one of the most prevalent causes of nosocomial infections (Zervos et al., 1987; Schaberg, 1991). These resistant bacteria which were typically treated through synergism of a penicillin and an aminoglycosides can no longer be killed by this combination therapy (Moellering et al., 1979; Moellering et al., 1980; Wennersten and Moellering, 1980; Eliopoulos et al., 1988).

A number of mechanisms have been used by bacteria to help them survive the onslaught of aminoglycoside antibiotics. One such mechanism is the mutation of the ribosomal site to which the aminoglycoside binds. A second mechanism to fight the aminoglycoside antibiotics is to prevent influx of the compound or alternatively efflux the drug. Either of these can create the resistance phenotype, however these mechanisms are rarely observed in clinical isolates that possess aminoglycoside resistance. The third

and most common form of clinical resistance to aminoglycosides is the enzymatic covalent modification of the antibiotic (Zimmerman et al. 1971; Krogstad et al. 1978a, Shaw et al. 1993). Widespread aminoglycoside resistance is due to the transfer of the genetic material encoding the aminoglycoside modifying enzymes. Transfer can occur from one strain of bacteria to another on plasmids or other pieces of transferable DNA, such as transposons (Courvalin et al., 1978; Krogstad et al., 1978b; Chen and Williams, 1985; reviewed in Davies, 1994; reviewed in Davies and Wright 1994).

The aminoglycoside modifying enzymes fall into three distinct categories based on the reaction that they catalyze. The first class of enzymes are the aminoglycoside phosphotransferases (APH). These enzymes are kinases which act by phosphorylating the aminoglycosides on hydroxyl functional group, ATP is the source of this phosphate. There are a large number of APH enzymes with a variety of substrate ranges and regiospecificities (Shaw et al., 1993). The different classes of APH's are discussed in Wright et al. (1999) and Wright and Thompson (1999). The three dimensional structure of a single member of this family APH(3')-IIIa has been determined (Hon et al., 1997). This structure showed significant structural homology to the eukaryotic protein kinase, with very little sequence homology (Hon et al., 1997). This result suggests the possibility of divergent evolution of these enzymes with the APH enzymes sharing an ancient ancestor with the eukaryotic protein kinases. APH(3')-IIIa will be further discussed in Chapter 3.

A second group of enzymes that can modify the aminoglycosides are the aminoglycoside acetyltransferases (AAC). These enzymes transfer an acetyl group from an acetyl-CoA cofactor to various amino or in one case a hydroxyl group on the aminoglycoside antibiotics. The AAC's like the APH's have numerous members in the family, and each member has a specific substrate range and regiospecificity (Shaw et al., 1993). The various members of this family of enzymes are reviewed by Wright et al. (1999). Two members of this family of enzymes have had their three dimensional structure determined AAC(3) (Wolf et al., 1998) and AAC(6')-Ii (Wybenga-Groot et al., 1999). Both of these members have been shown to possess a very similar fold to the histone acetyltransferase and to the GNAT superfamily of acetyltransferases (Wybenga-Groot et al., 1999). This finding suggests the possibility of divergent evolution relating the aminoglycoside acetyltransferase to eukaryotic enzymes that also carry-out an acetyltransfer reaction.

The third class of proteins that modify aminoglycosides are the aminoglycoside nucleotidyltransferases (ANT). These enzymes act by the transfer of a nucleotide group from ATP to a hydroxyl group on the aminoglycoside antibiotic. The ANT family, like the APH and AAC family, has a number of members, each having a specific range of substrates and regiospecificity (Shaw et al., 1993). The various members of this family are reviewed by Wright et al. (1999). The three dimensional structure of one member of the ANT family, ANT(4') has been solved by x-ray crystallography (Pedersen et al., 1995). This enzyme appears to share a fold similar to that of a DNA polymerase β

(Holm and Sander, 1995; Wright et al., 1999). Again this suggests the possibility of divergent evolution between the ANTs and the NMP transferases.

The three groups of aminoglycoside-modifying enzymes provide significant therapeutic targets for the treatment of bacterial infections caused by aminoglycoside resistant bacteria (Coleman et al., 1994). In order to develop therapeutic agents that are able to avoid or overcome the enzymatic modification by aminoglycoside modifying enzymes it is important to understand the function and structure of the enzymes (Gootz, 1990). This fact has led to studies that have been undertaken in the research presented here. Through understanding the structure of the enzyme the possibility is created to develop either enzyme inhibitors or new antibiotic compounds which are not susceptible to the enzymatic modification. Various structural studies have been undertaken on two specific aminoglycoside-modifying enzymes. These studies include crystallization efforts on a bifunctional aminoglycoside acetyltransferase and aminoglycoside phosphotransferase as well as computational modeling of an aminoglycoside phosphotransferase.

Chapter 2: Purification and Crystallization of Bifunctional AAC(6')-APH(2")

2.1 Introduction

2.1.1 Bifunctional Aminoglycoside 6'-N-Acetyltransferase - Aminoglycoside 2"-O-Phosphotransferase (AAC(6')-APH(2"))

The gentamicin C complex of aminoglycosides is extensively used in clinical practice (Davies, 1991, Edson and Terrell, 1991). Most high-level gentamicin C resistance in gram-positive bacteria is caused by the AAC(6')-APH(2") protein (Kaufhold et al., 1992). The enzyme is encoded by the gene *aac(6')-aph(2")* which is found extensively in both *Enterococci* and *Staphylococci*. This 57 kDa protein possesses both an AAC and an APH domain (Daigle et al., 1999). The gene for the bifunctional enzyme is located on transposable genetic elements or an R-plasmid (Rouch et al. 1987; Hodel-Christian and Murray, 1991), likely allowing for the widespread dissemination of the resistance profile (Patterson and Zervos, 1990). The gene that encodes the AAC(6')-APH(2") protein arose from a gene fusion event with the N-terminus encoding the acetyltransferase activity and the C-terminus encoding the phosphotransferase activity (Ferretti et al., 1986). In addition, the two activities can be individually expressed as separate 21 and 36 kDa proteins (private communication, Denis Daigle, McMaster University, Department of Biochemistry).

The bifunctional AAC(6')-APH(2") enzyme itself has an extensive range of substrates, it can modify virtually any 2-deoxystreptamine aminoglycoside (Daigle et al., 1999). The enzyme has recently been shown to possess the ability to N and O- acetylate aminoglycosides at the 6' site. In addition, phosphorylation has been observed at the 2", 3', 5", and 3"" sites depending on the aminoglycoside that is being modified (Figure 2.1.1) (Daigle et al., 1999). The full-length bifunctional enzyme from Enterococci has been expressed in *Bacillus subtilis* under the control of the constitutive *vegII* promoter (Daigle et al., 1999). The individual AAC(6') and APH(2") domains have been expressed in *Escherichia coli* (*E.coli*) under the control of a IPTG inducible promoter (private communication, Denis Daigle, McMaster University, Department of Biochemistry). Purification procedures for the full-length enzyme (Daigle et al., 1999) and the individual domains (private communication, Denis Daigle, McMaster University, Department of Biochemistry) have been established. However, the purity of the protein obtained from these procedures was not sufficient for crystallization efforts. Optimization of these procedures was therefore necessary before crystallization experiments could take place.

The clinical importance of this enzyme is obvious with its broad range of substrates and activities. For this reason structural studies were undertaken in an effort to begin structure based rational drug design. This chapter describes the efforts that have been taken in order to obtain a three-dimensional structure of either the full-length bifunctional protein or the APH(2") domain alone.

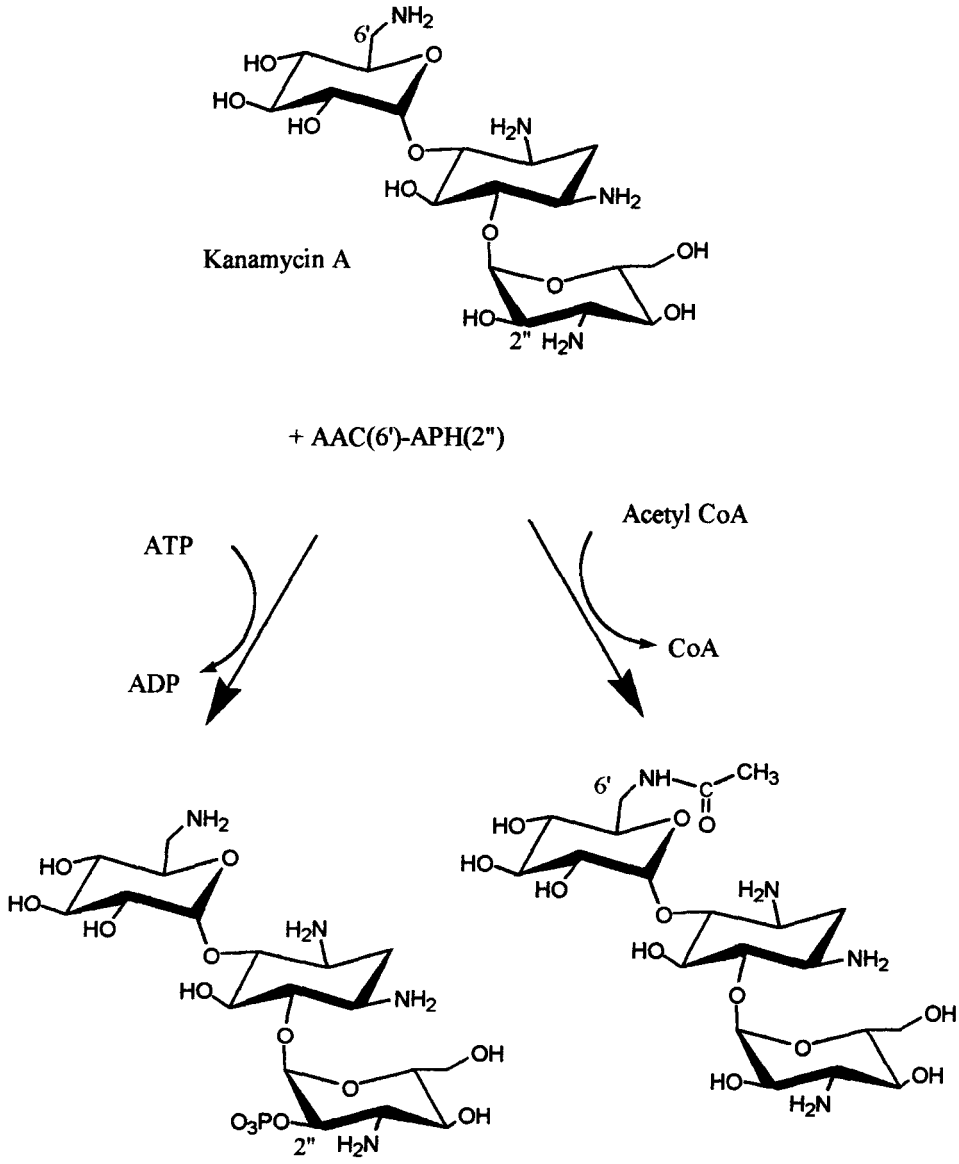


Figure 2.1.1: Typical sites of aminoglycoside modification by bifunctional AAC(6')-APH(2"). The 6' amino group is acetylated and the 2" hydroxyl group is phosphorylated. Additional sites of modification are discussed in Daigle et al. (1999).

2.1.2 Dynamic Light Scattering and Protein Crystallization

Dynamic Light Scattering (DLS) is a relatively new method used to assist in the selection of conditions to be used for protein crystallization trials. This technique measures the translational diffusion coefficient of a macromolecule undergoing Brownian motion in solution. This measurement allows for the determination of the hydrodynamic radius using the Stokes Einstein equation. The radius can be used for providing a molecular weight estimate based on a standard curve for globular proteins. Using the molecular weight estimates one can determine if the solution is made up of a single species or of multiple species through analysis of the standard deviation of the estimates. In this way, using DLS can allow one to predict the aggregation state and the molecular weight of the macromolecule (Ferré-D'Amaré and Burley, 1997). The degree of aggregation or polydispersity has been used to predict a protein's likelihood to crystallize (Zulauf and D'Arcy, 1992; D'Arcy, 1994). Estimating the likelihood of a protein to crystallize and which conditions are most likely to produce crystals prior to performing crystal growth experiments, can save time, effort, and materials (Ferré-D'Amaré and Burley, 1997).

Macromolecular crystallization involves achieving supersaturation conditions for the protein, and then thermodynamically promoting the exclusion of molecules from the solution leading to a net accumulation of the solid state (McPherson, 1999). Crystal growth involves three distinct stages: nucleation, growth, and cessation of growth

(McPherson, 1999). The objective in crystallization trials is to determine the optimum conditions to allow all these steps to occur and produce diffraction quality crystals. The most common method used for protein crystallization is the vapour diffusion method using either hanging or sitting drop techniques. This method works on the principle of adding a precipitate to the protein sample and slowly evaporating water in this mixture raising the concentration of both the protein and the precipitant (Rhodes, 1993). If the evaporation is sufficiently slow, the protein will precipitate leading to either crystals or an amorphous precipitate. Crystals form when the protein molecules pack together in an ordered three-dimensional lattice, the various molecules held together through a variety of non-covalent interactions (Rhodes, 1993). Numerous factors effect crystal formation and each protein behaves distinctly with respect to these factors (McPherson, 1999). These factors are often used as variables when optimizing the crystallization conditions. These include precipitant concentration, temperature, pH, and protein concentration (McRee, 1993). Initial crystallization experiments often take advantage of a matrix screen. One form is the sparse matrix (Jancarik and Kim, 1991) that employs a large variety of conditions that can hopefully produce a few conditions suitable for optimization. A second matrix is the grid matrix that looks at a specific precipitant at various concentrations and pHs. The grid matrix is useful for determining specific properties of the protein in question and may be able to lead to optimal conditions quickly (McPherson, 1999). In general, protein crystallography has no specific rules that need to be followed, the conditions that will produce crystals for a specific protein are unpredictable and anything is worth trying (McRee, 1993). The simple addition of a

compound or a change in how the protein is prepared could produce high-quality crystals and a three-dimensional structure.

2.2 Materials and Methods

2.2.1 General Materials and Methods

Table 2.2.1: Table of Reagents used in Purification and Crystallization experiments

Reagent	Supplier	Purpose
Yeast Extract	Becton Dickson	Media
Tryptone	Becton Dickson	Media
NaCl	ACP Chemicals	Media, Buffer B
IPTG	BioShops Canada	Induction of APH(2") Growth
Q-Sepharose Resin	Amersham Pharmacia Biotech	Purification
Sephadex G100 Resin	Amersham Pharmacia Biotech	Purification
Affigel-15 (N-Hydroxysuccinamide Activated) Resin	Bio-Rad	Purification
HEPES	GIBCO BRL	Purification, AAC Activity, DLS
EDTA	Sigma	Purification
PMSF	Sigma	Lysis
DTT	Sigma	Lysis
Gentamicin Sulfate	Sigma	APH Activity, Media
Tris	Boehringer Manheim	APH Activity, DLS
KCl	BDH	APH Activity
$MgCl_2$	BDH	APH Activity
ATP	Sigma	APH Activity
PK/LDH	Sigma	APH Activity
NADH	Sigma	APH Activity
PEP	Sigma	APH Activity
DTDP	Sigma	AAC Activity
Kanamycin A	Sigma	AAC Activity, Crystallization
Acetyl Coenzyme A (Acetyl CoA)	Amersham Pharmacia Biotech	AAC Activity
CoA	Sigma	DLS, Crystallization
H-9	Research Biochemicals International	Crystallization
ADP	Sigma	DLS, Crystallization
AMPPNP	Sigma	Crystallization

(**Table 2.2.1** continued)

Reagent	Supplier	Purpose
MES	Sigma	DLS
MOPS	BioShop	DLS
Na Cacodylate	Sigma	DLS
TAPS	Sigma	DLS
Bicine	Sigma	DLS
HCl	Fisher	pH Adjustment
NaOH	BDH	pH Adjustment
Crystal Screens I and II	Hampton Research	Crystallization
Additive Screens I, II, and III	Hampton Research	Crystallization
MPD Grid Screen	Hampton Research	Crystallization
Ammonium Sulfate Grid Screen	Hampton Research	Crystallization
PEG 6000 Grid Screen	Hampton Research	Crystallization
Wizard I and II Screens	Emerald BioStructures	Crystallization
Cryo I and II Screens	Emerald BioStructures	Crystallization
IZIT™ Dye	Hampton Research	Crystallization
Bradford Reagent Dye	Bio-Rad	Bradford protein Assay
BSA protein	Bio-Rad	Bradford protein Assay

Media and buffers used in the expression and the purification of the full-length bifunctional protein and the APH(2") domain were prepared according to the following specifications.

LB Media: 10 g Tryptone, 5 g Yeast Extract, and 10 g NaCl in 1 L of Deionized Water (supplemented with 1/1000 of 50 mg/ml Gentamicin Sulfate at time of innoculation)

Buffer A: 50 mM HEPES pH 7.5, 1 mM EDTA

Buffer B: 50 mM HEPES pH 7.5, 1 mM EDTA, 1 M NaCl

Lysis Buffer: 50 mM HEPES pH 7.5, 1 mM EDTA, 1 mM PMSF, 0.1 mM DTT

Activity assays for both phosphotransferase and acetyltransferase were employed at various times during the purification procedure to locate active fractions. The assays were prepared according to the following specifications.

Phosphotransferase Activity (as described in McKay et al., 1994)

Buffer: 50 mM TRIS pH 7.5, 40 mM KCl, and 10 mM $MgCl_2$

9.28 ml of Buffer was combined with 2 µl 500 mM Gentamicin Sulfate, 20 µl 500 mM ATP, 50 µl PK/LDH, 2 mg NADH and 6 mg PEP. This provided with sufficient quantity of enzyme assay reagent for ten assays. The slope at 340 nm was determined after enzyme was added. This was translated into the activity of the fraction. Kanamycin was substituted for by gentamicin to provide specifity for the APH(2") activity. The reactions involved in the phosphotransferase activity assay are shown in Figure 2.2.1.

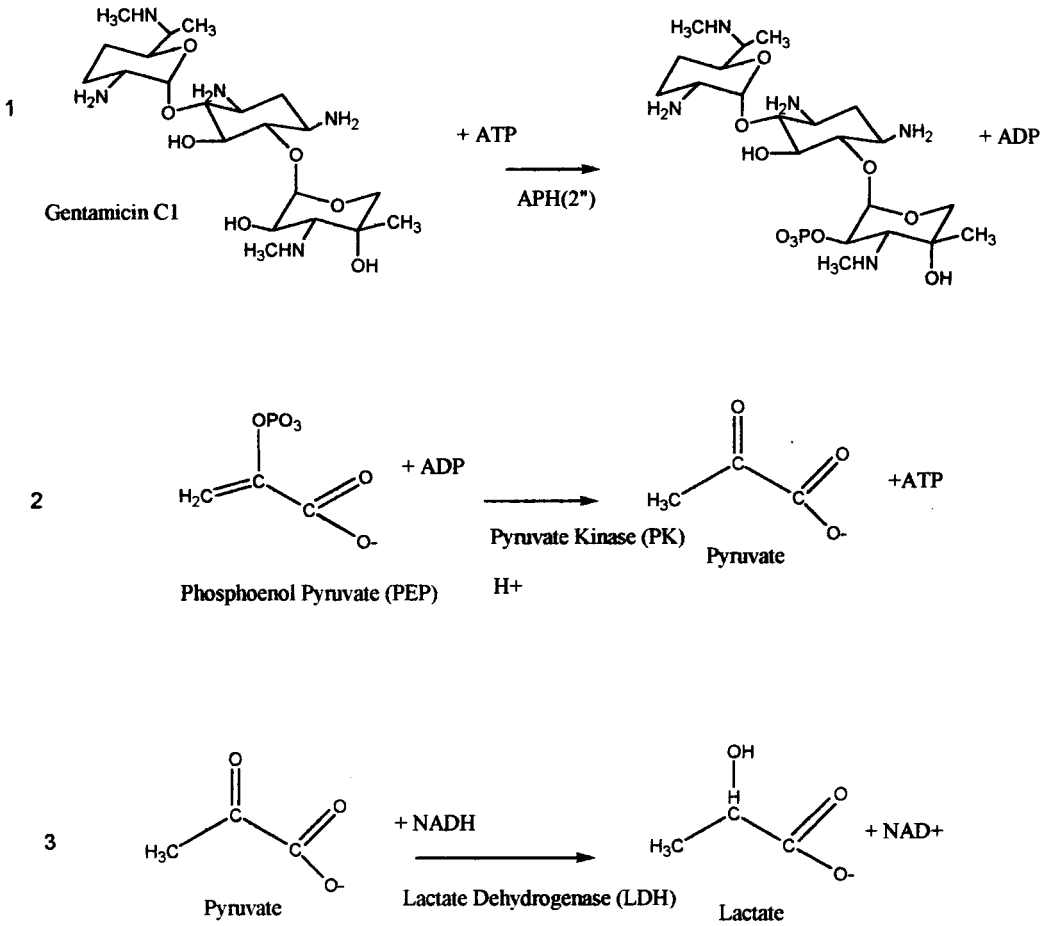


Figure 2.2.1: Reactions involved in phosphotransferase activity assay. The Assay monitors the disappearance of NADH at 340 nm by visible spectroscopy.

Acetyltransferase Activity (as described in Williams and Northrop, 1978)

Buffer: 25mM HEPES pH 7.5, 1mM EDTA

Each assay, contained 600 µl of buffer was combined with 80 µl 20 mM DTDP, 80 µl 0.7 mm Acetyl Coenzyme A, and 20 µl 1.0 mM Kanamycin A. The slope at 324 nm was

determined after the enzyme was added. This slope was translated into the activity of the fraction. The reactions involved in the acetyltransferase activity assay are shown in Figure 2.2.2.

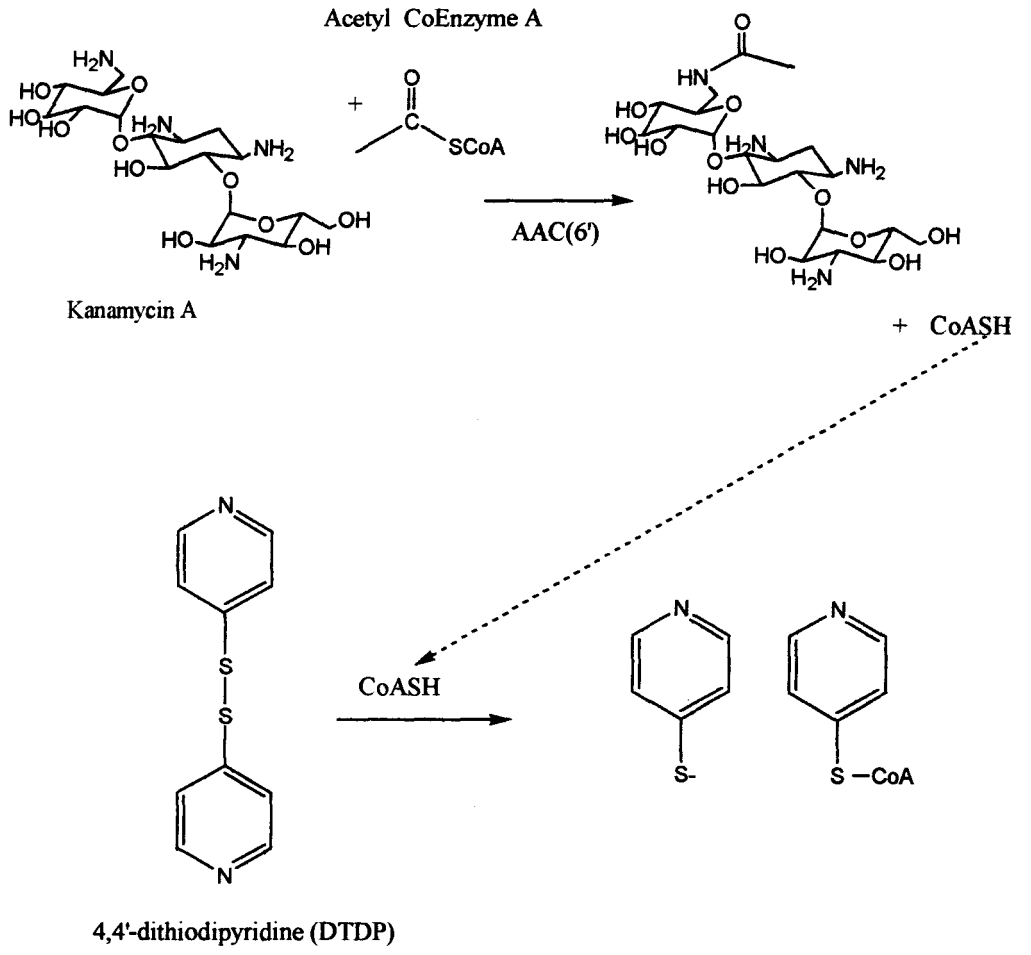


Figure 2.2.2: Reactions involved in acetyltransferase activity assay. Product formation of the second reaction was monitored at 324 nm.

Buffers used for DLS (Protein Solutions) were prepared in stock solutions to 0.5 M or 1.0 M. Stock buffers at 0.5 M were prepared for MES at pH 6.0 and 6.5 and MOPS at pH 7.0 and 7.5. The remaining buffers were prepared to a 1.0 M stock, Na

Cacodylate at pH 6.5 and 7.0, HEPES at pH 7.0 and 7.5, TRIS, TAPS and Bicine all at pH 8.0 and 8.5. Substrates used for DLS experiments included Kanamycin A, Coenzyme A, H-9, and ADP.

2.2.2 Purification Optimization

Purification protocols for the AAC(6')-APH(2") (Daigle et al., 1999) and the APH(2") (private communication, Denis Daigle, McMaster University, Department of Biochemistry) were established. However, optimization of these protocols was necessary for crystallization as the purity of the protein was not sufficient for crystallization.

AAC(6')-APH(2") was expressed in *B. subtilis* under the control of the constitutive *vegII* promoter, the plasmid construct pBF14 was obtained from Gerry Wright and Denis Daigle (Daigle et al., 1999). Two 25 ml cultures of LB media containing 25 µl of 50 mg/ml gentamicin sulfate were inoculated from a frozen stock of bacteria stored at –80 °C. These cultures were grown overnight and used as innoculum for two 1 L cultures of LB media containing 1 ml of 50 mg/ml gentamicin sulfate. The large cultures were allowed to grow for 14-16 hours to an $O.D._{600\ nm}$ of 1.6. Cells were harvested by centrifugation at 5,000 g for 10 minutes, washed with 0.85% NaCl and respun at 5,000 g for 10 minutes. The wet cell paste was resuspended in 50 mM HEPES pH 7.5, 1mM EDTA, 1mM PMSF, and 0.1mM DTT, and lysed by three consecutive

passes through 20,000 psi French Pressure cell. The cell debris was removed by centrifugation at 10,000 g for 20 minutes.

