Synthetic, Spectrometric and Computer Modelling Studies of Novel ATP Analogues

Thesis

Submitted in Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Of Rhodes University

By

Babalwa Siliziwe Blossom Gxoyiya

M.Sc (Rhodes University)

October 2007

ABSTRACT

This study has been concerned with the design and synthesis of ATP analogues with the potential to act as inhibitors of glutamine synthetase – a novel target for therapeutic intervention in the treatment of tuberculosis. Using a structural-analogy approach, various 3-indolylalkanoic acid, benzimidazole and pyrazolo[3,4-*d*]pyrimidine derivatives have been prepared and characterized.

Alkylation of the heterocyclic bases using 4-(bromomethyl)-2,2-dimethyl-1,3-dioxolane, 2-(bromomethoxy)ethyl acetate and 2-(chloroethoxy)ethanol in the presence of either NaH or Bu^tOK afforded the corresponding *N*-alkylated derivatives of benzimidazole and 4-aminopyrazolo[3,4-*d*]pyrimidine (4-APP). Similar reactions with 3-indolylalkanoic esters resulted in *O*-alkyl cleavage with the formation of new esters. Alkylation of benzimidazole with allyl bromide, 4-bromobutene and 2-methylbut-2-ene has also been shown to afford the corresponding 1-alkenylbenzimidazoles in moderate to excellent yield (43-96%). Subsequent oxidation of these products using CTAP, gave the dihydroxy derivatives in poor to good yields (26-77%). Phosphorylation of various hydroxy derivatives of benzimidazole and 4-APP has been achieved using diethyl chlorophosphate to afford the corresponding monophosphate and 1,2-diphosphate esters.

Glycosylation of each of the heterocyclic bases has been successfully achieved using 1,2,3,4,6-penta-*O*-acetyl-D-glucopyranose and $SnCl_4$ in acetonitrile, while methanolysis of the resulting tetraacetates, using methanolic NaOMe, afforded the hydroxy derivatives in good yield (50-70%).

Various 1- and 2-dimensional NMR spectroscopic methods (*e.g.*, ^1H, ^{13}C, ^{31}P, COSY, HSQC and HMBC) have been used to confirm the structures of the synthesized ATP analogues. The application of NMR prediction programmes has been explored, permitting assessment of their agreement with the experimental data and confirmation of assigned structures. High-resolution electron impact (EI) mass spectrometric data have

been used to explore the mass fragmentation pathways exhibited by selected derivatives, and certain common fragmentations have been identified.

Molecular modelling of selected products as potential glutamine synthetase ligands has been performed on the Accelrys Cerius^2 platform, and interactive receptor-ligand docking studies have been conducted using the Ligand-Fit module. These studies have revealed possible hydrogen-boding interactions between the selected analogues and various amino acid residues in the glutamine synthetase active site.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my heavenly father, through him every thing is possible.

I would like to express my gratitude to my supervisor Prof. P. T. Kaye for his support, patience, encouragement, and teachings and for allowing me to conduct this research under his supervision.

Thanks to Tommie van de Walt and Martin Brits of the University of Witwatersrand mass spectrometric unit and Mr Aubrey Sonemann for high- and low-resolution MS analysis. Many thanks to Dr Irvy Gledhill of CSIR for providing the glutamine synthetase enzyme, and Kevin Lobb, Dokes Ntebe and Vuyisile Dondashe for their contribution.

To my darling daughter, Buhlebuvelile Gxoyiya, thank you baby doll for the love, laughter, hugs and kisses. You gave me something to look forward to after the hard days work.

To my mother, Mrs Nyameka Gxoyiya, ndiswele imilomo eyabaliwa Bhayi, Khetsha, Camsholo, Gobhozi, nkonjan`emnyama ebhabhemafini. You are the wind beneath my wings, without your support and prayers I wouldn`t have made it this far.

To my cousin Rev. Siyabonga Sibeko, ndiyabulela Nxuba, thole lo Mthwakazi for all what you have done for me to make sure I achieve my goals including the reverends Ntuthuzelo Lutya and Deborah Meyer and their families for looking after my angel. I would also like to thank my family and friends at large for the encouragement and support especially Dudu, Lisa, my aunt Mrs Zodi Sibeko, ndiyabulela Bhelekazi, and my cousin Nosipho, ndiyabulela Rhudulu.

Finally I would like to thank DAAD, CSIR and Rhodes University for the financial support.

LIST OF ABREVIATIONS

APCI	Atmospheric pressure chemical ionisation
BuLi	Butyllithium
^{t}BuOH	tert-Butyl alcohol
CBz	Carbobenzoxy
COSY	Correlation spectroscopy (^{1}H-^{1}H)
DMF	Dimethyl formamide
DMSO	Dimethyl sulfoxide
EDCI	1-Ethyl-3(3-dimethylaminopropyl)cabodiimide
HPLC	High perfomance liquid cromatography
HOAt	1-Hydroxy-7-azabenzotriazole
HMBC	Heteronuclear multiple bond coherence (^{1}H-^{13}C)
HSQC	Heteronuclear single quantum coherence (^{1}H-^{13}C)
PTSA	para-Toluenesulfonic acid
TEAB	Triethylammonium bicarbonate
TEMPO	2,2,6,6-Tetramethyl-1-piperidinyloxy
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran

CONTENTS

Page

1. Introduction

Tuberculosis (TB) is a very old disease, but the causative factors were only identified at the turn of the 19th century.[1,2] While the disease was almost eradicated by the introduction of improved living standards and effective TB chemotherapy, it has now re-emerged as a serious public health problem worldwide. This is due to the increase in multiple-drug-resistant (MDR) TB, which is defined as resistance to isoniazid and rifampicin, and to the association between HIV and *Mycobacterium tuberculosis* (*M.Tb*) infection.[3,4] The main reasons for the increasing burden of TB globally are:- poverty and the widening gap between rich and poor in various populations; neglect of the disease (inadequate case detection, diagnosis and cure); collapse of the health infrastructure in countries experiencing severe economic crisis or civil unrest; and the impact of the HIV pandemic.[4] Recently there has been the complicating development of extensive or extreme drug resistant (XDR) TB, which is MDR-TB that is also resistant to three or more of the six classes of second-line drugs.[5]

1.1 Background to Tuberculosis

Mycobacteria are free-living organisms and are omnipresent in the environment. They can be found in soil, salt and fresh water, algae and grasses, and are potentially pathogenic to many animals, including cattle, pigs, fish and reptiles.[6] Tuberculosis is a chronic disease and is caused by the bacterium, *Mycobacterium tuberculosis*. The bacterium is characterized as an acid-fast bacillus due to its highly hydrophobic cell wall (resistant to alcohol, acid, alkali and some disinfectants) that cannot be decolourised by acid after staining with a dye.[7] The cell wall is composed of polysaccharides, protein, lipids and glycolipids, and the most important components are mycolyl-arabinogalactans (mAG), peptidoglycans, glycolipids and lipoarabinomannan (LAM).[8-11] The existence of mycobacteria is dependent on the ability of the organism to produce an intact cell wall.[8,9] The disease is spread by airborne droplet nuclei primarily through the respiratory route by coughing, sneezing and talking, and it generally affects the lungs resulting in pulmonary TB. Alternatively, it may escape through the bloodstream and lymph nodes

and affect other parts of the body, *e.g.*, the brain, skin, bones and eyes, resulting in extra-pulmonary TB. The most common type is pulmonary TB, which is associated with about 80-85% of TB cases.[7] After infection with *M.Tb*, an individual can have the disease in one or all of the following forms:- primary, latent and post-primary. Primary TB is an acute and infectious disease, which occurs shortly after an individual has been infected by *M. Tb*.[12] Latent TB is a clinical syndrome that occurs after an individual has been exposed to *M.Tb*, the infection has been established and an immune response has been generated to control the bacteria and force it into a dormant state.[13] Unlike patients with "active" TB, individuals with latent TB do not transmit the disease and do not pose a public health risk.[12] Post-primary TB is an active disease, which results from the reactivation of latent TB due to changes in the individual's immune system through, for example, malnutrition, HIV infection, immunodepressant therapy or from re-infection with *M.Tb*.[12]

1.2 History of TB Chemotherapy

TB chemotherapy was begun in 1944 with the discovery, in 1943, of streptomycin (SM) from the soil bacterium *Streptomyces griseus,* by Waksman and co-workers,[14] The drug showed efficient inhibition of *Mycobacterium tuberculosis* with low toxicity. In 1946, Lehmann[15] introduced the use of *p*-aminosalicylic acid (PAS) based on the observation that salicylic acid stimulates the use of oxygen in tubercle bacilli.[16] Tubercle bacilli are aerobic organisms and result in dormant bacilli in an anaerobic environment. Thioacetazone, the first sulfur-containing anti-TB compound, was introduced in 1950, although sulfonamide compounds were already in use for the treatment of other infectious diseases. This was followed, in 1952, by the discovery of the two citically important first-line drugs isoniazid (INH) [17-19] and pyrazinamide (PZA).[20] These drugs were developed following the observation that nicotinamide showed weak activity against tubercle bacilli. Although INH and PZA have similar structures, they have different mechanisms of action. In 1955 the Merck Research Laboratories[21] isolated cycloserine from metabolic products of *Streptomyces orchidaceus*. Various aminoglycosides, such as kanamycin and amikacin were introduced in 1957 followed, in

1962, by ethambutol, a synthetic drug developed by the Lederle laboratories.[22] Rifampin (Rifamycin), produced by *Streptomyces mediterranei* was discovered in 1963, and was introduced into TB chemotherapy in 1965.[23,24]

The second generation of sulfur-containing antibiotics was marked by the introduction of the synthetic drug, ethionamide, in 1966,[24] and another aminoglycoside, capreomycin, isolated from *streptomyces capreolus*,[25] in 1967. It was observed during the early years of chemotherapy that treatment with a single drug led to the growth of drug-resistant bacilli, but treatment with a combination of two or more drugs greatly reduced the emergence of drug resistant strains. In the late eighties another class of antibiotics having different structures from the existing class of antibiotics was introduced. These compounds contain a quinolone backbone with an attached fluorine atom and are widely known as the fluoroquinolone antibiotics, members of which include ciprofloxacin introduced in 1987, ofloxacin introduced in 1990, levofloxacin introduced in 1996 and moxifloxacin introduced in 1999.[26] The chemistry of important compounds used in TB chemistry will be reviewed in Section 1.3.

1.2.1 Current TB Treatments and the Associated Problems

It is well known that the TB infection can remain dormant for long period periods of time, and the main objectives in treating the disease are to ensure cure without relapse and to prevent transmission and the development of acquired resistance.[3,4,27,28] The treatment typically involves a six-month regime administered in two phases. In the first phase, termed the "initial phase," a combination of four drugs, Rifamycin (Rif), Ethambutol (EMB) or Streptomycin (SM), Pyrazinamide (PZA) and Isoniazid (INH), are taken twice daily for two months, while in the second phase, termed the "continuation phase," two drugs PZA and Rif are taken for four months. This method of treatment is administered under the Daily Observed Treatment Short-course (DOTS) programme, which was approved by the World Health Organisation (WHO) in 1993, and the cure rate is 95%.[4] However, due to the length of the treatment and the number of tablets to be swallowed, patient compliance is a problem because after two weeks of treatment the

patients may feel well and then default. This problem, combined with the use of wrong drug combinations, results in the development of drug resistant bacilli, which are much harder to treat and require longer periods of medication with much more expensive and highly toxic drugs.[3,4,27-29]

1.3 Anti-TB Drugs

The discovery of broad-spectrum aminoglycoside antibiotics with antitubercular activity from the *Streptomyces* species constituted a break-through in the chemotherapeutic treatment of tuberculosis and was followed by the development of various totally synthetic drugs. The drugs can be classified according to their level of toxicity and bacteriocidal activity, *viz.*, first-line or essential drugs and second-line or reserve drugs.[4]

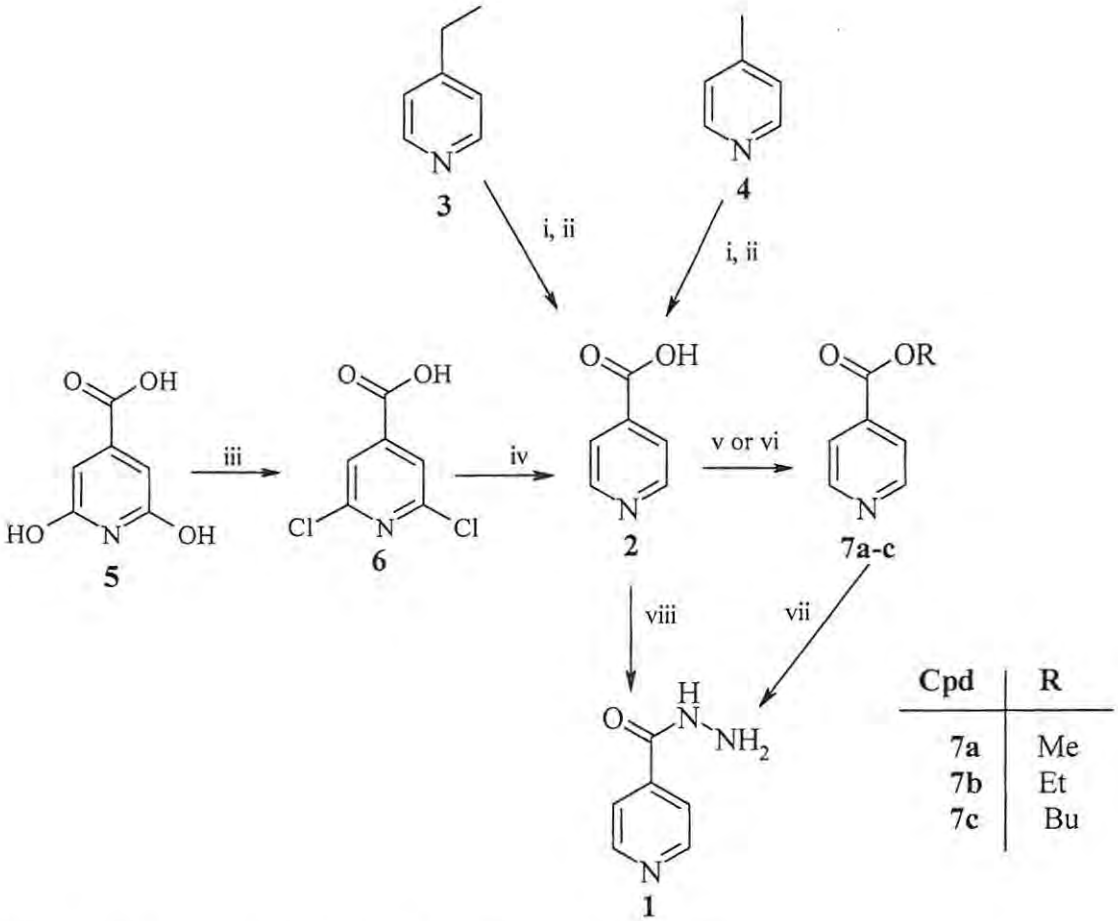
1.3.1 First-line Drugs

These drugs are given as the first course of action in the treatment of TB. They have lower levels of toxicity and are highly bacteriocidal. They act, as is the case with isoniazid, pyrazinamide and ethambutol, by inhibiting cell wall biosynthesis and, as with rifampicin and streptomycin, by intercepting other intracellular targets.[8]

1.3.1.1 Isoniazid

Isoniazid (INH) **1** is the most active of all the antitubercular drugs and is totally synthetic. Its spectrum of activity is limited to and selective for *M.Tb* infections and it has bacteriocidal effect against active infections.[4,27] It is a pro-drug that requires activation by the bacterial catalase-peroxidase (*KatG*) enzyme to generate a range of reactive organic radicals,[3,7,11,30] which then attack multiple targets in the bacteria. The most well known targets are the cell wall mycolic acid synthesis pathway and cellular constituents such as DNA, carbohydrates and lipids.[3,7,11]

Isoniazid **1** is an isonicotinic acid derivative and can be synthesized on a large scale by hydrazinolysis of the esters of isonicotinic acid **2**, which may be prepared in turn by oxidation of 4-ethylpyridine **3** or 4-methylpyridine **4** with potassium permanganate (**Scheme 1**).[31,32] The acid **2** can also be prepared from citrazinic acid **5**,[33] *via* chlorination with $POCl_3$ to form 2,6-dichloroisonicotinic acid **6**, followed by catalytic hydrogenation with Raney Nickel. The acid **2** is then protected as the methyl,[34] ethyl[31,33,35] or butyl[36] ester **7** and then hydrazonylised with hydrazine hydrate to the hydrazide **1**.



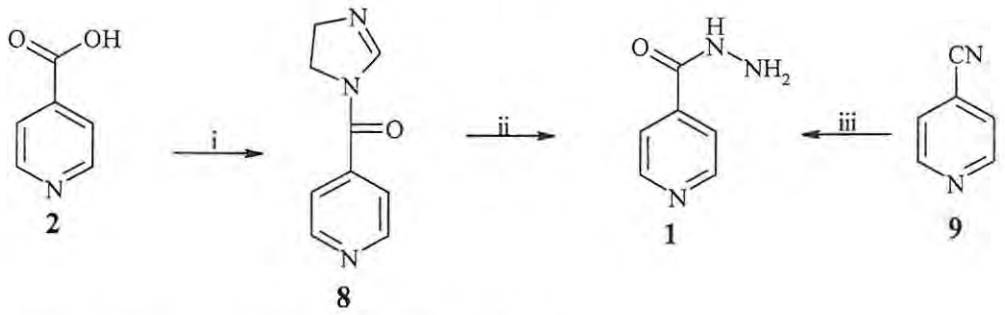
Cpd	R
7a	Me
7b	Et
7c	Bu

Scheme 1: Preparation of INH **1** from isonicotinic acid **2**.

Reagents: i) $KMnO_4$, KOH, H_2O; ii) EtOH, HCl then H_2O; iii) $POCl_3$; iv) Raney Ni; v) $SOCl_2$ and EtOH, MeOH; vi) exchange resin (KU-2-8) in the H^+ form, BuOH; vii) $N_2H_4.H_2O$; viii) dioxane, $ClCO_3Et$ and $N_2H_4.H_2O$.

The acid **2** can also be protected as the imidazolide **8** (**Scheme 2**), by treating with *N,N'*-carbonyldiimidazole (CDI), which is then hydrazinolysed to form the hydrazide **1**.[37] INH

1 can also be synthesized by treating 4-cyanopyridine **9** with hydrazine hydrate in the presence of sodium hydroxide at 100°C (**Scheme 2**).[38]



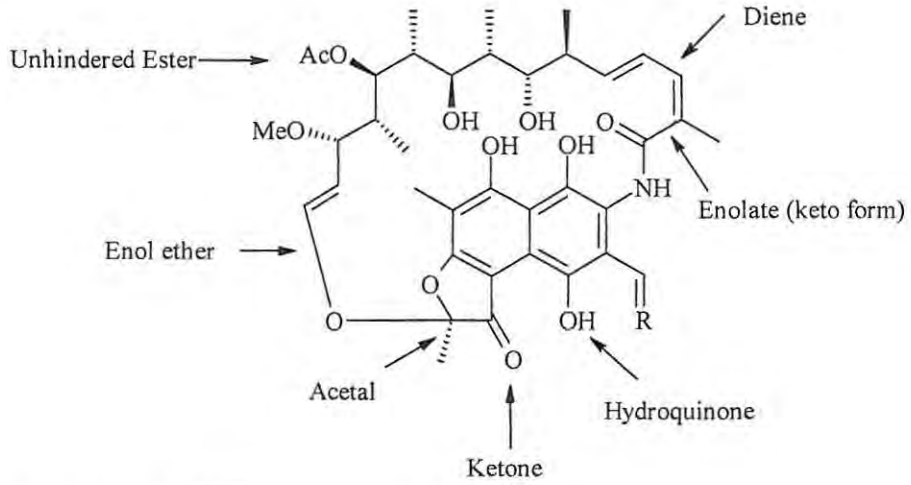
Scheme 2: Synthesis of INH **1** from **8** or **9**.

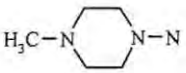
Reagents: i) CDI, THF; ii) $N_2H_4.H_2O$; iii) KOH, $N_2H_4.H_2O$, 100°C, 7h.

1.3.1.2 Rifampicin

Rifampicin **10** is a semi-synthetic drug, which is highly bacteriocidal and second only to isoniazid for the treatment of infections. It is a complex macrocyclic compound with several functional groups, including ester, acetal, ketone, diene, hydroquinone and ether moieties (**Figure 1**). It was discovered amongst a number of synthetic derivatives of rifamycin B, a broad-spectrum antibiotic with antitubercular activity, from *Streptomyces mediterranei*. It is active against semi-dormant and active forms of *M.Tb*, has a potent sterilizing effect[4,27,30] and acts by inhibiting RNA polymerase and thus disrupting mRNA synthesis.[3,7,30]

Rifampicin **10** is generally[39] synthesized by condensation of 3-formylrifamycin SV **11** and 1-amino-4-methylpiperazine in a haloaliphatic solvent, *e.g.* chloroform. Recently Wessjohann and co-workers[40] have reported the synthesis of rifampicin employing solid-phase organic chemistry (SPOC) methodology and using $PEGA_{1900}$ **12** modified as benzylhydrazide PEGA **13** (Figure 2) as the solid support. 3-Formylrifamycin SV is condensed on the hydrazide polymer to form the imine **14** which, on treatment with 1-amino-4-methylpiperazine in THF and conc. HCl, results in the formation of rifampicin **10** *via* transimination and cleavage from the solid support.



10: **R** = 

11: **R** = O

Figure 1: Structure of 3-formylrifamycin SV **11** and Rifampicin **10** showing the main functional groups.[40]

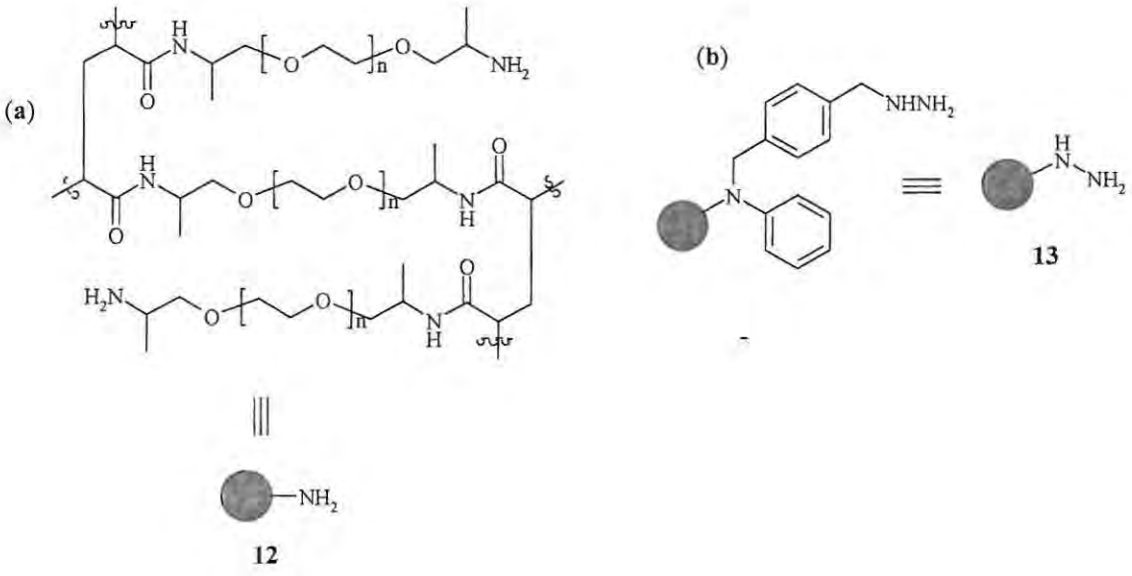
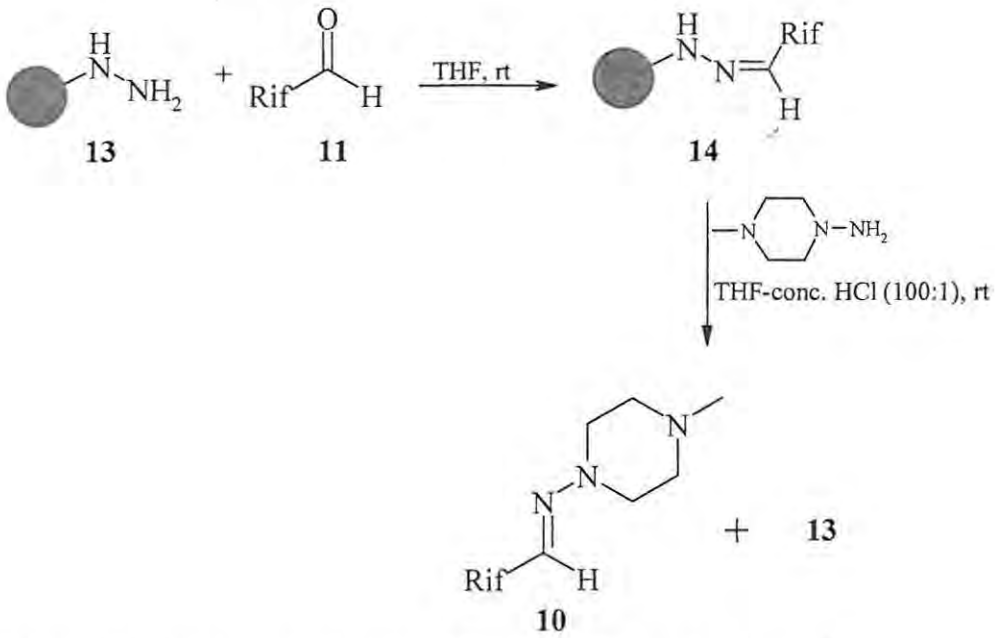


Figure 2: a) $PEGA_{1900}$ ($n = 43$) **12** and b) benzylhydrazyl $PEGA_{1900}$ **13**.[40]



13 **11** **14** **10**

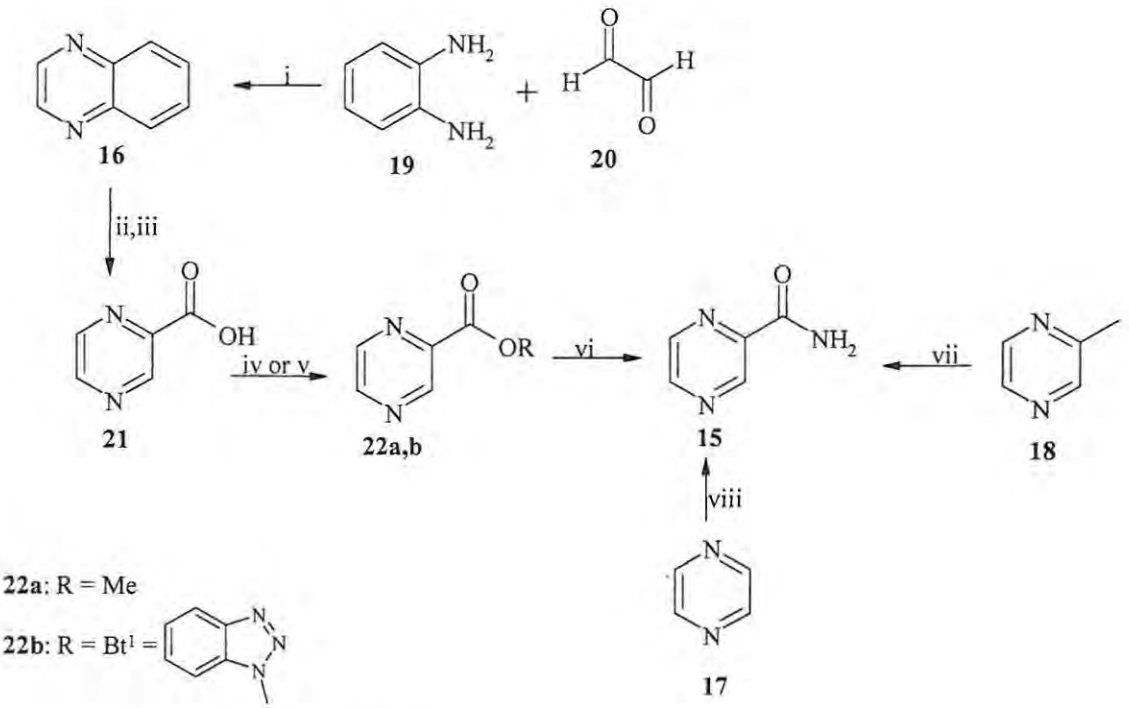
Scheme 3: Synthesis of rifampicin **10** using SPOC methodology.[40]

1.3.1.3 Pyrazinamide

Pyrazinamide (PZA) is a totally synthetic drug with a spectrum of activity limited to *M.Tb* infections. It is active against latent forms of TB and is responsible for the reduction of treatment time from nine months to six months.[28,30] It is an analogue of nicotinamide and is a pro-drug that requires activation by the Pzase/nicotinamidase enzyme to generate its active form, pyrazinoic acid (POA).[3,7,11,30] It acts in an acidic environment by disrupting membrane function and energy metabolism, and this is believed to occur by the inhibition of fatty acid synthase.[3,7,11,30]

PZA **15** is prepared by oxidation methods, starting from quinoxaline **16**, pyrazine **17** or 2-methylpyrazine **18** (**Scheme 4**). Quinoxaline **16**, obtained from the condensation of phenylenediamine **19** with glyoxal **20**, is oxidised and monodecarboxylated to form pyrazinecarboxylic acid **21**;[41,42] esterification,[41-44] followed by ammonolysis with ammonium hydroxide finally affords pyrazinamide **15**.[41-44] Bondareva and co-workers[45] have synthesized PZA **15** by ammoxidation of 2-methylpyrazine **18** in a mixture of ammonia, oxygen, steam and nitrogen using a P_2O_5 modified vanadium–titanium catalyst, in reported yields of 80-85% at 250^{O}C. Mnisci and co-workers[46] have also

employed carbamoylation to prepare PZA **15**. Thus, pyrazine **17** was reacted with carbamoyl radicals formed by the oxidation of formamide with either $FeSO_4$ or cerium(IV) ammonium nitrate-*N*-hydroxypthalimde [Ce(IV)-NHPI].



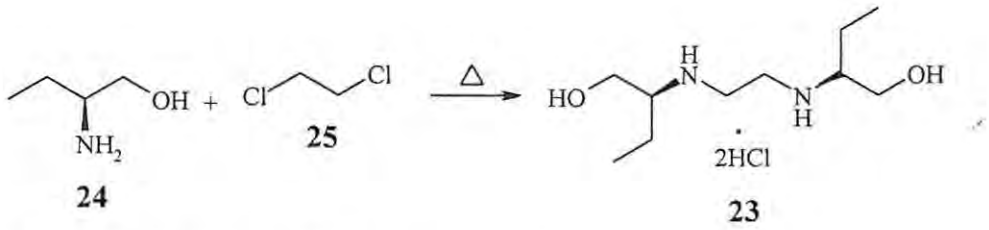
Scheme 4: Synthesis of PZA **15**

Reagents: i) EtOH; ii) $KMnO_4$, KOH; iii) 210°C *in vacuo*; iv) MeOH, HCl; v) $Bt^1SO_2CH_3$, Et_3N; vi) NH_4OH; vii) NH_3, P_2O_5-modified V-Ti catalyst; viii) $HCONH_2$, Ce(IV)-NHPI or H_2O_2 or $FeSO_4$.

1.3.1.4 Ethambutol

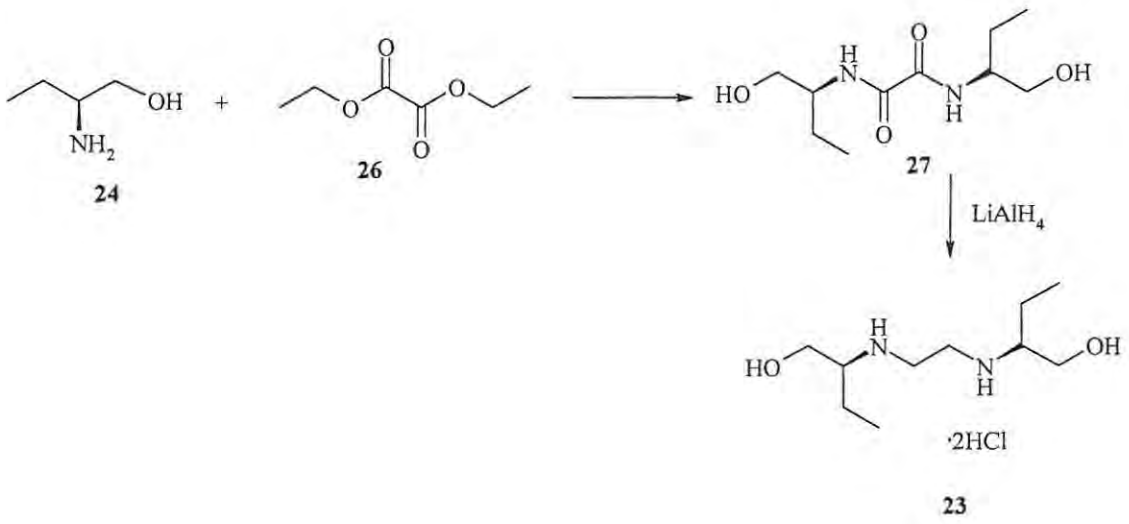
The totally synthetic antibiotic, ethambutol, was first prepared in the Lederle laboratories.[22] Its spectrum of activity is selective and limited to *MTb* infection and it is effective against active TB.[4] It acts on the cell wall by inhibiting arabinan biosynthesis and is also useful in preventing the development of resistant bacilli.[3,7,8,11,47]

Ethambutol (EMB) has been prepared as the dihydrochloride salt **23** by Wilkinson[22] by alkylating (+)-2-aminobutanol **24** with ethylene dichloride **25** (**Scheme 5**).



Scheme 5: Alkylation of 2-aminobutanol.[22]

Kritsyn[48] has also synthesized EMB as the dihydrochloride salt from 2-aminobutanol **24** and diethyl oxalate **26**, reducing the intermediate oxalamide **27** with $LiAlH_4$ (**Scheme 6**).



Scheme 6: Synthesis of EMB by Kritsyn.[48]

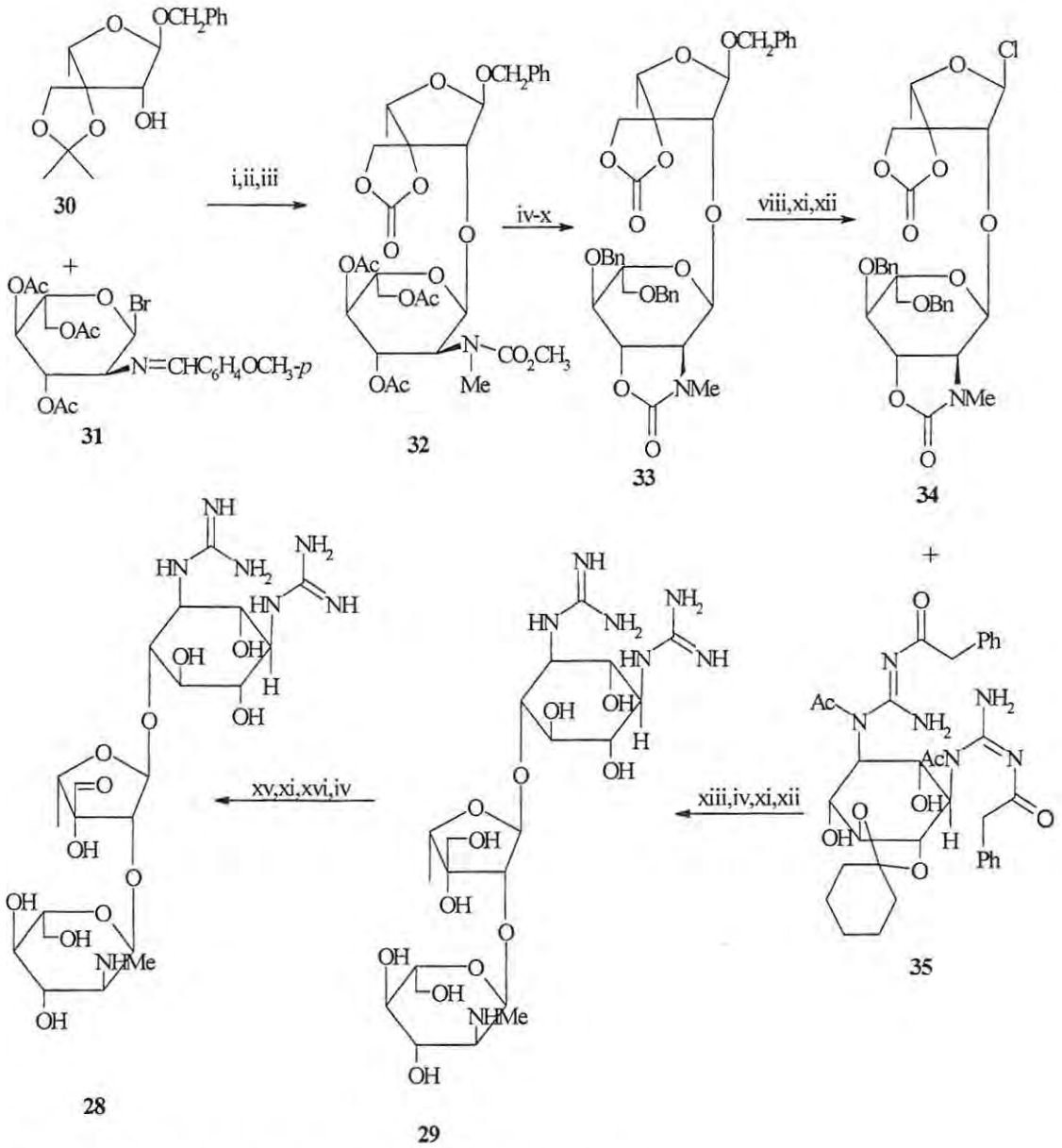
1.3.1.5 Streptomycin

The aminoglycoside, streptomycin (SM) **28**, is considered to be the most important broad-spectrum antibiotic after penicilin. It is effective against actively growing bacteria and it acts by binding on the ribosomal S12 protein and 16S rRNA thereby inhibiting protein synthesis.[3,7]

The aminoglycoside can be synthesized from dihydrostreptomycin (DSM) **29** which was previously obtained by fermentation or by hydrogenation of streptomycin **28**. The total synthesis of DSM and SM was reported by Umezawa and co-workers in 1974.[49,50] The

synthesis starts from the condensation of the isopropylidene derivative of benzyl α-L-dihydrostreptoside **30** with the L-glucosamine derivative **31** using mecuric cyanide followed by:- hydrolysis of the Schiff base with aqueous acetic acid; acylation with methyl chloroformate; and *N*-methylation with methyl iodide to form the disaccharide **32**. This is followed by several de-protection and re-protection steps to afford the derivative **33**. Hydrogenation followed by chlorination affords the chloro derivative **34** which, on condensation with the streptidine derivative **35**, leads *via* hydrolysis and hydrogenation to DSM **29**.

SM **28** is synthesized in several steps from DSM **29**. The first step involves protection of the hydroxyl groups as isopropylidine or acetate ester derivatives. Selective hydrolysis of the isopropylidine followed by Pfitzner-Moffatt oxidation and deacetylation then affords SM **28**.



Scheme 7: Total synthesis of DSM **29** and SM **28**

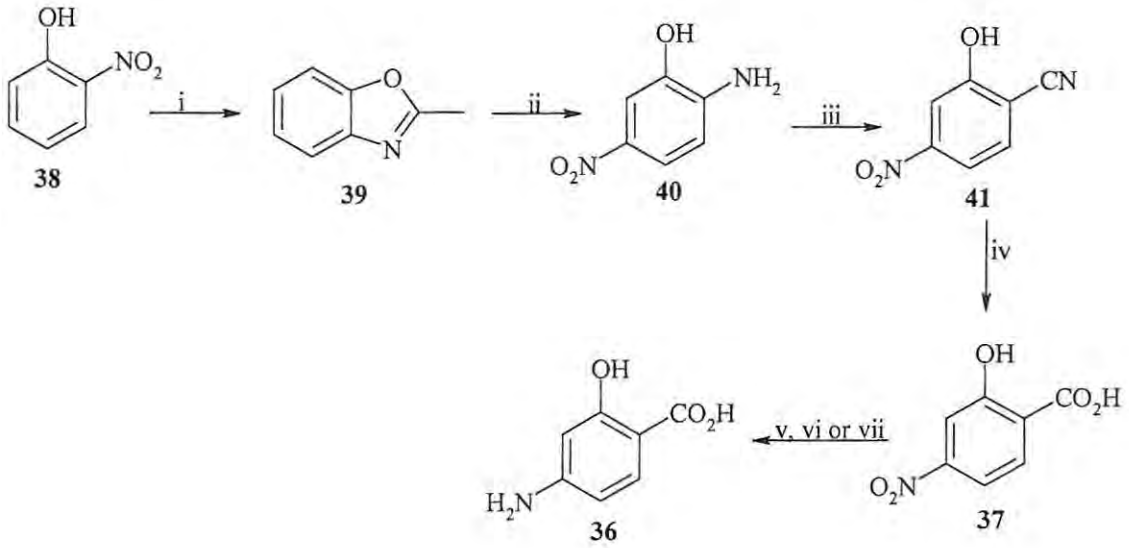
Reagents: i) HgCN, C_6H_6; ii) 50% aq. AcOH, $ClCO_2CH_3$; iii) MeI, Ag_2O, DMF; iv) NaOMe, MeOH; v) 1M-HCl, MeOH; vi) 10% $Ba(OH)_2$; vii) PTSA, acetone, 5Å molecular sieves; viii) NaOH, acetone; ix) 25% aq. AcOH, MeOH; x) PhCOCl, pyridine; xi) 75% aq. AcOH; xii) Pd, dioxane, $SOCl_2$; xiii) Ag_2CO_3-$AgClO_4$, 3Å molecular sieves, benzene; xiv) 50% aq. AcOH; xv) Ac_2O, PTSA; xvi) DMSO, DCC, TFA, pyridine.

1.3.2 Second-line Drugs

These drugs are given to patients who have relapsed and have developed bacterial strains that are resistant to the front-line drugs, such as isoniazid and rifampicin. The second-line drugs are highly toxic and have poor bacteriocidal activity.

1.3.2.1 p-*Aminosalicylic acid*

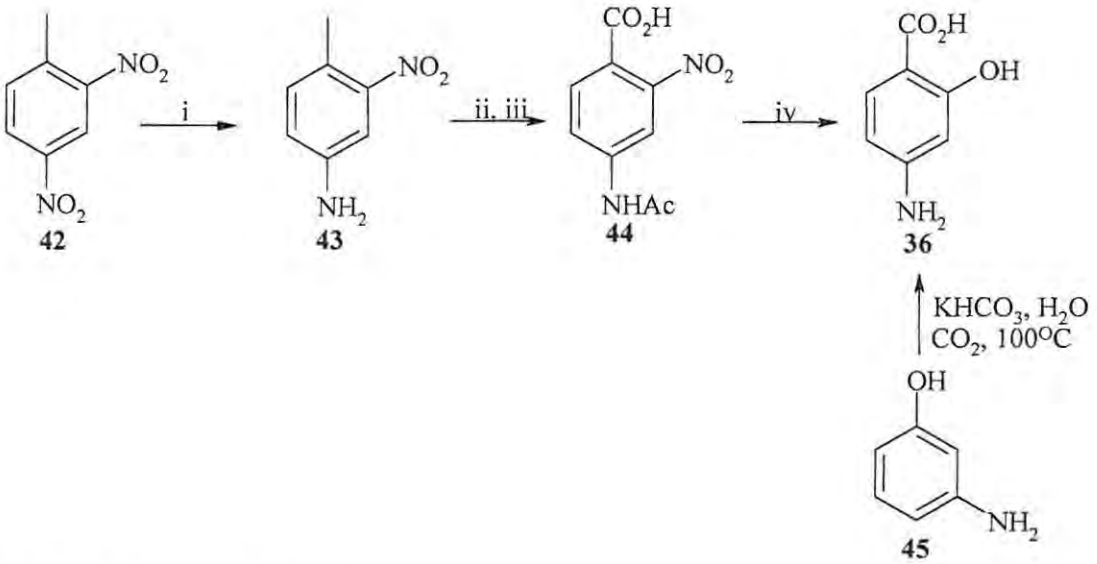
Following the observation that benzoic acid and salicylic acid stimulate the use of oxygen in tubercle bacilli,[16] Lehmann investigated various derivatives of benzoic acid and found that *p*-aminosalicylic acid (PAS) was the most active, with 50-75% growth inhibition at a concentration of 0.15mg/100ml.[15] PAS **36** can be synthesized by the reduction of 4-nitrosalicylic acid **37**, which can be synthesized, in turn, in several steps from 2-nitrophenol **38**, as illustrated in **Scheme 8**.[51] However, this method is not suitable for large scale production and alternative methods were investigated.[52-55]



Scheme 8: Synthesis of PAS

Reagents: i) Raney Ni, H_2, Ac_2O; ii) HNO_3, H_2SO_4; iii) $CuSO_4$, KCN; iv H_2SO_4, H_2O; v) H_2, PtO; vi) NaOH, Raney Ni, H_2; vii) NH_4OH, $FeSO_4$.

PAS **36** can also be synthesized from 2,4-dinitrotoluene **42** by partial reduction to 2-nitro-*p*-toluidine **43**, which on acetylation and oxidation forms the intermediate **44**. Subsequent reduction affords PAS **36** (**Scheme 9**).[52] The most successful method, however, has been the direct carboxylation of 3-aminophenol **45** as illustrated in **Scheme 9**.[53-55]



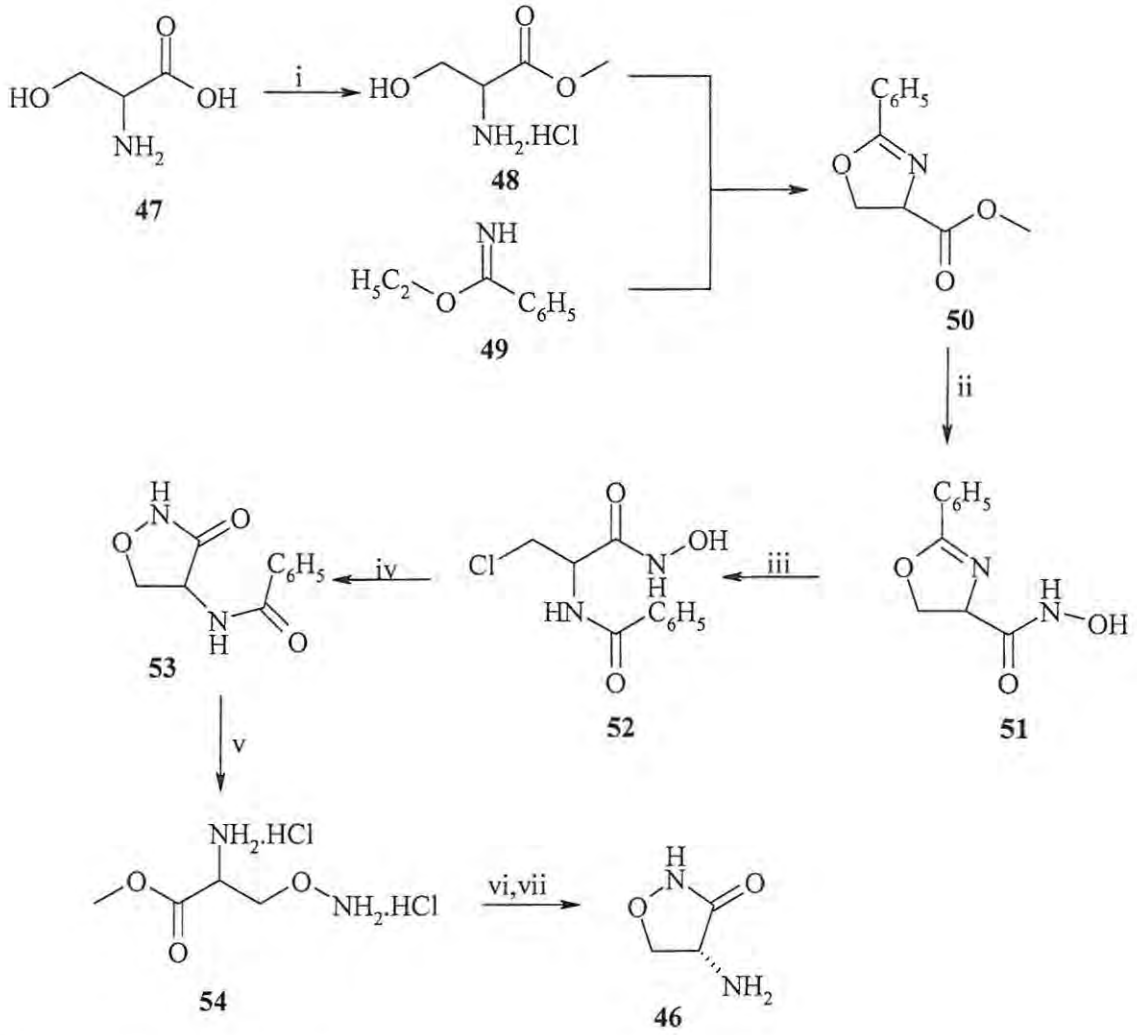
Scheme 9: Synthesis of PAS

Reagents and conditions: i) EtOH, $Na_2S \cdot 9H_2O$, reflux 6h then dil. HCl; ii) Ac_2O, reflux 0.5h; iii) $MgSO_4$, H_2O, $KMnO_4$, 95°C, 4h, HCl; iv) NH_4OH, $FeSO_4 \cdot 7H_2O$, H_2O, 1h then $NaNO_2$, H_2SO_4, Et_2O.

1.3.2.2 Cycloserine

This natural broad-spectrum antibiotic was determined to be D-4-amino-3-isoxazolidone (oxamycin) **46**[21] and has been given the generic name, cycloserine. The compound has been synthesized in several steps from DL-serine **47** as indicated in **Scheme 10**.[21,56-59] The first step involves protection of the amino acid by Fischer esterification to form the methyl ester hydrochloride **48** which, on treatment with ethyl iminobenzoate **49**, yields DL-4-carbomethoxy-2-phenyl-2-oxazoline **50**. Conversion to the 4-carbohydroxamido-2-phenyl-2-oxazoline **51** is achieved by treatment of the precurser **50** with hydroxylamine and sodium methoxide in methanol at room temperature. Heating the hydroxamic acid **51** in dry dioxane with one equivalent of hydrogen chloride results in rearrangement and ring opening of the oxazoline hydrochloride to afford α-benzamido-β-

cloropropionohydroxamic acid **52**,[58,59] which cyclises to 4-benzamido-3-isoxazolidone **53** when treated with warm aqueous alkali.[58] Removal of the benzoyl group with boiling ethanolic HCl affords the dihydrochloride salt of DL-β-aminoxyalanine ethyl ester **54,** which recyclises when treated with aqueous potassium hydroxide. Resolution with a chiral resolving agent (*e.g.* tartaric acid) affords D-cycloserine **46**.

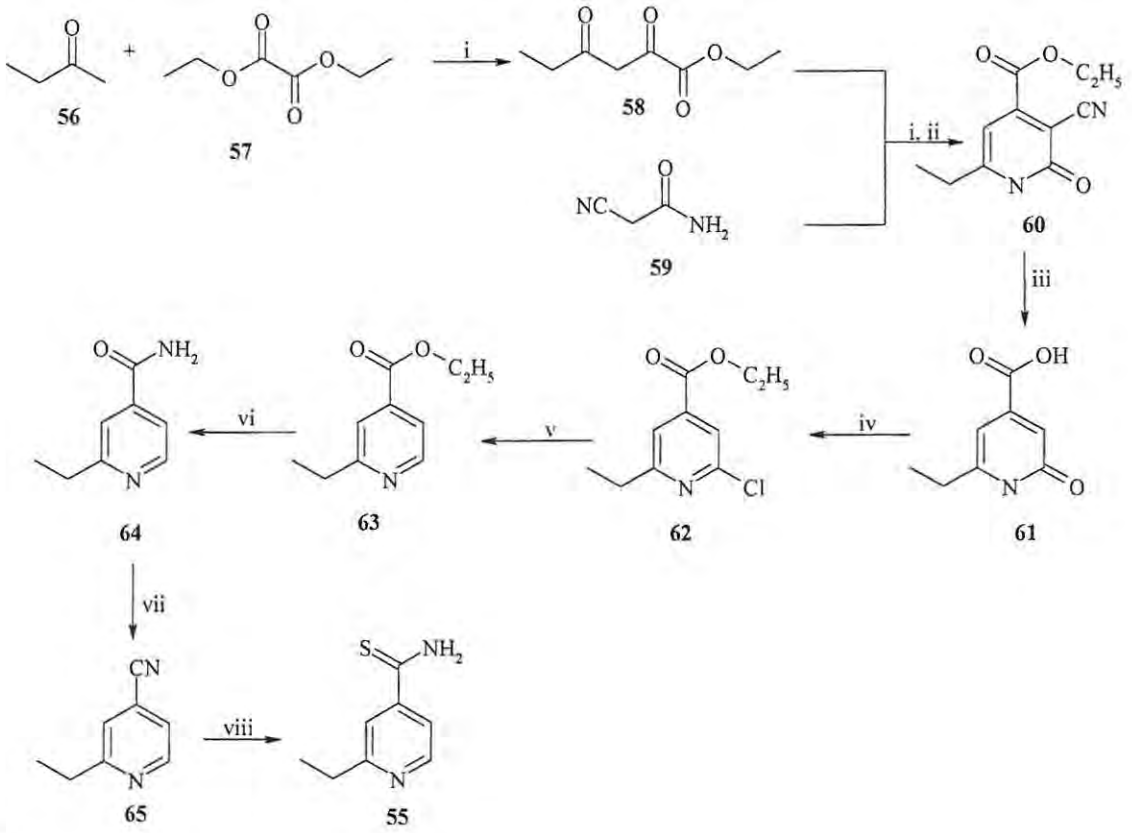


Scheme 10: Synthesis of cycloserine

Reagents: i) MeOH, HCl; ii) NH_2OH, $NaOCH_3$; iii) HCl, dioxane; iv) NaOH then HCl; v) C_2H_5OH, HCl; vi) KOH, AcOH; viii) tartaric acid, Amberlite.

1.3.2.3 Ethionamide

2-Ethylisonicotinic acid thioamide (ethionamide) **55** is a second-generation sulfur-containing antituberculostic agent, which was introduced in 1966 and which is structurally similar to isoniazid (INH). Its mechanism of action is also believed to be similar to that of INH.[3,7] It is a pro-drug that requires activation by the enzyme *ethR, via S*-oxidation, instead of the enzyme *katG* and it acts on multi-drug resistant strains of *MTb* infection by inhibiting mycolic acid biosynthesis.[3, 7,11]



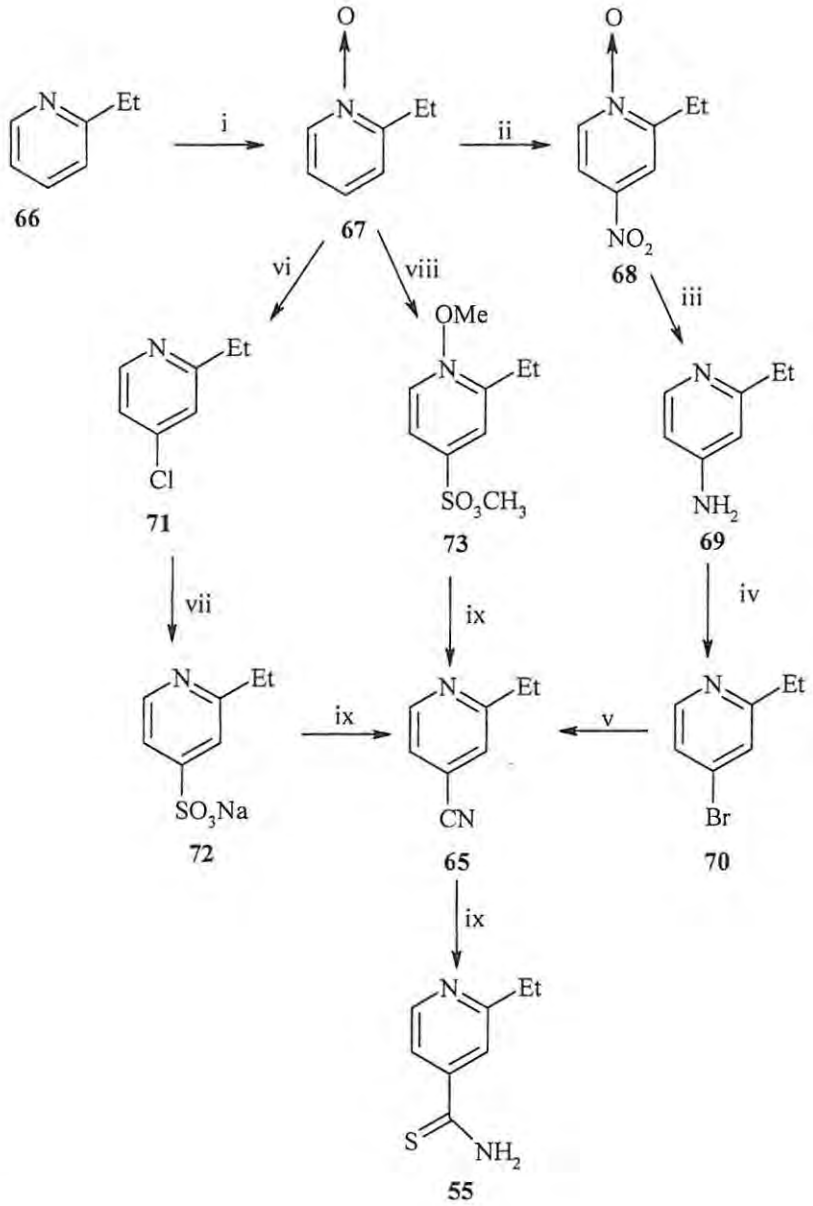
Scheme 11:

Reagents: i) sodium ethoxide; ii) reflux in 95% EtOH; iii) HCl or H_2SO_4; iv) $POCl_3$ or PCl_5, EtOH; v) H_2, Pd/C; vi) NH_4OH; vii) P_2O_5; viii) Et_3N, MeOH, H_2S.

Several methods have been developed for the synthesis of ethionamide,[23,60-64] the first of which (**Scheme 11**) involves condensation of methyl ethyl ketone **56** with oxalic acid diethyl ester **57** to form the diketo acid ester **58** which, on condensation with cyano-acetamide **59**, forms 3-cyano-4-carbethoxy-6-ethyl-2-pyridone **60**. Decarboxylative

hydrolysis affords precursor **61** which, on treatment with $POCl_3$ in ethanol, forms the ester **62**. Hydrogenolysis followed by aminolysis affords 2-ethylisonicotinamide **64**. Dehydration results in the formation of 2-ethylisonicotinonitrile **65** from which 2-ethylisonicotinic thioamide **55** is precipitated by the addition of hydrogen sulphide. However, this method is not suitable for industrial synthesis as it involves a large number of steps.

Other methods[61-63] of preparing ethionamide **55** make use of 2-ethylpyridine **66** as the starting material (**Scheme 12**). This is oxidised with peracetic acid to form the pyridine-*N*-oxide **67**, nitration of which gives 2-ethyl-4-nitropyridine-*N*-oxide **68**. Compound **68** is then reduced to the amine **69** and converted by diazotization and bromination to 4-bromo-2-ethylpyridine **70** which, upon heating with copper cyanide, produces 4-cyano-2-ethylpyridine **65**.[61] This method for producing compound **65** is not ideal as it involves the use of Cu(II)CN which has to be prepared by a special method. However, this approach can be avoided by either converting the *N*-oxide **67** to 4-chloro-2-ethylpyridine **71** using chlorine dioxide and sulfur dioxide[62] followed by sulfonation with sodium metabisulfite to form the sulfate **72** or by treating compound **67** with dimethyl sulfate to form the sulfate **73**.[63] Treatment of the sulfonate **72** and the sulfate **73** with sodium cyanide then leads to the intermediate **65** which forms ethionamide **55** on treatment with hydrogen sulfide.

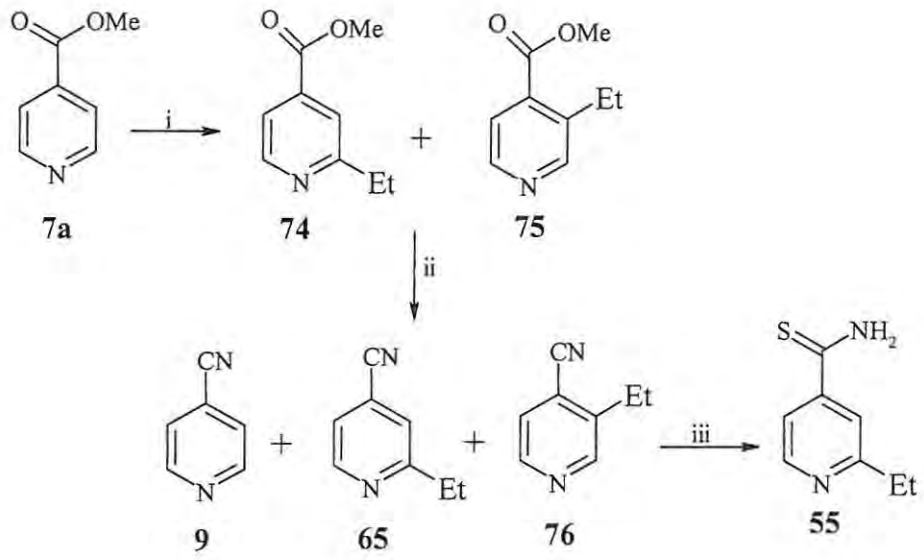


Scheme 12:

Reagents: i) CH_3CO_3H; ii) H_2SO_4 and HNO_3; iii) Fe, AcOH-H_2SO_4; iv) Br_2, $NaNO_2$, NaOH; v) $Cu(CN)_2$; vi) chlorine dioxide, SO_2 ; vii) $Na_2S_2O_5$; viii) Me_2SO_4; ix) NaCN; x) Et_3N, MeOH, H_2S.

Another method (**Scheme 13**)[64] for preparing ethionamide **55** involves radical ethylation of isonicotinic acid methyl ester **7a** with dipropionyl peroxide (prepared from propionic anhydride, H_2SO_4 and 60% H_2O_2) to form the 2-ethyl- and 3-ethylisonicotinic acid methyl esters (**74** and **75**). These are treated with ammonium hydroxide and dehydrated

with alumina to form the isonicotinonitriles **9**, **65** and **76** which are reacted with hydrogen sulphide to form a mixture of thioamides from which ethionamide **55** can be isolated.

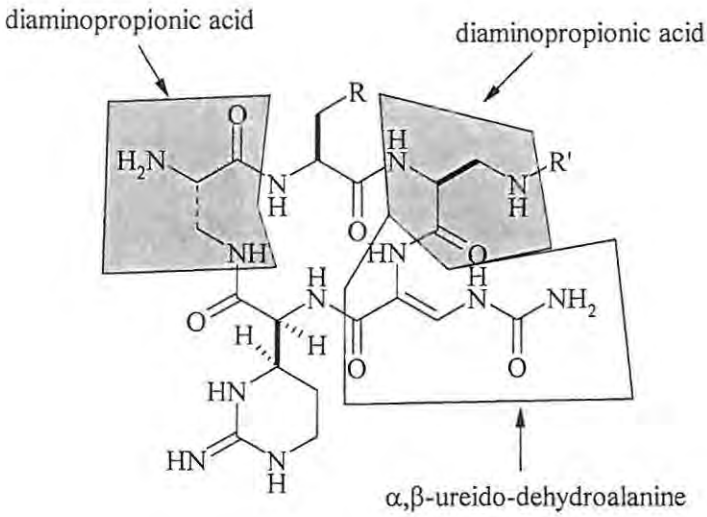


Scheme 13:

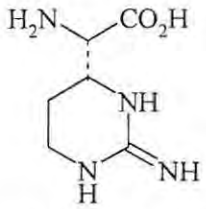
Reagents: i) dipropionyl peroxide, Na_2CO_3, ii) NH_4OH, Al_2O_3, NH_3; iii) Et_3N, MeOH, H_2S.

1.3.2.4 Capreomycin

Capreomycin (Cpm) **77** is a cyclic polypeptide antibiotic comprising four congeners, *viz.*, Cpm IA, IB, IIA and IIB (**Figure 3**), each of which are composed of two diaminopropionic acid residues, an α,β-unsaturated amino acid residue, a ureido-dehydroalanine moiety and a cyclic guanidino amino acid (2*S*,3*R*)-capreomycidine **78**.[65,66] Both Cpm IA and IIA contain a serine residue in the cyclic peptide structure (R = OH), while Cpm IB and IIB contain an alanine residue (R = H). In addition Cpm IA and IB have a β-lysine residue attached to one of aminopropionic acid residues. Cpm is active against MDR-TB and this is very important because of the increase in the number of individuals with compromised immune systems due to HIV infection. It acts by inhibiting protein biosynthesis, inhibiting the dissociation of ribosomal subunits and the translocation of peptidyl tRNA.[66]

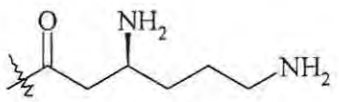


77a: Cpm IA; R= OH, R' = β-lysine
77b: Cpm IB; R= H, R' = β-lysine
77c: Cpm IIA; R= OH, R' = H
77d: CpmIIB; R= H, R' = H



(2*S*, 3*R*)-capreomycidine

78



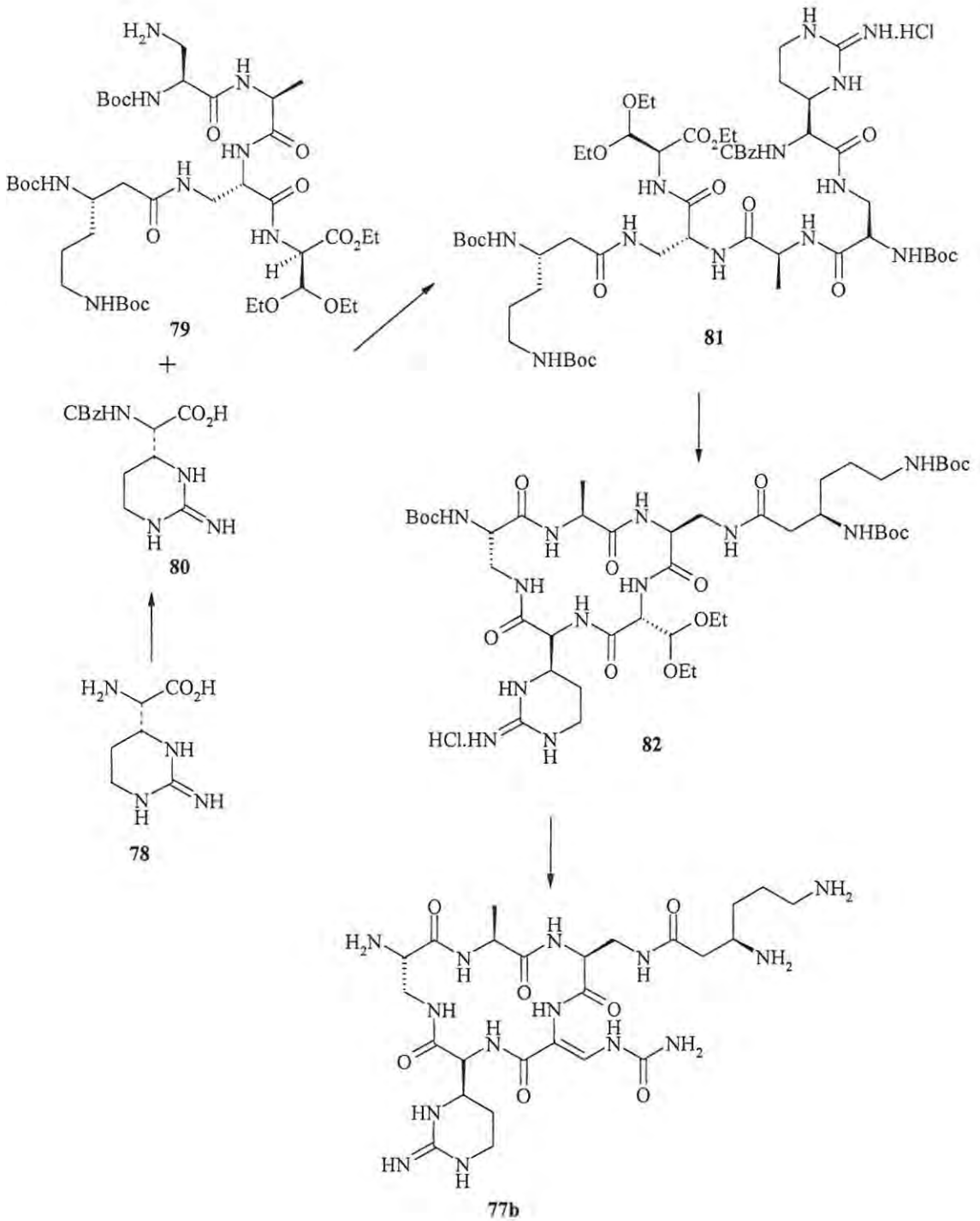
β-lysine

Figure 3: Structures of capreomycins **77a-d** and capreomycidine **78**.

The synthesis of Cpm IB **77b** involves several phases, including:-

i) the multistep synthesis of capreomycidine **78**;[65,66]
ii) the multistep synthesis of the protected pentapeptide **79** from α-formylglycine;[65,66] and
iii) coupling of the amino acids.[65,66]

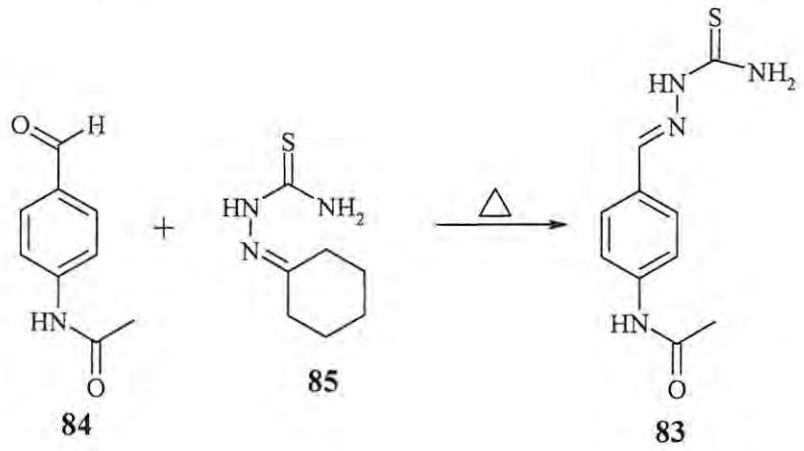
Protection of the capreomycidine **78** amino group as the CBz derivative **80** is followed by condensation with pentapeptide **79** to form the linear hexapeptide **81**. Hydrolysis of the CBz group and the ethyl ester, followed by treatment with EDCI and HOAt results in cyclization of the linear peptide **81** to afford the macrocyclic peptide **82**. Deprotection of **82** and treatment with urea finally affords Cpm IB **77b**.[66]



Scheme 14: Synthesis of Cpm 1B.[66]

1.3.2.5 Thiacetazone

Thiacetazone **83** is a first-generation sulfur-containing antibiotic used in the treatment of TB. It is synthesized by reacting the aromatic aldehyde, *p*-acetamidobenzaldehyde **84**, with the thiosemicarbazine **85**, thiocarbazide or ammonium thiocyanate under reflux conditions (**Scheme 15**).[67-69]



84 **85** **83**

Scheme 15: Synthesis of thiacetazone **88**.