With the cell debris removed the bifunctional protein was passed over a Fast Flow Q-Sepharose column. This column was attached to the GRADI-FRAC (Amersham Pharmacia Biotech) column chromatography system and the gradient described in Table 2.2.2 was used to elute the proteins.

Table 2.2.2: Gradient information used for Q-Sepharose column

Breakpoint (ml)	% Buffer B	Flow Rate (ml/min)	Fraction Size (ml)
0.0	0	1.0	0
200.0	0	1.0	8.0
350.0	20	1.0	8.0
450.0	20	1.0	8.0
650.0	40	1.0	8.0
780.0	50	1.0	8.0
830.0	50	1.0	8.0
950.0	100	1.0	0
1020.0	100	1.0	0
1021.0	0	1.0	0
1200.0	0	1.0	0

-Gradient is based on a 100 ml column

The active protein from this column was concentrated using a PM30 membrane on a Nitrogen gas concentrator (Amicon) followed by a Ultra-Free centrifugal concentrator (Millipore) to ~2 ml. The concentrated protein was loaded onto a G100 molecular sizing column and eluted with Buffer A. The protocol established by Daigle et al. (1999) uses a linear gradient to elute the protein from a gentamicin agarose affinity column as the final purification step. Attempts to optimize the final step involved using a different column, namely neomycin agarose. Also, other conditions explored included using the gentamicin agarose column but eluting with $MgCl_2$ rather than NaCl (LeGoffic et al.,

1973), loading the protein onto the gentamicin agarose in the presence of ATP, and finally altering the NaCl gradient used on the gentamicin agarose column. The best results were obtained by alteration of the NaCl salt gradient on the gentamicin agarose column. The gentamicin agarose affinity column was prepared as per the steps described in Daigle et al. (1999). The column was run on the GRADI-FRAC chromatography system with a step gradient (see Table 2.2.3).

Table 2.2.3: Gradient Information for gentamicin agarose column.

Breakpoint (ml)	% Buffer B	Flow Rate (ml/min)	Fraction Size (ml)
0.0	0	0.6	0
150.0	0	0.6	7.5
275.0	20	0.6	7.5
375.0	20	0.6	7.5
550.0	50	0.6	7.5
650.0	50	0.6	7.5
750.0	100	0.6	7.5
800.0	100	0.6	0
850.0	100	0.6	0
851.0	0	0.6	0
1000.0	0	0.6	0

- Gradient is based on a 75 ml column volume

After the purified protein was obtained, the purity was analyzed using SDS-PAGE with silver staining. The purified protein was then concentrated using Ultra-Free centrifugal concentrators (from Millipore) and the concentration determined using the Bradford Assay (Bradford, 1977).

The APH(2") protein was purified using a protocol similar to that used for the full-length bifunctional protein. The first major difference in the protocols is bacteria used for expression. The APH(2") domain was overexpressed in *E. coli* using plasmid

pBF16 obtained from Denis Daigle rather than *B. subtilis* (private communication, Denis Daigle, McMaster University, Department of Biochemistry). In addition, the expression of this domain was under the control of an IPTG inducible promoter. 1 mM IPTG was therefore used to induce protein production when the $O.D._{600nm}$ of the culture reached 0.6. The length of the induction was 3.5 hours. The next major difference between the bifunctional and APH(2") purification protocols was the columns that were used in the purification. During the APH(2") purification the G100 molecular sizing column was not used. The active protein was dialyzed against Buffer A in preparation for loading onto the gentamicin agarose column.

2.2.3 Crystallization

The DLS instrument (from Protein Solutions) was used to analyze the protein in variety of storage buffers for the likelihood to remain monodisperse and therefore have a greater potential to crystallize. A dilute protein (0.5 mg/ml) sample was prepared with a final buffer concentration of 50 mM. The resultant sample was then analyzed and the molecular weight and polydispersity index estimated for each sample. The bifunctional protein was analyzed as the APO (unbound) form as well as in the presence of a five-fold molar excess of ADP and CoA, and ADP, CoA, and kanamycin A.

Based on DLS of the bifunctional protein, two buffers were chosen to use as the storage buffer for subsequent crystal experiments. The protein was exchanged into these

buffers using Ultra-Free centrifugal concentrators. The buffers used were 50 mM Tris pH 8.5 and 50 mM sodium cacodylate pH 7.0. The Crystal Screens I and II, the MPD, Ammonium Sulfate, and PEG 6000 Grid Screens, Wizard Screens I and II and the Cryo Screens I and II were all set up using TRIS pH 8.5 as the buffer for the apo condition at both 4 °C and 22 °C. Only crystal screens I and II and the three Grid Screens were set up on a protein substrate mixture stored in a buffer of Sodium Cacodylate pH 7.0. The substrates used for these trials included ADP, CoA, and kanamycin A. Conditions yielding promising crystalline material were subjected to fine screening efforts. As an additional variable for the Hampton Crystal Screens with the apo protein the protein concentration was varied. Additionally, crystal trials were set up with the bifunctional protein stored in 25 mM HEPES pH 7.5. Crystal Screens I and II were set up at both 22 °C and 4 °C. Conditions yielding crystalline material were fine screened, and subsequently Additive Screens were set up with a single condition.

DLS analysis of the apo form of the APH(2") domain indicated preferred buffers for maintaining the enzyme in a monodisperse state. The protein was exchanged in 50 mM TRIS pH 8.5 using Ultra-Free centrifugal concentrators. This buffer was used for subsequent crystal trials using the Sparse Matrix (Jancarik and Kim 1991) method adopted in the Hampton Crystal Screen I and II kits using the hanging drop vapour diffusion set up. Fine screens were carried by altering the pH and concentration of the various components in the condition of interest. In addition, APH(2") crystal trials were set up using the Hampton Crystal Screens I and II with the protein stored in 25 mM

HEPES pH 7.5. These trials were set up with a variety of substrates present including ADP, AMPPNP, H-9, and kanamycin A.

2.3 Results and Discussion

2.3.1 Purification Optimization

Optimization of the purification protocols was necessary in order to obtain protein that was essentially homogeneous. Homogeneity is far more important for protein crystallography than for other biochemical work such as kinetics (McPherson, 1999). The purification of the bifunctional AAC(6')-APH(2") was optimized for use with the GRADI-FRAC programmable chromatography system. This system allows for programming of gradients for use with columns. The established protocol used the FPLC for the Q-Sepharose column with a flow rate of 3 ml/min, however a reduction in the flow rate to 1 ml/min improved the resolution of the peaks. In addition, the application of a non-linear gradient for the gentamicin agarose affinity column resulted in a significant improvement in resolution between impurities and the pure protein. When a simple linear gradient was used, the fractions containing pure protein were part of a tailing hump on the peak containing not only the desired protein but many impurities as well. However, with the more advanced gradient clear separation can be seen between the impurity peak and the peak with pure protein, this can be seen in Figures 2.3.1 and 2.3.2, which show the profiles for the gentamicin agarose column for the bifunctional and the APH(2")

respectively. Through modifications of the gradient for the Q-Sepharose and the gentamicin agarose columns, the resolution of protein from the impurities was greatly improved as is seen in Figures 2.3.1 and 2.3.2.

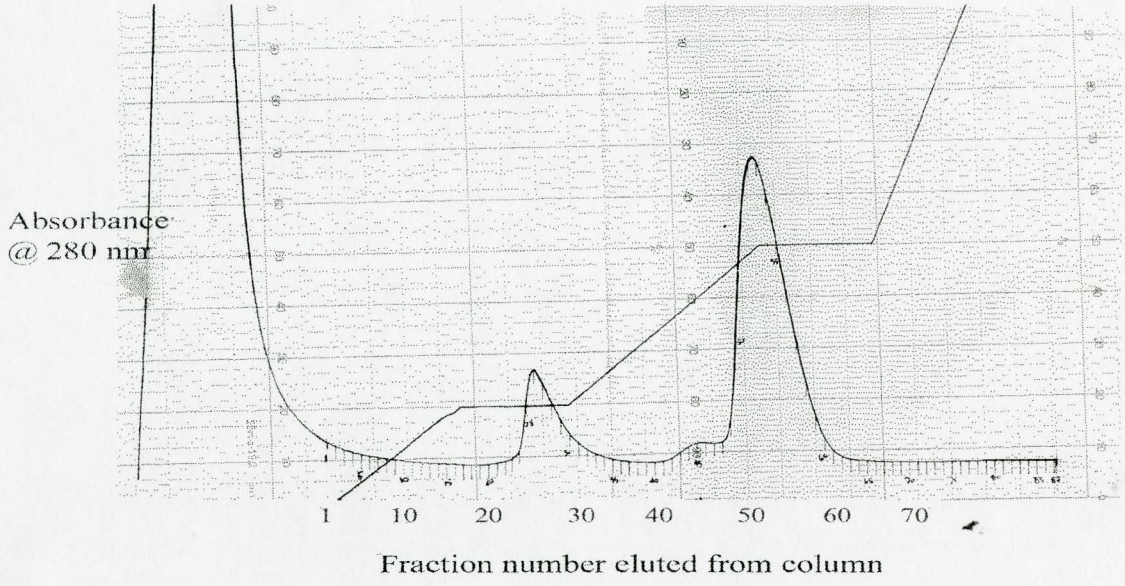


Figure 2.3.1: Profile of bifunctional AAC(6')-APH(2") run on gentamicin agarose affinity column. Pure bifunctional protein elutes at 45 % buffer B between fractions 48 and 62.

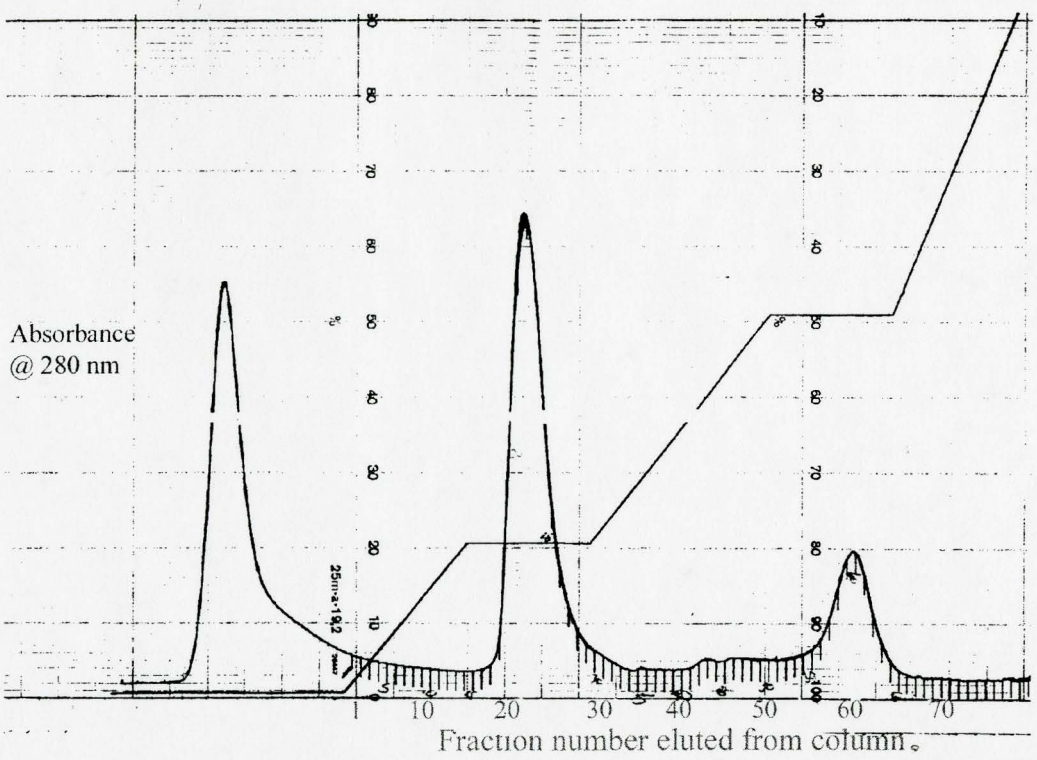


Figure 2.3.2: Profile of APH(2") domain from gentamicin agarose affinity column. Pure APH(2") elutes from the column at 50 % buffer B between fractions 55 and 65.

Figures 2.3.3 and 2.3.4 show the pure AAC(6')-APH(2") and APH(2") respectively, obtained when the optimized purification protocol was employed.

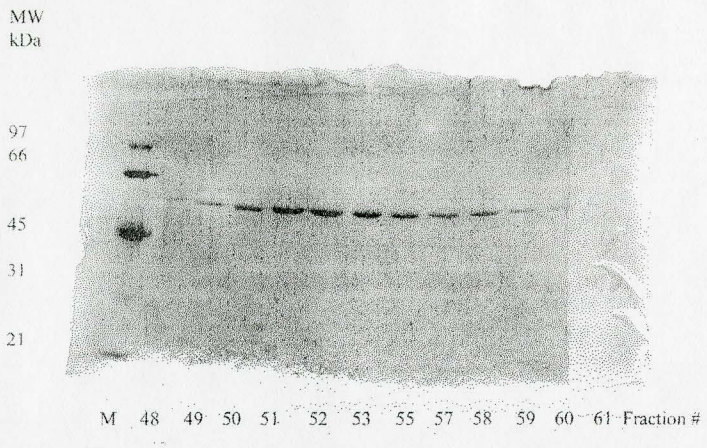


Figure 2.3.3: Silver Stained SDS-PAGE gel of AAC(6')-APH(2") gentamicin agarose fractions. The fraction #'s represent the fraction from the gentamicin agarose affinity and correspond to fraction #'s in Figure 2.3.1 column M represents molecular weight markers (from Bio-Rad).

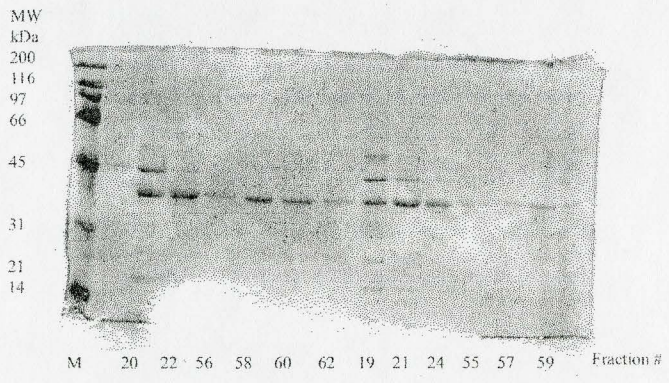


Figure 2.3.4: Silver Stained SDS-PAGE gel of APH(2") gentamicin agarose fractions. Fraction numbers from two consecutive gentamicin agarose column runs are indicated beneath the gel M represents the molecular weight markers (Bio-Rad) fraction #'s correspond to those shown in Figure 2.3.2.

The gel shown in Figure 2.3.3 shows the pure AAC(6')-APH(2") in lanes 48-60. The single band in these corresponds to a molecular weight of 57 kDa, as expected for

the full-length bifunctional protein. Figure 2.3.4 shows pure APH(2") domain in lanes 56-62 from the first gentamicin agarose run and in lanes 55-61 from the second gentamicin agarose run. The single band in these lanes corresponds to a molecular weight of 36 kDa. In addition, this gel shows the impurities present along with the APH(2") protein, a band at 36 kDa appears along with various other bands fractions 20-22 and 19-24. These fractions are representative of those found in the peak eluting at 20% NaCl seen in the profile from the gentamicin agarose column (Figure 2.3.2). The resolution between the impurities and the pure protein is clear from both the gel (Figure 2.3.4) and the profile (Figure 2.3.2). The yield of pure protein from two liters of culture was estimated to be ~30 mg for the APH(2") domain and ~13 mg for the bifunctional protein. The bifunctional protein yield is similar to that achieved using the original protocol of 7 mg/liter of culture (private communication, Denis Daigle, McMaster University, Department of Biochemistry), however the purity as indicated by SDS-PAGE (Figure 2.3.3) is greatly improved.

2.3.2 Dynamic Light Scattering

DLS was used in the cases of the APH(2") domain and the full length bifunctional AAC(6')-APH(2") in an effort to determine buffers that would maintain the protein in a monodisperse state. Buffers that provided readings for the protein that corresponded well with the actual molecular weight and also with a % Polydispersity of less than 30 were desirable (Ferré-D'Amaré and Burley 1997). In cases where no % polydispersities were

under 30 then the buffer that gave a value closest to 30 % was selected. In addition, the reproducibility of the individual measurements was observed, monodisperse solutions will show little variation in the translation diffusion coefficient and the determined molecular weight (Ferré-D'Amaré and Burley 1997).

Table 2.3.1: Dynamic Light Scattering data for bifunctional AAC(6')-APH(2") with substrates.

Buffer	**Estimated Molecular Weight (kDa)**		**Polydispersity Index** *		**% Polydispersity** **	
Substrates Added	CoA, ADP, Kan	CoA, ADP	CoA, ADP, Kan	CoA, ADP	CoA, ADP, Kan	CoA, ADP
MES pH 6.0	22.0	NR	0.51	NR	71	NR
MES pH 6.5	44.0	NR	0.22	NR	47	NR
Na Cacodylate pH 6.5	58.5	53.4	0.08	0.33	28	57
Na Cacodylate pH 7.0	55.3	56.2	0.07	0.04	26	20
HEPES pH 7.0	34.4	39.8	0.37	0.37	61	61
HEPES pH 7.5	42.2	59.9	0.31	0.11	56	33
MOPS pH 7.0	46.2	170.1	0.24	0.48	49	69
MOPS pH 7.5	53.5	128.5	0.20	0.32	45	57
Tris pH 8.0	55.7	81.7	0.09	0.12	30	35
Tris pH 8.5	65.5	82.4	0.12	0.15	35	39
TAPS pH 8.0	49.4	NR	0.37	NR	61	NR
TAPS pH 8.5	42.4	ND	0.22	ND	47	ND
Bicine pH 8.0	61.6	NR	0.20	NR	45	NR
Bicine pH 8.5	NR	ND	NR	ND	NR	ND

All substrates added in 5 fold molar excess compared to the protein. CoA = Coenzyme A, ADP = Adenosine diphosphate, Kan = Kanamycin A.
NR – No Readings – DLS gave no results when sample was tested
ND – Not Determined – Sample was not tested
* Polydispersity Index represents the standard deviation of a Gaussian distribution based on a monomodal fit.
** % Polydispersity is the square root of the Polydisperisty index.

Table 2.3.2: Dynamic Light Scattering for APO AAC(6')-APH(2") bifunctional protein.

Buffer	Estimated Molecular Weight (kDa)	Polydispersity Index	% Polydispersity
MES pH 6.0	NR	NR	NR
MES pH 6.5	125.9	0.19	44
Na Cacodylate pH 6.5	53.6	0.07	26
Na Cacodyalte pH 7.0	55.5	0.08	28
HEPES pH 7.0	32.7	0.47	69
HEPES pH 7.5	48.3	0.25	50
MOPS pH 7.0	51.6	0.10	32
MOPS pH 7.5	53.6	0.12	35
Tris pH 8.0	57.5	0.15	33
Tris pH 8.5	56.9	0.05	22
TAPS pH 8.0	52.0	0.31	56
TAPS pH 8.5	51.0	0.34	58
Bicine pH 8.0	53.0	0.12	35
Bicine pH 8.5	62.0	0.16	40

All results are averages for between 12 and 17 individual measurements.
NR – No Readings – DLS gave no results when sample was tested.

Table 2.3.3: Dynamic Light Scattering for APO APH(2") Domain.

Buffer	Estiamated Molecular Weight (kDa)	Polydispersity Index	% Polydispersity
MES pH 6.0	NR	NR	NR
MES pH 6.5	30.2	0.50	71
Na Cacodylate pH 6.5	53.1	0.14	37
Na Cacodylate pH 7.0	55.1	0.21	46
HEPES pH 7.0	21.0	0.45	67
HEPES pH 7.5	37.8	0.43	66
MOPS pH 7.0	43.1	0.54	73
MOPS pH 7.5	48.3	0.38	62
Tris pH 8.0	44.6	0.15	39
Tris pH 8.5	41.4	0.11	33
TAPS pH 8.0	45.3	0.47	69
TAPS pH 8.5	45.8	0.41	64
Bicine pH 8.0	48.0	0.36	60
Bicine pH 8.5	37.2	0.26	51

All results are averages for between 12 and 17 individual measurements. Protein concentration was 0.5 mg/ml and final concentration of each buffer was 50mM.
NR – No Readings – DLS gave no results when sample was tested

Dynamic Light Scattering experiments for both the full-length bifunctional protein and the APH(2") domain indicated that the preferred buffer for crystallization was 50 mM TRIS pH 8.5 (Tables 2.3.2-2.3.3). It was also determined that with substrates the bifunctional enzyme remained most monodisperse in a buffer of 50 mM Na cacodylate pH 7.0 (Table 2.3.1). However, upon SDS-PAGE analysis of protein stored in this buffer it was determined that proteolysis was occurring (Figure 2.3.5). Panel B of Figure 2.3.5 shows that after four days of storage in Na cacodylate the AAC(6')-APH(2") has undergone significant proteolysis, this proteolysis continues to progress at six and eleven days of storage (Panels C and D, Figure 2.3.5). When the AAC(6')-APH(2") protein was stored in the Tris or HEPES buffer no proteolysis was observed at four and six days (Panel B and C), and only limited proteolysis had occurred at eleven days (Panel D). This result led to the decision to ignore co-crystallization of AAC(6')-APH(2") for the time being and focus solely on crystallizing the apo form of the enzyme. The reason for the proteolysis of the bifunctional AAC(6')-APH(2") when stored in cacodylate is unknown. It is possible that the stock buffer that was prepared contained some trace proteases that could result in the degradation of the protein over time. The fact remains that cacodylate is the preferred buffer for the full bifunctional enzyme when combined with substrates. By performing DLS on a sample of protein that has been stored in cacodylate for some time, it may be possible to see if this buffer maintains the monodisperse behavior or if a more polydisperse signature is seen. If proteolysis is occurring, one would expect the signature of the protein to become more polydisperse. Another possible solution that could allow for the crystallization of the bifunctional

enzyme in the presence of substrates would be to try other buffers that were not used in these DLS trials. Also a variety of other substrates could be tested using DLS, or the substrates could be soaked in the crystals after apo crystals have been grown.

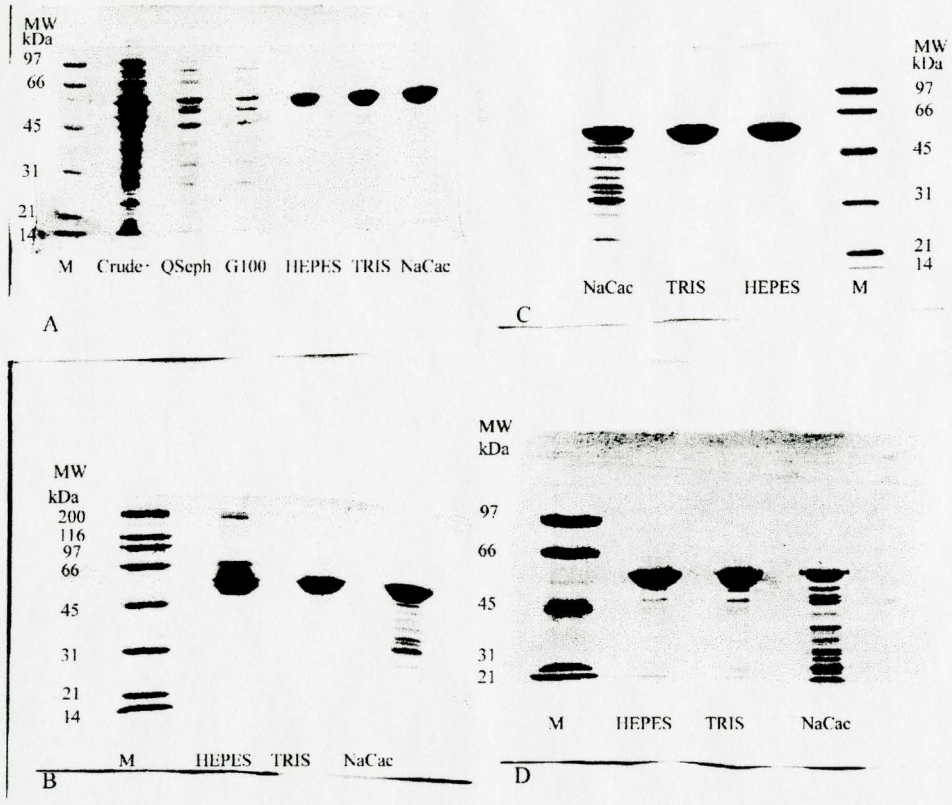


Figure 2.3.5: Coomassie Blue stained SDS-PAGE gels showing AAC(6')-APH(2"). In all gels M represents the molecular weight marker, HEPES represents protein stored in 25 mM HEPES pH 7.5, TRIS represent protein stored in 50 mM TRIS pH 8.5, and NaCac represents protein stored in Na Cacodylate pH 7.0. Panel A was run on the day the purification was completed and includes samples of the crude lysate, pooled fraction from the Q-Sepharose and G100. Panel B was from four days after completing the purification. Panel C was from six days following the purification. The gel in Panel D was run eleven days after the purification was completed.

2.3.3 Crystallization

Initial crystal trials were performed using the sparse matrix method (Jancarik and Kim, 1991). From the various conditions that were tested for both the AAC(6')-APH(2") enzyme and the APH(2") domain, a number of conditions were selected for more extensive screening. In addition to the sparse matrix screens, grid screens were used for the bifunctional enzyme. Conditions that yielded crystalline material and were analyzed further are presented in Table 2.3.4 – 2.3.7.

Table 2.3.4: Conditions from Initial Crystal Trials that produced Crystalline Material of AAC(6')-APH(2") stored in 25 mM HEPES pH 7.5.

Condition	Protein Concentration	Temperature
30% v/v PEG 400, 0.1M HEPES pH 7.5, 0.2 M $MgCl_2 \cdot 6H_2O$, Co-crystallized with 5x molar excess Kanamycin, AMPPNP, and CoA	15 mg/ml	4°C

Fine screens on the above Crystal Screen Condition led to improved quality crystals, additive screens were performed finding ethylene glycol to be an effective additive. The resulting crystal was mounted on the x-ray source and produced no diffraction.

Table 2.3.5: Conditions from Initial Crystal Trials that produced Crystalline Material of AAC(6')-APH(2") stored in 50 mM TRIS pH 8.5.

#	Condition	Protein Concentration	Temperature
1*	12 % w/v PEG 20,000, 0.1M MES pH 6.5	14 mg/ml	4°C
2	1.0 M 1,6 Hexandiol, 0.1 M Na Acetate pH 4.6, 0.01 M $CoCl \cdot 6H_2O$	14 mg/ml	4°C
3#	0.1 M Citric Acid pH 4.0, 65% v/v MPD	9 mg/ml	4°C
4	0.1 M NaHEPES pH 7.5, 1.4 M Na Citrate $\cdot 2H_2O$	14 mg/ml	22°C
5	2.0 M Ammonium Formate, 0.1 M HEPES pH 7.5	14 mg/ml	22°C
6	0.2 M Na Acetate $\cdot 3H_2O$, 0.1 M Na Cacodylate pH 6.5, 30% w/v PEG 8000	14 mg/ml	4°C
7	30 % v/v PEG 400, 0.1 M Na Acetate pH 4.6, 0.1 M CdCl	14 mg/ml	22°C
8	10 % w/v PEG 20000, 0.1 M Bicine pH 9.0, 2% v/v Dioxane	14 mg/ml	22°C
9	1.0 M Imidazole pH 6.5	14 mg/ml	22°C

	Condition	Protein Concentration	Temperature
10	0.2 M Na Acetate·$3H_2O$, 0.1 M TRIS pH 8.5, 30% w/v PEG 4000	14 mg/ml	4°C
11	35% v/v 2-propanol, 0.1 M Acetate pH 4.5	10 mg/ml	4°C
12	35% MPD, 0.1 M HEPES pH 7.5, 0.2 M NaCl	10 mg/ml	22°C

Conditions #2-10 were fine screened at protein concentrations of 8 mg/ml and 14 mg/ml. A large portion of the wells resulted in precipitation forming. An attempt was made to mount the crystal in condition #3, however the crystal dissolved upon opening the well. A number of the initial crystal were stained blue with IZIT™ dye indicating that they were protein crystals including condition #1, #4 and #5. These crystals were rod shaped with lengths of 0.41 mm, 0.48 mm and 0.3 mm respectively.

*-Photograph of crystal #1 is shown in Figure 2.3.6

#-Photograph of crystal #3 is shown in Figure 2.3.7

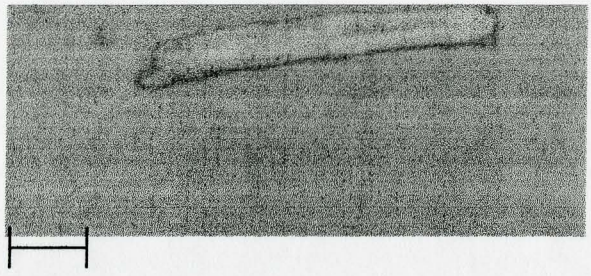


Figure 2.3.6: Crystalline material of AAC(6')-APH(2") protein produced from condition #1 in Table 2.3.5.

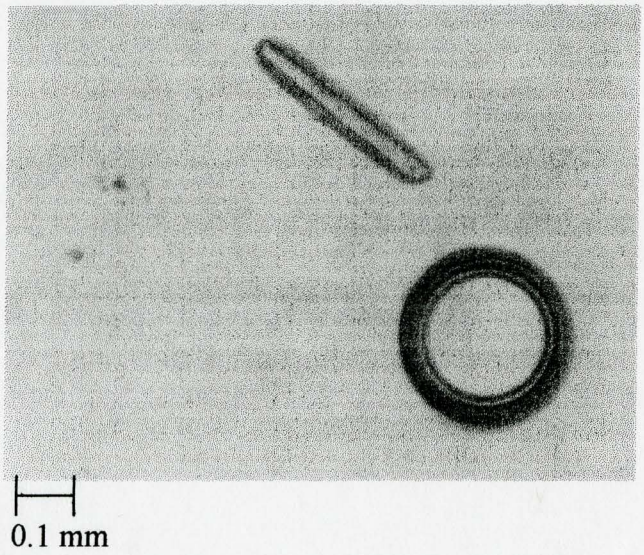


Figure 2.3.7: Crystal of bifunctional AAC(6')-APH(2"), from condition #3 listed in Table 2.3.5. Length of rod shaped crystal was 0.3 mm. Circle in bottom right corner is an air bubble.

Table 2.3.6: Conditions from Initial Crystal Trials that produced Crystalline Material of APH(2") Domain stored in 25 mM HEPES pH 7.5.

Condition	Protein Concentration	Temperature
25 % v/v tert-butanol, 0.1 M Tris pH 8.5, 0.1 M $CaCl_2 \cdot 2H_2O$	11.5 mg/ml	4°C

Fine screens attempted on this condition did not yield any results

Table 2.3.7: Conditions from Initial Crystal Trials that produced Crystalline Material of APH(2") Domain stored in 50 mM TRIS pH 8.5.

#	Condition	Protein Concentration	Temperature
1*	0.1 M Na HEPES pH 7.5, 0.8 M K, Na Tartrate	15 mg/ml	22°C
2	20% Jeffamine M-600, 0.1 M HEPES pH 7.5	15 mg/ml	22°C
3	0.1 M MES pH 6.0, 40 % v/v MPD	15 mg/ml	22°C
4	0.1 M HEPES pH 7.0, 65 % v/v MPD	15 mg/ml	22°C
5	4.0 M Na Formate	15 mg/ml	22°C

Fine Screens of condition #1 were performed yielding large rod shaped crystals up to 0.95 mm in length these were severely twinned however. Subsequent screening was performed using additives and oils to overlay the well solution. Condition #2 yielded crystals that were twinned and fine screening the condition produced no new crystals. Condition #3 produced crystals that were stained blue with IZIT™ dye, these crystals were very small and fine screening did not generate any improved crystals.

*-Photograph of crystal #1 is shown in Figure 2.3.8

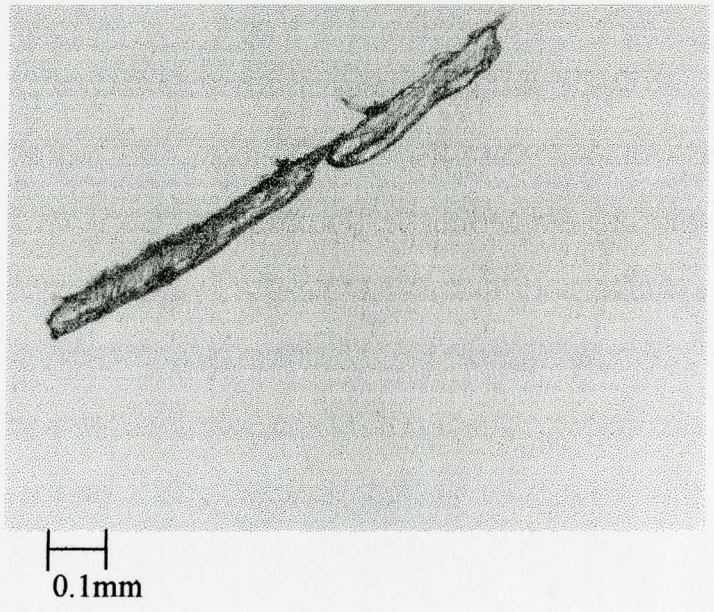


Figure 2.3.8: Crystal of APH(2") domain, fine screen of condition #1 listed in Table 2.3.7. Length of entire crystal totals 0.95 mm.

Very little success has been realized during attempts to crystallize either the AAC(6')-APH(2") protein or the APH(2") domain. Only a few of the crystalline samples obtained were tested for diffraction on the home or synchrotron x-ray sources. In every case little or no diffraction was observed. To this point, no crystals of sufficient quality for data collection have been found, and the reproducibility of the crystals that have been found has been poor. This may be due to the use of different batches of protein or because the solutions used for fine screen conditions may be prepared differently from those used for the crystal screen kits.

Chapter 3: Docking of Aminoglycosides Antibiotics to APH(3')-IIIa

3.1 Introduction

3.1.1 Aminoglycoside Phosphotransferase (3")-IIIa

The 3' aminoglycoside phosphotransferase (APH(3')-IIIa) is a plasmid encoded bacterial enzyme that is responsible for broad range aminoglycoside resistance in Enterococci and other gram-positive bacteria. Similar to the AAC(6')-APH(2") protein, the presence of the gene for the APH enzyme on a plasmid has lead to widespread dissemination of aminoglycoside resistance. APH(3')-IIIa possesses the broadest aminoglycoside substrate range of all APH enzymes (Shaw et al., 1993) but also shows remarkable regiospecificity. APH(3')-IIIa catalyzes the phosphate transfer from a molecule of ATP to the 3' hydroxyl group of both 4,5 and 4,6-disubstituted deoxystreptamine aminoglycoside antibiotics. In addition, this enzyme has the ability to transfer the phosphate moiety to the 5" hydroxyl of 4,5-disubstituted aminoglycoside (Figure 3.1.1) as was shown by the ability to modify lividomycin which lacks the 3' hydroxyl (Thompson et al., 1996a). The transfer of a single phosphate group to an aminoglycoside in effect renders the drug useless, as it presumably no longer has the ability to bind to the ribosome and therefore will not possess it's bactericidal activity.

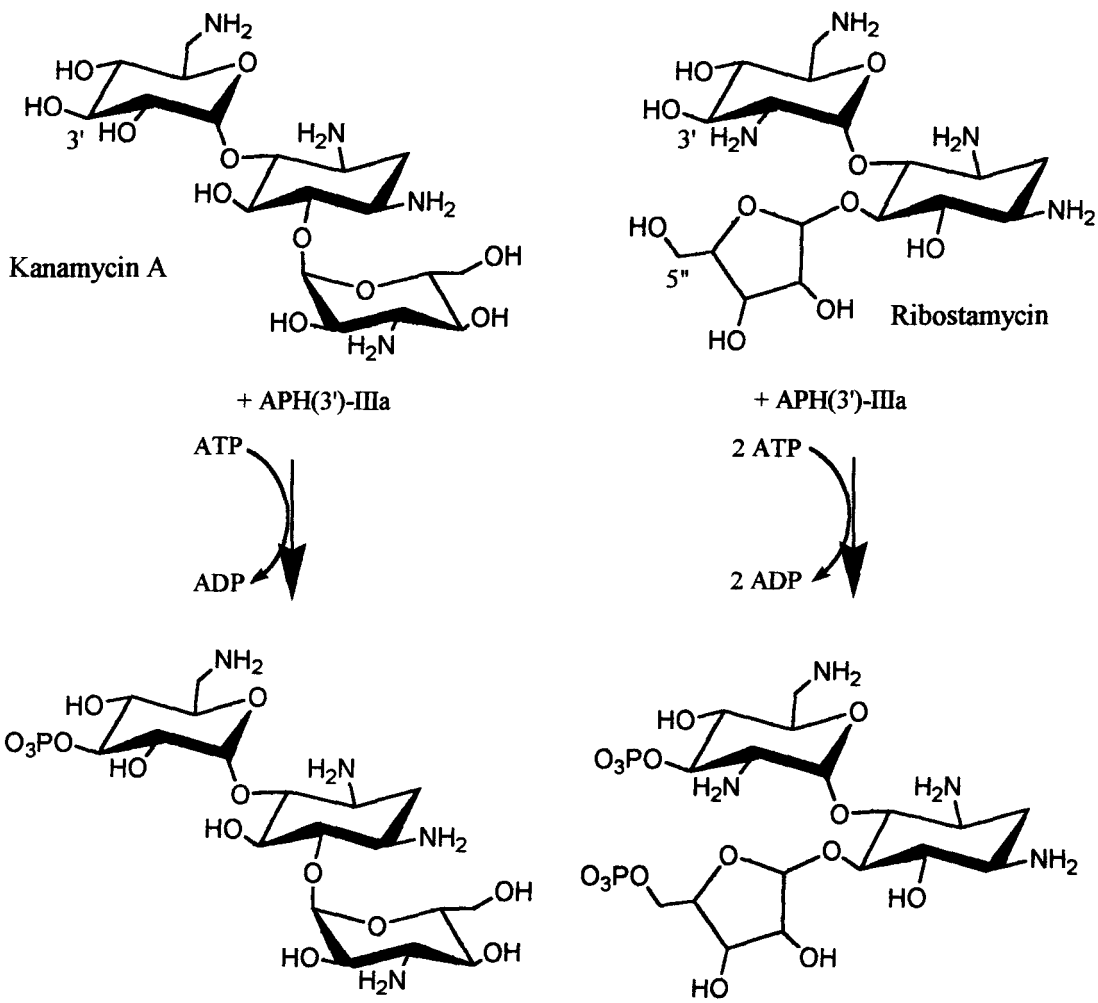


Figure 3.1.1: Aminoglycoside modification by APH(3')-IIIa. 3' site of kanamycin A is phosphorylated and 3' and 5" sites of ribostamycin are phosphorylated.

APH(3')-IIIa has been extensively studied by numerous biochemical and biophysical methods. The kinetic mechanism has been determined to be a Theorell-Chance mechanism with ATP binding first followed by the aminoglycoside, which is rapidly phosphorylated and released, and the final step being the rate limiting step of ADP release (McKay and Wright, 1995 and 1996). The chemical mechanism appears to be a direct phosphoryl transfer and not via a phopho-enzyme intermediate (Thompson et al., 1996b). The APH enzyme has been studied by mutagenesis to determine residues important not only in catalysis but also in binding of ATP and the aminoglycoside substrates (Hon et al., 1997, Thompson et al., 1999). NMR studies have been used in attempts to describe the conformation of the bound aminoglycosides (Cox et al., 1996, Cox and Serpersu, 1997). Finally, x-ray diffraction studies have revealed the three dimensional structure of APH(3')-IIIa and found that this enzyme possesses a fold similar to that of the eukaryotic protein kinases (Hon et al., 1997). The structure of APH(3')-IIIa has been determined bound to ADP (Hon et al., 1997) as well as in the unbound or apo form (private communication, Adelaine Leung, Department of Biochemistry, McMaster University). The work described here involves the use of computational methods in an attempt to describe the interactions of aminoglycosides and APH(3')-IIIa in the complex.

3.1.2 Computational Methods

Computational methods have proven fruitful in recent history for analysis of three-dimensional protein structures. Electrostatics are known to be a key factor in the

attraction of a protein to its substrate (Nakamura, 1996, Honig and Nicholls, 1995). These forces can be studied using a variety of computational tools. One use for electrostatics is to combine the information derived from calculating the electric field surrounding a protein with information about the protein's molecular surface. This can aid in understanding molecular recognition and also assist in assessing surface complementarity (Sharp et al., 1987).

A second computational tool that has recently seen a marked increase in usage is molecular docking. Docking has the ability to provide detailed information about the interactions between two molecules without the requirement for structure determination by either NMR or x-ray crystallography. Many docking algorithms have been developed and many have been tested in various competitions such as the CASP2 competition (Dixon, 1997). The number of different algorithms and scoring methods available for docking studies is too great to discuss all but the DockVision package used for this research will be discussed. In the CASP2 competition, DockVision was able to correctly predict three of the six structures attempted. Two of the remaining ligands were also correctly predicted but not chosen by the scoring functions (Hart et al. 1997). The Results of the CASP2 competition ranked DockVision among the top docking programs, based on its record for correct structure prediction (Dixon, 1997). The correct structures were selected in half the cases tested and in the remaining cases the correct answer was present in the list. The application of additional information to the problem could have assisted in selecting these structures.

DockVision employs a hybrid Monte Carlo algorithm, which employs a simulated annealing and is a score function driven procedure (Hart et al., 1992, Cummings et al., 1995, Hart et al. 1997). The simulated annealing schedules were created to produce a coarse sampling of conformation space during the first stage of docking and a finer sampling during the refinement stages. The Monte Carlo portion of the algorithm allows for the random selection of the next model remaining within the constraints of the annealing schedule. In this way there is less chance of biasing the docking into a local energy minimum rather than a global minimum. The DockVision algorithm involves the generation of multiple ligand conformations. Each individual trial in the docking run involves first the selection of a random ligand conformation, followed by a floating into the grid of the protein avoiding steric hindrance. The floating stage is followed by the Monte Carlo simulated annealing, where a new state of the ligand is randomly chosen. The initial state is then compared to the new state and if the new state is preferred it becomes the starting point for the next step of the annealing schedule. Each docking trial consists of multiple steps as outlined in the annealing schedule. This schedule describes the number of steps to be taken at a given temperature and the maximum rotation and translation allowed for each of those steps.

DockVision uses one of two force fields for scoring the results of a docking trial. The first is the standard RESEARCH force field (Hart et al. 1992) that uses hydrogen bonding and van der Waals interactions for calculating the empirical energy score. The

second is the Merck Molecular Force Field (MMFF) (Halgren, 1996), using hydrogen bonding, van der Waals and electrostatics in calculating the docking score.

3.2 Methods

(The methods for this chapter are also described in Thompson et al. (1999))

Using the previously determined 2.2 Å resolution crystal structure of APH(3')-IIIa with bound ADP a variety of computational studies have been performed to explore the interactions with the aminoglycoside antibiotics. However, it was more desirable to examine this enzyme with ATP bound rather than ADP as the interactions between the aminoglycoside and the enzyme would involve ATP as the co-substrate. Therefore, a model of APH(3')-IIIa bound to ATP was constructed. This involved first the selection of one of the crystallographically independent APH(3')-IIIa molecules and the removal of the ADP molecule. The next step was the modeling of the ATP molecule into the nucleotide binding pocket. This modeling was based on the conformation of the bound ADP molecule in the APH(3')-IIIa structure and the conformation of ATP molecules in the crystal structures of two protein kinase molecules (catalytic domain of cAMP depedent protein kinase, PDB code 1ATP (Knighton et al. 1991), and catalytic domain of phosphorylase kinase, PDB code 1PHK (Owen et al., 1995)). The use of these structures as models for the ATP molecule is justified by the structural homology that is shared between the protein kinases and APH(3')-IIIa (Hon et al., 1997). In addition, the

conformation of the two ATP molecules used as models were essentially identical, the γ-phosphates were within 0.3 Å of one another when overlaid.

The molecular surface of APH(3')-IIIa·ATP complex was calculated using the program GRASP, which stands for Graphical Representation and Analysis of Surface Properties (Nicholls and Honig, 1991). GRASP uses the algorithm described by Connolly (1983) for determining the protein surface. In addition, the surface properties of the enzyme were evaluated using GRASP, including the electrostatics and the electric field surrounding the protein (Sharp et al., 1987). A more detailed electrostatic analysis was performed using the Poisson-Boltzmann equation as implemented in DelPhi (Nicholls and Honig 1991, Honig and Nicholls, 1995, Biosym Technologies, San Diego, California). DelPhi calculations used dielectric constants of 80 for the solvent and 2 for the interior of the protein and partial charges for the protein and co-factor atoms were adopted from the AMBER force field (Weiner, 1984). The electrostatic analysis was used as a basis for the selection of the aminoglycoside binding site.

In order to analyze the APH(3')-IIIa·ATP·aminoglycoside ternary complex, molecular docking studies were undertaken. Four aminoglycosides were chosen for docking studies: kanamycin, amikacin, ribostamycin, and butirosin. These aminoglycosides were chosen because of the structural differences that these compounds possess. Kanamycin and amikacin are both 4,6-disubstituted 2-deoxystreptamine aminoglycosides, whereas ribostamycin and butirosin are 4,5-disubstituted. In addition,

amikacin and butirosin both possess a hydroxybutyrate substitution at the N1 position of the 2-deoxystreptamine ring. In the six models constructed, kanamycin and amikacin each have one productive binding mode (3' hydroxyl phosphorylation), and ribostamycin and butirosin each have two productive binding modes (3' and 5" hydroxyl phosphorylation) (Thompson et al., 1996a).

Molecular docking studies were performed using the program DockVision (Hart et al., 1997). This program was used to generate between 100-250 possible models from which a single model for each ternary complex was selected. The precise process involved three stages of docking. Stage one involved 3000 trials per ligand, each trial being a random ligand conformation which was flexibly (torsion angles were allowed to change) docked to the binding pocket of APH(3')-IIIa. These models were sorted and clustered resulting in between 200-500 unique aminoglycoside strucutres. These structures were then applied to the second stage of docking, refinement. During the refinement procedure each structure was submitted to a single trial and allowed to dock flexibly again. The result from the first stage of docking was the starting conformation for the refinement docking. The results from this refinement were clustered and sorted resulting in 100-300 distinct models. A final stage involved another round of refinement and clustering, leading to 100-250 unique ternary complex models. Simulated annealing schedules for both the initial and refinement docking runs are shown in Tables 3.2.1 and 3.2.2. The initial stage used high temperature and large maximum allowed rotations,

whereas the refinement stages used lower temperatures and smaller maximum allowed rotations.

Table 3.2.1: Simulated Annealing Schedule used for Initial Stage of Aminoglycoside Docking

Temperature	Number of Steps	Maximum Rotation	Maximum Translation
800	100	10.0	2.0
350	100	6.0	1.0
1.0	150	5.0	1.0
1.0	450	2.5	0.5

Table 3.2.2: Simulated Annealing Schedule used for Refinement Stages of Aminoglycoside Docking

Temperature	Number of Steps	Maximum Rotation	Maximum Translation
1.0	1500	2.0	0.8
1.0	3500	1.5	0.4

The scoring function used in the docking minimization procedure that generated the ternary complex models was a modified Merck Molecular Force Field (MMFF) (Halgren, 1996). The force field was optimized for use with DockVision by Richard Gillilan (private communication, Richard Gillilan, Cornell Theory Centre, Cornell University), and incorporates electrostatic and van der Waals terms in the score function. The partial charges on the protein and ATP atoms were based on AMBER parameters, and the aminoglycoside partial charges for the various amino groups were based on pKa values determined by NMR (Botto and Coxon, 1983, Cox and Serpersu, 1997, DiGiammarino et al., 1997). An additional term of a restraint penalty was employed in the docking score function. This penalty put a distance constraint between two atoms, specifically the γ-phosphate of ATP and the oxygen of the hydroxyl group (3' or 5")

being phosphorylated. The restraint penalty was large for the first stage of the docking procedure, was cut in half for the first refinement procedure, and entirely removed for the second refinement. The purpose of the restraint penalty was to remove models that were inconsistent with phosphoryl transfer and therefore allowed for more extensive sampling of the models consistent with productive aminoglycoside binding.

After the docking procedure was completed, the 100-250 possible models were assessed based on a number of criteria. Two types of tests were used for evaluation of models: a) properties indicative of favorable protein-ligand interaction and b) properties consistent with functional studies. The first set of tests involved properties such as the reduction in non-polar surface area upon complex formation, the number of hydrogen bonds, and the strength of these hydrogen bonds assessed by geometric considerations. These properties were assessed by a number of programs including Outrank (Hart et al. 1997) which was designed specifically for use in conjunction with DockVision. This program generates a list of all the ligands with numerical values relating to the surface burial and hydrogen bond formation when a protein-ligand complex is formed. The second group of tests examined the distance between the γ-phosphate of ATP and the phosphate accepting hydroxyl group, and the proximity of the ligand to residues known to interact with the substrate based on mutagenesis studies on the following residues: Tyr55, Arg211, Asp261, Glu262, and Phe 264 (Thompson et al., 1999). Scoring of the models was performed for each criteria, with weighting towards the functional data. This permitted the selection of 20 models to be visually inspected using the DockVision

associated viewer DockCam. Graphical analysis was used for the final selection of a single model for each of the six ternary APH(3')-IIIa·ATP·aminoglycoside complexes.

3.3 Results and Discussion

3.3.1 Electrostatic and Surface Analysis

The electrostatic analysis of the molecular surface of APH(3')-IIIa indicates a large negatively charged region. This region forms a shallow depression and is the putative aminoglycoside binding site. The depression can clearly be seen in Figure 3.3.1 and is indicated by the dark red region of the protein surface. This region was chosen as the aminoglycoside binding pocket for a couple of reasons. First the close proximity of this region to the modeled γ-phosphate of ATP. The chemistry of the reaction catalyzed by APH(3')-IIIa imposes an in-line attack of the phosphate by the aminoglycoside (Thompson et al., 1996b). Thus, the aminoglycoside must bind in close proximity to the γ-phosphate. Secondly, the presence of the large amounts of negative charge would provide a strong electrostatic force on the cationic aminoglycoside antibiotics. The binding region is lined by a number of acidic amino acid residues (Asp 104, Asp153, Asp155, Glu157, Asp190, Glu230, Asp231, Asp261, and Glu262) in addition to the C-terminus of the polypeptide chain (Phe264), producing a number of negative charges compared to very few positive charges in the area (Arg139, Arg211, Arg219). In addition, the C-terminus of the protein forms an α–helix (Hon et al., 1997), creating the possibility of a dipole effect at the c-terminus generating additional negatively charged

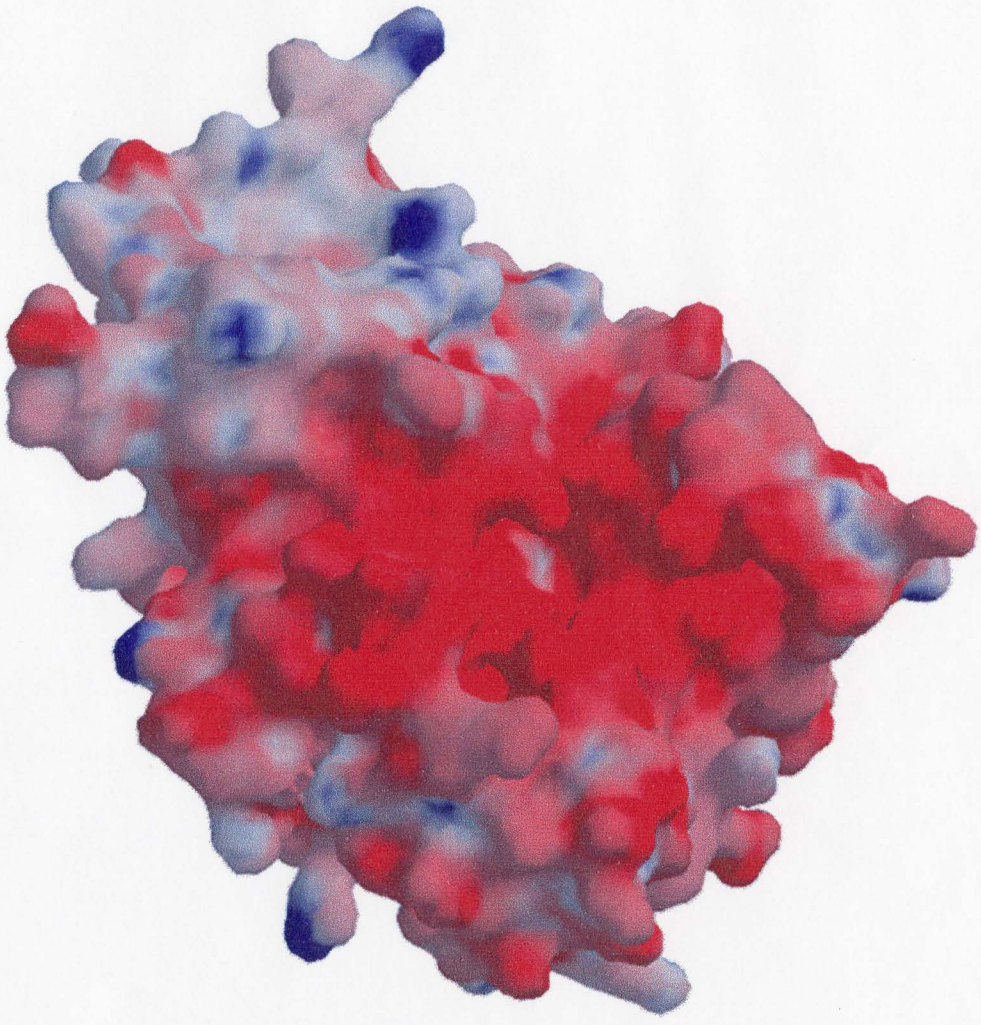


Figure 3.3.1: Surface of APH(3')-IIIa coloured by electrostatic potential. Red regions indicate negative charge, and blue indicates positive charge. The dark red depression in the surface is the putative aminoglycoside binding site.