1.3.2.6 Fluoroquinolones

As resistance to the familiar antibiotics grows, it has become important to develop new classes of antibiotics with new structures; the quinolone-based antibiotics were developed to fulfill such a role. These antibiotics show a broad spectrum of activity against Gram-positive, Gram-negative and mycobacterial organisms.[26] The compounds consist of a 4-quinolone nucleus with:-

i) a carboxylic acid moiety at position 3;

ii) a fluorine atom at position 6 for biological activity; and

iii) other substituents at position 7 and/or 8 to fine-tune the properties of the antibiotic (**Figure 4**).[26]

The class includes ciprofloxacin **86**,[26,70-72] levofloxacin **87**,[26,73,74] ofloxacin **88**[26,70,73] and moxifloxacin **89**[26,75] (**Figure 5**).

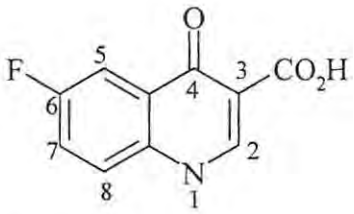
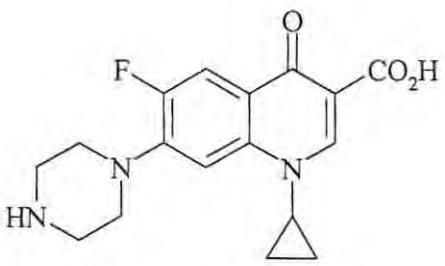
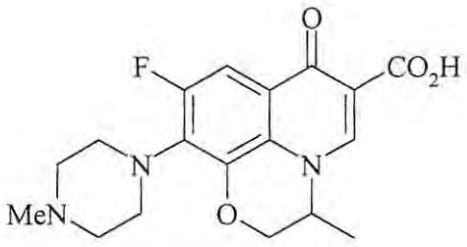


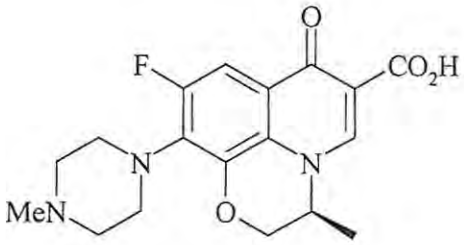
Figure 4: The fluoroquinolone nucleus.



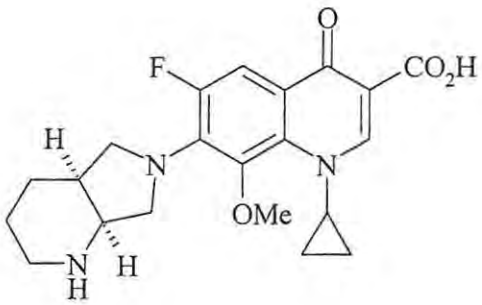
ciproflxacin 86



ofloxacin 88



levofloxacin 87



moxifloxacin 89

Figure 5: Examples of fluoroqinolone antibiotics used in TB Chemotherapy.

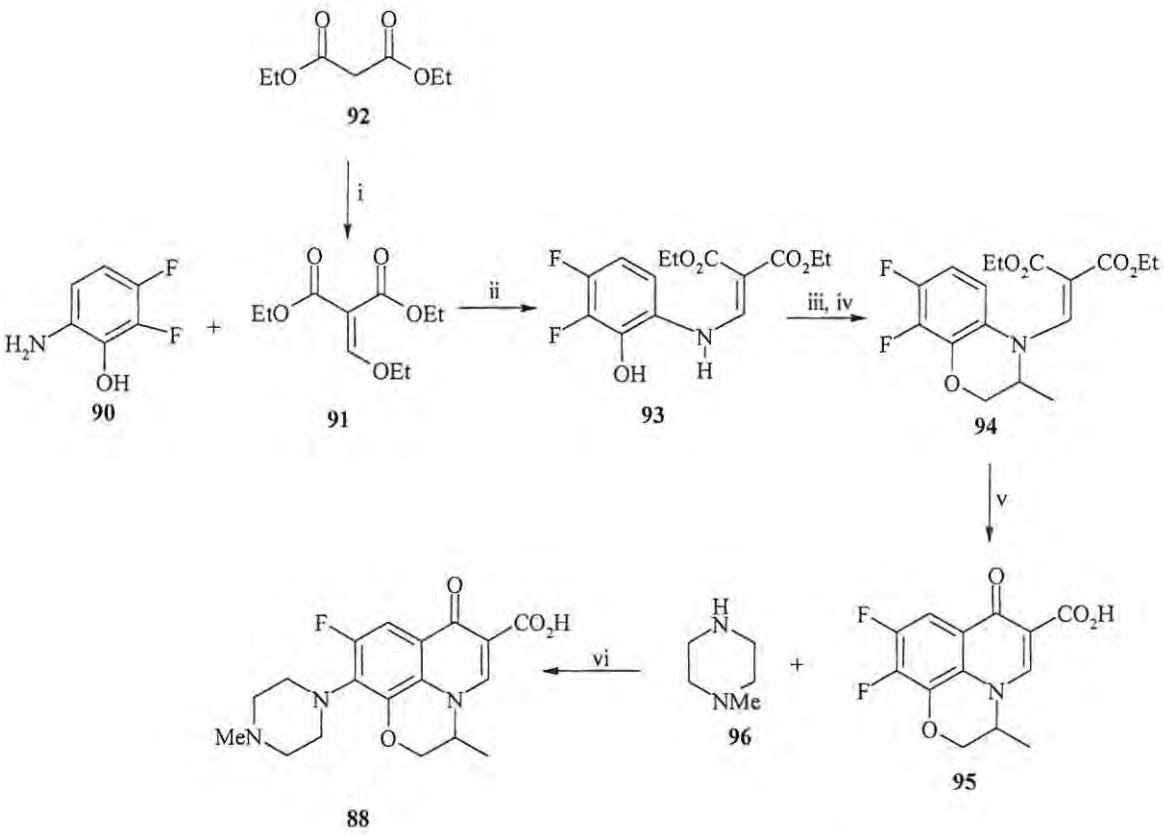
The synthesis of these compounds involves two parts:-

i) synthesis of the fluoroquinolone nucleus; and

ii) the introduction of the substituent at position 7.[26,70-75]

The quinolone nucleus can be constructed by either of the two major methods, the Bayer or the Gould-Jacobs method, which is a modification of the Skraup`s method; both methods involve an enamine intermediate.[26,70] In the Gould-Jacobs methodology[70,73] a fluorinated aromatic amine **90** reacts with an enol ether **91** (formed by reaction between diethyl malonate **92** and ethyl orthoformate) in an addition-elimination process to form

the enamine **93**. This is converted to the preccusor **94** which undergoes an acid- or heat-induced cyclization to form the quinolone **95** which, on treatment with *N*-methylpiperizine **96** followed by hydrolysis, affords ofloxacin **88**.

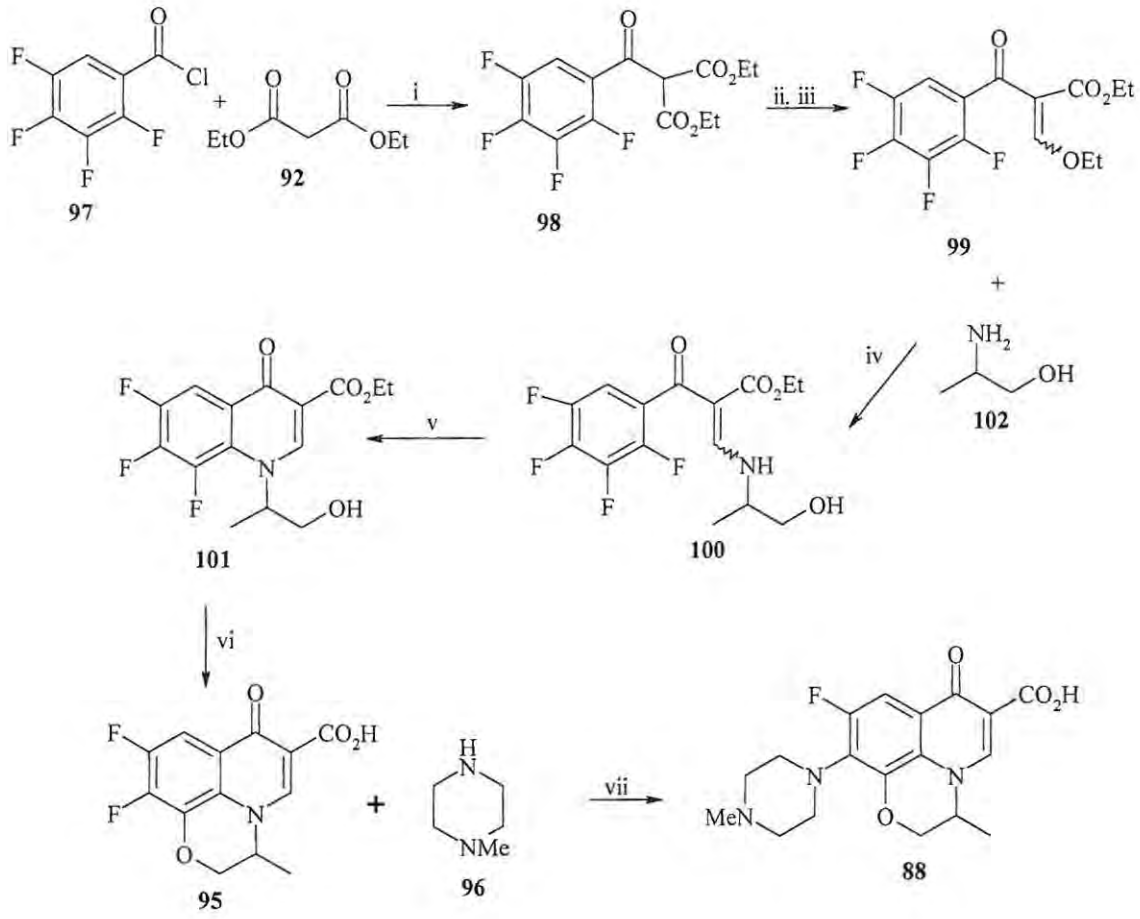


Scheme 16: Synthesis of ofloxacin by the Gould-Jacobs method.[73]

Reagents: i) $CH(OEt)_3$, Ac_2O; ii) EtOH, heat; iii) $LiClO_4$, NaH, propylene oxide; iv) DEAD, PPh_3, AcOEt; v) Ac_2O, H_2SO_4; vi) DMSO

The Bayer methodology[70,72,74-76] (illustrated in **Scheme 17**) generally involves the acylation of diethyl malonate **92** with tri- or tetrafluorinated benzoyl chloride **97** to form the acyl malonate **98**, which is then decarboxylated to remove the redundant ester and reacted with ethyl orthoformate to form the enol ether **99**. This is followed by conjugate addition of an amine to the double bond to form the enamine **100** which on addition of a base undergoes cyclization *via* intramolecular aromatic substitution to form the quinolone **101**. In the case of ofloxacin and levofloxacin the amine can have a hydroxy substituent, *e.g.* 2-aminopropanol **102**, so as to facilitate formation of the tricyclic derivative **95**.

Whereas in ciprofloxacin and moxifloxacin cyclopropylamine is used as the amine. The quinolone **95** then undergoes nucleophilic substitution with *N*-methylpiperizine **96** to afford ofloxacin **88**.



Scheme 17: Synthesis of ofloxacin by Bayer Method.[70,74-76]

Reagents: i) NaH; ii) NaOH then HCl, heat; iii) $CH(OEt)_3$, Ac_2O; iv) EtOH, heat; v) KOH, THF, 0^oC; vi) KOH, THF-H_2O, heat; vii) DMSO.

1.4 Targets for New Drug Development

The emergence of drug-resistant TB poses a real challenge in the effective treatment and complete eradication of TB in society and, hence, there is a need for the development of new drugs that inhibit different targets. It is also important for the new agents to be able to shorten the length of therapy, an aim that could be achieved by the development of drugs that inhibit targets involved in dormancy or persistance. When choosing targets for drug development it is crucial that the target is involved in vital aspects of bacterial growth, metabolism, viability or survival and persistance *in vivo*. Since a complete sequence of the *MTb* genome has been established,[77] a wide variety of enzymes have been targeted by researchers as inhibition sites (structure-based design). These targets include:- glutamine synthetase, which is believed to be responsible for virulence (see Section 1.5); isocitrate lyase, which is believed to be responsible for persistance;[3,78,79] and the Antigen 85 complex, which is composed of three proteins (antigens 85A, B and C) and is responsible for the synthesis of the cell wall mycolic acids.[3,80]

1.5 Glutamine Synthetase

Glutamine is an amino acid found in all living organisms. It is a constituent of proteins and is used in the biosynthesis of several amino acids *via* transamination; its amide nitrogen is used in several biosynthetic pathways, including the synthesis of purines, pyrimidines, glucosamine and carbamoyl phosphate.[81] In bacteria, the glutamine is mainly used in the production of histidine, tryptophan and *p*-aminobenzoic acid.[81]

Glutamine Synthetase [GS (EC 6.3.1.2)] is an enzyme, which is also found in living organisms where it plays an important role in the control of nitrogen. There are at least three types of GS. GSI is found only in eubacteria and archaebacteria, GSII is found in eukaryotes and some bacterial species, while GSIII is found in a few bacterial species.[82,83] GS is an ATP-dependent enzyme that catalyses the condensation of glutamate and ammonia to form glutamine, ADP, and energy in the form of a free phosphate. The biosynthetic reaction may be written as;

$$\text{glutamate} + NH_3 + \text{ATP} \xrightleftharpoons{2 \times M^{2+}} \text{glutamine} + \text{ADP} + \text{Pi} \qquad (1)$$

where M^{2+} is either Mg^{2+} or Mn^{2+} at binding sites n_1 and n_2 and Pi is the free phosphate.

M.Tb GS belongs to the GSI type and is encoded in the gene, *gln*A. It is expressed in four of these genes, known as *gln*A1, *gln*A2, *gln*A3 and *gln*A4, but most of the GS activity is observed in the *gln*A1.[77,84-85] GSI is a dodecamer composed of twelve identical subunits that are arranged in two face-to-face hexagons. Each subunit has:- an active site located at the interface of two adjacent subunits, which can be likened to a "bifunnel"; an ATP binding site at the top of the bifunnel; and a glutamate binding site at the opposite end of the bifunnel with two metal ion binding sites, n_1 and n_2, at the joint of the bifunnel that are always saturated for activity expression (**Figure 6**).[80,86-90]

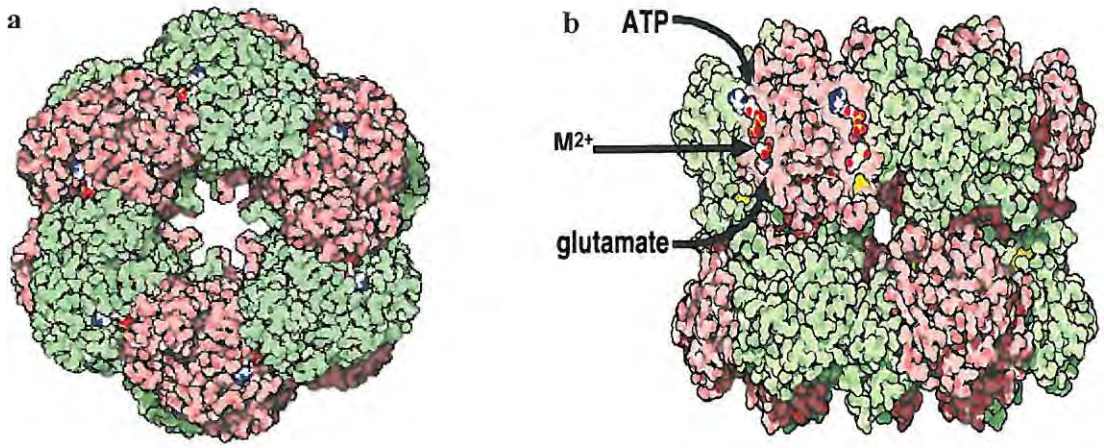


Figure 6: X-Ray structure of bacterial Glutamine Synthetase obtained from the PDB site http://www.pdb.org/pdb/static.do?p=education_discussion/molecule_of_the/pdb30_1.html; **a**) showing the twelve subunits and **b**) showing the binding of ATP and glutamate with the two metal ions shown in purple.

The n_2 metal cation is involved in phosphoryl transfer while the n_1 metal cation stabilizes the active GS and plays a role in binding glutamate.[89] The structure of the dodecamer is held together mostly by hydrogen bonding and hydrophobic interactions between the hexameric rings.[83,86] There are also several exposed loops or "passive sites," which are believed to have functional importance.[86,91] These include:-

a) the Glu-327 loop, which is a flap that consists of residues 324-329, which "guards" the glutamate entrance to the active site and protects intermediates from hydrolysis;[86,89,90]
b) the Tyr-179 loop, which consists of residues 153-188 extending from an active site into the central solvent channel or toward the 6-fold axis of symmetry;[89]
c) the Asn-264 loop which is a flexible loop that consists of residues 255-266, which is located near the glutamate entrance at the lower end of the bifunnel and is adjacent to the Glu-327 flap; [86,89,90]
d) the Asp50' loop, which is a latch consisting of residues 50-64 and is located on the *N*-terminal domain;[86,89,90]
e) the Tyr-397 loop, which consists of residues 388-411 and which is located outside the bottom entrance of the bifunnel, is a site for adenylation and is thus called the adenylation loop;[86,89.91] and
f) the last loop, called the "central loop," which consist of hydrophilic residues 156-173 and which extends into the central channel and contains an arginine residue (Arg-172), which is believed to be a site for covalent modification by ADP-ribosylation.[86,91]

Unlike non-pathogenic bacteria, *M.Tb* and other pathothogenic mycobacteria are unusual in that they secrete a number of proteins in large quantities into the extracellular environment whether growing in broth culture or in human mononuclear phagocytes.[88,92] One of the proteins which is abundantly secreted by *M. Tb* is glutamine synthetase (GS), and this is unusual because this regulatory enzyme is normally situated in the cytoplasm of the bacteria.[92] This suggest that the enzyme is not limited to a regulatory function and hence it is believed that the extracellular enzyme is responsible for the biosynthesis of the poly-L-glutamine-glutamate (P-L-glx) component of the cell wall which is only found in pathogenic bacteria.[88] Extracellular GS is also associated with the ability of the bacteria to inhibit lysosome-phagosome fusion and phagosome acidification and hence virulence.[88]

GSII is composed of eight identical subunits, which are arranged in two face-to-face tetramers, and human GS belongs to this type.[93,94] In humans, the enzyme is mostly found in the brain, liver and kidneys, and in small quanties in muscle tissue. In the brain, GS is responsible for the regulation of the excitory neurotrasmitter, L-glutamate, and for detoxification of neurotoxic ammonia. Brain GS has been associated with neurological disorders such as Alzheimer's disease[95] and schizophrenia.[96] Liver GS is mostly responsible for detoxification of ammonia which escapes from the Kreb's cycle.

1.5.1 Enzyme Regulation

Depending on the function of the GS in the particular tissue (in mammals) or organism, it may be regulated by feedback inhibition or covalent modification, the latter including mechanisms like divalent cation specificity and adenylation.[81,86,93]

1.5.1.1 Feedback Inhibition

Studies have shown that both bacterial and mammalian GS can be inhibited by any of the end-products of glutamine metabolism, *viz.*, adenosine monophosphate (AMP), tryptophan, histidine, alanine, glycine, carbamoyl phosphate, serine or glycine.[81,86,94] Before structural studies had been conducted on the enzyme, it was thought that inhibition occurred by a proccess called 'cumulative feedback inhibition', whereby each inhibitor blocks a different site, but when acting together they almost shut the enzyme down. Later studies have shown non-competitive binding of alanine, serine and glycine to the glutamate site, while AMP and histidine bind to the ATP site. When these metabolites are available in low concentrations they stimulate enzyme activity, but high concentrations inhibit activity. This is observed in mammalian liver and bacterial GS. Brain GS, however, is only inhibited by carbamyl phosphate.[86,94]

1.5.1.2 Covalent modification

1.5.1.2.1 Adenylation-Deadenylation

This mechanism of regulation has been shown to be applicable only in bacterial GS[86] as adenylation occurs at a certain tyrosyl residue (Tyr-397 numbered relative to *E. coli*[86] or *Salmonella*[90,97] or Tyr-406 numbered accoding to *MTb*GS[97,98]), which is absent in GSII. Adenylation-deadenylation regulates the catalytic activity of GSI by the covalent addition or removal of an AMP group to a tyrosine residue in each subunit, which progressively inactivates or activates the enzyme and this occurs as shown in the following reactions.[81,82]

$$GS + 12\ ATP \xrightarrow{ATase,\ Mg^{2+}} GS{\cdot}(AMP)_{12} + 12\ PPi \qquad (2)$$

$$GS{\cdot}(AMP)_{12} + 12\ Pi \xrightarrow{ATase,\ Mg^{2+}} GS + 12\ ADP \qquad (3)$$

Where ATase is the enzyme adenyltransferase.

Adenylation (Reaction 2) occurs in response to an increase in the levels of intracellular nitrogen such that as the cell becomes more nitrogen sufficient, the activity of GSI is progressively reduced. In *MTbGS* the process has been shown to be glutamate dependent.[98] Adenylation of a subunit of GSI inactivates only that subunit, so that the enzyme can exist in a range of activity states in the cell. While deadenylation (Reaction 3) occurs in response to a decrease in the levels of intracellular nitrogen such that as the cell becomes nitrogen defficient, GSI activity is progressively increased. Both these processes are activated by the enzyme adenyltrasferase (ATase) encoded in the *glnE* gene.[99] Hence this enzyme also needs to be regulated, because coupling of the reactions (2) and (3) may occur resulting in the GS fluctuating between the adenylated and unadenylated forms which will result, in turn, in the the aimless phosphorylation of ATP as shown in Reaction 4.[81]

$$12\ ATP + 12\ Pi \longrightarrow 12\ ADP + 12\ PPi \qquad (4)$$

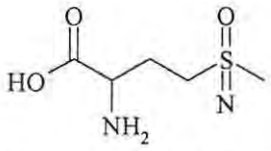
1.5.1.2.2 Divalent Cation Specificity

Both mammalian and bacterial GS need two divalent metal ions for each subunit to maintain structure and activity and, hence, changes in the environmental free concentration of specific divalent ions in the cell result in changes in enzyme activity.[81] The absence of Mn^{2+}or Mg^{2+} causes conformational changes in the structure of the enzyme, a state known as relaxation,[81] leading to the structural flexibility of the protein and thereby exposing the tryptophan and tyrosyl residues surrounding the active site. This leads to a decrease in sedimentation as well as dissociation of the subunits, resulting in inactivation of the enzyme.[81,86,88,89] Addition of the metal ions results in a process known as "tightening," which reverses the conformational changes caused by the relaxation process.[81,89] The human brain GS has been shown to be more active when bound to Mg^{2+} than Mn^{2+} although the enzyme has a high affinity for Mn^{2+}.[86]

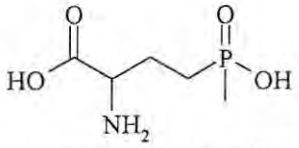
1.5.1 Effects of Synthetic Inhibitors

Extensive research has been undertaken on GS, especially bacterial GS, to investigate enzyme kinetics, to characterize the mechanisms of regulation *in vivo* and to establish the effects of enzyme inhibition in different organisms.[86] So far, research shows that there are three distinct substrate binding sites, *viz.*, nucleotide, ammonia and the amino acid residues in the active site of the enzyme. The most studied of these substrate sites is the amino acid site, which has been found to be inhibited by the naturally occuring phosphinothrin (PPT), L-methionine-(*S,R*)-sulfoximine (MetSox or MSO) and its synthetic derivatives.[86,90] Of all the inhibitors investigated, MetSox is the most active and widely studied for the inhibition of *MTb*GS. MetSox has been found to irreversibly inhibit both GSI and GSII.[92,94,100] It competes with L-glutamate for binding in the active site and can be phosphorylated to produce an intermediate that binds tightly to the enzyme.[94,101] Although MetSox inhibits both GSI and GSII, the inhibitory effect is greater with GSI.[92] MetSOx has been shown to inhibit bacterial growth by inhibiting the extracellular GS (GSI) [92] thus preventing death, weight loss (a symptom of the disease) and spreading of the disease.[100] Anti-sense oligonucleotides modified with phosphorothionate (PS-ODNs) have also been investigated and shown to reduce

intracellular and extracellular GS activity and reduce the formation of poly-glutamine-glutamate complexes and thereby inhibit cell growth.[102]



Methionine sulfoximine (MetSox)



Phosphinothricin (PPT)

Figure 7: Strucutures of GS inhibitors.

Given the belief that:-

i) GSI and GSII do not come from the same class of GS proteins because of differences in their primary structure[93] and in their inhibition mechanisms; and
ii) *M.Tb* has both GSI and GSII because of the differences in regulation of the intracellular and the extracellular *MTb*GS,[85,99]

it has been presumed that extracellular *MTb*GS might be a possible target for the development of new anti-TB drugs.

1.6 Heterocyclic Compounds with Medicinal Activity

In recent years scientists have intensified their efforts in the search for new and more effective antibacterial and antiviral agents. The search has included the synthesis of various acyclonucleosides with different side chains and aglycones.[103]

1.6.1 Indolyl Derivatives

Compounds containing the indole nucleus are abundant in nature, and melatonin **102** (**Figure 8**), a hormone secreted by the pineal mammalian gland, is one of such compounds. The hormone plays an important role in the regulation of circadian and seasonal rhythms, tumor inhibition, immune function and retinal physiology and is also a free radical scavenger and antioxidant.[104-106] Other compounds with an indole nucleus,

which have medicinal applications include:- indomethacin **103** (an anti-inflammatory and the analgesic used against masculoskeletal and joint disorders like rheumatoid arthritis)[107,108] and the anti-depressant agent etryptamine **104** (a tryptamine derivative).[108] Indole-3-acetic acid **105**, a plant growth hormone which is not harmful to humans, has been investigated as a pro-drug in cancer therapy.[108-113] The pro-drug is not oxidised by human peroxidase but, when reacted with Horseradesh peroxidase (HRP), it produce metabolites that are cytotoxic to cancer cells.[108-113] Indole-3-propionic acid **106** has been shown to be a better free radical scavenger than melatonin[114] and it minimises cell damage after injury,[115,116] while indole-3-butyric acid **107** derivatives have been investigated as potential acytylcholine inducers in the treatment of Alzheimer's disease.[117]

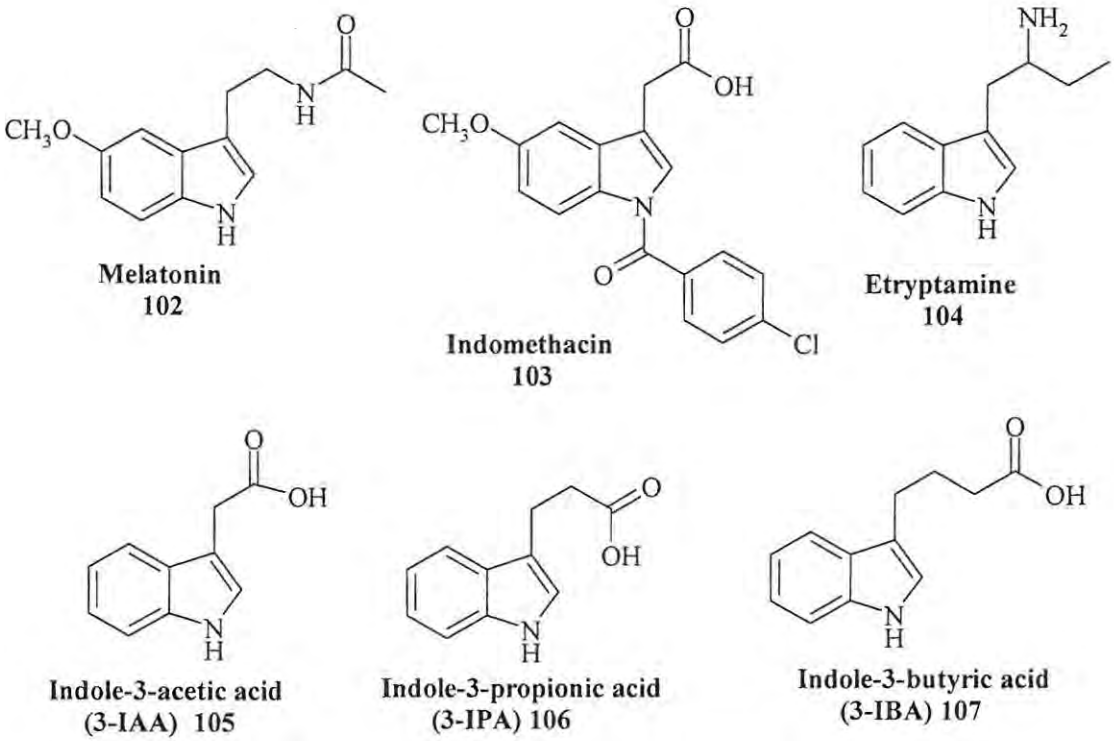
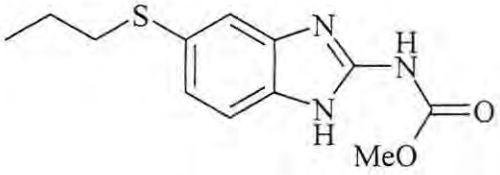


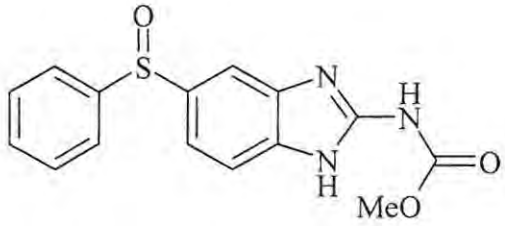
Figure 8: Structures of indole derivatives with medicinal applications.

1.6.2 Benzimidazole Derivatives

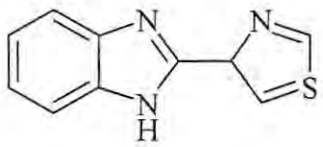
The benzimidazole-based drugs, such as albendazole (albenza), mebendazole and thiabendazole (**Figure 9**), have enjoyed widespread application as microtubule inhibitors in the treatment of parasitic related diseases.[118] They are used in the treatment of diseases such as alveolar echinococcosis hydatid disease and neurocysticercosis, caused by nematode, trematode and cestode parasites.[118-122] Thiabendazole has also been investigated as a potential anti-HIV and anti-cancer agent,[123] while benzimidazole has been investigated as a potential base in the development of neucleotides that can inhibit DNA primase, an enzyme which synthesizes the short-chain RNA oligonucleotides needed to initiate DNA synthesis.[124]



Albendazole



Mebendazole

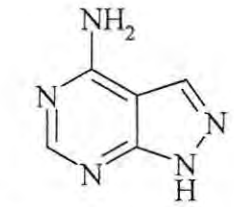


Thiabendazole

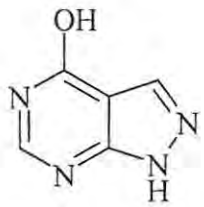
Figure 9: Structures of benzimidazole-based drugs.

1.6.3 Pyrazolo[3,4-*d*]pyrimidine

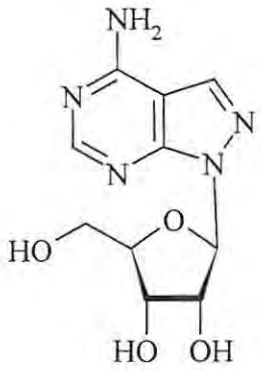
The 4-substituted pyrazolo[3,4-*d*]pyrimidine systems, 4-aminopyrazolo[3,4-*d*]pyrimidine (4-APP) **108a** and 4-hydroxypyrazolo[3,4-*d*]pyrimidine (allopurinol) **109a** have attracted attention because of their structural similarity to adenine, a purine that is widely distributed in nature. Allopurinol **109a**, a xanthine oxidase inhibitor, has been widely used in controlling gout and related metabolic disorders.[125] Recently it has been shown to exhibit anti-psychotic effects against schizophrenia[126] and to prevent the development of chemotherapy induced stomatosis.[127-129] Both 4-APP **108a** and allopurinol **109a,** and their ribonucleoside and truncated acyclic derivatives (compounds **108b-d-109b-d** in **Figure 10**) have been shown to posses antiparasitic activity[130-133] as well as antitumor activity against leukimia cells.[134] 4-APP derivatives substituted at either N-1 or N-2 have been investigated as potential adenosine deaminase inhibitors,[135] while those substituted at C-6 have been investigated as potential DNA gyrase inhibitors.[136]



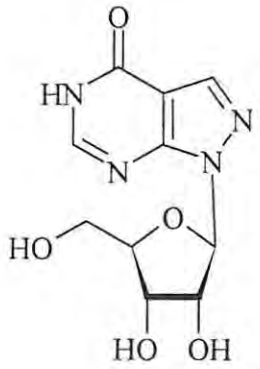
4-aminopyrazolo[3,4-*d*]pyrimidine 108a



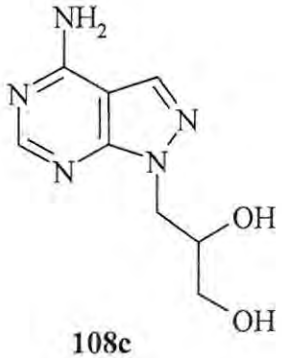
Allopurinol 109a



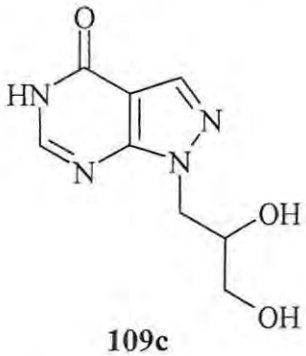
108b



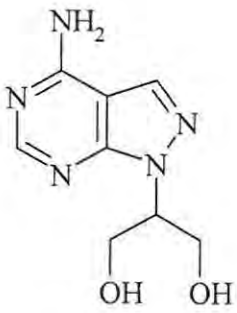
109b



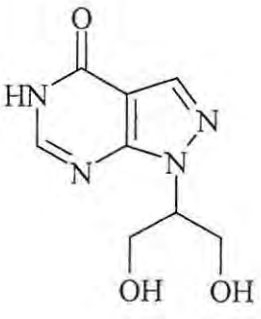
108c



109c



108d



109d

Figure 10: Stuctures of 4-APP, allopurinol and their derivatives with medicinal properties.

1.7 Aims of the Present Investigation

Of all diseases, tuberculosis remains the world's number one killer even though a number of effective anti-TB drugs are on the market. The development of resistance, in the form of MDR-TB and XDR-TB, and the association of the disease with HIV/AIDS has emphasized the need for the development of new drugs that exhibit a different mode of action (target) to those already in use.

In our efforts to develop new anti-TB drugs, glutamine synthetase was identified as the target. Previous work on the enzyme has shown that inhibition of the enzyme results in the inhibition of bacterial growth and produces non-virulent strains. Consequently, this work has focussed on the design and synthesis of ATP analogues as potential GS inhibitors. More specifically, attention has been given to:-

i) the use of various heterocyclic scaffolds for the construction of novel ATP analogues;
ii) the application of 1- and 2-dimensional NMR spectroscopic methods in the characterisation of these products;
iii) the application of high-resolution mass spectrometric analysis to elucidate the fragmentation patterns exhibited by representative systems; and
iv) computer modelling to explore the docking of the ATP analogues into the active site of the enzyme.

2. Discussion

In the following discussion attention will be given to:- the design, synthesis and characterization of ATP analogues (Section 2.1); NMR studies of various analogues (Section 2.2); molecular modelling (Section 2.3) and mass spectral analysis (Section 2.4).

2.1 Synthesis of ATP Analogues

2.1.1 Design of ATP Analogues

ATP is a nucleotide composed of:- a purine, adenine; a ribose sugar; and three phosphate ester groups (**Figure 11**). It binds to the active site of glutamine synthetase and plays a role in the catalytic conversion of glutamate to glutamine as described in Section 1.5. In our attempts to develop novel enzyme inhibitors as potential anti-tuberculosis agents, we have aimed to explore the development of inhibitors, which might be expected to behave as structural mimics of ATP. ATP has three major components that can be modified to produce new analogues, *viz.*, the heterocyclic base, the ribose linker and the polyhydroxy triphosphate group.

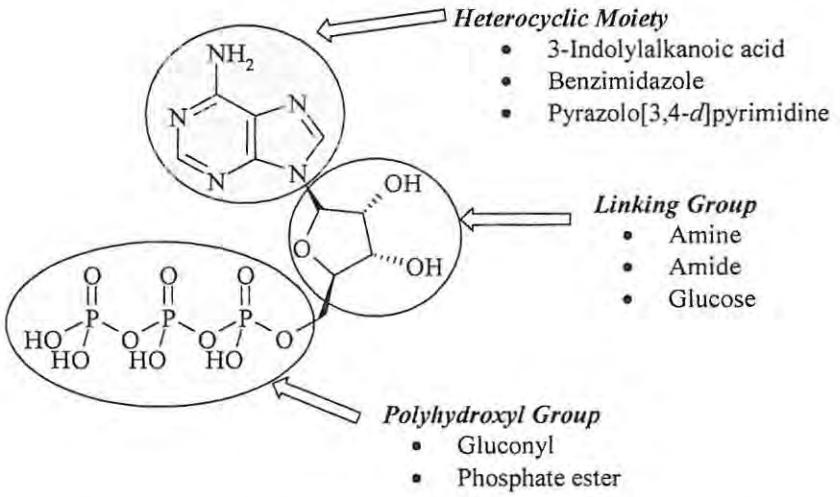


Figure 11: Structure of ATP showing sites for modification and potential replacement options.

In this study, heterocyclic systems, which can undergo normal Watson-Crick hydrogen bonding, such as 4-substituted pyrazolo[3,4-*d*]pyrimidines, or those which cannot, such as benzimidazole and 3-indolylalkanoic acids, have been considered as replacements for the base. Glucopyranosyl and alkyl ether groups have been explored as the linking group and polyhydroxy or phosphorylated groups as replacements for the triphosphate ester moiety. The hydroxyl groups have been reported to be essential for coordination to the divalent metal ion.[137]

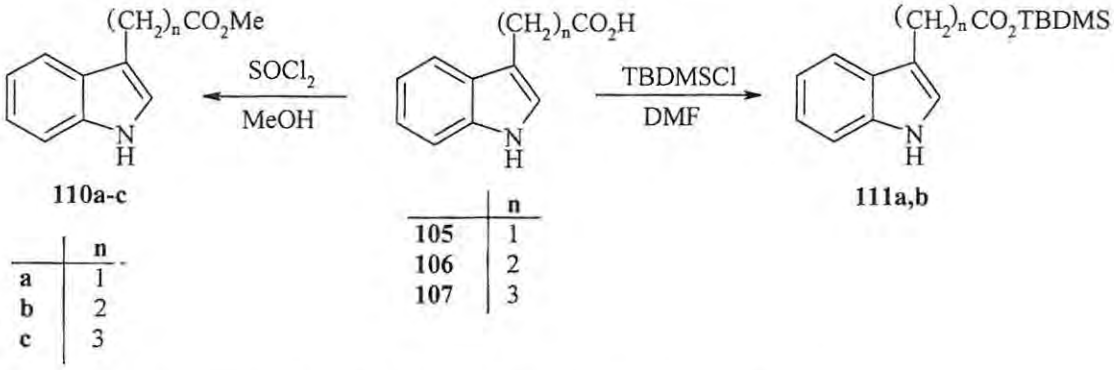
2.1.2 Synthesis of 3-Indolylalkanoic Acid Derivatives

As indicated in Section 1.6.1, 3-indolylalkanoic acids **105-107** have pharmacological activity and can be tolerated by humans. They are structurally similar to the amino acid tryptophan, and have been used as scaffolds in the development of glycoprotein IIb/IIIa (GP IIb/IIIa)[138] and thromboxane[116] antagonists, which are required to activate platelet aggregation. The nitrogen on the indole nucleus has a lone pair of electrons that are not readily available for reactions with electrophiles as they are involved in stabilizing the indole ring. However, removal of the attached proton by strong bases such as NaH, BuLi or $NaNH_2$, affords an activated amino anion that can then be alkylated or acylated to introduce a linking group. The resulting amide or amine can undergo further reactions to introduce polyhydroxy groups. Hence it was hoped that application of these compounds (**105-107**) as scaffolds in the development of ATP analogues would afford mimics capable of participating in hydrogen bonding with the amino acids in the active site of the enzyme.

2.1.2.1 Synthesis of 3-Indolylalkanoic acid esters

The chosen indole scaffold has three obvious reaction sites, *viz.*, the carboxylic acid acyl function, the indole nitrogen and the α-carbon of the acyl group. Esterification of the acid and acylation or alkylation of the ring nitrogen at low temperatures was expected to prevent undesirable side reactions at the α-carbon.[138-140]

Synthesis of the indolyl derivatives was accomplished *via* protection of the commercially available 3-indolylalkanoic acids **105-107** as esters using literature methods.[139-142] The acids were stirred with thionyl chloride in methanol under a stream of nitrogen[139,140] (**Scheme 18**). Purification by flash chromatography afforded the methyl esters **110a-c** in good yields (57-93%; **Table 1**). The structures of the esters were verified by spectroscopic methods. The ^{1}H NMR spectrum illustrated for compound **110b** (**Figure 12**) shows a sharp singlet at *ca.* 3.7ppm integrating for 3 protons and corresponding to the methoxy group, while the NH signal shifts from *ca.* 11ppm, in the precursor, to *ca.* 8ppm; the methylene triplets at *ca.* 3ppm are also clearly evident. The ^{13}C NMR spectrum showed a new signal at *ca.* 52ppm, which was assigned to the methoxy group, and the carbonyl carbon signal was observed at *ca.* 173ppm.



Scheme 18: Synthesis of methyl and silyl 3-indolylalkanoate esters.

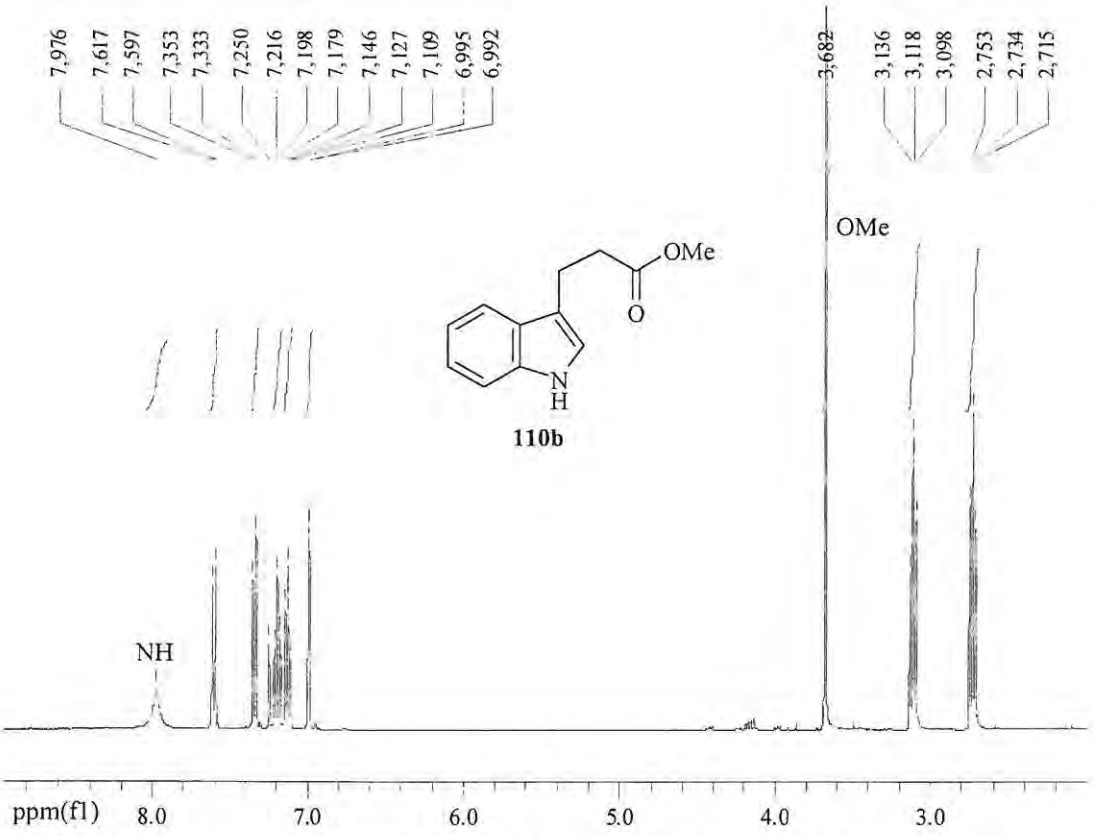


Figure 12: 400MHz ^{1}H NMR spectrum of compound **110b** in $CDCl_3$.

Although the methyl esters were relatively easy to synthesize and isolated in good yields, the literature[143] reveals that they are susceptible to base-catalysed hydrolysis at high temperatures and, consequently, synthesis of silyl ester derivatives was also attempted. *tert*-Butyldimethylsilyl (TBDMS) esters are much more stable than the trimethylsilyl (TMS) esters and can withstand harsh reaction conditions.[141] The 3-indolylalkanoic acids **105** and **106** were treated with 1.1 equivalents of *t*-butyldimethylsilyl chloride (TBDMSCl) in dry DMF in the presence of 2.5 equivalents of imidazole (**Scheme 18**).[142] The reaction mixture was purified by flash chromatography to afford the silyl esters **111a** and **b** in low yields (44% and 47% respectively; **Table 1**). The ^{1}H NMR spectrum illustrated for compound **111a** (**Figure 13**) shows two singlets at *ca.* 0.3 and 0.9ppm integrating for 6 and 9 protons respectively and these were assigned to the silyl methyl and *tert*-butyl protons respectively. Due to lower yields of the silyl esters **111a,b,** the methyl esters **110a-c** were used for further synthesis. It is anticipated that, *in vivo*, the

ester group would be hydrolysed revealing a carboxylic acid group capable of hydrogen-bonding or metal coordination in the GS receptor cavity.

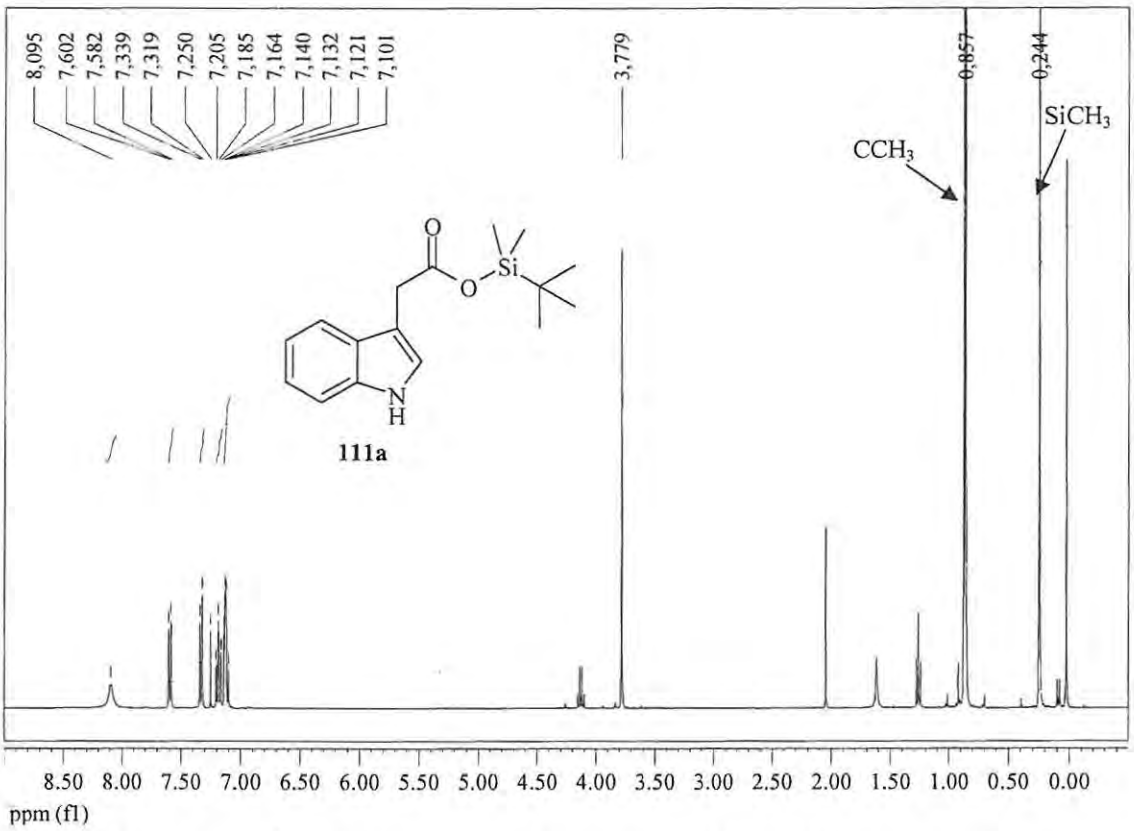
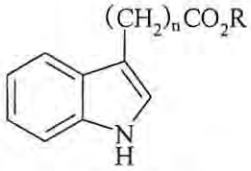


Figure 13: 400MHz ^{1}H NMR spectrum of compound **111a** in $CDCl_3$.

Table 1: Data for the indolylalkanoate esters **110a-111b**.



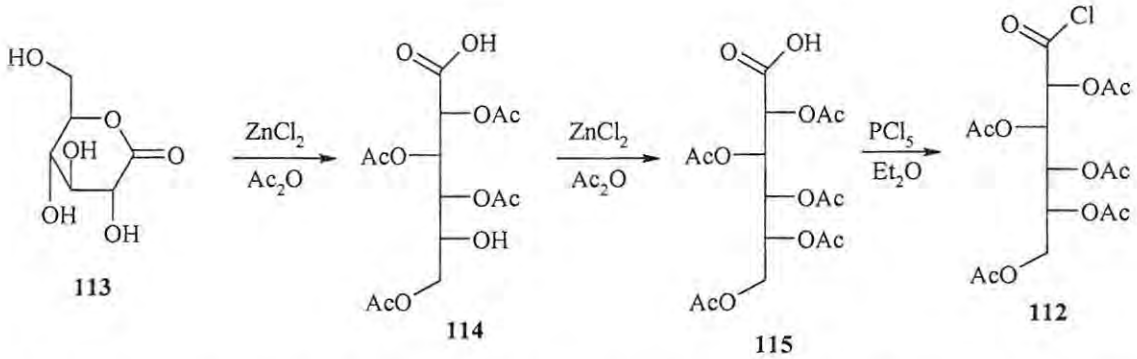
Compound	n	R	Yield [a]/%	Mp.[b]/°C
110a	1	Me	96	oil
110b	2	Me	68	78-80[c]
110c	3	Me	88	61-63
111a	1	TBDMS	44	oil
111b	2	TBDMS	47	oil

[a] Isolated material. [b] Following flash chromatography. [c] Reported as oil; lit.[139,140]

2.1.2.2 Attempted Acylation of the 3-Indolylalkanoic Acid Esters

Several commercially available compounds with hydroxyl substituents were explored as replacements for the ribofuranosyl group of ATP. However, they needed to be protected before coupling to the heterocyclic nucleus to prevent undesirable side reactions and/ or to permit activation. One such compound is the acyclic sugar derivative, gluconic acid, which exists naturally as the cyclic structure **113** due to its tendency to lactonise. The lactone is not highly reactive and cannot be used directly as an acylating agent. The acid chloride **112** was therefore required and was synthesized from D-glucono-δ-lactone **113**, by initial protection of the hydroxyl groups in D-glucono-δ-lactone **113** as esters using the method described by Braun and Cook,[144] followed by treatment with phosphorous pentachloride. Thus D-glucono-δ-lactone **113** was acetylated with acetic anhydride in the presence of a Lewis acid catalyst, anhydrous zinc (II) chloride. Interestingly acetylation is achieved in two steps (**Scheme 19**) *via* the intermediate 2,3,4,6-tetra-*O*-acetyl-D-gluconic acid monohydrate **114**, which was isolated and characterized (**Table 2**); further treatment under the same conditions affords 2,3,4,5,6-penta-*O*-acetyl-D-gluconic acid **115**. The penta-acetate was then converted to the acid chloride **112**, which was used in the

acylation of the heterocylic bases, by treating with either phosphorous pentachloride[144] or oxalyl chloride.[145] However, initial attempts to form gluconamide derivatives proved to be very frustrating.



Scheme 19: Synthesis of penta-acetylated gluconyl chloride **112** from D-glucono-δ-lactone **113**.

Table 2: Comparative yields and ^{1}H and ^{13}C NMR data for the gluconyl derivatives **112, 114** and **115**

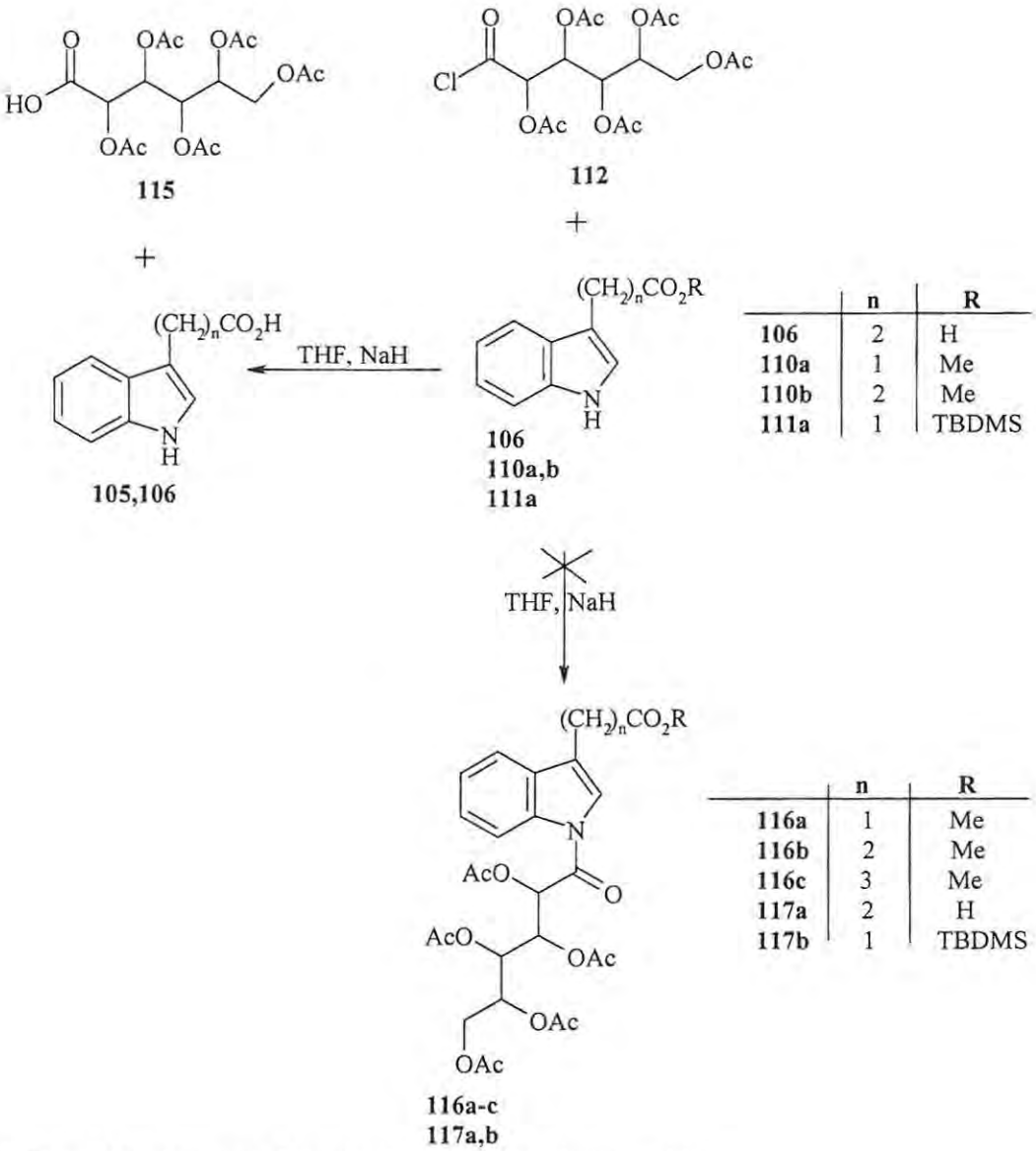
Compound	Yield [a]/%	Mp. ^{O}C	δ_H[e] CH_3	δ_C[e] CO
112	62.8	60-62 (68-71)[b]	2.04-2.20	169.7-170.6[f]
114	23	82-84 (113-117)[c]	2.07-2.16	169.1-170.3[g]
115	93	110-113 (110-111)[d]	2.02-2.16	169.8-170.7[f]

[a] Isolated material. [b] Crystallized from Et_2O. [c] Crystallized from water. [d] Crystallized from toluene; lit.[9] values in parentheses. [e] In $CDCl_3$. [f] 6 x carbonyl signals, [g] 5 x carbonyl signals

Acylation of both series of 3-indolylalkanoic acid esters **110a-c** and **111a,b** was attempted using reagents such as *N,N'*-dicyclohexyldicarbodiimide (DCC), as coupling agent, and NaH, BuLi and NaOH as bases to deprotonate the substrate and enhance acylation .[138,146-149]

Our first attempt was to acylate the esters **110** and **111** with penta-*O*-acetylgluconyl chloride **112** in the presence of NaH in THF as outlined in **Scheme 20**. When the

acylation was attempted in THF at room temperature under a stream of nitrogen, the reaction did not occur; instead, both the ester **110a** and the acyl chloride **112** were hydrolysed as shown by spectroscopic analysis of the products. Addition of chloroform to the crude reaction mixture precipitated the 3-indolylalkanoic acid **105** while the oil isolated from the chloroform solution was shown by MS analysis to contain ester **110a** and acid **115,** but not the expected amide **116a**. The same pattern was observed with ester **111a**. The reaction conditions were therefore changed. Firstly, the experiment was conducted at low temperature (–78^{O}C) under argon, and butyl lithium was used to enhance the nucleophilicity of the indole nitrogen as described by Teranishi *et al.*[146] and, secondly, the experiment was conducted at room temperature in pyridine with freshly prepared acid chloride **112**. Unfortunately, both these methods also failed to produce the desired product.



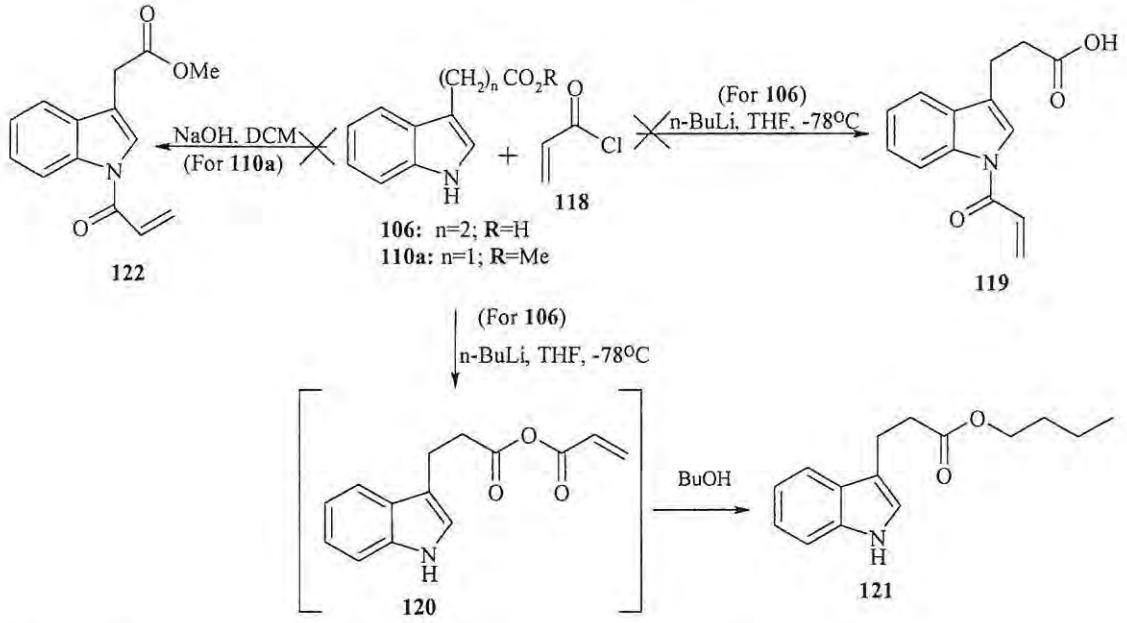
	n	R
106	2	H
110a	1	Me
110b	2	Me
111a	1	TBDMS

	n	R
116a	1	Me
116b	2	Me
116c	3	Me
117a	2	H
117b	1	TBDMS

Scheme 20: Attempted synthesis of indolyl gluconamides.

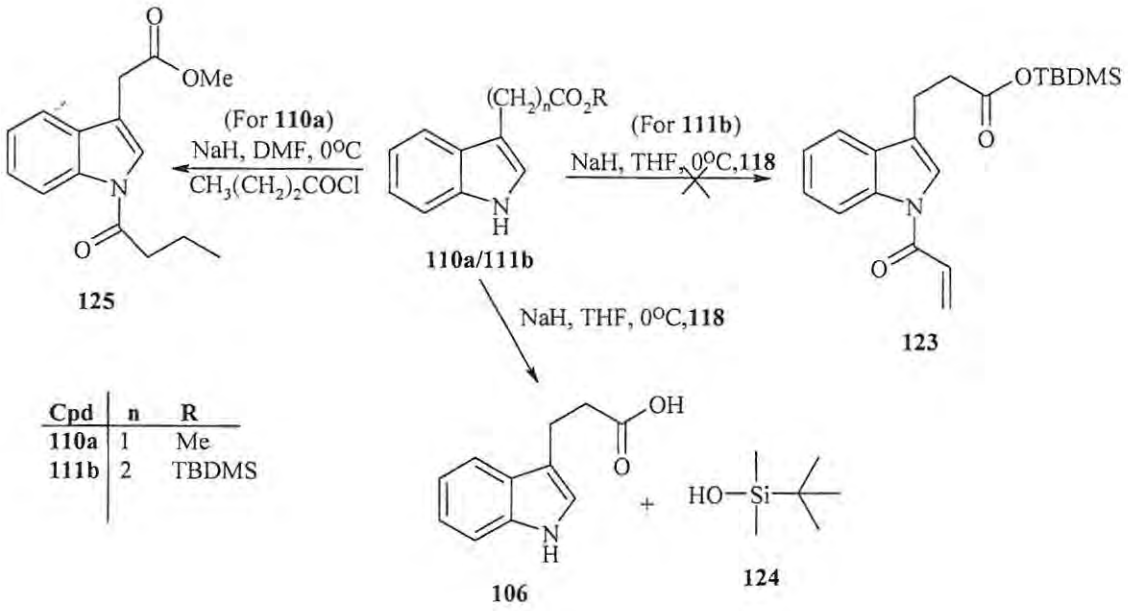
Due to the fact that the acid chloride **112** is highly susceptible to hydrolysis, gluconic acid **115** itself was explored as the acylating agent. Treatment of the indolylalkanoic ester **110c** with acid **115** in the presence of DCC in dichloromethane was also unsuccessful,[150] and smaller molecules (such as acryloyl chloride **118,** with potential for chain elongation *via* the Baylis-Hillman reaction, and butyryl chloride) were considered as the acylating agents. Since the methyl esters were hydrolysed during acylation, the 3-indolylalkanoic acid (3-IPA) **106** was used directly instead. Following the method reported by Teranishi

et al.,[146] the acid **106** was treated with acryloyl chloride **118** and two equivalents of BuLi in THF under argon (**Scheme 21**). This failed to produce the desired indole amide **119**, affording instead ester **121**, the formation of which was indicated by the absence of vinylic nuclei signals in both the ^{1}H and ^{13}C NMR spectra. Formation of the ester **121** may be due to reaction of the lithium carboxylate with acryloyl chloride to form a mixed anhydride **120**, which underwent acyl substitution with BuOH during work-up. Further attempts at this acylation were made with ester **110a** following the method described by Ottoni *et al*,[147] but these also failed to produce the desired product **122.**



Scheme 21: Attempted acylation with acryloyl chloride **118.**

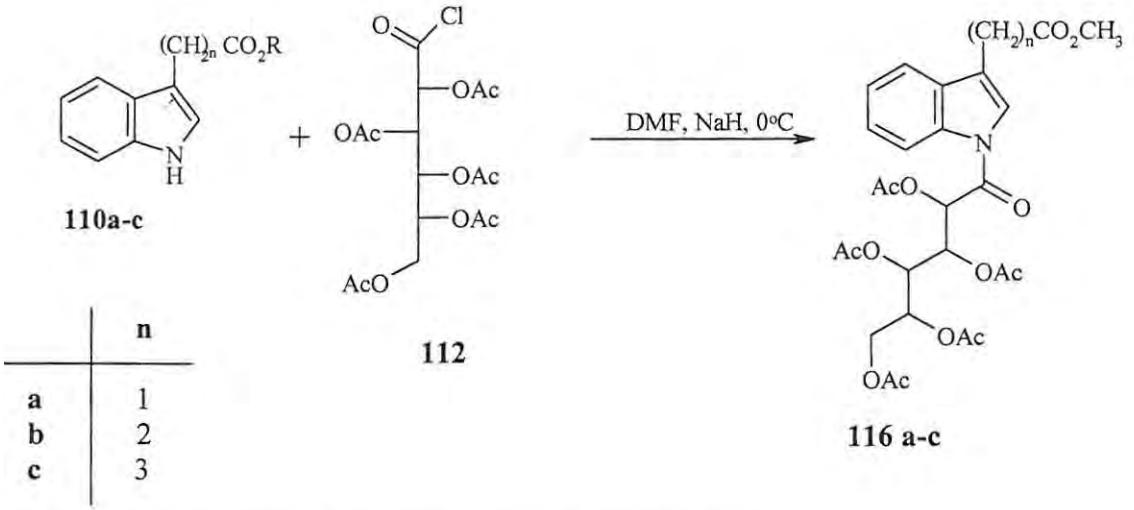
Acylation of the TBDMS ester **111b** with acryloyl chloride and NaH in THF (at $0^{O}C$-room temperature)[149] afforded a fraction, following flash chromatography, which was thought to contain the indole amide **123**. After HPLC purification and analysis by ^{1}H NMR, however, this fraction was shown to be a mixture containing the hydroxysilane **124** and the acid **106** (**Scheme 22**). Our final attempt to acylate the indole system involved the method described by Grummel,[138] using butyryl chloride. Thus, the methyl ester **110a** was treated with butyryl chloride and NaH in DMF at ($0^{O}C$-room temperature) for 24h and the 3-indolyl amide **125** was isolated in low yield as a brown oil following flash chromatography (**Scheme 22**).



Cpd	n	R
110a	1	Me
111b	2	TBDMS

Scheme 22: Attempted synthesis of compounds **123** and **125**.

This method was then used to synthesize the gluconamide derivatives **116a-c** (**Scheme 23)** which were isolated, following flash chromatography, as brown oils in very low yields (0.1-0.5%). Compound **116c** was also purified by HPLC and the ^{1}H, ^{13}C NMR and HSQC spectra are illustrated in **Figures 14, 15** and **16**, respectively. The disappearance of the indole NH singlet at *ca.* 8ppm and the presence of the methoxy singlet at *ca.* 3.7ppm in the ^{1}H NMR spectrum (**Figure 14**), confirms *N*-acylation. Further support for the proposed structure is provided by the downfield shift of the indole protons from *ca.* 7.1-8.1 to 7.3-8.4ppm, while the ^{13}C NMR spectrum (**Figure 15**) shows extra signals in the carbonyl (*ca.* 165-174ppm) and acetyl methyl (*ca.* 22ppm) regions.



	n
a	1
b	2
c	3

Scheme 23: Synthesis of 3-indolyl gluconamide derivatives

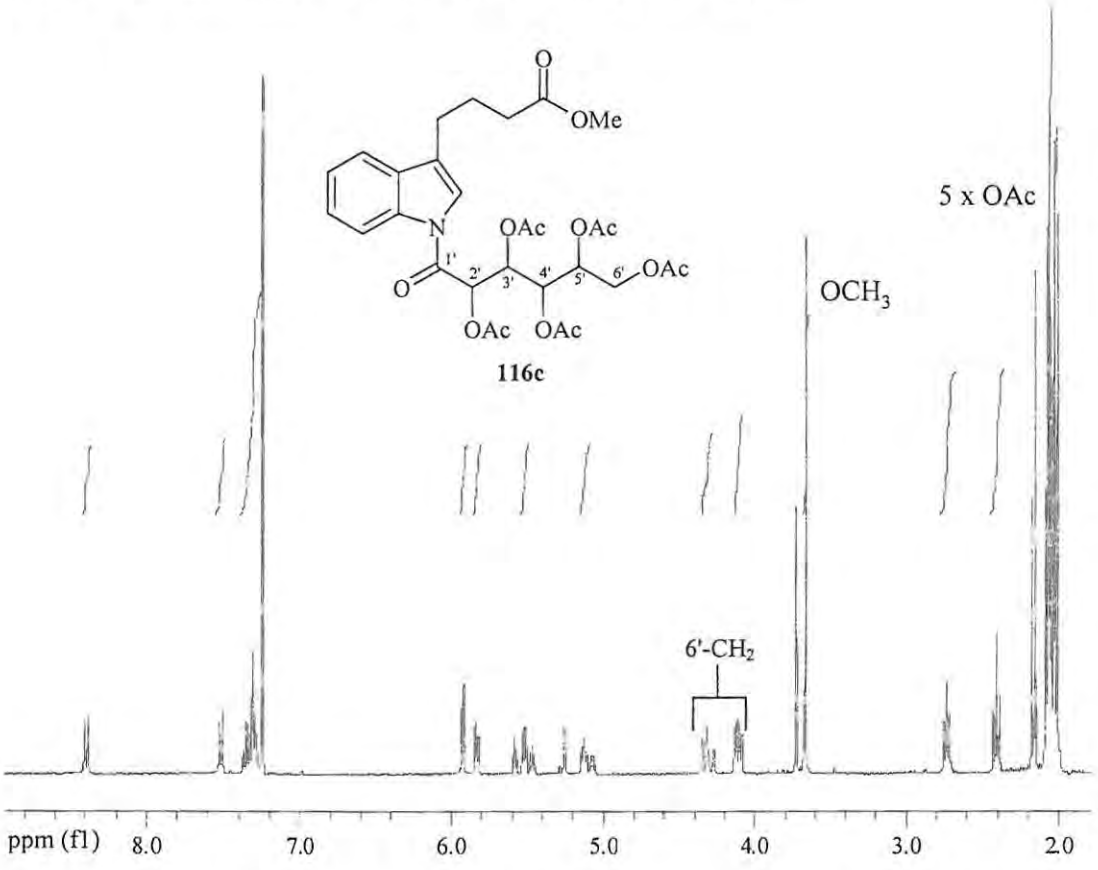


Figure 14: 400MHz ^{1}H NMR spectrum of the gluconamide derivative **116c** in $CDCl_3$.