This figure was prepared using GRASP (Honig and Nicholls, 1995), MOLSCRIPT (Kraulis, 1991), and Raster3D (Merritt and Bacon, 1997).

character here. The importance of electrostatics for this enzyme has been shown by analyzing aminoglycosides with reduced positive charge that were synthetically prepared (McKay et al., 1996, Roestamadji et al., 1995). The discovery of this putative binding site for aminoglycosides based on surface and electrostatic analysis led to experiments that attempted to explore how the aminoglycosides would actually bind to APH(3')-IIIa. These experiments were two fold, first mutagenesis was performed on the C-terminus of APH(3')-IIIa and second molecular docking studies were performed in an effort to provide three dimensional models for the interactions that were occurring at this site (Thompson et al., 1999).

3.3.2 Aminoglycoside Docking

The six models for the APH(3')-IIIa·ATP·aminoglycoside ternary complexes showed that each aminoglycoside has a distinct binding mode (Figure 3.3.2). This fact is also confirmed by the biochemical data produced through mutagenesis studies (Thompson et al., 1999). A single point mutation in the case of one aminoglycoside would have a large impact on the binding or catalytic ability of the enzyme, but for a closely related aminoglycoside no such effect would be seen. The effect of the Glutamate 262 to Alanine mutation on kanamycin is very large in terms of K_m but no effect is seen on amikacin when this mutation is made. The selected docking models, are shown in Figure 3.3.2, in the active site of APH(3')-IIIa with potential hydrogen bonding interactions indicated to the five amino acid residues mutated by Thompson et al. (1999).

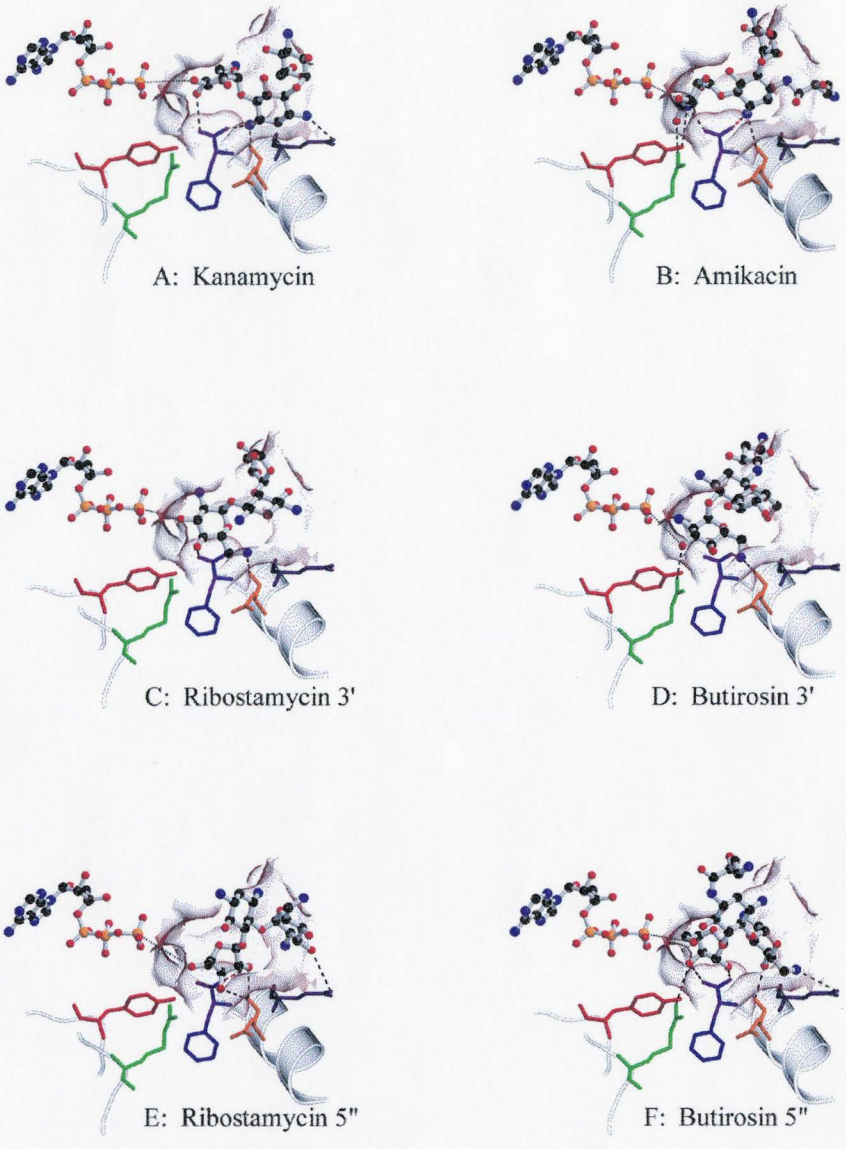


Figure 3.3.2: Docked models of aminoglycoside antibiotics. Active site residues are indicated Tyr55(red), Arg211(green), Asp 261(yellow), Glu262 (purple), Phe264(blue). The surface of the active site is shown transparently.

This figure prepared using MOLSCRIPT (Kraulis) and Raster3D (Merritt and Bacon, 1997)

All six models created from the docking studies were biased by the use of the biochemical mutagenesis data. The mutagenesis data allowed for distance constraints between a particular amino acid residue and any hydrogen bonding functional group on the aminoglycoside to be established. Model selection was facilitated by the use of these distance constraints.

The model created for kanamycin binding (Panel A Figure 3.3.2) shows potential interactions to Asp261, Glu262 and the carboxylate of Phe264. The deletion of Phe264, and therefore the carboxylate produced a dramatic effect on the catalysis of kanamycin by APH(3')-IIIa, whereas the substitution of Phe264 to Ala had no effect. In addition, the substitution of Glu262 to Ala saw a dramatic increase in K_m for kanamycin, showing importance for this residue in binding the aminoglycoside. The effect of mutating Asp261 to Ala was most significant on the catalysis of kanamycin by the enzyme. The model seems to show the interaction of the backbone of this residue with the aminoglycoside. Therefore the actual role of the Asp261 may not be a direct interaction but rather to maintain the stability of the C-terminal helix (Thompson et al., 1999). The model created provides us with the ability to visualize what the biochemical data may in fact represent. This facilitates conclusions to be drawn concerning the specific role of the various amino acids.

The amikacin model from the molecular docking experiments shows interactions with Tyr55, Arg211, Asp261 and the carboxylate of Phe264 (Panel B Figure 3.3.2). The

deletion of Phe264 has a dramatic effect on amikacin binding and catalysis, whereas the substitution of this residue to Ala has an effect only on binding. Therefore a potential role for the phenyl ring of Phe264 in amikacin binding exists. This could be accounted for by the large effect that the Tyr55 to Ala mutation has on amikacin. The phenyl ring of Tyr55 is in close proximity to that of Phe264 and the potential for a hydrophobic interaction occurring to correctly position the Tyr55 exists. The mutation of Arg211 to Ala showed a small effect on the K_m of amikacin. Asp261, as in the case of kanamycin, effects the catalytic ability of the enzyme on amikacin, and the interaction is shown to the backbone. Again helix stability is the likely role for this residue (Thompson et al., 1999).

The models presented for the aminoglycoside ribostamycin (Panels C and E Figure 3.3.2) show interactions with the carboxylate of Phe264, Asp261, and Glu262. The interaction with Glu262 is exclusively found in the 5" phosphorylation model (Panel E Figure 3.3.2). Mutagenesis determined that in the monophosphorylated product of the Glu262Ala mutant only the 3' hydroxyl group was phosphorylated (Thompson et al., 1999). Both models show interactions between ribostamycin and the carboxylate of Phe264 and Asp261. Phe264 does not have the same level of effect on ribostamycin as on the 4,6-disubstituted aminoglycosides, but Asp261 has a larger effect.

The models for butirosin (Panels D and F Figure 3.3.2) show interactions with Arg211, Asp261, Phe264 and Glu262. Like ribostamycin, the interaction with Glu262 only exists in the 5" productive mode. The interactions with the other three residues are

however present in both models. The two binding modes of the 4,5-disubstituted aminoglycosides such as ribostamycin and butirosin make it difficult to account for all the observations made from the mutagenesis data because a specific mutant may effect one binding mode but not the other. This creates difficulties in model selection based on this data as the role of a particular residue in the two binding modes may not be equivalent.

In an attempt to validate the docked models of the ternary APH(3')-IIIa·ATP·aminoglycoside complexes, a comparison to the NMR structures of butirosin and amikacin bound to APH(3')-IIIa·Cr^{3+}ATP was performed (Cox et al. 1996). A least-squares minimization was performed to compare the modeled conformation to the experimental model and the rms difference for both the 3' and 5" butirosin models was 2.3Å, the rms difference for the amikacin model was 3.4Å (Figure 3.3.3). The largest deviations for the amikacin model were in the hydroxybutyrate chain. The similarities between the model structures and the experimental structures support the conformations predicted from the docking. In addition the position of the aminoglycoside models relative to the phosphates of ATP was compared to the position determined in the NMR studies with Cr^{3+} ATP. A least-squares refinement was used to overlay the three phosphate groups of the ATP models. In both the amikacin and butirosin models all three of the phosphates aligned nearly perfectly, however the aminoglycoside models were positioned in entirely different orientations relative to the γ-phosphate of ATP (Figure 3.3.4). The NMR model has the aminoglycoside positioned in

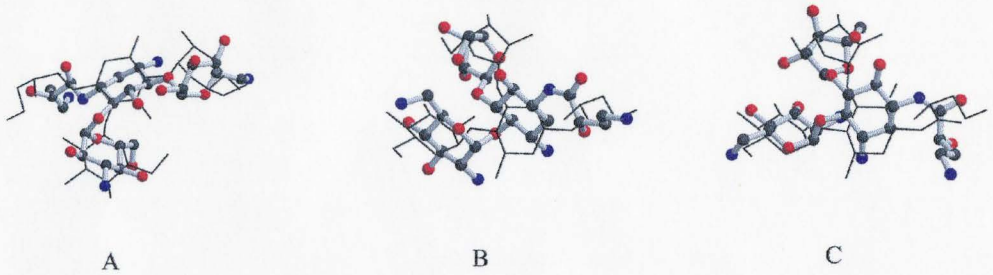


Figure 3.3.3: Conformational overlay of docked model (ball-and-stick) with NMR model (wireframe). Panel A shows the conformations of amikacin, panel B the model of the 3' productive butirosin and panel C the 5" productive model of butirosin.

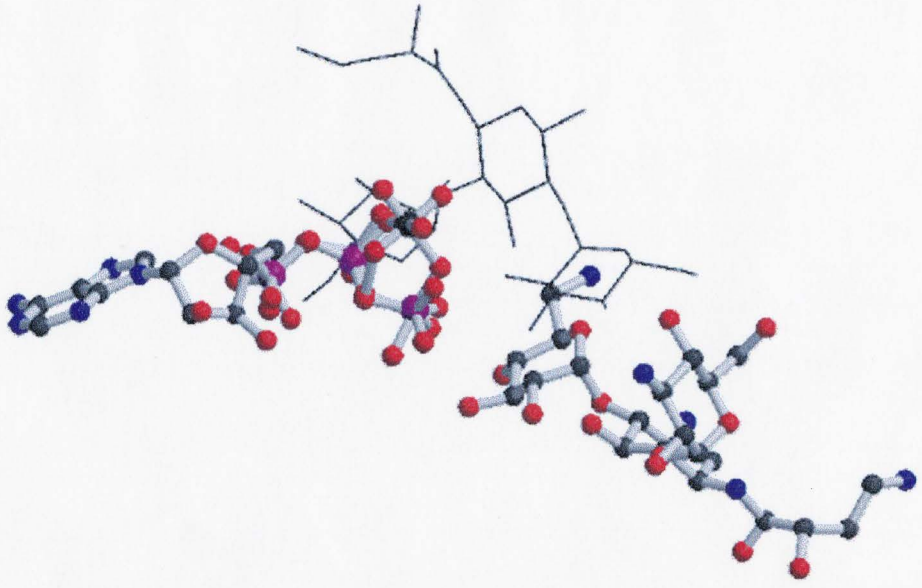


Figure 3.3.4: Positional overlay of docked model of amikacin (ball-and-stick) with NMR model (wireframe). Overlay was made between the phosphate of the ATP model used in docking studies and the three phosphate from the Chromium ATP NMR model.

These figures prepared using MOLSCRIPT (Kraulis, 1991) and Raster3D (Merritt and Bacon, 1997)

such a way that the antibiotic is embedded within the protein itself. A binding mode with part of the aminoglycoside inside the protein surface would not be possible as steric hindrance would prevent this from occurring. One problem that may have influenced this result is whether the Cr^{3+} ATP complex mimics the two Mg^{2+} or one Mg^{2+} ATP state. The APH(3')-IIIa structure indicates that this would bind the two Mg^{2+} ATP coordination state. It is not clear which coordination state of Mg^{2+}ATP Cr^{3+}ATP actually mimics. However, studies using Cr^{3+} ATP with pyruvate kinase have indicated a two metal coordination state (Gupta et al., 1976). A separate analysis involved the overlaying of the aminoglycoside and comparing the position of the phosphates of ATP between the docking ternary complex and the NMR model. When this was done the phosphates were not at all close to one another. It is obvious from these analyses that the NMR models generated using Cr^{3+} ATP (Cox et al., 1996, Cox et al., 1997) cannot be brought into agreement with the studies presented here. These models would place a portion of the aminoglycoside in direct contact with protein atom. In addition the models are not likely to fit the biochemical data generated from mutagenesis experiments (Thompson et. al, 1999). The models presented from the docking experiment were created by combining computational minimization with biochemical data. Without the crystal structure of the ternary complex, it is impossible to know which, if either, model is correct, however the inclusion of biochemical data in model determination should serve to increase confidence in the presented model.

The quality of the structure predictions from docking studies is highly dependent on the ability of the scoring functions to accurately predict the binding between the two molecules. Without prior knowledge of the binding site or of the chemistry involved in the interactions between a target and ligand, molecular docking could end up being a futile effort. The score functions for DockVision alone have been shown to be insufficient for the selection of the aminoglycoside models. The models selected using all the analysis tools ranked very poorly by docking energy alone. Amikacin ranked 37th, kanamycin 125th, 3' ribostamycin 36th, 5" ribostamycin 171st, 3' butirosin 49th, and 5" butirosin 86th. Many of the docked models that were generated produced scores that were similar or better than the finally selected model but were obviously incorrect. This was judged by the position of the hydroxyl group that was to be phosphorylated. The APH enzyme is highly regiospecific and therefore one would expect to see an energetic benefit to this binding mode. This benefit has not been observed with the docking studies undertaken and the reasons behind the selection of the 3' or 5" hydroxyl group for phosphoryl transfer are unclear with the presence of a number of chemically equivalent hydroxyl moieties on the aminoglycosides.

Docking studies were performed on kanamycin with two separate force fields each one with and without restraints. The distance between the γ-phosphate of ATP and each hydroxyl group on the various models was measured to determine if the force fields alone could display regiospecificity, or if the restraints were required to orient the aminoglycosides in a productive conformation. The results are shown in Table 3.3.1. In

addition to examining the ability of the score functions to deal with the regiospecifity of the APH enzyme, these studies allow for a comparison of the two force fields and to see the effect of restraints on the docking procedure.

Table 3.3.1: Docking studies of Kanamycin with RESEARCH and MMFF force fields with and without restraints.

Hydroxyl group		RESEARCH	RESEARCH + Restraints§	MMFF	MMFF + Restraints§
Total	# of Models	508	77	867	125
	Score	9.8	46.5	-204.9	-216.0
2' OH*	% of Models	2.4% (12)	64% (49)	3.3% (29)	62% (78)
	Score	15.3	42.1	-249.5	-193.6
3' OH*	% of Models	2.8% (14)	97% (75)	3.0% (26)	100% (125)
	Score	7.9	45.7	-259.6	-216.0
4' OH*	% of Models	1.8% (9)	61% (47)	3.3% (29)	48% (60)
	Score	13	48.5	-279.2	-201.1
2" OH*	% of Models	2.6% (13)	0% (0)	2.5% (22)	0% (0)
	Score	18.2	-----	-221.5	-----
4" OH*	% of Models	2.4% (12)	0% (0)	3.2% (28)	0% (0)
	Score	14.8	-----	-181.1	-----
5 OH*	% of Models	0.4% (2)	3% (2)	0.7% (6)	4% (5)
	Score	-10.2	26.5	-249.1	-207.5

*The cutoff distance between the γ-phosphate of ATP and the hydroxyl group for selection was 6Å.
§The restraints used in these docking runs were between the γ-phosphate of ATP and the 3' hydroxyl group of kanamycin.
The Score represents the average docking energy for all the models that have the desired hydroxyl group within the distance constraint.
The # of Models that met each distance cutoff criteria are shown in ( ).

The results presented in Table 3.3.1 show the importance of the restraints in the docking experiments. The restraints have a two-fold effect on the docking data. First, the restraints greatly reduce the total number of structures that are generated as potential models this greatly increases the speed at which analysis can be performed. The second effect is the increase in the percentage of the total structures that have the correct

orientation. Without the restraints in place, the force fields do not show a preference for any of the hydroxyl groups to be positioned for phosphorylation, but once the restraint is used then not surprisingly the 3' hydroxyl group is consistently placed in position for phosphoryl transfer. This effectively increases the chance of finding a correct structure as the number of models with the proper orientation is significantly increased. The use of the MMFF rather the RESEARCH force field has the effect of lowering the average docking energy. The MMFF does not have an effect on the ligand orientation but effectively increases the number of outputted structures because of the lower energy term. When the MMFF and restraints were used in combination, the average energy of the correctly oriented models was lower than the average energy that had any other hydroxyl oriented close to the γ-phosphate of ATP. This indicates that the MMFF with restraints is the most effective force field for separating the correctly oriented molecules from incorrectly oriented models. The MMFF was used in an effort to allow electrostatics to play a larger role in the docking energy score function. One factor that was not examined that could have improved the ability of the MMFF to predict accurately oriented structures alone is the dielectric constant used in docking. The effect of changing this constant has not been examined in the case of the aminoglycosides and APH(3')-IIIa.

The poor quality of the docking score functions resulted in the use of other functions to score the models generated in the docking trials. These functions mostly in the program Outrank (Hart et al. 1997) provided information on surface burial and

hydrogen bonding. Selection of aminoglycoside models that ranked at one extreme of various Outrank functions led to models that were in agreement with biochemical data. Smaller values were chosen for terms relating to surface burial, whereas large values for terms relating to hydrogen bonding were chosen. Despite the usefulness of the Outrank scoring function, without first selecting models that were oriented correctly, Outrank like the docking scoring function would have proved nearly useless.

Molecular docking studies were employed in order to model the ternary complex of APH(3')-IIIa·ATP·aminoglycoside. The docking method alone was not able to correctly model the aminoglycoside binding. However, through the unique combination of computational and biochemical data that was presented here models for this system have been generated. These models can only be validated through the resolution of the three-dimensional structure of the complex. The most useful biochemical information involves the reaction mechanism of the enzyme and the effect of mutating various amino acids in the enzymes binding pocket. The hybrid of docking with biochemical data for model determination should be applicable to other systems in order to study interactions between a macromolecule and its ligand. In the case of APH(3')-IIIa multiple binding modes were determined for the enzyme to its various substrates. If this enzyme in fact binds each of the aminoglycosides differently, the prospects for structure based drug design become more difficult. Without a single consistent binding mode, no general rules can be applied to the design of an inhibitor.

Chapter 4: Database Docking to APH(3')-IIIa

4.1 Introduction

4.1.1 Database Docking

Molecular docking not only has the ability to model the binding conformation of know substrates for an enzyme (as described in Chapter 3) but also can be useful in the search for enzyme inhibitors. Searching a large list of compounds for inhibitors using docking is referred to as database docking which is an example of database mining. Database docking can screen a large number of compounds and provide a score for each of these compounds. This process may in fact be useful in limiting the search for inhibitors to only a few hundred compounds rather than thousands.

Database docking has recently become a popular tool for researchers involved in structure based drug design (Marrone et al., 1997, Burkhard et al., 1998, Shoichet et al., 1993, Gschwend et al., 1997, Schnecke et al., 1998). A number of databases exist that have been used for the purpose of database docking and a number of different algorithms with varied approaches to the problem also exist. In general, the methods of database docking have met with success in extracting good scoring compounds from a database that were already known inhibitors of the macromolecule under investigation (Shoichet et

al., 1993). Complex structures predicted by database docking have also been confirmed through structure determination of the enzyme inhibitor complex (Burkhard et al., 1998). The method of database docking has proven itself useful for the design of inhibitors by providing an initial structure of a complex between the target and the inhibitor. This combined with experimental evidence to prove that the compounds extracted from the database actually bind to the target could provide a lead much quicker than random screening of compounds.

To perform docking experiments, with a database of small molecules, knowledge of the three dimensional structure of the target is required, also some type of structural information about the ligands is needed. This can be as simple as the SMILES string or the structural formula as a three-dimensional coordinate set can be generated from this information. In the case of the National Cancer Institute (NCI) database the conversion has been performed and is readily available for download from the Internet.

Normal docking algorithms perform a comprehensive search on a single ligand and use an advanced score function in an attempt to predict the interactions between the protein and the ligand. Normal docking provides multiple answers for the single ligand and the score functions can be used to compare these models and determine the most probable model. Database docking however, performs a quick search of multiple ligands, only simple calculations are performed and therefore only simple interactions are included in the score function. Database docking provides a single answer for each

compound in the database therefore, comparing the models is very difficult because each individual compound has a distinct structure. The goal of normal docking is to model the interactions between the macromolecule in question and a known ligand or inhibitor. Whereas, the goal of database docking is to quickly search for compounds which may act as an inhibitor of the macromolecule. Database docking is designed to sample a large number of very different compounds.

4.1.2 NCI Database

The NCI database is available free for download on the Internet and contains ~127,000 compounds. Each of the compounds is available in standard pdb format and is listed with a standard identifier. The compounds are also available for laboratory testing on 96 well plates as well as individually. This fact makes this database an excellent tool for drug discovery, as any hits from the docking studies can be obtained for experimental testing.

The NCI database when compared to other commercially available databases offers many advantages. First is the relative size of the database, the NCI database contains ~127,000 unique compounds versus ~53,000 for the Fine chemical directory. In addition, a conversion program is included with DockVision to convert the NCI database to the required pdb format.

One drawback of the NCI database and its use in docking studies is the number of compounds that are chemically unstable within the database. Some of these are intermediate compounds from synthesis of other compounds. In addition, some of the molecules in the NCI database are of unusual structures and do not represent good potential pharmaceuticals based on their large size and odd structures. Many of the compounds are highly uncommon or not well known as seen by the inability to locate them in other databases based on their standard identifier.

4.1.3 Database Docking using DockVision

The DockVision software package includes an algorithm for performing database docking, the RSDB (Research database) algorithm. RSDB like the regular docking discussed in Chapter 3 uses the Monte Carlo method for searching and optimizing (private communication, Trevor Hart, Department of Medical Microbiology and Infectious Diseases, University of Alberta). The database method of DockVision uses two annealing schedules rather than just one used in normal docking. These schedules perform first a search for general position of the compound and second a refinement of the determined conformation. RSDB uses a fast score function that is only based on hydrogen bonding and near grid contacts. The near grid provides an accurate picture of where protein atoms are located in space and where they are not, the score function prevents the ligand from locating it's atoms within the same space as a protein atom. The

score function itself is simple in order to increase the speed of the database docking procedure.

4.2 Materials and Methods

Database docking was performed with the APH(3')-IIIa-ADP complex structure determined by Hon et al. (1997) as the target molecule. The nearly 127,000 compounds were docked in three separate experiments each with a different constraint file. The first docking used a constraint file that only allowed molecules to dock in the aminoglycoside binding pocket (defined in Chapter 3). The second target had the ADP molecule removed and the compounds were docked to a combination of both the ATP binding and aminoglycoside binding pockets. Finally, with the ADP removed from APH(3')-IIIa compounds were permitted to dock within the ATP binding pocket alone.

The docking to the aminoglycoside pocket alone used 300 trials per ligand, the other two docking runs each only used 100 trials per ligand. This decision was made due to CPU availability and length of time required for the first round of database docking. Each trial involved a randomization of the starting configuration of the ligand. Monte Carlo simulated annealing was used and the schedules used in all three docking runs were the same, see Tables 4.2.1 and 4.2.2. The first annealing schedule docked the compound with random conformations equal to the number of trials in order to find the best general position. The second annealing schedule chose the preferred conformation from the first

docking and refined the docking of this molecule to the target. A single solution for each compound was output along with the docking scoring function based on the hydrogen bonding ability of the ligand to the target. Upon completion of docking the entire database in each of the three cases, analysis of the docked structures was performed. This analysis involved using Outrank to calculate scoring functions based on surface burial and more advanced hydrogen bonding calculations.

Table 4.2.1: Annealing Schedule used in first stage of Database Docking

Temperature	Steps	Max Rotation	Max Translation
150	100	6	1
35	100	4	1

Table 4.2.2: Annealing Schedule used in the second stage of Database Docking

Temperature	Steps	Max Rotation	Max Translation
100	300	5	1
1	400	2.5	0.5

For the energy term and all the outrank parameters, distribution histograms were generated. From these histograms ranges for each term were determined. Compounds that fell into all these ranges were selected and extracted from the docking output file. These compounds were visually inspected using Dockcam to check for feasibility and quality of fit. Compounds passing visual inspection were listed for future reference and potential *in vitro* experimentation.

Database docking was performed using a total of five Silicon Graphics Inc. (SGI) processors. For the docking to the aminoglycoside pocket a SGI Indigo II, SGI Indigo,

SGI Indy and a SGI O_2 R5000 were used in parallel on various sections of the database. For docking to the combined ATP Aminoglycoside pocket a SGI Indy and a SGI O_2 R5000 were used in parallel on the database split into two sections. Lastly, for the ATP pocket a SGI Indy and a SGI O_2 R10000 were used in parallel to dock the entire database. All analysis and visualization was performed on a SGI O_2 R10000.

4.3 Results and Discussion

Compounds of interest for further study in each case of the database docking were selected based on a variety of scoring functions and a visual inspection. Generally from all the compounds extracted solely based on the score functions approximately 50% were kept after visual analysis. The ranges used in extracting compounds for each of the score functions were determined after analysis of histograms which were generated based on the distribution of compounds for each term (Figures 4.3.1-4.3.4). The ranges used for the various terms were as follows: the delta NPFE (empirical calculation of Non-Polar Free Energy) range was (–8.45) - (–7.45), the delta FE (empirical calculation of Free Energy) range was 10.2 - 16.6, the Hydrogen bond score (based on number and geometry of hydrogen bonds) range was 2 – 4, and the docking energy (empirical calculation of binding energy) range was (-20) – 12. The largest weighting in determining the ranges was the shape of the histograms, an effort was made to select a particular region of the histogram peaks based on the general position of the aminoglycoside models determined in Chapter 3 when they were compared to other potential models. General lower

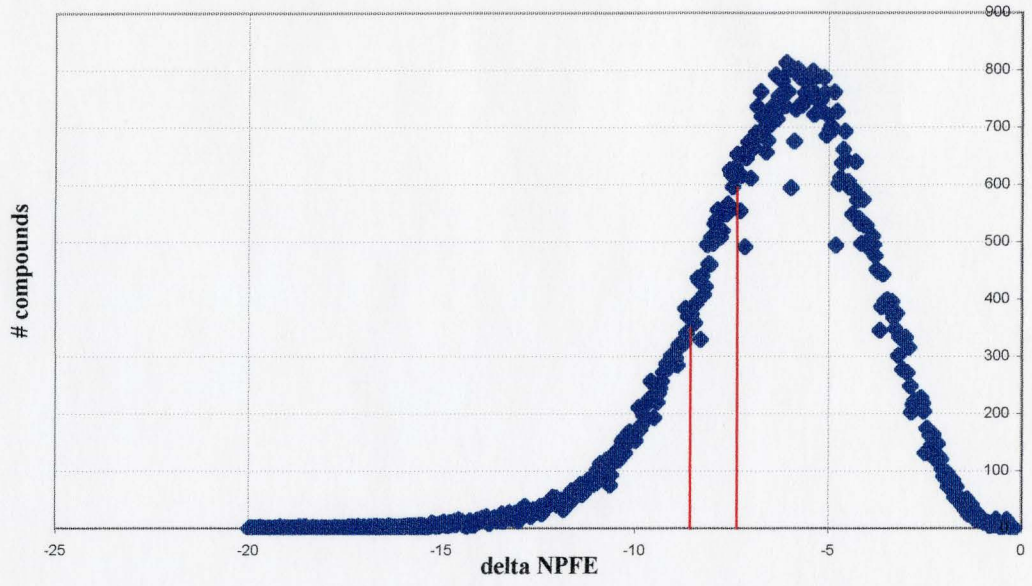


Figure 4.3.1: Histogram distribution of compounds based on delta NPFE for database docking to aminoglycoside pocket. Area between red lines indicates selected region

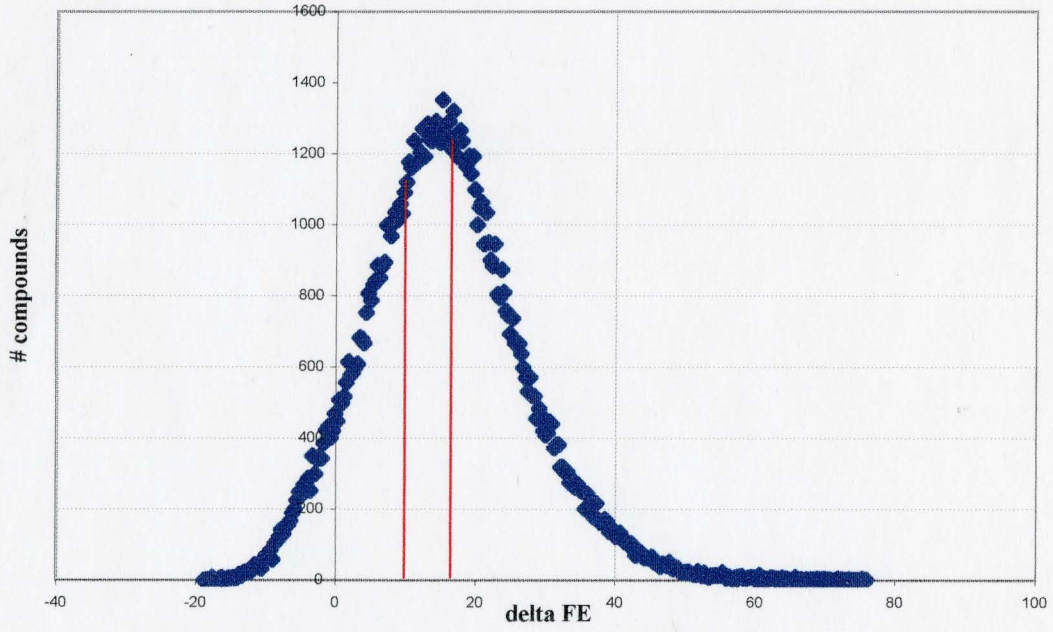


Figure 4.3.2: Histogram distribution of compounds based on delta FE for database docking to aminoglycoside pocket. Area between red lines indicates selected region.

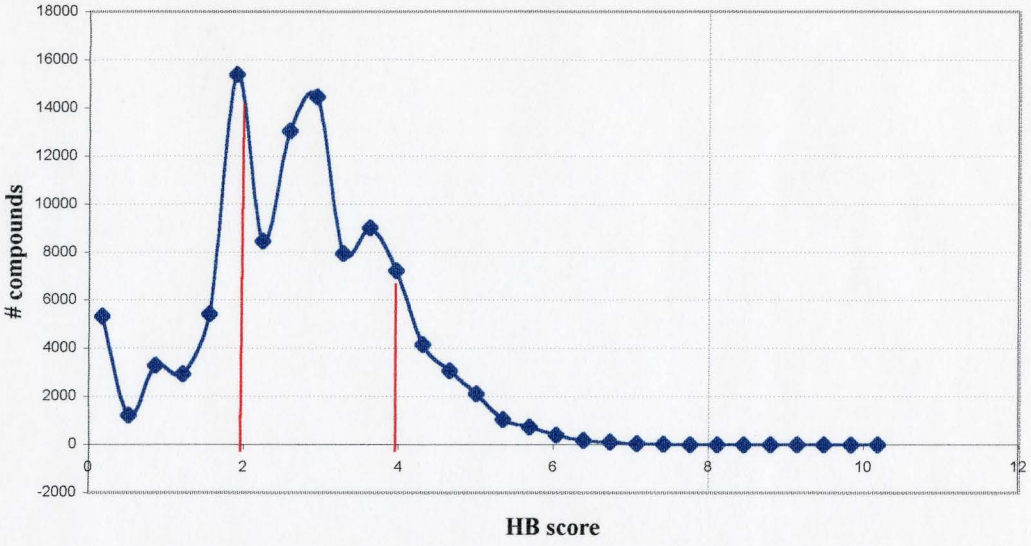


Figure 4.3.3: Histogram distribution of compounds based on Hydrogen Bond Score for database docking to aminoglycoside pocket. Area between red lines indicates selected region.

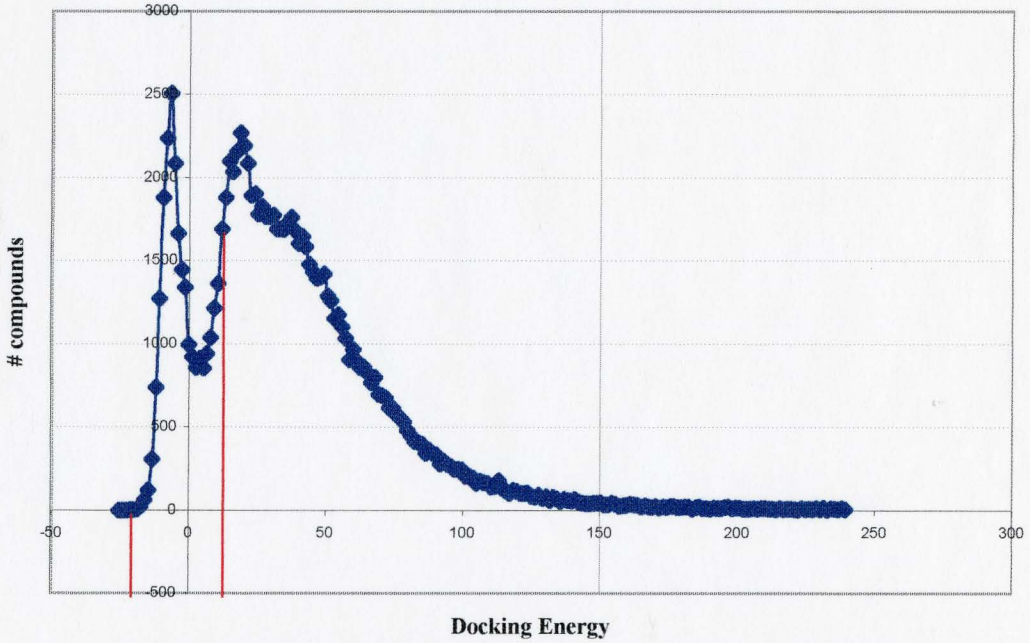


Figure 4.3.4: Histogram distribution of compounds based on Docking Energy for database docking to aminoglycoside pocket. Area betwwen red lines indicates selected region.

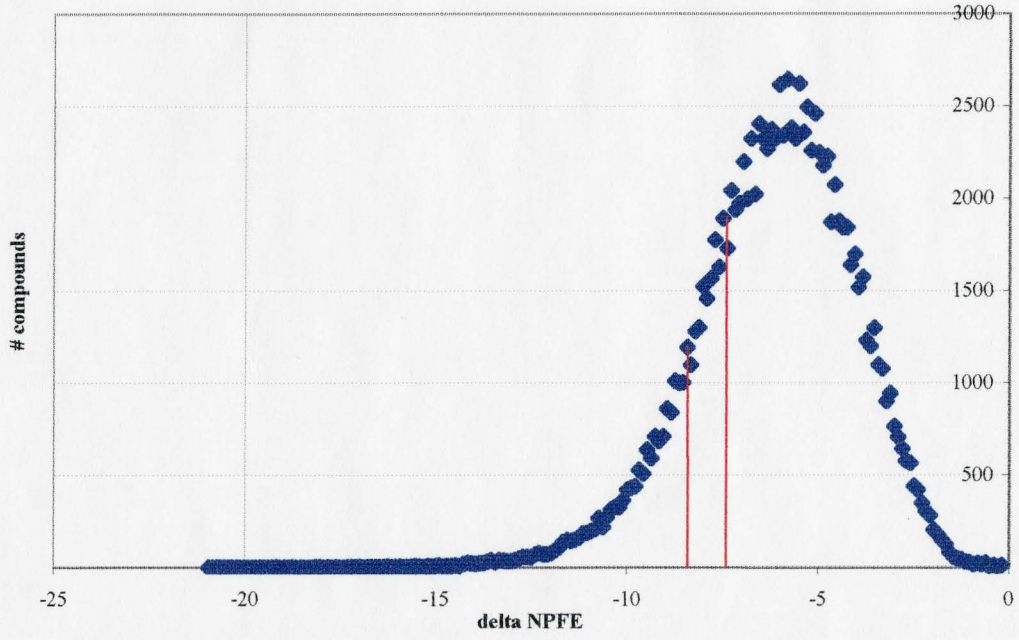


Figure 4.3.5: Histogram distribution of compounds based on delta NPFE for database docking to combined ATP and Aminoglycoside binding regions. Area between red lines indicates selected region.

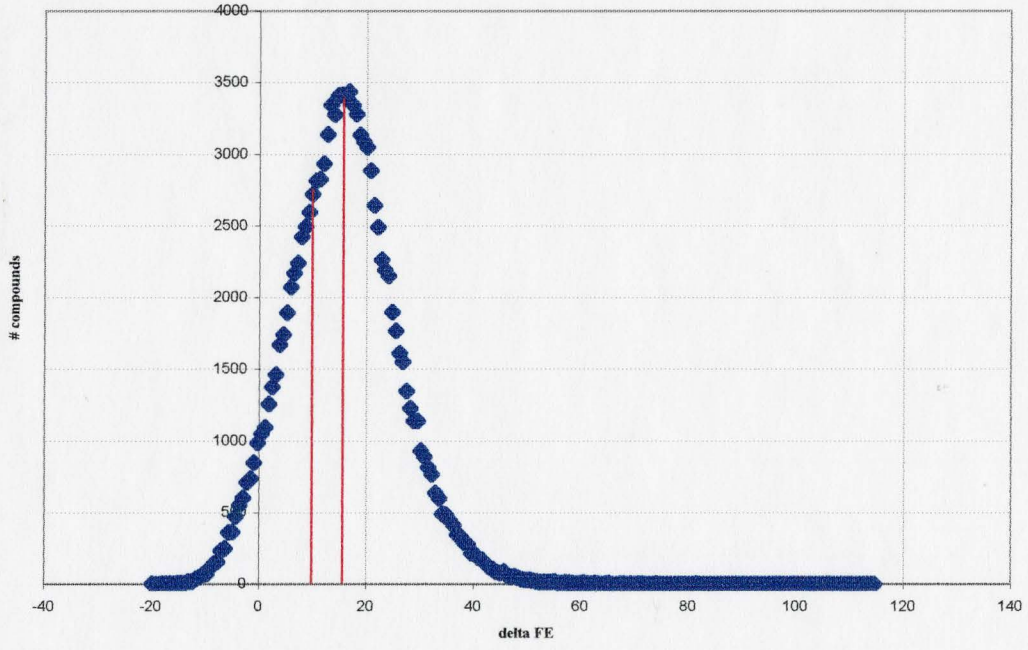


Figure 4.3.6: Histogram distribution of compounds based on delta FE for database docking to combined ATP and Aminoglycoside binding regions. Area between red lines indicates selected region.

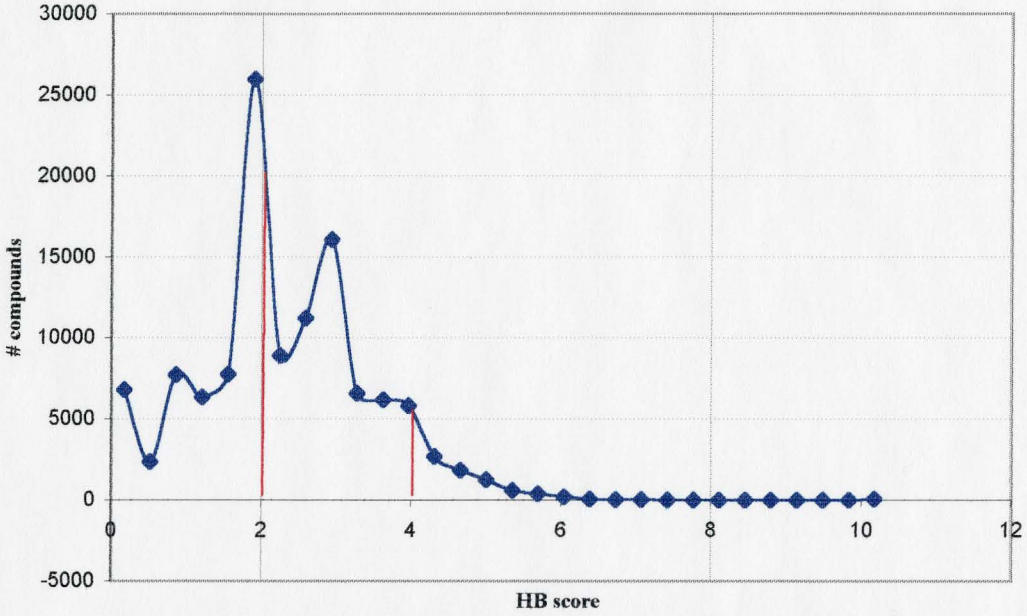


Figure 4.3.7: Histogram distribution of compounds based on HB Score for database docking to combined ATP and Aminoglycoside binding regions. Area between red lines indicates selected region.

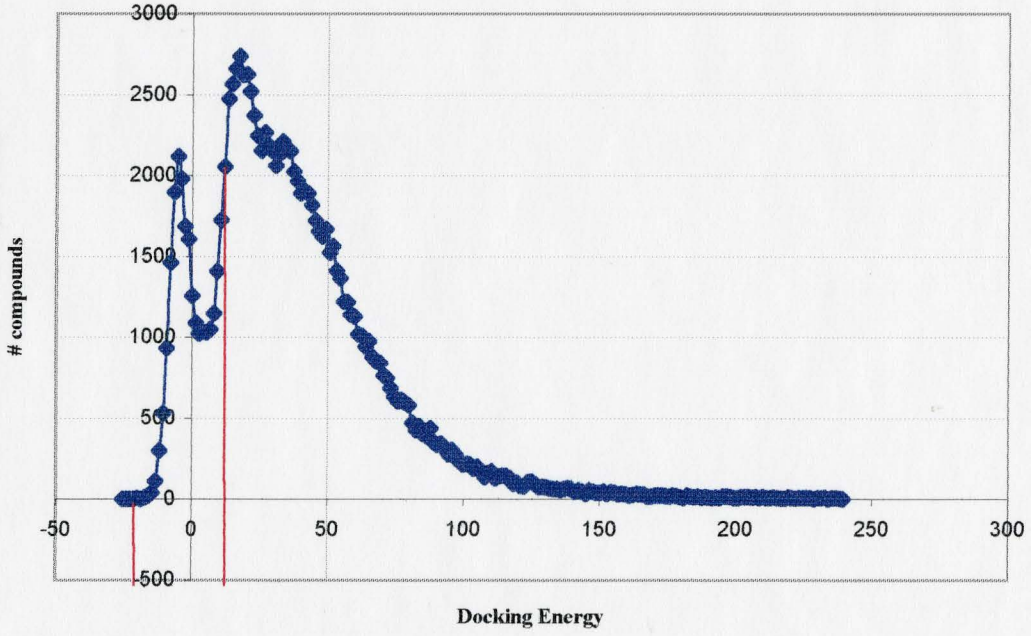


Figure 4.3.7: Histogram distribution of compounds based on Docking Energy for database docking to combined ATP and Aminoglycoside binding regions. Area between red lines indicates selected region.

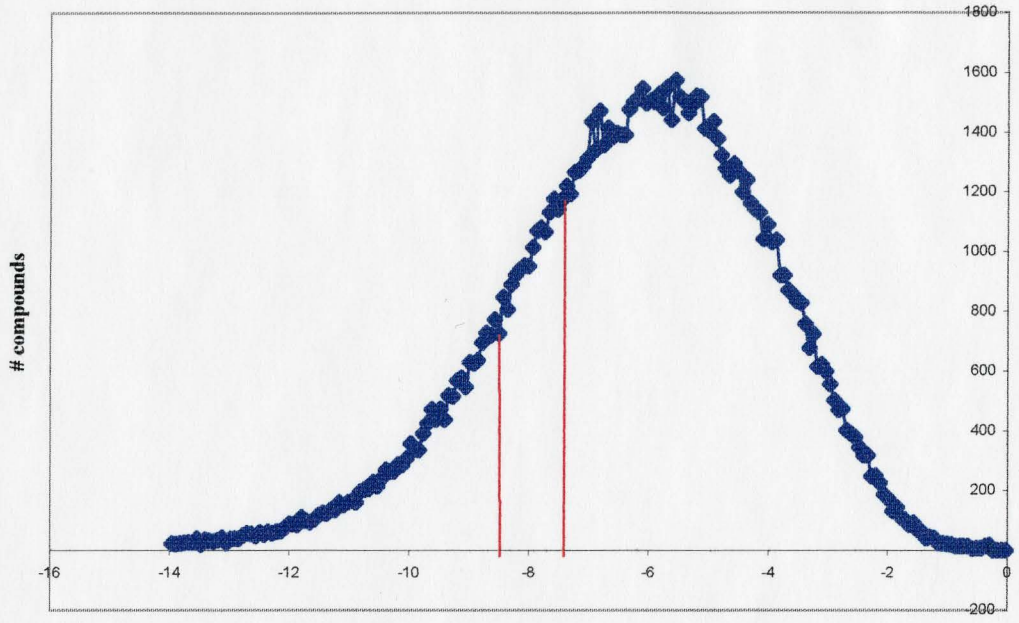


Figure 4.3.9: Histogram distribution of compounds based on delta NPFE for database docking to ATP binding region. Area between red lines indicates selected region

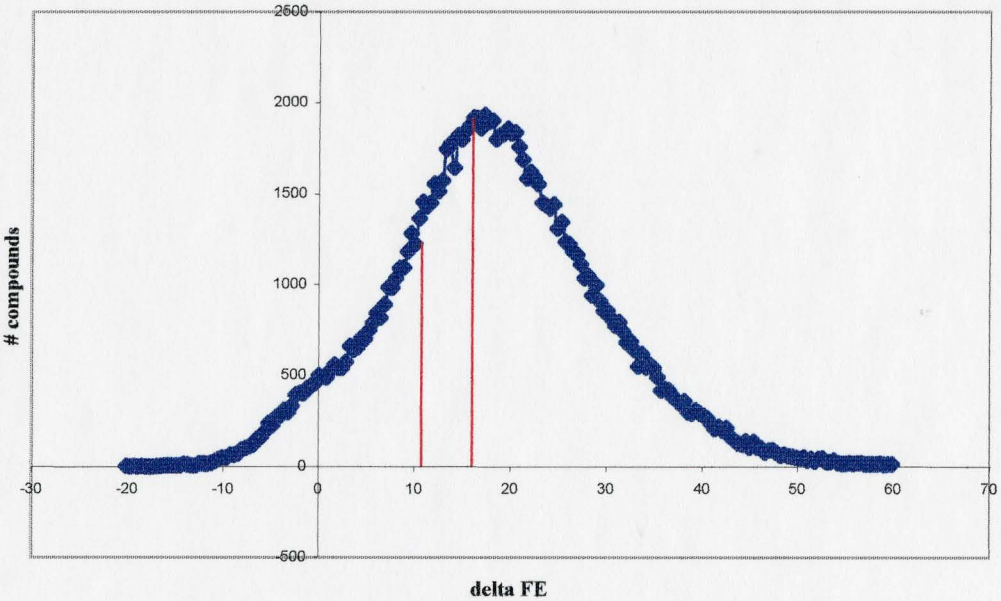


Figure 4.3.10: Histogram distribution of compounds based on delta FE for database docking to ATP binding region. Area between red lines indicates selected region

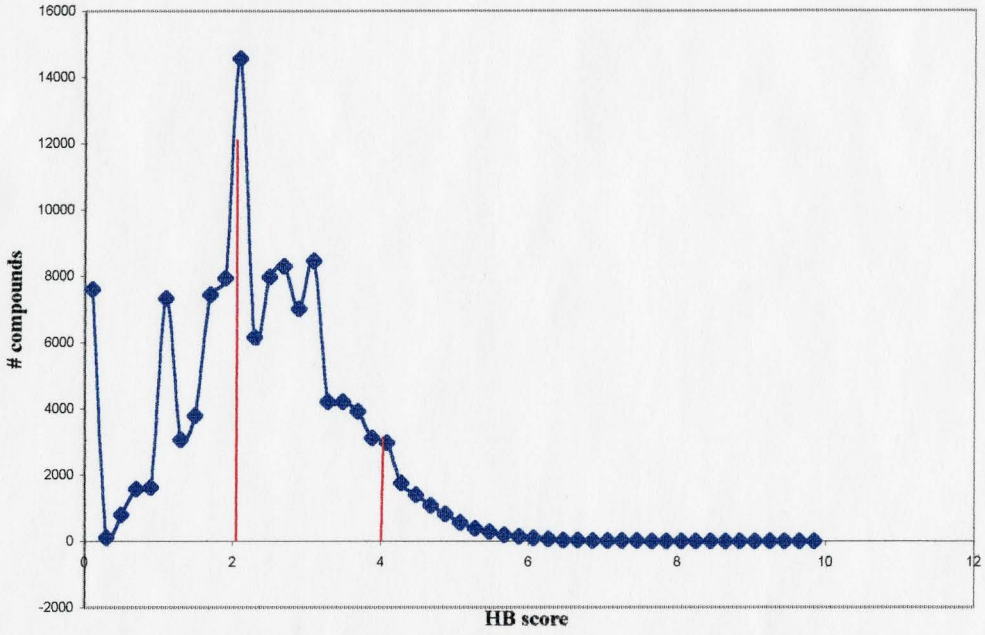


Figure 4.3.11: Histogram distribution of compounds based on delta NPFE for database docking to ATP binding region. Area between red lines indicates selected region

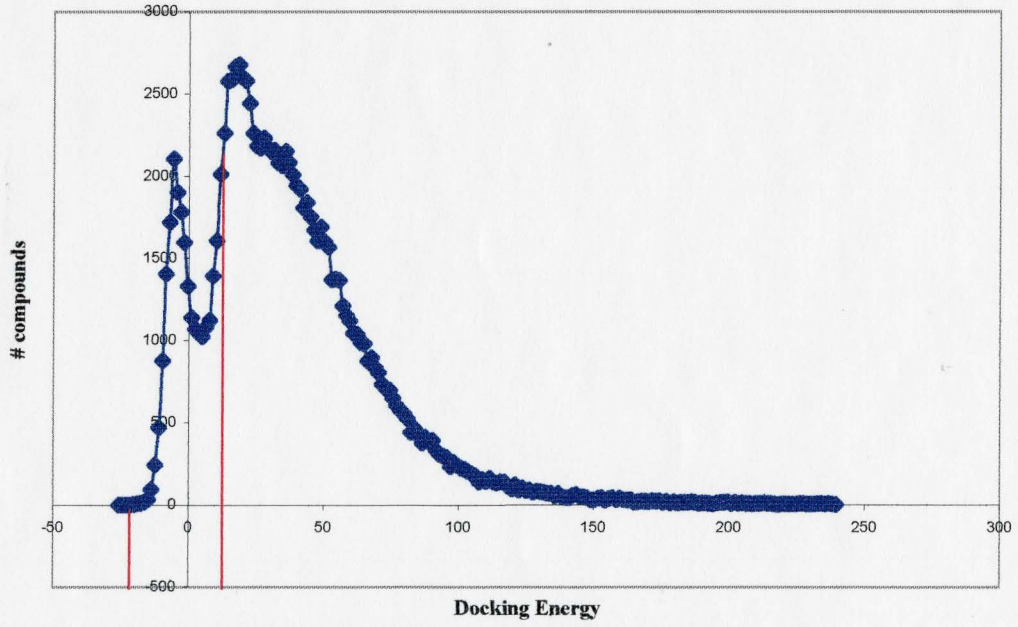
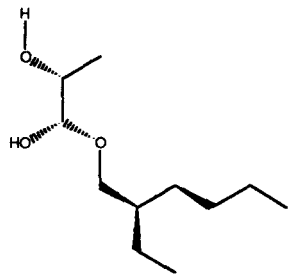


Figure 4.3.12: Histogram distribution of compounds based on delta NPFE for database docking to ATP binding region. Area between red lines indicates selected region

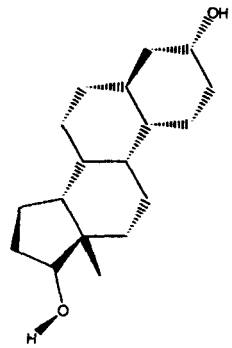
midrange values of delta NPFE and delta FE were chosen, middle to high values of Hydrogen bond score, and lower values of docking energy were chosen. The goal of this selection was to limit the number of models for visual inspection to between 50 and 150. The ranges that were selected were based on values for substrates which are know to bind to the enzyme however, it may be more useful to select ranges on the limits of the histograms rather than in the middle. If the empirical terms that are used in the range selection are meaningful the compounds with the lowest delta NPFE, delta FE and docking energies and highest HB Scores would bind the tightest to the enzyme in question. Further analysis of compounds that are present in the tails of the histograms should be performed as these regions may be more likely to contain potential inhibitors.