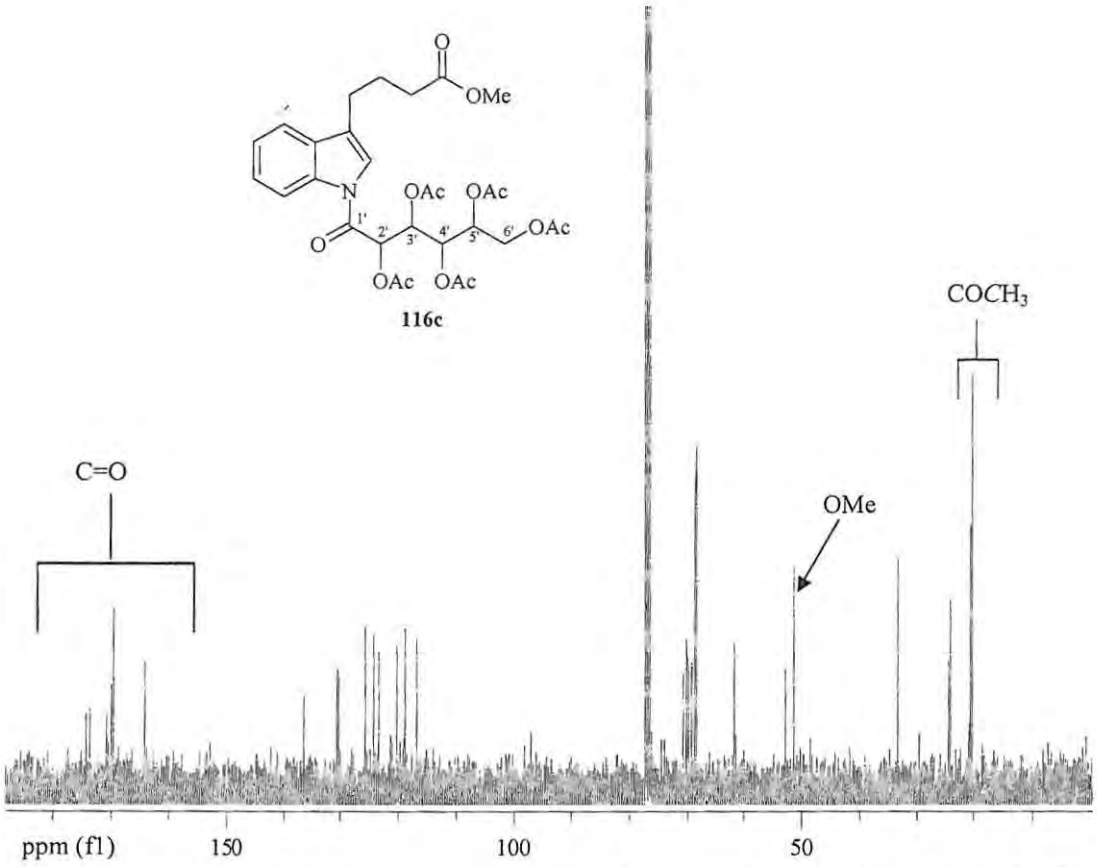


Figure 15: 100MHz ^{13}C NMR spectrum of the gluconamide derivative **116c** in $CDCl_3$.

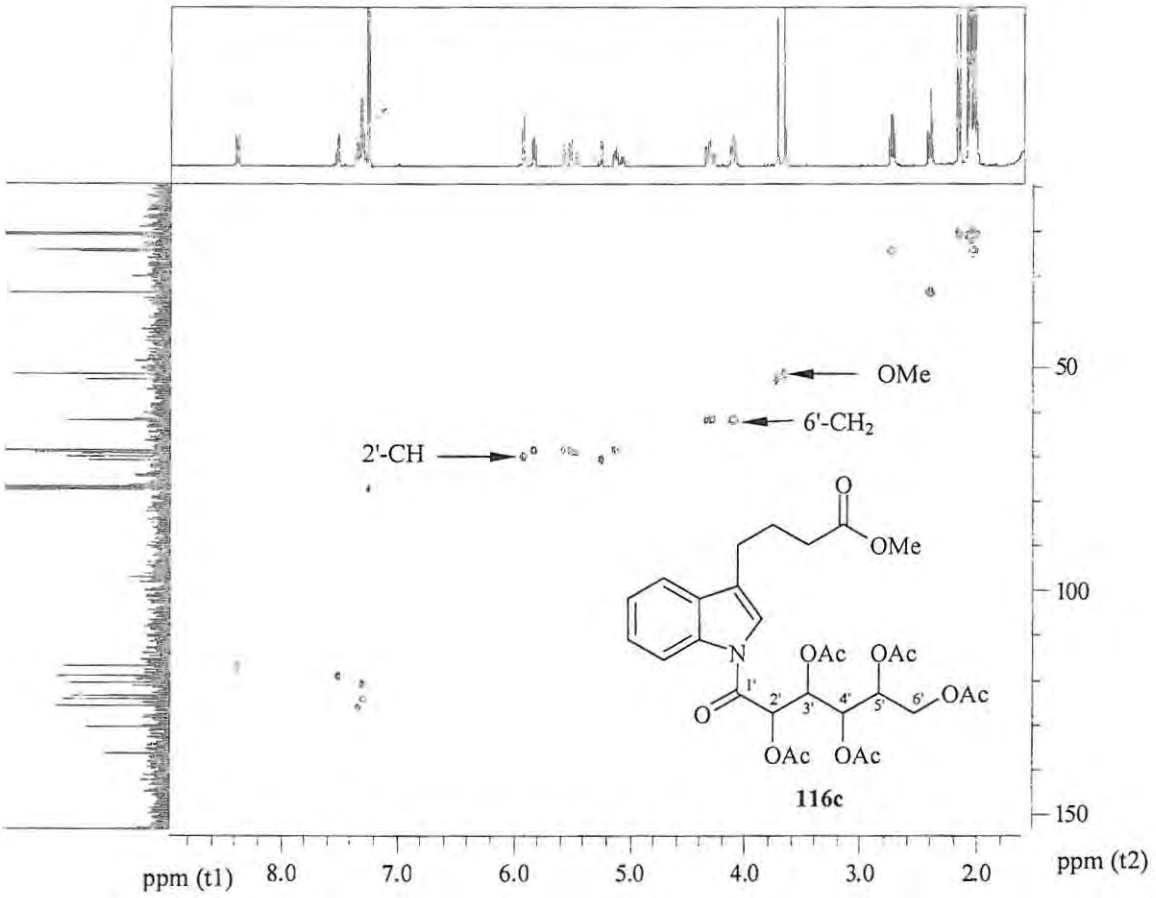
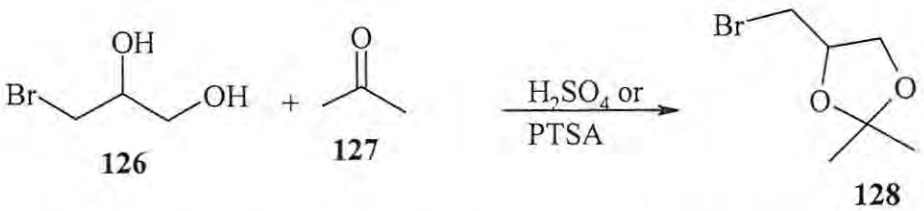


Figure 16: HSQC spectrum of the gluconamide derivative **116c** in $CDCl_3$.

2.1.2.3 Alkylation of the 3-Indolylalkanoic Acid Esters

Due to the difficulties encountered in introducing the amide linkage it was decided to explore an amine linkage using the alkyl halide derivatives, 3-bromo-1,2-dihydroxypropane **126** and 2-(bromomethoxy)ethanol. The hydroxy groups in both of these compounds needed protection to avoid undesirable side reactions. Protection of the diol **126** as an acetal was attempted using two of the known literature methods.[151,152] In the first, dry acetone **127** and the diol **126** were refluxed in the presence of catalytic amounts of *p*-TSA using a Dean-Stark apparatus. The acetal **128** was isolated, as an oil in poor yield (18%), after distillation (**Scheme 24**). In the second method, a mixture of the diol **126** and acetone was stirred at room temperature for 36h in the presence of catalytic amounts of concentrated sulphuric acid. After solvent extraction with diethyl ether, the acetal **128** was isolated as a colourless oil in good yield (70%; **Scheme 24**). ^{1}H NMR analysis revealed two sharp singlets at *ca.* 1.3 and 1.4ppm corresponding to the diastereotopic methyl groups, while the ^{13}C NMR spectrum exhibits three extra signals at *ca.* 25, 27 and 110ppm, which were assigned to the two methyl groups and the quaternary acetal carbon respectively.

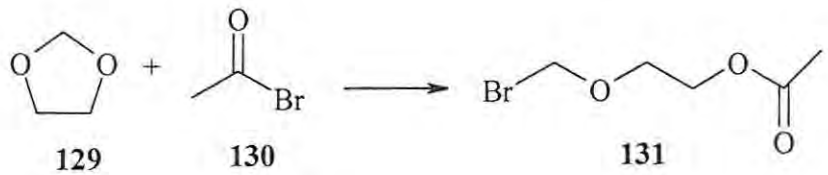


126 **127** **128**

Scheme 24: Protection of 3-bromo-1,2-dihydroxypropane **126** as an acetal

Introduction of the 2-methoxyethanol moiety to guanine has produced an important acyclic nucleoside, acyclovir, which is a very potent inhibitor of the herpes virus.[153,154] The halogenated 2-methoxyethanol derivative **131**, required for *N*-alkylation of the indole systems, can be accessed from the cyclic acetal 1,3-dioxolane as described by Robins[155] (**Scheme 25**). Reaction of 1,3-dioxolane **129** with acetyl bromide **130** at 0^{O}C and purification of the crude material by distillation afforded (2-acetoxyethoxy)methyl bromide **131** in good yield (79%). Formation of compound **131** was verified ^{1}H NMR

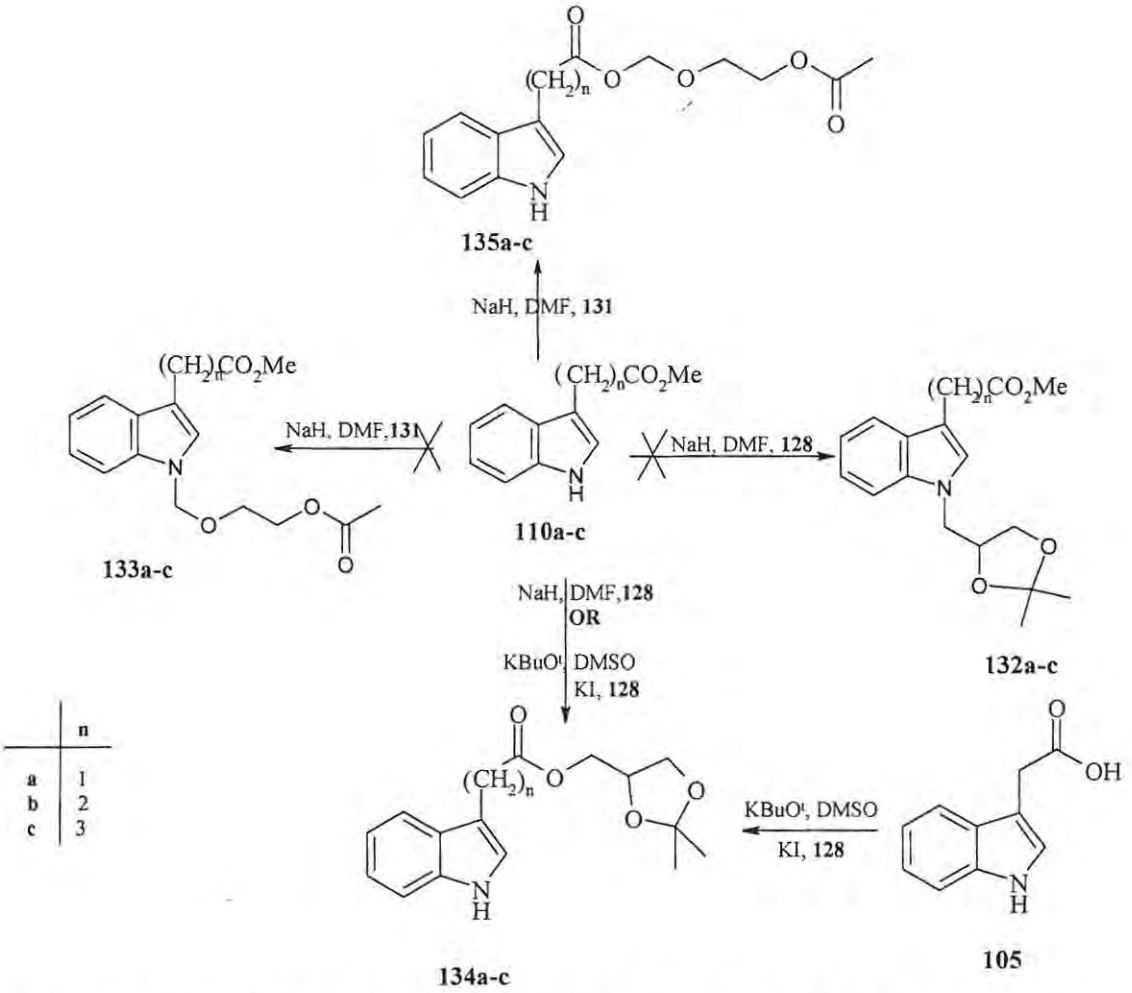
and ^{13}C NMR spectroscopy. The ^{1}H NMR spectrum shows two triplets at 3.8 and 4.3ppm corresponding to the two pairs of chemically non-equivalent ethylene protons and two singlets at *ca* 2.1 and 5.7ppm assigned to the acetyl and bromomethyl protons, respectively. The ^{13}C NMR spectrum shows a distinct carbonyl signal at 171ppm.



129 **130** **131**

Scheme 25: Synthesis of (2-acetoxyethoxy)methyl bromide **131**

The literature[156,157] reveals that alkylation of the 3-indolylalkanoic acids can be achieved in relatively good yields when NaH or potassium *tert*-butoxide ($KOBu^{t}$) is used to deprotonate and thus enhance the nucleophilicity of the indole nitrogen. Alkylation of esters **110a-c** with the bromoketal **128** and with (2-acetoxyethoxy)methyl bromide **131** was attempted following the method described by Faul *et al.*[156] The esters **110a-c** were treated with 1.5 equivalents of NaH and each of the alkyl bromides **128** and **131** in DMF at temperatures of $0^{o}C$ to room temperature (**Scheme 26**). Purification of the reaction mixtures by flash or radial chromatography gave unreacted starting material and products shown by ^{1}H NMR analysis to retain the indole NH moiety. In fact it was apparent that, in each case, the methyl group of the ester moiety had been replaced by the respective alkyl groups to afford the corresponding esters **134** and **135**. The ^{13}C NMR spectrum of ester **135c** (**Figure 17**) exhibits two carbonyl signals and the presence of additional signals corresponding to the ketal group.



Scheme 26: Attempted synthesis of *N*-alkyl indolylalkanoic acid derivatives

The reactions were then repeated following the method described by Winterfeldt *et al.*[157] Ester **110a** was treated with KOBu^t, and three equivalents of ketal **128** and KI in dry DMSO. After chromatographic purification, the isolated product was analysed by NMR and mass spectrometric methods. The ^{1}H NMR data revealed a singlet at 8.1ppm due to the indole NH proton and no significant changes in the aromatic region (*ca.* 7.1-8.1ppm; **Figure 18**), suggesting that *N*-alkylation had not occurred. The absence of the methoxy singlet at *ca.* 3.7ppm suggests *O*-alkyl cleavage of the methyl ester **110a** and formation of the ester **134a**. This conclusion was supported by the ^{13}C NMR spectrum, which exhibited 16 signals (instead of the 17 signals expected for compound **132a**) and the absence of the methoxy signal at *ca.* 51ppm (**Figure 19**). The existence of compound

134a was also confirmed by HMBC spectrum (**Figure 20**) which shows a correlation between the multiplet assigned to the 6'-methylene protons of the acetal substituent and the carbonyl carbon. Moreover, the low-resolution mass spectrometric data indicated a molecular mass corresponding to that of ester **134a** (14amu lower than the expected mass for **132a**).

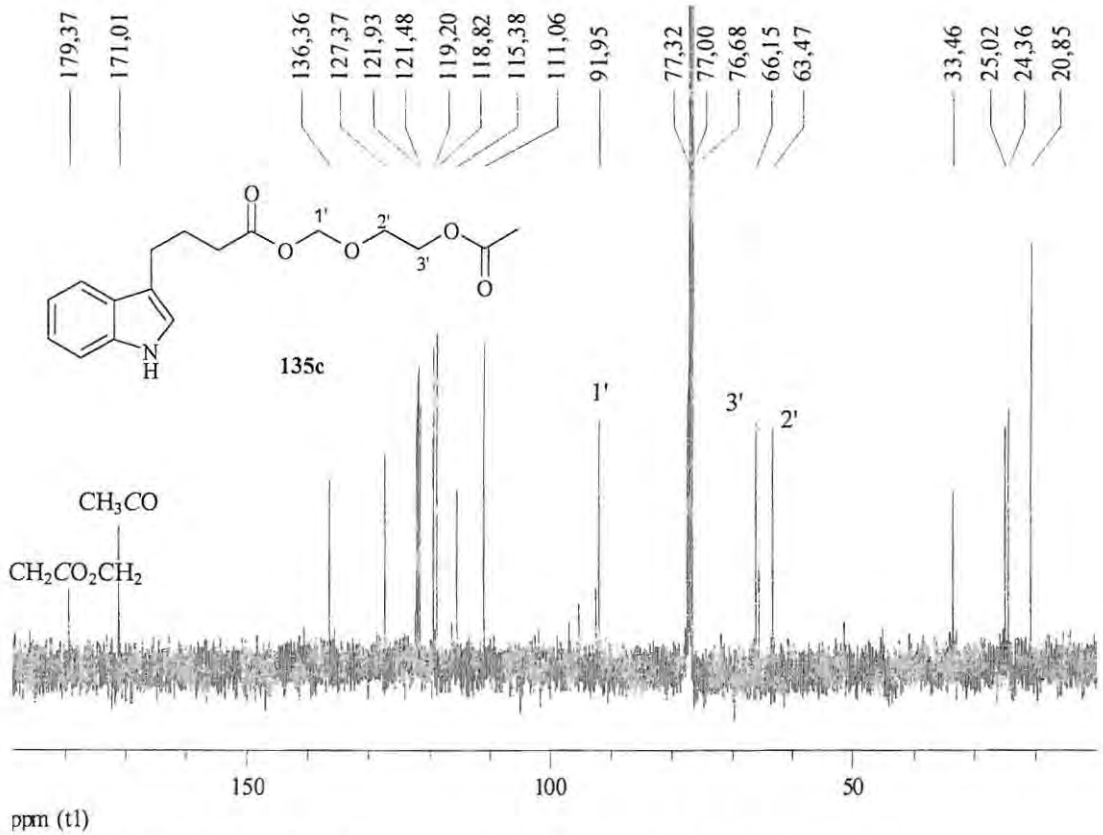


Figure 17: 100MHz ^{13}C NMR spectrum of compound **135c** in $CDCl_3$

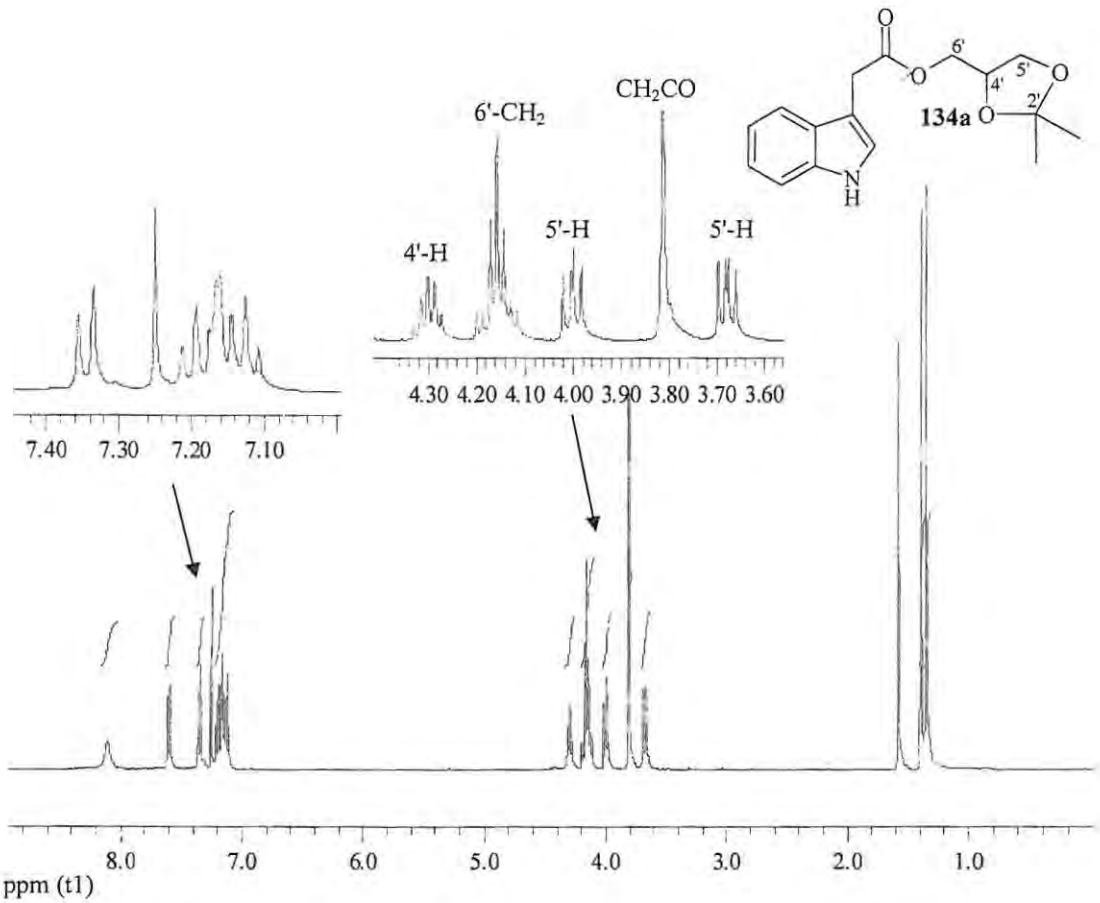


Figure 18: 400MHz ^{1}H NMR spectrum of compound **134a** in $CDCl_3$.

Finally the free acid **105** was treated with two equivalents of KOBu^t and three equivalents of the ketal **128** and KI in DMSO. After chromatographic purification only the ketal **128**, the acid **105** and compound **134a** were isolated. Clearly, reaction of the acid with KOBu^t affords the carboxylate anion, which then attacks the bromoketal **128** displacing the halide to afford the ester **134a**. It is, however, unclear as to how the methyl ester precursors **110a-c** afford the corresponding esters **134** and **135**. It is possible that nucleophilic attack by potassium butoxide or hydroxide ion at the methyl carbon of the ester displaces the nucleophilic carboxylate anion *via* *O*-alkyl cleavage. However, the presence of hydroxide ions would require the presence of water in the dried reactants and solvent.

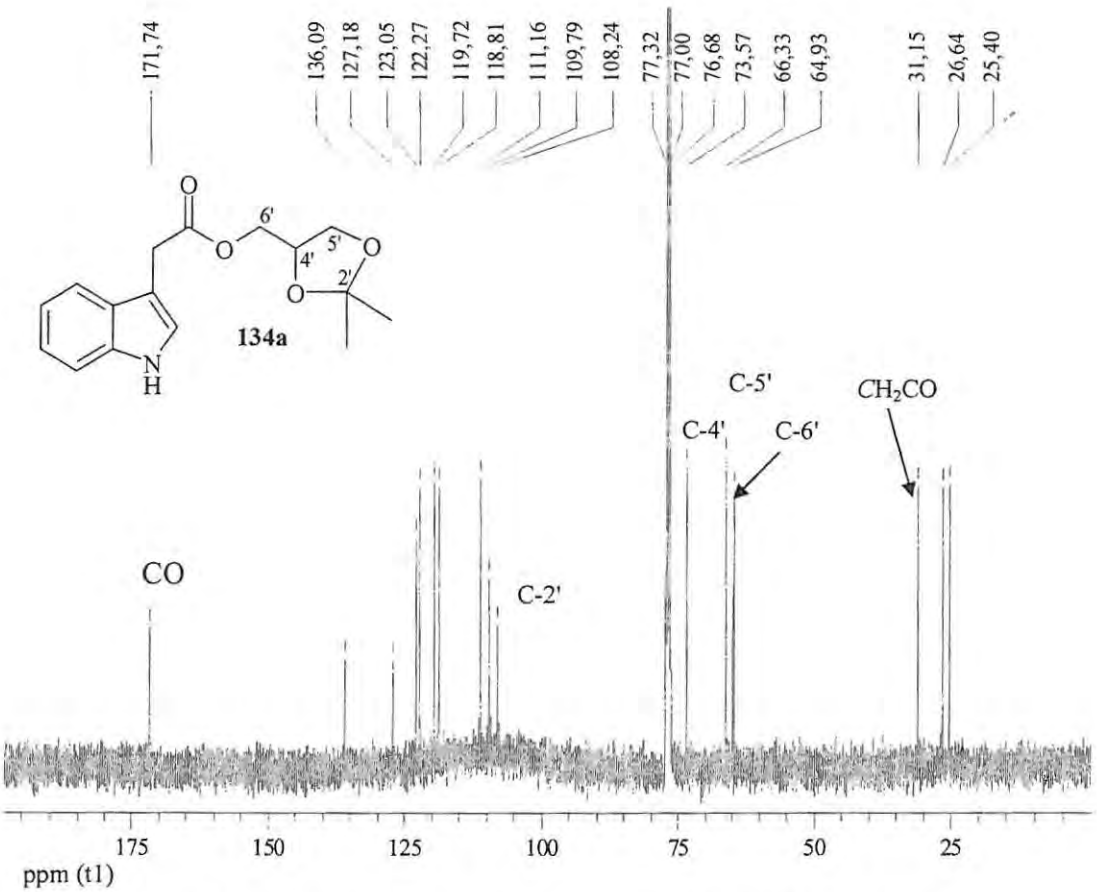


Figure 19: 100MHz ^{13}C NMR spectrum of compound **134a** in $CDCl_3$.

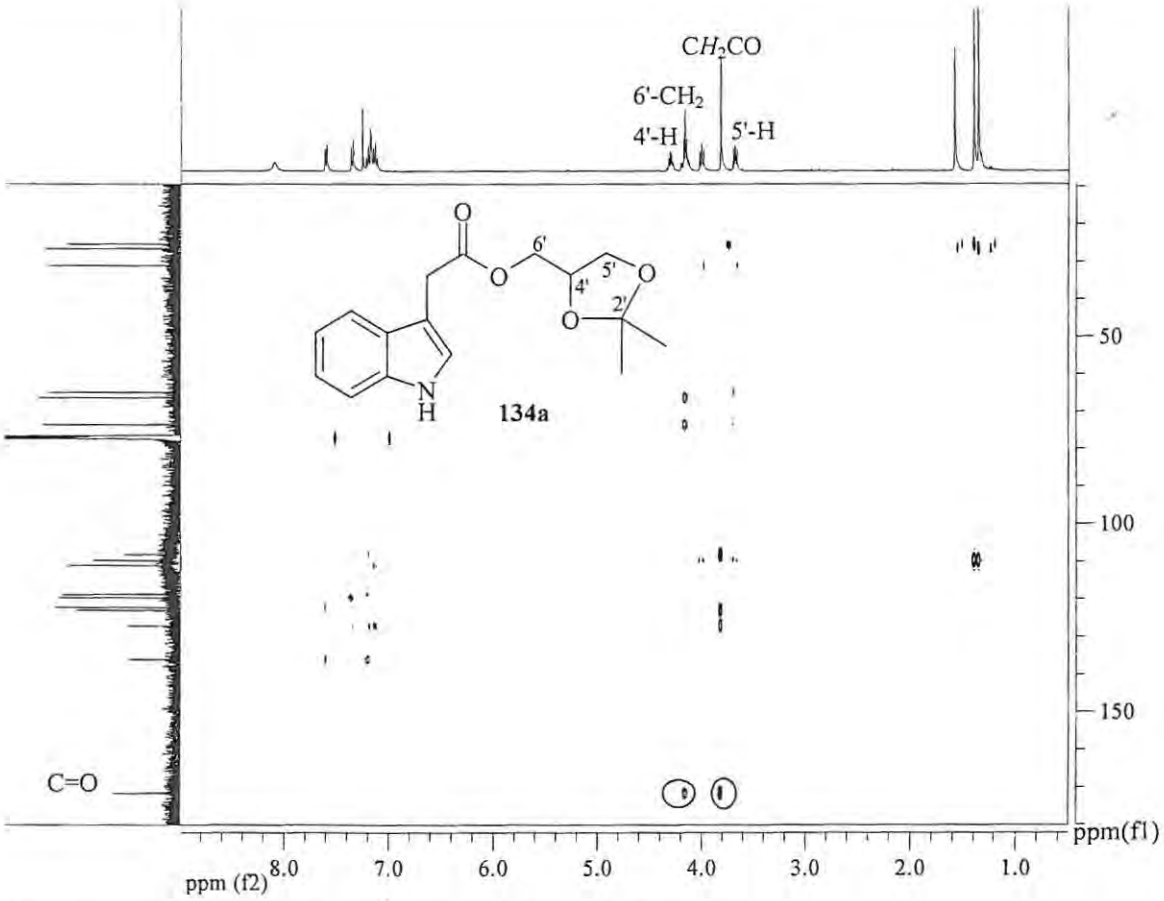
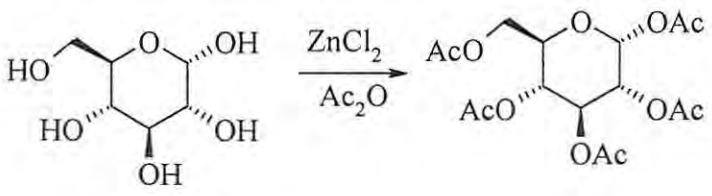


Figure 20: HMBC spectrum of compound **134a** in $CDCl_3$.

2.1.2.4 Glycosylation of Indolylalkanoic Acid Derivatives

Given the difficulties encountered in introducing acyclic polyoxy moieties, attempts were made to link the cyclic glucose derivative **137** to the indole-3-alkanoic esters **110**. Interestingly individuals who have a defective tryptophan metabolism also produce the plant indole glucoside, indican,[158] which contains an *O*-substituted glycoside at position 3 of the indole ring, as a by-product. Glucose has several hydroxyl groups and under basic conditions it is able to undergo various degradation reactions; hence protection of the hydroxyl groups is needed to ensure regioselective derivatisation. Protection of the hydroxyl groups in α-D-glucopyranose **136** was effected by acetylation, following an established procedure[159] (**Scheme 27**). α-D-Glucopyranose **136** was stirred in acetic

anhydride in the presence of anhydrous $ZnCl_2$, and the isolated product was confirmed, by spectroscopic analysis, to be 2,3,4,5,6-penta-*O*-acetyl-α-D-glucopyranose **137**.

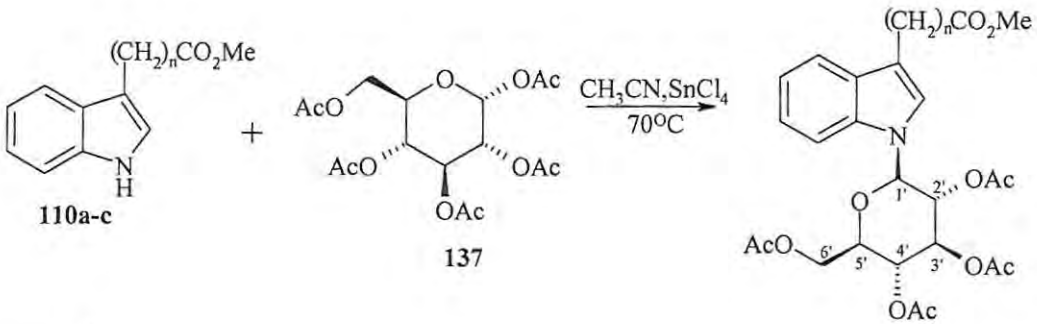


136 **137**

Scheme 27: Acetylation of glucose

A search of the literature[104,158] revealed that *N*-glycosylation of indole and its derivatives is generally achieved through the "indoline-indole" method or by reaction of an indole derivative with an anhydrosugar in the presence of NaH.[158,160] Glycosylation was carried out using a modification of the procedures described by Jain *et al.*[125] and Kraybill *et al.*[161] The esters **110a-c** were each treated with the glucopyranose derivative **137** at 70°C in the presence of tin tetrachloride and acetonitrile (**Scheme 28**). This is a Lewis acid-catalysed nucleophilic displacement of acetate at the anomeric centre (C-1) by the indolyl nitrogen. The resulting crude product was purified by flash chromatography and the fractions that produced ^{1}H NMR spectra corresponding to the expected products **138a-c** were further purified by HPLC. The pure products were isolated in disappointing yields (7-26%) and were characterized spectroscopically. The ^{1}H NMR spectrum illustrated for compound **138b** in **Figure 21** shows the presence of the methoxy signal at *ca.* 3.7ppm; the absence of a broad NH singlet at *ca.* 8ppm confirms binding through the indole nitrogen. The four singlets at *ca.* 1.7-2.1ppm integrating for 12 protons were assigned to the acetoxy protons and the upfield doublet at δ 5.6ppm was assigned to the C-1' methine proton. The ^{1}H NMR spectrum also shows two double doublets at 4.1 and 4.3ppm integrating for one proton each, and these were assigned to the diastereotopic methylene protons of the glucopyranose substituent. These assignments are supported by the HSQC data illustrated for the analogue **138a** in **Figure 22**, which shows that the two double doublets are attached to a single carbon. The ^{13}C NMR spectrum shows the expected four methyl

carbon signals at *ca.* 20ppm and the five carbonyl carbon signals at *ca.*170ppm. The large coupling constants ($J_{1,2}$ = 9.2Hz) observed as a result of trans diaxial coupling in a β-glucoside, confirms the formation of a β-glucosides.



Cpd	n
a	1
b	2
c	3

Scheme 28: Glycosylation of esters **110a-c**

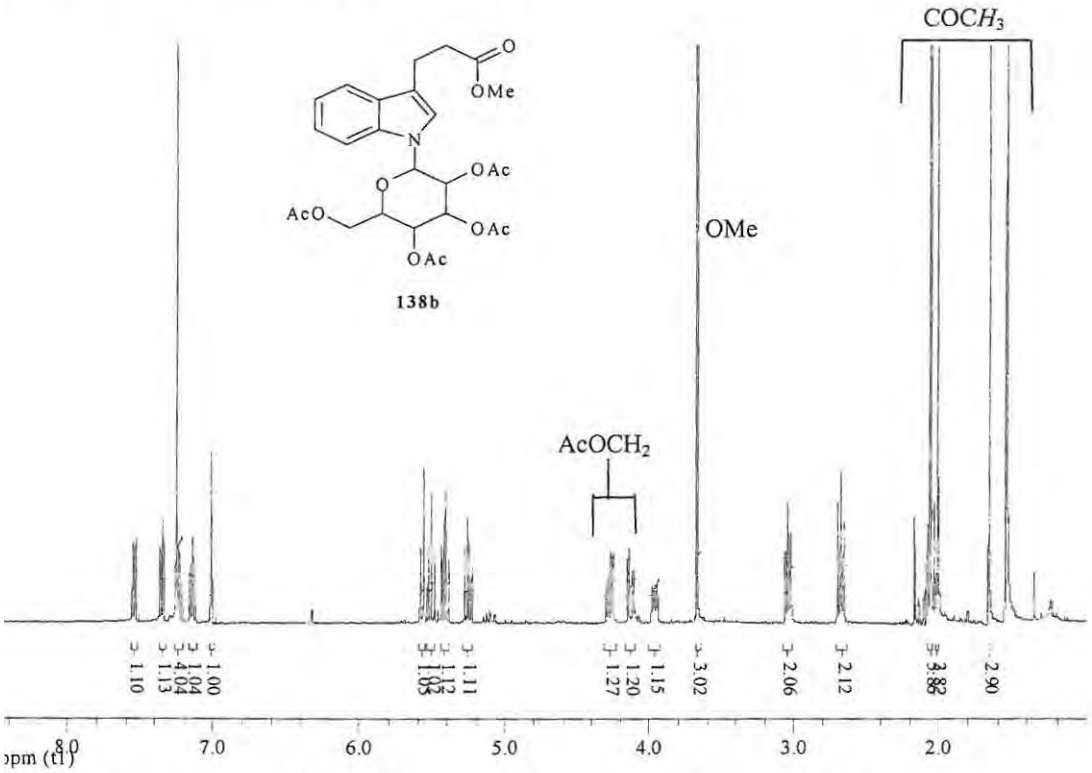


Figure 21: 400MHz ^{1}H NMR spectrum of compound **138b** in $CDCl_3$.

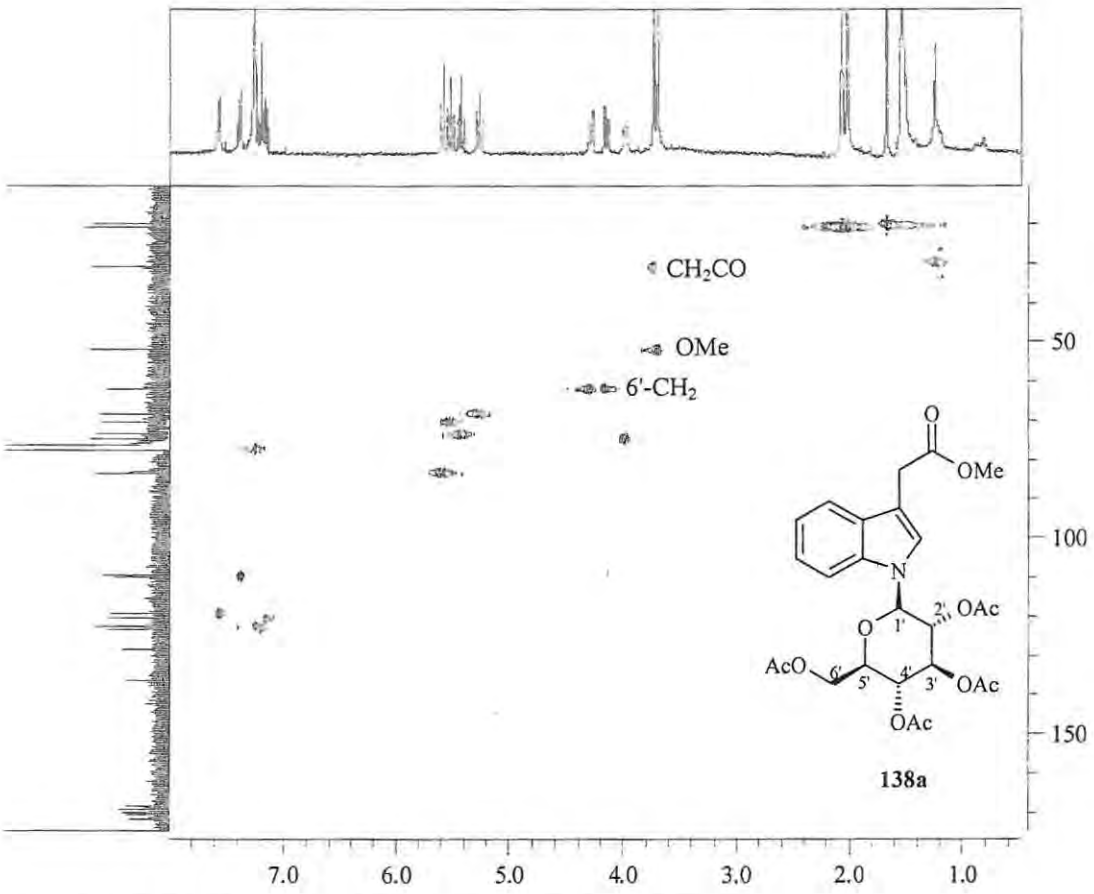
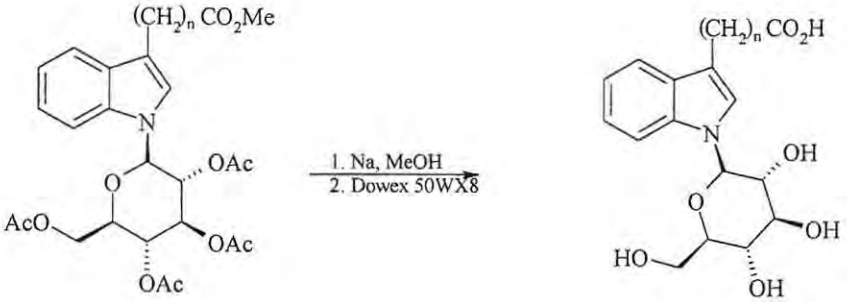


Figure 22: HSQC spectrum of compound **138a** in $CDCl_3$

The final steps in the synthesis of the indolylalkanoic acid derivatives involved removal of the acetyl protecting groups and phosphorylation the 6'-hydroxyl group to form mono or triphosphate esters. Methanolysis[162-164] of the acetyl groups was effected by treating compounds **138a-c** with sodium methoxide, prepared *in situ*, in methanol and then neutralizing the basic mixture using an ion-exchange resin (Dowex 50WX8) (**Scheme 29**),[162] The crude material was separated, in each case, by HPLC to afford the pure compounds **139a-c** in reasonable yields (50-70%). The structures of these compounds were verified by NMR spectroscopy. The absence of the of the acetyl methyl signals at *ca.* 1.7-2.1ppm and the methoxy signal at *ca* 3.7ppm in the ^{1}H NMR spectra (illustrated for compound **139c** in **Figure 23a**) confirmed the absence of the ester groups in compounds **139a-c**. Further confirmation was provided by the ^{13}C NMR spectra (illustrated for compound **139c** in **Figure 23b**) which showed the absence of the methoxy carbon signal at *ca.* 51ppm as well as the methyl and carbonyl carbon signals

corresponding to the four acetyl groups at *ca.* 20 and 175ppm, respectively. The indolyl acid carbonyl signal was shifted downfield to *ca.* 182ppm due to the loss of the methyl ester. Due to time constraints phosphorylation of the indolyl system was not investigated.



138a-c **139a-c**

Cpd	n
a	1
b	2
c	3

Scheme 29: Hydrolysis of the glucosyl esters **138a-c**.

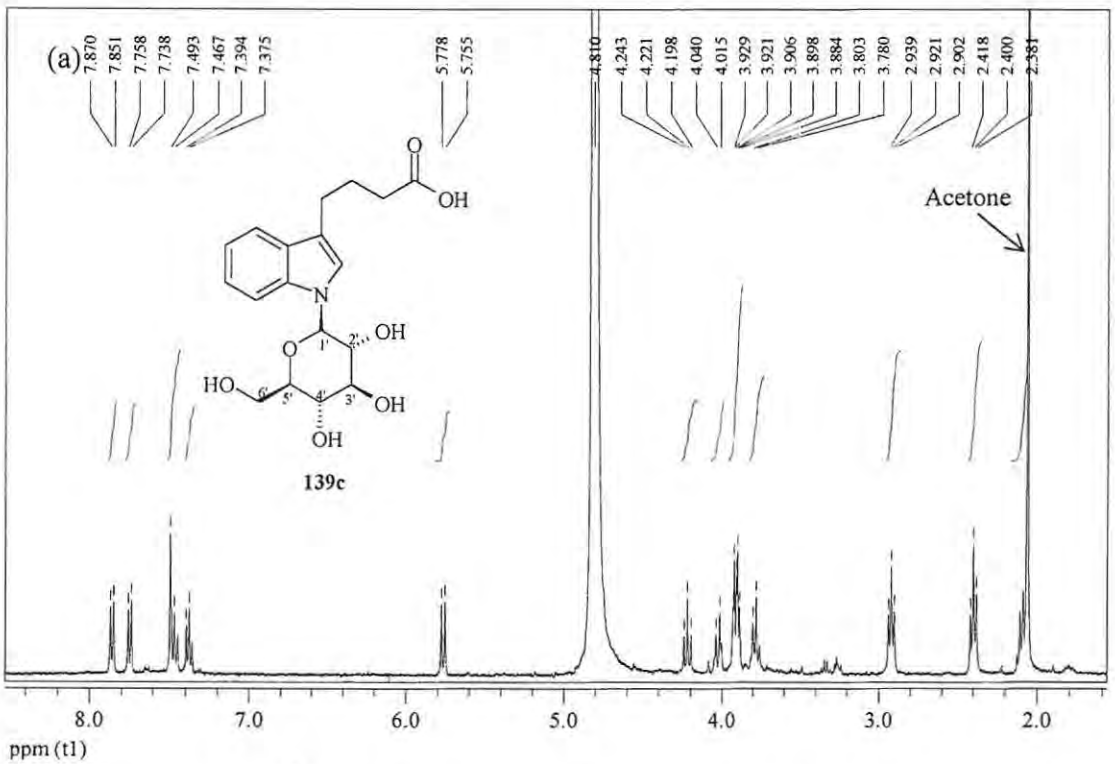


Figure 23a: 400 MHz ^{1}H NMR spectrum of compound **139c** in D_2O.

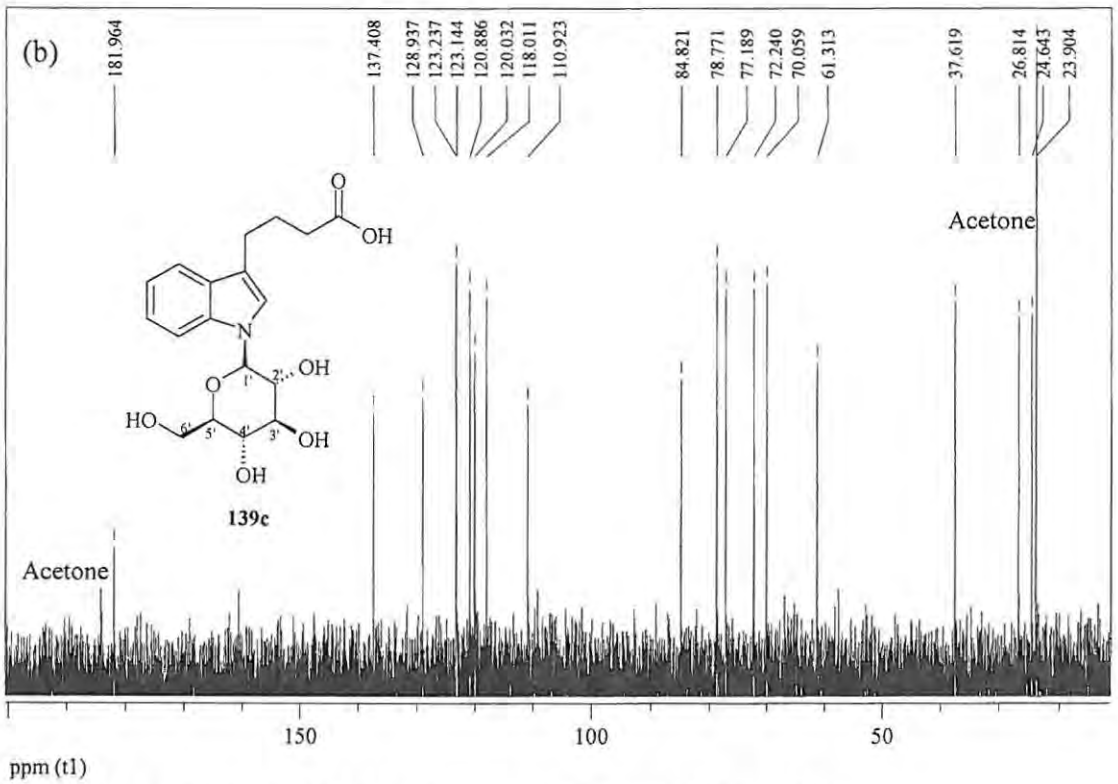


Figure 23 b: 100MHz ^{13}C NMR spectra of compound **139c** in D_2O.

2.1.2.5 Structural comparison of the 3-indolylalkanoic acid derivatives with ATP

Selected 3-indolylalkanoic acid derivatives, designed as ATP analogues, were explored using a molecular dynamics simulations routine to obtain their minimum energy conformers. The ten highest and lowest energy conformers were selected and energy minimised to find the global minimum structure.

In order to determine the similarities and differences between the global minimum structures of these analogues and ATP, the structures were manually aligned, by matching the atoms of the heterocylic base in each case. The energy minimized structures of the 3-indolylalkanoic acid derivatives **116c**, **134a**, **135c**, **138b, 139c** and ATP (**Figure 24**) were aligned with ATP as illustrated in **Figure 25**. In general, however, the aligned structures of 3-indolylalkanoic acid derivatives exhibit relatively poor structural

homology with ATP. Of course, there is no necessary correlation between the global minimum and enzyme binding conformations of ATP and any potential GS ligands. The issue of binding conformations will be explored in the Docking Studies section (Section 2.3.1).

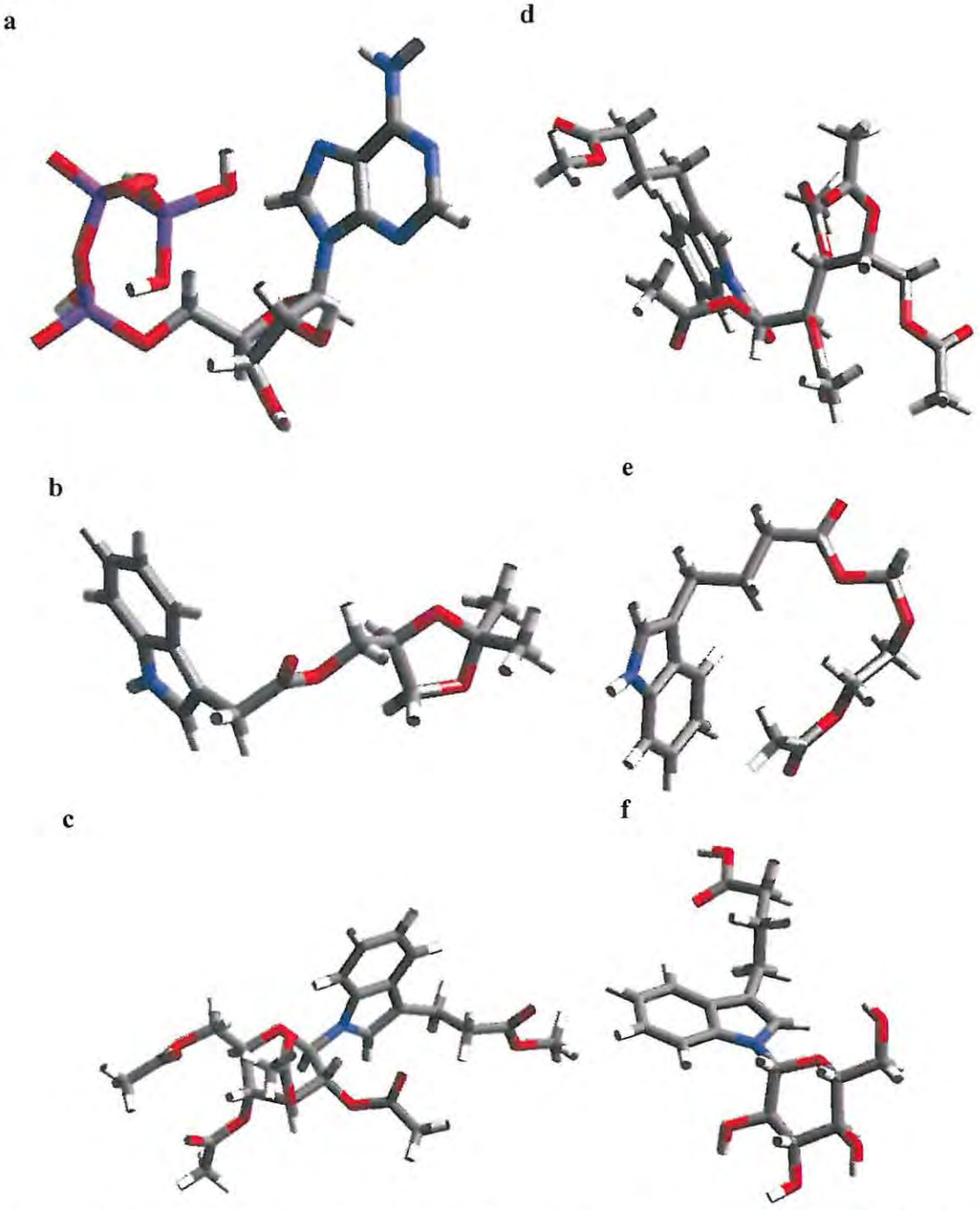


Figure 24: Energy-minimized structures of: a) ATP and compounds **134a** (b), **138b** (c), **116c** (d), **135c** (e) and **139c** (f)

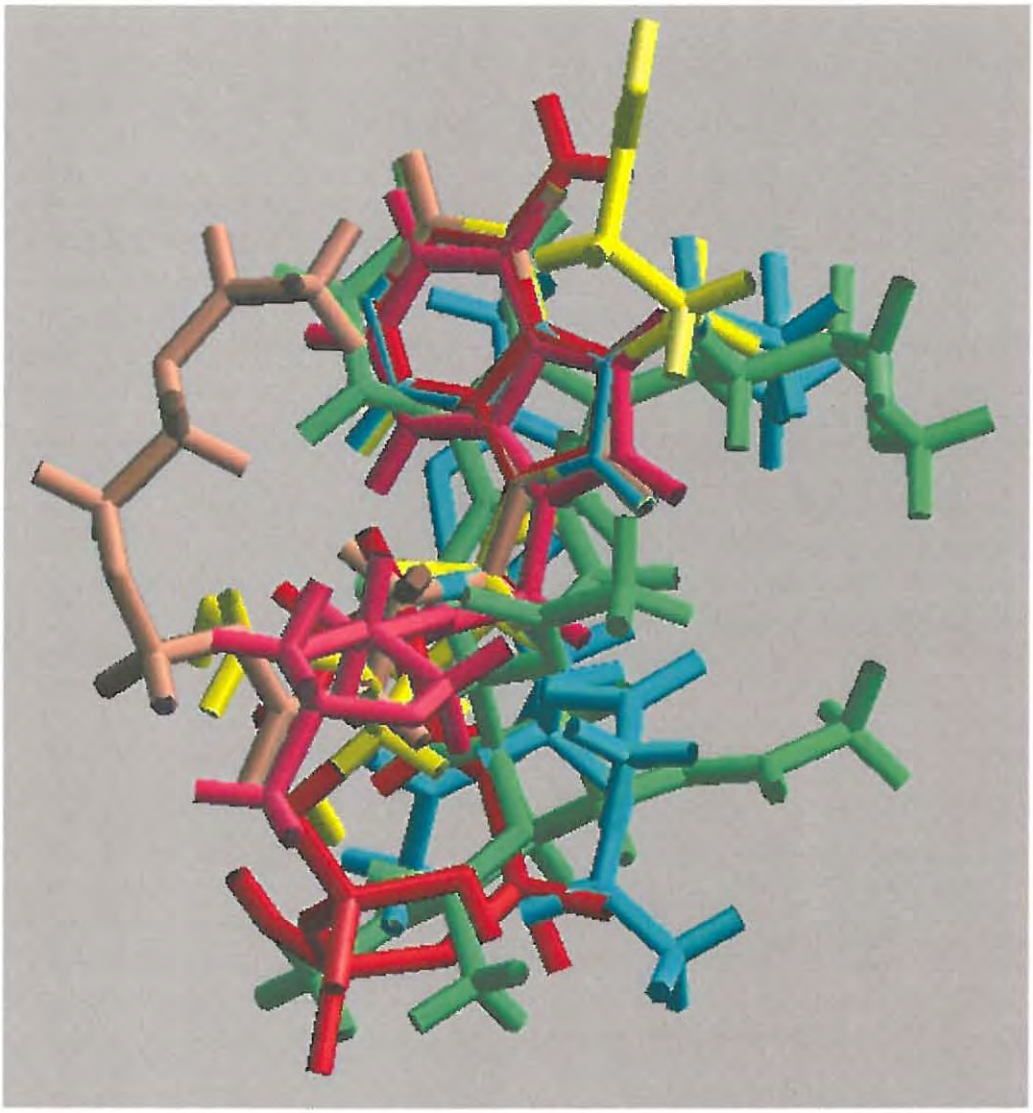


Figure 25: Alignment of the energy-minimized structures of compounds **116c** (light-green), **134a** (pink), **135c** (light-brown), **138b** (light-blue) and **139c** (yellow) and ATP (red).

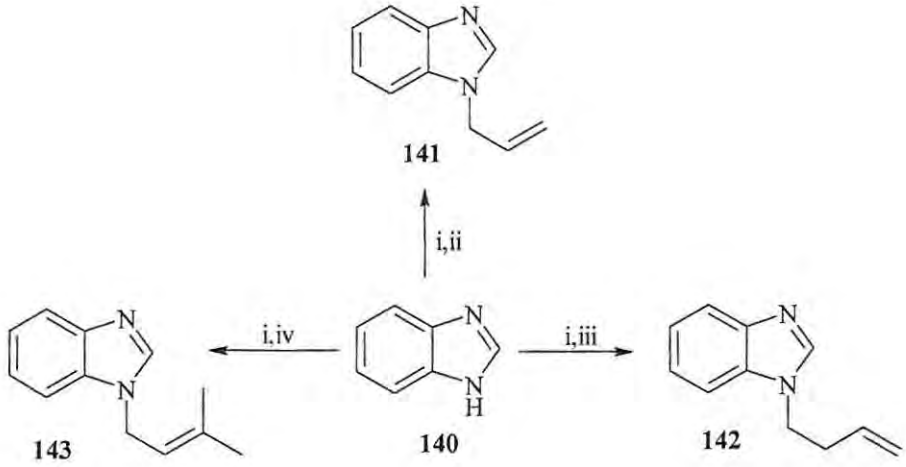
2.1.3 Synthesis of Benzimidazole Derivatives

Interest in the synthesis of benzimidazole nucleosides and nucleotides arose from the discovery of 5,6-dimethyl-l-(α-D-ribofuranosyl)benzimidazole (α-ribazole) as a component of the chemical structure of vitamin B_{12},[158] a vitamin which plays an important role in metabolic processes within the human body. Benzimidazole **140** has been reported to inhibit some phosphorylase enzymes in plants, *e.g.* wheat,[158] and, as indicated in Section 1.6.2, a number of benzimidazole derivatives are pharmacologically active.

The natural nucleoside ATP has an amino substituent at position 4, which is capable of hydrogen-bonding to amino acid residues in the receptor site. Benzimidazole lacks this substituent and, hence cannot undergo normal Watson-Crick hydrogen-bonding. However, this does not preclude the possibility that benzimidazole derivatives may fit well into the GS receptor cavity and bind sufficiently strongly to inhibit the normal functions of the enzyme.

2.1.3.1 Alkylation of Benzimidazole

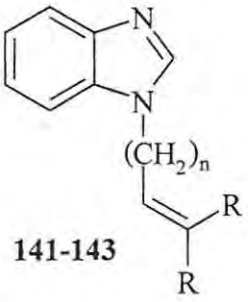
In our attempts to synthesize benzimidazole derivatives, *N*-alkylation was investigated using a variety of alkenyl bromides. Thus, benzimidazole **140** was treated with 4-bromo-2-methylbut-2-ene, 4-bromobutene and allylbromide, in the presence of sodium hydride, following the procedure described by Tan *et al.*[165] (**Scheme 30**). Purification of the crude material by flash chromatography afforded the alkenylbenzimidazoles, *N*-allylbenzimidazole **141**, *N*-(but-3-enyl)benzimidazole **142** and *N*-(3-methylbut-2-enyl) benzimidazole **143** in moderate to excellent yields 43-96%, and their structures were confirmed by NMR spectroscopy (**Table 3**).



Scheme 30: Alkylation of benzimidazole

Reagents and Reaction Conditions: i) NaH, THF, 0°C, 15min; ii) allyl bromide, 30min reflux; iii) 4-bromobutene, 4h reflux; iv) 4-bromo-2-methyl-2-butene, 40min reflux.

Table 3: Comparative yields and selected NMR data for the alkenylbenzimidazoles **141-143**

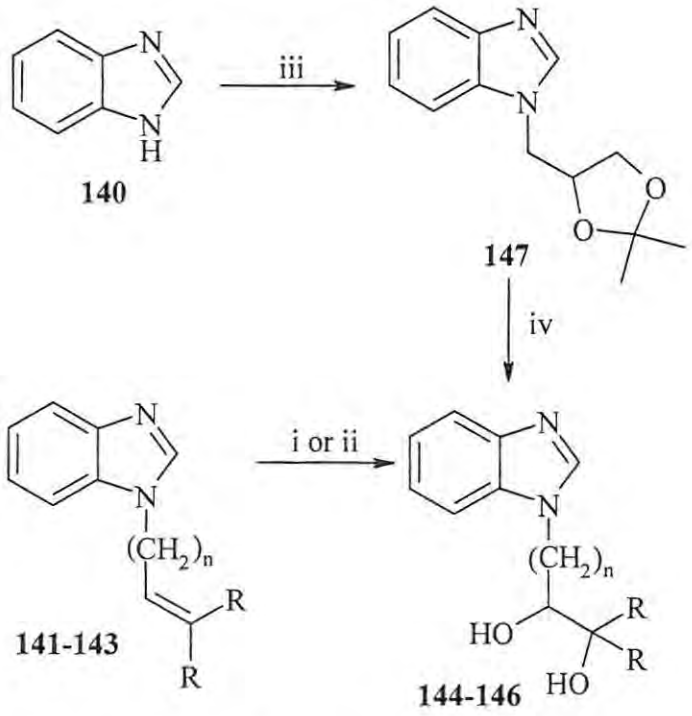


141-143

Cmpd.	n	R	Yield[a]/ %	Time /h	δ_H[b] CH=C*H*$_2$	δ_H[b] C*H*=CR$_2$	δ_C[b] *C*H=CR$_2$	δ_C[b] CH=*C*R$_2$	δ_C[b] NCH$_2$
141	1	H	96	0.5	5.15 (2xd)	5.91 (m)	131.7	118.3	47.1
142	2	H	43	4	4.98 (m)	5.68 (m)	133.5	118.2	44.4
143	1	CH_3	67	0.67	-	5.38 (t)	118.1	133.8	42.9

[a] Isolated material. [b] In $CDCl_3$; signal multiplicity in parenthesis.

The next step involved oxidation of the alkenyl derivatives **141-143** in order to introduce the hydroxyl groups. This oxidation was attempted with $KMnO_4$[166] and with cetyltrimethylammonium permanganate (CTAP), which was prepared from $KMnO_4$ and cetyltrimethylammonium bromide.[167] Oxidation of compound **141** with $KMnO_4$ was carried out in water at 0^oC (**Scheme 31**) and the crude material was purified by radial chromatography. The desired product, *N*-(2,3-dihydroxypropyl)benzimidazole **144**, was isolated as colourless oil in 15% yield. This low yield was attributed to the fact that compound **144** is water-soluble and some of the product might have been discarded with the solvent. To improve the yields the oxidation was carried out using CTAP, which is soluble in organic solvents.[167] The compounds **141-143** were therefore treated with CTAP in Bu^{t}OH for several hours at 20^oC, and the crude material was purified by radial chromatography. The pure dihydroxy compounds **144-146** were isolated in poor to good yields (26-77%; **Table 4**) and their structures were verified by spectroscopic methods.

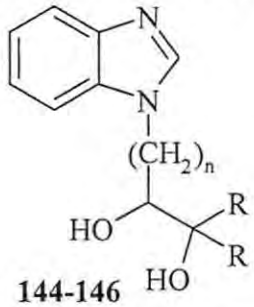


Cpd	n	R
141/144	1	H
142/145	2	H
143/146	1	CH_3

Scheme 31: Synthesis of compounds **143-146**

Reagents and conditions: i) Bu^{t}OH, $KMnO_4$, H_2O, 0^oC then NaOH, r.t; ii) CTAP, ^{t}BuOH, H_2O, 20^oC then $CHCl_3$, NaOH; iii) NaH, THF, 0^oC then **128**, 18h reflux; iv) 75% AcOH, 1h reflux.

Table 4: Comparative yields of compounds **144-146** (**Scheme31**).



144-146

Compound	n	R	Time/h	Yield[a]/%
144	1	H	9	77
145	2	H	6.5	41
146	3	CH_3	6.5	26

[a] Isolated material

Table 5: ^{1}H- and ^{13}C NMR chemical shift data for the dihydroxy compounds **144-146** in DMSO-d_6

Cmpd.	n	R	δ_H			δ_C		
			CR_2OH	$CHOH$	NCH_2	CR_2OH	$CHOH$	NCH_2
144	1	H	4.83	5.06	4.13 & 4.35	70.0	63.1	47.4
145	2	H	4.60	4.83	4.34	65.7	68.3	41.0
146	1	CH_3	4.65	5.11	4.02 & 4.52	71.0	75.8	46.6

The ^{1}H NMR spectra of compounds **144** and **145** (illustrated for compound **144** in **Figure 26a**) show a doublet and a triplet in the region, 4.6-5.1ppm, each signal integrating for one proton. These signals were assigned to the hydroxyl protons on the methine and methylene carbons of the alkyl side chains of compounds **144** and **145**, respectively. The ^{1}H NMR spectrum of compound **146** however, shows a singlet and a doublet at 4.6 and 5.1ppm for the hydroxyl groups on the methylene and methine carbons, respectively. The methyl groups in compound **146** are also diastereotopic and resonate as two singlets at 1.1 and 1.2ppm. The COSY spectra illustrated for compound **144** in **Figure 26b**) confirmed these assignments revealing correlations between the hydroxyl protons and the methylene and methine protons.

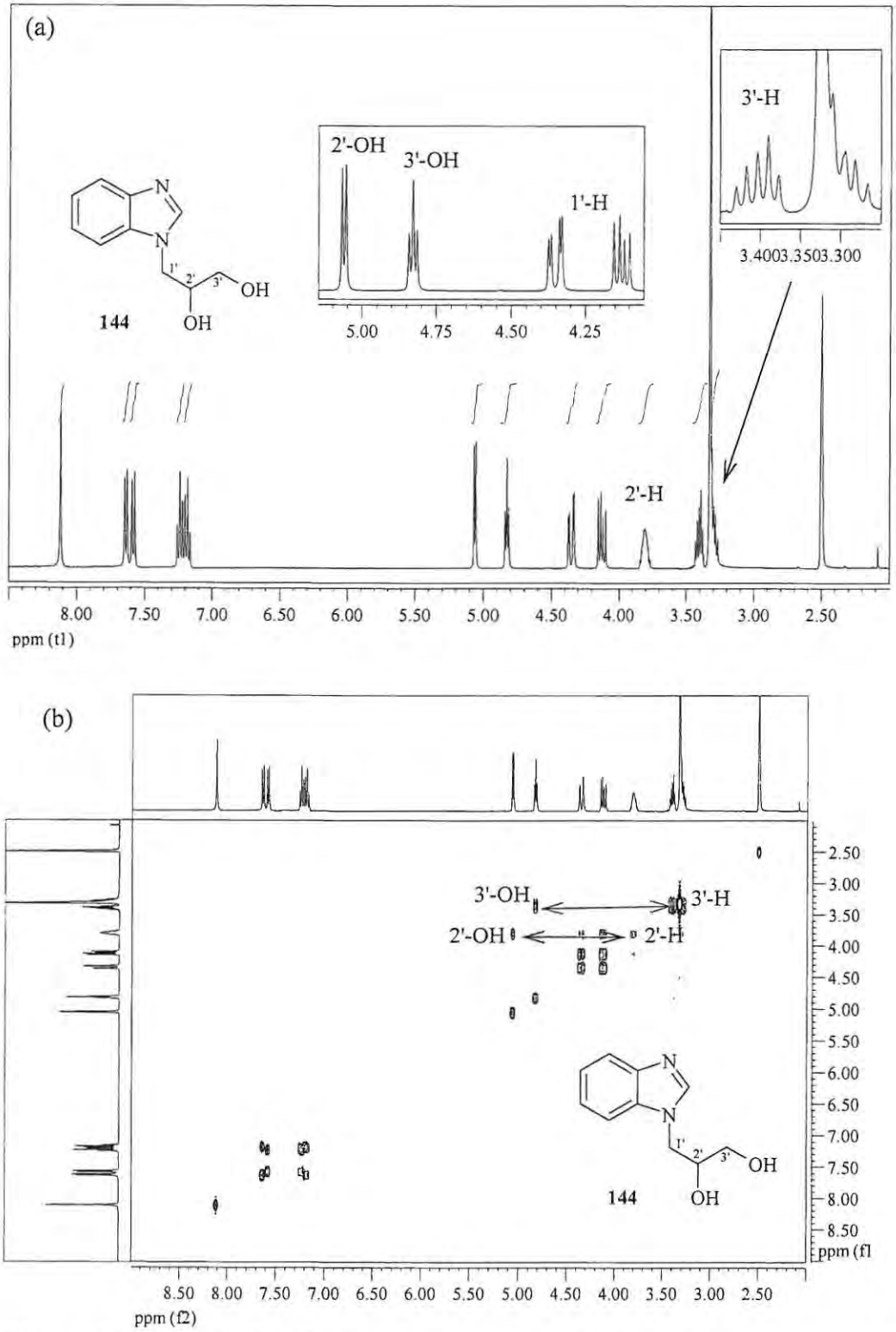


Figure 26: **a**) 400MHz ^{1}H NMR spectrum and **b**) COSY spectrum of *N*-(2,3-dihydroxypropyl)benzimidazole **144** in DMSO-d_6

The ^{13}C NMR spectra for each of the dihydroxy compounds **144-146** (illustrated for compound **145** in **Figure 27**) showed two signals in the region δ 63-76ppm, which were assigned to the hydroxy methylene and hydroxy methine carbons.

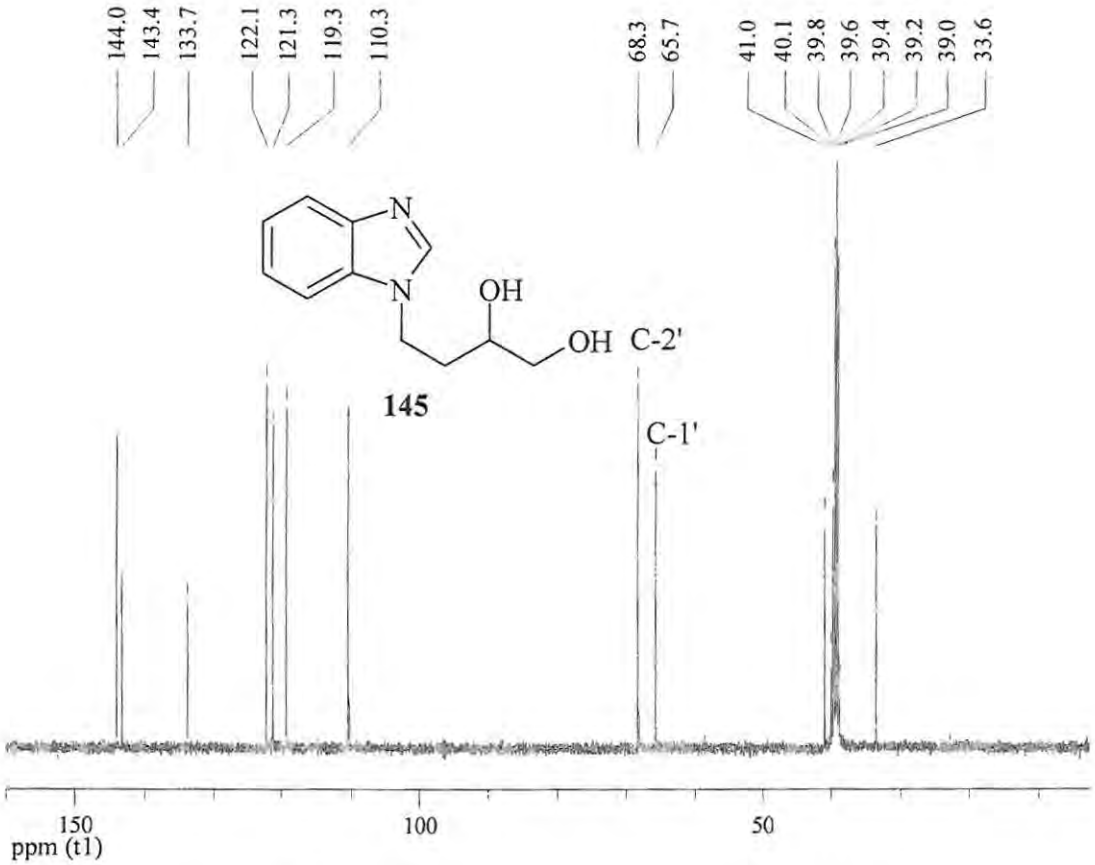


Figure 27: 100MHz ^{13}C NMR spectrum of compound **145** in DMSO-d_6.

In an attempt to improve the yields of diols **144-146**, alkylation of benzimidazole **140** with acetal **128** (**Scheme 31**) was explored following the same procedure used in the synthesis of the precursors **141-143**. After flash chromatographic purification, pure 1-[(2,2-dimethyl-1,3-dioxolan-4-yl)methyl]-1*H*-benzimidazole **147** was isolated as cream crystals melting at 89-91^{O}C in 38% yield. The ^{1}H and ^{13}C NMR and mass spectroscopic data for the product agreed with the expected structure. The ^{1}H NMR spectrum illustrated in **Figure 28** shows a series of double doublets at 3.7 and 4.0, and at 4.2 and 4.3ppm for the diastereotopic NCH_2 and OCH_2 methylene protons, respectively; the

diastereotopic methyl protons resonate at 1.3 and 1.4ppm. Compound **147** was then hydrolysed with aqueous acetic acid to afford compound **144** in 80% yield.

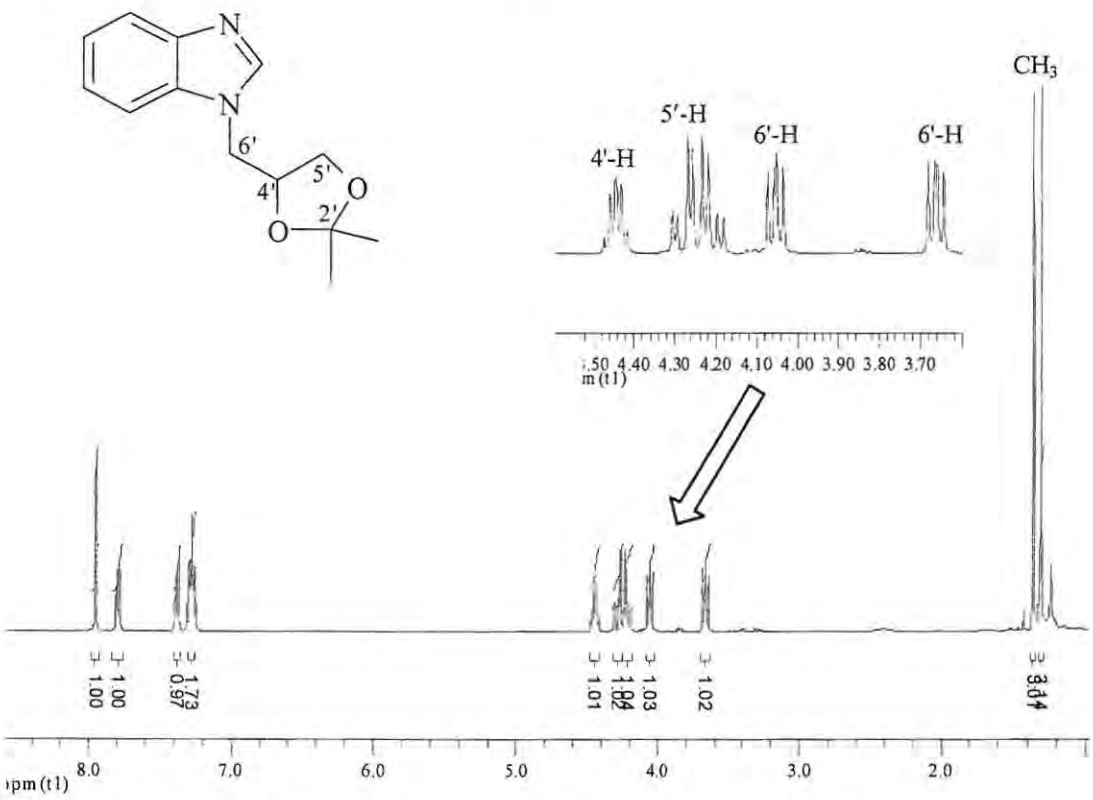
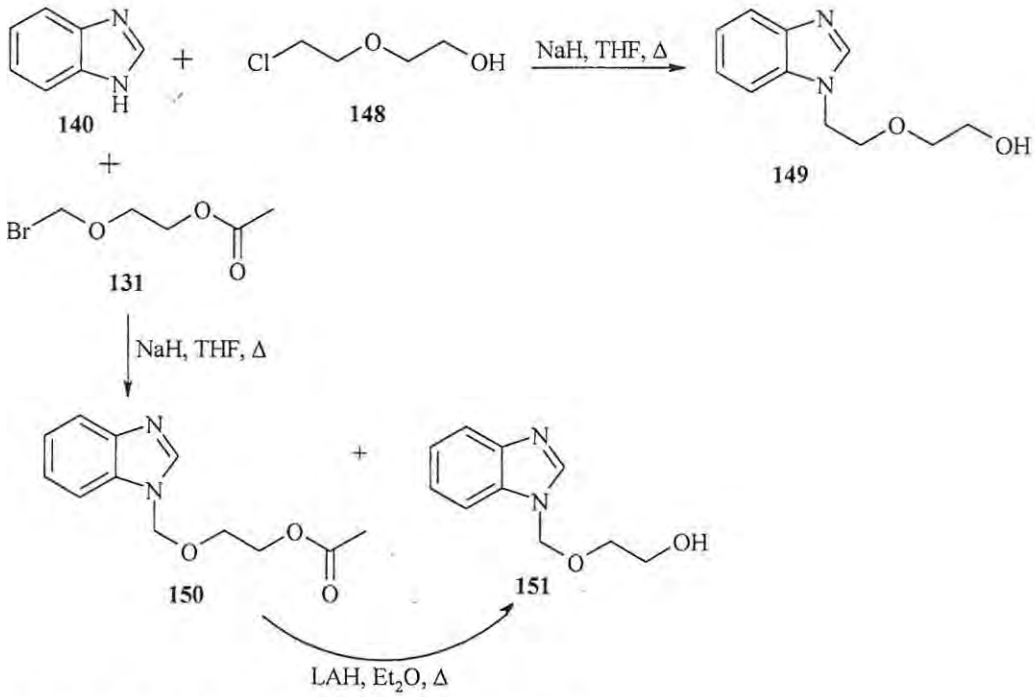


Figure 28: 400MHz ^{1}H NMR spectrum of compound **147** in $CDCl_3$.

Alkylation of benzimidazole **140** was also attempted with (2-bromomethoxy)ethyl acetate **131** and 2-choloroethoxyethanol **148** following the procedure described by Tan[165] (**Scheme 32**). After radial chromatography and HPLC purification, compound **149** was isolated in 25% yield. Purification of *N*-[(2'-acetoxyethoxy)methyl]benzimidazole **150** proved to be a particular challenge and the product had to be purified several times by flash and column chromatography on silica gel. ^{1}H NMR analysis of the material isolated in 14% yield showed the expected peaks for compound **150** but, on MS analysis, a peak with *m/z* 192 was observed, corresponding to the molecular mass of the deacetylated product **151**. Further purification of this material by column chromatography afforded compounds **150** and **151**.



Scheme 32: Alkylation of benzimidazole **140**

The formation of the alcohol **151** can be attributed to the use of excess sodium hydride and subsequent base-catalysed hydrolysis of the acetyl group during workup. Since compound **150** was susceptible to hydrolysis the alkylation of benzimidazole **140** with compound **131** was repeated and the crude product was subjected, without further purification, to reduction with LAH in diethyl ether (**Scheme 32**), affording the alcohol **151** in 21% after flash chromatography. Comparison of the ^{1}H and ^{13}C NMR spectra of compound **151** prepared by the reduction and the compound isolated following the alkylation reaction confirmed their identity. Comparison of the ^{1}H NMR spectra (**Figure 29**) of compounds **150** and **151** show similar signals except for the sharp singlet at *ca* 1.9ppm due to the acetyl protons in compound **150** and the broad singlet at *ca* 4.7ppm due to the hydroxyl proton in compound **151**. The proton NMR spectrum of alcohol **151** also showed an upfield shift in the ethylene triplets, reflecting the absence of the deshielding acetyl group. The ^{13}C NMR spectra of both products (**Figure 30**) also showed similar signal patterns except for the two extra signals observed for compound **150** at *ca.* 171 and 21ppm due to the acetyl carbonyl and methyl groups, respectively.

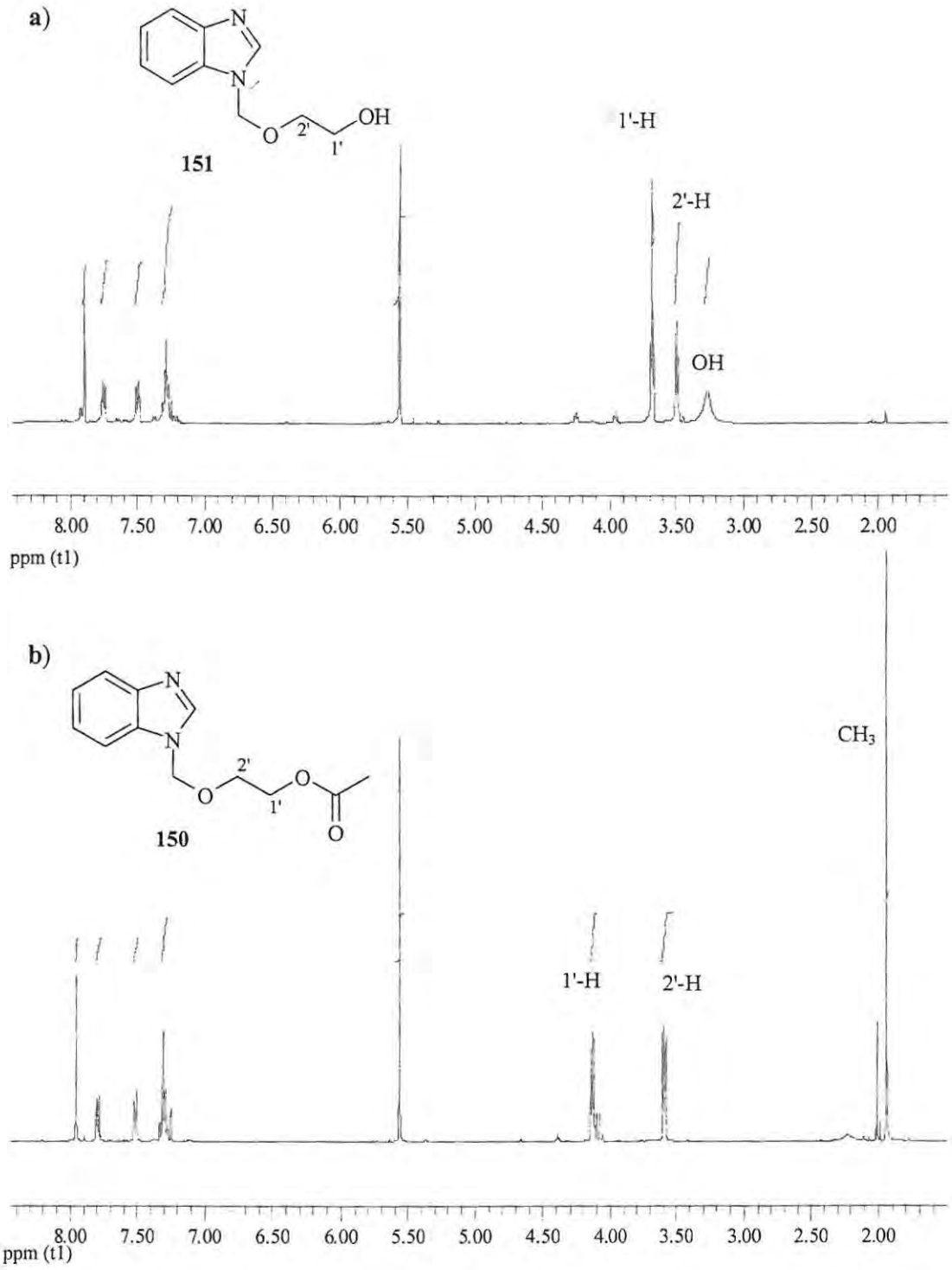


Figure 29: 400MHz ^{1}H NMR spectra of: **a**) compound **151** and **b**) compound **150** in $CDCl_3$.

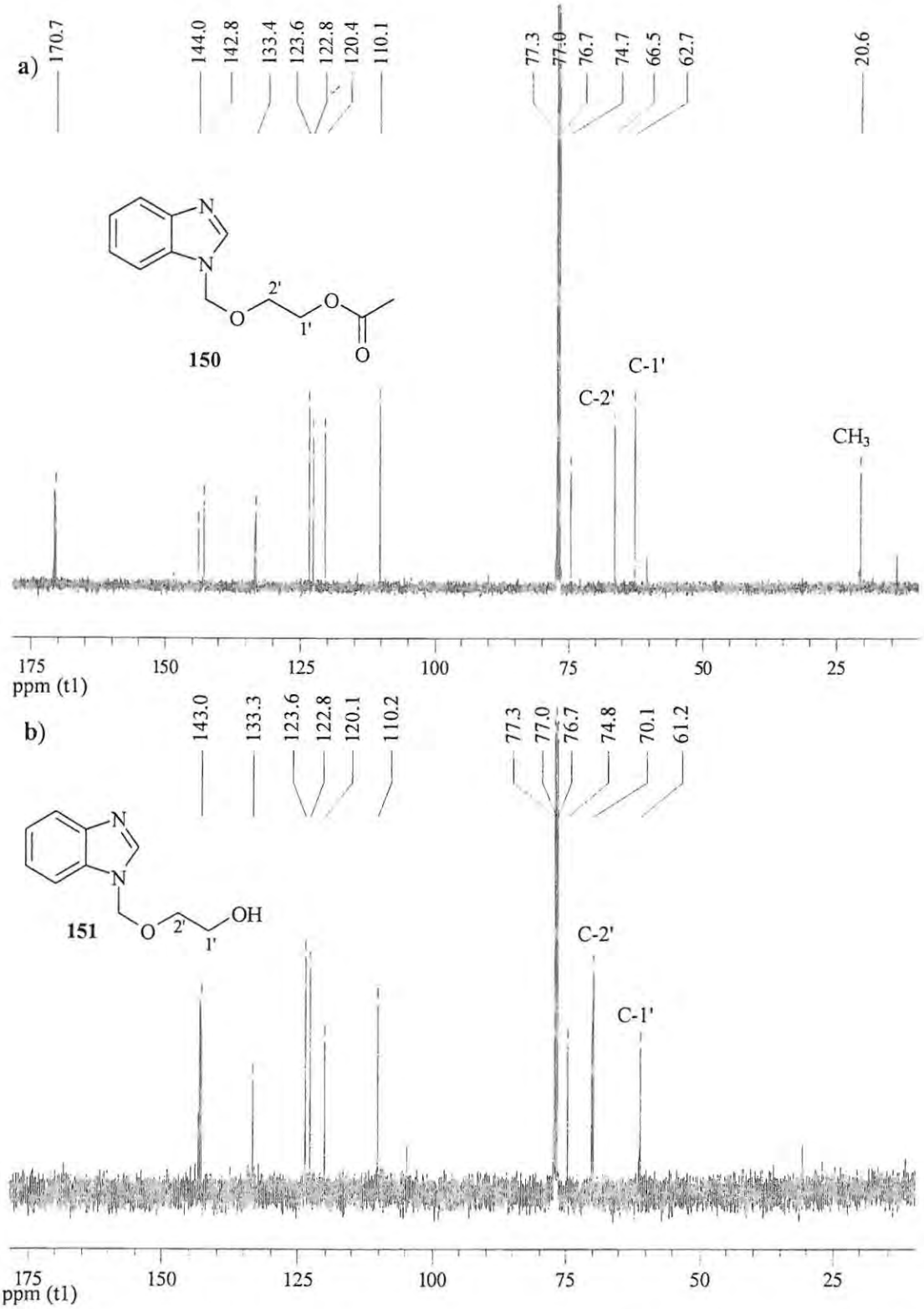
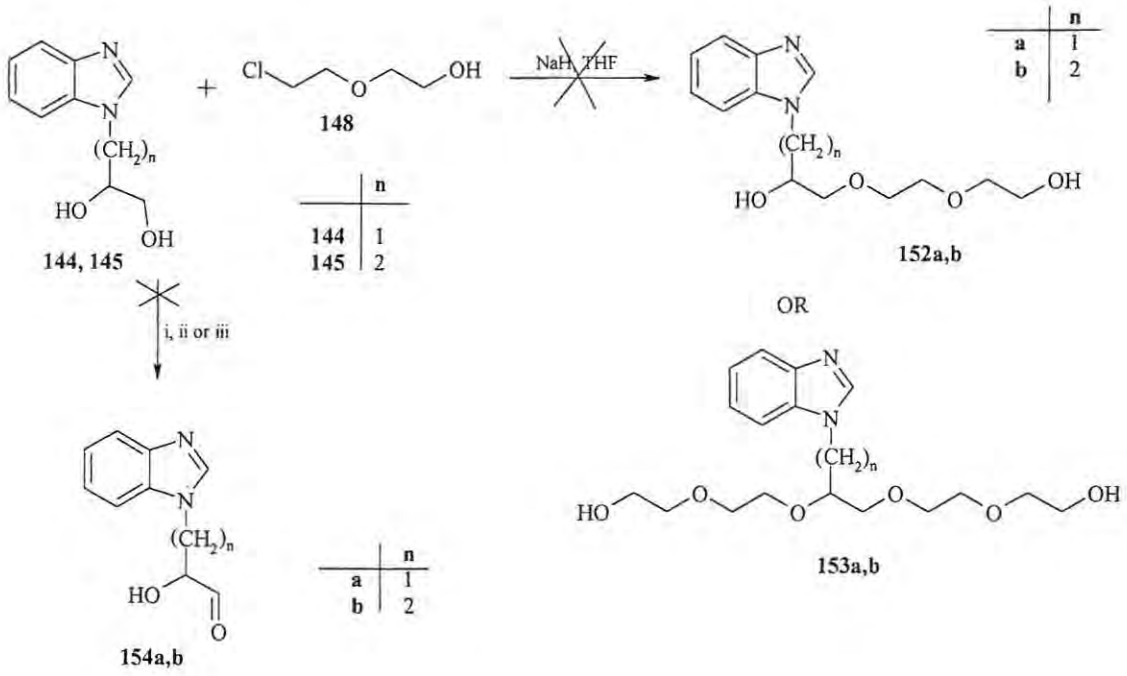


Figure 30: 100MHz ^{13}C NMR spectra of: **a**) compound **150** and **b**) compound **151** in $CDCl_3$.

2.1.3.2 Attempted side-chain elongation

Attempts were made to elongate the alkyl chain in the *N*-alkylated benzimidazole derivatives using two approaches. In the first approach, compounds **144** and **145** were treated with NaH and compound **148** in THF (**Scheme 33**). After solvent extraction, the crude material was analysed, in each case, by ^{1}H NMR spectroscopy but, unfortunately, the spectra did not show the signals corresponding to either the desired products or to the starting compounds **144** and **145**.



Scheme 33: Attempted synthesis of polyhydroxy compounds **152** or **153**.

Reagents and conditions: i) trichloroisocyanuric acid, TEMPO, CH_2Cl_2; ii) MnO_2, $CHCl_3$, Reflux 24h; iii) DCC, DMSO.

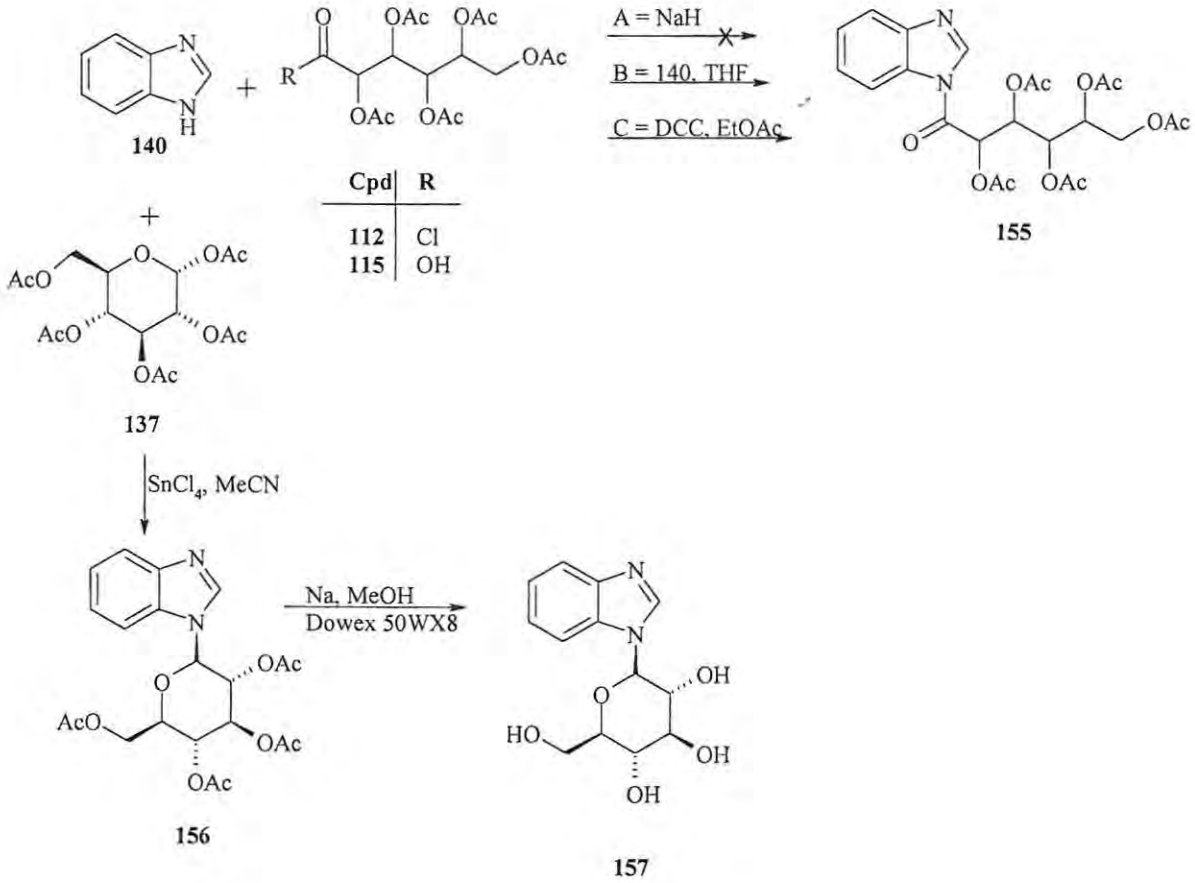
A second approach was to selectively oxidize the primary alcohol to an aldehyde using various reagent systems, *viz.*, manganese oxide, DCC and DMSO or trichloroisocyanuric acid and the free radical TEMPO. It was expected that the aldehyde could undergo a Baylis-Hillman reaction to produce a longer alkyl chain. Oxidation of compounds **144** and **145** was then attempted following literature methods.[168-170] Thus, compound **144** was

treated with trichloroisocyanuric acid and TEMPO in dichloromethane following the method described by Luca *et al*,[168] while the oxidation of compounds **144** and **145** was attempted using the Pfitzner-Moffatt oxidant following the method described by Giudici.[169] In all cases, NMR analysis of the material obtained after work-up, indicated that the reaction had been unsuccessful.. Finally, oxidation was attempted, using manganese dioxide following the method described by Ziesssel *et al.*[170] Compounds **144** and **145** were each treated with MnO_2 in chloroform under reflux overnight, but ^1H and ^{13}C NMR analysis of the crude material failed to reveal the expected aldehyde signals. In view of the difficulties encountered, the chain elongation experiments were abandoned.

2.1.3.3 Glycosylation of Benzimidazole

In order to introduce the amide-linking group acylation of benzimidazole **140** with gluconyl chloride **112** or acryloyl chloride **118** was explored, using sodium hydride, dicyclohexylcarbodiimide (DCC) or benzimidazole **140** itself to enhance the reaction following three of the methods found in literature.[165,171,172] In method A[165] (**Scheme 34**), benzimidazole **140** was treated with two equivalents of NaH and the acyl chloride **112**. However, this reaction was unsuccessful as the substrate **140** and the gluconic acid derivative **115** were isolated instead of the desired product **155**. Since it was evident that the gluconyl chloride **112** was susceptible to hydrolysis in the presence of NaH, other methods that did not involve the use of NaH were investigated. Following the procedure described by Fife[171] (Method B; **Scheme 34**), compound **112** was treated with two equivalents of benzimidazole **140** in THF under reflux conditions. After flash chromatography benzimidazole **140** and trace amounts of the desired product **155** were isolated, the latter contaminated with benzimidazole. Attempts at further purification were unsuccessful, and the use of DCC (Method C;[172] **Scheme 34**) was investigated. Compound **140** was treated with gluconic acid **115** and DCC in ethyl acetate and the mixture was stirred for 24h at room temperature. After chromatographic purification benzimidazole **140**, gluconic acid **115** and trace amounts of the desired product **155** contaminated with benzimidazole **140** were isolated.

Glycosylation of benzimidazole was then attempted using cyclic glucosyl acetate **137** and $SnCl_4$ following the modified procedures described by the Jain[125] and Krybill.[159] Benzimidazole **140** was treated with $SnCl_4$ and compound **137** in acetonitrile and the reaction mixture was heated at 70°C for 18h (**Scheme 34**). After work-up, flash chromatography and HPLC afforded compound **156** in 10% yield. The ^{1}H NMR spectrum of the product **156** (**Figure 31**) reveals an upfield shift of the 1'-H signal from 6.3ppm in the precursor **137** to 5.6ppm due to the replacement of the acetyl group by benzimidazole. The The ^{13}C NMR spectrum reveals four carbonyl and four methyl signals at 170-171ppm and 20.5-20.7ppm, respectively, corresponding to the four acetyl groups. Deacetylation[162] of compound **156** was achieved using sodium methoxide prepared *in situ* from sodium metal and methanol (**Scheme 34**). After neutralising with the ion-exchange resin, Dowex 50WX8, the reaction mixture was purified by HPLC to afford the tetra-hydroxy compound **157** in 70% yield. ^{1}H NMR analysis of compound **157** confirmed the absence of the four acetyl methyl singlets at *ca.*2.0ppm.



140 155 137 156 157

Scheme 34: Glycosylation of benzimidazole.

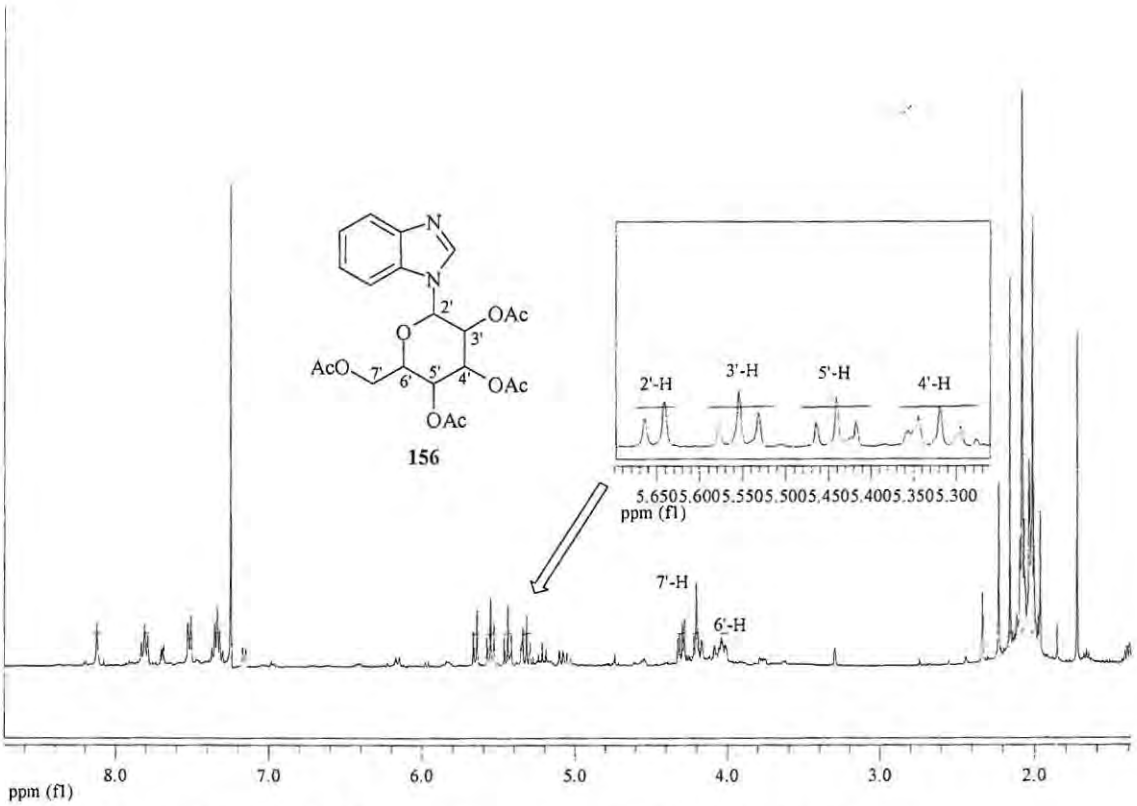
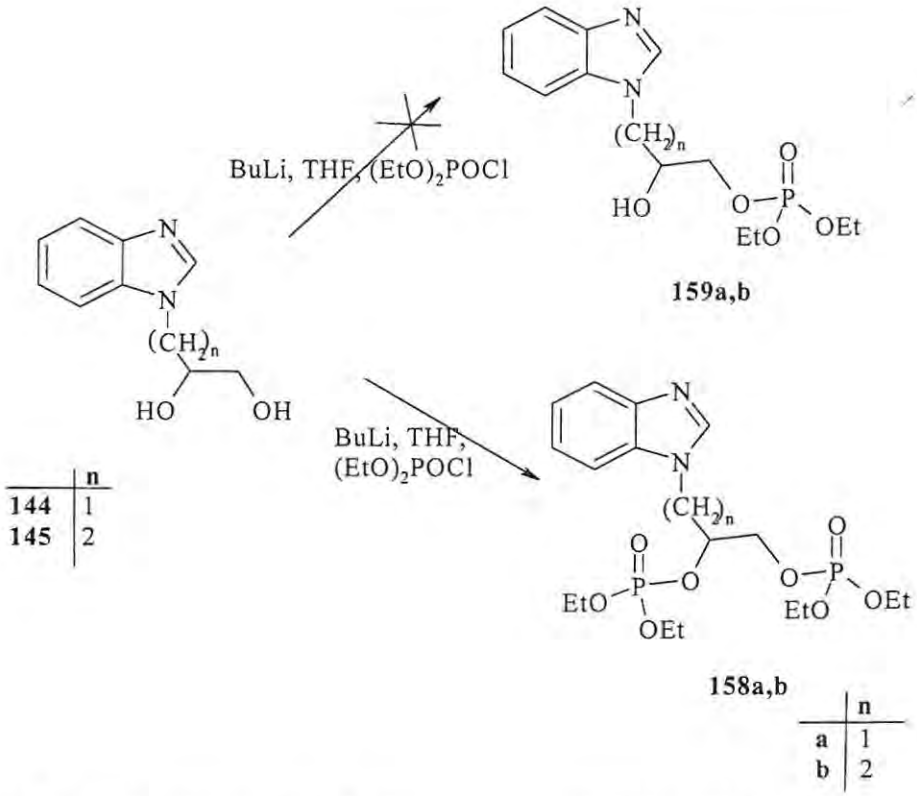


Figure 31: 400MHz ^{1}H NMR spectrum of compound **156** in $CDCl_3$.

2.1.3.4 Phosphorylation of the Benzimidazole Derivatives

The natural nucleotide, ATP, has phosphate ester groups, which play a role in the formation of glutamyl phosphate. Consequently, after the introduction of the hydroxyl substituents, the benzimidazole derivatives **144**, **145**, **149** and **151** were then subjected to phosphorylation reactions in order to generate phosphate esters.[173-176] Following the method reported by Takemoto,[173] compounds **144** and **145** were treated with one equivalent of BuLi and diethyl chlorophosphate at $0^{o}C$ (**Scheme 35**). After flash chromatography, the 1,2-diphosphate compounds **158a,b** were obtained as oils in very low yield, (2-11%), instead of the expected monophosphates **159a,b**.



	n
144	1
145	2

	n
a	1
b	2

Scheme 35: Phosphorylation of the hydroxy compounds **144** and **145**.

The ^{1}H NMR spectrum illustrated for compound **158a** (**Figure 32**) shows the presence of multiplets at *ca.* 4.1 and 1.2ppm corresponding to the methylene and methyl groups of the phosphate ester moieties. The ^{31}P NMR spectrum illustrated for compound **158a** (**Figure 33**) shows two distinct signals at –0.63 and –0.77ppm due to the chemically non-equivalent phosphorus nuclei. The ^{13}C NMR and DEPT135 spectra, illustrated for compound **158b** in **Figure 34**, show splitting of the methylene and methyl carbons by the phosphorus nuclei.[177] The low-resolution mass spectrometric data reveal peaks corresponding to the molecular ions for compounds **158a** and **b**. In the case of compound **158b**, the molecular ion is responsible for the base peak in the low-resolution mass spectrum.

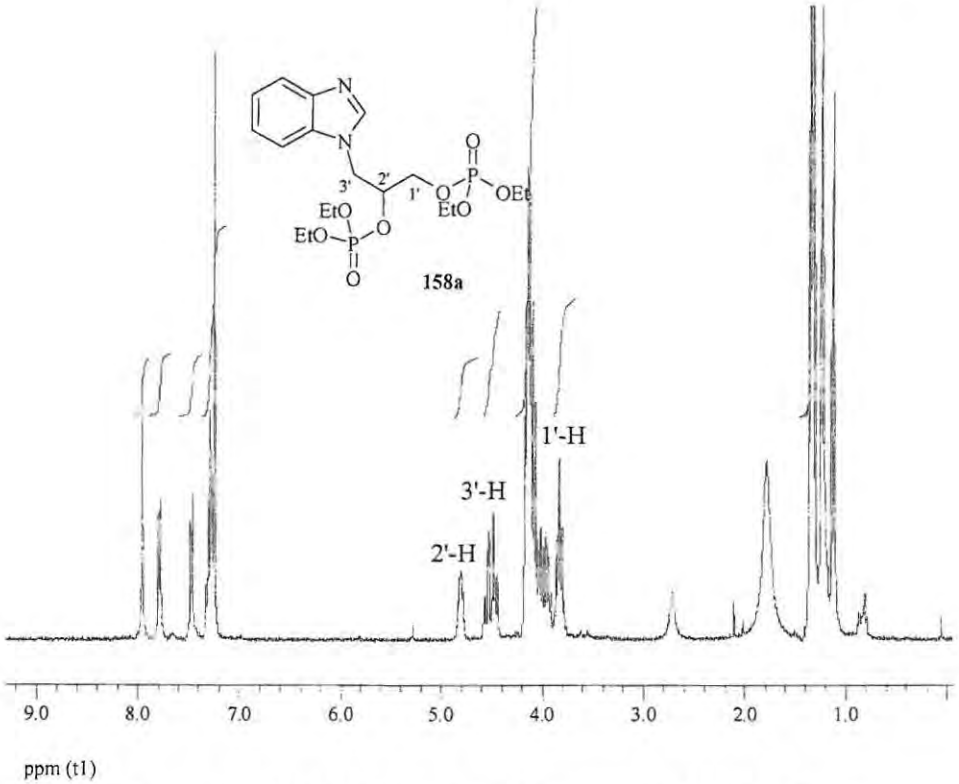


Figure 32: 400MHz ^{1}H NMR spectrum of compound **158a** in $CDCl_3$.

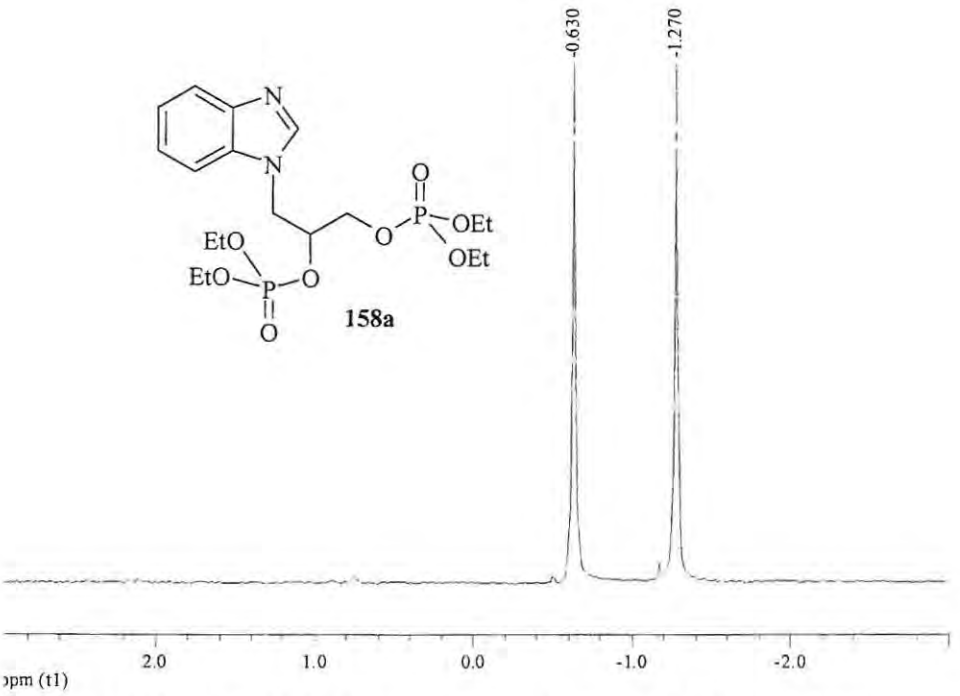


Figure 33: 162MHz ^{31}P NMR spectrum of compound **158a** in $CDCl_3$.

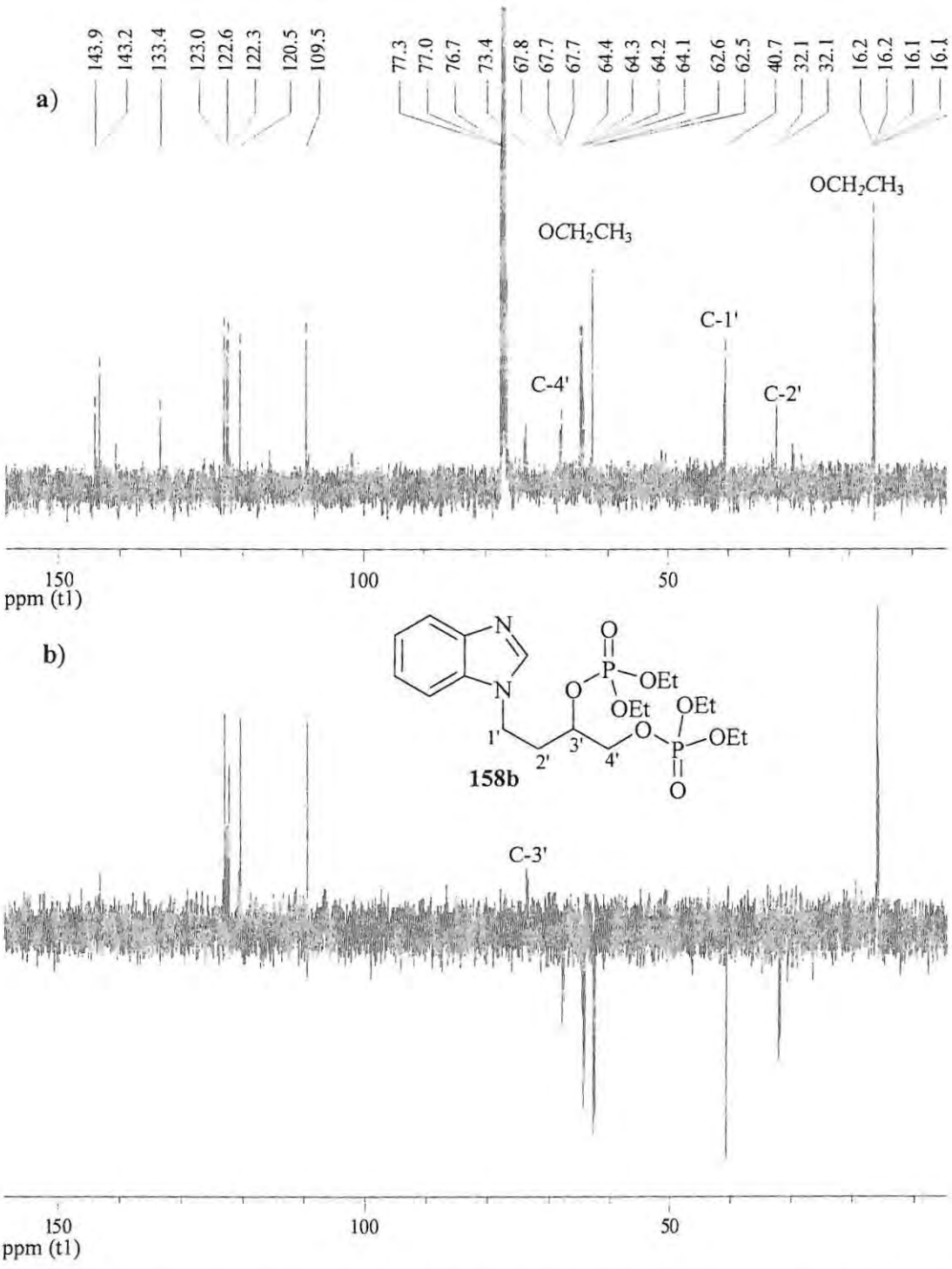
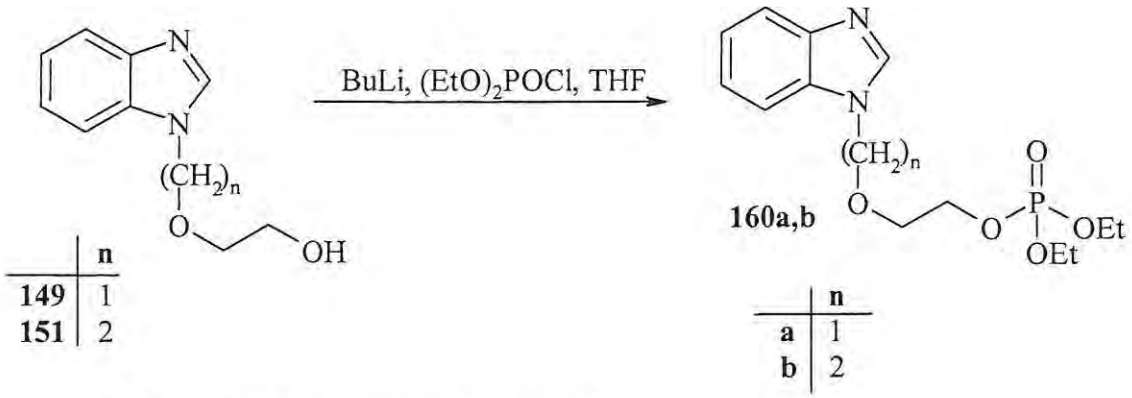


Figure 34: **a**) 100MHz ^{13}C NMR and **b**) DEPT 135 spectra of compound **158b** in $CDCl_3$.

The monohydroxy compounds **149** and **151** were also phosphorylated following the method reported by Takemoto[173] and, after flash chromatography, compounds **160a** and **b** (**Scheme 36**) were isolated as oils in moderate yield (23-35%) and their structures were verified by spectroscopic methods. The ^{1}H NMR spectrum, illustrated for compound **160b** in **Figure 35**, shows a multiplet at *ca.* 4ppm integrating for six protons due to the overlapping phosphate ester methylene nuclei. This assignment was confirmed by the HSQC spectrum (**Figure 36**) which shows a correlation of these methylene protons with two different methylene carbon atoms. The ^{31}P NMR spectrum for compound **160b** shows a single signal at -0.48ppm, while the low- resolution mass spectrometric data reveal a peak corresponding to the molecular ion and the HRMS data confirms the molecular formula of compound **160b** and provides evidence for the assigned structure.



	n
149	1
151	2

	n
a	1
b	2

Scheme 36: Synthesis of compounds **160a** and **b**.

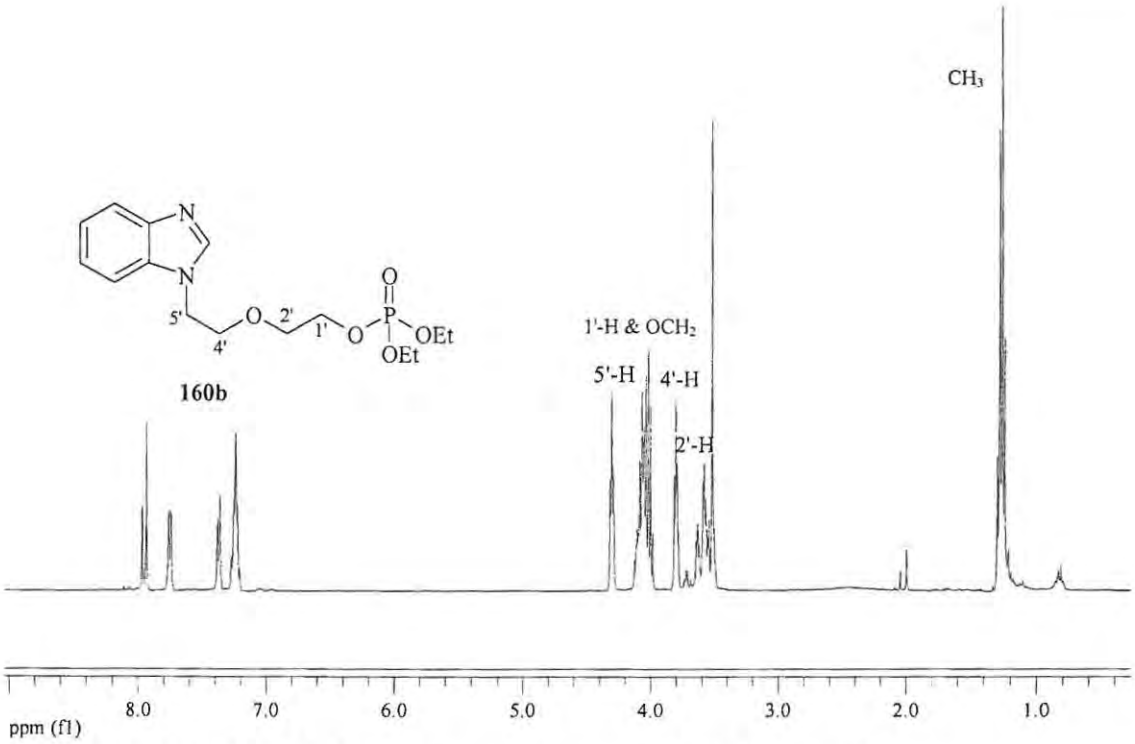


Figure 35: 400MHz ^{1}H NMR spectrum of compound **160b** in $CDCl_3$.

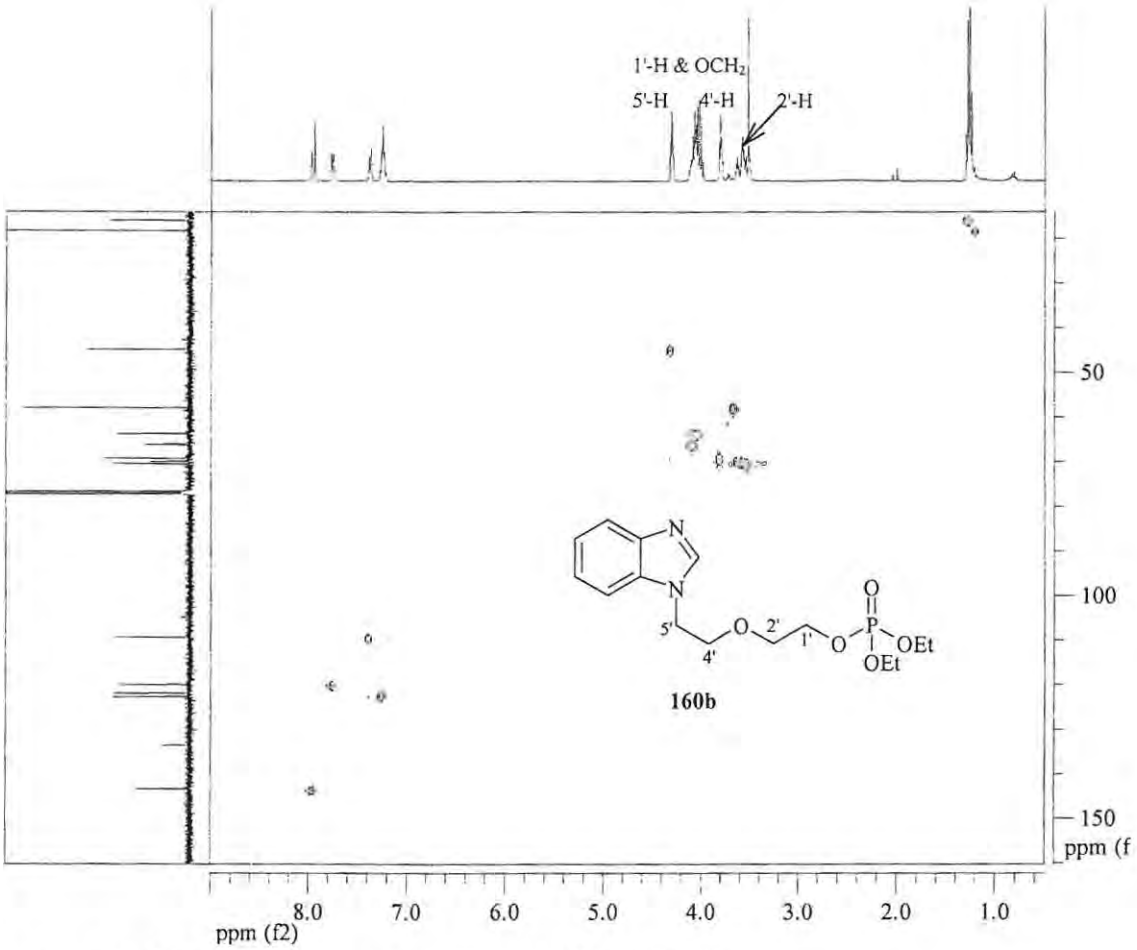
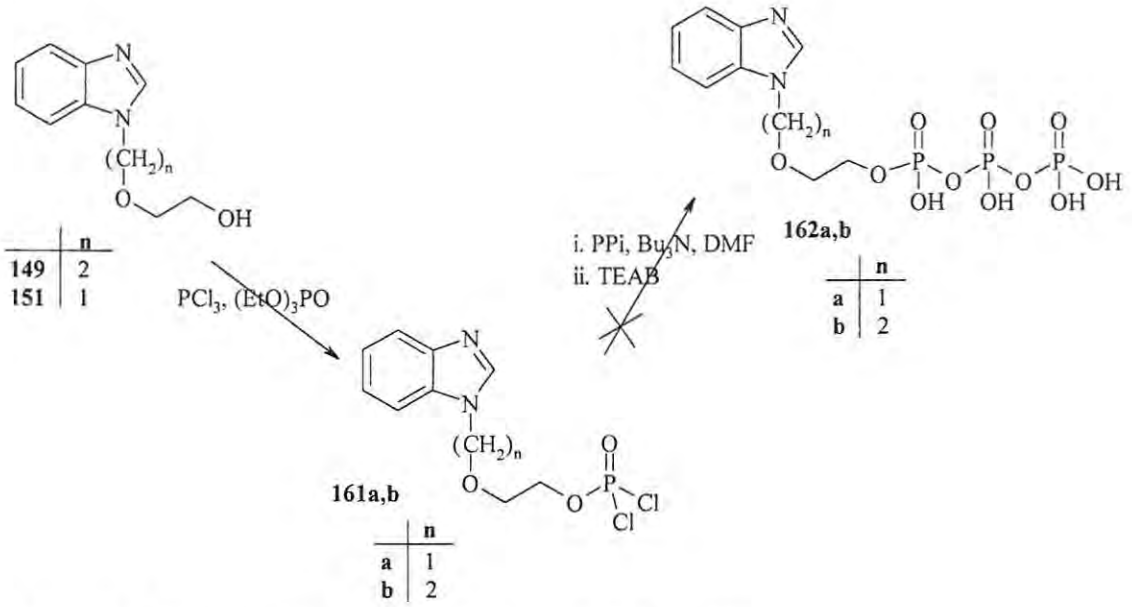


Figure 36: HSQC spectrum of compound **160b** in $CDCl_3$.

Attempts were also made to synthesize triphosphate ester derivatives of the benzimidazole derivatives following methods reported in literature.[172-174] Compounds **149** and **151** were treated with phosphoryl chloride in triethyl phosphate under argon to form the intermediates **161a** and **b** (**Scheme 37**). However, attempts to access the corresponding triphosphates **162a** and **b**, using five equivalents of tributylammonium pyrophosphate (PPi) and tributylamine in DMF proved unsuccessful.



Scheme 37: Attempted synthesis of compounds **162a** and **b**.

2.1.3.5 Structural Comparison of the Benzimidazole Derivatives with ATP

Selected benzimidazole derivatives, designed as ATP analogues, were explored using a molecular dynamics simulation routine to obtain their minimum energy conformers. The energy minimized structures of the benzimidazole derivatives **149, 157, 158a** and **b** and **160a** and **b** (**Figure 37**) were aligned with ATP to examine their similarities and differences. The preliminary overlay (**Figure 38**) shows that these derivatives exhibit reasonably good structural homology with ATP although the phosphate ester groups in compounds **158a,b** and **160b** lie outside the alignment. Use of the Cerius^2 LIGAND FIT module to determine the most favourable bound conformations and investigate the interaction of various benzimidazole derivatives with the amino acids in the enzyme active site is addressed in Section 2.3.2.

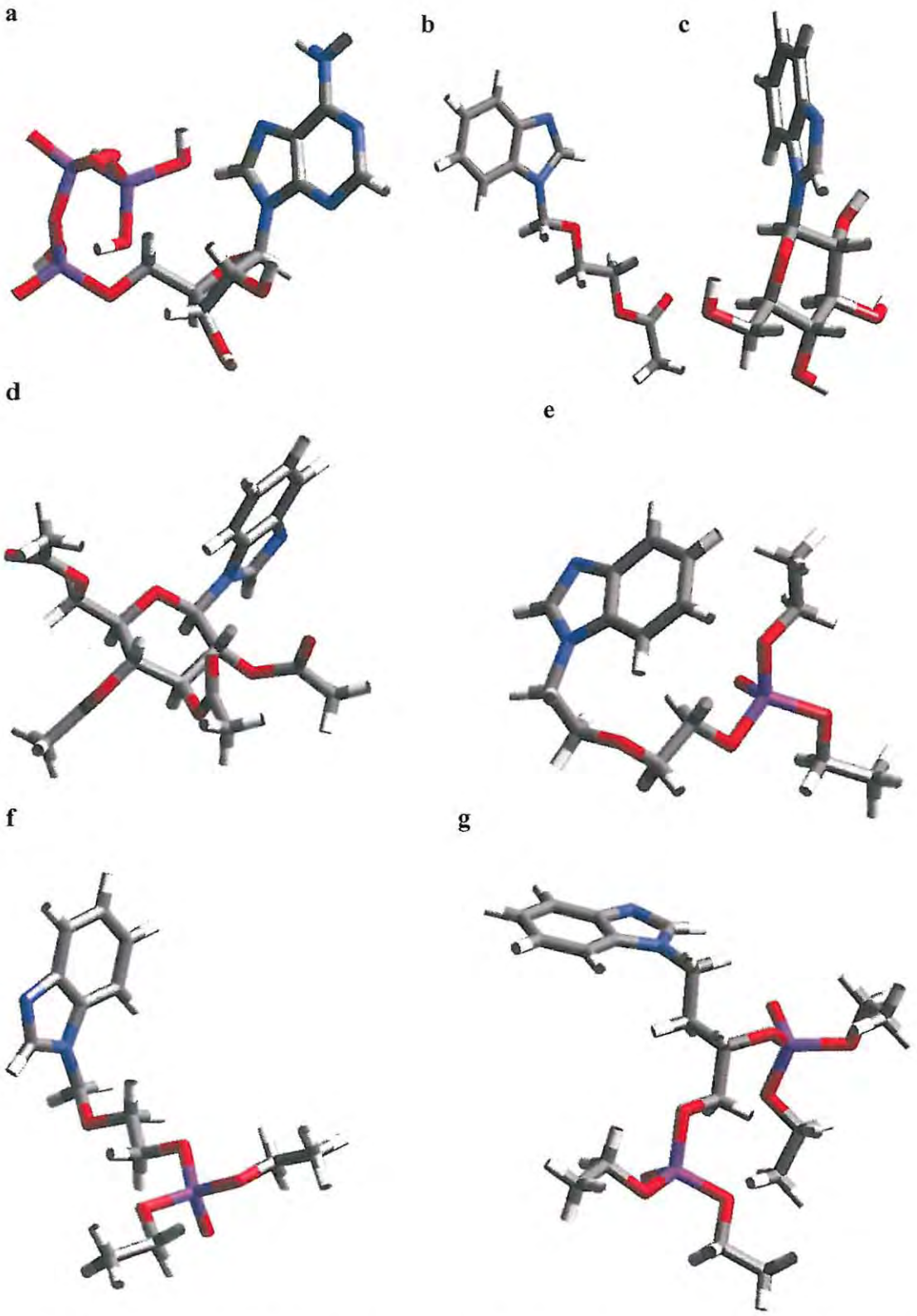


Figure 37: Energy minimized structures of ATP (**a**), and compounds **150** (**b**), **157** (**c**), **156** (**d**), **160b** (**e**), **160a** (**f**) and **158b** (**g**).

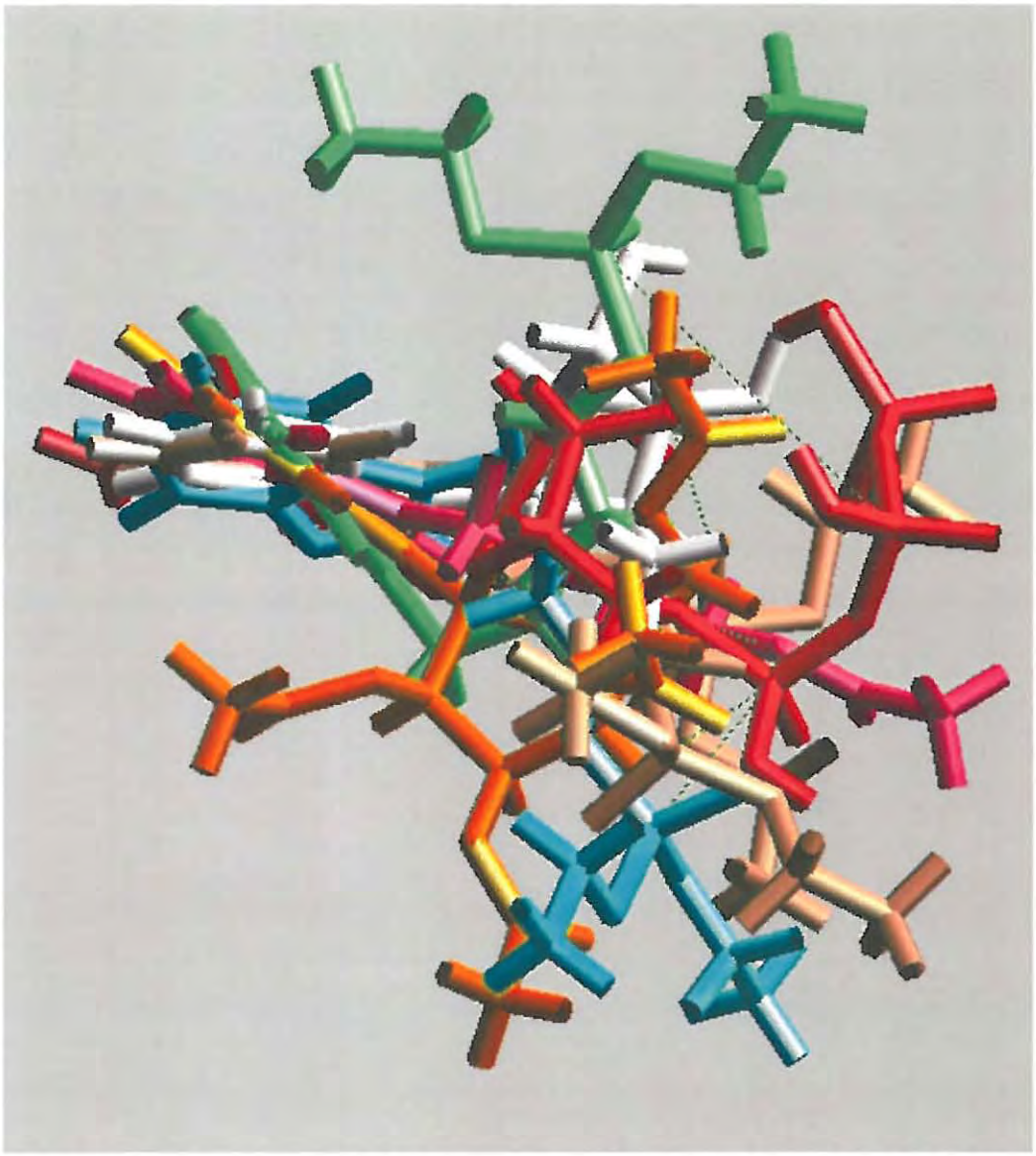
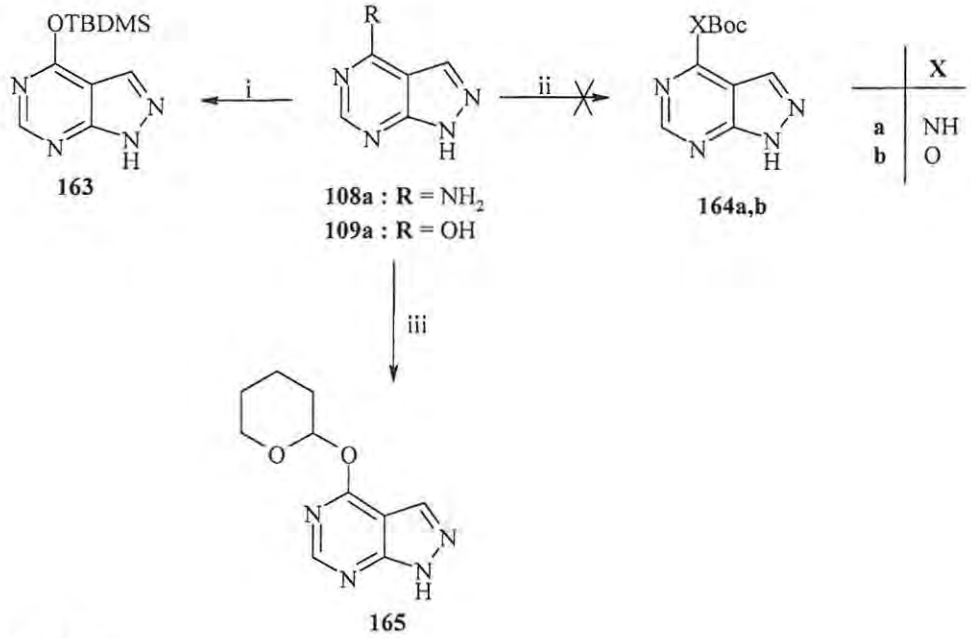


Figure 38: Alignment of the energy-minimized structures of compounds **149** (pink), **157** (white), **158a** (yellow), **158b** (light brown), **160a** (light green) and **160b** (light blue) with ATP (red).

2.1.4 Synthesis of pyrazolo[3,4-*d*]pyrimidine derivatives

Both 4-aminopyrazolo[3,4-*d*]pyrimidine (4-APP) **108a** and 4-hydroxypyrazolo[3,4-*d*]pyrimidine (allopurinol) **109a**, an analogue of hypoxanthine, are structurally similar to the heterocyclic base of the natural nucleotide ATP and, as indicated in **Section 1.6.3**, both allopurinol and 4-APP-1-ribonucleotide and ribonucleoside have anti-viral and antitumor[125,130,134] properties, while both allopurinol and 4-APP and their ribonucleoside and acyclic riboside derivatives have anti-parasitic activity.[130-132] Keto-enol tautomerism of allopurinol has been observed in both solid and neutral liquid state,[178,179] but the HMBC spectrum from the NMR spectrometric analysis of the commercial reagent revealed that it was 4-hydroxypyrazolo[3,4-*d*]pyrimidine, the enol tautomer of allopurinol **109a**. Both the amino group in 4-APP **108a** and the hydroxy group in allopurinol **109a** are susceptible to electrophilic attack and, hence, protection of these substituents was attempted to ensure selective substitution at the N-1 position.



Scheme 38:

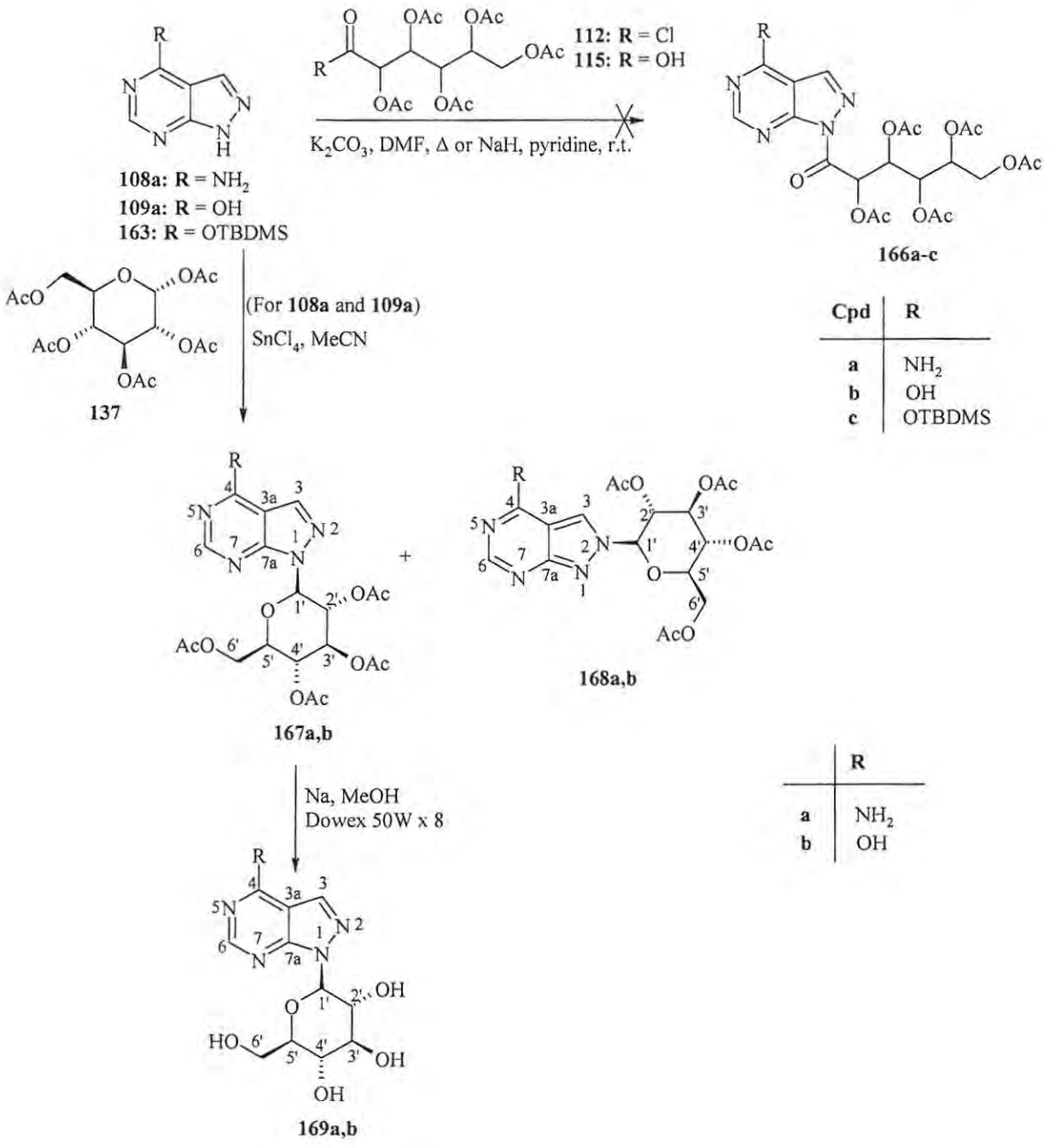
Reagents and conditions: i) TBDMSCl, imidazole, DMF; ii) $(Boc)_2O$, THF, 70^{O}C; iii) 2,3-dihydropyran, DMF, PTSA, benzene.