The histograms for each of the three database docking experiments were very similar which was interesting since each docking was to a different area of the protein. Perhaps this was a result of the scoring terms being more dependent on the compound itself and not the interaction with the protein. However, despite the similarity of the histograms and the ranges used to extract compounds only two of the selected compounds were found in common between the aminoglycoside pocket and the ATP pocket and none were found in common with the combined pocket. This indicates that the score functions are indeed an indication of how the compound interacts with the protein in the various binding pockets.

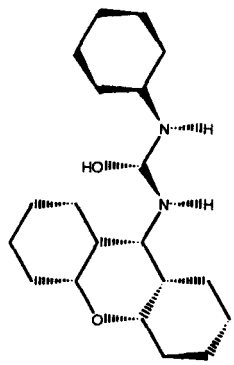
After visual analysis of the compounds a list of compounds was generated that not only met the scoring ranges but also appeared to be feasible for binding to the protein in the desired region. Since extensive electrostatic and surface knowledge has been determined for APH(3')-IIIa (Chapter 3), compounds possessing amino groups that could be cationic were selected as well as compounds that appear to have a complementary structure with the enzyme's binding pocket. In addition, compounds with rings in their structure were selected both the aminoglycosides and ATP, have multiple rings. Thus it is reasonable to assume that potential inhibitors could contain one or more rings in their structure. The selected compounds for the three regions of the protein are shown in Figures 4.3.13 – 4.3.15. Compounds that were hydrophobic in nature with long fatty acid chains were excluded after visual analysis despite their being selected in the ranges of the histograms. It is highly unlikely that these hydrophobic compounds would bind to the negatively charged APH(3')-IIIa. The fact that these compounds scored well based on the score functions but were did not make sense with the known biology and chemistry of the enzyme may be a result of the ranges selected from the distributions or perhaps an inherent problem with database docking. Large compounds have the ability to make extensive contacts with the protein surface and therefore may produce a better score than other compounds despite being hydrophobic.



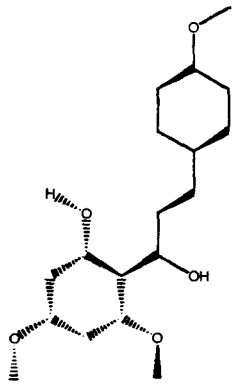
1. 6283-86-9



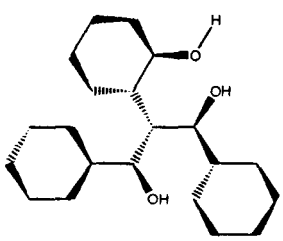
2. 1089-78-7



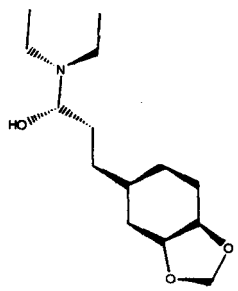
3. 6331-74-4



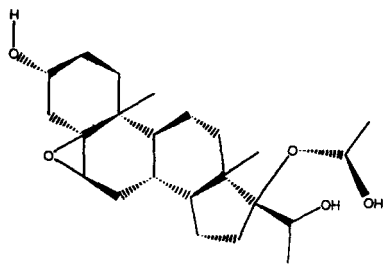
4. 3420-72-2



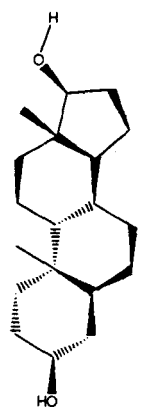
5. 6274-70-0



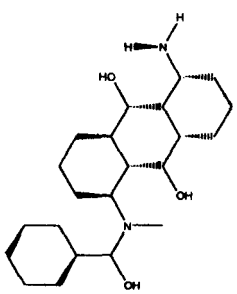
6. 40951-05-1



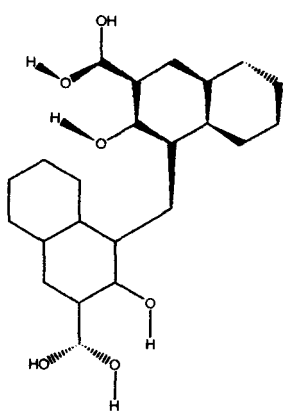
7. 34553-66-7



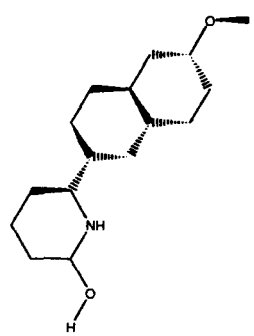
8. 65-06-5



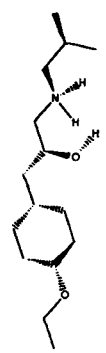
9. 6337-20-8



10. 130-85-8

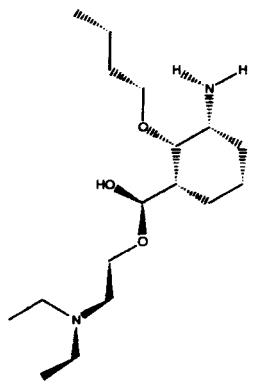


11. 93008-89-0

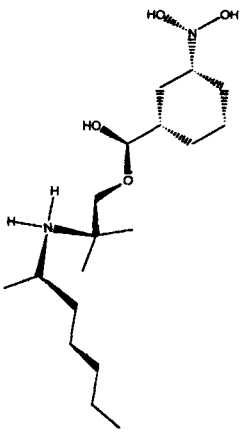


12. 6302-19-8

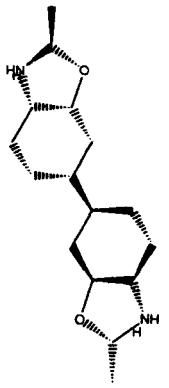
Figure 4.3.13 (Page 1)



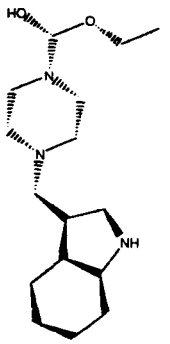
13. 550-01-6



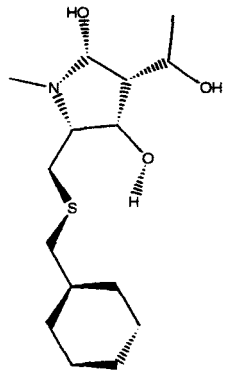
14. 6625-13-4



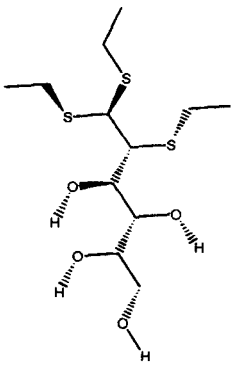
15. 6309-66-6



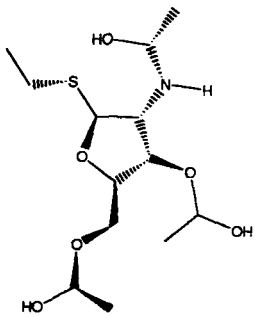
16. 6308-64-1



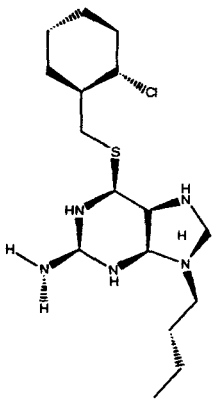
17. 94252-91-2



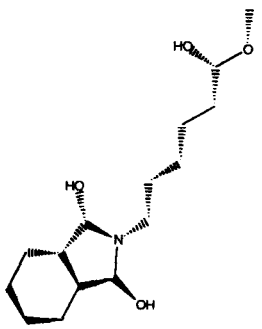
18. 15356-41-9



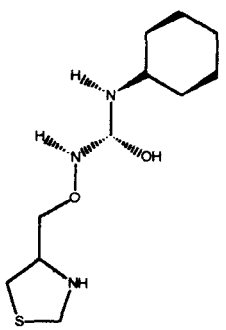
19. 24807-83-8



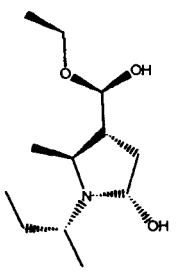
20. 735-18-2



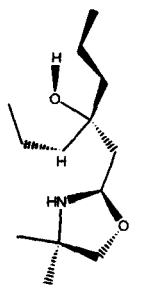
21. 19991-87-8



22. 24885-86-7

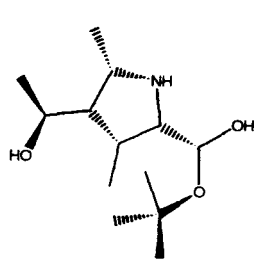


23. 6946-41-4

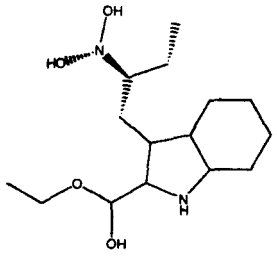


24. 51869-20-6

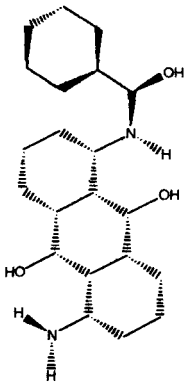
Figure 4.3.13 (Page 2)



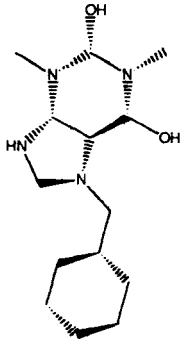
25. 63040-83-5



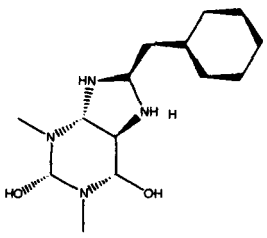
26. 92700-75-9



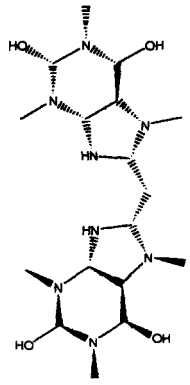
27. 117-06-6



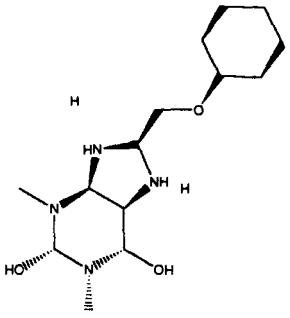
28. 1807-85-8



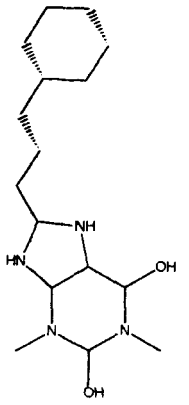
29. 2879-15-4



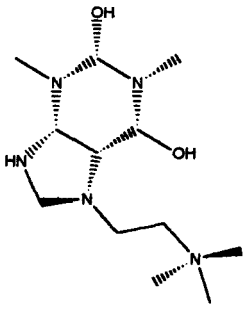
30. 5429-47-0



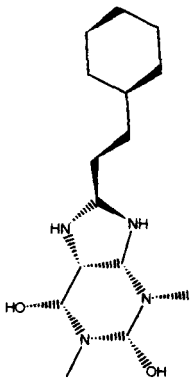
31. 5426-70-0



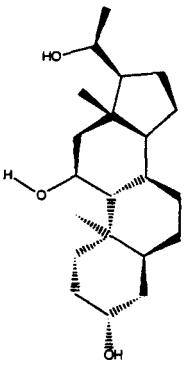
32. 74039-67-1



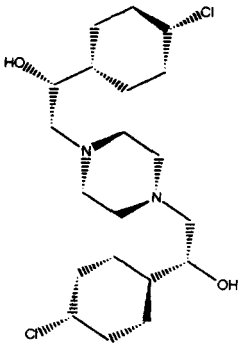
33. 7145-94-0



34. 5426-90-4



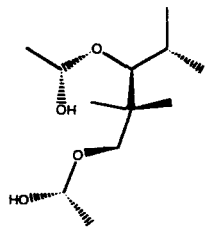
35. 600-57-7



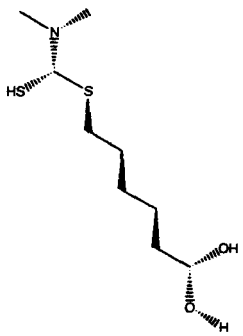
36. 5447-51-8

Figure 4.3.13 (Page 3)

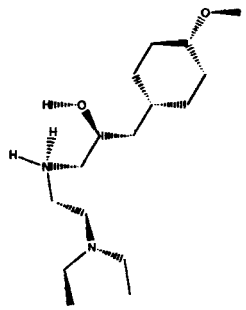
Figure 4.3.13: Compounds selected from database docking to aminoglycoside binding region of APH(3')-IIIa. The standard identifier is listed beneath each compound. Figure continued from previous pages. The 36 compounds were selected by visual inspection from a list of 105 compounds.



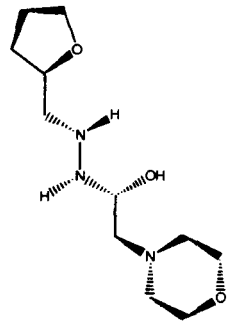
1. 4100-09-8



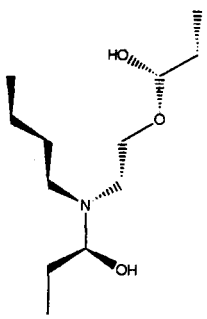
2. 94646-69-2



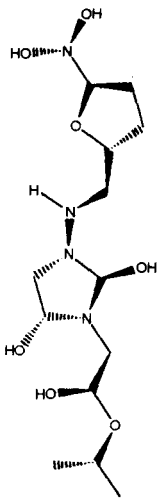
3. 6304-08-1



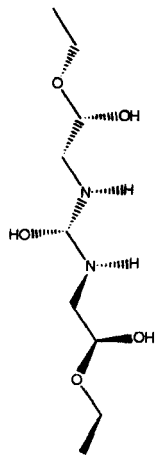
4. 61221-07-6



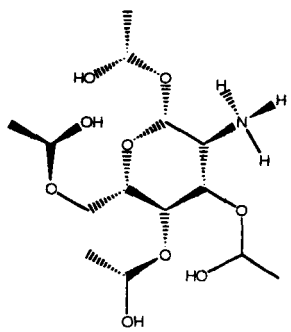
5. 6946-53-8



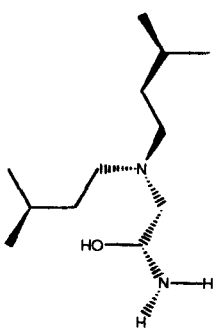
6. 1749-96-8



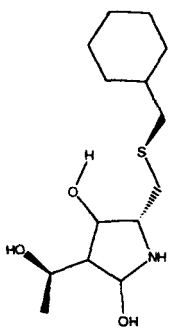
7. 7150-63-2



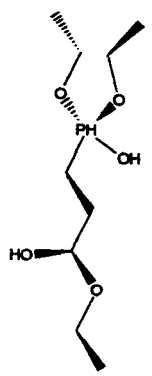
8. 10034-20-5



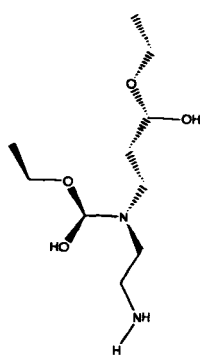
9. 7409-53-2



10. 92248-75-4

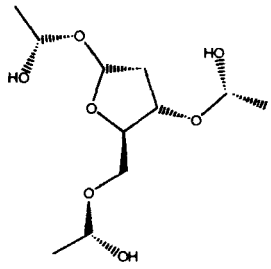


11. 3699-67-0

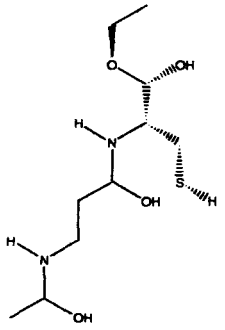


12. 1438-40-0

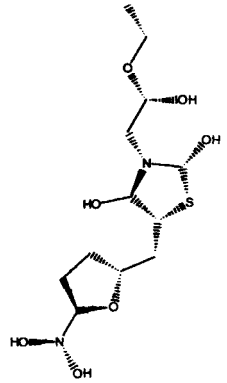
Figure 4.3.14 (Page 1)



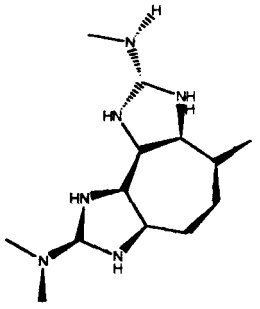
13. 4594-52-9



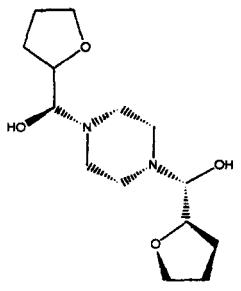
14. 64949-81-1



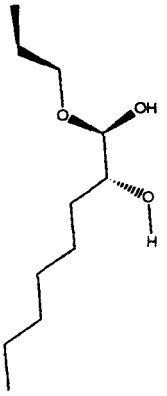
15. 27550-27-2



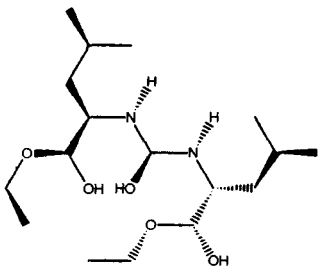
16. 57695-32-6



17. 31350-27-3



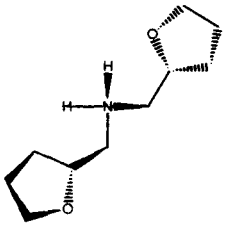
18. 7472-56-2



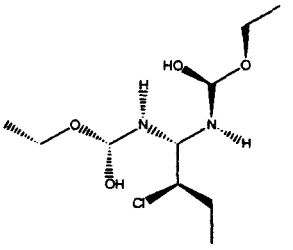
19. 70779-90-7

Figure 4.3.14 (Page 2)

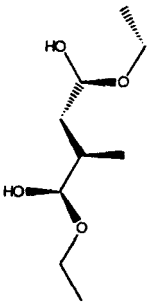
Figure 4.3.14: Compounds selected from database docking to combined aminoglycoside and ATP binding regions of APH(3')-IIIa. The standard identifier is listed beneath each compound. Figure continued from previous page. The 19 compounds were selected by visual inspection from a list of 45 compounds.



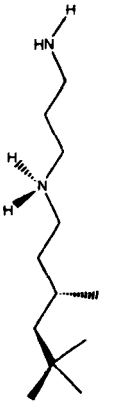
1. 5343-16-8



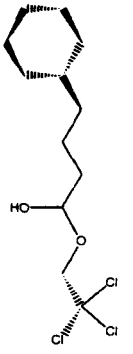
2. 5336-12-9



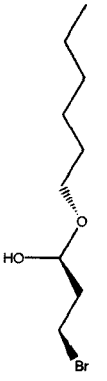
3. 2409-52-1



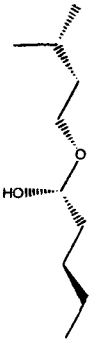
4. 6281-54-5



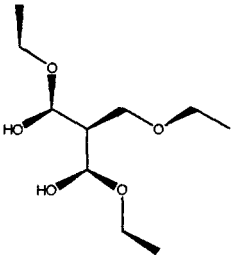
5. 5432-70-2



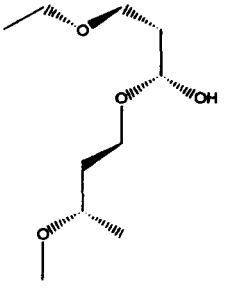
6. 6282-50-4



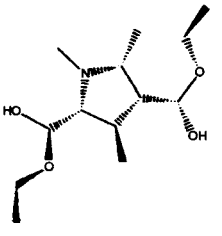
7. 2050-09-1



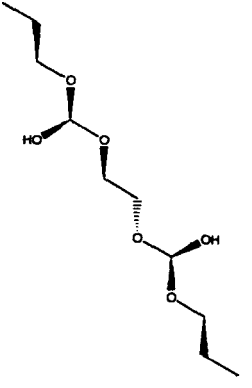
8. 87-13-8



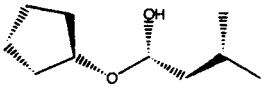
9. 5434-56-0



10. 5462-32-8

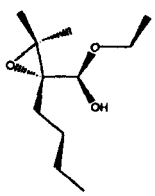


11. 4074-91-3

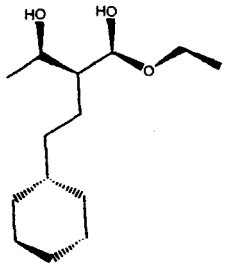


12. 5452-04-0

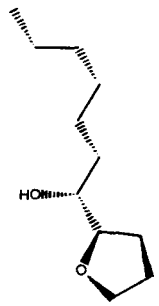
Figure 4.3.15 (Page 1)



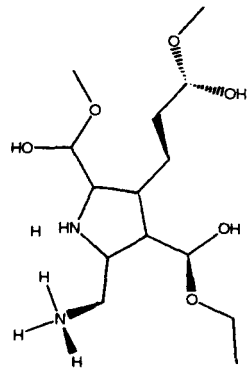
13. 5445-36-3



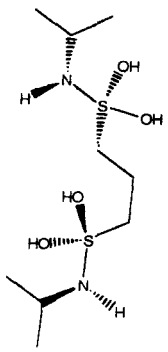
14. 6949-93-5



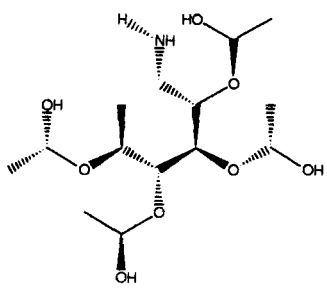
15. 5466-40-0



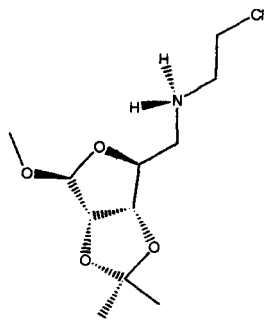
16. 6960-59-4



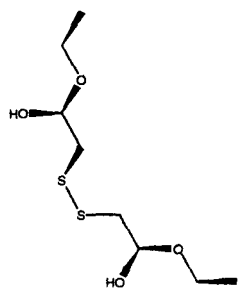
17. 6941-62-4



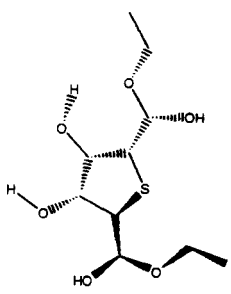
18. 6303-62-2



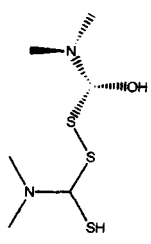
19. 54946-35-9



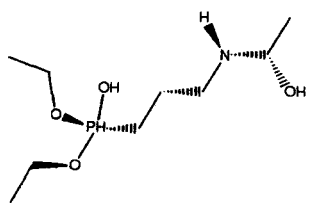
20. 1665-65-2



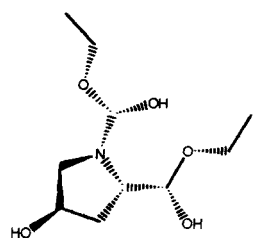
21. 1822-66-8



22. 1115-06-6

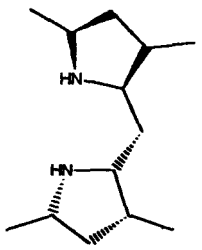


23. 54008-28-5

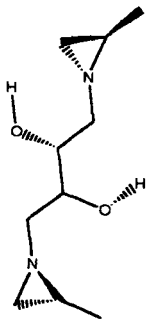


24. 4801-24-5

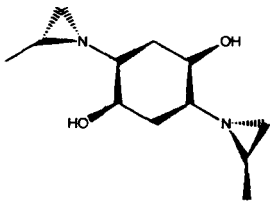
Figure 4.3.15 (Page 2)