Thus, allopurinol **109a** was treated with imidazole and TBDMSCl in dry DMF (**Scheme 38**)[142] and, after recrystallisation, compound **163** was isolated as an off-white solid in very low yield (10%). Attempts were then made to protect the amino and hydroxy substituents as the *tert*-butyloxycarbonyl (Boc) amide or ester, but treatment of 4-APP **108a** and allopurinol **109a** with $(Boc)_2O$ in THF under reflux conditions for seven days failed to produce the desired compounds **164a** or **b** (**Scheme 38**).[180]

Protection of the hydroxy substituent of allopurinol was also attempted using dihydropyran following the method described by Hussain.[181] Allopurinol **109a** was treated with *p*-toluenesulfonic acid (PTSA) in DMF and, after recrystallisation from petroleum ether, compound **165** was isolated in moderate yield (51%). Due to the poor yield of the protected systems it was decided to explore reactions using the unprotected pyrazolo[3,4-d]pyrimidine systems **108a** and **109a**.

2.1.4.1 Glucosylation of pyrazolo[3,4-d]pyrimidine derivatives

Efforts were made to incorporate an amide-linking group by acylating 4-APP **108a**, allopurinol **109a** and the allopurinol derivative **163** with gluconyl chloride **112** or gluconic acid **115**. 4-APP **108a** was treated with the acid chloride **112** and potassium carbonate in dry DMF under reflux conditions (**Scheme 39**).[182] Analysis of the isolated fractions by ^1H NMR spectroscopy, however, showed none of the signals expected for the desired product **166a** nor any that corresponded to either of the starting materials. It was assumed that the high temperatures might have resulted in product decomposition and hence the reaction conditions were altered. Both compounds **108a** and **109a** were treated with NaH in pyridine at room temperature (**Scheme 39**), but this also failed to produce the desired compounds **163a** and **b**; instead, the starting materials were isolated. Attempts were made to introduce an amide linkage on the allopurinol system by treating compound **163** with gluconic acid **115** and triethylamine in the presence of thionyl chloride and dry benzene.[183] This reaction also proved unsuccessful, as only the starting compounds were isolated and attention was then focused on introducing an amine linkage instead.



Cpd	R
a	NH_2
b	OH
c	OTBDMS

	R
a	NH_2
b	OH

Scheme 39: Glycosylation of pyrazolo[3,4-*d*]pyrimidine systems.

Given the difficulties experienced in introducing an acyclic sugar derivative into the system, attempts were made to introduce the cyclic glucose derivative **137**. Jain *et al.*[125] reported the synthesis of compound **167a** from silylated allopurinol and isolated mono- and diglycosylylated derivatives. We attempted the synthesis of compound **167b** starting from the unprotected allopurinol **109a** using a modification of the procedures described by Jain[125] and Krybill.[161] Allopurinol **109a** and compound **137** were treated with $SnCl_4$ in dry acetonitrile at 70°C (**Scheme 39**). After purification by flash chromatography the fraction, which revealed signals corresponding to the desired product **167b**, was further purified by recrystalization from ethanol to afford compound **168b,** as shown by NMR spectrometry, in low yield (6%). The ethanol filtrate was further purified by HPLC to afford compound **167b**, as shown by NMR spectrometry, also in low yield (1.4%). The ^{1}H NMR spectra illustrated for both isomers in **Figure 39** revealed similar signals except for the slight shift of the 1'-H and 2'-H signals respectively, which are observed overlapping as a multiplet at *ca.* 6.0ppm in compound **167b** and as a doublet and triplet at 6.09 and 5.68ppm, respectively, in compound **168b**. The ^{13}C NMR spectra illustrated for both isomers in **Figure 40** also revealed an upfield shift (to 127.7ppm) of the signal for carbon 3 due to the enamine bond in compound **168b** compared to 136.6ppm due to the imine bond in compound **167b**. The large $J_{1,2}$ value 8.8Hz in isomer **168b** confirms the presence of a β-glucoside.

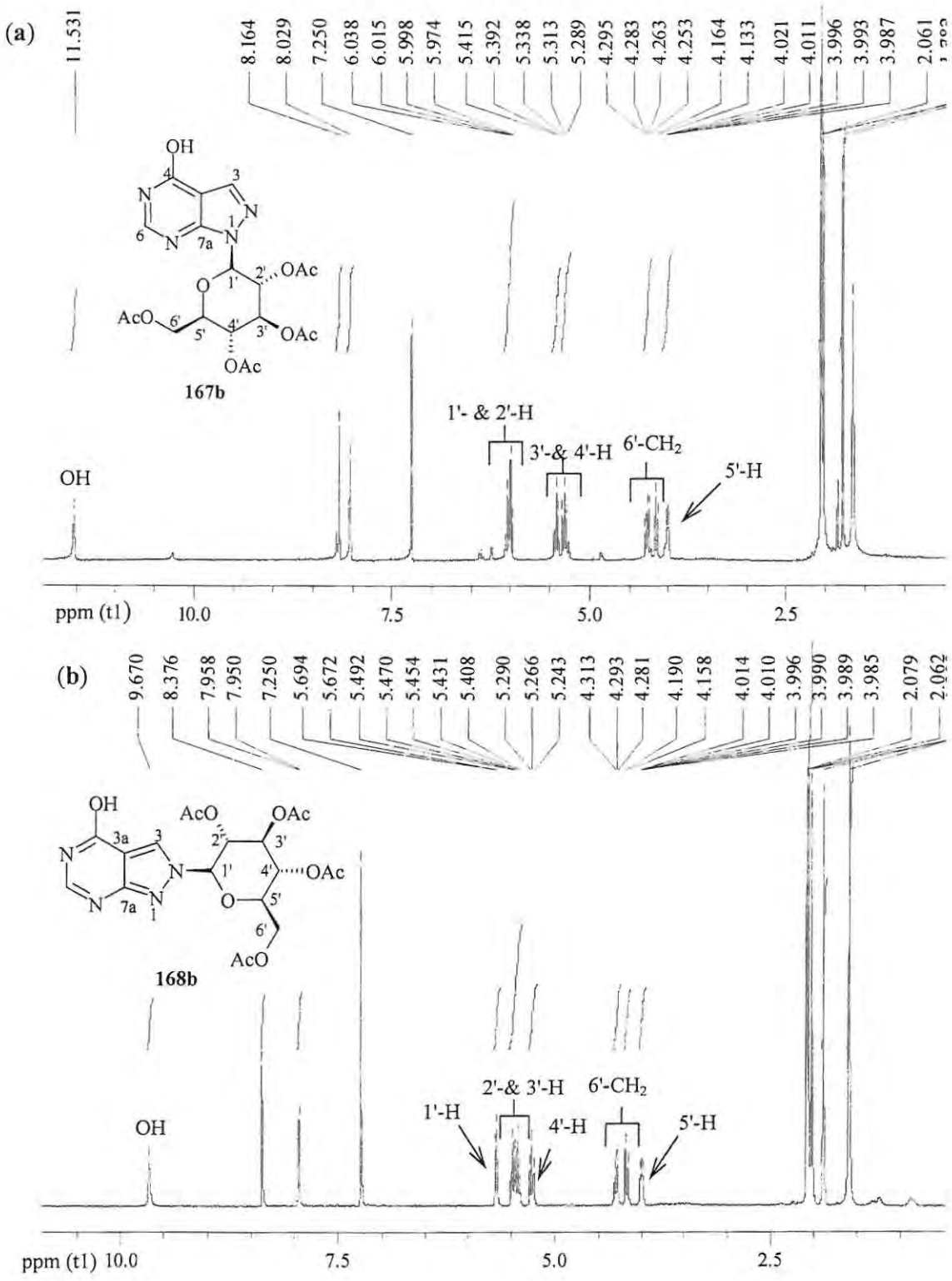


Figure 39: 400MHz ^{1}H NMR spectra of compounds; **a) 167b** and **b) 168b** in $CDCl_3$

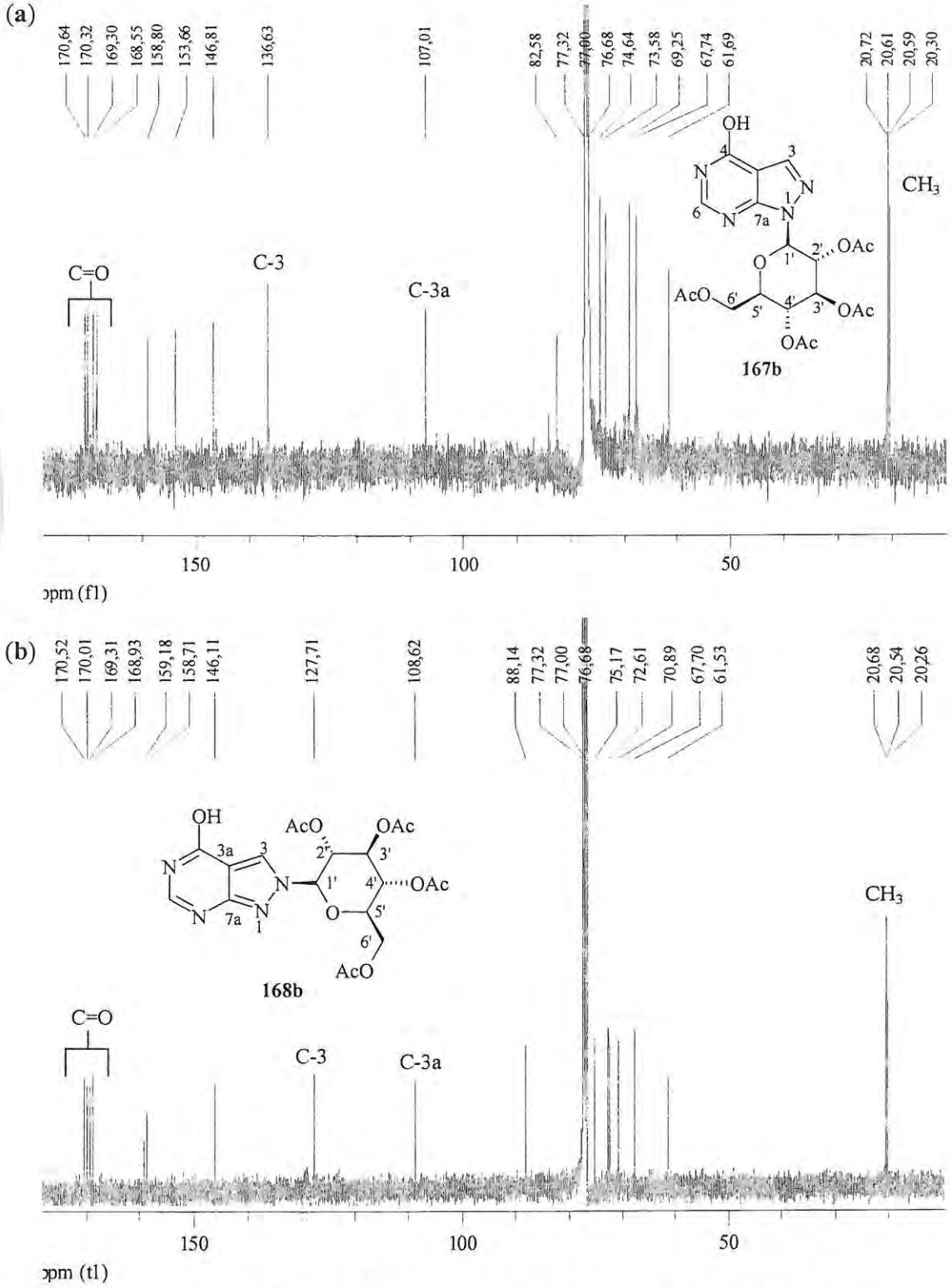


Figure 40: 100MHz ^{13}C NMR of compound **a**) **167b** and **b**) **168b** in $CDCl_3$.

Similarly, 4-APP **108a** and compound **137** were reacted with tin tetrachloride in acetonitrile under reflux conditions (**Scheme 39**). Following radial chromatography, two fractions were isolated as golden oils and characterized by NMR and low- and high-resolution mass spectrometry. The ^{1}H NMR spectra, illustrated for both isomers in **Figure 41**, revealed similar signals, *viz.*, four singlets at *ca.* 2ppm integrating for 12 protons, corresponding to the four acetyl groups, and a broad singlet in the region δ 6-7ppm integrating for 2 protons and corresponding to the amino group. The presence of the amino group in each case indicates *N*-glycosylation of the pyrrole ring. It was also observed that the signal corresponding to the glucosyl 1'- and 2'-methine protons are observed at *ca.* 5.8 and 5.5ppm, respectively, in one isomer and at 6.1 and 6.0ppm in the other. Based on these observations, it was initially presumed that the two fractions were the α- and β-anomers of compound **167**. However, this presumption was refuted by the data from the coupling constants $J_{1,2}$, ^{13}C NMR and HMBC spectra and the Modgraph NMR prediction programme (Section 2.2.2). The observed large $J_{1,2}$ values, 9.3Hz for both isomers, are an indicative of trans diaxial coupling in β-glucosides. The ^{13}C NMR spectra, illustrated for both isomers in **Figure 42**, revealed four methyl and four carbonyl signals at *ca.* 20 and 170ppm, respectively, confirming the presence of the glucosyl moiety in each case. But discrepancies were noted for the glucosyl C-1' signal which appears at 81ppm on one isomer and at 88ppm in the other, while the heterocyclic 4-APP C-3 and C-3a signals appear at 124 and 103ppm, respectively, in one isomer and at 134 and 101ppm for the other. The HMBC data illustrated for both isomers in **Figure 43** shows that the glucopyranose 1'-methine proton correlates to the 4-APP C-7a signal in one isomer and to C-3a in the other. These observations indicate that the two fractions are indeed the constitutional isomers **167a** and **168a**.

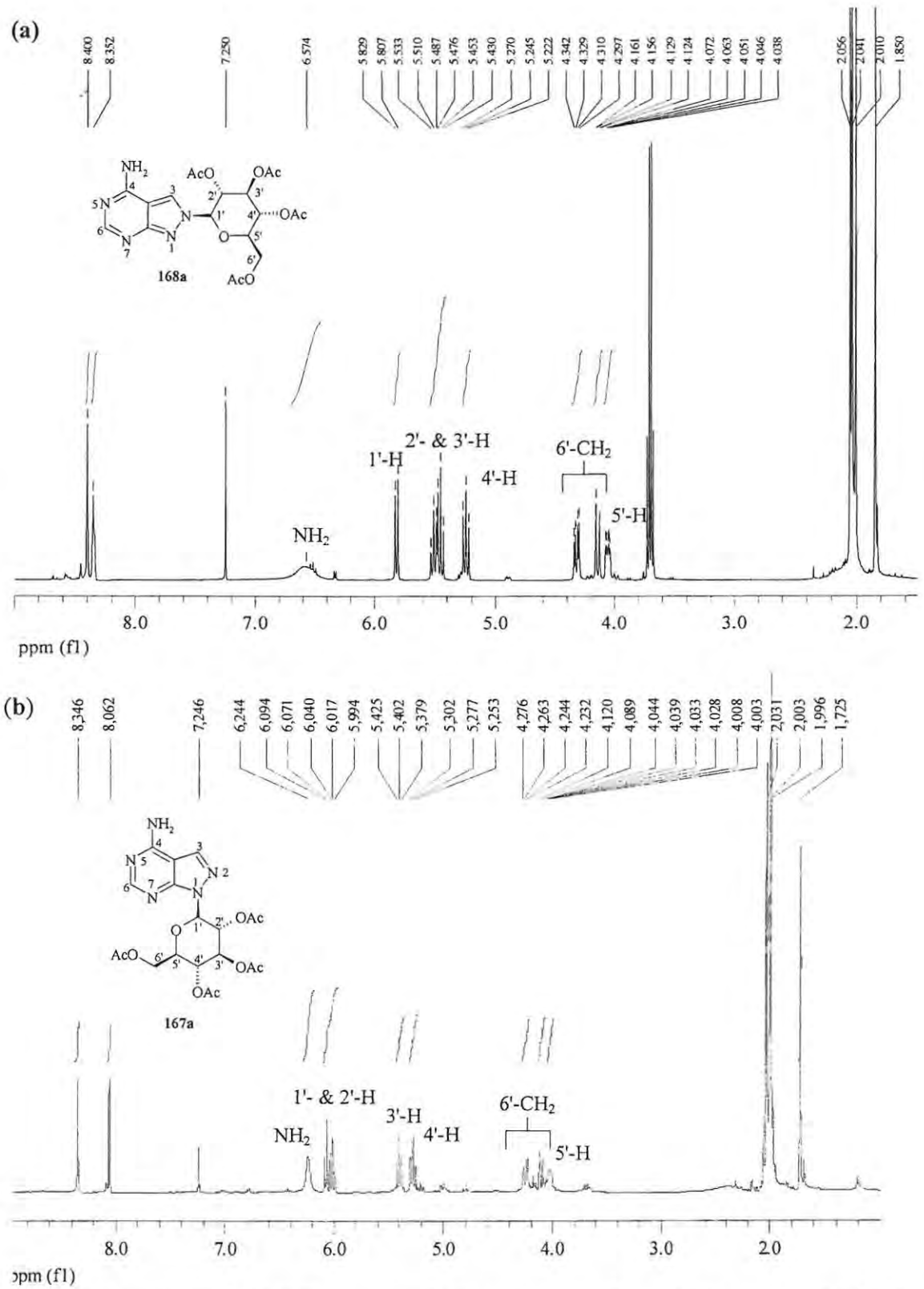


Figure 41: 400MHz ^{1}H NMR spectra of: **a**) compound **168a** and **b**) compound **167a** in $CDCl_3$.

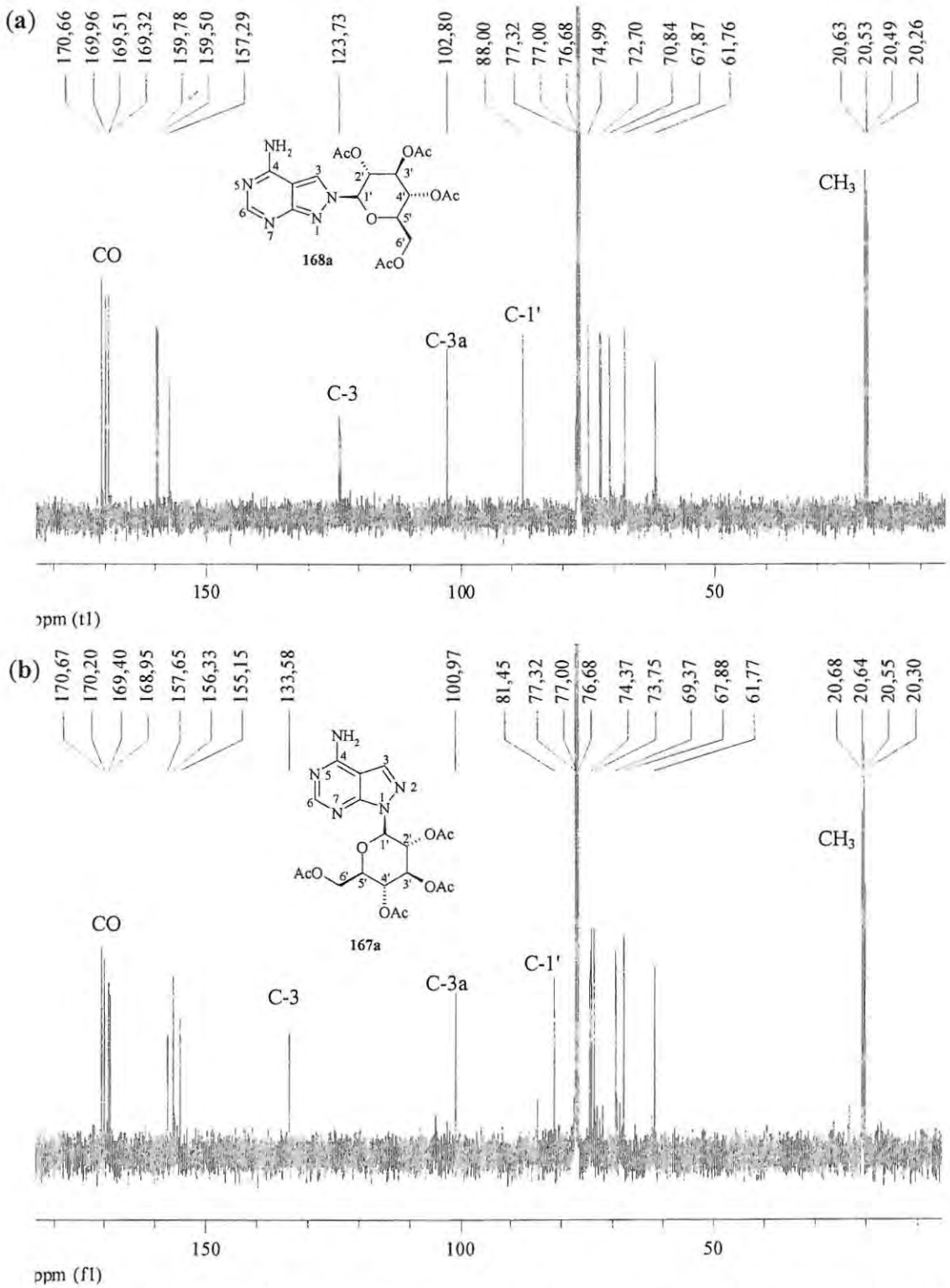


Figure 42: 100MHz ^{13}C NMR spectra of: **a**) compound **168a** and **b**) compound **167a** in $CDCl_3$.

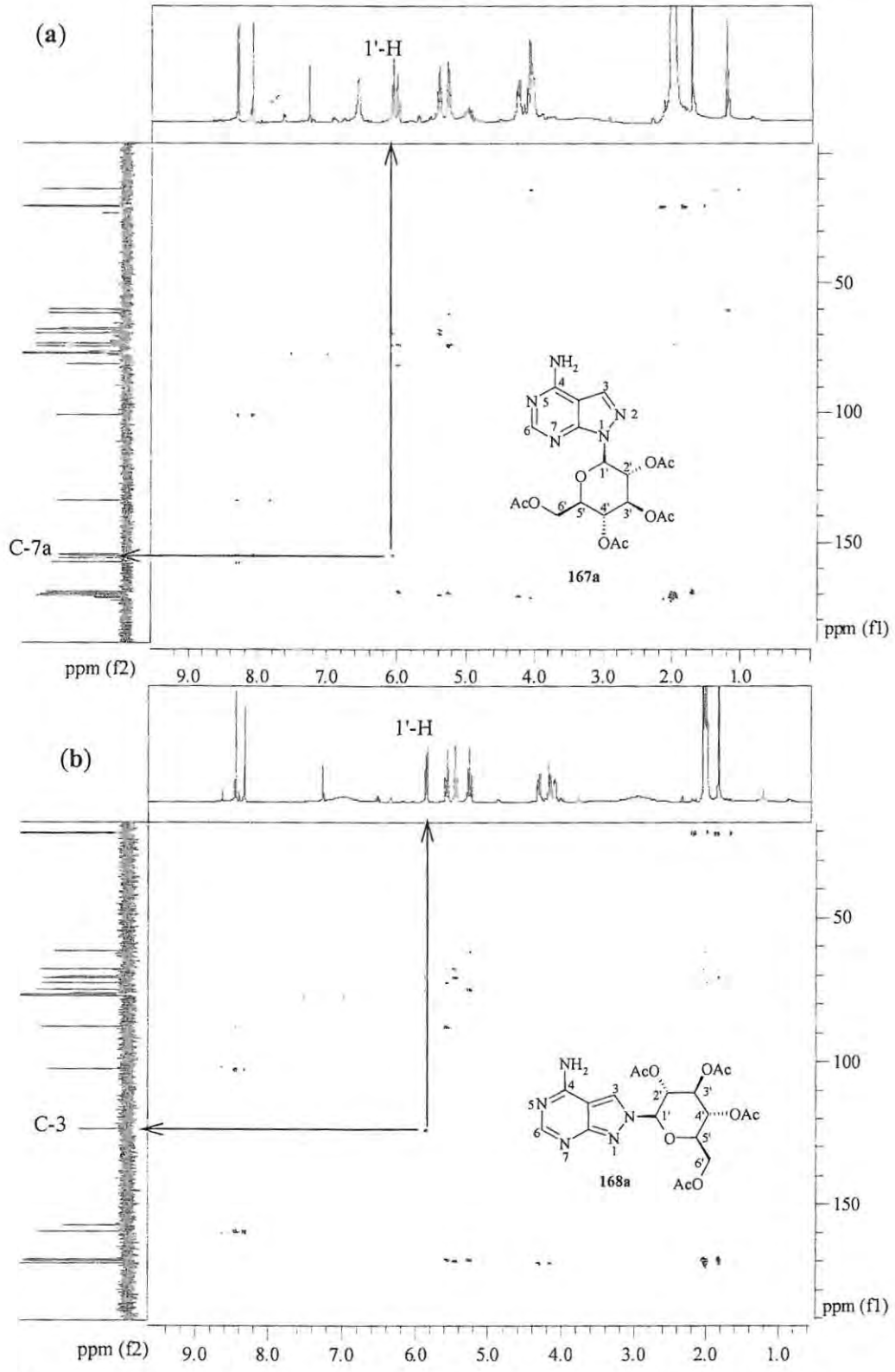


Figure 43: HMBC spectra of compounds **167a** (**a**) and **168a** (**b**) in $CDCl_3$.

Compounds **167a** and **b** were then deacetylated to afford the polyhydroxy derivatives **169a** and **b (Scheme 39)**. Deacetylation was carried out using sodium methoxide (prepared *in situ*) in methanol. After neutralization (using the cation exchange resin, Dowex 50Wx8) and HPLC purification, compounds **169a** and **b** were obtained as white powders in very good yields (60% and 94%, respectively). ^{1}H NMR analysis, in each case, revealed the absence of the four acetyl methyl singlets at *ca.* 2ppm, while the ^{13}C NMR spectra (illustrated for compound **169a** in **Figure 44**) also revealed the absence of the acetyl methyl and carbonyl signals at *ca.* 20 and 170ppm, respectively, thus confirming formation of compounds **169a** and **b**.

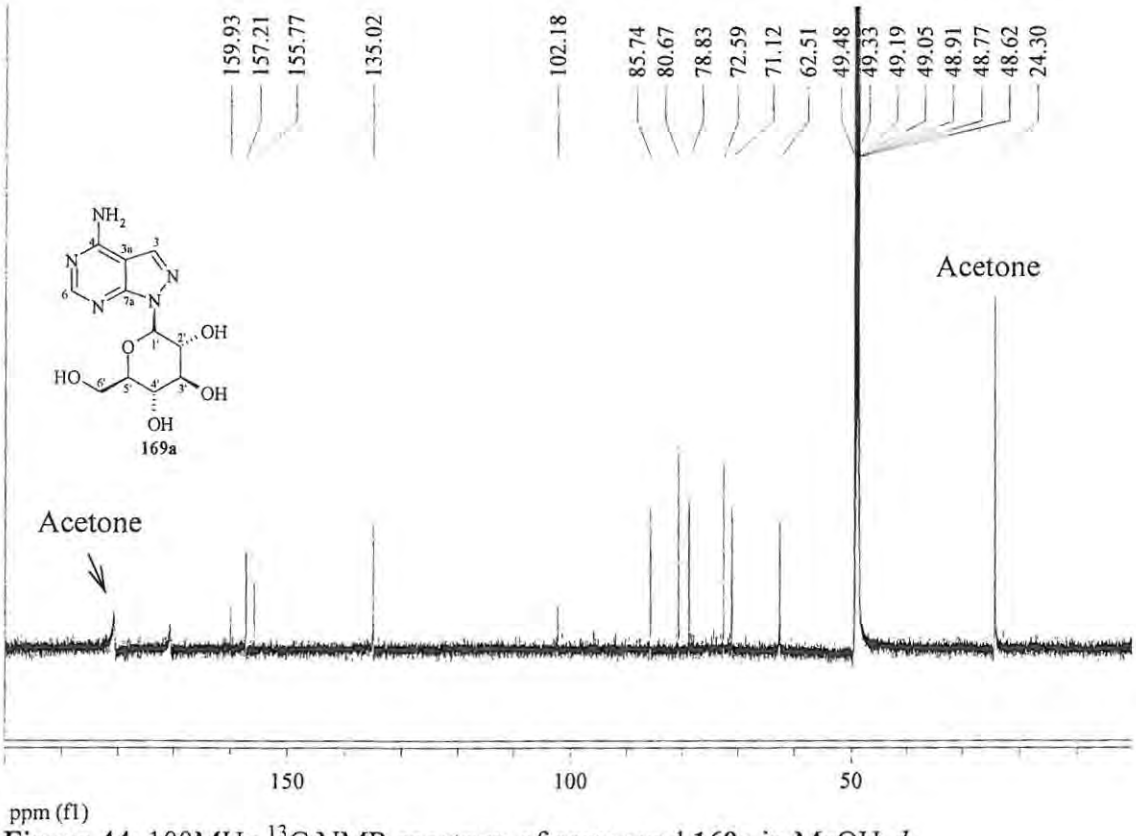
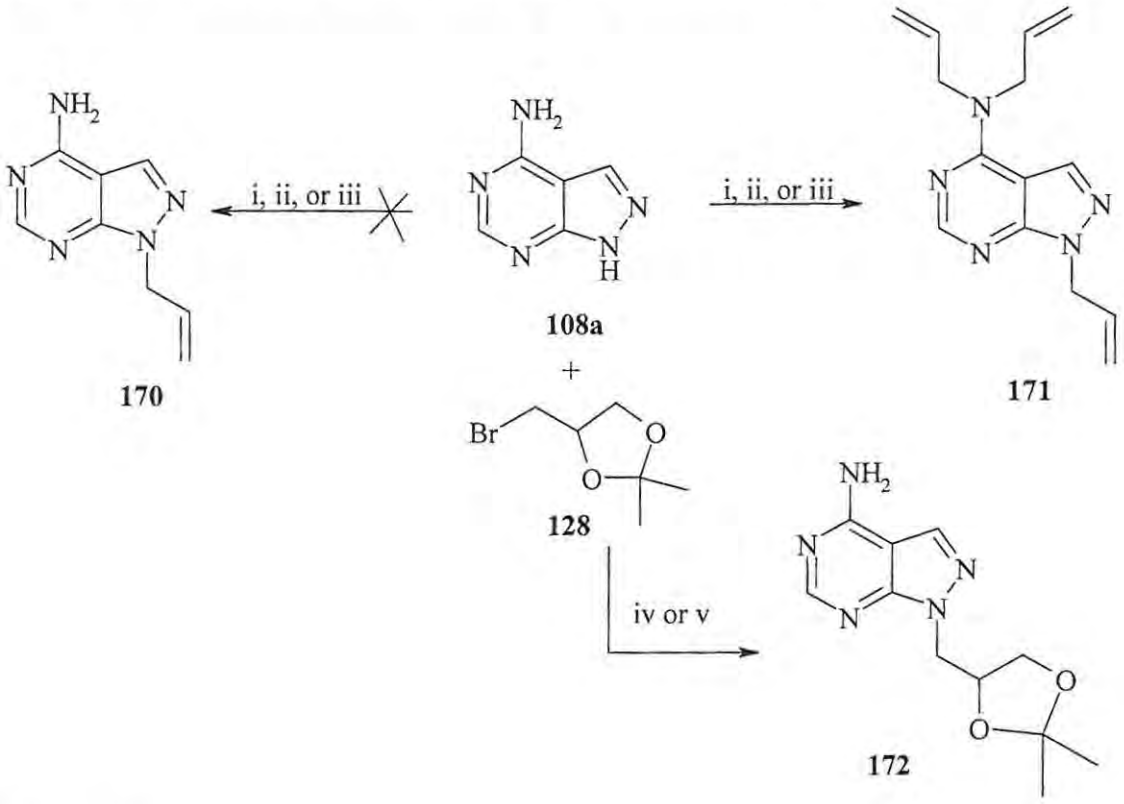


Figure 44: 100MHz ^{13}C NMR spectrum of compound **169a** in MeOH-d_4.

2.1.4.2 Alkylation of pyrazolo[3,4-d]pyrimidine system

Following the success at introducing the cyclic glucosyl derivative to the pyrazolo[3,4-*d*]pyrimidine system, attempts were then made to introduce other groups. The first attempt was to alkylate 4-aminopyrazolo[3,4-*d*]pyrimidine **108a** with excess allyl bromide in the presence of potassium bicarbonate following the method described by Holy *et al.*[184] (**Scheme 40**). Purification of the mixture following workup by flash chromatography and HPLC afforded a fraction, which was indicated by mass spectrometry to be the trialkylated product **171**. Attempts to produce the monoalkylated product **170** by using one equivalent of allyl bromide or by changing the base to sodium hydride were unsuccessful as compound **171** was again isolated in both cases.



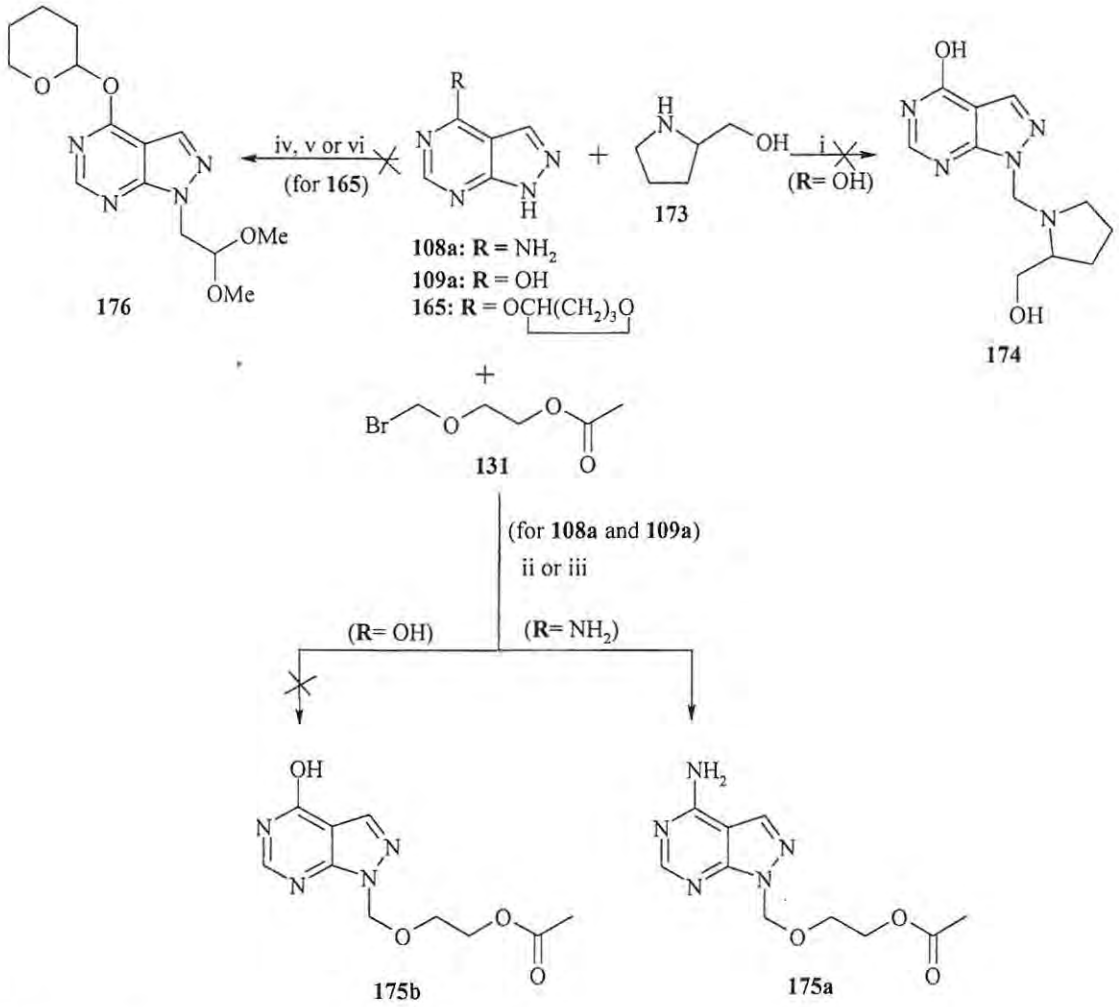
Scheme 40:

Reagents and conditions: i) K_2CO_3, DMF, allylbromide (3x), 140°C; ii) K_2CO_3, DMF, allylbromide, 140°C, iii) NaH, DMF, allylbromide, 140°C; iv) K_2CO_3, DMF, 140°C; v) NaH, DMF, 140°C.

However, selective alkylation by the bromoketal **128** at the N-1 position of 4-APP **108a** was achieved using the method reported by Holy *et al.*[184] Thus 4-APP **108a** was treated with 3 equivalents of the bromoketal **128** and with K_2CO_3 in DMF under reflux (**Scheme 40**). After purification by column chromatography the desired product **172** was isolated as an oil in 22% yield. In an attempt to improve the yield, the reaction was repeated replacing K_2CO_3 with NaH and refluxing for 24h (**Scheme 40**). This method gave a slight improvement in the yield, with compound **172** being isolated in 28% yield after purification by column chromatography. The structure of compound **172** was verified by ^1H and ^{13}C NMR spectroscopy. The ^1H NMR spectrum revealed a broad singlet at 5.8ppm and two sharp singlets at *ca.* 1.3ppm corresponding to the amino and diastereotopic methyl protons, respectively. The ^{13}C NMR spectrum revealed 11 signals, which correspond to the expected signals for compound **172**.

Attachment of prolinol **173** at the N-1 position of allopurinol **109a** was explored using the Mannich reaction (**Scheme 41**).[185] However, this proved unsuccessful affording formaldehyde polymer and an oil, which showed none of the expected ^1H NMR signals, for the desired product **174**. Attempts were then made to link a polyoxy moiety to 4-APP **108a** and allopurinol **109a** using 2-(bromomethoxy)ethyl acetate **131**. Following the method reported by Robins,[155] 4-APP and allopurinol were each treated with NaH and compound **131** in dry DMF, initially at -63^{O}C and then at, room temperature (**Scheme 41**). After purification by flash chromatography, compound **175a** was isolated as off-white crystals in low yield (14%). The structure of the pure compound was confirmed by spectroscopic methods. Thus the ^1H NMR spectrum (**Figure 45a**) reveals the presence of a broad singlet at *ca.* 6.0ppm, integrating for two protons and corresponding to the amino group, and two sharp singlets at 5.9 and 2.0ppm corresponding to the *N*-methylene and methyl protons, respectively. The ^{13}C NMR spectrum (**Figure 45b**) reveals the presence of a carbonyl signal at 171ppm and methyl signal at 21ppm. However, compound **175b** was not isolated from the corresponding reaction with allopurionol **109a**. Instead the starting allopurinol was isolated. Attempts were made to either improve the yield of compound **175a** by varying the temperature, but these resulted in compound **175a** being isolated in even lower yield. Efforts to access the compounds **175a** and **b** by employing

microwave technology were also unsuccessful as 4-APP and allopurinol were isolated unchanged following preparative layer chromatography of the crude product in each case.



Scheme 41:

Reagents and conditions: i) 40% formaldehyde, MeOH; ii) NaH, DMF, -63^{O}C, r.t.; iii) NaH, DMF, 0-70^{O}C; iv) K_2CO_3, DMF, microwave; v) NaH, THF, chloroacetyldehyde dimethyl acetal, r.t.; vi) NaH, THF, chloroacetyldehyde dimethyl acetal, reflux; vii) NaH, DMF, chloroacetaldehyde dimethyl acetal, reflux.

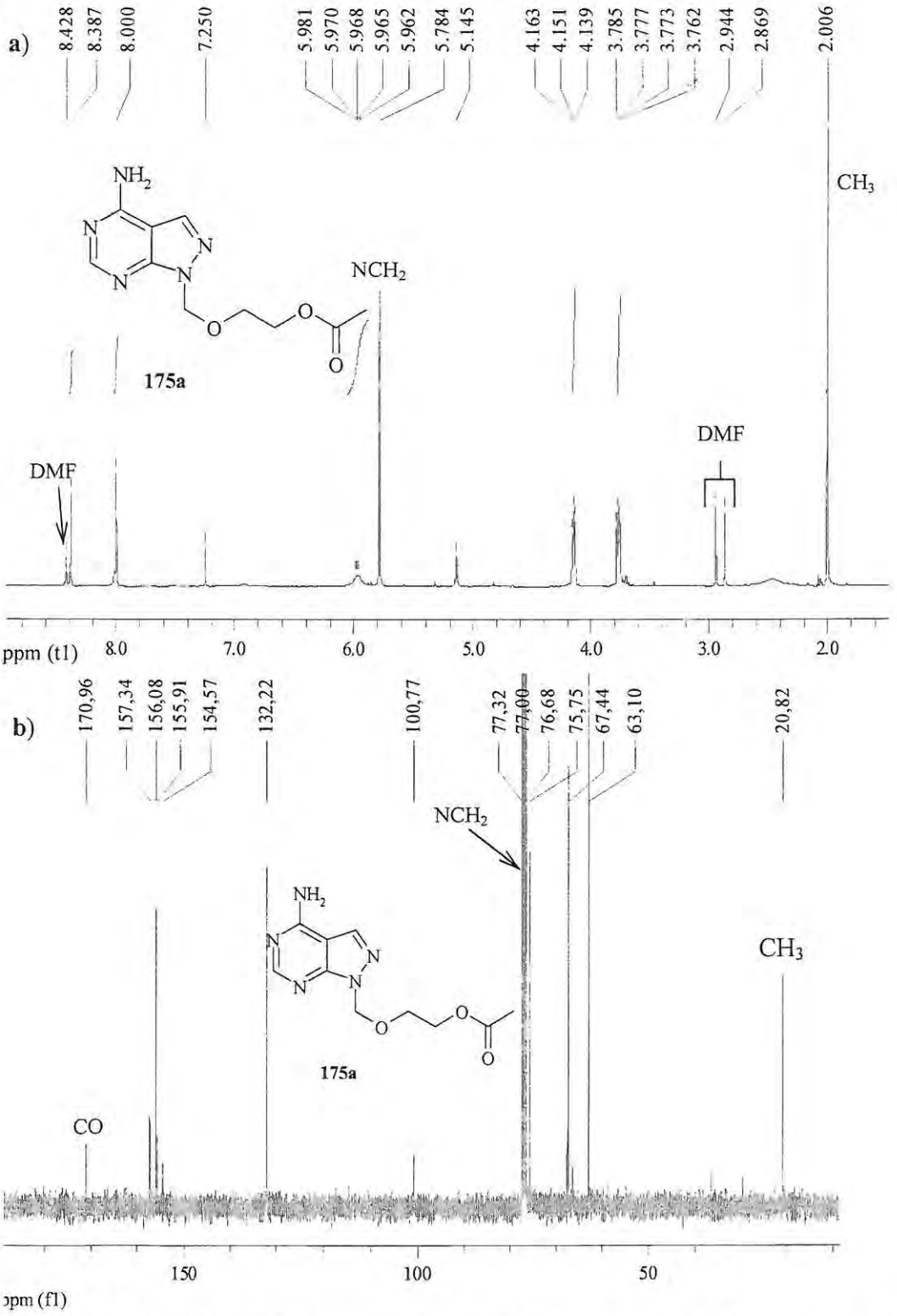
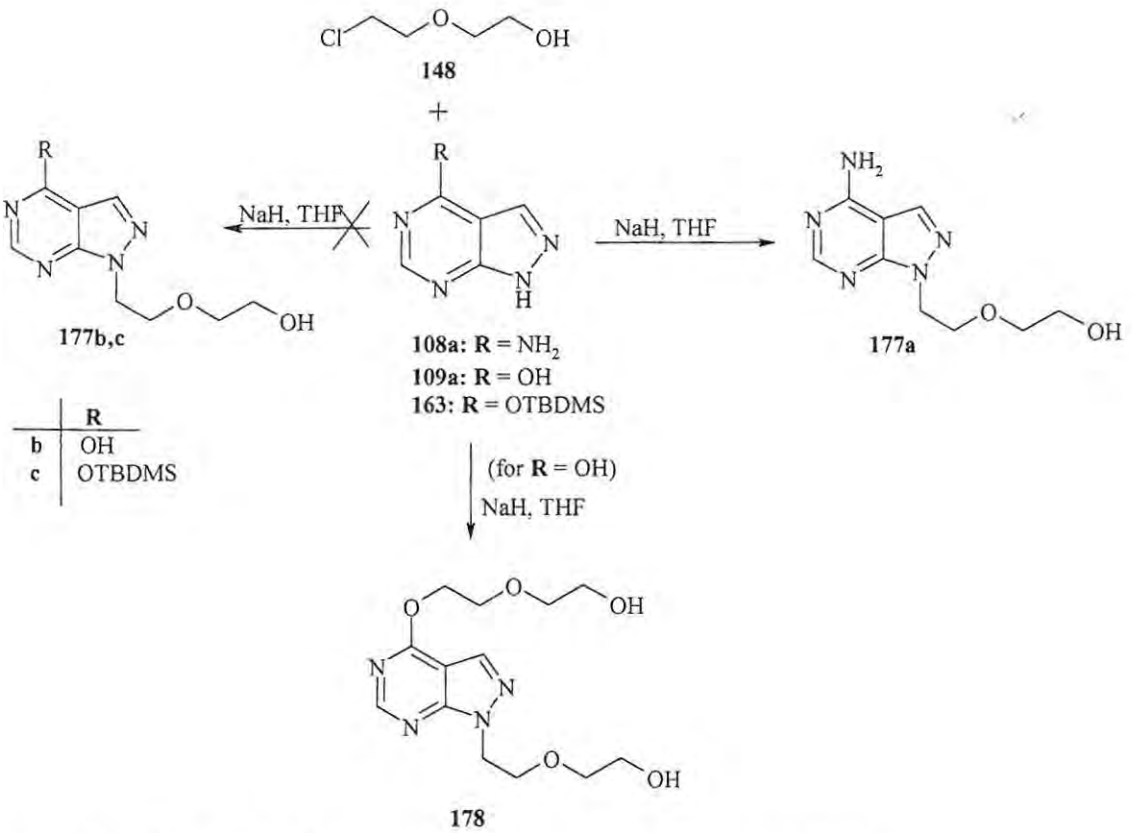


Figure 45: **a**) 400MHz ^{1}H and **b**) 100MHz ^{13}C NMR spectra of compound **175a** in $CDCl_3$.

Given the difficulties encountered in introducing the polyoxy moiety into the allopurinol system, it was decided to use the protected acetaldehyde, chloroacetaldehyde dimethyl acetal, on the allopurinol derivative **165**. It was expected that, after deprotection, the resulting aldehyde could be subjected to the Baylis-Hilman reaction to afford a longer side-chain substituent. Compound **165** was treated with chloroacetaldehyde dimethyl acetal in the presence of NaH and dry THF at room temperature. Following work-up, instead of the desired product **176**, the pure precursor **165** was isolated, as shown by ^{1}H NMR analysis. Further unsuccessful attempts were made to obtain compound **176** by varying the temperature and changing the solvent but the precusor **165** was isolated in each case.

Since compounds **172** and **175a** were isolated in low yield no attempts were made to convert them into their corresponding hydroxy derivatives. Instead, the hydroxyl-bearing substituent was introduced directly by alkylating 4-APP **108a**, allopurinol **109a** or the TBDMS derivative **163** with 2-(chloroethoxy)ethanol **148**. 4-APP, allopurinol and compound **163** were each treated with sodium hydride and compound **148** in dry DMF under a stream of nitrogen (**Scheme 42**). Following flash chromatography, the desired compound **177a** was isolated as a pale yellow oil in low yield (12%). The structure of the product was verified by NMR spectroscopic methods. The ^{1}H NMR spectrum (**Figure 46**) reveals a broad singlet at 7.6ppm integrating for two protons, which corresponds to the amino substituent on the pyrimidine ring, and two triplets at 3.8 and 4.4ppm, integrating for two protons each and corresponding to the amine and hydroxyethylene protons; the remaining methylene proton signals overlap with the DMSO-d_6 H_2O signal. The spectrum also revealed that the compound was contaminated with traces of 4-APP. However, the expected products **177b** and **c**, from the reaction of allopurinol and derivative **163**, were not isolated. Flash chromatography of the crude material obtained from the reaction with allopurinol afforded a fraction, spectroscopic analysis of which indicated formation of the bis-alkylated derivative **178**. The ^{13}C NMR spectrum revealed 13 signals which suggested that both *N*- and *O*-alkylation had occurred, while the low-resolution mass spectrometric analysis revealed a peak corresponding to the molecular ion for compound **178**.



Scheme 42: Alkylation of the pyrazolo[3,4-*d*]pyrimidine systems with (2-chloroethoxy)ethanol **148**.

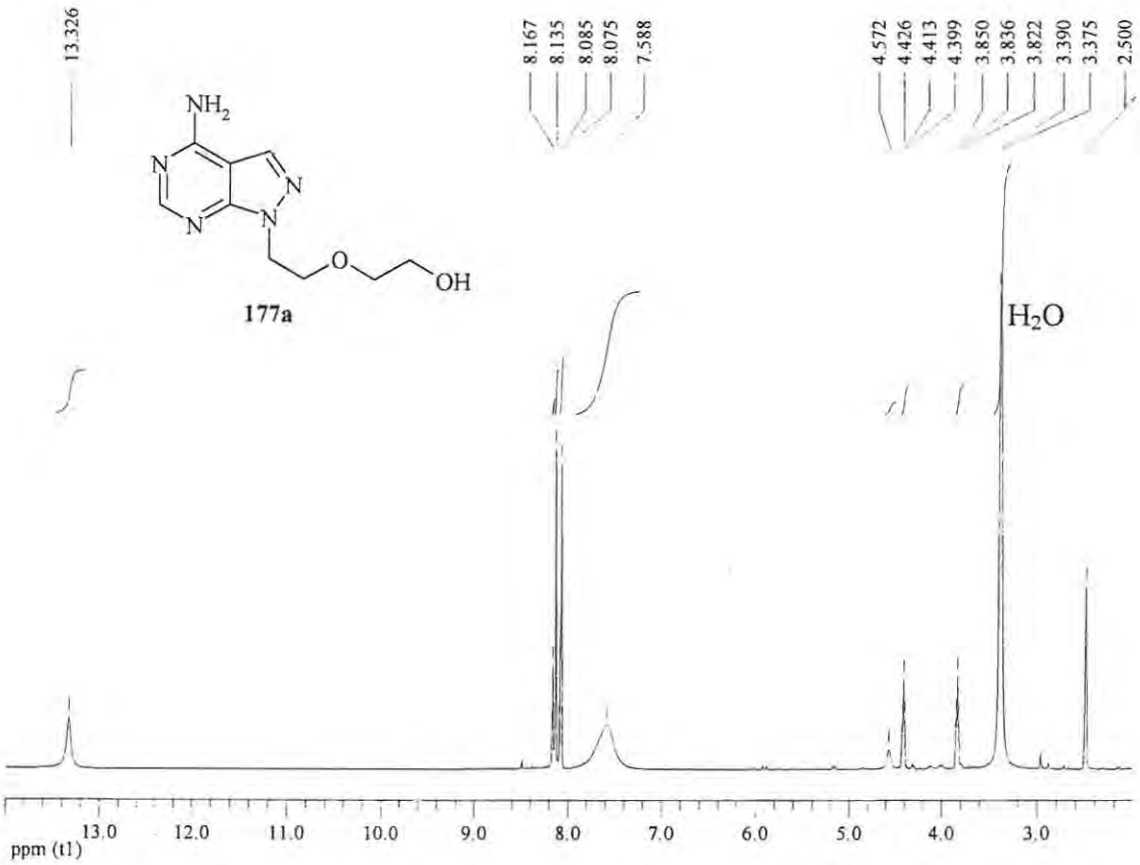
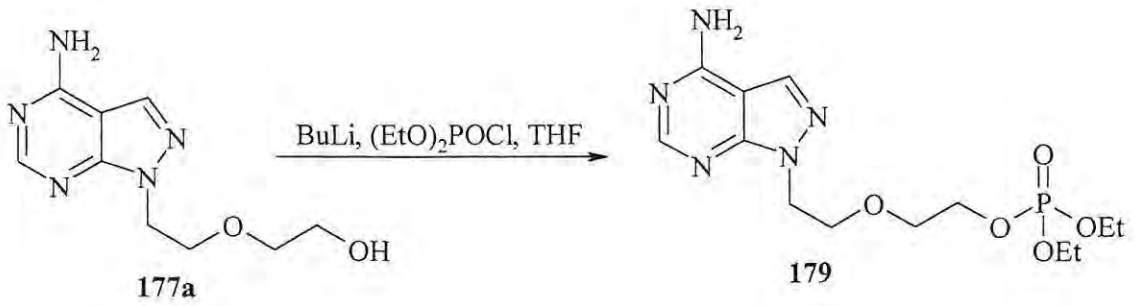


Figure 46: 400MHz ^{1}H NMR spectrum of compound **177a** in DMSO-d_6.

2.4.1.3 Phosphorylation of the pyrazolo[3,4-d]pyrimidine derivatives

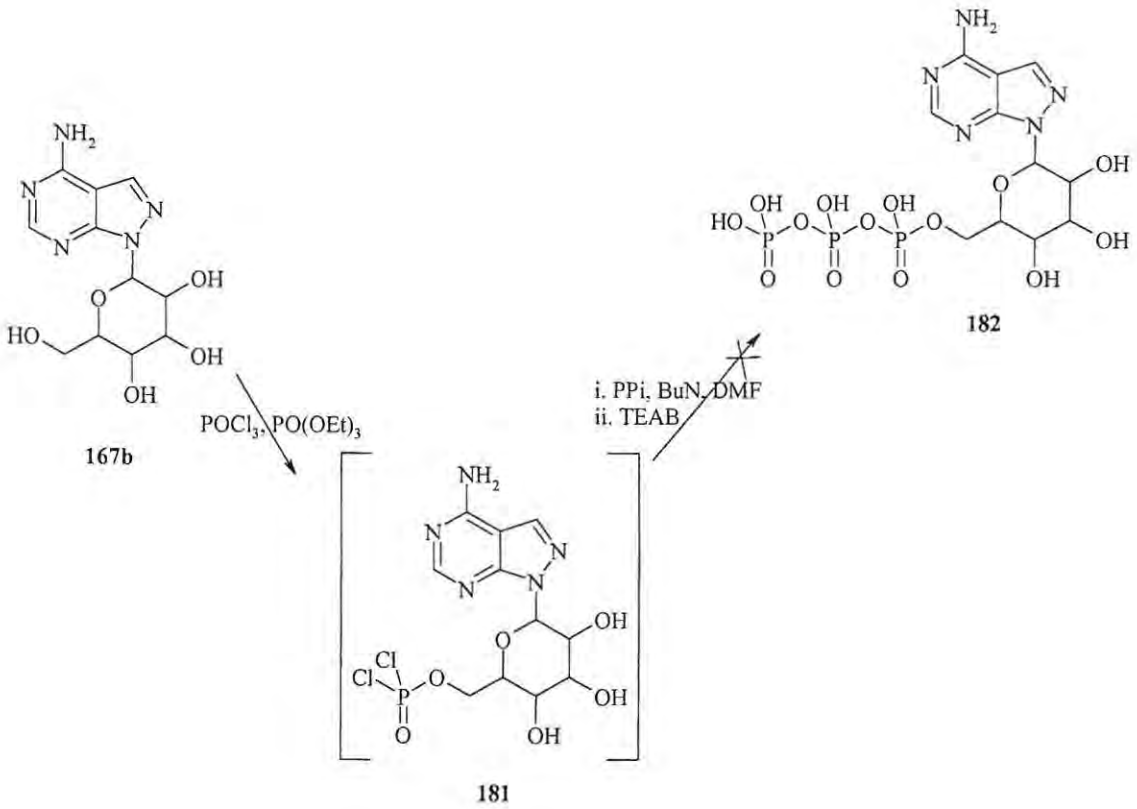
Following the successful formation of the hydroxylated derivatives **169** and **177a**, attempts were made to introduce the phosphate ester moiety. The monophosphate derivative **179** (**Scheme 43**) was prepared following a method described by Takemoto.[173] Compound **177a** was treated with diethyl chlorophosphate and BuLi in the presence of THF at 0^{O}C. After column chromatographic purification, compound **179** was isolated as a golden oil in low yield (14%). The structure of compound **179** was confirmed by NMR spectroscopy. The ^{1}H spectrum revealed multiplets at *ca* 1.3 and 3.5-4.3ppm corresponding to the phosphate ester methyl and methylene protons, respectively, while the ^{31}P NMR spectrum showed a single signal at -17.4ppm.



177a **179**

Scheme 43: Synthesis of the phosphate ester derivative **179**

Attempts were also made to synthesize a triphosphate ester derivative of the 4-aminopyrazolo[3,4-*d*]pyrimidine derivative following methods reported in literature.[32-35] Thus, compound **167a** was treated with "Proton Sponge" and $POCl_3$ in triethyl phosphate at $0^{o}C$ to form the intermediate **181** (**Scheme 44**). However attempts to access the corresponding triphosphate ester **182** using tributylammonium pyrophoshate and tributylamine in dry DMF proved unsuccessful.



Scheme 44: Attempted synthesis of the triphosphate ester derivative **182.**

2.4.1.4 Structural Comparison of the Pyrazolo[3,4-d]pyrimidine Derivatives with ATP

The various energy-minimized structures (**Figure 47**) of 4-aminopyrazolo[3,4-d]pyrimidine and allopurinol glucosyl derivatives **167b, 167a, 169b, 169a** as well as the 4-aminopyrazolo[3,4-d]pyrimidine alkyl derivative **175a** and the phosphorylated derivative **179** were aligned with ATP to check their similarities. The overlayed (**Figure 48**) global-minimum structures show reasonable structural homology with ATP. The Cerius^2 Ligand Fit module was used to determine the most favourable bound conformations and investigate the interaction of these derivatives with the amino acids in the enzyme active site (Section 2.3.2.4).

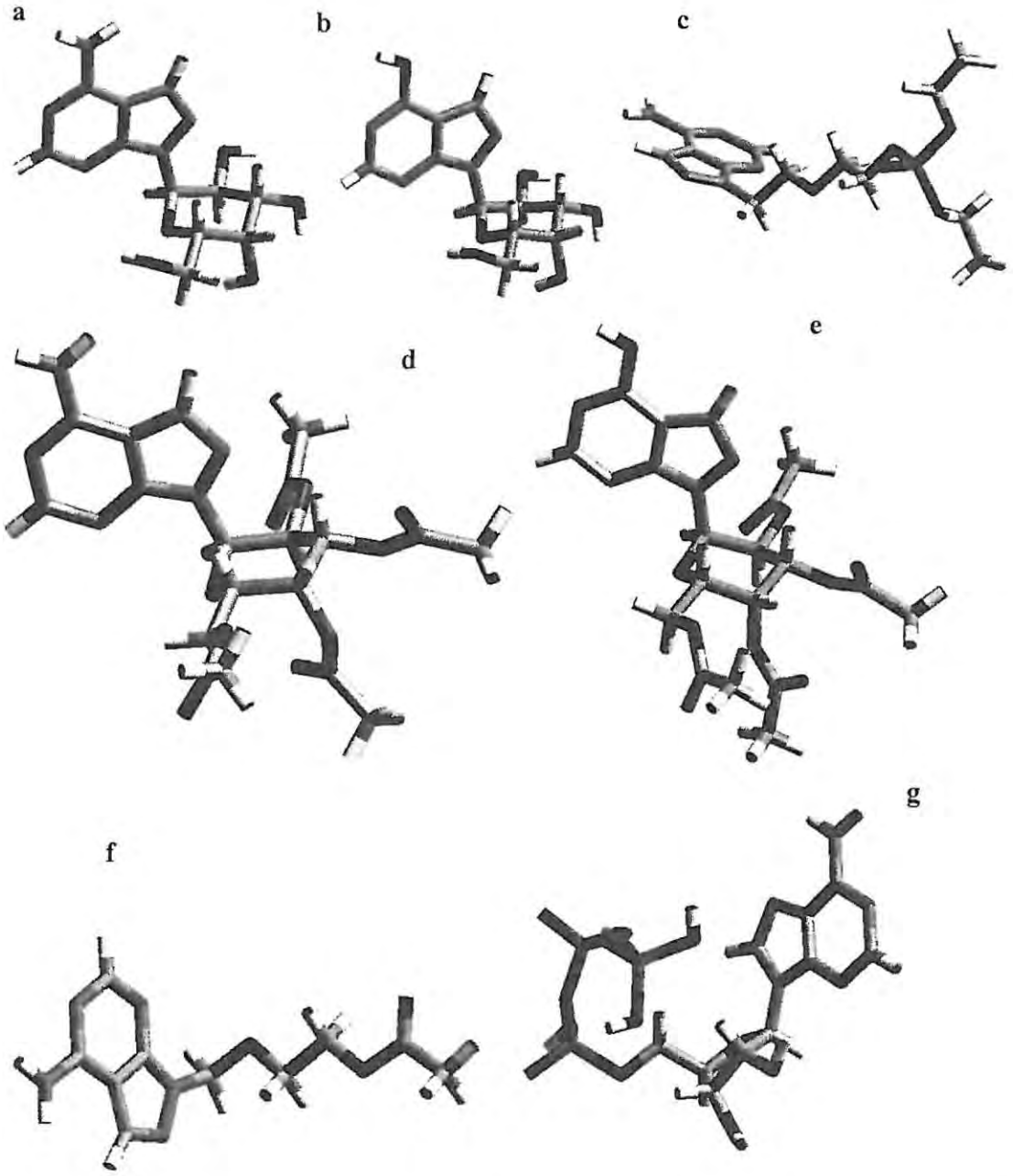


Figure 47: Energy-minimized structures of compounds **169b** (a), **169a** (b), **179** (c), **167b** (d), **167a** (e), **175a** (f) and ATP (g).

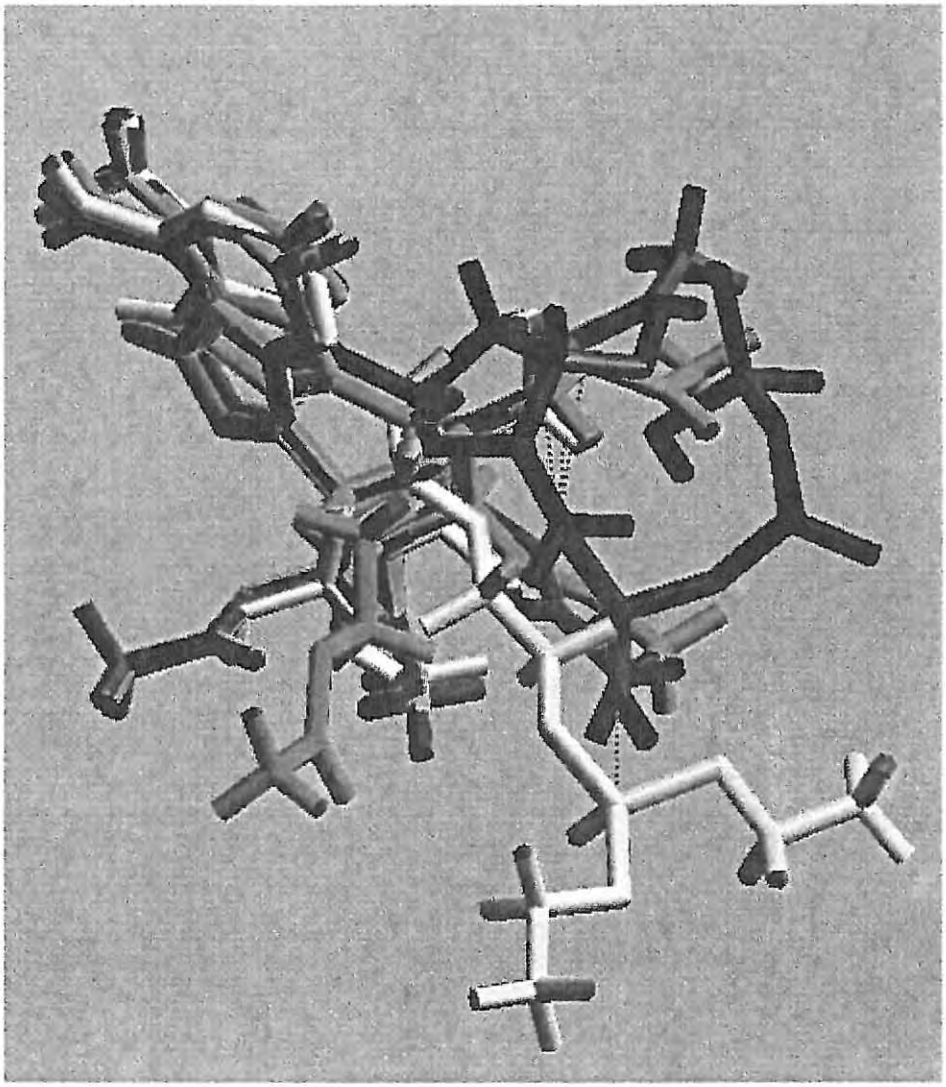


Figure 48: Alignment of the energy-minimized structures of compounds **167b** (pink), **167a** (light-blue), **179** (yellow), **169b** (white), and **175a** (light-green), **169a** (orange) with ATP (red).

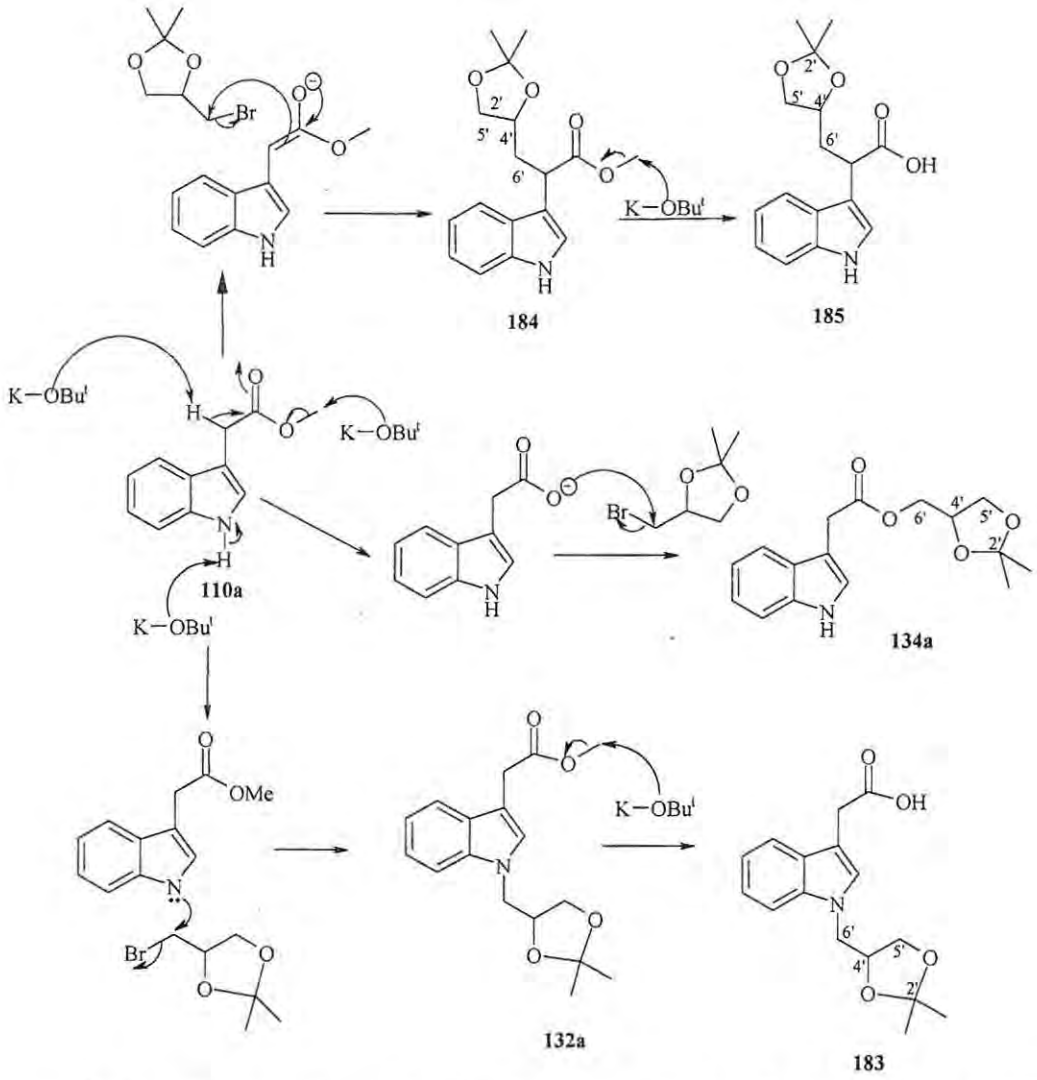
2.2 NMR Chemical shift prediction for various analogues

2.2.1 NMR chemical shift prediction for 3-indolylalkanoic acid derivatives

As indicated in Section 2.1.2.3, NMR spectroscopy was used to verify the structure of compound **134a**. The alkylation of methyl 3-indolylacetanoate **110a** with 4-bromomethyl-2,2-dimethyl-1,3-dioxolane **128** in the presence of a base can, in principle, result in the formation of one or more of the three products, compound **132a** *via* *N*-alkylation, compound **184** *via* *C*-alkylation and compound **134a** *via* *O*-alkylation. However, it has been reported that under anhydrous conditions and at low to room temperature only *N*-alkylation occurs.[155] The low-resolution mass spectrometric analysis data for the compound isolated from this reaction revealed a peak with *m/z* 289, corresponding to compound **134a**, which suggests *O*-alkylation. Other possibilities are that either compound **132a** or **184** was formed and then reacted with $KOBu^t$ to form compound **183** or **185**, respectively (**Scheme 45**), both of which are isomeric with compound **134a**.