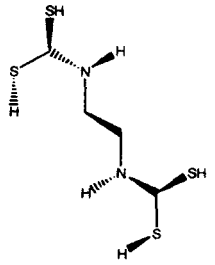
25. 15536-55-7



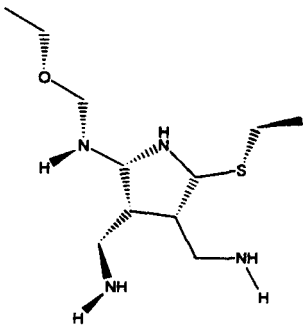
26. 15870-20-9



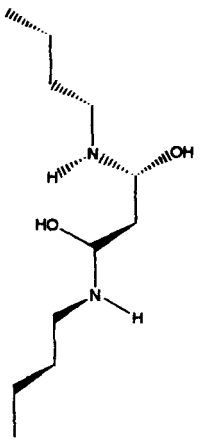
27. 23650-98-8



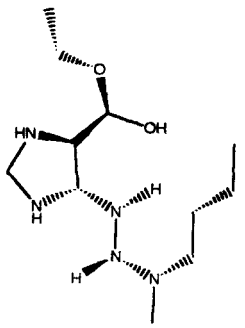
28. 91859-15-3



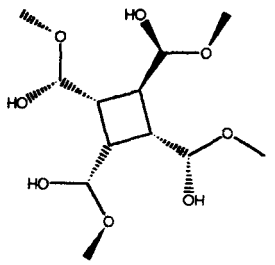
29. 1500-89-6



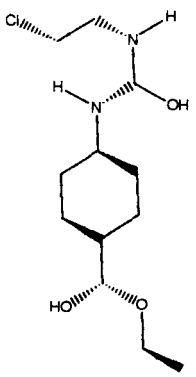
30. 10174-66-0



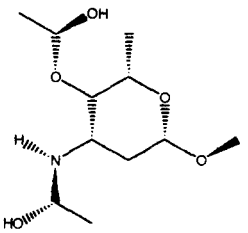
31. 36137-86-7



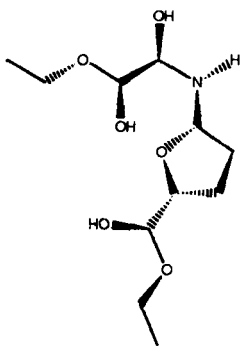
32. 3999-67-5



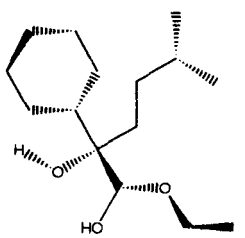
33. 13908-22-0



34. 18981-63-0

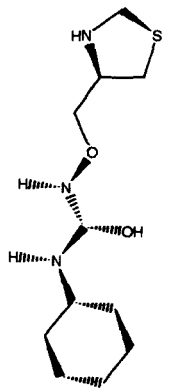


35. 67318-13-2

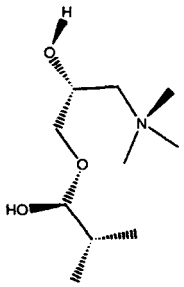


36. 92957-02-3

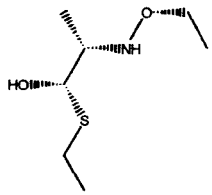
Figure 4.3.15 (Page 3)



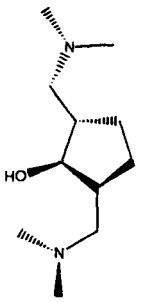
37. 24885-86-7



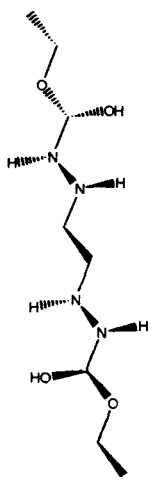
38. 13052-11-4



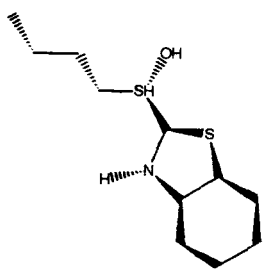
39. 18241-70-8



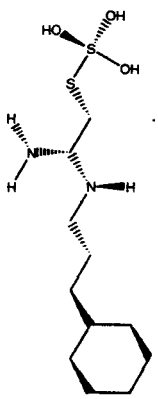
40. 6333-22-8



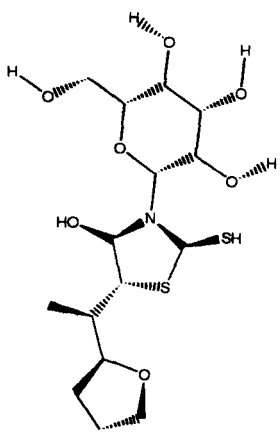
41. 80805-23-8



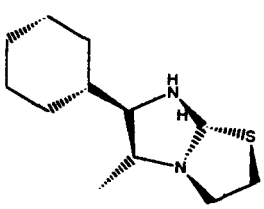
42. 79252-73-6



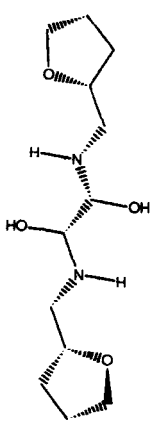
43. 40283-58-7



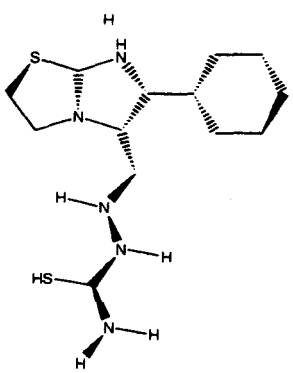
44. 65562-46-1



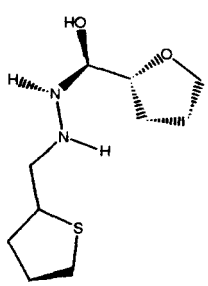
45. 51226-38-1



46. 69010-90-8

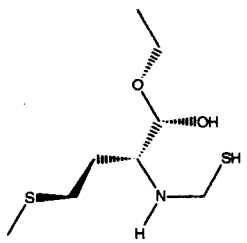


47. 75224-72-5

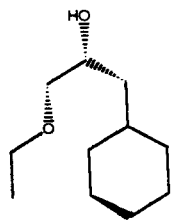


48. 60943-76-2

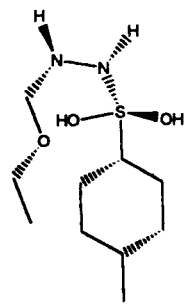
Figure 4.3.15 (Page 4)



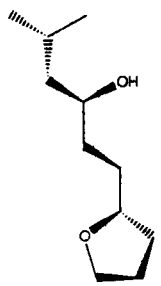
49. 58511-07-2



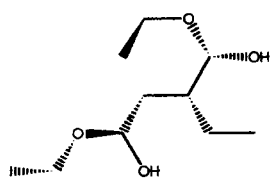
50. 51149-72-5



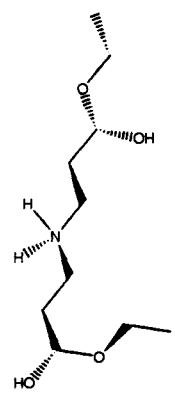
51. 3898-95-1



52. 4196-96-7



53. 4795-18-0



54. 3518-88-5

Figure 4.3.15 (Page 5)

Figure 4.3.15: Compounds selected from database docking to ATP binding region of APH(3')-IIIa. The standard identifier is listed beneath each compound. Figure continued from previous pages. The 54 compounds were selected by visual inspection from a list of 118 compounds.

Figure 4.3.16 shows compound # 3 from Figure 4.3.13 bound to the aminoglycoside pocket as an example of a good scoring compound that passed visual inspection. This compound had values of –8.36 for delta NPFE, 11.92 for delta FE, 2.98 for HB Score, and -9.92 for docking energy. Figure 4.3.17 shows an example of a compound that passed the histogram selection procedure but failed the visual inspection. This compound had values of –8.49 for delta NPFE, 14.06 for delta FE, 3.00 for HB Score, and –10.57 for docking energy. The docking results show this compound with a long hydrophobic tail inside the negatively charged aminoglycoside binding pocket. The prospect of binding a hydrophobic compound to a charged region of a protein is highly unlikely.

Compound #46 from Figure 4.3.15 is shown in Figure 4.3.18 bound to the ATP binding region of APH(3')-IIIa. This compound has value of –8.18, 14.32, 2.50, and -0.81 for delta NPFE, delta FE, HB Score, and docking energy respectively. The Figure shows the compound mostly within the ATP binding region and part of the compound within the aminoglycoside binding region. Figure 4.3.19 shows the structure of a good scoring compound that failed the visual inspection. This compound had scores of –7.98, 11.72, 2.00, and –3.02 for delta NPFE, delta FE, HB Score, and docking energy respectively. The structure of this compound is a very long straight chain hydrophobic compound. It extends from the ATP pocket to the Aminoglycoside pocket but has only two amino groups at the extreme terminus. The non-polar nature and length of this compound resulted in its rejection after visual inspection.

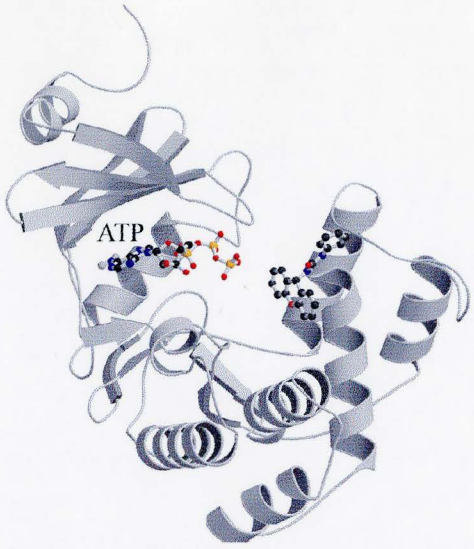


Figure 4.3.16: Selected compound bound to Aminoglycoside binding region of APH(3')-IIIa (ribbon). Compound # 3 from Figure 4.3.13 (ball-and-stick) was selected by score functions and visual inspection.

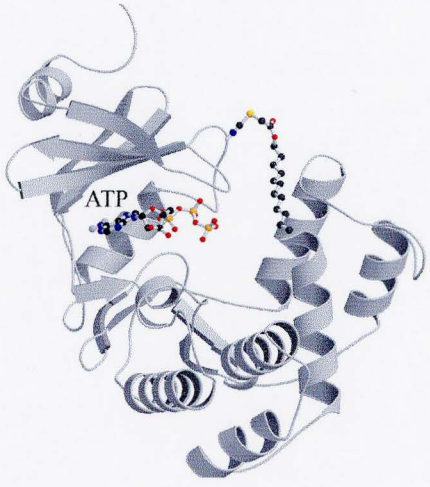


Figure 4.3.17: Rejected compound bound to Aminoglycoside binding region of APH(3')-IIIa (ribbon). This compound (ball-and-stick) was selected based on score functions but rejected after visual inspection.
Figures prepared using MOLSCRIPT (Kraulis, 1991) and Raster3D (Merritt and Bacon, 1997).

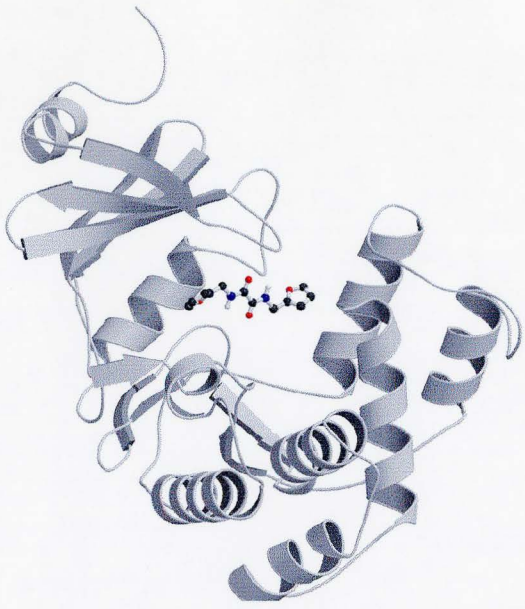


Figure 4.3.18: Selected compound bound to ATP binding region of APH(3')-IIIa (ribbon). Compound # 46 from Figure 4.3.15 (ball-and-stick) was selected based on score functions and visual inspection.

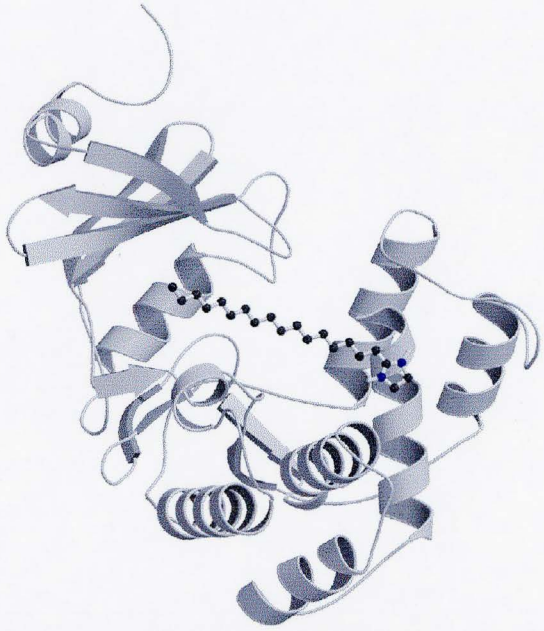


Figure 4.3.19: Rejected compound bound to ATP binding region of APH(3')-IIIa (ribbon). This compound (ball-and-stick) was selected based on score functions but rejected after visual inspection.
Figures prepared using MOLSCRIPT (Kraulis, 1991 and Raster3D (Merritt and Bacon, 1997).

The structures selected from the database docking experiments have provided models by which these compounds may bind to the various regions of APH(3')-IIIa. If these models have been correctly predicted the listed compounds may possess the ability to inhibit the aminoglycoside modifying ability of this enzyme. In order to determine the validity of any of the models, kinetic analysis is required to confirm the ability of the enzyme to bind the specific compound, in addition structural studies would be needed to confirm the actual binding mode of the compound to the enzyme.

As a control for the RSDB algorithm, database docking was performed on ADP to the ATP pocket. Two separate docking trials were performed with different constraint files. The first constraint was larger and extended into the aminoglycoside pocket, the second was smaller and contained only space within the ATP binding pocket. The Outrank values are shown in Table 4.3.1.

Table 4.3.1: Outrank values for ADP docking using RSDB algorithm

Compound	delta NPFE	delta FE	HB Score	Docking Energy
ADP (big constraint)	-4.35	13.81	5.36	13.1
ADP (small constraint)	-7.02	14.02	4.01	17.4
ADP (crystal structure)	-7.09	6.92	3.90	ND

ND – Not Determined

The results with the big constraint file do not overlay at all with the crystallographically determined structure. The ADP in this case is located inside the aminoglycoside pocket.

This is not suprising since the aminoglycoside pocket is much larger thereby reducing the steric clashes with the protein. However, the smaller constraint file docking structure overlays nicely with the crystal structure as shown in Figure 4.3.20. The similarity of these structures makes sense when we compare the Outrank values. These results show that the Outrank terms are useful for selection of structures from database docking. In addition, these results show the role that the constraint file plays on correctly predicting the position of the ligand. When the constraint file is large the algorithm has difficulty in finding the correct position. Usually when performing database docking experiments the exact binding position of the ligand would not be known and therefore the allowed region would be made large to allow for many possibilities. This could create a problem of missing potential good inhibitors because they were docked to an incorrect position. A similar problem was noticed with the combined pocket docking in that the majority of the good scoring compounds were positioned mostly within the aminoglycoside binding pocket, again the larger size here creates less chances for steric hindrance.

The database docking algorithm was also employed on various aminoglycosides, kanamycin, amikacin, butirosin, and ribostamycin. The Outrank analysis showed that the aminoglycosides did not fall within the ranges used for selecting the compounds from the database docking (Table 4.3.2). This is not surprising when the structure of the kanamycin generated from database docking was compared to that of the aminoglycoside docking from Chapter 3 (Figure 4.3.21).

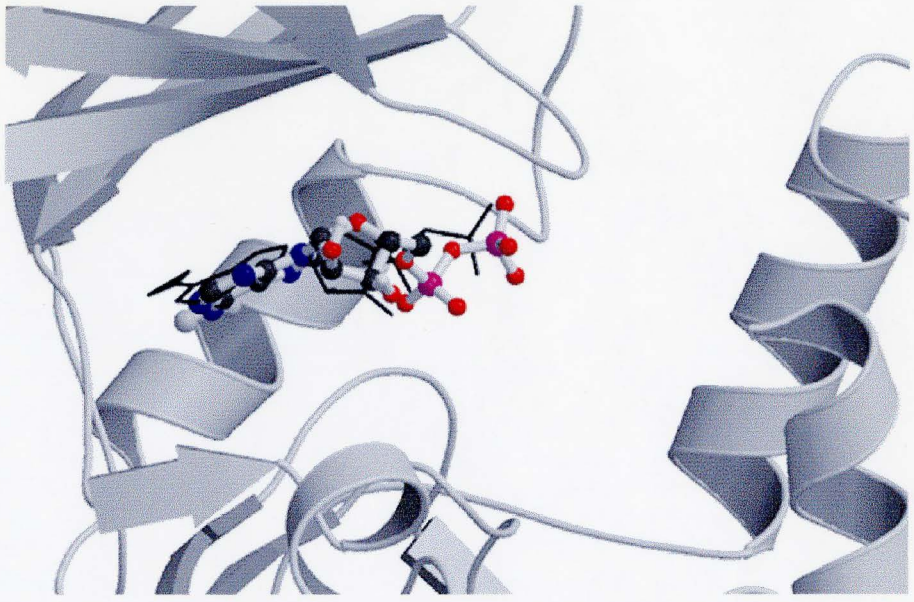


Figure 4.3.20: Overlay of database docked ADP (wireframe) with ADP determined in crystal structure of APH(3')-IIIa (ball-and-stick).

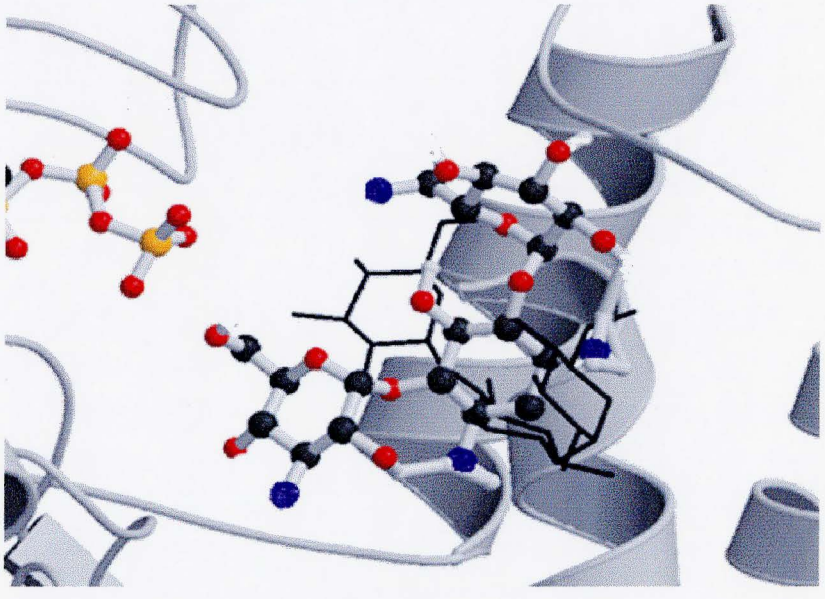


Figure 4.3.21: Overlay of database docked kanamycin (wireframe) with kanamycin determined in Chapter 3 (ball-and-stick).

Figures prepared using MOLSCRIPT (Kraulis, 1991) and Raster3D (Merritt and Bacon, 1997).

Table 4.3.2: Outrank results from the database docking of aminoglycoside antibiotics

Compound	Delta NPFE	delta FE	HB Score	Docking Energy
Kanamycin	-0.78 (-6.10)	25.12 (16.91)	0.0 (5.1)	21.0
Amikacin	-0.84 (-7.61)	29.26 (12.88)	0.0 (4.2)	33.6
Butirosin	-0.64 (-7.03)	28.06 (15.74)	0.0 (7.7)	11.7
Ribostamycin	-0.69 (-6.02)	27.01 (34.04)	0.0 (3.2)	9.4

Values in brackets indicate the Outrank values determined for the aminoglycoside model determined in Chapter 3.

One major problem with database docking is the score functions that are available to evaluate the complex structures of the database compound with the macromolecular target. Because of the speed required for each compound in order to complete the docking of the entire database in a reasonable amount of time the interactions that are simulated and calculated need to be limited. In the case of DockVision this is limited to a hydrogen bond function and a steric function. Thus some factors that may be of importance in binding a ligand may be missed.

The poor quality of the database docking scoring functions has led to the possibility of false positives (private communication, Trevor Hart, Department of Medical Microbiology and Infectious Diseases, University of Alberta). These are compounds that score well during the docking simulations but in a real situation do not bind to the target molecule. The problem of false positives is accepted, as the number of compounds required to test experimentally is greatly reduced versus the entire database of molecules. A more serious problem is the one of false negatives this is when a compound that does inhibit the protein in question is not selected by the docking scoring

function. This problem can be severe since the database docking may overlook a potential inhibitor resulting in a great chance it will never be experimentally tested. It is impossible to determine how many times a database docking algorithm has missed a good inhibitor of a target molecule, unless the inhibitor is already known. The analysis of the aminoglycosides and ADP by the RSDB algorithm show how common a false negative could be. In the case of the aminoglycosides, they would not have been selected from the database docking based on any criteria, yet they are the natural substrates of the enzyme. For ADP, using a large constraint file as would be standard in database docking would have resulted in the rejection of this compound. Thus two factors seem to be important in database docking, the first is the score functions ability to predict the binding mode of the compound and the second is how the space allowed for docking is selected. This second point is discouraging because this means that for accurately determining a model by database docking a small allowed space would be required.

The results of database docking analysis must always be taken in combination with experimental evidence (either structural or kinetic) to provide a complete picture of the inhibitory capabilities of the compounds in question. Selection of compounds from a database docking study needs to use all available information. The score function alone provided with DockVision is not sufficient for selecting compounds, the ADP models that were generated showed that the ADP that overlaid with the crystallographically determined ADP scored worse than the ADP that was docked in the wrong portion of the protein (Table 4.3.1). The Outrank terms on the other hand provide information

regarding how well the model fits into the protein. It seems that the most important term in Outrank is the delta NPFE term. In the case of ADP the delta FE and HB Score terms for the wrong model and the model that was similar to the experimental structure were not significantly different but the delta NPFE was very different (Table 4.3.1). This term had the tightest range of value during the selection procedure. The use of the histograms to look at the distributions of the compounds based on the various terms was a novel approach to investigate the ranges of values that were important for selecting a workable number of compounds. Finally the graphical analysis was essential to make sure that the selected compounds made chemical sense with what is known about the protein and that they were bound to the correct region of the protein. Overall, if using the method of database docking for structure based drug design all the information available for selection of compounds must be used. The information presented in Outrank which quantifies buried and exposed surface area of the protein and ligand, as well as hydrogen bonding is the most useful information for compound selection. A method is also required to determine which range of values is desirable, the histogram method is one way of performing this task.

Generally, database docking can be a useful tool for structure based drug design efforts. Cases exist where selected compounds are able to bind to the desired protein and do exhibit some level of inhibition (Shoichet et al., 1993). The largest area that still requires improvement is in the selection of the compounds for further testing. The tools that were applied in the research presented here including Outrank and the histograms for

selection purposes have provided some new ideas into how selection can be accomplished.

Concluding Remarks

In summary, the structural studies of the aminoglycoside modifying enzyme AAC(6')-APH(2") have provided the initial ground work for structure based drug design. These studies have provided initial crystallization conditions that may be refined to provide diffraction quality crystals. In addition, the analysis of this protein by DLS has provided information regarding the aggregation state of the protein in a variety of storage buffers. This information should be applicable to producing higher quality protein crystals.

The molecular docking studies involving the aminoglycoside modifying enzyme APH(3')-IIIa have produced models for the binding of various aminoglycosides to the protein. The most significant information derived from this research involves the methods used for model selection. The first consideration is the force field that is used for performing the docking and producing a score for the models. The most important consideration for model selection from a docking study is the application of biochemical data to the problem. This research applied some biochemical data at the beginning of the docking in the form of restraints to bias the models towards the correct chemistry. During the model selection process biochemical data was employed in order to

select a model that could explain the data. Using biochemical data to bias the docking procedure was a novel procedure and was a significant improvement towards the use of docking in model selection.

Database docking with the aminoglycoside modifying enzyme APH(3')-IIIa was performed in an effort to determine potential inhibitors for the enzyme. Three areas of the protein were studied and lists of potential inhibitors were generated for each area. The method used for selecting the potential inhibitors was a novel procedure. The method involved using a variety of different score functions and looking at the distribution of compounds based on these score functions. The distributions allowed the selection of a range of values for the various score functions, and compounds that fell within all these ranges were selected for visual screening.