The Modgraph NMRPredict version 3.2.2 programme was employed to differentiate between the three isomers. The programme, is simple to use and can predict both ^1H and ^{13}C NMR spectra.[186] For ^1H NMR both the frequency and solvent must be specified and, in this case, the data were predicted for chloroform at 400MHz, while in ^{13}C NMR this information is not provided. The ^{13}C NMRPredict facility uses two approaches for ^{13}C NMR prediction, the HOSE (Hierarchical Organisation of Spherical Environments) code approach which was developed by Bremser[187] and a unique Nnet (Neural Network) algorithm. The HOSE code assumes that the molecule is spherical and measures increasing radii starting at the carbon atom whose chemical shift is to be predicted. The larger the number of spheres (three spheres or more) the more reliable is the data, and the NMRPredict program goes to a maximum of five spheres. However, if the query structure is not well represented in the database and the atom can only be predicted to one

or two spheres the prediction is not reliable. The Nnet is a general approach and is not restricted by ring size.



Scheme 45: Possible reaction pathways for the synthesis of the compound with a molecular ion peak (*m/z*) of 289

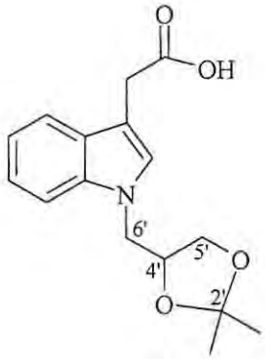
The predicted 400MHz ^{1}H NMR spectra for structures **183** and **185** indicated a sharp singlet at *ca.* 11ppm corresponding to the carboxylic proton, while the spectrum of compound **134a** lacks this signal. The predicted spectrum of structure **185** also reveals two triplets at *ca.* 2.4ppm corresponding to the diastereotopic 6'-methylene protons, while in structures **134a** and **183** these diastereotopic protons are expected as double doublets at

ca. 4.5ppm. The acetyl protons are predicted to resonate at *ca.* 4ppm as a double doublet in structure **185** and as two doublets in structures **134a** and **183**. The aromatic protons are predicted at 7.2-7.8, 7.2-7.6 and 7.2-7.7ppm in structures **134a**, **185** and **183**, respectively. Comparing the predicted data with the experimental ^{1}H NMR spectral data for the isolated compound reveals the absence of signals at *ca.*2 and 11ppm, which rules out the presence of compound **185**. Based on the previous observations, the broad signal at 8ppm was assigned to the amine proton instead of the carboxylic acid proton since the latter can be shifted downfield under certain conditions. The predicted ^{13}C NMR spectra of the three isomers reveals 15 carbon signals instead of 16 because the ketal methyl signals are presumed to be equivalent. The 6'-methylene signal is expected at 65, 48 and 32ppm for structures **134a**, **183** and **185**, respectively, while the ^{13}C NMR spectrum of the isolated compound revealed 16 carbon signals (since the two ketal methyl carbons are not equivalent) and the 6'-methylene carbon signal was observed at 65ppm, which is the region for ethers. These correlations confirmed that the isolated material is, in fact, the *O*-alkylated compound **134a**.

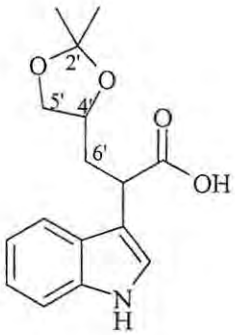
Table 6: ^{13}C NMR data for the possible compounds with *m/z* 289

^{13}C	**134a**[a]	**134a**[b]	**183**[b]	**185**[b]
CH_3	25.4	25.9	26.2	26.3
CH_3	25.6	25.9	26.2	26.3
2''	31.1	33.4	31.3	34.1
6'	64.9	65.1	48.2	32.0
5'	66.3	66.7	66.7	69.7
4'	73.6	73.6	72.7	72.6
3	108.3	105.2	112.3	109.7
2'	109.8	109.7	109.1	109.1
4	111.2	119.6	117.8	120.8
7	118.8	111.6	109.3	112.7
6	119.7	121.0	120.9	123.3
5	122.3	121.1	126.6	122.7
2	123.0	124.8	128.2	121.9
3a	127.2	128.1	128.3	127.6
7a	136.1	137.6	139.8	137.5
C O	171.7	171.2	173.6	179.2
RMSE		**3.04**	**5.57**	**9.05**

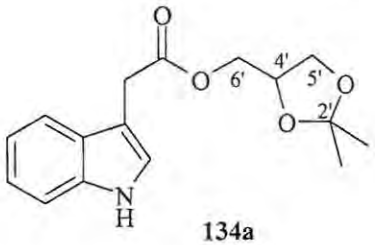
[a] Experimental data; [b] predicted data



183

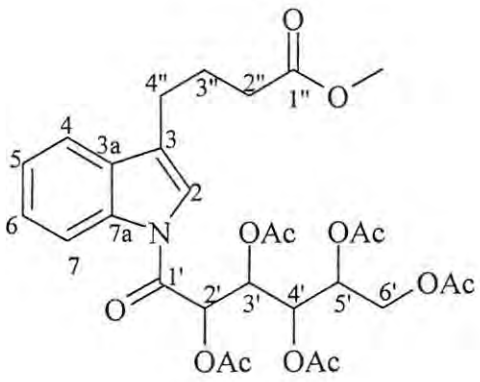


185



134a

The ^{13}C NMR spectra of the glycosylated indolylalkanoic acid and benzimidazole derivatives were also predicted using Modgraph and ChemWindow NMRPredict programmes, and their RMSE (root mean square error) values were compared to determine the programme affording the most reliable data. The results are summarized in **Tables 7-11** and the errors are less than 5.0 for the indolylalkanoic derivatives and 6.0 for the benzimidazole derivative. In general, the Hose data are more comparable to the experimental data than the Nnet and ChemWindow data, except for compound **116c** where the experimental data is more comparable to the Nnet data. The ChemWindow data is the least comparable since the programme does not differentiate the acetyl and methyl carbons.



116c

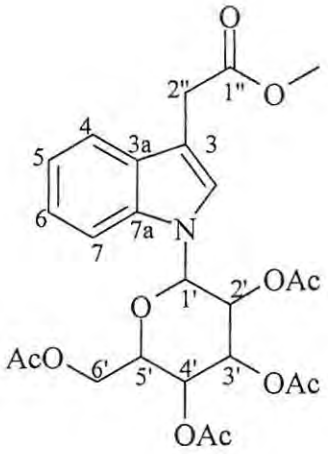
Table 7: Experimental and predicted ^{13}C NMR data for derivative **116c**

^{13}C	A*	B*	C*	D*	^{13}C	A*	B*	C*	D*
3''	24.1	24.0	24.6	22.5	**5**	123.6	122.9	123.8	121.7
4''	24.3	26.0	29.3	34.6	**3**	124.3	108.2	117.1	112.1
2''	33.1	31.5	33.4	33.7	**6**	125.7	125.4	123.8	119.6
OCH_3	51.5	51.5	51.9	50.4	**3a**	130.6	130.0	129.0	131.6
6'	61.8	62.3	62.7	64.7	**7a**	136.4	136.0	135.7	136.5
3'	68.6	68.8	70.5	69.1	**1'**	164.1	171.0	170.4	177.6
5'	68.6	68.2	69.2	70.2	**3'CO**	169.5	170.2	171.3	171.0
4'	68.9	68.5	67.4	68.5	**4'CO**	169.6	170.2	171.3	171.0
2'	70.1	71.2	71.6	76.5	**2'CO**	169.7	169.8	170.8	171.0
4	116.9	118.1	120.6	120.5	**5'CO**	169.9	170.6	171.0	171.0
7	119.0	115.3	114.6	111.0	**6'CO**	170.7	170.7	171.7	171.0
2	120.4	124.2	125.2	122.8	**1''**	173.6	173.6	174.7	172.0
					RMSE		**3.80**	**2.89**	**4.95**

*A: experimental; B: HOSE; C: Nnet and D: ChemWindow

Table 8: Experimental and predicted ^{13}C NMR data for derivative **138a**

^{13}C	A*	B*	C*	D*
2'CH_3	20.1	20.4	21.2	17.6
3'CH_3	20.6	20.7	21.2	17.6
4'CH_3	20.6	20.7	21.2	17.6
6'CH_3	20.7	20.6	20.7	17.3
CH_2CO	31.0	30.6	33.1	38.5
OCH_3	51.9	52.0	51.9	50.5
6'	61.9	61.9	62.6	68.5
4'	68.2	67.9	68.4	69.3
2'	70.3	69.3	69.4	72.8
3'	73.4	72.8	70.8	69.0
5'	74.6	74.3	73.1	67.9
1'	83.3	84.0	84.6	85.9
4	109.8	118.6	121.7	120.5
3	110.2	108.8	116.2	113.1
7	119.4	111.7	111.6	111.0
6	120.7	126.6	125.0	119.6
5	122.7	121.0	121.9	121.7
2	123.3	119.3	124.2	125.6
3a	128.5	127.5	129.2	132.6
7a	136.3	135.3	137.9	139.3
2'CO	168.7	168.5	171.2	171.0
4'CO	169.4	169.7	171.2	171.0
3'CO	170.2	169.9	171.2	171.0
6'CO	170.6	171.6	171.8	171.0
1''	171.9	172.8	173.0	172.0
RMSE		**2.81**	**3.42**	**4.25**

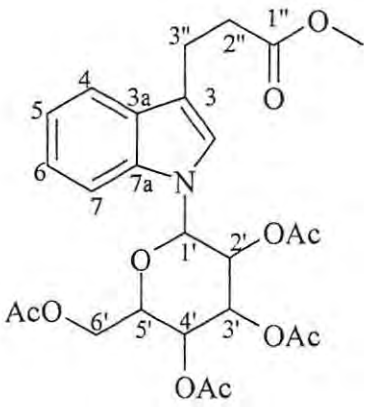


138a

*A: experimental; B: HOSE; C: NNet and D: ChemWindow

Table 9: Experimental and predicted ^{13}C NMR data for derivative **138b**

^{13}C	A*	B*	C*	D*
2'CH_3	20.1	20.4	21.2	17.6
3''	20.5	22.0	25.9	30.6
3'CH_3	20.6	20.7	21.2	17.6
4'CH_3	20.6	20.7	21.2	17.6
6'CH_3	20.7	20.6	20.7	17.3
2''	34.4	33.7	35.3	30.4
OCH_3	51.6	51.6	51.8	50.4
6'	61.9	61.9	62.6	68.5
4'	68.2	67.9	69.1	69.3
2'	70.2	69.3	69.4	72.8
3'	73.5	72.8	66.4	69.0
5'	74.6	74.3	73.1	67.9
1'	83.1	84.0	84.6	85.9
4	109.6	118.6	121.3	120.5
3	116.7	113.5	113.9	113.1
7	119.2	111.8	111.1	111.0
5	120.4	121.0	122.1	121.7
2	121.4	119.3	120.8	125.6
6	122.6	126.6	123.9	119.6
3a	128.5	127.5	129.6	132.6
7a	136.6	135.3	137.3	139.3
2'CO	168.7	168.5	171.1	171.0
4'CO	169.4	169.7	171.1	171.0
3'CO	170.2	169.9	171.1	171.0
6'CO	170.6	171.6	171.7	171.0
1''	173.5	173.7	173.8	172.0
RMSE		**2.60**	**3.49**	**4.57**

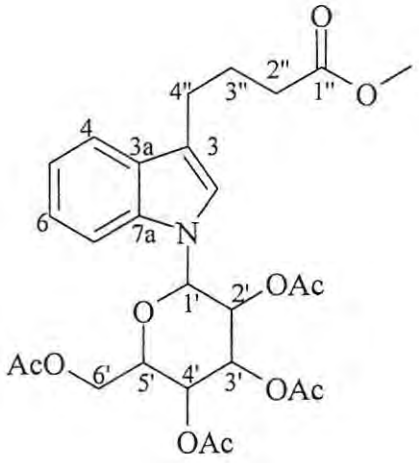


138b

*A: experimental; B: HOSE; C: Nnet and D: ChemWindow

Table10: Experimental and predicted ^{13}C NMR data for derivative **138c.**

^{13}C	A*	B*	C*	D*
2'CH_3	20.1	20.4	21.2	17.6
3'CH_3	20.6	20.7	21.2	17.6
4'CH_3	20.6	20.7	21.2	17.6
6'CH_3	20.7	20.6	20.7	17.3
4''	24.4	26.0	29.3	34.9
3''	25.0	24.6	24.0	22.5
2''	33.5	31.5	33.4	33.7
OCH_3	51.5	51.5	51.8	50.4
6'	62.0	61.9	62.6	68.5
4'	68.2	67.9	69.1	69.3
2'	70.3	69.3	69.4	72.8
3'	73.4	72.8	66.4	69.0
5'	74.5	74.3	73.1	67.9
1'	83.0	84.0	84.6	85.9
4	109.5	118.6	121.3	120.5
3	117.3	113.5	112.8	113.1
7	119.4	111.7	111.1	111.0
5	120.2	121.0	122.1	121.7
2	121.5	119.3	119.0	125.6
6	122.4	126.6	123.9	119.6
3a	126.7	127.5	129.6	132.6
7a	136.7	135.3	137.3	139.3
2'CO	168.7	168.5	171.1	171.0
4'CO	169.4	169.7	171.1	171.0
3'CO	170.1	169.9	171.1	171.0
6'CO	170.6	171.6	171.7	171.0
1'	174.0	173.6	174.6	172.0
RMSE		**2.67**	**3.56**	**4.57**

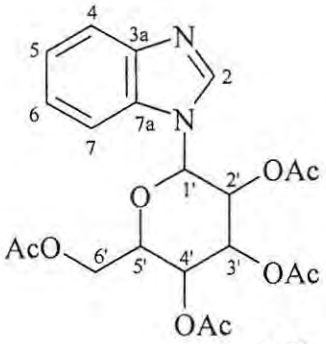


138c

*A: experimental; **B**: HOSE; **C**: Nnet and **D**: ChemWindow

Table 11: Experimental and predicted ^{13}C NMR data for derivative **156**

^{13}C	A*	B*	C*	D*
2'CH_3	20.1	20.4	21.2	17.6
3'CH_3	20.6	20.7	21.2	17.6
4'CH_3	20.6	20.7	21.2	17.6
6'CH_3	20.7	20.6	20.7	17.3
6'	61.9	61.9	62.6	68.5
4'	68.2	67.9	69.1	69.3
2'	70.3	69.3	69.4	72.5
3'	73.4	72.8	66.4	68.6
5'	74.6	74.3	73.1	67.5
1'	83.3	84.0	82.1	82.0
4	109.8	120.5	119.6	115.4
7	119.4	111.7	113.9	115.4
6	120.7	126.6	123.9	122.9
5	122.7	124.0	124.1	122.9
3a	128.5	139.0	138.8	137.9
7a	136.3	135.3	136.6	139.3
2	123.3	139.9	144.7	141.5
2'CO	168.7	168.5	171.1	171.0
4'CO	169.4	169.7	171.1	171.0
3'CO	170.2	169.9	171.7	171.0
6'CO	170.6	171.6	171.7	171.0
RMSE		**5.35**	**6.06**	**5.56**



156

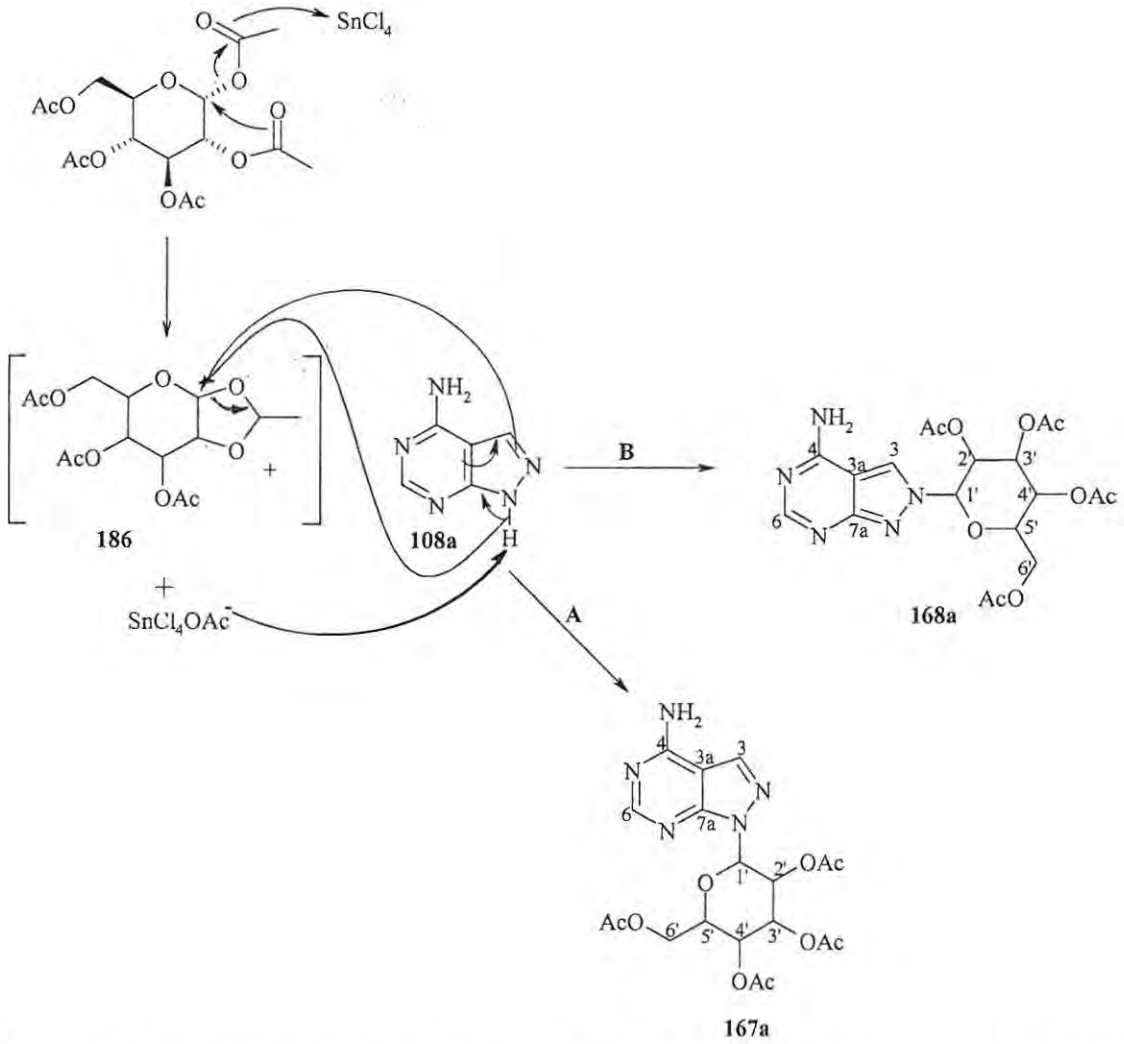
*A: experimental; B: HOSE code; C: Nnet and D: ChemWindow

2.2.2 NMR Chemical Shift Prediction of 4-Aminopyrazolo[3,4-*d*]pyrimidine Glycosyl Derivatives

In the glycosylation reaction, the isolation of two fractions with similar ^{1}H NMR signals prompted the application of the NMRPredict programme. Based on the proposed mechanism (**Scheme 46**), an intramolecular reaction would result in the elimination of the C-1 acetyl and formation of the oxonium ion intermediate **186** and an anionic tin species, which then removes the proton from the pyrazole nitrogen. The resulting anionic heterocyclic system can then react *via* two possible pathways, **A** and **B**. In path **A** (black arrows), the resulting nucleophilic nitrogen (N-1) can then attack C-1 on either face of oxonium ion **186** resulting in the α- or β-anomers of compound **167a**. In path **B** (red arrows) the negative charge on N-1 can be delocalized in the pyrazole towards N-2, which then attacks intermediate **186** resulting in the α- or β-anomers of compound **168a**. NMRpredict version 3.2.2 was employed to differentiate the isomeric products. The ^{1}H NMRpredict spectra for the α- and β-anomers of compound **167a** revealed similar signals except for the downfield shift of the 5'-methine proton from 4.62ppm to 4.99ppm in the β-anomer and the upfield shift of the 1'-methine proton from 7.34ppm to 7.03ppm in the β-anomer. The predicted ^{13}C NMR spectra revealed the same chemical shift for all the carbons in both anomers. The ^{1}H NMRPredict spectra for compound **168a** revealed overlapping of signals for the 5'-methine and 6'-methylene protons in both anomers and for the 2'- and 3'-methine protons in the β-anomer, while the predicted ^{13}C NMR revealed the same chemical shift for all the carbons in both anomers. Compared to the predicted spectra for both isomers, which revealed the aromatic protons at *ca.* 8 and 9ppm, the ^{1}H NMR spectrum of the isolated compounds revealed an upfield shift of the aromatic protons to *ca.* 8ppm. The 1'-methine proton was also observed to resonate upfield (*ca.* 6ppm) in the recorded spectrum compared to the predicted NMR shift (*ca.* 7ppm). ^{1}H NMRpredict could thus not differentiate effectively between the two isomers.

Comparison of the predicted and the observed ^{13}C NMR spectra also revealed similar carbon shifts. However, the chemical shift for C-3 and C-7a, differ by *ca* 10 and 4ppm, respectively from isomer **167a** to isomer **168a**. Using the HOSE or Nnet prediction data,

it may be concluded that the structures of the isomeric glycosylates **167a** and **168b** have been correctly assigned. The ^{13}C NMR data are summarized in **Tables 12** and **13**. These patterns carbon C-3 and 7a were also observed for the 4-hydroxy analogues **167b** and **168b**, and their ^{13}C NMR data are summarized in **Tables 14** and **15**. Comparison of the experimental and HOSE and Nnet data thus permits confirmation of the assigned structures.

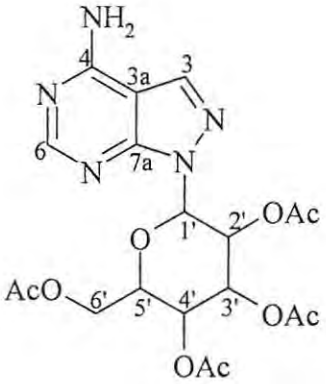


Scheme 46: Proposed mechanistic pathways for the formation of compounds **167a** and **168a**

Generally, the Modgraph prediction facility gave results that appeared to be more consistent than the ChemWindow programme, with the HOSE approach giving the smallest apparent errors.

Table 12: Experimental and predicted ^{13}C NMR data for derivative **167a**

^{13}C	A*	B*	C*	D*
2'CH_3	20.3	20.4	21.2	17.6
3'CH_3	20.6	20.7	21.2	17.6
4'CH_3	20.6	20.7	21.2	17.6
6'CH_3	20.7	20.6	20.7	17.3
6'	61.8	61.9	62.6	68.5
4'	67.9	67.9	68.5	69.3
2'	69.4	69.3	68.2	70.3
3'	73.8	72.8	70.8	68.6
5'	74.4	74.3	73.1	67.5
1'	81.4	85.8	81.0	86.6
3a	101.0	100.4	99.2	92.7
				154.7
6	156.3	152.7	158.7	156.8
				171.0
4	157.6	157.4	161.1	170.1
2'CO	169.0	168.4	171.2	171.0
4'CO	169.4	169.7	171.2	171.0
3'CO	170.2	169.9	171.2	171.0
6'CO	170.7	171.6	171.7	171.0
RMS		**1.78**	**1.76**	**7.57**

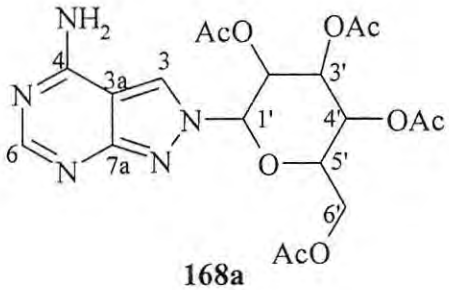


167a

*A: experimental; B: HOSE; C: Nnet and D: ChemWindow

Table 13: Experimental and predicted ^{13}C NMR data for derivative **168a**

^{13}C	A*	B*	C*	D*
2'CH_3	20.3	20.4	21.2	17.6
3'CH_3	20.5	20.7	21.2	17.6
4'CH_3	20.5	20.7	21.2	17.6
6'CH_3	20.6	20.6	20.7	17.3
6'	61.8	61.9	62.6	68.5
4'	67.9	67.9	69.1	69.3
2'	70.8	69.3	68.0	70.8
3'	72.7	72.8	66.4	68.7
5'	75.0	74.3	73.1	67.6
1'	88.0	89.0	84.1	84.2
3a	102.8	101.1	92.7	105.1
[illegible]	[illegible]	[illegible]	[illegible]	130.3
6	157.3	156.4	156.7	162.8
[illegible]	[illegible]	[illegible]	[illegible]	138.4
4	159.8	158.3	164.2	163.7
2'CO	168.5	168.5	171.1	171.0
4'CO	169.5	169.7	171.1	171.0
3'CO	170.0	169.9	171.1	171.0
6'CO	170.7	171.6	171.7	171.0
RMS		**0.77**	**4.63**	**6.85**

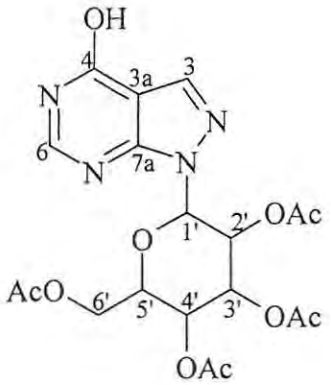


168a

*A: experimental; B: HOSE; C: Nnet, D: ChemWindow

Table 14: Experimental and predicted ^{13}C NMR data for derivative **167b**

^{13}C	A*	B*	C*	D*
2'CH_3	20.3	20.4	21.2	17.6
3'CH_3	20.6	20.7	21.2	17.6
4'CH_3	20.6	20.7	21.2	17.6
6'CH_3	20.7	20.6	20.7	17.3
6'	62.0	61.9	62.6	68.5
4'	67.7	67.9	68.4	69.3
2'	69.2	69.3	68.1	70.3
3'	73.6	72.8	70.8	68.6
5'	74.6	74.3	73.1	67.5
1'	82.6	85.8	80.9	86.6
3a	107.0	100.0	106.0	103.9
[illegible]	[illegible]	[illegible]	[illegible]	154.7
6	146.8	153.7	160.7	149.5
[illegible]	[illegible]	[illegible]	[illegible]	167.8
4	158.8	159.0	163.2	174.7
2'CO	168.5	168.4	171.2	171.0
4'CO	169.3	169.7	171.2	171.0
3'CO	170.3	169.9	171.2	171.0
6'CO	170.6	171.6	171.8	171.0
RMS		**2.71**	**3.65**	**7.18**

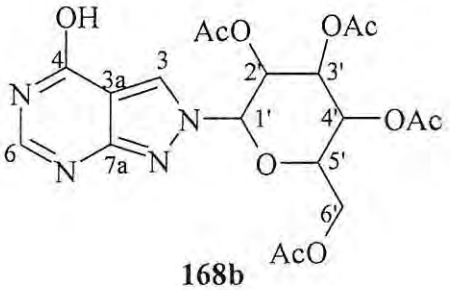


167b

*A: experimental; B: HOSE code; C: Nnet and D: ChemWindow

Table 15: Experimental and predicted 13C NMR data for compound **168b.**

^{13}C	A*	B*	C*	D*
2'CH_3	20.3	20.4	21.2	17.6
3'CH_3	20.5	20.7	21.2	17.6
4'CH_3	20.7	20.7	21.2	17.6
6'CH_3	20.7	20.6	20.7	17.3
6'	61.5	61.9	62.6	68.5
4'	67.7	67.9	68.4	69.3
2'	70.9	69.3	67.9	70.8
3'	72.6	72.8	70.8	68.7
5'	75.2	74.3	73.1	67.6
1'	88.1	89.1	84.1	84.2
3a	108.6	102.0	94.3	105.1
				130.3
6	146.1	155.2	156.9	162.8
				138.4
4	159.2	158.3	141.9	163.7
2'CO	168.9	168.5	171.2	171.0
4'CO	169.3	169.7	171.2	171.0
3'CO	170.0	169.9	171.2	171.0
6'CO	170.5	171.6	171.8	171.0
RMS		**2.88**	**6.72**	**6.94**



168b

*A: experimental, B: HOSE, C: NNet and D; chemdraw

2.3 Molecular Modelling of the Various ATP Analogues

The development of resistance to many antiviral and antibacterial agents has led to the development of structure-based approaches, using the X-ray crystal structure of enzymes and molecular modelling techniques, to evaluate and design potential inhibitors of disease-specific enzymes. This rational approach to drug design has been successful in the generation and development of various antiviral and antibacterial drugs including ritonavir, an HIV-1 protease inhibitor.[188] This strategy has also been employed in our research to evaluate the potential binding of the various ATP analogues to the glutamine synthetase (GS) active site. Two approaches were followed. In the first (Section 2.1.2.5), alignment of the global minimum conformations of ATP and various synthetic analogues was examined, while in the second (Section 2.3.1), attention was given to the docking of ATP and the synthetic analogues into the GS active site. The latter approach is concerned with the structural homology of binding- rather than global minimum conformations.

2.3.1 Computer Modelling of the Docking of ATP and Synthetic Analogues in the GS Receptor Site

The ACCELRYS Cerius^2 LIGANDFIT module was used to determine the preferred binding conformations of our synthetic compounds and the nature of their interactions with the GS receptor site. The X-ray crystal structure of the glutamine synthetase (GS) enzyme as determined by Eisenberg[90] was used for the docking studies. As indicated in the introduction (Section 1.5) the GS enzyme is a dodecamer with twelve ATP active sites, but the X-ray crystal structure of GS from the Brookhaven Protein Databank database used in our study, was determined using GS containing ADP in the active sites. More specifically, the structure used in this study is a recontructed structure of model 1F52 comprising of four adenylated sub-units of the *MTb* GS structure, each containing two Mg^{2+} ions. A 5Å binding pocket around ADP in one of the sub-units was selected and defined as the 'active site' (**Figure 49a and 49b**). Attention was then focused on amino acid residues in the active site (**Figure 50**) and, in particular those residues which might interact directly with ATP and other ligands. What is important is the strength and nature of interactions between the relative binding and the plasma solvation energies.

Figure 49a. Full structure of the enzyme glutamine synthetase downloaded from the PDB database: http://www.rgb.pdb.org

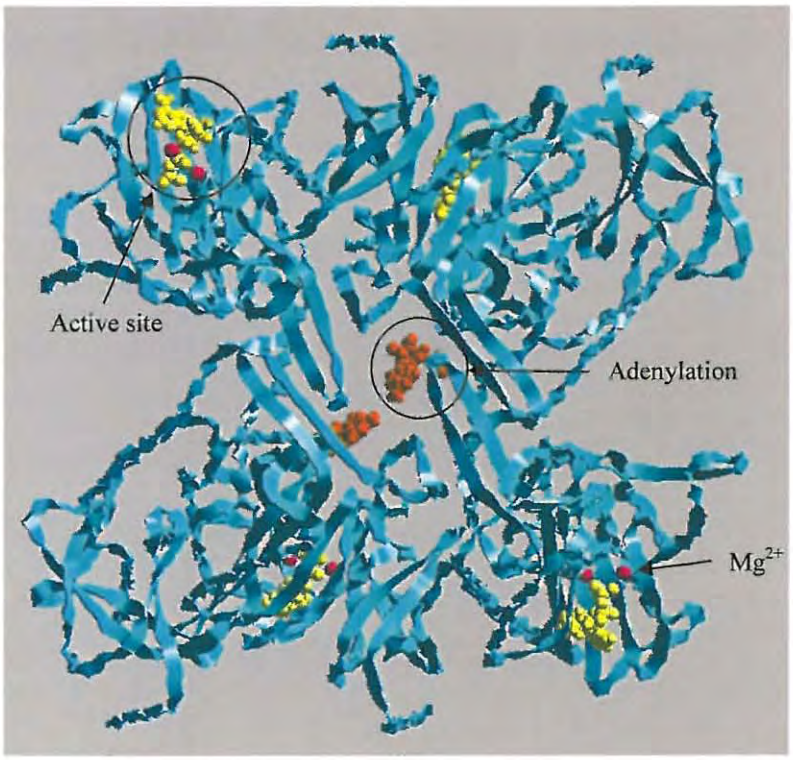


Figure 49b. Reconstructed structure of the adenylated four subunit GS (1F52) high-lighting the active and adenylation sites.[189]

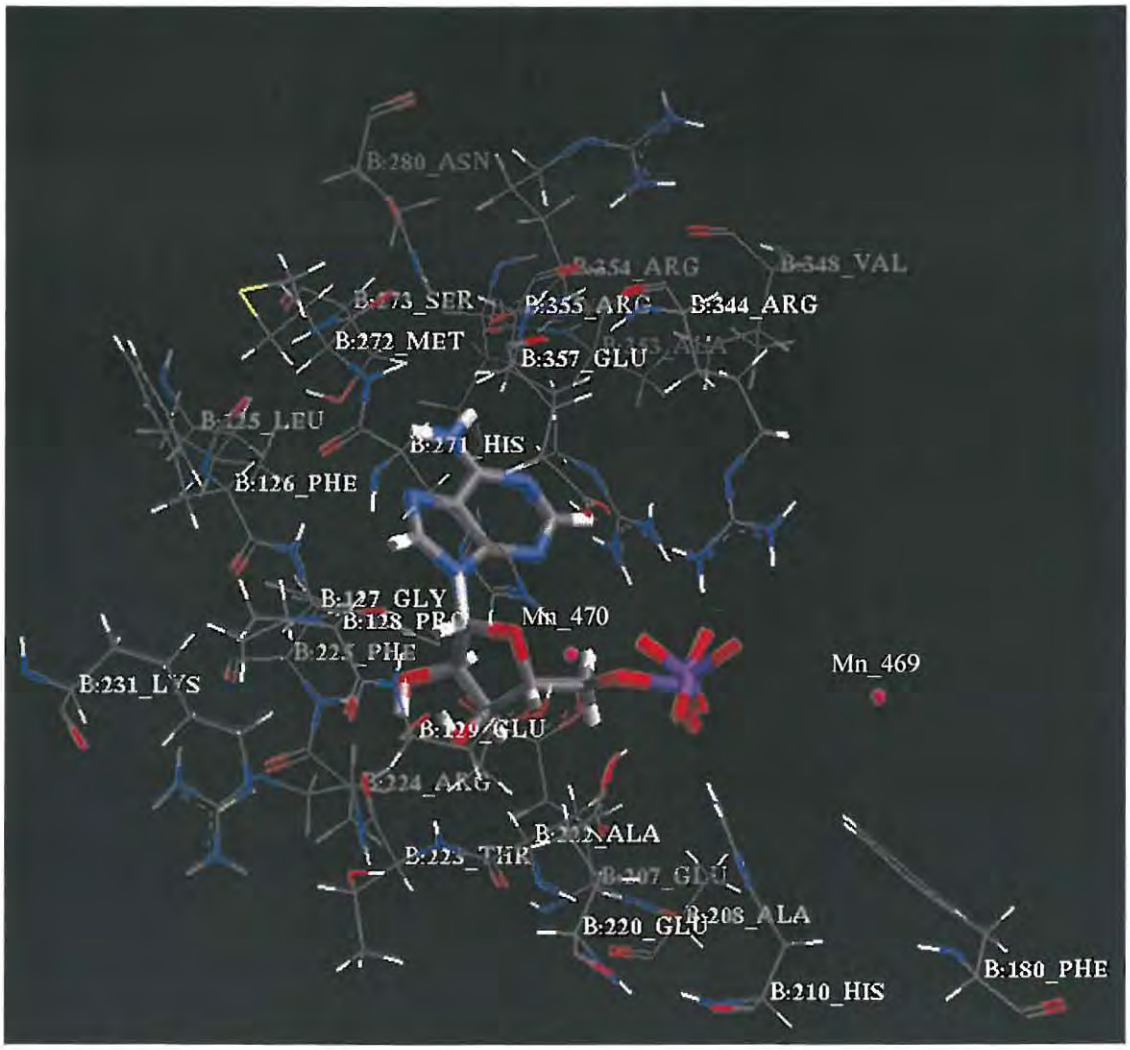


Figure 50: Amino acid residues in the defined active site of GS

In order to identify the important amino acid residues in the active site, it was assumed that amino acid residues within 3Å will have significant interactions with ADP and, by implication, with ATP and any potential inhibitors. These amino acid residues (illustrated in **Figure 51**) were found to be:-

i) Arg 344, which interacts with the phosphate side chain *via* donor hydrogen-bonding;

ii) Arg 355, which interacts with both the phosphate side chain and the adenine ring *via* donor hydrogen-bonding as well as *via* acceptor hydrogen-bonding interactions;

iii) Glu 220, His 210 and 271, which interact with the phosphate side chain;

iv) Glu 129, Phe 225, Thre 223 and Glu 207, which interact with the riboside moiety through hydrogen-bonding;

v) Gly127, Arg 224 and Ala 222, which interact with the riboside moiety *via* acceptor hydrogen-bonding interactions; and

vi) Ser 273 and Arg 354, which interact with the amino group of the adenine ring- the former through donor hydrogen-bonding, the latter through acceptor hydrogen-bonding interactions.

Furthermore the metal ion Mn 470 also interacts with the amino acid residues Glu 129, His 210 and Glu 220.

These findings agree with the observations made by Liaw *et al.*,[190] whose studies of GS from *Salmonella Tryptium* indicated that the important amino acids in the active site are Phe 225, Ser 273, Glu 207, Arg 355 and His 271, although the specific interactions were somewhat different.

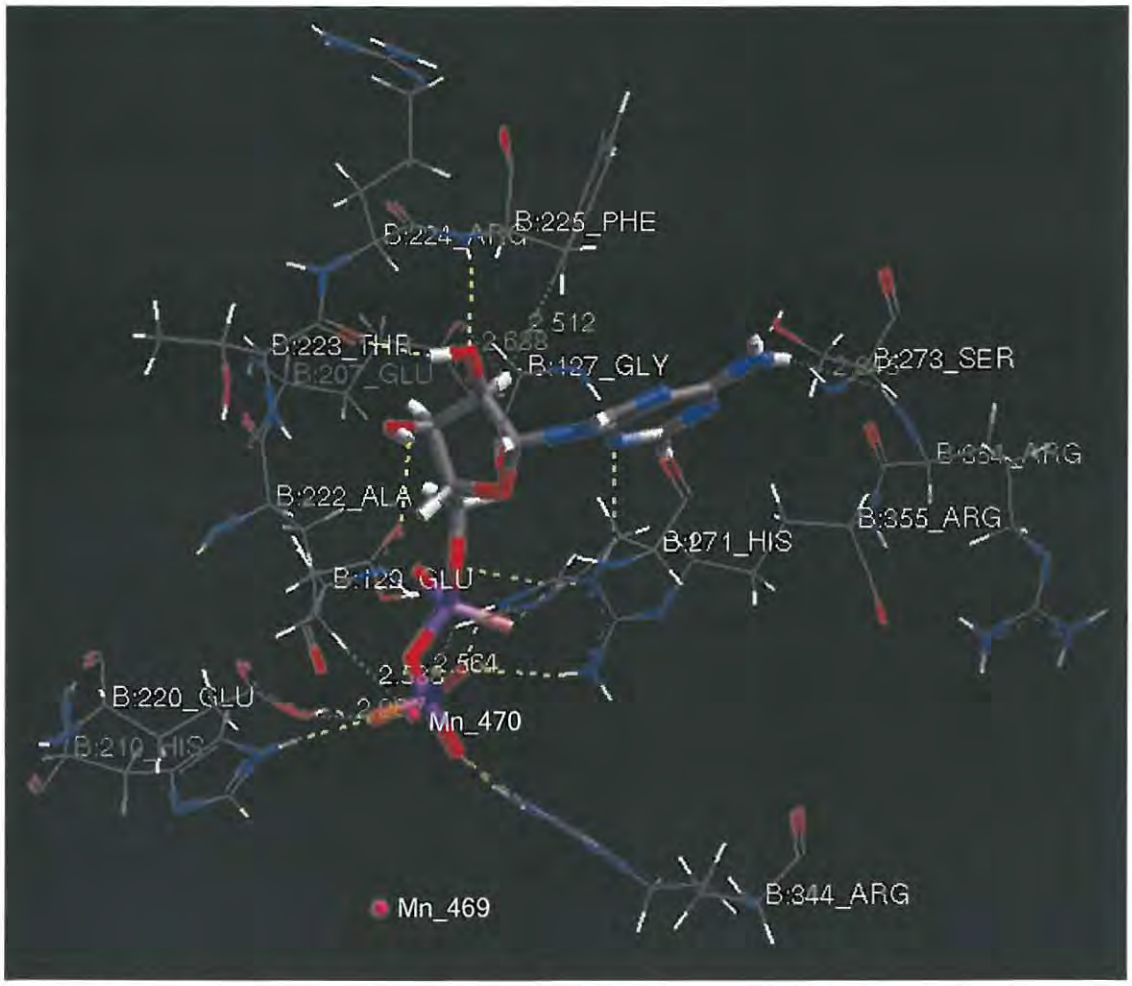


Figure 51: Interactions between ADP and active site amino acids within a distance of 3Å; the yellow lines depict amino acid residues with potential for donor hydrogen-bonding with ADP, while the green lines depict potential acceptor hydrogen-bonding interaction.

2.3.1.1 Modelling ATP in the GS active site

In order to determine the interactions between ATP and the enzyme active site, the shape of the enzyme "binding site" was determined using SITE SEARCH parameters in the Ligand Fit Module with ADP as the the docked ligand (**Figure 52**). It is important to note that the docking studies were conducted *in vacuo*. The energy minimized structure of ATP was then docked into the enzyme active site and the most favourable bound conformer was determined by scoring the different possible conformations to obtain the corresponding "Dock score", "Ligscore" and van der Waals ("vdW") energies (**Table**

16). The "Dock score" is the negative value of the non-bonded intermolecular energy between the ligand and protein. The "Ligscore" is the measure of protein-ligand interaction, and vdW is the Leornard-Jones soft potential. ATP conformer 1 (shown in green in **Figure 52**) was chosen as the best bound conformation as it exhibited the best docking score and has vdW and ligscore values most comparable to ADP [total ligand-protein binding affinities (ligscore): 1.12; and van der Waals energy: -31.47 kcal/mol].

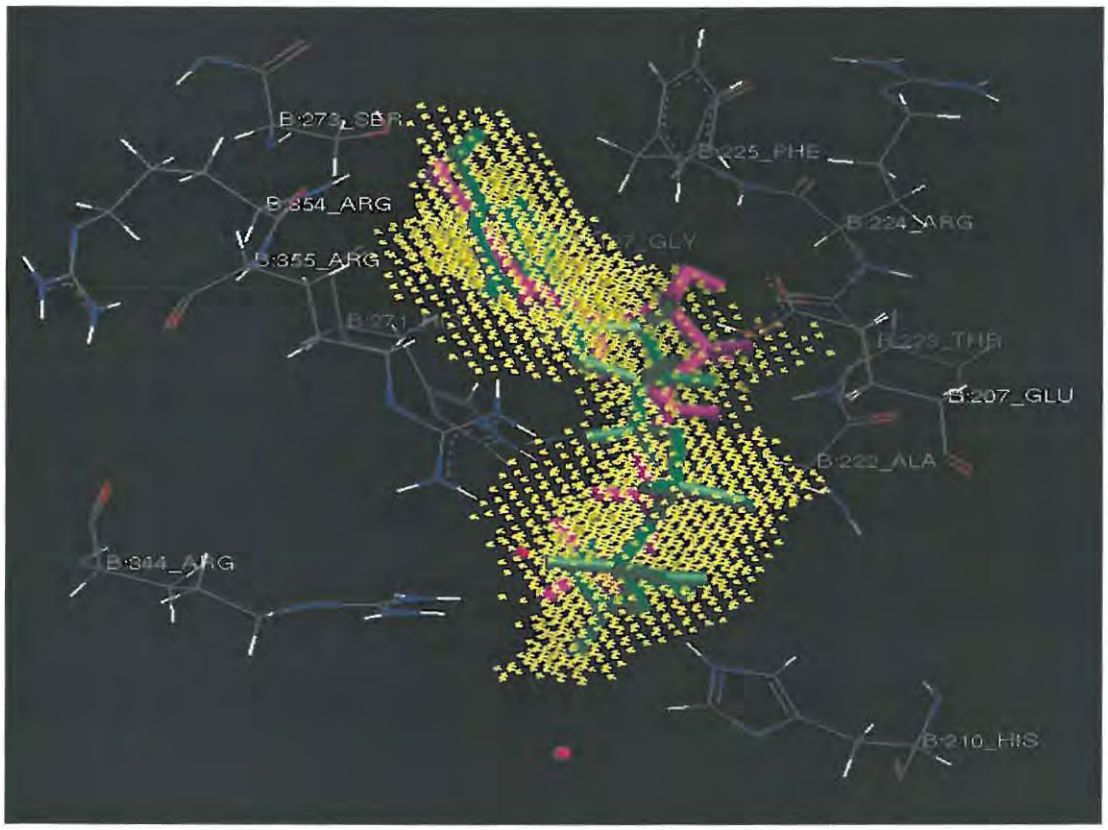


Figure 52: The best bound conformations of ATP (green) and ADP (magenta) in the GS "active site" (yellow).

Table 16: Data showing the scores of the twenty best binding conformations of ATP in the GS enzyme active site in descending order.

Conformer	Dock score (kcal/mol)	Ligscore (kcal/mol)	vdW energy (kcal/mol)
1	34.52	1.33	-35.95
2	29.06	0.43	-16.77
3	28.93	-0.15	-4.25
4	28.32	1.37	-36.76
5	27.69	0.87	-26.16
6	26.68	0.30	-13.88
7	24.39	0.33	-14.53
8	24.06	0.81	-24.85
9	22.46	-2.37	43.14
10	19.16	-1.30	20.33
11	14.42	-1.77	30.46
12	13.99	-1.05	14.95
13	12.80	-1.32	20.71
14	11.95	-0.38	0.75
15	11.54	-2.74	51.13
16	11.17	-0.26	-1.91
17	11.12	-1.23	18.89
18	7.68	-3.73	72.26
19	6.89	-3.41	65.52
20	6.69	-0.35	0.11

The important amino acid residues which interact with ATP were found to be similar to those determined for ADP, except that ATP does not appear to interact with Arg 224, Ser 273 and Thre 223; instead it interacts with Lys 352 through hydrogen-bonding with N-3 of the adenine ring (**Figure 53**). It was also observed that the metal ion Mn 469 interacts with the hydroxyl oxygen attached to the γ phosphorous atom. These observations are consistent with previous studies,[81,86] which suggest that, while ADP binds at the ATP

site, the small differences in the set of interacting amino acid residues accounts for its capacity to inhibit the activity of the enzyme.

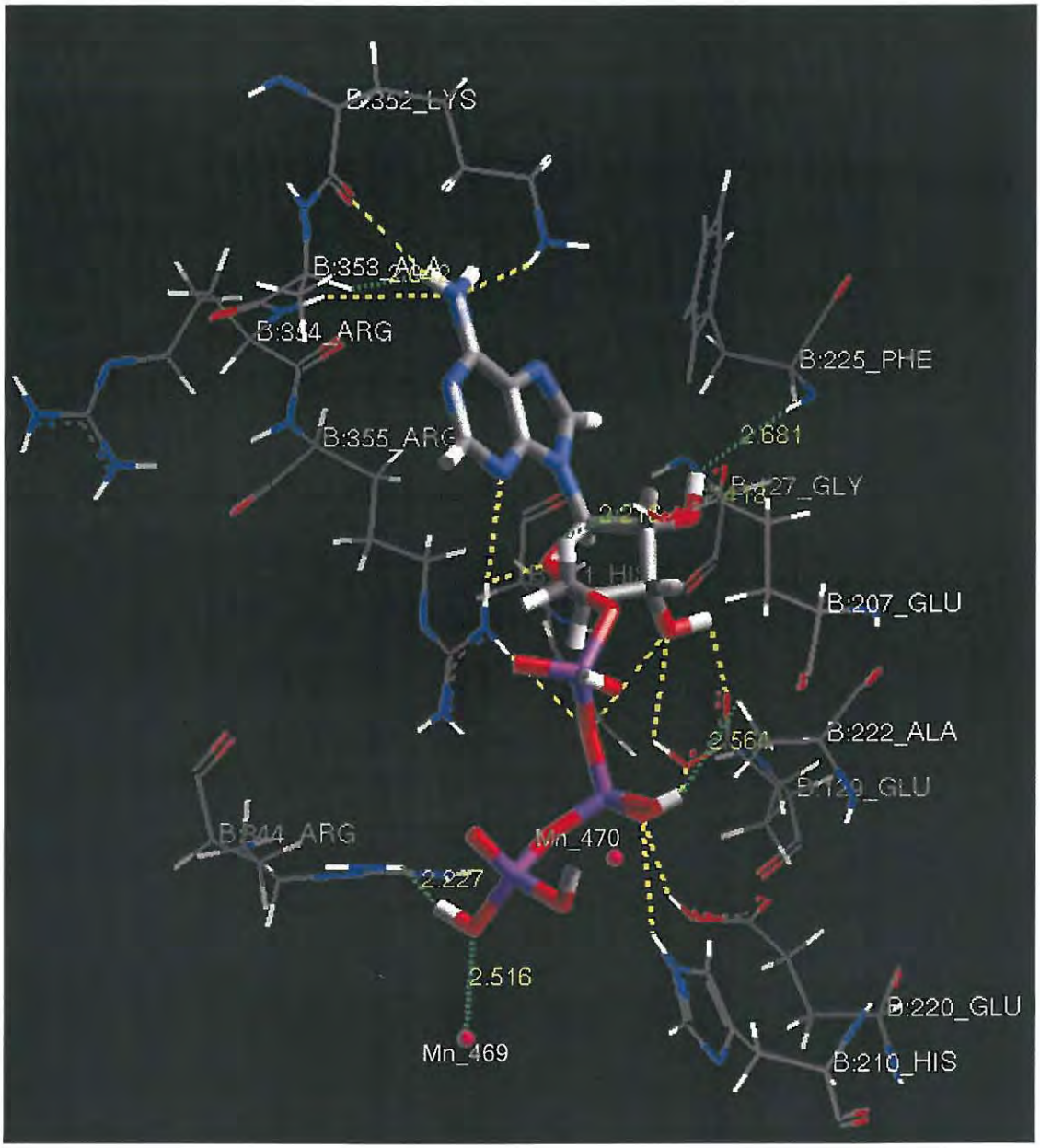


Figure 53: Interactions between ATP and the GS active site amino acid residues within a distance of 3Å from the ligand. The yellow lines depict amino acid residues with donor hydrogen-bonding potential, while the green lines depict acceptor hydrogen-bonding interactions.

2.3.1.2 Modelling 3-Indolylalkanoic Acid Derivatives in the GS Active Site

The energy minimized structures of the indolylalkanoic acid derivatives (Section 2.1.2.5) were each docked into the enzyme active site and the data for the best-bound conformation in each case, are summarised in **Table 17**. Based on all three parameters (Dockscore, Ligscore and vdW energy), it is expected that the acetylated glucosyl derivatives **116c** and **138b** will not exhibit significant interaction with the enzyme active site as the binding energy values for all these parameters differ greatly from those found for ATP. It is possible, of course, that these derivatives could undergo *in vivo* hydrolysis to afford the corresponding polyhydroxy derivatives with far better binding potential. The energy values for compounds **134a**, **135c** and **139c**, on the other hand, compare reasonably well with those found for ATP, and these compounds are expected to exhibit significant binding interactions with the receptor site. It was also observed that the best bound conformations for each of these derivatives also assumes the same general arrangement with the polyoxy moieties orientated towards the Mn metal ions.

Table 17: Data showing the scores for the best binding conformation of ATP and each of the indolylalkanoic acid derivatives in the enzyme active site.

Compound	Dockscore (kcal/mol)	Ligscore (kcal/mol)	vdW energy (kcal/mol)
ATP	34.52	1.33	-35.95
116c	-631.96	-25.08	592.26
134a	37.90	1.60	-41.66
135c	29.89	1.17	-32.62
138b	-530.08	-18.77	394.22
139c	36.10	1.03	-29.49

The amino acid residues that interact with these compounds were also investigated. **Figures 54-58** illustrate the amino acid residues that are within a binding distance of 3Å for compounds **116c**, **134a**, **135c**, **138b** and **139c**, respectively. It was observed that these derivatives exhibit potential hydrogen-bonding interactions with some of the important amino acid residues noted for both ATP and ADP. Thus, compounds **116c**, **134a**, **135c**, **138b** and **139c** all interact with Arg 355 through the oxygen atoms of the polyoxy moiety and, in some cases, (**138b** and **139c**), with the nitrogen of the indole ring. Compounds **116c**, **134a**, **135c** and **138b** also exhibit interactions with Glu 129 through the oxygen atoms of the polyoxy moiety. However, only the glucosylylated derivatives, **116c**, **138b** and **139c,** appear to interact with Glu 207, while only the acetylated derivatives **116c** and **138b** exhibit interactions with His 210 and Lys 352. Further hydrogen-bonding interactions are observed with His 271 for compounds **116c, 135c** and **138b**, with Arg 354 for compound **135c** and with Gly 127 for compound **139c**.

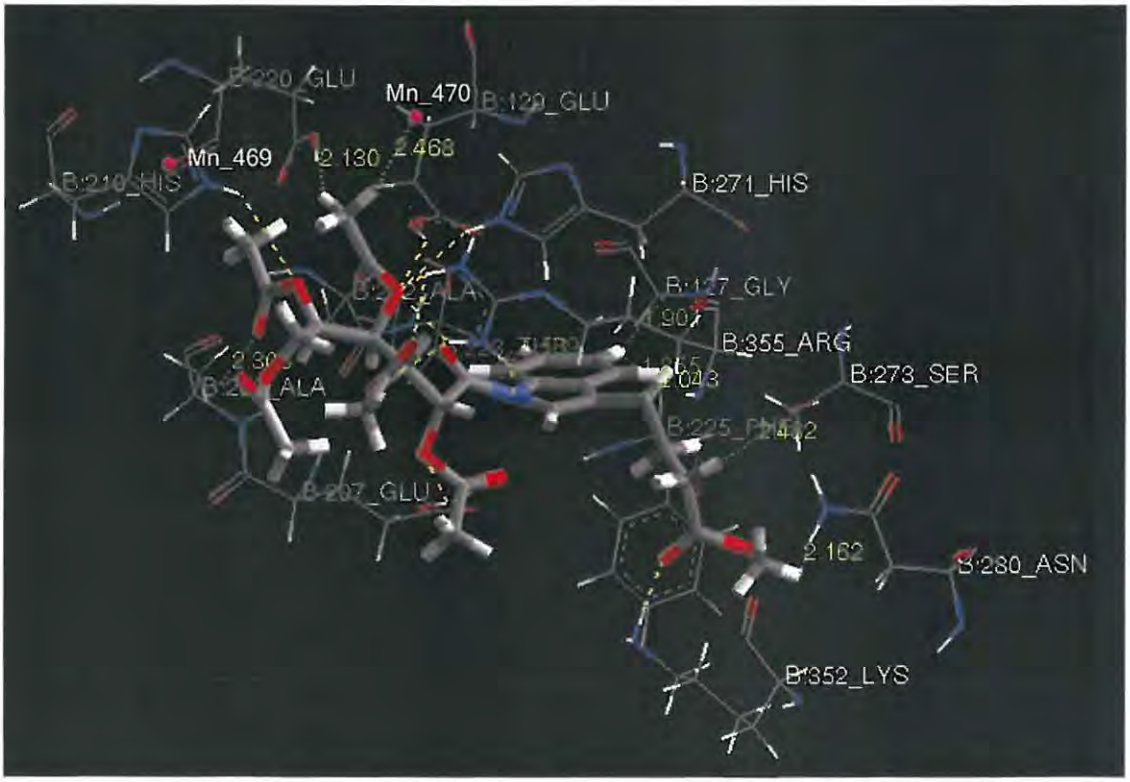


Figure 54: Favoured bound conformer of compound **116c,** showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential hydrogen-bonding interactions.

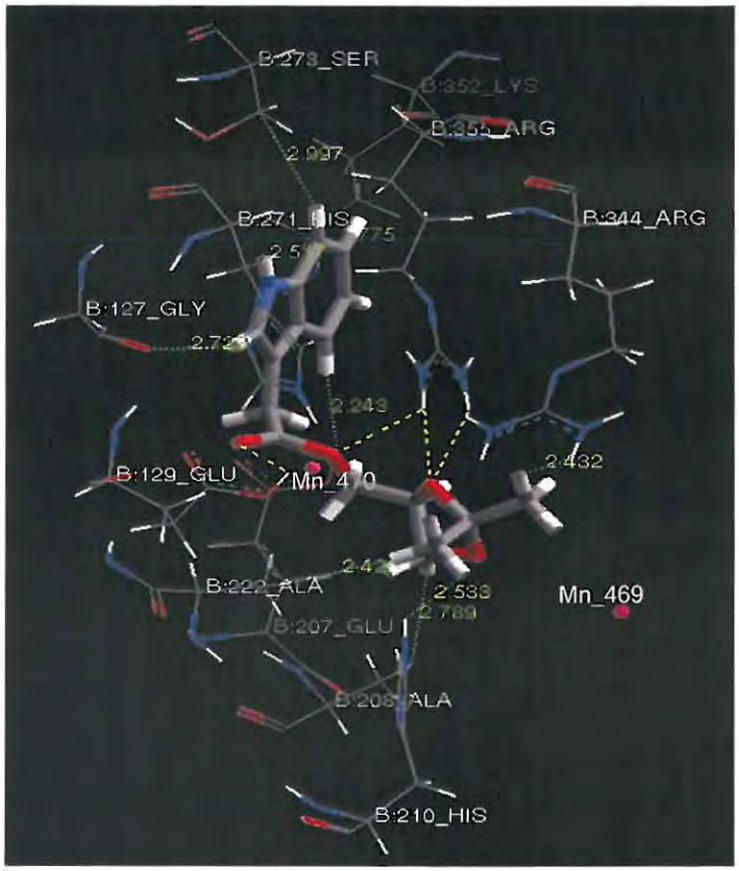


Figure 55: Favoured bound conformer of compound **134a**, showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential hydrogen-bonding interactions.

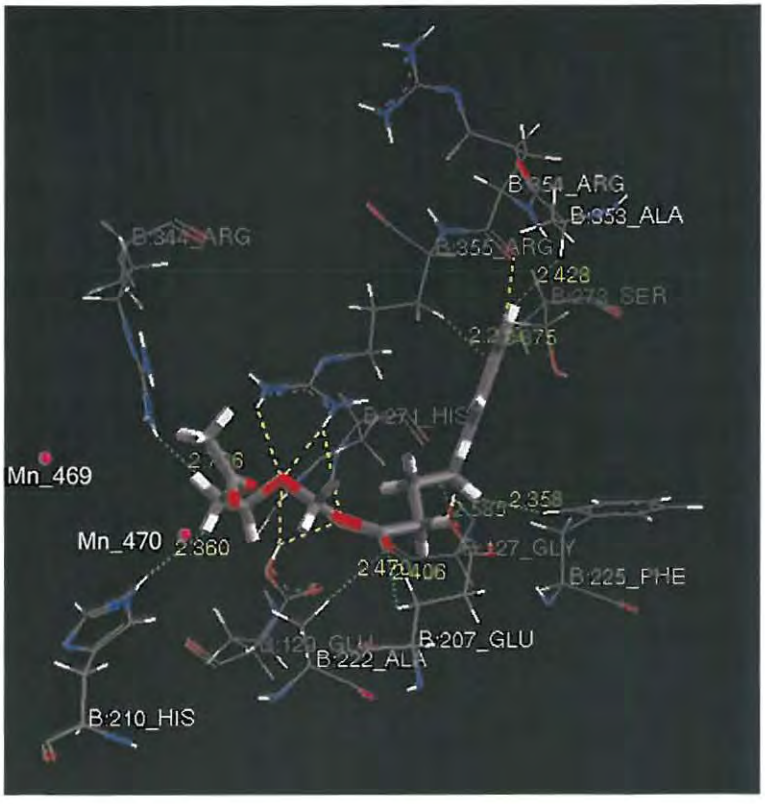


Figure 56: Favoured bound conformer of compound **135c**, showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential hydrogen-bonding interactions.

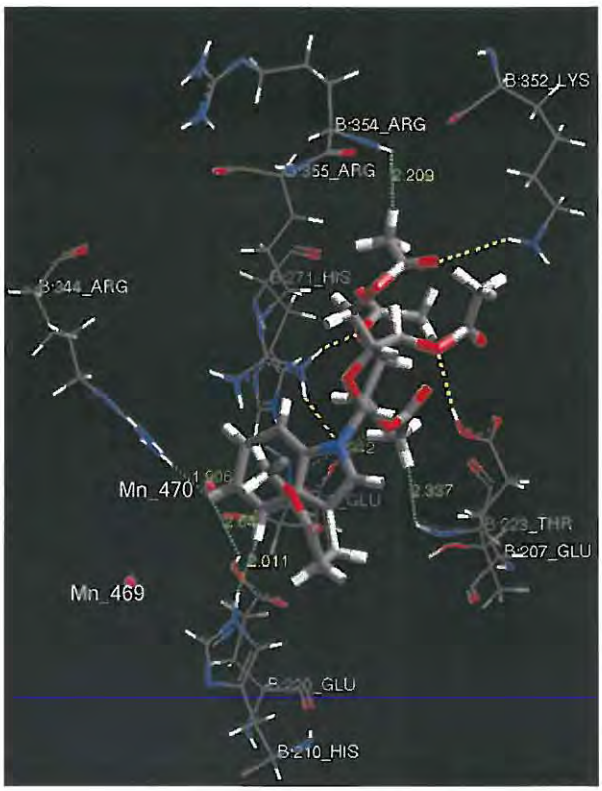


Figure 57: Favoured bound conformer of compound **138b**, showing the amino acid residues, which are within a 3Å distance from the ligand, the yellow lines depict potential hydrogen-bonding interactions.

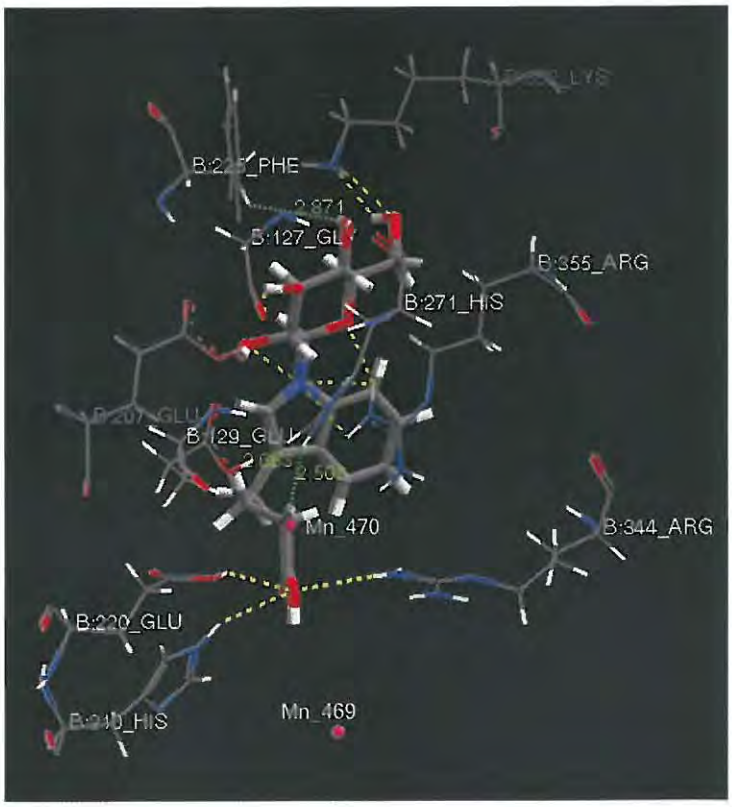


Figure 58: Favoured bound conformer of compound **139c**, showing the amino acid residues, which are within a 3Å distance from the ligand, the yellow lines depict potential hydrogen-bonding interactions.

2.3.1.3 Modelling Benzimidazole Derivatives in the GS Active Site

The energy minimized structures of the benzimidazole derivatives (**Section 2.1.3.5**) were each docked into the enzyme active site and the data for the best bound conformations are summarised in **Table 17**. Based on these data, it seems that the glucosyl compound **156** and **157** and the phosphates **158b** and **160a** are unlikely to bind effectively with the enzyme active site as their binding energy values differ greatly from those found for ATP. However, the hydrolysed glucosyl derivative **157** and compounds **150** and **160b** have binding energy values that are more comparable with those found for ATP and are thus expected to exhibit significant binding interactions with the receptor site. In this series, it was observed that, in general, the best bound conformation of each derivative assumes the same general arrangement, with the benzimidazole ring oriented towards the Mn metal ions. The exception is compound **160b**, whose polyoxy moiety is oriented towards the metal ions, as was the case for the 3-indolylalkanoic acid derivatives (**Section 2.3.1.2**).

Table 17: Data showing the scores for the best binding conformation of ATP and each of the benzimidazole derivatives in the enzyme active site.

Compound	Dockscore (kcal/mol)	Ligscore (kcal/mol)	vdW energy (kcal/mol)
ATP	34.52	1.33	-35.95
150	32.44	1.02	-29.24
156	-167.94	-7.38	150.39
157	-14.08	-1.22	18.68
158b	-272.46	-11.57	249.08
160a	0.13	0.16	-10.61
160b	-88.10	1.21	-33.9

The amino acid residues that interact with these compounds were also investigated and these are illustrated in **Figures 59-64**. It was noted that these derivatives could exhibit potential-hydrogen bonding interactions with some of the important amino acid residues

identified for both ATP and ADP. Thus, compounds **150**, **157**, **156**, **160a** and **160b** all interact with Arg 355 through the oxygen atom of the polyoxy moiety and the N-1 nitrogen of the imidazole ring. Compound **158b** also appears to interact with Arg 355 but not *via* the imidazole ring. In compound **150**, Arg 355 also interacts with N-3 of the imidazole ring. Furthermore, all of these compounds (**150**, **157**, **156**, **160a**, **160b** and **158b)** interact with Glu 207 through the oxygen moiety of the polyoxy side chain. Compounds **156** and **157** also exhibit interactions with Glu 129 through the sugar moiety as well as the benzimidazole N-3 in compound **157**. Compound **160b** also exhibit interactions with Lys 352 and only compound **160b** interaction with Arg 344.

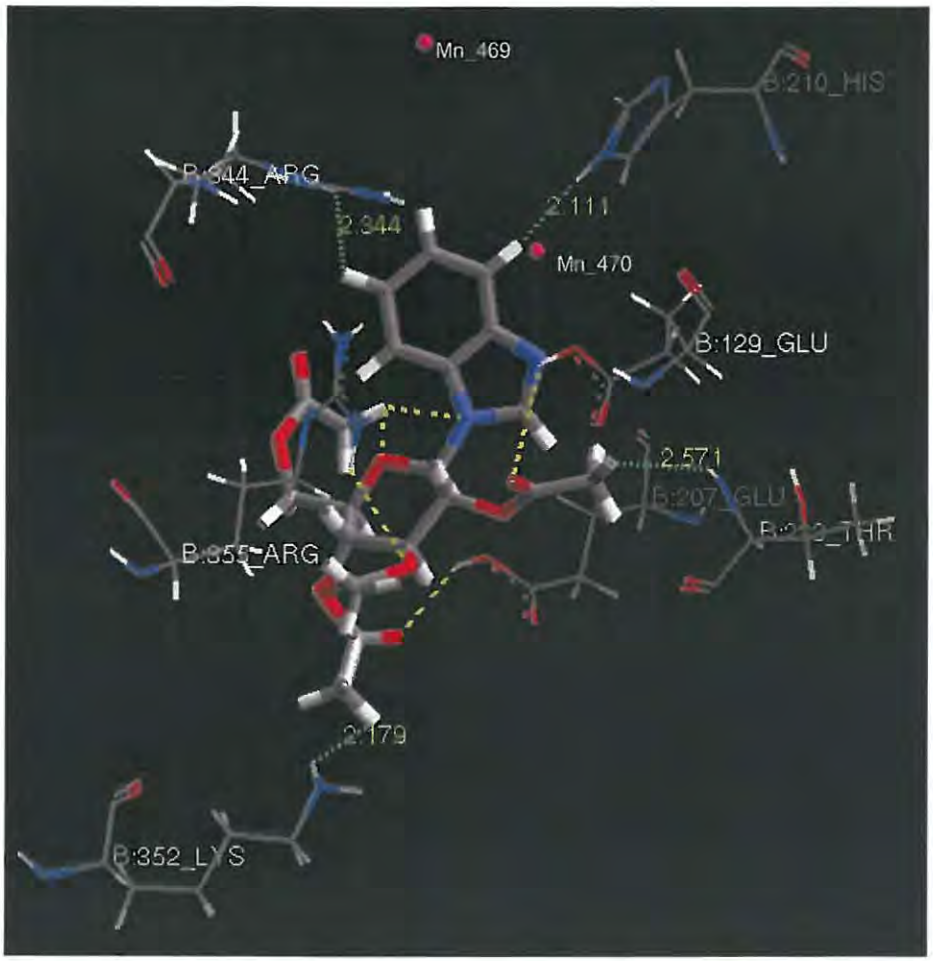


Figure 59: Favoured bound conformer of compound **156**, showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential hydrogen bonding interactions.

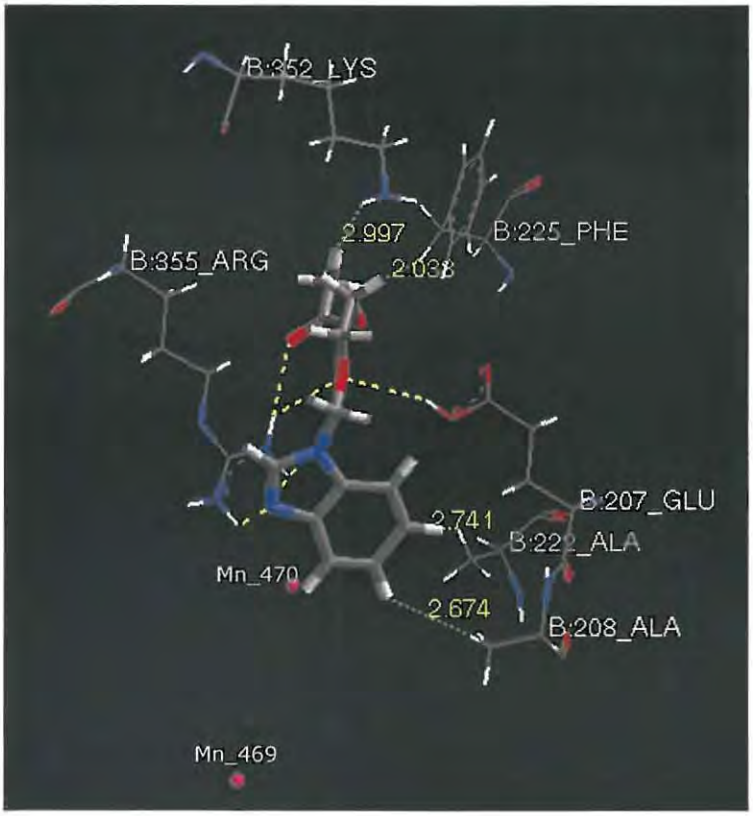


Figure 60: Favoured bound conformer of compound **150**, showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential donor hydrogen-bonding, while the green lines depict potential acceptor hydrogen-bonding interactions.

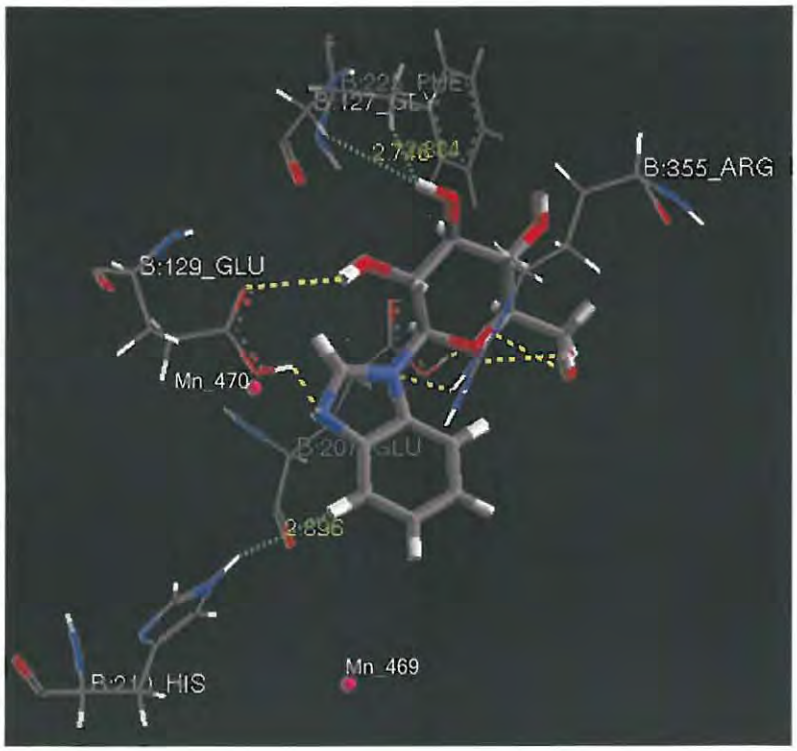


Figure 61: Favoured bound conformer of compound **157** showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential donor hydrogen-bonding, while the green lines depict potential acceptor hydrogen-bonding interactions.

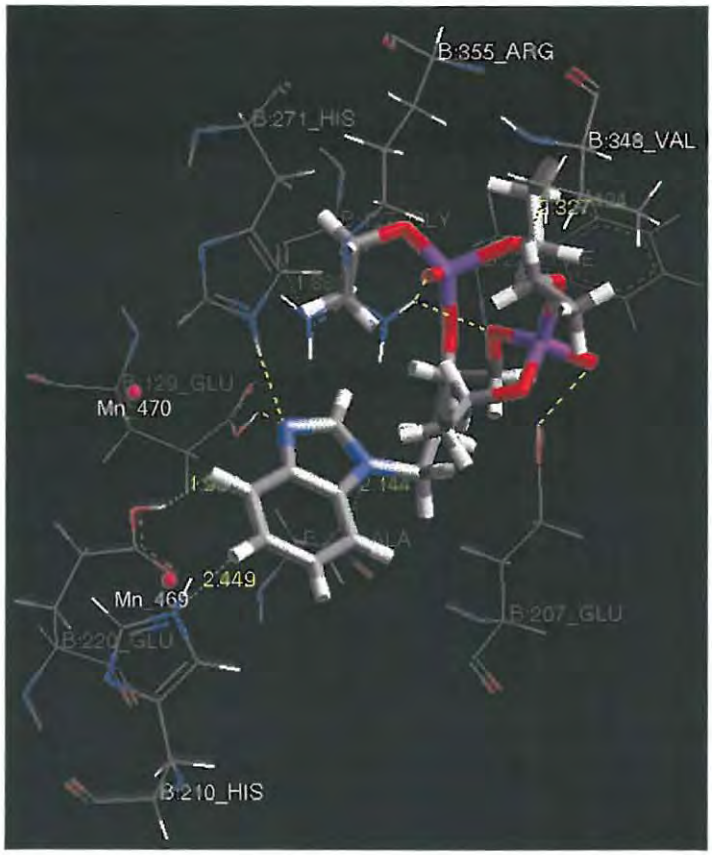
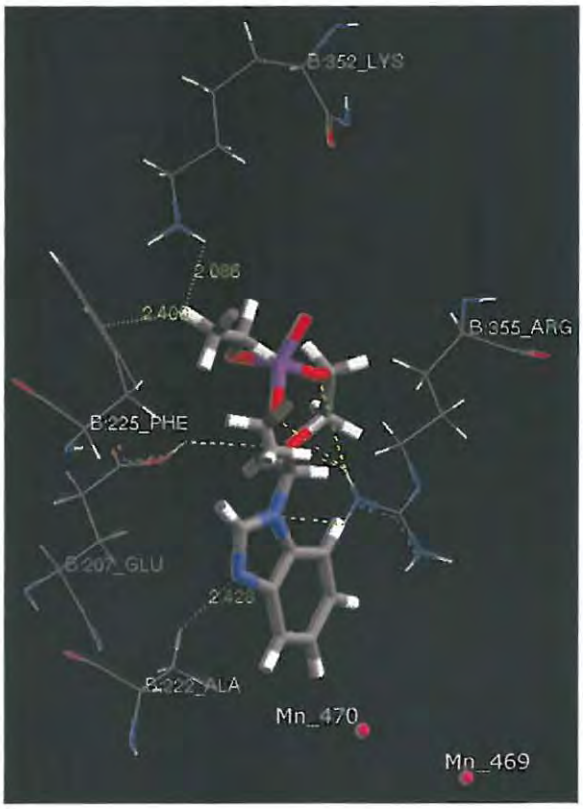


Figure 62: Favoured bound conformer of compound **158b** showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential donor hydrogen-bonding interactions.

Figure 63: Favoured bound conformer of compound **160a** showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential donor hydrogen-bonding interactions.



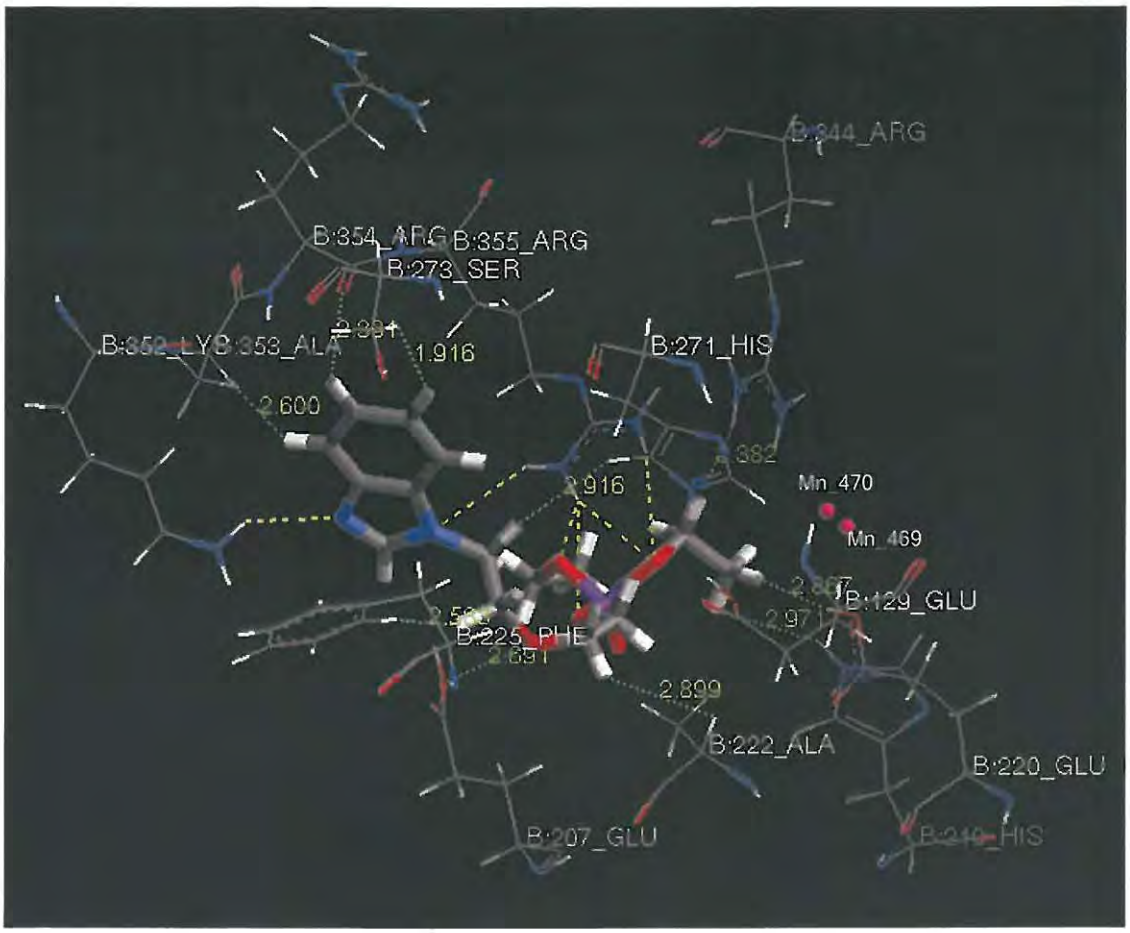


Figure 64: Favoured bound conformer of compound **160b** showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential donor hydrogen-bonding, while the green lines depict potential acceptor hydrogen-bonding interactions.

2.3.1.4 Modelling Pyrazolo[3,4-d]pyrimidine Derivatives in the GS Active Site

The energy minimized structures of the pyrazolo[3,4-d]pyrimidine derivatives (**Section 2.1.4.4**) were each docked in the enzyme active site and the data for the best bound conformations are summarised in **Table 18**. Based on a comparison with the data from ATP, compounds **167a, b** and **169b** are all expected to exhibit poor interactions with the amino acid residues in the enzyme active site. However, compounds **169a**, **175a** and **179** are expected to exhibit significant interactions with the residues in the enzyme active site because their data is comparable to that found for ATP. It was observed that the cyclic glucosyl derivatives assume a similar general arrangement with the heterocyclic base orientated towards the Mn metal ions, while the acyclic derivatives assume an arrangement in which the polyoxy moiety is orientated towards the Mn metal ions.

Table 18: Data showing the scores for the best binding conformation of ATP and each of the pyrazolo[3,4-d]pyrimidine derivatives in the enzyme active site.

Compound	Dockscore (kcal/mol)	Ligscore (kcal/mol)	vdW energy (kcal/mol)
ATP	34.52	1.33	-35.95
175a	32.52	1.09	-30.74
167a	-96.99	-6.32	127.73
167b	-166.62	-9.30	191.44
169b	-143.22	-8.07	165.16
169a	11.46	-0.18	-3.62
179	-15.12	0.77	-23.93

The amino acid residues that interact with these derivatives are illustrated in **Figures 65-70**. It seems that these derivatives could also exhibit hydrogen-bonding interactions with some of the important amino acid residues in the active site. Thus, compounds **175a**, **167b** and **169b** all appear to interact with Arg 355 through the pyrazole nitrogen as well

as the oxygens of the polyoxy moiety, whereas compounds **167a** and **169a** interact with Arg 355 through the pyrazole and pyrimidine nitrogens. Compounds **167a**, **169a** and **179** also appear to interact with Arg 355 through oxygens of the polyoxy moiety. Further interactions are observed with Glu 207 through a pyrazole nitrogen in compound **175a** and through a glucosyl oxygen in the glucosyl derivatives **167a, 167b** and **169b**. Compounds **179**, **167a,b** and **169a** also interact with Glu 129 through the oxygens of the polyoxy moiety and, in some cases (**167b** and **169a,b**) through the pyrazole nitrogen. Interactions with His 210 were observed for compound **175a**, through an oxygen of the polyoxy moiety, and for the acetylated glucosyl derivatives **167a,b** through the 4-hydroxy or 4-amino substituent on the pyrimidine ring; interestingly these interactions were not observed with the hydrolysed derivatives **169a,b**. Interactions with His 271 were also observed for the glucosyl derivatives **167a** and **169b** through a pyrazole nitrogen. The acyclic derivatives **175a** and **179** appear to interact with Lys 352 through the amino substituent and the pyrimidine nitrogen, respectively. However, compound **175a** interact with Arg 344 through the oxygen of the polyoxy moiety, while the glucosyl derivative **167a** interact through the pyrimidine nitrogen. Only the acyclic sugar derivative **179** Arg 354 exhibit interaction with through the 4-amino substituent on the pyrimidine ring.

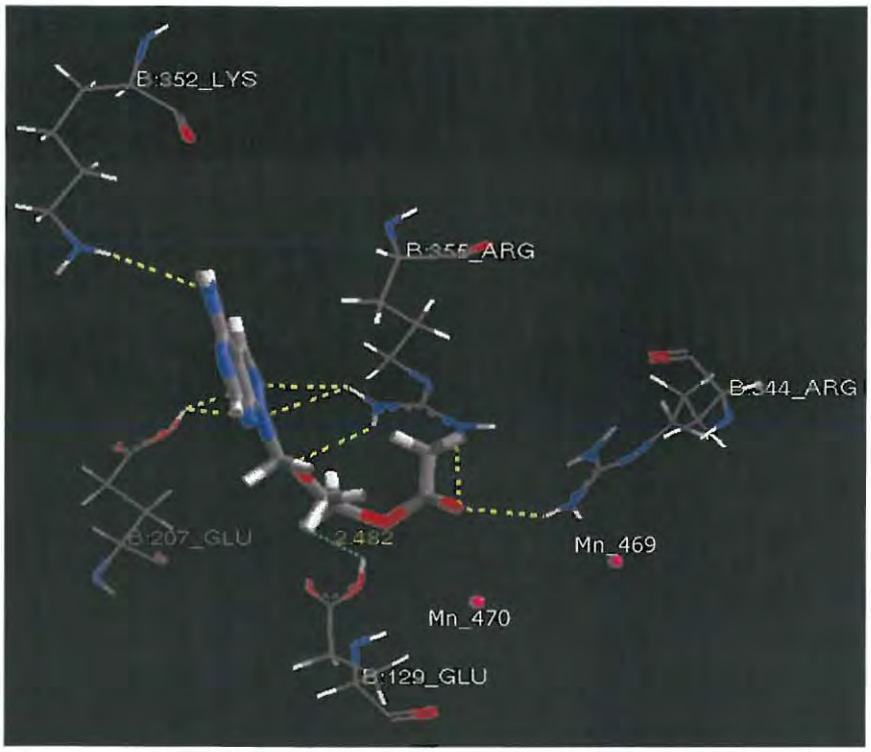


Figure 65: Favoured bound conformer of compound **175a**, showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential hydrogen-bonding interactions.

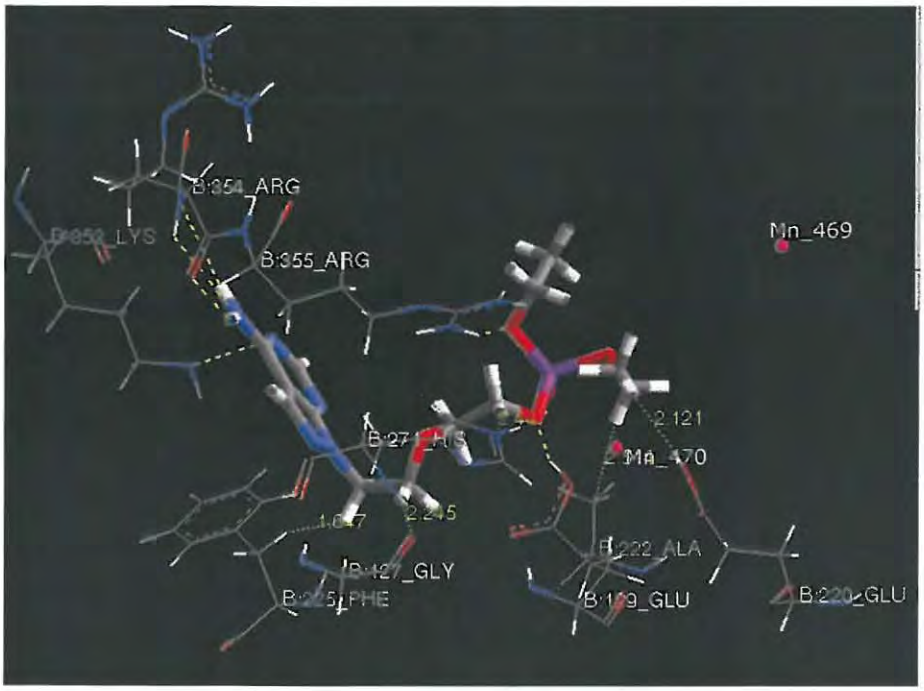


Figure 66: Favoured bound conformer of compound **179**, showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential hydrogen-bonding interactions.

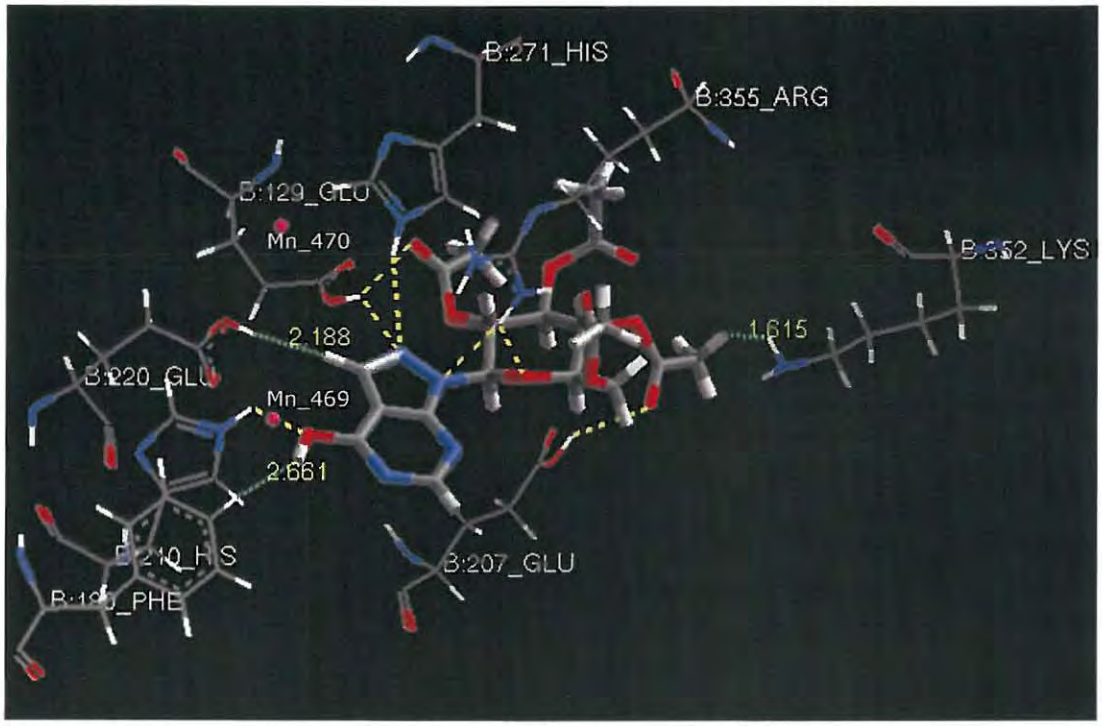


Figure 67: Favoured bound conformer of compound **167b** showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential hydrogen-bonding interactions.

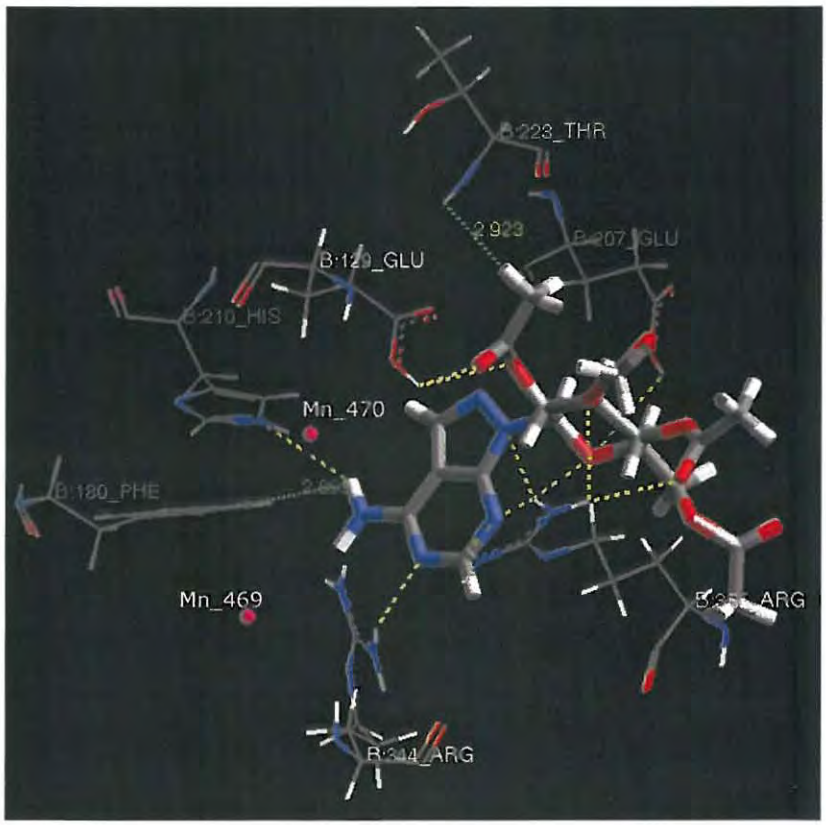


Figure 68: Favoured bound conformer of compound **167a** showing the amino acid residues, which are within a 3Å distance from the ligand, the yellow lines depict potential donor hydrogen-bonding

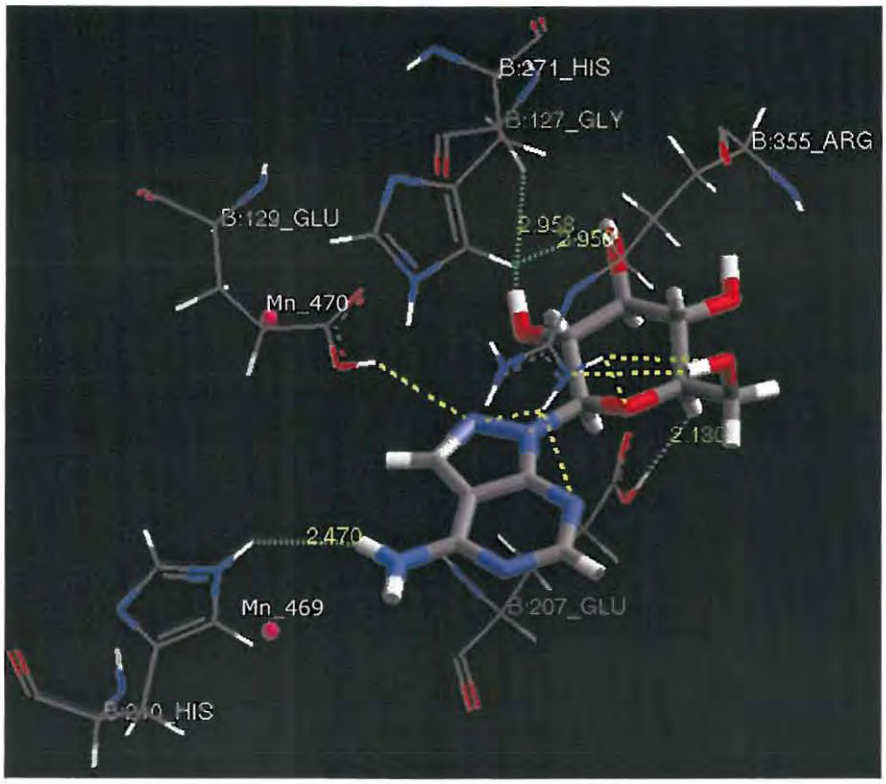


Figure 69: Favoured bound conformer of compound **169a** showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential hydrogen-bonding interactions.

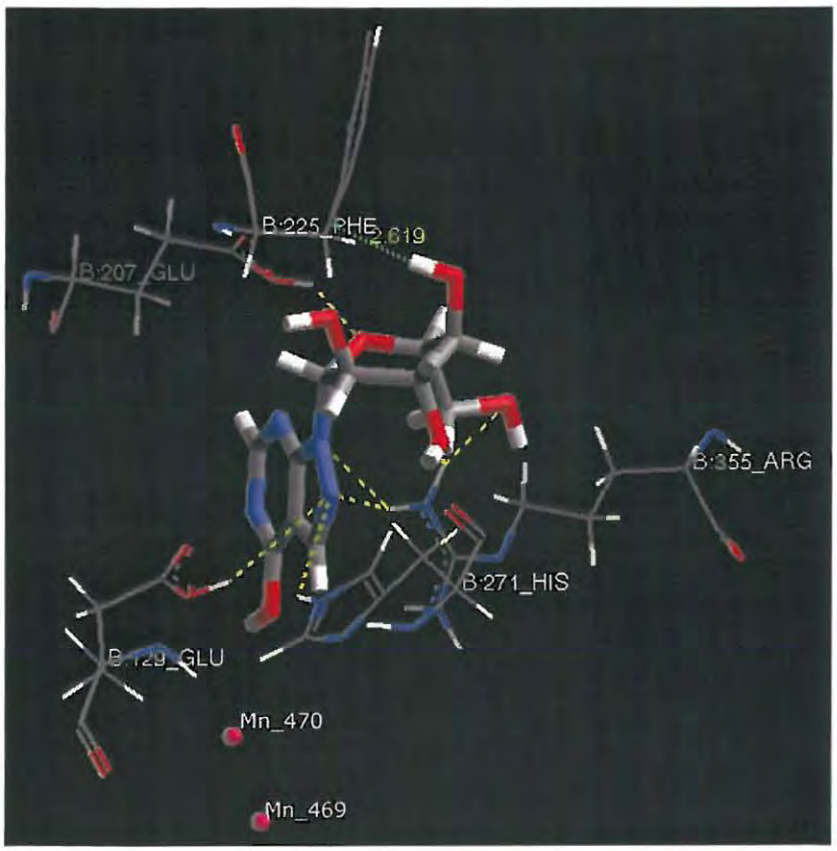


Figure 70: Favoured bound conformer of compound **169b** showing the amino acid residues, which are within a 3Å distance from the ligand; the yellow lines depict potential hydrogen-bonding interaction.

2.4 Mass Spectrometric Studies of Selected ATP Analogues

Various ATP analogues, prepared in this study, with different heterocyclic bases, *viz.*, indole, benzimidazole and pyrazolo[3,4-*d*]pyrimidine, were selected for mass spectrometric studies. In each case the analysis has focused on the more intense peaks in the respective electron-impact (EI) mass spectra. High-resolution data has been used to establish the elemental composition of each of the fragments; support for the proposed pathways, however, must await the results of metastable or B/E linked scan analysis.

2.4.1 Glycosylated 3-Indolylalkanoic Acid Derivatives.

A combination of high- and low-resolution EI mass spectrometric data were used to determine the possible mass fragmentation pathways for the glycosylated indolylalkanoic acid derivatives, compounds **138a-c**. The proposed fragmentation pathways are outlined in **Schemes 51-53**, from which it may be noted that the fragmentations, although somewhat different, exhibit similar patterns involving the indole moiety present in all of them. The glucosyl moiety common to compounds **138a** and **138c** also exhibits similar fragmentation.

The EI mass spectrum of compound **138a** is illustrated in **Figure 71**, while the proposed fragments corresponding to the significant peaks are detailed in **Scheme 51**. Fragmentation of the molecular ion **I** (*m/z* 519), which appears as a weak peak (0.2%), from compound **138a** (**Scheme 51**) is proposed to occur *via* two pathways, **A** and **B**, involving the glucosyl and the indolylalkanoate moieties, respectively. In path **A**, the molecular ion **I** loses a heterocyclic radical to yield the even-electron ion **A1** (*m/z* 331), which in turn fragments *via* loss of methyl acetate and a methyl radical to afford the odd-electron species **A4** (*m/z* 242). This species is observed in all three derivatives **138a-c**. Fragmentation of the cation **A1** also affords cation **A8** (*m/z* 73), while loss of an acetic acid molecule affords the cation **A2** (*m/z* 271); successive loss of acetic acid, ketene and another acetic acid molecule then leads to the even-electron fragments **A5** (*m/z* 211), **A6**

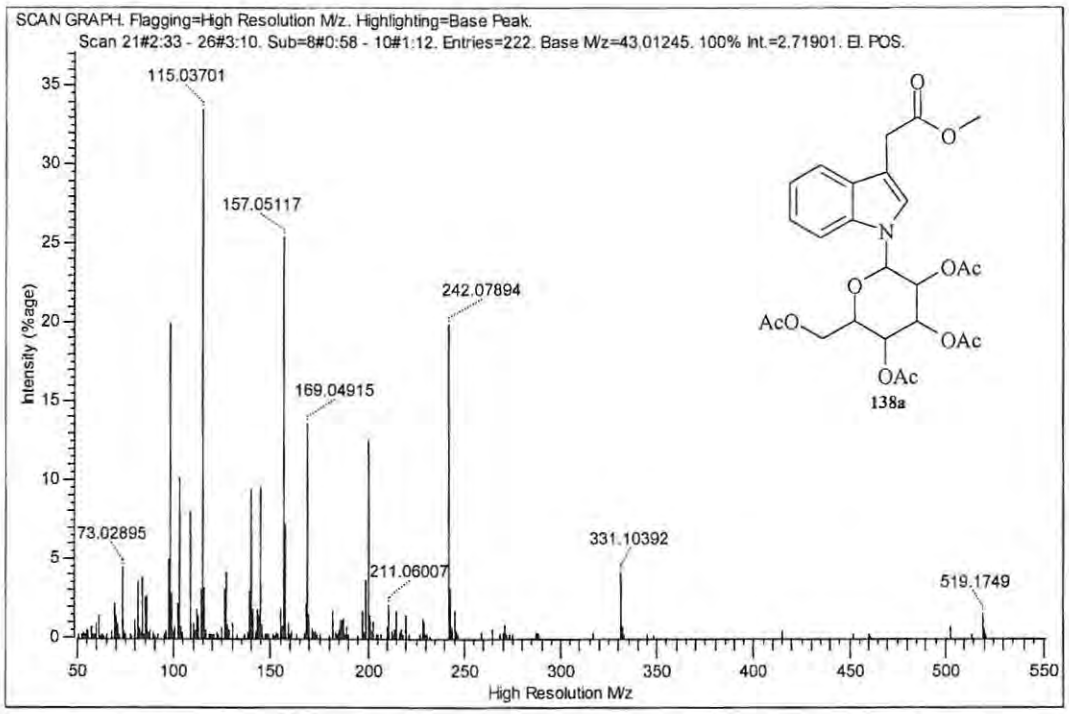
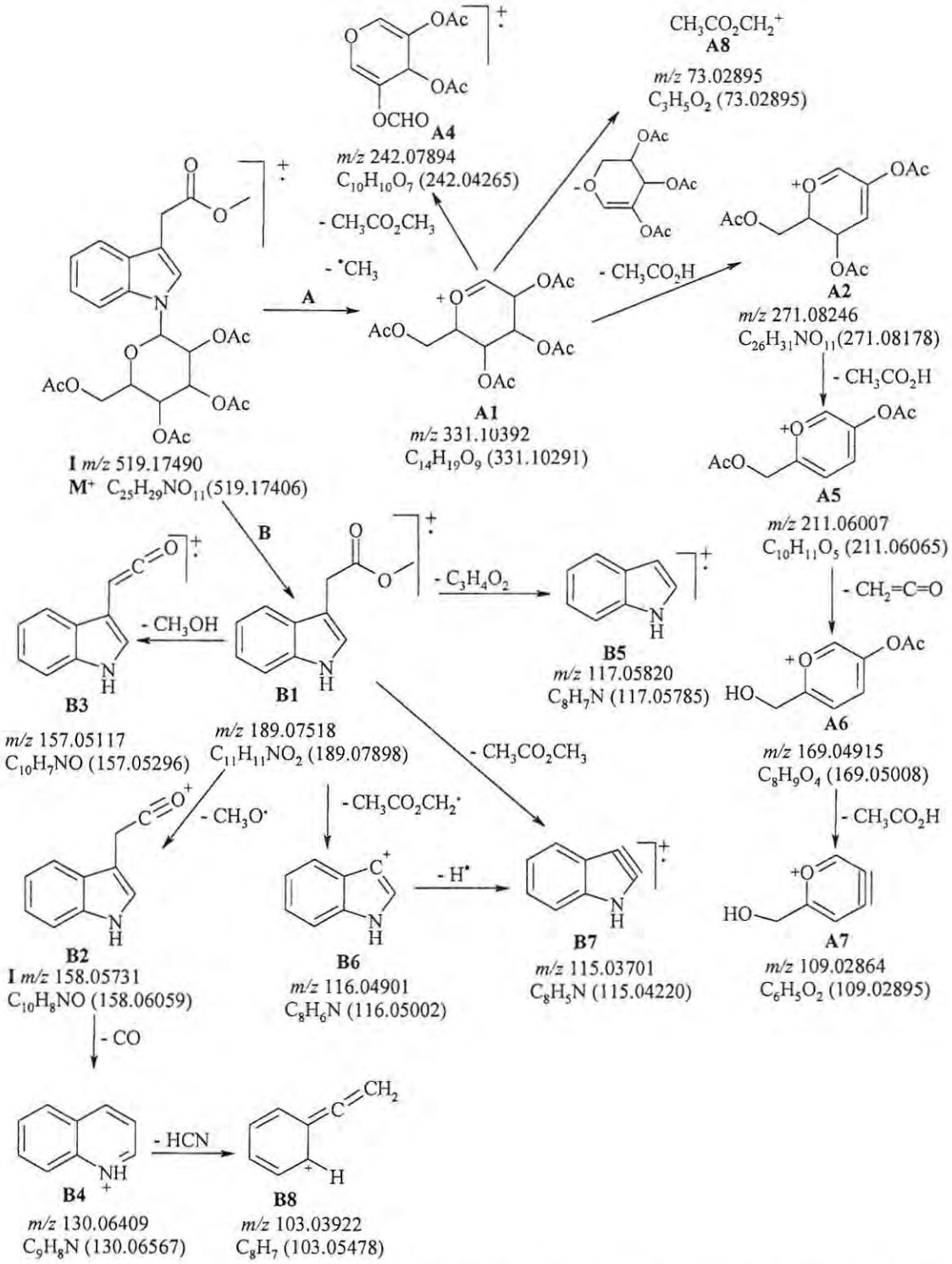


Figure 71: High-resolution EI mass spectrum of the indole derivative **138a**.

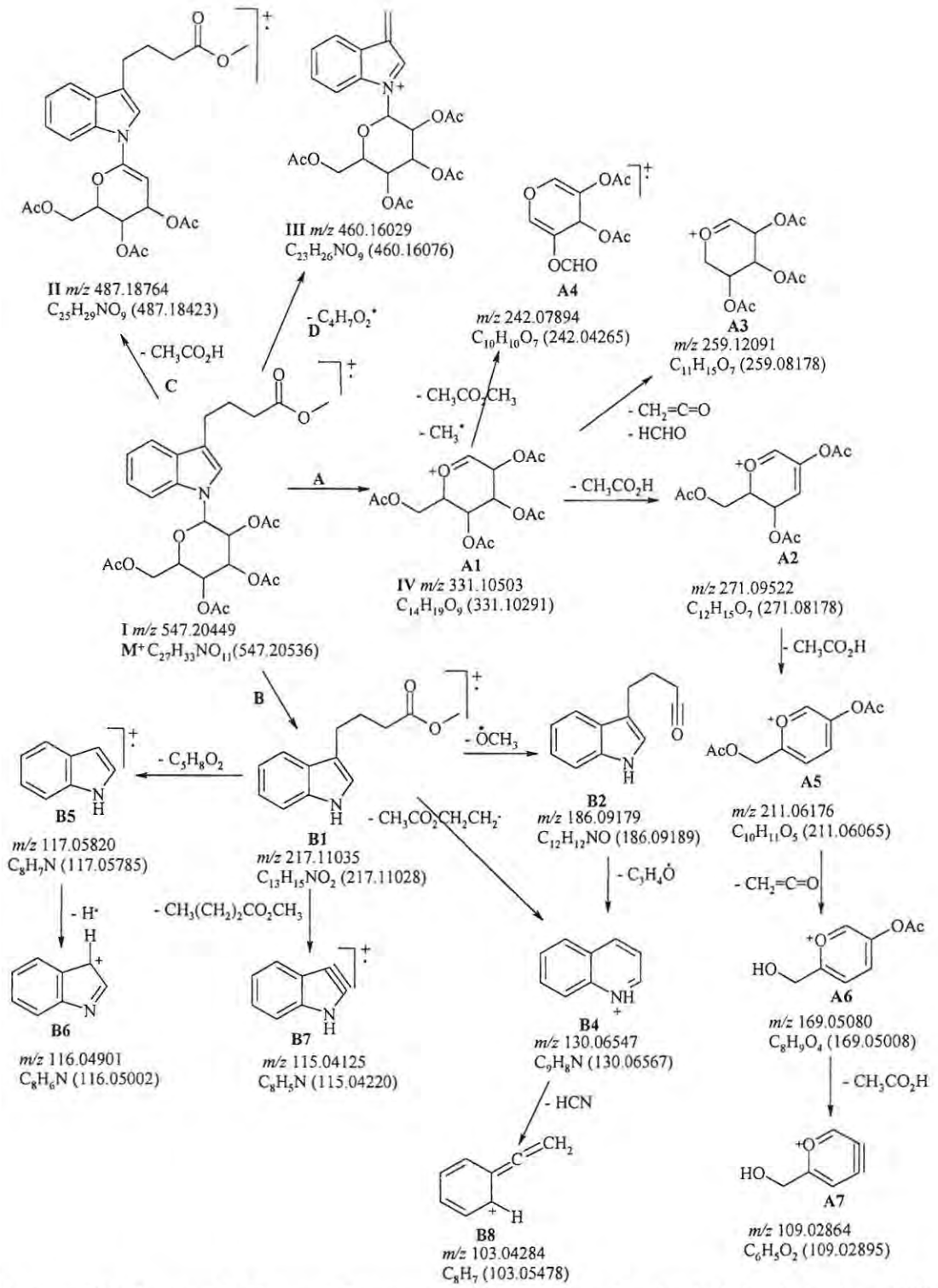
(*m/z* 169) and **A7** (*m/z* 109), respectively. This pattern is also observed in the fragmentation of compound **138c**.

Fragmentation *via* path **B** involves a loss of the glucosyl moiety to afford the methyl 3-indolylacetate radical-cation **B1** (*m/z* 189). Fission of fragment **B1** involves:- i) loss of methanol to afford the ketene radical-cation **B3** (*m/z* 157); ii) loss of a methoxy radical to afford the acylium ion **B2** (*m/z* 158), which undergoes successive loss of carbon monoxide and hydrogen cyanide to afford the even-electron species **B4** (*m/z* 130) and **B8** (*m/z* 103), respectively; iii) formation of the indole radical-cation **B5** (*m/z* 117); iv) formation of an even-electron species **B6** (*m/z* 116); and v) loss of methyl acetate to afford fragment **B7** (*m/z* 115), formulated as the heteroaryne species **B7**. The loss of a methoxy radical (*m/z* 31; **B1**⟶**B2**) and HCN (*m/z* 27; **B4**⟶**B8**) is indicative of the presence, in compound **138a**, of a methyl ester and a nitrogen, respectively, while the presence of fragments with *m/z* 130, 117, 116 and 115 are commonly observed in the decomposition of indole derivatives substituted with an alkyl group at position 3 of the pyrole ring.[191-193]



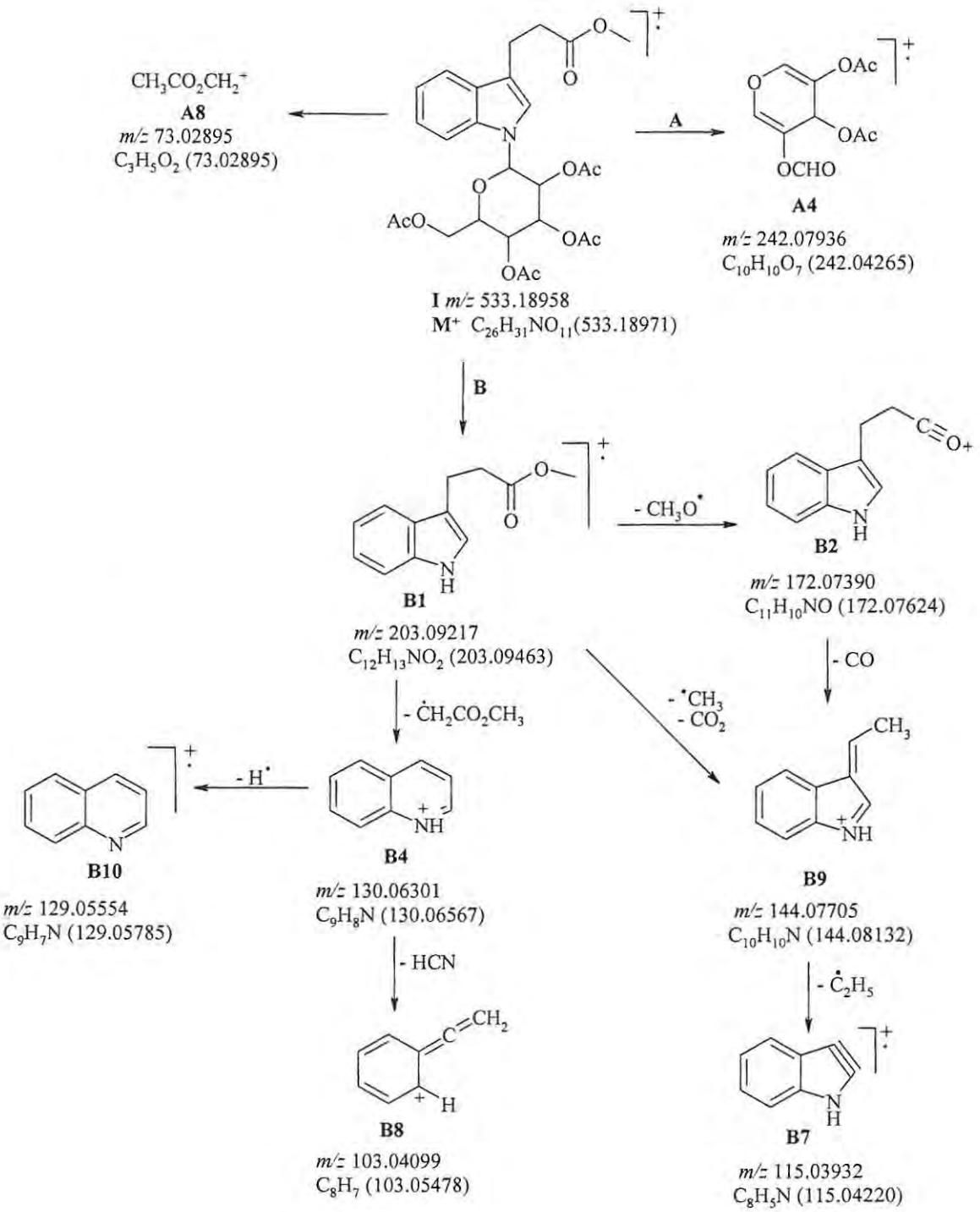
Scheme 51. Proposed EI mass fragmentation pathways for the indole derivative **138a**. High-resolution (*m/z*) data are followed, in parentheses, by calculated masses.

Fragmentation of the molecular ion **I** (*m/z* 547) from compound **138c** (**Scheme 52**), which appears as a very intense peak (96% relative abundance), is proposed to occur *via* four pathways, *viz.*, **A**, **B**, **C** and **D**. Fragmentation through paths **A** and **B** are similar to those observed for compound **138a**, leading to the glucosyl and indolylalkanoate fragments, respectively. The glucosyl cation **A1** (*m/z* 331) fragments *via*:- loss of acetic acid to give the even-electron species **A2** (*m/z* 271); loss of ketene and formaldehyde to give a second even-electron cation **A3** (*m/z* 259); and loss of methyl acetate and a methyl radical to give the odd-electron species **A4** (*m/z* 242). Successive fission of the cation **A2**, involving loss of acetic acid, ketene and acetic acid, affords the three cations **A5** (*m/z* 211), **A6** (*m/z* 169) and **A7** (*m/z* 109) also observed in the mass spectrum of compound **138a**. The cationic species **A7** accounts for the base peak with 100% relative abundance. Loss of the glucosyl moiety from the molecular ion gives rise to the methyl 3-indolylbutanoate radical-cation **B1** (*m/z* 217), which leads to a set of fragments **B4-B8**, that parallel the corresponding fragments arising from fission of the radical-cation **B1** (m/z 189) in the mass spectrum of compound **138a**. The fragment types **B5-B7** with *m/z* values of 117, 116 and 115, respectively, indicate the presence of the indole moiety,[189] while the acylium cation **B2** (*m/z* 186) differs from the corresponding fragment in the mass spectrum of compound **138a** by 28 mass units equating to two methylene groups. The molecular ion **I** also appears to lose acetic acid or a methyl propanoate radical to afford the fragments **II** (*m/z* 487) and **III** (*m/z* 460) in which both the indolyl and glucosyl moieties are retained (paths **C** and **D**).



Scheme 52. Proposed EI mass fragmentation pathways for the indole derivative **138c**. High-resolution (*m/z*) data are followed, in parentheses, by calculated masses.

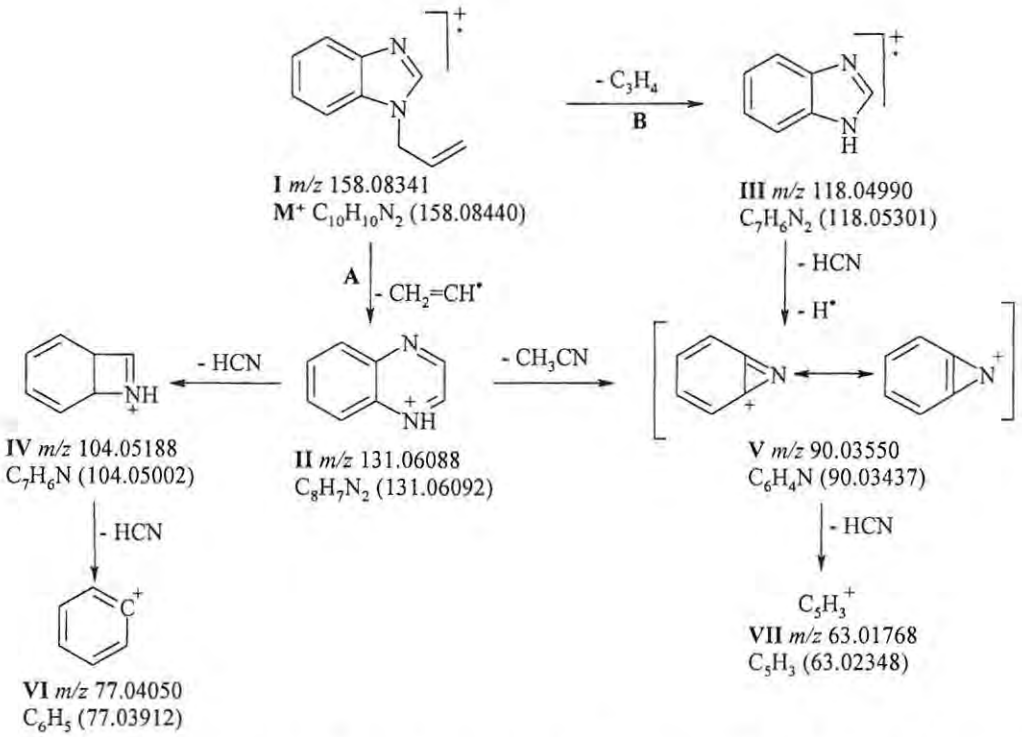
Fragmentation of the molecular ion **I** (*m/z* 533) from compound **138b** (**Scheme 53**), which appears as an intense peak, exhibits a slightly different pattern from that observed in the decomposition of molecular ions of compounds **138a** and **c**. The presence of the odd-electron species **A4** (*m/z* 242), as well as the cation **A8** (*m/z* 73) suggests that, while the glucosyl ion fragment type **A1** (as in **Schemes 51** and **52**) may have been formed, it rapidly decomposes to species **A4** before it can be detected. Fragmentation of the methyl 3-indolylpropanoate radical-cation **B1** (*m/z* 203), which arises from the loss of the glucosyl moiety from the molecular ion follows a similar pattern to that observed in both compound **138a** and **c**. Loss of the methoxy radical from the radical-cation **B1** affords the acylium cation **B2** (*m/z* 172), which loses carbon monoxide to form the cation **B9** (*m/z* 144). The cationic species **B9** may also arise from the loss of an acetate radical (or $CH_3^{\cdot}$ and CO_2) from the radical-cation species **B1**. Loss of an ethyl radical from species **B9** affords the radical-cation **B7** (*m/z* 115). The quinolinium cation **B4** (*m/z* 130) arises from the loss of a methyl acetate radical from species **B1** followed by ring expansion; loss of a hydrogen atom then affords the quinoline radical-cation **B10** (*m/z* 129). The cationic species **B8** (*m/z* 103) is proposed to arise from the loss of hydrogen cyanide from the quinolinium cation **B4**.



Scheme 53. Proposed EI mass fragmentation pathways for the indole derivative **138b**. High-resolution (*m/z*) data are followed, in parentheses, by calculated masses.

2.4.2 Benzimidazole Derivatives

High-resolution EI mass spectrometric studies were also undertaken to investigate the fragmentation pathways of the alkenyl and alkyl benzimidazole derivatives **141**, **144**, **149-151**, **158b** and **160b**. The proposed fragmentation patterns are outlined in **Schemes 54-60**. Fragmentation of the molecular ion **I** (*m/z* 158), which accounts for the base peak in compound **141** (**Figure 72**; **Scheme 54**), occurs *via* two possible pathways, **A** and **B**. In path **A**, the molecular ion **I** loses a vinylic radical to form the quinoxolium cation **II** (*m/z* 131), which then loses hydrogen cyanide or acetonitrile to form the even-electron fragments **IV** (*m/z* 104) or **V** (*m/z* 90), respectively. Both fragments **IV** and **V** then lose hydrogen cyanide to form the phenyl cation **VI** (*m/z* 77) and cation **VII** (m/z 63), respectively. In fragmentation through path **B**, the molecular ion **I** affords the benzimidazole radical cation **III** (*m/z* 118) which can lose HCN and a hydrogen radical to provide an alternative pathway to cation **V** (*m/z* 90).[194]



Scheme 54: Proposed EI mass fragmentation pathways for *N*-allylbenzimidazole **141**. High-resolution (*m/z*) data are followed, in parentheses, by calculated masses.

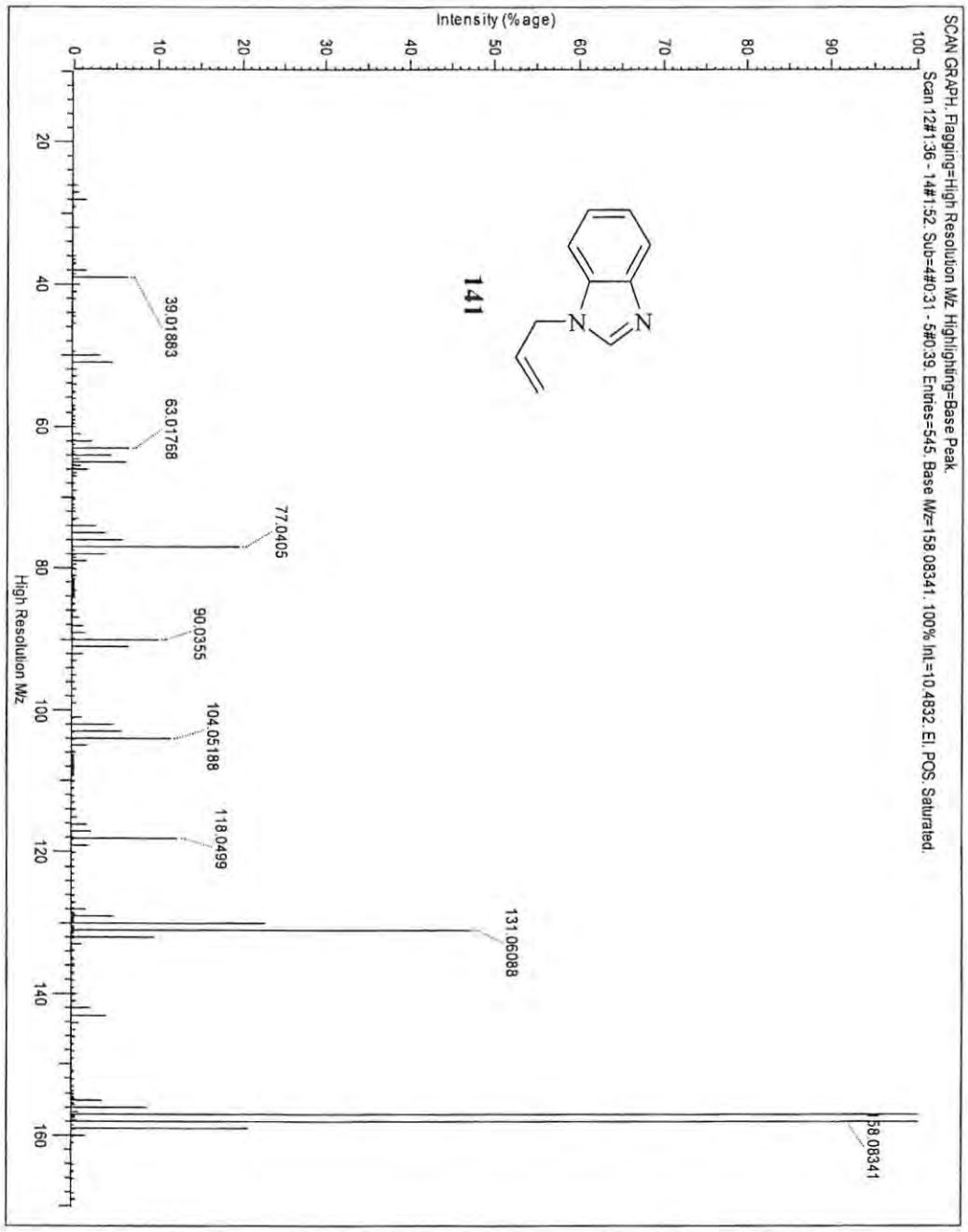
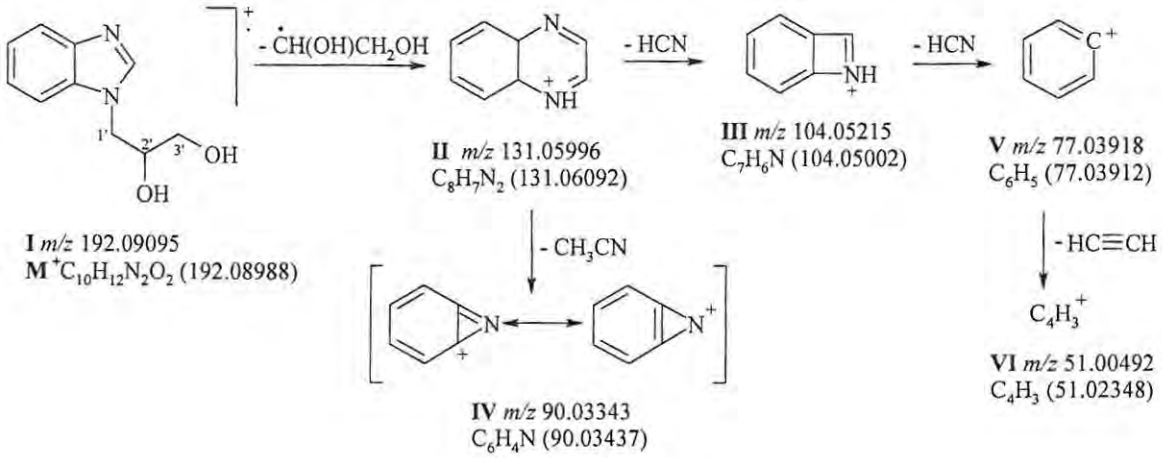


Figure 72: High-resolution EI mass spectrum of the benzimidazole derivative **141**.

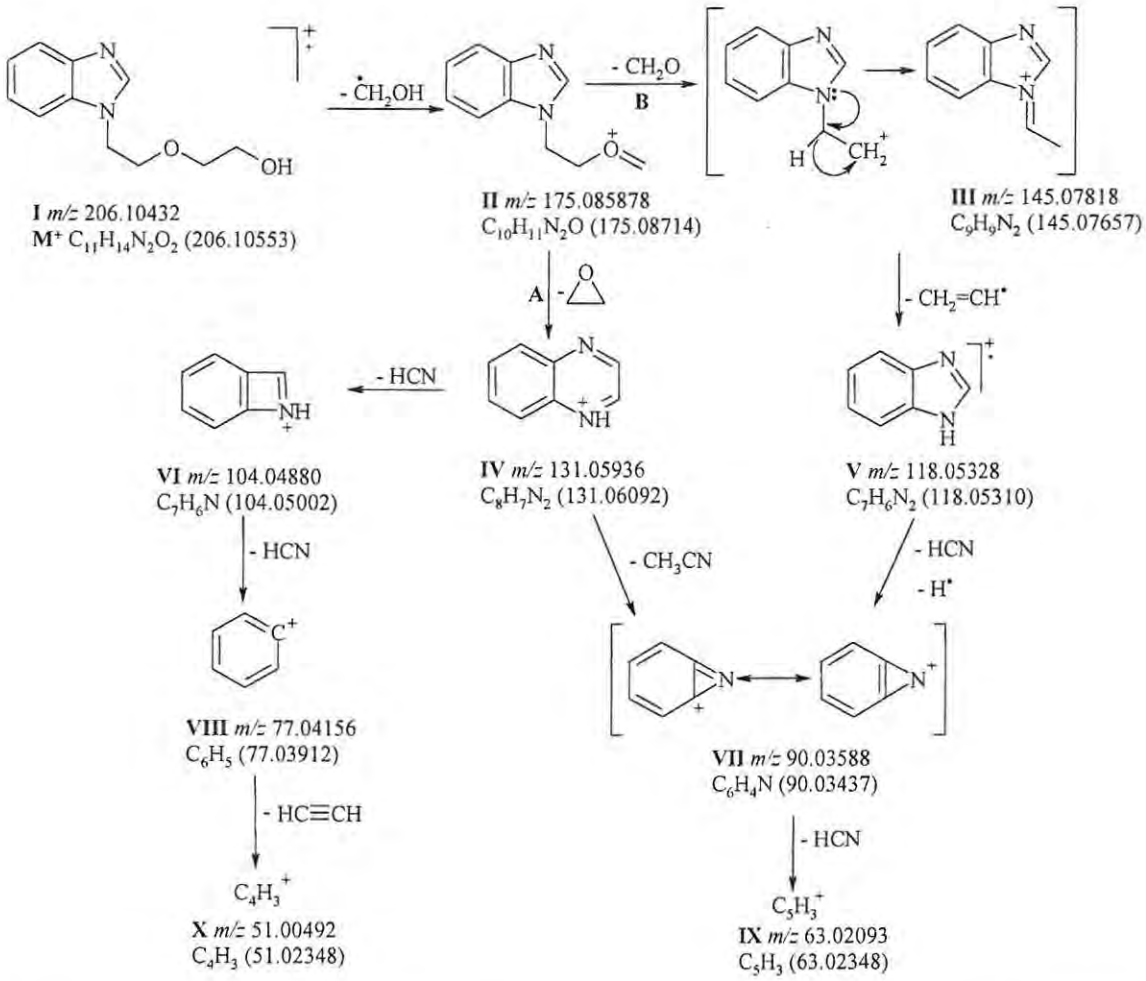
Fragmentation of the molecular ion **I** (*m/z* 192), which appears as a very strong peak (65% relative abundance) in the mass spectrum of compound **144** (**Scheme 55**) shows a similar pattern to the fragmentation, *via* path **A**, of compound **141**. Thus, loss of an ethylene glycol radical affords the quinoxolium cation **II** (*m/z* 131), which accounts for the base peak and which fragments to form the even-electron ions **III** and **IV**. The cation **VI** (*m/z* 51) is attributed to the loss of an ethyne molecule from the phenyl cation **V** (m/z 77).



Scheme 55: Proposed EI mass fragmentation pathways for *N*-(2',3'-dihydroxypropyl)benzimidazole **144**. High-resolution (*m/z*) data are followed, in parentheses, by calculated masses.

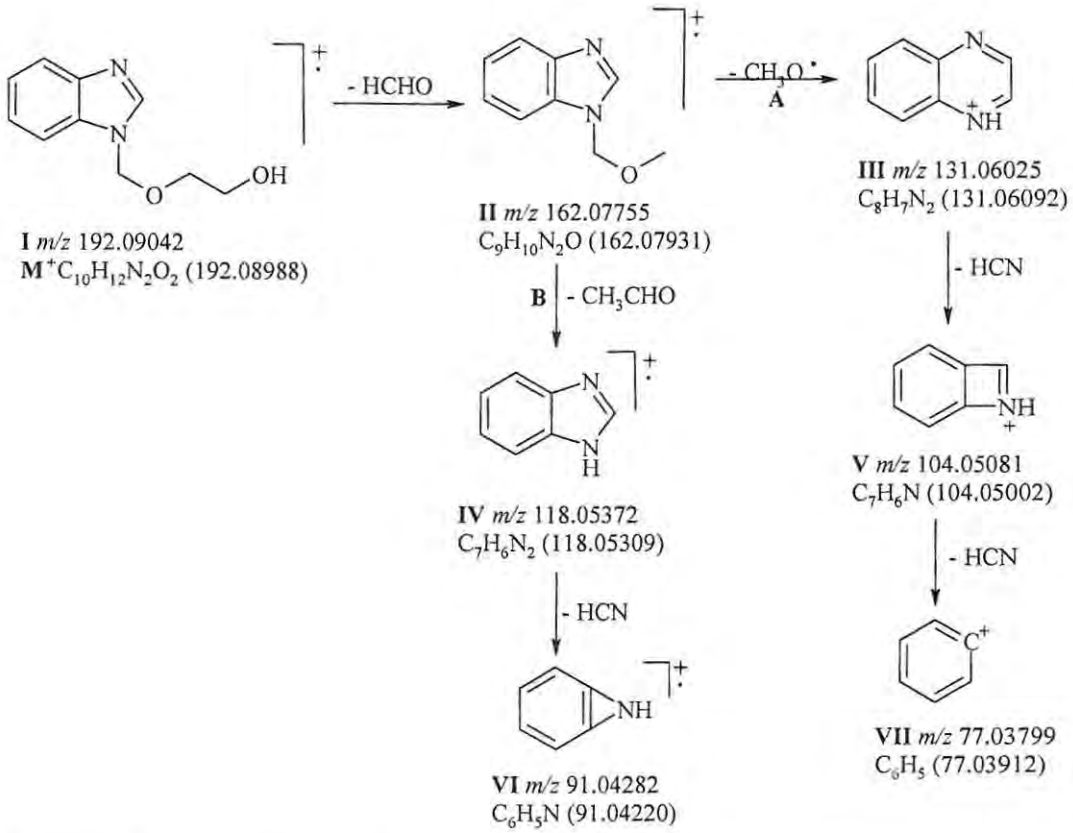
Compounds **149** and **151**, whose structures are very similar, appear to exhibit similar fragmentations of the oxygenated side-chain. Loss of $^{\cdot}CH_2OH$ from the molecular ion **I** (*m/z* 206), which appears as a strong peak (59% relative abundance) in the mass spectrum of compound **149** (**Scheme 56**) affords the even-electron ion **II** (*m/z* 175) which fragments, in turn, *via* two possible pathways, **A** and **B**. Fragmentation through path **A** involves the loss of ethylene oxide to afford a fragment with *m/z* 131, formulated as the quinoxolium species **IV**. This fragment accounts for the base peak and in turn loses HCN and CH_3CN to afford the cations **VI** and **VII**, respectively.

Fragmentation *via* path **B** involves the loss of formaldehyde from fragment **II** to afford the benzimidazolium species **III** (*m/z* 145) arising, we propose, following rearrangement of a 1^{O} carbocation species; loss of a vinylic radical then affords the odd-electron ion **V** (*m/z* 118), which provides access to the aziridinium species **VII** (*m/z* 90). It is apparent that a number of the fragments observed in the mass spectrum of compound **150** are also encountered in the spectra of the analogues **141**, **144**, **150**, **151**, **158b** and **160b**; these include the quinoxalium cation **IV**, the benzimidazolium radical-cation **V**, the benzazete cation **VI**, the benzaziridine fragment **VII** and the phenyl cation **VIII**.



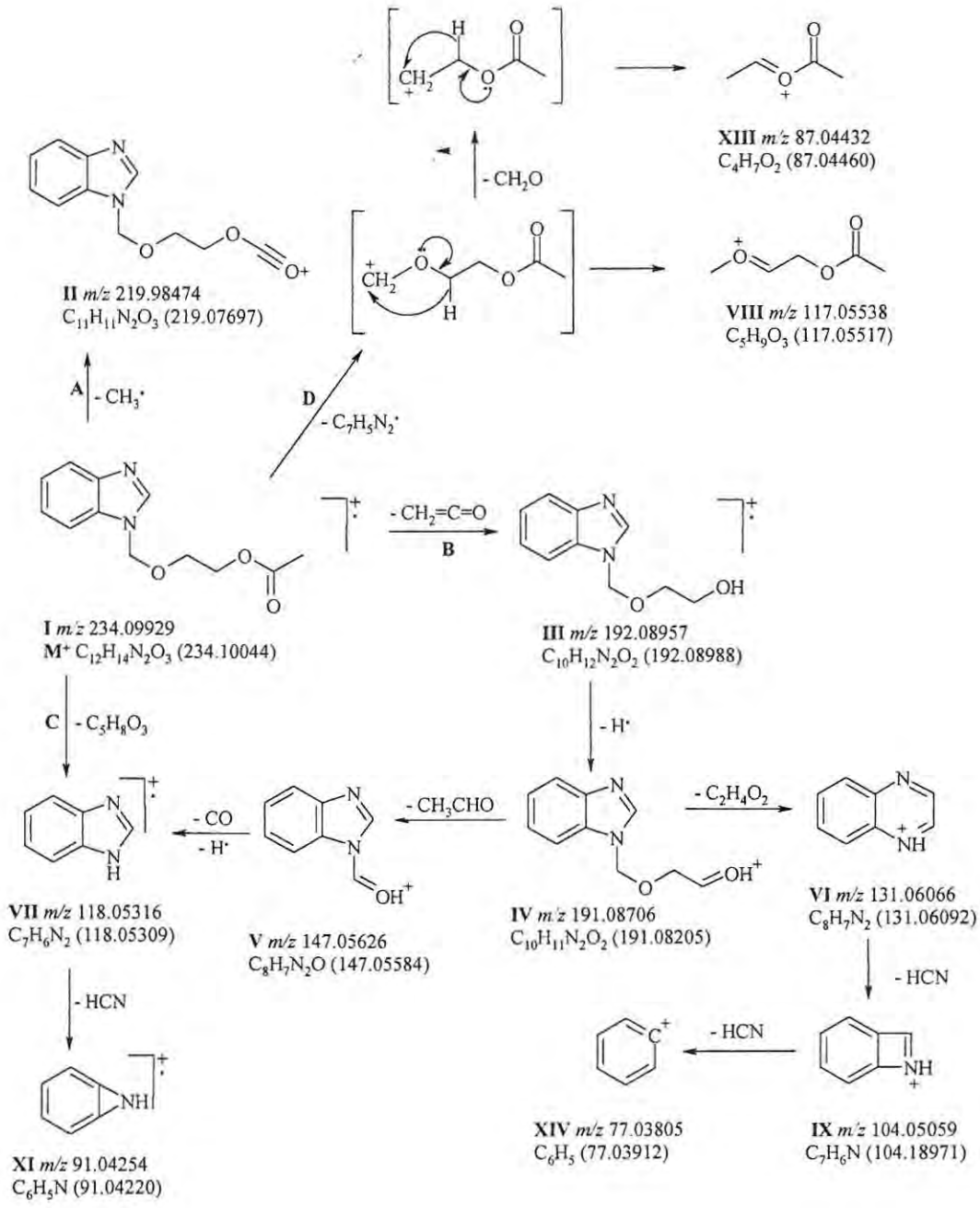
Scheme 56: Proposed EI mass fragmentation pathways for *N*-[2-(2-hydroxyethoxy)ethyl]benzimidazole **149**. High-resolution (*m/z*) data are followed, in parentheses, by calculated masses.

Fragmentation of the molecular ion **I** (*m/z* 192), which appears as a strong peak (71% relative abundance) in the mass spectrum of compound **151** (**Scheme 57**) involves a loss of formaldehyde to afford the odd-electron ion **II** (*m/z* 162) which fragments, in turn, *via* two possible pathways **A** and **B**. Loss of a methoxy radical, in path **A**, affords the even-electron species **III** (*m/z* 131), while loss of acetaldehyde, in path **B**, affords the odd-electron species **IV** (*m/z* 118), which accounts for the base peak. Both fragments **III** and **IV** lose hydrogen cyanide to afford the previously encountered fragments **V**, **VI** and **VII**.



Scheme 57: Proposed EI mass fragmentation pathways for *N*-[(2-hydroxyethoxy)methyl]benzimidazole **151**. High-resolution (*m/z*) data are followed, in parentheses, by calculated masses.

Interesting patterns were noted in the fragmentation of the molecular ion **I** (*m/z* 234), which appears as a very strong peak (93% relative abundance) in the mass spectrum of compound **150** (**Scheme 58**). The molecular ion appears to undergo fission *via* at least four possible pathways **A**, **B**, **C** and **D**.



Scheme 58: Proposed EI mass fragmentation pathways for *N*-[(2-acetoxyethoxy)methyl]benzimidazole **150**. High-resolution (*m/z*) data are followed, in parentheses, by calculated masses.