The research presented here has provided some initial steps towards structure based drug design. The production of pure AAC(6')-APH(2") and the crystallization trials provide the initial steps towards the resolution of a three-dimensional structure of this protein, eventually leading to drug design. Additionally, models generated in the docking studies should provide a basis for future work towards structure based drug design on the APH(3')-IIIa protein. More importantly, the novel methods used in the docking studies could provide a precedent for future structure based drug design efforts.

Future Work

For the crystallization of the bifunctional AAC(6')-APH(2") a number of areas remain to be explored. Generally, in crystallization it is desirable to have the protein stored in as close to pure water as possible (McPherson, 1999). Therefore it may be useful to explore storage buffers at a lower ionic strengths. In addition, it could be useful to further explore the protein in combination with various substrates. Only a small number of substrates have been attempted to this point. These studies should involve using DLS to examine the protein's behavior in the presence of the various substrates. In using additional substrates the amount of substrate to be used can be derived based on the known affinities between the enzyme and substrate rather than simply a random molar excess of substrate being used. Another area than can be explored is the two individual domains of the enzyme. Initial crystallization experiments with the APH(2") domain have been performed, however DLS experiments with substrates have not been performed and extensive crystallization has not yet been performed. No crystallization of the AAC(6') domain has been attempted to this point, experiments to examine the structure of this part of the bifunctional enzyme should be undertaken.

Future efforts for the aminoglycoside models would require the resolution of a three-dimensional structure of APH(3')-IIIa with the aminoglycoside bound.

In this way the model from docking could be compared to the crystallographic structure, thereby proving or disproving the docking models. Another experiment that could be attempted with docking would involve altering the dielectric constant used during the docking procedure this would alter the effect of the electrostatics on the overall score function and may improve the ability of the docking energy score to select an appropriate model.

Future work on the database docking of the NCI database to APH(3')-IIIa should involve a reanalysis of the data that has been generated. The method of using the histograms to select ranges of values for each score function is useful. However, the ranges that have been selected may not be correct for determining potential inhibitors. It may be more meaningful to select ranges that exist at the tails of the distributions rather than in the middle of the histograms. The compounds at the extremes (low delta NPFE, low delta FE, low docking energy, and high HB Score) may actually be better potential inhibitors than the ones that have been selected. The compounds at these extremes may overall have better affinities for the enzyme than ones whose score function values are similar to the known enzyme substrates. Therefore, a reanalysis of the database docking data and selection of compounds based on new score function values could prove useful for discovering a lead compounds that may act as inhibitors.

References

Bartz, Q. R., 1948, Isolation and Characterization of chlormycetin, *Journal of Biological Chemistry* **172**: 445-450.

Berkelman, R. L., and Hughes, J.D., 1993, The Conquest of Infectious Diseases: Who Are we Kidding?, *Ann. Intern. Med.* **119**: 426-428.

Bloom, B. R., and Murray, C.J., 1992, Tuberculosis: commentary on a reemergent killer, *Science* **257** (5073): 1055-1064.

Bottcher, H. M., 1963, *Miracle Drugs: A History of Antibiotics* (William Heinemann Ltd., London).

Bradford, M. M., 1976, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal Biochem* **72**: 248-254.

Burkhard, P., Taylor, P., and Walkinshaw, M.D., 1998, An Example of a Protein Ligand Found by Database Mining: Description of the Docking Method and its Verification by a 2.3 A X-ray Structure of a Thrombin-Ligand Complex, *J. Mol. Biol.* **277**: 449-466.

Chen, H. Y., and Williams, J.D., 1985, Transferable Resistance and Aminoglycoside-Modifying Enzymes in Enterococci, *Journal of Medical Microbiology* **20**: 187-196.

Coleman, K., Athalye, M., Clancey, A., Davison, M., Payne, D.J., Perry, C.R., and Chopra, I., 1994, Bacterial resistance mechanisms at therapeutic targets, *Journal of Antimicrobial Chemotherapy* **33**: 1091-1116.

Connolly, M. L., 1983, Solvent-accessible surfaces of proteins and nucleic acids, *Science* **221** (4612): 709-713.

Courvalin, P. M., Shaw, W.V., and Jacob, A.E., 1978, Plasmid-Mediated Mechanisms of Resistance to Aminoglycoside-Aminocyclitol Antibiotics and to Chloramphenicol in Group D Streptococci, *Antimicrobial Agents and Chemotherapy* **13** (5): 716-725.

Cox, J. R., McKay, G.A., Wright, G.D., and Serpersu, E.H., 1996, Arrangement of Substrates at the Active Site of an Aminoglycoside Antibiotic 3'-Phosphotransferase As Determined by NMR, *J. Am. Chem. Soc.* **118**: 1295-1301.

Cox, J. R., and Serpersu, E. H., 1997, Biologically Important Conformations of Aminoglycoside Antibiotics Bound to an Aminoglycoside 3'-Phosphotransferase as Determined by Transferred Nuclear Overhauser Effect Spectroscopy, *Biochemistry* **36**: 2353-2359.

Cummings, M. D., Hart, T.N., and Read, R.J., 1995, Monte Carlo docking with ubiquitin, *Protein Science* **4**: 885-899.

D'Arcy, A., 1994, Crystallizing Proteins - a Rational Approach, *Acta. Cryst.* **D50**: 469-471.

Daigle, D. M., Hughes, D.W., and Wright, G.D., 1999, Prodigious substrate specificity of AAC(6')-APH(2"), an aminoglycoside antibiotic resistance determinant in enterococci and staphylococci, *Chemistry & Biology* **6**: 99-110.

Davies, J., Gilbert, W., and Gorini, L., 1964, Streptomycin, Suppression, and the Code, *PNAS* **51**: 883-890.

Davies, J. E., 1991, Aminoglycoside-Aminocyclitol Antibiotics and Their Modifying Enzymes, in *Antibiotics in Laboratory Medicine* (ed. Lorian) 691-713 (Williams & Wilkins, Baltimore).

Davies, J., 1994, Inactivation of Antibiotics and the Dissemination of Resistance Genes, *Science* **264**: 375.

Davies, J., and Wright G. D., 1997, Bacterial resistance to aminoglycoside antibiotics, *Trends Micorbiology* **5** (6): 234-240.

Davis, B. D., Chen, L., and Phang, C.T., 1986, Misread protein creates membrane channels: An essential step in the bactericidal action of aminoglycosides, *Proc. Natl. Acad. Sci. USA* **83**: 6164-6168.

Davis, B. D., 1987, Mechanism of Bactericidal Action of Aminoglycosides, *Microbiological Reviews* **51** (3): 341-350.

Davis, B. D., 1988, The lethal action of aminoglycosides, *Journal of Antimicrobial Chemotherapy* **22**: 1-3.

DiGiammarino, E. L., Draker, K.a., Wright, G.D., and Serpersu, E.H., 1998, Solution Studies of Isepamicin and Conformational Comparisons between

Isepamicin and Butirosin A When Bound to an Aminoglycoside 6'-*N*-Acetyltransferase Determined by NMR Spectroscopy, *Biochemistry* **37**: 3638-3644.

Dixon, J. S., 1997, Evaluation of CASP2 Docking Section, *PROTEINS: Structure, Function, Genetics* **Supplement 1**: 198-204.

Edson, R. S., and Terrell, C.L., 1991, The Aminoglycosides, *Mayo Clinical Proceedings* **66**: 1158-1164.

Eliopoulos, G. M., Wennersten, C. Zighelboim-Daum, S., Reiszner, E., Goldmann, D., and Moellering, R.C. Jr., 1988, High-Level Resistance to Gentamicin in Clinical Isolates of *Streptococcus (Enterococcus) faecium*, *Antimicrobial Agents and Chemotherapy* **32** (10): 1528-1532.

Ferré-D'Amaré, A. R., and Burley, S.K., 1997, Dynamic Light Scattering in Evaluating Crystallizability of Macromolecules, *Methods in Enzymology* **276**: 157-166.

Ferretti, J. J., Gilmore, K. S., and Courvalin, P., 1986, Nucleotide sequence analysis of the gene specifying the bifunctional 6'-aminoglycoside acetyltransferase 2"-aminoglycoside phosphotransferase enzyme in *Streptococcus faecalis*and identification and cloning of gene regions specifying the two activities, *Journal Bacteriology* **167**: 631-638.

Fleming, A., 1929, On the antibacterial action of cultures of a Penicillium with special reference to their use in the isolation of *H. influenzae*, *Br J Exp Pathol* **10**: 226-236.

Fourmy, D., Recht, M.I., Blanchard, S.C., and Puglisi, J.D., 1996, Structure of the A site of *Escherichia coli*16S ribosomal RNA complexed with an aminoglycoside antibiotic, *Science* **274**: 1367-1371.

Gootz, T. D., 1990, Discovery and development of new antimicrobial agents, *Clin. Microbiol. Rev.* **3** (1): 13-31.

Gshwend, D. A., Sirawaraporn, W., Santi, D.V., and Kuntz, I.D., 1997, Specificity in structure-based drug design: identification of a novel, selective inhibitor of Pneumocystis carinii dihydrofolate reductase, *Proteins: Structure, Functions, and Genetics* **29** (1): 59-67.

Gupta, R. K., Fung, C.H., and Mildvan, A.S., 1976, Chromium(III)-Adenosine Triphosphate as a Paramagnetic Probe to Determine Intersubstrate Distances on Pyruvate Kinase, *The Journal of Biological Chemistry* **251** (8): 2421-2430.

Halgren, T. A., 1996, Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94, *Journal of Computational Chemistry* **17** (5&6): 490-519.

Hart, T. N., and Read, R.J., 1992, A Multiple-Start Monte Carlo Docking Method, *Proteins: Structure, Function, and Genetics* **13**: 206-222.

Hart, T. N., Ness, S. R., and Read, R. J., 1997, Critical Evaluation of the Research Docking Program for the CASP2 Challenge, *PROTEINS: Structure, Function, Genetics* **Supplement 1**: 205-209.

Hinshaw, H. C., and Feldman, W.H., 1945, Streptomycin in Treatment of Clinical Tuberculosis: A Preliminary Report, *Proceedings of the Staff Meetings of the Mayo Clinic* **20** (18): 313-318.

Hodel-Christian, S. L., and Murray, B.E., 1991, Characterization of the gentamicin resistance transposon Tn*5281* from *Enterococcus faecalis* and comparison to the staphylococcal transposons Tn*4001* and Tn*4031*, *Antimicrob. Agnets Chemother.* **35**: 1147-1152.

Holm, L., and Sander, C., 1995, DNA Polymerase b belongs to an ancient nucleotidyltransferase superfamily, *Trends Biol. Chem.* **20**: 345-347.

Hon, W.-C., McKay, G.A., Thompson, P.R., Sweet, R.M., Yang, D.S.-C., Wright, G.D., and Berghuis, A.M., 1997, Structure of an Enzyme Required for Aminoglycoside Antibiotic Resistance Reveals Homology to Eukaryotic Protein Kinases, *Cell* **89** (6): 887-895.

Honig, B., and Nicholls, A., 1995, Classical Electrostatics in Biology and Chemistry, *Science* **268**: 1144-1149.

Hunter, T. H., 1947, Use of Streptomycin in the Treatment of Bacterial Endocarditis, *American Journal of Medicine* **2**: 436-442.

Jancarik, J., and Kim, S-H., 1991, Sparse matrix sampling: a screening method for crystallization of proteins, *Journal of Applied Crystallography* **24**: 409-411.

Jawetz, E., Gunnison, J.B., and Coleman, V.R., 1950, The Combined Action of Penicillin with Streptomycin or Chloromycetin on Enterococci *in Vitro*, *Science* **111**: 254-256.

Kapusnik-Uner, J. E., Sande, M.A., and Chambers, H.F., 1996, Antimicrobial Agents: Tetracyclines, Chloramphenicol, Erythromycin and Miscellaneous Antibacterial Agents, in *Goodman & Gilman's The Pharmacological Basis of Therapeutics* (ed. Hardman, J. G., Limbird, L.E., Molinoff, P.B., Ruddon, R.W., Gilman, A.G.) 1123-1153 (McGraw Hill, New York).

Kaufhold, A., Podbielski, A., Horaud, T., and Ferrieri, P., 1992, Identical genes confer high-level resistance to gentamicin upon *Enterococcus faecium,* and *Streptococcus agalactiae, Antimicrob. Agents Chemother.* **36**: 1215-1218.

Knighton, D. R., Zheng, J.H., TenEyck, L.F., Ashford, V.A., Xuong, N.H., Taylor, S.S., and Sowadski, J.M., 1991, Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase, *Science* **253**: 407-414.

Kraulis, P. J., 1991, MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *Journal of Applied Crystallography* **24**: 946-950.

Krogstad, D. J., Korfhagen, T.R., Moellering, R.C. Jr., Wennersten, C., Swartz, M.N., Perzynski, S., and Davies, J., 1978a, Aminoglycoside-Inactivating Enzymes in Clinical Isolates of *Streptococcus Faecalis*An Explanation for Resistance to Antibiotic Synergism, *J. Clin. Invest.* **62** (2): 480-486.

Krogstad, D. J., Korfhagen, T.R., and Moellering, R.C. Jr., 1978b, Plasmid Mediated Resistance to Antibiotic Synergism in Enterococci, *J. Clin. Invest.* **61** (6): 1645-1653.

Kunin, C. A., 1993, Resistance to Antimicrobial Drugs - A Worldwide Calamity, *Ann. Intern. Med.* **118**: 557-561.

Le Goffic, F., Moreau, N., et Chevereau, M., 1973, Purification par chromatographie d'affinité d'une acétyltranférase et d'une phosphotransférase inactivant les antibiotiques aminosidiques, *Biochimie* **55**: 1183-1186.

Lortholary, O., Tod, M., Cohen, Y., and Petitjean, O., 1995, Aminoglycosides, *Medical Clinics of North America* **79** (4): 761-787.

Mandell, G. L., and Petri, W.A.Jr., 1996a, Antimicrobial Agents: Penicillins, Cephalosporins and Other β-lactam Antibiotics, in *Goodman & Gilman's The Pharmacological Basis of Therapeutics* (ed. Hardman, J. G., Limbird, L.E., Molinoff, P.B., Ruddon, R.W., Gilman, A.G.) 1073-1101 (McGraw Hill, New York).

Mandell, G. L., and Petri, W.A.Jr., 1996b, Antimicrobial Agents: Sulfonimides, Trimethoprim-Sulfamethoxazole, Quinolones, and Agents for Urinary Tract Infections, in *Goodman & Gilman's The Pharmacological Basis of Therapeutics* (ed. Hardman, J. G., Limbird, L.E., Molinoff, P.B., Ruddon, R.W., Gilman, A.G.) 1057-1072 (McGraw-Hill, New York).

Marrone, T. J., Briggs, J.M., McCammon, J.A., 1997, Structure-Based Drug Design: Computational Advances, *Annu. Rev. Pharmacol. Toxicol.* **37**: 71-90.

Matsunaga, K., Yamaki, H., Nishimura, T., Tanaka, N., 1986, Inhibition of DNA replication initiation by aminoglycoside antibiotics, *Antimicrob. Agents Chemother.* **30**: 468-474.

McCormick, M. H., Stark, W.M., Pittenger, G.E., Pittenger, R.C., and McGuire, J.M., 1956, Vancomycin, a new antibiotic. I. Chemical and biologic properties, in *Antibiotics Annual, 1955-1956* 606-611 (Medical Encyclopedia, Inc., New York).

McKay, G. A., Thompson, P.R., and Wright, G.D., 1994, Broad Spectrum Aminoglycoside Phosphotransferase Type III from *Enterococcus*: Overexpression, Purification, and Substrate Specificity, *Biochemistry* **33**: 6936-6944.

McKay, G. A., and Wright, G.D., 1995, Kinetic Mechanism of Aminoglycoside Phosphotransferase Type IIIa, *The Journal of Biological Chemistry* **270** (42): 24686-24692.

McKay, G. A., and Wright, G.D., 1996, Catalytic Mechanism of Enterococcal Kanamycin Kinase (APH(3')-IIIa): Viscosity, Thio, and Solvent Isotope Effects Support a Theorell-Chance Mechanism, *Biochemistry* **35**: 8680-8685.

McKay, G. A., Roestamadji, J., Mobashery, S., and Wright, G.D., 1996, Recognition of Aminoglycoside Antibiotics by Enterococcal-Staphylcoccal Aminoglycoside 3'-Phosphotransferase Type IIIa: Role of Substrate Amino Groups, *Antimicrob. Agents Chemother.* **40** (11): 2648-2650.

McPherson, A., 1999,*Crystallization of Biological Macromolecules* (Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York).

McRee, D. E., 1993,*Practical Protein Crystallography* (Academic Press, Inc., San Diego).

Merritt, E. A., and Bacon, D. J., 1997, Raster3D: Photorealistic Molecular Graphics, *Methods in Enzymology* **277** (B): 505-524.

Mingeot-Leclercq, M., Glupczynski, Y., and Tulkens, P.M., 1999, Aminoglycoside: Activity and Resistance, *Antimicrobial Agents and Chemotherapy* **43** (4): 727-737.

Moazed, D., and Noller, H.F., 1987, Interaction of antibiotics with functional sites in the 16S ribosomal RNA, *Nature* **27**: 389-394.

Moellering, R. C. Jr., Wennersten, C., and Weinberg, A. N., 1971, Studies on antibiotic synergism against enterococci, *J. Lab. Clin. Med.* **77** (5): 821-828.

Moellering, R. C. Jr., Wennersten, C., and Weinstein, A.J., 1973, Penicillin-Tobramycin Synergism Against Enterococci: a Comparison with Penicillin and Gentamicin, *Antimicrobial Agents and Chemotherapy* **3** (4): 526-529.

Moellering, R. C. Jr., Korzeniowski, O.M., Sande, M.A., and Wennersten, C.B., 1979, Species-Specific Resistance to Antimicrobial Synergism in *Streptococcus faecium* and *Streptococcus faecalis*, *The Journal of Infectious Diseases* **140** (2): 203- 208.

Moellering, R. C. Jr., Murray, B.E., Schoenbaum, S.C., Adler, J., and Wennersten, C.B., 1980, A Novel Mechanism of Resistance to Penicillin-Gentamicin Synergism in *Streptococcus faecalis*, *The Journal of Infectious Diseases* **141** (1): 81-86.

Musser, J. M., 1995, Antimicrobial agent resistance in mycobacteria: molecular genetic insights, *Clin. Microbiol. Rev.* **8**: 496-514.

Nakamura, H., 1996, Roles of electrostatic interactions in proteins, *Quaterly Review of Biophysics* **29** (1): 1-90.

Neu, H. C., 1992, The Crisis in Antibiotic Resistance, *Science* **257**: 1064-1073.

Nicholls, A., and Honig, B., 1991, A Rapid Finite Difference Algorithm, Utilizing Successive Over-Relaxation to Solve the Poisson-Boltzmann Equation, *Journal of Computational Chemistry* **12** (4): 435-445.

Owen, D. J., Noble, M.E., Garman, E.F., Papageorgiou, A.C., and Johnson, L.N., 1995, Two structures of the catalytic domain of phosphorylase kinase: an active protein kinase complexed with substrate analogue and product, *Structure* **3**: 467-482.

Patterson, J. E., and Zervos, M. J., 1990, High-level gentamicin resistance in Enterococcus: Microbiology, genetic basis, and epidemiology, *Review Infectious Diseases* **12** (4): 644-652.

Pederson, L. C., Benning, Matthew M., and Holden, Hazel M., 1995, Structural Investigation of the Antibiotic and the ATP-Binding Sites in Kanamycin Nucleotidyltransferase, *Biochemistry* **34** (41): 13305-13311.

Perl, T. M., 1999, The threat of vancomycin resistance, *Am. J. Med.* **106** (5A): 26S-37S.

Price, K. E., 1986, Aminoglycoside Research 1975-1985: Prospects for Development of Improved Agents, *Antimicrobial Agents and Chemotherapy* **29** (4): 543-548.

Recht, M. I., Douthwaite, S., and Puglis, J.D., 1999, Basis for prokaryotic specificity of action of aminoglycoside antibiotics, *The EMBO Journal* **18** (11): 3133-3138.

Rhodes, G., 1993,*Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models* (Academic Press, Inc., San Diego).

Roestamadji, J., Grapsas, I., and Mobashery, S., 1995, Loss of individual electrostatic interactions between aminoglycoside antibiotics and resistance enzymes as an effective means to overcoming bacterial drug resistance, *J. Am. Chem. Soc.* **117**: 11060-11069.

Rouch, D. A., Byrne, M. E., Kong, Y. C., and Skurray, R. A., 1987, The aacA-aphD gentamicin and kanamycin resistance determinant of Tn*4001* from *Staphylococcus aureus*: Expression and nucleotide sequence analysis, *J. Gen. Microbiol.* **133**: 3039-3052.

Schaberg, D. R., Culver, D.H., and Gaynes, R.P., 1991, Major Trends in the Microbial Etiology of Nosocomial Infection, *The American Journal of Medicine* **91** (Supplement 3B): 72S-75S.

Schatz, A., Bugie, E., and Waksman, S.A., 1944, Streptomycin, a substance exibiting antibiotic activity against Gram-positive and Gram-negative bacteria, *Proc. Soc. Exp. Biol. Med.* **55**: 66-69.

Schnecke, V., Swanson, C.A., Getzoff, E.D., Tainer, J.A., and Kuhn, L.A., 1998, Screening a peptidyl database for potential ligands to proteins with side-chain flexibility, *Proteins: Structure, Function, and Genetics* **33** (1): 74-87.

Sharp, K., Fine, R., and Honig, B., 1987, Computer simulations of the diffusion of a substrate to an active site of an enzyme, *Science* **236**: 1460-1463.

Shaw, K. J., Rather, P.N., Hare, R.S., and Miller, G.H., 1993, Molecular Genetics of Aminoglycoside Resistance Genes and Familial Relationships of the Aminoglycoside-Modifying Enzymes, *Microbiological Reviews* **57** (1): 138-163.

Shoichet, B. K., Stroud, R.M., Santi, D.V., Kuntz, I.D., and Perry, K.M., 1993, Structure-Based Discovery of Inhibitors of Thymidylate Synthase, *Science* **259**: 1445-1450.

Standiford, H. D., de Maine, J.B., and Kirby, W.M.M., 1970, Antibiotic Synergism of Enterococci, *Arch Intern Med* **126**: 255-259.

Tanaka, N., Matsunaga, K., Yamaki, H., and Nishimura, T., 1984, Inhibition of initiation of DNA synthesis by aminoglycoside antibiotics, *Biochem. Biophys. Res. Commun.* **122**: 460-465.

Thompson, P. R., Hughes, D.W., and Wright, G.D., 1996a, Regiospecificity of Aminoglycoside Phosphotransferase from Enterocooci and Staphylococci (APH(3')-IIIa), *Biochemistry* **35**: 8686-8695.

Thompson, P. R., Hughes, D. W., and Wright, G. D., 1996b, Mechanism of aminoglycoside 3'-phosphotransferase type IIIa: His 188 is not a phosphate-accepting residue, *Chemistry & Biology* **3**: 747-755.

Thompson, P. R., Schwartzenhauer, J., Hughes, D.W., Berghuis, A.M., and Wright, G.D., 1999, The C-terminus of aminoglycoside phosphotransferase (3')-IIIa is critical for antibiotic recognition and resistance, *J. Biol. Chem.* **274** (43): 30697-30721.

Waksman, S. A., and Lechevalier, H.A., 1949, Neomycin, a new antibiotic active against streptomycin-resistant bacteria, including tuberculosis organisms, *Science* **109**: 305-307.

Weiner, S. J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., and Weiner, P., 1984, A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins, *J. Amer. Chem. Soc.* **106**: 765-784.

Weinstein, A. J., and Moellering, R.C. Jr., 1973, Penicillin and Gentamicin Therapy For Enterococcal Infections, *JAMA* **223** (9): 1030-1032.

Wennersten, C. B., and Moellering, R.C. Jr., 1980, Mechanism of Resistance to Penicillin-Aminoglycoside Synergism in *Streptococcus faecium*, in *Current Chemotherapy and Infectious Diseases* (ed. Nelson, J. D., Grassi, C.) 710-712 (American Society for Microbiology, Washington D.C.).

Williams, J. W., and Northrop, D.B., 1978, Kinetic Mechanisms of Gentamicin Acetyltransferase I, *The Journal of Biological Chemistry* **253** (17): 5902-5907.

Wise, R., 1999, A review of the mechanisms of action and resistance of antimicrobial agents, *Canadian Respiratory Journal* **6** (Suppl. A): 20A-22A.

Wolf, E., Vassilev, A., Makino, Y., Sali, Andrej, N., Yoshihiro, and Burley, S. K., 1998, Crystal Structure of a GCN5-Related N-acetyltransferase: *Serratia marcescens* Aminoglycoside 3-N-acetyltransferase, *Cell* **94**: 439-449.

Woodcock, J., Moazed, D., Cannon, M., Davies, J., and Noller, H.F., 1991, Interaction of antibiotics with A- and P-site-specific bases in 16S ribosomal RNA, *The EMBO Journal* **10** (10): 3099-3103.

Wright, G. D., Berghuis, A. M., and Mobashery, S., 1999, Aminoglycoside Antibiotics: Structures, Functions, and Resistance, in *Resolving the Antibiotic Paradox: Progress in Drug Design and Resistance* (ed. Rosen, B. P., Mobashery, S.) 27-69 (Plenum Publishing Corporation, New York).

Wright, G. D., and Thompson, P.R., 1999, Aminoglycoside Phosphotransferases: proteins, structure, and mechanism, *Front. Biosci.* **4**: D9-21.

Wybenga-Groot, L. E., Draker, K.a., Wright, G.D., and Berghuis, A.M., 1999, Crystal structure of an aminoglycoside 6'-*N*-acetyltransferase: defining the GCN5-related *N*-acetyltransferase superfamily fold, *Structure* **7**: 497-507.

Zervos, M. J., Kauffman, C.A., Therasse, P.M., Bergman, A.G., Mikesell, T.S., and Schaberg, D.R., 1987, Nosocomial Infection by Gentamicin-Resistant *Streptococcus faecalis*, *Annals of Internal Medicine* **106**: 687-691.

Zimmermann, R. A., Moellering, R.C. Jr., and Weinberg, A.N., 1971, Mechanism of Resistance to Antibiotic Synergism in Enterococci, *Journal of Bacteriology* **105** (3): 873-879.

Zulauf, M., and D'Arcy, A., 1992, Light Scattering of proteins as criterion for crystallization, *Journal of Crystal Growth* **122**: 102-106.