Fragmentation of molecular ion **I**, *via* path **A**, involves the loss of a methyl radical to afford the acylium cation **II** (*m/z* 219). Loss of a ketene molecule, in path **B**, affords the odd-electron species **III** (*m/z* 192), which then loses a hydrogen atom to afford the even-electron species **IV** (*m/z* 191), which in turn fragments *via* two further pathways. Both fragments **II** and **IV**, which are barely visible (0.64% and 0.71% relative abundance, respectively), are the expected fragments from decomposition of an acetate ester.[193] Cation **IV** then fragments to afford the quinoxalium cation **VI** (*m/z* 131) which, in turn, undergoes successive loss of two hydrogen cyanide molecules to afford fragments **IX** (*m/z* 104) and **XIV** (*m/z* 77), respectively. The presence of the benzimidazolyl radical cation **VII** (*m/z* 118) may be attributed to the loss of either carbon monoxide and a hydrogen atom from fragment **V** or *via* direct fragmentation of the molecular ion **I** *via* path **C**. Elimination of acetaldehyde from cation IV would account for the formation of fragment **V**. Loss of a heterocyclic radical, *via* path **D**, affords cation **VIII** (*m/z* 117) which, we propose, arises from the rearrangement of the initial 1^{O} carbocation species, which can also lose formaldehyde to form cation **XIII** (*m/z* 87), following a similar intramolecular rearrangement. Fragment **VIII** accounts for the base peak.

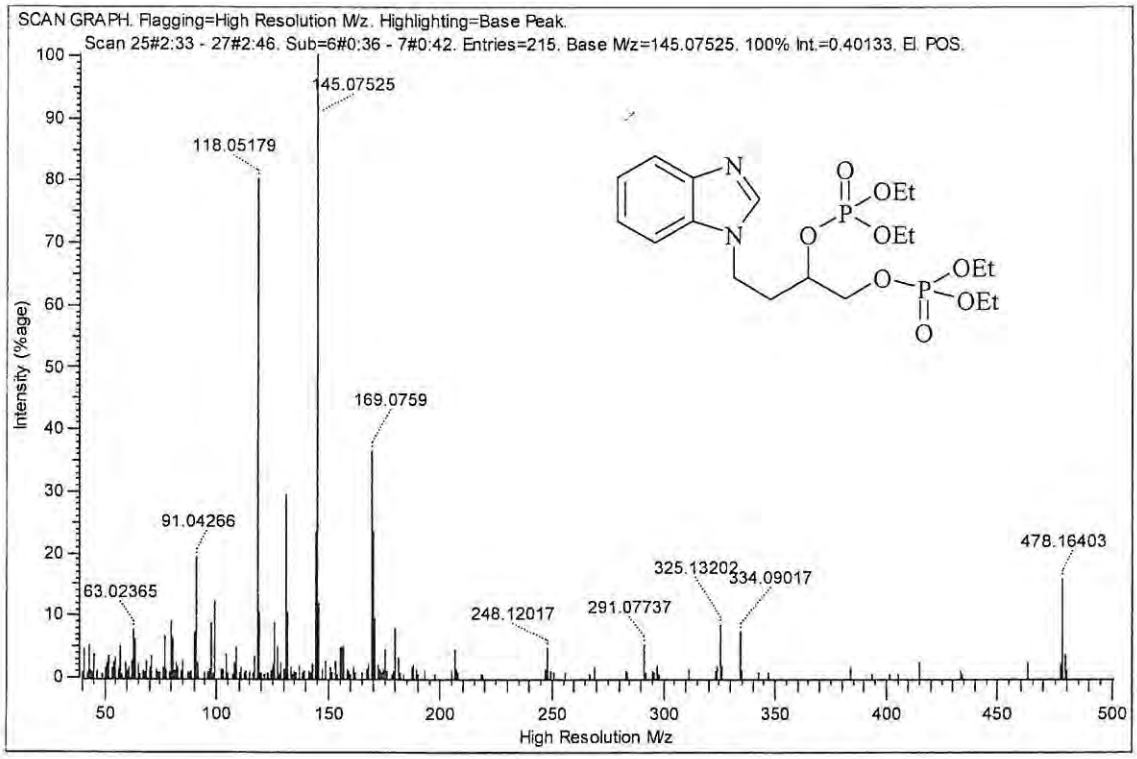
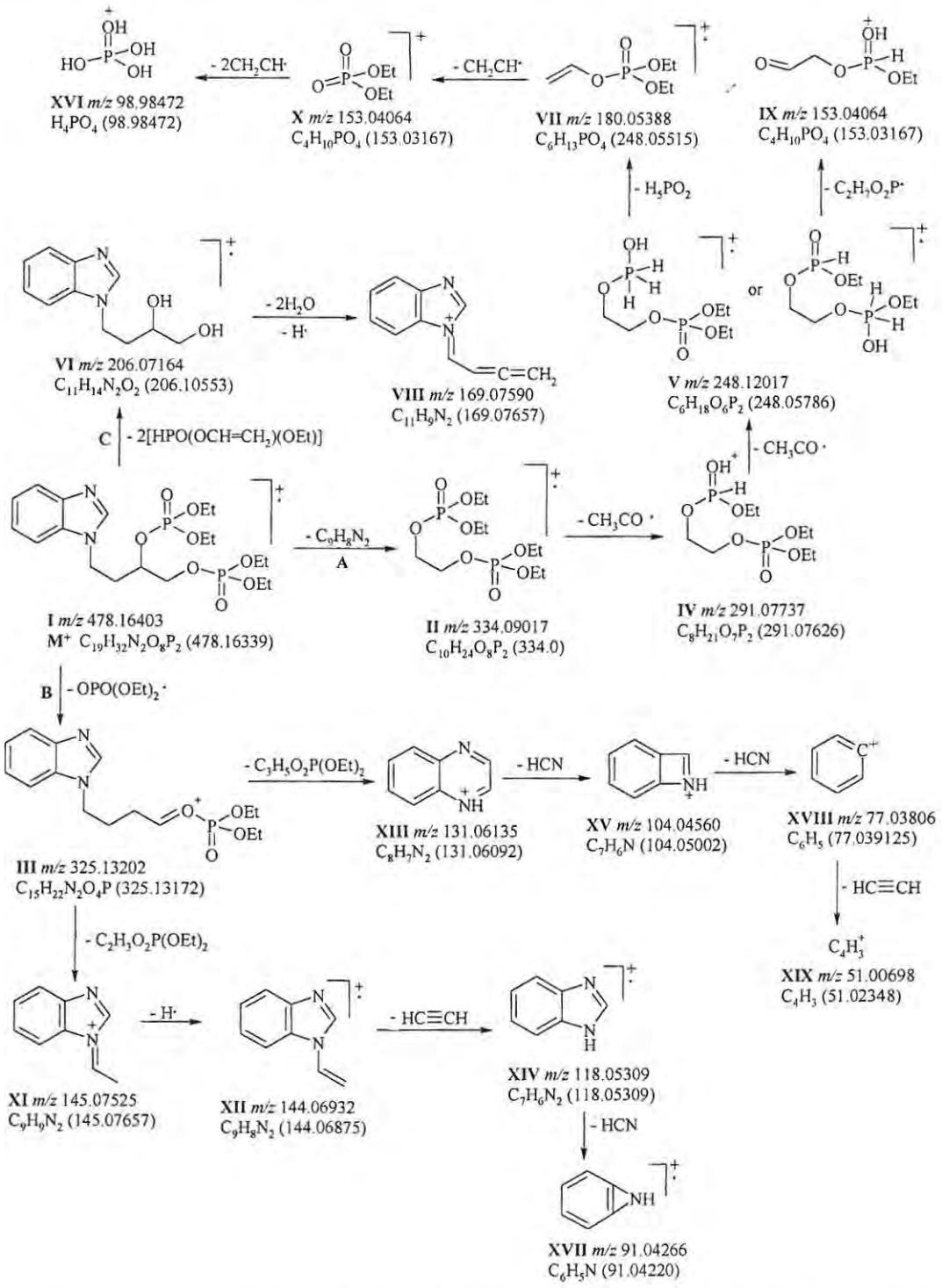


Figure 73: High-resolution EI mass spectrum of 4-(benzimidazol-1-yl)butane-1,2-diol bis(diethyl phosphate) **158b.**

Pentavalent phosphoric ester derivatives have been reported to give fragments arising from a McLafferty rearrangement involving the transfer of two hydrogens to the phosphorus oxygen bond,[195-197] and this pattern was also observed with our phosphorylated benzimidazole derivatives **158b** and **160b**. Fragmentation of molecular ion **I** (*m/z* 478), which appears as a weak peak (16% relative abundance) in the mass spectrum of compound **158b** (**Figure 73**; **Scheme 59**) is proposed to occur *via* three possible pathways **A**, **B** and **C**. Loss of *N*-vinylbenzimidazole, in path **A**, affords the odd-electron species **II** (*m/z* 334), which then undergoes successive loss of two acetyl radicals and the migration of two hydrogens from each of the two leaving ethoxy groups to the phosphorous moiety to afford ions **IV** (*m/z* 291) and **V** (*m/z* 248), which can be represented as either structure **A** or **B**. Loss of H_2PO_2Et from fragment **V** affords cation **IX** (*m/z* 153), while loss of H_5PO_2 from fragment **V** affords the radical cation **VII** (*m/z* 180) which, on elimination of a vinylic radical, leads to the cation **X** (*m/z* 153). Successive loss of two vinylic radicals from fragment **X** affords cation **XVI** (*m/z* 98).

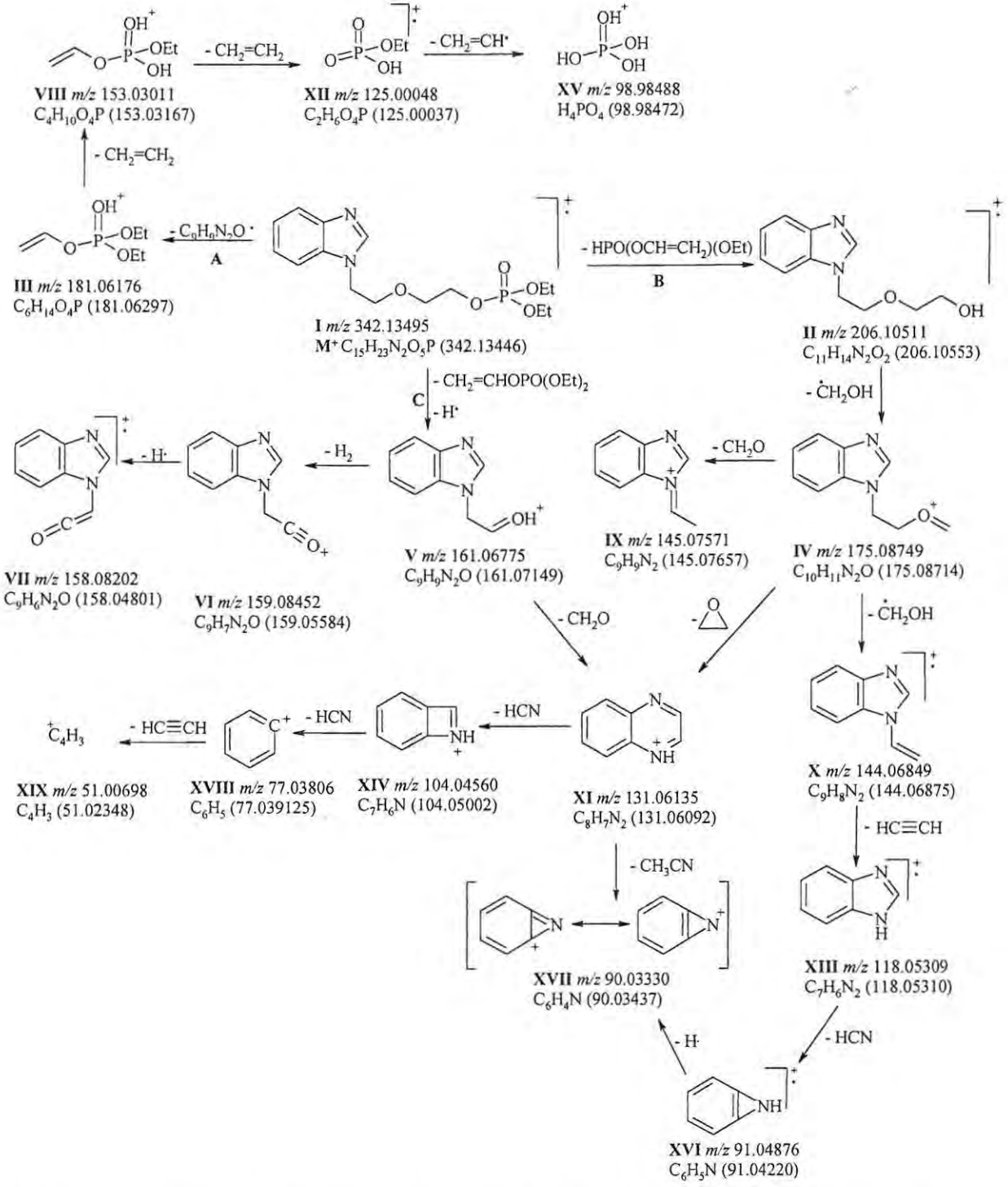
Fragmentation *via* path **B** involves successive loss of phosphate moieties to afford the odd-electron species **VI** (*m/z* 206), which then loses two water molecules and a hydrogen atom to afford the even-electron species **VIII** (*m/z* 169). The molecular ion **I** may also lose a phosphate radical to afford ion **III** (*m/z* 325) which, in turn, undergoes elimination involving:- i) loss of $H_2O_2P(OEt)_2$ to afford the cation **VIII** (*m/z* 169); ii) loss of $CH_3CHCHOPO(OEt)_2$ (diethyl propenylphosphate) to afford the quinoxalium cation **XIII** (*m/z* 131); and iii) loss of $CH_2CHOPO(OEt)_2$ (diethyl vinylphosphate) to afford the cation **XI** (*m/z* 145), which accounts for the base peak. Loss of a hydrogen atom from fragment **XI** followed by the elimination of acetylene affords ions **XII** (*m/z* 144) and **XIV** (*m/z* 118), respectively. Both fragments **XII** and **XIV** lose hydrogen cyanide and follow the previously observed patterns for the formation of fragments **XV**, **XVII**, **XVIII** and **XIX**.



Scheme 59: Proposed EI mass fragmentation pathways for 4-(benzimidazol-1-yl)butane-1,2-diol bis(diethyl phosphate) **158b**. High-resolution (*m/z*) data are followed, in parentheses, by calculated masses.

Similarly fragmentation of the molecular ion **I** (*m/z* 342), which appears as a weak peak (10% relative abundance) in the mass spectrum of compound **160b** (**Scheme 60**), seems to follow three possible pathways. Loss of the benzimidazolyl radical, *via* path **A**, affords the cation **III** (*m/z* 181), which undergoes successive elimination of two ethylene molecules to afford cations **VIII** (*m/z* 153) and **XII** (*m/z* 125); subsequent loss of $CH_2CH\cdot$ affords cation **XV** (*m/z* 98). Loss of the phosphate moiety as CH_2CHO_2PHOEt (ethyl vinyl hydrogen phosphate), *via* path **B**, affords the odd-electron species **II** (*m/z* 206), which corresponds to the molecular ion **I** from compound **149** and which follows the same fragmentation pattern (**Scheme 56**). Loss of a hydroxymethyl radical from fragment **IV** (*m/z* 175) affords the odd-electron species **X** (*m/z* 144) which loses acetylene to provide an alternative pathway to the benzimidazolyl radical cation **XIII** (*m/z* 118).

Fragmentation *via* path **C** involves the successive loss of the phosphate moiety as diethyl vinyl phosphate and a hydrogen atom to afford the even-electron species **V** (*m/z* 161), which may either lose formaldehyde and thus provides an alternative pathway to the quinoxalium cation **XI** (*m/z* 131) or H_2 and a hydrogen atom to afford cations **VI** (*m/z* 159) and **VII** (*m/z* 158), respectively. Unlike compound **149** whose base peak is attributed to fragment **XIII** (*m/z* 118), the base peak for compound **160b** corresponds to fragment **X** (*m/z* 144).



Scheme 60: Proposed EI mass fragmentation pathways for 2-[2-(benzimidazol-1-yl)ethoxy]ethyl diethyl phosphate **160b**. High-resolution (*m/z*) data are followed, in parentheses, by calculated masses.

2.4.3 Pyrazolo[3,4-*d*]pyrimidine Derivatives

High-resolution EI mass spectrometric data was also used to explore the possible fragmentation pathways for the glycosylylated pyrazolo[3,4-d]pyrimidine derivatives. The fragmentation patterns observed with 1-[3,4,5,6-tetra-*O*-acetyl-D-glucopyranosyl]-4-aminopyrazolo[3,4-*d*]pyrimidine **167b** (**Figure 74**) shows a similar pattern to that observed with the indolylalkanoic acid derivatives **138a-c**. Fragmentation of the molecular ion **I** (*m/z* 465), which appears as a very weak peak (**Scheme 61**) appears to involve three possible pathways, **A**, **B** and **C**. Path **A** involves the loss the 4-aminopyrazolo[3,4-*d*]pyrimidine radical to form the glucosyl cation **A1** (*m/z* 331), which then follows the same fragmentation pattern observed for the indolylalkanoic acid derivatives **137a** and **c**. Fragmentation of the molecular ion **I** *via* path **B** involves the loss of an acetate radical to afford cation **II** (*m/z* 306), followed by successive loss of two molecules of acetic acid and a molecule of acetaldehyde to afford ions **III** (*m/z* 346), **IV** (*m/z* 286) and **V** (*m/z* 242), respectively. The even-electron fragment **V** (*m/z* 242) is, in fact, the base peak. Fragmentation through path **C** involves the loss of a glucosyl radical or a neutral glucosyl derivative to form fragments **VI** (*m/z* 136) and **VII** (*m/z* 135), following the general pattern observed in fragmentation of purines.[198] Fragment **VII** (*m/z* 135) could also arise from the loss of H· from cation **VI**. The radical-cation **VII** (*m/z* 135) then loses two hydrogen cyanide to afford the radical-cation species (*m/z* 81) tentatively formulated as the triazine **VIII**. Due to the low volatility of the allopurinol derivative **167a**, the high-resolution EI mass spectrum could not be obtained, but the fast atom bombardment (FAB) mass spectrometric data indicates that its fragmentation is similar to that observed for compound **167b**.

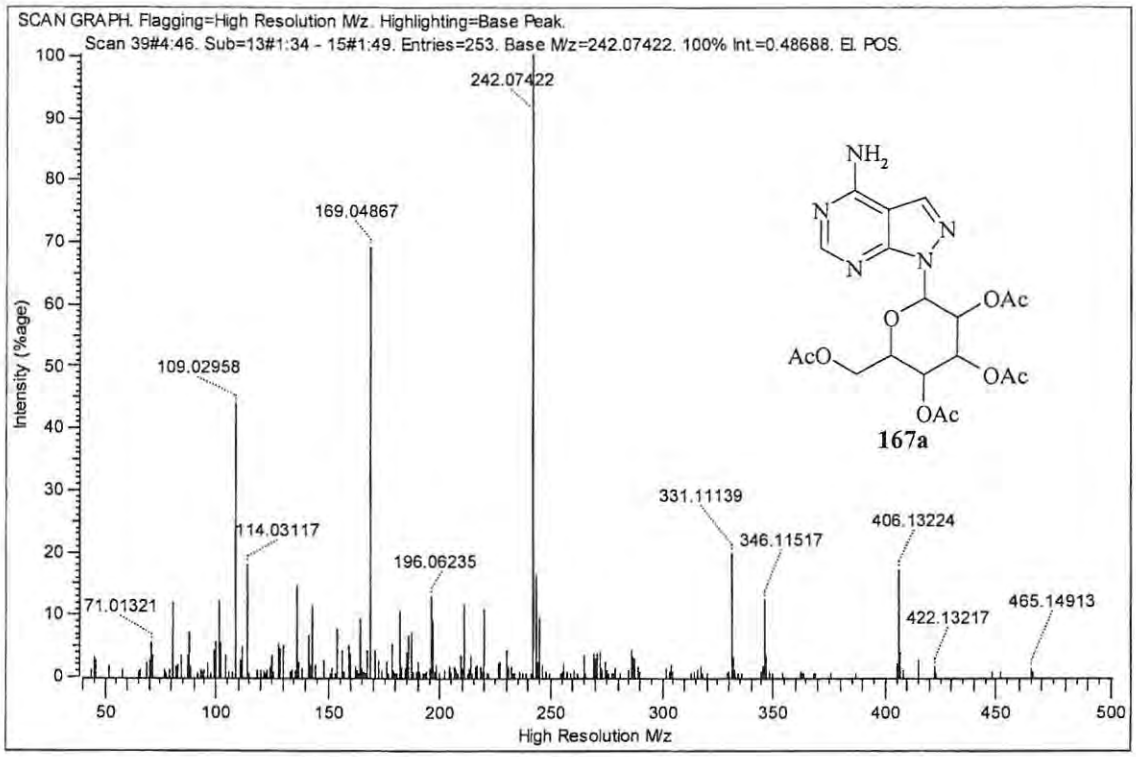
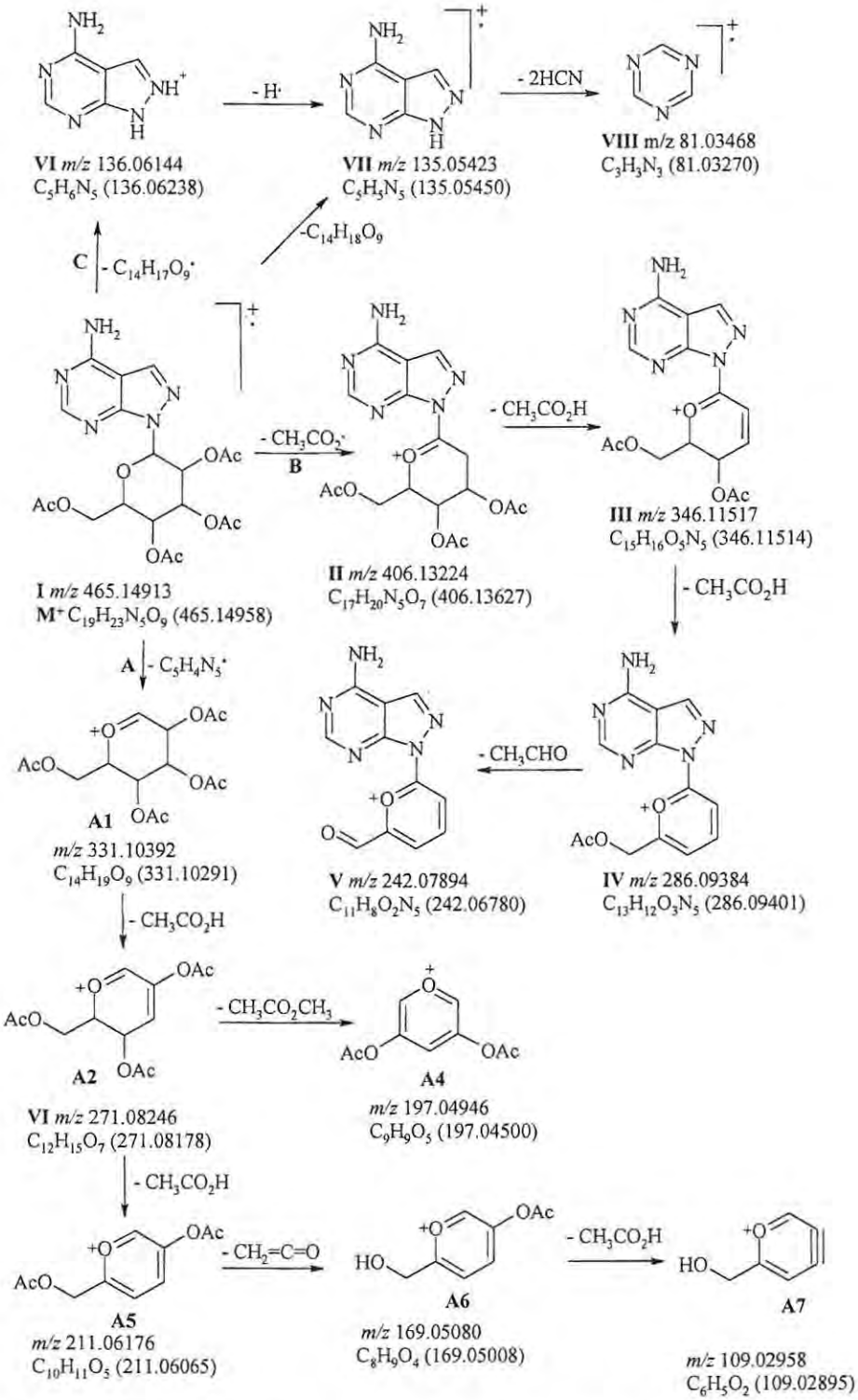


Figure 74: High-resolution EI mass spectrum of the 4-aminopyrazolo[3,4-*d*]pyrimidine derivative **167b.**



Scheme 61: Proposed EI mass fragmentation pathways for the 4-aminopyrazolo[3,4-*d*]pyrimidine derivative **167b**. High-resolution (*m/z*) data are followed, in parentheses, by calculated masses.

2.5 Conclusions

Compounds containing 3-indolylalkanoic acid, benzimidazole, 4-aminopyrazolo[3,4-*d*]pyrimidine and allopurinol nuclei have been shown to have various pharmacological properties,[104-134] and appeared to provide promising scaffolds for the synthesis of ATP analogues, designed as potential GS inhibitors. It was hoped that screening-assay data would permit an activity-driven approach but, in the absence of such data, the initial structural analogy approach had to be continued throughout the project.

The cyclic glucosylated derivatives of 3-indolylalkanoic acid, benzimidazole, 4-aminopyrazolo[3,4-*d*]pyrimidine and allopurinol, were successfully prepared by the Lewis acid-catalysed nucleophilic displacement of the acetate at the anomeric centre (C-1) by the heterocyclic base. These reactions were conducted at elevated temperatures, as it was observed that at room temperature, only the precursors were isolated.

N-Alkylated benzimidazole and 4-aminopyrazolo[3,4-*d*]pyrimidine derivatives were successfully obtained by treating each of these heterocyclic bases with the bromoketal **128**, (2-bromomethoxy)ethyl acetate **131** or (2-chloroethoxy)ethanol **148** in the presence of NaH. However, under the same conditions the 3-indolylalkanoic acid esters **110a-c** failed to exhibit *N*-alkylation, forming the corresponding esters **134a-c** and **135a-c** instead. These products appear to arise *via* nucleophilic *O*-alkyl cleavage of the ester groups. Although these esters were not originally targeted, their lowest-energy bound conformations, obtained from the interactive docking studies, exhibit a degree of structural homology with ADP and ATP in the GS active site, with the polyoxy moieties orientated towards the magnesium metal ions.

Phosphorylation of selected of benzimidazole and 4-aminopyrazolo[3,4-*d*]pyrimidine derivatives was achieved, in dissappointing yields, as either mono- or 1,2-diphosphates. The required 1,2-dihydroxybenzimidazole derivatives were obtained in three steps by treating benzimidazole with an alkenyl bromide followed by oxidation of the alkenyl moiety with CTAP; the resulting diols were reacted with diethyl chlorophosphate in the

presence of BuLi. Monohydroxy benzimidazole and 4-aminopyrazolo[3,4-*d*]pyrimidine derivatives were treated directly with diethyl chlorophosphate in the presence of BuLi. The best, bound conformations of the monophosphate derivatives (from the interactive docking studies) assume the general arrangement of ADP and ATP, with the polyoxy moieties orientated towards the metal ions. However, with the 1,2-diphosphates, the heterocyclic base is orientated towards the metal ions.

A wide series of ATP analogues have thus been prepared, albeit in variable yields, providing a 'library' of compounds for screening as potential GS inhibitors, selected examples of which have been subjected to *in silico* docking into the GS active site. The interactive docking studies, using the Accelrys Ligand Fit module, suggest that these derivatives could exhibit various hydrogen-bonding interactions with the amino acid residues in the GS active site.

The experimental ^{13}C NMR chemical shift data for selected glucosylated 3-indolylalkanoic acid, benzimidazole, 4-aminopyrazolo[3,4-*d*]pyrimidine and allopurinol structures (compounds **116c**, **138a-c**, **156**, **167a,b** and **168a,b**) were compared with data obtained from Modgraph and ChemWindow NMRpredict programmes. The NMR prediction data proved useful in confirming certain structural assignments, although the Modgraph HOSE data generally gives results which are more comparable with the experimental data than the Modgraph Nnet and ChemWindow data.

High-resolution EI mass spectrometric studies of selected 3-indolylalkanoic acid, benzimidazole and pyrazolo[3,4-*d*]pyrimidine derivatives were undertaken revealing common fragmentation patterns in each class of derivatives. Final confirmation of the proposed fragmentation pathways, however, awaits the availability of metastable peak and/or B/E linked scan data.

While the general aims of the present study have been realized, further work is expected to address some of the following issues.

i) The problem of aqueous insolubility, by generating water-soluble hydrochloride salts of the targeted products and thus facilitating biological testing in aqueous media.

ii) A programmed biological evaluation of the library compounds prepared.

iii) Optimization of the reaction conditions and, hence, the yields for interesting biologically active analogues.

iv) Obtaining B/E linked scan or metastable peak MS data to confirm the proposed fragmentation pathways.

v) Complexation studies of the synthesized ATP analogues with either Mg or Mn metal ions, thus permitting complex geometry to be explored both *in silico* (using DFT computational methods) and experimentally (using X-ray crystallographic data).

3. Experimental

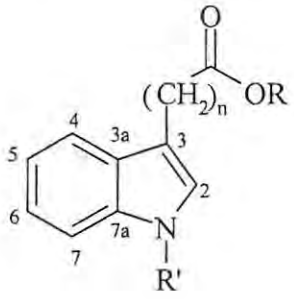
General

Reagents were used as supplied by the manufacturers. Solvents were re-distilled before use except for DMF, diethyl ether, THF, acetonitrile, pyridine and methanol, which were dried by literature methods[199-201] and distilled before use. The reactions prepared with dry solvents were carried out in an inert atmosphere (either nitrogen or argon) unless otherwise stated. Plastic plates, pre-coated with silica gel $60F_{254}$ (as supplied by Merck), were used for thin layer chromatography (TLC) and visualization was effected by inspection under UV light. Silica gel 60 (particle size 0.040-0.063nm) was used as the stationary phase for flash chromatography.[202] A Phenomenex C-18 LUNA semi-preparative column was used for reverse-phase HPLC and a Whatmans Partisil 10 semi-preparative column was used for normal phase HPLC.

^{1}H, ^{13}C and ^{31}P NMR spectra were recorded on a Bruker AVANCE 400MHz spectrometer. Chemical shifts are reported relative to the solvent peaks (δ_H: 7.25ppm for $CDCl_3$, 4.81ppm for D_2O, 3.30ppm for $MeOD_4$ and 2.50ppm for DMSO-d_6; δ_C: 77.0ppm for $CDCl_3$, 49.0ppm for CD_3OD and 39.4ppm for DMSO-d_6). Infrared spectra were recorded on a Perkin-Elmer spectrum 2000FT-IR spectrometer either as KBr discs or as thin films between NaCl windows (mid-infrared). Melting points were determined on a Reichert hot-stage apparatus and are uncorrected.

Low-resolution mass spectra were recorded on a Finnigan-MAT GCQ mass spectrometer, while low-resolution mass spectrometric data were obtained by Martin Brits using APCI and high- resolution mass analysis was obtained by Tommie van de Walt (University of Witwatersrand Mass Spectrometry Unit).

3-Indolylalkanoic Acid Derivatives



Methyl 3-indolylacetate 110a[139]

Thionyl chloride (7.18g, 60.4mmol) was added drop-wise to stirred dry methanol (50ml) cooled to –30°C. After stirring for 30min, 3-indolylacetic acid **105** (5.26g, 30.0mmol) was added in one portion. The mixture was stirred at room temperature for 4h before adding H_2O (50ml) and evaporating the MeOH *in vacuo*. The aqueous phase was extracted with EtOAc (3 x 15ml) and the combined organic solutions were dried over anhydrous Na_2SO_4. The solvent was evaporated *in vacuo* and the residual oil was purified by flash chromatography [on silica gel; elution with EtOAc-hexane (1:2)] to afford, as a pale yellow oil, methyl 3-indolylacetate **110a** (5.45g, 96.1%); ν_{max} (thin film)/cm^{-1} 3410.2 (NH) and 1730.9 (C=O); δ_H (400MHz; $CDCl_3$) 3.71 (3H, s, CH_3), 3.79 (2H, s, CH_2), 7.07 (1H, s, 2-H), 7.15 (1H, t, J = 7.8Hz, 5-H), 7. (1H, t, J = 7.8Hz, 6-H), 7.31 (1H, d, J = 7.8Hz, 4-H); 7.63 (1H, d, J = 7.8Hz, 7-H) and 8.19 (1H, s, NH); δ_C (100MHz; $CDCl_3$) 31.1 (CH_2), 52.0 (OCH_3), 108.3 (C-3), 111.2 (C-4), 118.8 (C-7), 119.6 (C-5), 122.2 (C-6), 123.1 (C-2), 127.2 (C-3a), 136.1 (C-7a) and 172.6 (C=O)

Methyl 3-indolylpropanoate 110b[140]

The experimental employed for the preparation of methyl 3-indolylacetate **110a** was followed, using thionyl chloride (0.655g, 5.51mmol), 3-indolylpropanoic acid **106** (0.590g, 3.12mmol) and dry methanol (5ml). The crude product was purified by flash chromatography [on silica gel; elution with EtOAc-hexane (1:2)] to afford, as off-white

crystals, methyl 3-indolylpropanoate **110b** (0.426g, 67.7%), mp. 78-80^{O}C (lit.[140] oil); ν_{max} (thin film)/cm^{-1} 3413.1 (NH) and 1726.2 (C=O); δ_H (400MHz; $CDCl_3$) 2.75 (2H, t, J = 7.6Hz, CH_2CH_2CO), 3.14 (2H, t, J = 7.6Hz, CH_2CH_2CO), 3.70 (3H, s, OCH_3), 6.97 (1H, s, 2-H), 7.15 (1H, t, J = 7.4Hz, 5-H), 7.21 (1H, t, J = 7.4Hz, 6-H), 7.34 (1H, d, J = 8.0Hz, 4-H), 7.62 (1H, d, J = 8.0Hz, 7-H) and 8.10 (1H, s, NH); δ_C (100MHz; $CDCl_3$) 20.6 (CH_2CH_2CO), 34.7 (CH_2CH_2CO), 51.5 (OCH_3), 111.1 (C-4), 114.7 (C-3), 118.6 (C-7), 119.2 (C-5), 121.4 (C-2), 121.9 (C-6), 127.1 (C-3a), 136.2 (C-7a) and 173.9 (C=O).

Methyl 3-indolylbutanoate 110c[203]

The experimental employed for the preparation of methyl 3-indolylacetate **110a** was followed, using thionyl chloride (4.18g, 35.2mmol), 3-indolylbutanoic acid **107** (3.52g, 17.3mmol) and dry methanol (10ml). The crude product was purified by flash chromatography [on silica gel; elution with EtOAc-hexane (1:2)] to afford, as off white crystals, methyl 3-indolylbutanoate **110c** (3.26g, 88.3%), mp. 61-63^{O}C; ν_{max} (thin film)/cm^{-1} 3419 (NH) and 1717.4 (C=O); δ_H (400MHz; $CDCl_3$) 2.06 (2H, quintet, J = 7.4Hz, $CH_2CH_2CH_2CO$), 2.40 (2H, t, J = 7.4Hz, $CH_2CH_2CH_2CO$), 2.81 (2H, t, J = 7.4Hz, $CH_2CH_2CH_2CO$), 3.66 (3H, s, OCH_3), 6.97 (1H, s, 2-H), 7.11 (1H, t, J = 7.4Hz, 5-H), 7.19 (1H, t, J = 7.4Hz, 6-H), 7.34 (1H, d, J = 8.0Hz, 4-H), 7.61 (1H, d, J = 8.0Hz, 7-H) and 8.01 (1H, s, NH); δ_C (100MHz; $CDCl_3$) 24.9 ($CH_2CH_2CH_2CO$), 25.8 ($CH_2CH_2CH_2CO$), 34.1 ($CH_2CH_2CH_2CO$), 51.9 (OCH_3), 111.0 (C-7), 115.5 (C-3), 118.8 (C-4), 119.1 (C-5), 121.4 (C-2), 121.9 (C-6), 127.4 (C-3a), 136.4 (C-7a) and 174.2 (C=O).

tert-*Butyldimethylsilyl 3-indolylacetate* 111a

To a solution of 3-indolylacetic acid **105** (0.198g, 1.13mmol) in dry DMF (6ml), TBDMSCl (0.186g, 1.24mmol) was added in one portion and the mixture stirred until dissolution was complete. To this solution, imidazole (0.165g, 2.42mmol) was added in one portion and the mixture was stirred at 25^{O}C for 48h. The reaction was quenched with H_2O (10ml) and extracted with petroleum ether [bp.40-60^{O}C (2 x 4ml). The organic layer

was washed with saturated solution of $NaHCO_3$ and dried over $MgSO_4$. The solvent was removed *in vacuo* to afford, as colourless crystals, tert-*butyldimethylsilyl 3-indolylacetate* **111a** (0.138g, 44%), (Found: M^+, 289.14933 $C_{16}H_{23}NO_2Si$ requires *M*, 289.14981), mp 80-82^{O}C; ν_{max} (thin film)/cm^{-1} 3386.4 (NH) and 1717.0 (C=O); δ_H (400MHz; $CDCl_3$) 0.25 (6H, s, $SiCH_3$), 0.87 (9H, s, CCH_3), 3.79 (2H, s, CH_2), 7.11-7.17 (2H, m, 2- and 5-H), 7.20 (1H, t, *J* = 7.6Hz, 6-H), 7.33 (1H, d, *J* = 8.0Hz, 4-H), 7.59 (1H, d, *J* = 8.0Hz, 7-H) and 8.10 (1H, s, NH); δ_C (100MHz; $CDCl_3$) –4.8 ($SiCH_3$), 17.6 [Si*C*$(CH_3)_3$], 24.5 [C(*C*$H_3)_3$], 32.9 (CH_2), 109.0 (C-7), 111.1 (C-3), 118.8 (C-4), 119.5 (C-6), 122.0 (C-5), 122.9 (C-2), 127.3 (C-3a), 136.1 (C-7a) and 172.4 (C=O); *m/z* 290 (M + 1, 100%).

tert-*Butyldimethylsilyl 3-indolypropanoate* 111b

The experimental procedure employed for the synthesis of *tert*-butyldimethylsilyl 3-indolylacetate **111a** was followed, using 3-indolylpropionic acid **106** (1.001g, 5.296mmol) in dry DMF (6ml), TBDMSCl (0.963g, 6.39mmol) and imidazole (0.914g, 13.4mmol). Work-up afforded, as colourless oil, tert-*butyldimethylsilyl 3-indolylpropanoate* **111b** (0.755g, 47.0%); ν_{max} (thin film)/cm^{-1} 3386.0 (NH) and 1717.6 (C=O); δ_H (400MHz; $CDCl_3$) 0.25 (6H, s, $SiCH_3$), 0.91 (9H, s, CCH_3), 2.74 (2H, t, *J* = 7.6Hz, $CH_2C\mathit{H}_2CO$), 3.09 (2H, t, *J* = 7.6Hz, $C\mathit{H}_2CH_2CO$), 6.98 (1H, s, 2-H), 7.11 (1H, t, *J* = 7.2Hz, 5-H), 7.18 (1H, t, *J* = 7.2Hz, 6-H), 7.33 (1H, d, *J* = 8.0Hz, 7-H), 7.60 (1H, d, *J* = 8.0Hz, 4-H) and 8.14 (1H, s, NH); δ_C (100MHz; $CDCl_3$) -3.6 ($SiCH_3$), 17.6 [Si*C*$(CH_3)_3$], 20.8 ($C\mathit{C}H_2$), 25.6 [C(*C*$H_3)_3$], 36.6 (CH_2CO), 111.1 (C-3), 115.0 (C-7), 118.7 (C-4), 119.2 (C-6), 121.3 (C-5), 121.9 (C-2), 127.2 (C-3a), 136.3 (C-7a) and 173.9 (C=O).

2,3,4,5,6-Penta-*O*-acetyl-D-gluconyl chloride 112[144]

Method A

PCl_5 (3.172g, 15.23mmol) was added slowly to a vigorously stirred solution of 2,3,4,5,6-penta-*O*-acetyl-D-gluconic acid **115** (5.035g, 12.40mmol) in dry diethyl ether (37ml) in a flask fitted with a calcium chloride guard tube. After stirring overnight at room

temperature, un-dissolved solid material was removed by filtration through a sintered glass funnel into a round-bottomed flask. The ethereal solution was evaporated to half volume *in vacuo* and left overnight in a freezer. The resulting colourless crystals were filtered and washed with ice-cold petroleum ether (bp. 40-60^{O}C) to afford 2,3,4,5,6-penta-*O*-acetyl-D-gluconyl chloride **112** (3.302g, 62.8%), mp. 60-62^{O}C (lit.[144] 68-71^{O}C), ν_{max} (thin film)/cm^{-1} 1717.0 (C=O); δ_H (400MHz; $CDCl_3$) 2.04 (3H, s, 6-CO_2CH_3), 2.07-2.08 (9H, 3 x s, 3-, 4- and 5-CO_2CH_3), 2.20 (3H, s, $COCHOCH_3$), 4.11 (1H, dd, J = 5.6 and 12.4Hz, 6-H $_a$), 4.30 (1H, dd, J = 3.6 and 12.4Hz, 6-H $_b$), 5.04-5.08 (1H, m, 5-H), 5.28 (1H, d, J = 4.0Hz, 2-H), 5.50 (1H, t, J = 5.6Hz, 4-H) and 5.61 (1H, t, J = 4.4Hz, 3-H); δ_C (100MHz; $CDCl_3$) 20.3 (6-CO_2CH_3), 20.4 (4-CO_2CH_3), 20.6 (3- and 5-CO_2CH_3), 20.7 (2-CO_2CH_3), 61.4 (C-6), 68.4 (C-3 and C-5), 69.3 (C-4), 70.0 (C-2), 169.7-170.1 (5 x CO_2CH_3) and 170.6 (COCl).

Method B[145]

2,3,4,5,6-Penta-*O*-acetyl-D-gluconic acid **115** (1.002g, 2.569mmol) was vigorously stirred in an excess of oxalyl chloride (1.5ml) at room temperature until a precipitate was formed. The excess oxalyl chloride was removed *in vacuo* and pure 2,3,4,5,6-penta-*O*-acetyl-D-gluconyl chloride **112** was isolated as white crystals (1.084g, 99%).

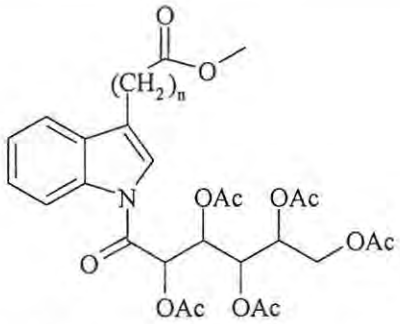
2,3,4,6-tetra-*O*-acetyl-D-gluconic acid monohydrate 114[144]

To a vigorously stirred solution of anhydrous zinc chloride (10.07g, 73.89mmol) in acetic acid (125ml) cooled in an ice bath, D-glucono-δ-lactone **113** (24.99, 140.3mmol) was added slowly and the temperature kept below 10^{O}C. The mixture was left in the ice for an hour and stirred at room temperature for 24h. The mixture was poured into water (500ml) and stirred for 1h to effect complete hydrolysis of the acetic anhydride. It was then placed in a refrigerator until the product re-crystallized completely. The crude product was filtered and washed with ice water and dried *in vacuo* to afford as white crystals, 2,3,4,6-tetra-*O*-acetyl-D-gluconic acid monohydrate **114** (12.4g, 23.2%), mp. 114-117^{O}C (lit.[144] 113-117^{O}C); ν_{max} (thin film)/cm^{-1} 3448.8 (br, OH) and 1750.5 (C=O); δ_H (400MHz; $CDCl_3$) 2.07 (3H, s, 2-CO_2CH_3), 2.08 (3H, s, 4-CO_2CH_3), 2.11 (3H, s, 6-CO_2CH_3), 2.16

(3H, s, 3-CO_2CH_3), 4.25 (1H, dd, J = 2.4 and 12.8Hz, 6-H_a), 4.39 (1H, dd, J = 3.6 and 12.8Hz, 6-H_b), 4.60 (1H, td, J = 3.0 and 8.6Hz, 5-H), 5.11 (1H, d, J = 8.8Hz, 2-H), 5.35 (1H, t, J = 8.8Hz, 4-H) and 5.54 (1H, t, J = 8.8Hz, 3-H); δ_C (100MHz; $CDCl_3$) 20.3 (3-CO_2CH_3), 20.5-20.6 (6-, 4- and 2-CO_2CH_3), 61.3 (C-6), 66.6 (C-4), 70.3 (C-3), 70.4 (C-2), 75.8 (C-5), 164.5 (C-1), 169.1 (3-CO_2CH_3), 169.5 (4-CO_2CH_3), 169.9 (2-CO_2CH_3) and 170.3 (6-CO_2CH_3).

2,3,4,5,6-Penta-*O*-acetyl-D-gluconic acid 115[144]

To a solution of zinc chloride (2.267g, 16.63mmol) in acetic anhydride (25.92g, 253.9mmol) cooled to 0°C, 2,3,4,6-tetra-*O*-acetyl-D-gluconic acid monohydrate **114** (6.284g, 16.44mmol) was added slowly and the mixture was stirred for 1h in ice and for 24h at room temperature. The mixture was diluted with water (130ml) and extracted with $CHCl_3$ (4 x 13ml). Half of the solvent was removed by distillation and toluene (33ml) was added. A further 33ml of the mixture was distilled and more toluene (33ml) was added. The remaining solution was distilled to half its volume and then left at 0°C for 3days to crystallize. The resulting white crystals were filtered and washed with cold toluene followed petroleum ether (bp. 40-60°C) to afford 2,3,4,5,6-penta-*O*-acetyl-D-gluconic acid **115** (5.84g, 93%), mp. 110-113°C (lit.[144] 110-111°C); ν_{max} (thin film)/cm^{-1} 1751.7 (C=O); δ_H (400MHz; $CDCl_3$) 2.02 and 2.04 (12H, 2 x s, CH_3), 2.16 (3H, s, CH_3), 4.10 (1H, dd, J = 5.6 and 12Hz, 6-H_a), 4.27 (1H, dd, J = 3.8 and 12Hz, 6-H_b), 5.04 (1H, dd, J = 5.6 and 10Hz, 5-H), 5.27 (1H, d, J = 4Hz, 2-H), 5.46 (1H, t, J = 5.6Hz, 4-H), 5.60 (1H, t, J = 4.4Hz, 3-H) and 9.80 (1H, s, CO_2H); δ_C (100MHz; $CDCl_3$) 20.3 ($CH_2CO_2CH_3$ and $CHCOCH_3$), 20.5, 20.6 and 20.7 (3 x CH_3), 61.4 (C-6), 68.5 (C-3 and C-5), 69.3 (C-4), 70.1 (C-2), 169.8-170.1 (5 x $COCH_3$) and 170.7 (CO_2H).



Methyl N-(2,3,4,5,6-penta-O-acetyl-D-gluconyl)-3-indolylacetate **116a**

Method A[125]

To a stirred suspension of NaH (21.9mg, 0.912mmol) in dry DMF (1.5ml) cooled in an ice bath, methyl 3-indolylacetate **110a** (130mg, 0.687mmol) in dry DMF (0.5ml) was added drop-wise and the mixture stirred at 0^{o}C for 45min. The mixture was then added to a solution of 2,3,4,5,6-penta-*O*-acetyl-D-gluconyl chloride **112** (377mg, 0.889mmol) in dry DMF (1.5ml) at 0^{o}C. The mixture was warmed to room temperature and stirred for 3 days. The solvent was removed *in vacuo* and the residue was dissolved in a 1:1 mixture of EtOAc-H_2O to afford two layers. The top, organic layer was decanted and the bottom, aqueous layer extracted with EtOAc (3 x 10ml). The organic extracts were combined and dried with anhydrous $MgSO_4$ and the solvent was evaporated *in vacuo*. The residual oil was purified by flash chromatography [silica gel, elution EtOAc-DCM (1:49)] to afford, as a brown oil, methyl N-*(2,3,4,5,6-penta*-O-*acetyl-D-gluconyl)-3-indolylacetanoate* **116a** (10.2mg, 2.5%); δ_H (400MHz; $CDCl_3$) 2.02-2.16 (15H, 3 x s, 5 x $COCH_3$), 3.71 (3H, s, OCH_3), 3.79 (2H, s, CH_2CO), 4.12 (1H, dd, J = 2.6 and Hz, 6'-H), 4.34 (1H, dd, J = 2.7 and 12.5Hz, 6'-H), 5.13 (1H, m, 5'-H), 5.52 (1H, dd, J = 3.5 and 7.4Hz, 4'-H), 5.83 (1H, dd, J = 3.6 and 6.1Hz, 3'-H), 5.92 (1H, d, J = 6.2Hz, 2'-H), 7.28-7.37 (3H, m, 2-, 5- and 6-H), 7.52 (1H, d, J = 7.6Hz, 7-H) and 8.39 (1H, d, J = 8.0Hz, 4-H).

Method B

To a suspension of NaH (32.2mg, 1.34mmol) in dry THF (5ml) was added drop-wise a solution of methyl 3-indolylacetate **110a** (0.205g, 1.08mmol) in dry THF (1ml) and the solution stirred for 1h at room temperature. 2,3,4,5,6-Penta-*O*-acetyl-D-gluconyl chloride **112** (0.459g, 1.08mmol) was added in one portion and the reaction mixture stirred for

further 24h at room temperature before adding H_2O (1ml) and removing the THF *in vacuo*. The aqueous phase was extracted with EtOAc (4 x 2ml) and the organic extracts were dried over anhydrous $MgSO_4$ and the solvent removed *in vacuo*. $CHCl_3$ was added to the residual oil and the resulting precipitate filtered off and the filtrate evaporated to dryness. Both the residual oil and the solid material were analysed by ^1H NMR spectroscopy and shown to be impure 2,3,4,5,6-penta-*O*-acetyl-D-gluconic acid **115** and 3-IAA **105**, respectively.

Method C

To a solution of methyl 3-indolylacetate **110a** (306mg, 1.62mmol) in dry pyridine (2ml) at room temperature was slowly added NaH (39.6mg, 1.65mmol). To this solution was added 2,3,4,5,6-penta-*O*-acetyl-D-gluconyl chloride **112** (684mg, 1.61mmol) and the reaction mixture was stirred at room temperature for 24h. The reaction was quenched with water and the solvent removed *in vacuo*. The residual oil was re-dissolved in $CHCl_3$ (10ml), washed with saturated aq. $NaHCO_3$ (2 x 5ml) and brine (5ml), and the aqueous phase re-extracted with $CHCl_3$. The combined organic layers were dried with anhydrous Na_2SO_4 and the solvent was removed *in vacuo*. The residual oil was purified by flash column [on silica gel; elution with EtOAc-hexane (3:2)] but none of the isolated fractions showed ^1H NMR signals corresponding to the desired product. Instead pure gluconic acid **115** and 3-IAA **105** were isolated.

***Methyl* N-*(2,3,4,5,6-penta*-O-*acetyl-D-gluconyl)- 3-indolylpropanoate* 116b**

Method A[125]

The experimental procedure employed for the synthesis of methyl *N*-(2,3,4,5,6-penta-*O*-acetyl-D-gluconyl)-3-indolylacetate **116a** was followed using, NaH (27.2mg, 1.13mmol), methyl 3-indolylpropanoate **110b** (150mg, 0.739mmol), 2,3,4,5,6-penta-*O*-acetyl-D-gluconyl chloride (477mg, 1.11mmol) and dry DMF (4.5ml). The residual oil was purified by flash chromatography [on silica gel; elution with EtOAc-DCM (1:49)] to afford, as a brown oil, *methyl* N-*(2,3,4,5,6-penta*-O-*acetyl-D-gluconyl)-3-indolylpropanoate* **116b** (9.7mg, 2.2%); δ_H (400MHz; $CDCl_3$) 2.02-2.16 (15H, 3 x s, 5 x

$COCH_3$), 2.75 (2H, t, J = 7.2Hz, CH_2CH_2CO), 3.14 (2H, t, J = 7.2Hz, CH_2CH_2CO), 3.66 (3H, s, OCH_3), 4.12 (1H, dd, J = 2.6 and 12.5Hz, 6'-H), 4.34 (1H, dd, J = 2.7 and 12.5Hz, 6'-H), 5.13 (1H, m, 5'-H), 5.52 (1H, dd, J = 3.5 and 7.4Hz, 4'-H), 5.83 (1H, dd, J = 3.6 and 6.1Hz, 3'-H), 5.92 (1H, d, J = 6.2Hz, 2'-H), 7.28-7.37 (3H, m, 2-, 5- and 6-H), 7.52 (1H, d, J = 7.6Hz, 7-H) and 8.39 (1H, d, J = 8.0Hz, 4-H).

***Methyl* N-*(2,3,4,5,6-penta*-O-*acetyl-D-gluconyl)-3-indolylbutanoate* 116c**

Method A[125]

The experimental procedure employed for the synthesis of methyl *N*-(2,3,4,5,6-penta-*O*-acetyl-D-gluconyl)-3-indolylacetanoate **116a** was employed using, NaH (21.9mg, 0.912mmol), methyl 3-indolylbutanoate **110c** (130mg, 0.599mmol), 2,3,4,5,6-penta-*O*-acetyl-D-gluconyl chloride **112** (377mg, 0.889mmol) and dry DMF (3.5ml). The residual oil was purified by flash chromatography [on silica gel; elution with EtOAc-DCM (1:49)] and HPLC [on partisil 10; elution with EtOAc-hexane (1:9)] to afford as a golden brown oil methyl N-*(2,3,4,5,6-penta-O-acetyl-D-gluconyl)-3-indolylbutanoate* **116c** (14.4mg, 3.4%); $\nu_{max}/(cm^{-1})$ 1733.4, 1695.6 and 1652.5 (C=O); δ_H (400MHz; $CDCl_3$) 2.02-2.16 (17H, m, 5 x $COCH_3$ and $CH_2CH_2CH_2CO$) 2.41 (2H, t, J = 7.2Hz, $CH_2CH_2CH_2CO$), 2.73 (2H, t, J = 7.2Hz, $CH_2CH_2CH_2CO$), 3.66 (3H, s, OCH_3), 4.12 (1H, dd, J = 2.6 and 12.5Hz, 6'-H_a), 4.34 (1H, dd, J = 2.7 and 12.5Hz, 6'-H_b), 5.13 (1H, m, 5'-H), 5.52 (1H, dd, J = 3.5 and 7.4Hz, 4'-H), 5.83 (1H, dd, J = 3.6 and 6.1Hz, 3'-H), 5.92 (1H, d, J = 6.2Hz, 2'-H), 7.28-7.37 (3H, m, 2-, 5- and 6-H), 7.52 (1H, d, J = 7.6Hz, 7-H) and 8.39 (1H, d, J = 8.0Hz, 4-H); δ_C (100MHz; $CDCl_3$) 20.4-20.8 (4 x C, $COCH_3$), 24.1 ($CH_2CH_2CH_2CO$), 24.3 ($CH_2CH_2CH_2CO$), 33.1 ($CH_2CH_2CH_2CO$), 51.5 (OCH_3), 61.8 (C-6'), 68.6 (C-3' and C-5'), 68.9 (C-4'), 70.1 (C-2'), 116.9 (C-4), 119.0 (C-7), 120.4 (C-2), 123.6 (C-3), 124.3 (C-6), 125.7 (C-5), 130.6 (C-3a), 136.4 (C-7a), 164.1 (NCO), 169.5 (3'-CO_2CH_3), 169.6 (4'-CO_2CH_3), 169.7 (2'-CO_2CH_3), 169.9 (5'-CO_2CH_3), 170.7 (C-6'CO_2CH_3) and 173.6 ($CH_2CO_2CH_3$); *m/z* 605 (M^+, 26.5%) and 218 (100%).

Method B

The experimental procedure employed for the attempted synthesis of methyl *N*-(2,3,4,5,6-penta-*O*-acetyl-D-gluconyl)-3-indolylacetanoate **116a** (**method C**) was followed using, methyl 3-indolylbutanoate **110c** (300mg, 1.40mmol), 2,3,4,5,6-penta-*O*-acetyl-D-gluconyl chloride **112** (594mg, 1.40mmol), NaH (27mg, 1.12mmol) and dry pyridine (1.5ml). The residual oil was purified by flash chromatography, and the ^{1}H NMR spectra of the isolated fractions showed none of the peaks corresponding to the expected product; the ester **110c** and impure gluconic acid **115** were isolated instead.

Attempted synthesis of N-*(2,3,4,5,6-penta*-O-*acetyl-D-gluconyl)-3-indolylpropanoic acid* 117a

To a suspension of NaH (0.128g 5.32mmol) in dry THF (20ml), cooled to 0^{O}C, 3-IPA **106** (0.498g, 2.63mmol) in THF (1ml) was added drop-wise, and the reaction mixture was stirred for a further 30min at 0^{O}C. 2,3,4,5,6-Penta-*O*-acetyl-D-gluconyl chloride **112** (1.112g, 2.62mmol) was then added and the reaction mixture was stirred for 24h at room temperature. The solvent was removed *in vacuo* and the oily residue was re- dissolved in EtOAc (10 ml) and washed with brine (2 x 2ml). The organic layer was dried over anhydrous Na_2SO_4 and the solvent removed *in vacuo*. The residual oil was shown by ^{1}H NMR analysis to be 3-IPA **106**. The aqueous layer was acidified with 5% aq. HCl and re-extracted with EtOAc to afford 3-IPA **106** and impure gluconic acid **115**.

Attempted synthesis of tert-*butyldimethylsilyl* N-*(2,3,4,5,6-penta*-O-*acetyl-D-gluconyl)-3-indolylacetate* 117b

A solution of *tert*-butyldimethylsilyl 3-indolylacetate **111a** (99.1mg, 0.343mmol) in THF (1ml) was added drop-wise to a suspension of NaH (19.5mg, 0.812) in dry THF (5m) under a stream of nitrogen. After stirring for 1h at room temperature, 2,3,4,5,6-penta-*O*-acetyl-D-gluconyl chloride **112** (153mg, 0.361mmol) was added and the resulting mixture stirred for 24h. The reaction was quenched with water (1ml) and the solvent was evaporated *in vacuo*. The residual aqueous phase was extracted with EtOAc (4 x 2ml)

and the organic extracts were dried with anhydrous $MgSO_4$. The solvent was evaporated *in vacuo* and the residual oil was analysed by ^1H NMR spectroscopy, which revealed none of the signals expected for the desired product.

Reaction of 3-indolylpropionic acid 106 with acryloyl chloride to afford *butyl 3-indolylpropanoate* 121

1.6M BuLi in hexane (3.30ml, 5.31mmol) was added slowly to as solution of 3-indolylpropionic acid **106** (0.525g, 2.78mmol) in THF (15ml) cooled to –78°C under nitrogen. After stirring for 5min, acryloyl chloride **118** (0.223g, 2.46mmol) was added and the mixture was stirred for a further 15min at –78°C. The mixture was allowed to warm gradually and stirred for 15min at –50°C and then for 15min at –20°C before quenching with saturated aq. NH_4Cl. The solvent was removed *in vacuo* and the residual aqueous phase was extracted with EtOAc (3 x 10ml). The organic phase was washed with brine, dried with anhydrous $MgSO_4$ and evaporated *in vacuo*. The residual oil was purified by flash chromatography [on silica gel; elution with EtOAc-hexane (1:2)] to afford, as a yellow oil, *butyl 3-indolylpropanoate* **121** (12.8mg, 1.8%); δ_H (400MHz, $CDCl_3$) 0.91 (3H, t, J = 7.4Hz, $OCH_2CH_2CH_2CH_3$), 1.34 (2H, sextet, J = 7.4Hz, $OCH_2CH_2CH_2CH_3$), 1.58 (2H, quintet, J = 7.3Hz, $OCH_2CH_2CH_2CH_3$), 2.71 (2H, t, J = 7.6Hz, CH_2CH_2CO), 3.10 (2H, t, J = 7.6Hz, CH_2CH_2CO), 4.08 (2H, t, J = 7.4Hz, $OCH_2CH_2CH_2CH_3$), 7.00 (1H, s, 2-H), 7.12 (1H, t, J = 7.4Hz, 5-H), 7.19 (1H, t, J = 7.4Hz, 6-H), 7.34 (1H, d, J = 8.0Hz, 4-H), 7.60 (1H, d, J = 8.0Hz, 7-H) and 7.98 (1H, s, NH); δ_C 13.7 ($OCH_2CH_2CH_2CH_3$), 19.1 ($OCH_2CH_2CH_2CH_3$), 20.7 (CH_2CH_2CO), 30.7 ($OCH_2CH_2CH_2CH_3$), 35.0 (CH_2CH_2CO), 64.3 ($OCH_2CH_2CH_2CH_3$), 111.1 (C-4), 115.1 (C-3), 118.7 (C-7), 119.3 (C-5), 121.4 (C-2), 122.0 (C-6), 127.2 (C-3a), 136.3 (C-7a) and 173.5 (C=O).

Attempted synthesis of methyl 1-acryloyl-3-indolylacetate 122[147]

Acryloyl chloride **118** (0.167g, 1.85mmol) was added in one portion to a vigorously stirred mixture of methyl 3-indolylacetate **110a** (0.227g, 1.20mmol) and NaOH (56.2mg,

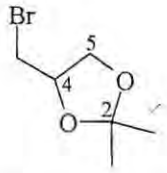
1.40mmol) in DCM (2ml), at room temperature. The mixture was vigorously stirred for a further 1h and then washed with water. The organic phase was dried with anhydrous Na_2SO_4 and the solvent evaporated *in vacuo*. The residual brown oil was purified by flash chromatography [on silica gel; elution with EtOAc-hexane (1:2)] to afford methyl 3-indolylacetate **110a** and an unidentified acryloyl adduct.

Attempted synthesis of *tert*-butyldimethylsilyl 1-acryloyl-3-indolylpropanoate 123

NaH (52.8mg, 2.20mmol) was added slowly (to control evolution of H_2) to a solution of *tert*-butyldimethylsilyl 3-indolylproppanoate **111b** (0.641g, 2.1mmol) in dry THF (10ml), cooled under N_2 to $0^{O}C$. After stirring for 15min, acryloyl chloride (0.209g, 2.31mmol) was added and stirring continued at room temperature for 24h. The reaction was quenched with H_2O (5ml) and the solvent evaporated *in vacuo*. The residual aqueous phase was extracted with EtOAc (3 x 10ml) and dried over anhydrous $MgSO_4$, and the solvent evaporated *in vacuo*. The residual oil was purified by flash chromatography [on silica gel; elution with EtOAc-DCM (2:3)] and then by HPLC [on partisil 10; elution with EtOAc-hexane (1:1)] to afford 3-indolylpropionic acid **106** and *tert*-butyldimethylsilyl alcohol.

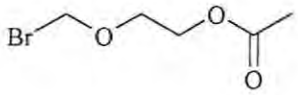
Synthesis of *methyl 1-butanoyl-3-indolylacetate* 125

The experimental procedure employed for the synthesis of methyl *N*-(2,3,4,5,6-penta-*O*-acetyl-D-gluconyl)-3-indolylacetate **116a** was followed using, methyl 3-indolylacetate **110a** (0.208g, 1.10mmol), butanoyl chloride (0.133g, 1.21mmol), NaH (28.9mg, 1.20mmol) and DMF (2.2ml). The residual oil was purified by flash chromatography [on silica gel; elution with EtOAc-DCM (1:49)] to afford, as a dark brown oil, methyl 1-butanoyl-3-indolylacetate **125** (6mg, 2.1%); δ_H (400MHz; $CDCl_3$) 1.06 (3H, t, J = 7.4Hz, $COCH_2CH_2CH_3$), 1.86 (2H, sextet, J = 7.4Hz, $COCH_2CH_2CH_3$), 2.88 (2H, t, J = 7.4Hz, $COCH_2CH_2CH_3$), 3.72 (3H, s, CH_2CO and OCH_3), 7.11-7.21 (3H, m, 2-, 5- and 6-H), 7.52 (2H, d, J = 8.0Hz, 4-H) and 8.45 (1H, d, J = 8.0Hz, 7-H).



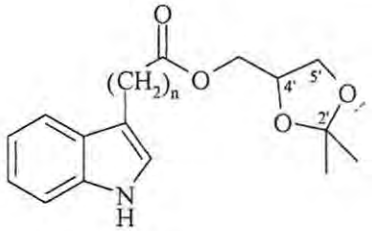
4-Bromomethyl-2,2-dimethyl-1,3-dioxolane 128[152]

To a solution of 3-bromopropane-1,2-diol **126** (8.855g, 57.1mmol) in dry acetone (60ml) was added H_2SO_4 (0.15ml), and the resulting solution was stirred for 36h at room temperature. The solution was then treated with solid K_2CO_3 (6.325g, 45.77mmol), stirred for 30min and filtered. The filtrate was evaporated and the residue was re-dissolved in diethyl ether (50ml) and washed with H_2O (3 x 50ml). The organic layer was dried over anhydrous Na_2SO_4 and evaporated *in vacuo* to afford, as a colourless oil, 4-bromomethyl-2,2-dimethyl-1,3-dioxolane **128** (7.762g, 69.7%). δ_H (400MHz; $CDCl_3$) 1.34 and 1.43 (6H, 2 x s, 2 x CH_3), 3.29 (1H, dd, J = 8.2 and 10.0Hz, $BrCH_a$) 3.42 (1H, dd, J = 4.7 and 10.0Hz, $BrCH_b$), 3.85 (1H, dd, J = 5.2 and 8.8Hz, 5-H_a), 4.11 (1H, dd, J = 6.2 and 8.6Hz, 5-H_b) and 4.33 (1H, ddd, J = 5.1, 8.2 and 11.1Hz, 4-H); δ_C (100MHz; $CDCl_3$) 25.4 and 27.0 (2 x CH_3), 32.7 (CH_2Br), 68.3 (C-5), 75.3 (C-4) and 110.3 (C-2)



2-(Bromomethoxy)ethyl acetate **131**[155]

1,3-Dioxolane **129** (8.48g, 0.115mol) was added drop-wise to freshly distilled acetyl bromide **130** (14.025g, 0.114mol) at 0^OC and the mixture was stirred for 2 min. The crude material was purified by distillation to afford, as a colourless oil 2-(bromomethoxy)ethyl acetate **131** (17.83g, 79.3%), bp. 57-59^OC/0.09mmHg (lit.[155] 58-60^OC/0.1mmHg); δ_H (400MHz; $CDCl_3$) 2.07 (3H, s, CH_3), 3.84 and 4.26 (4H, 2 x t, J = 4.6Hz, 2 x CH_2), 5.68 (2H, s, $BrCH_2$); δ_C (100MHz; $CDCl_3$) 20.8 (CH_3), 62.1 and 69.2 (OCH_2CH_2O), 75.4 ($BrCH_2O$) and 170.8 (C=O).



(2,2-Dimethyl-1,3-dioxol-4yl)methyl-2-(1*H*-indol-3-yl)acetate 134a[204,205]

Method A[157]

A solution of methyl 3-indolylacetate **110a** (0.408g, 2.16mmol) and potassium *tert*-butoxide ($KOBu^t$) (0.245g, 2.18mmol) in dry DMSO (4ml) was stirred N_2 for 1h at room temperature. A solution of KI (0.023g, 0.14mmol) in 4-bromomethyl-2,2-dimethyl-1,3-dioxolane **128** (1.251g, 6.417mmol) was then added. The reaction mixture was stirred for a further 4 days at room temperature and added to a mixture of EtOAc (9ml) and a half saturated solution of NH_4Cl (3ml). After extraction with EtOAc (3 x 9ml) the organic solutions were combined and washed with brine (2 x 5ml), and the aqueous phase was re-extracted with EtOAc (2 x 9ml). The combined organic layers were dried with anhydrous $MgSO_4$ and evaporated *in vacuo*. The residual oil was purified by radial chromatography [on silica gel; elution with EtOAc-hexane (2:3)] and by HPLC [on Partisil 10; elution with EtOAc-hexane (1:3)] to afford, as a golden oil, (2,2-dimethyl-1,3-dioxol-4-yl)methyl-2-(1*H*-indol-3-yl)acetate **134a** (0.069g, 11.1%); ν_{max} (thin film)/cm^{-1} 3405.1 (NH) and 1732.2 (C=O); δ_H (400MHz; $CDCl_3$) 1.35 and 1.40 (6H, 2 x s, 2 x CH_3), 2.71 (2H, t, J = 7.6Hz, CH_2CH_2CO), 3.10 (2H, t, J = 7.6Hz, CH_2CH_2CO), 3.68 (1H, dd, J = 6.2 and 8.6Hz, 5'-H), 4.00 (1H, dd, J = 6.2 and 8.4Hz, 5'-H), 4.12-4.20 (2H, m, OCH_2), 4.33 (1H, quintet, J = 6.0Hz, 4'-H), 7.11-7.21 (3H, m, 2-, 5- and 6-H), 7.35 (1H, d, J = 8.0Hz, 4-H), 7.60 (1H,d, J = 7.6 Hz, 7-H) and 8.10 (1H, s, NH); δ_C (100MHz; $CDCl_3$) 25.4 and 26.6 (2 x CH_3), 31.1 (CH_2CO), 64.9 (OCH_2), 66.3 (C-5'), 73.6 (C-4'), 108.3 (C-3), 109.8 (C-2'), 111.2 (C-4), 118.8 (C-7), 119.7 and 122.3 (C-5 and C-6), 123.0 (C-2), 127.2 (C-3a), 136.1 (C-7a) and 171.7 (C=O); *m/z* 289 (M^+ 79.9%) and 157 (100%).

Method B

The experimental procedure employed in **method A** was followed using, 3-indolylacetic acid **105** (0.503g, 2.87mmol), 4-bromomethyl-2,2-dimethyl-1,3-dioxolane **128** (1.671g, 8.571mmol), $KOBu^t$ (0.322g, 2.87mmol), KI (23.8mg, 0.144mmol) and DMSO (8ml). The residual oil was purified by radial chromatography [on silica gel; elution with EtOAc-hexane (2:3)] to afford, as a golden oil, (2,2-dimethyl-1,3-dioxol-4-yl)methyl-2-(1*H*-indol-3-yl)acetate **134a** (67.5mg, 8.13%).

Method C[13]

To a suspension of NaH (48.2mg, 2.01mmol) in dry DMF (1.8ml) under a stream of nitrogen at $0^{O}C$ was added a solution of methyl 3-indolylacetate **110a** (0.300g, 1.59mmol) in DMF (1.8ml), and the reaction mixture was stirred for 1h at room temperature. The reaction mixture was cooled to $0^{O}C$ and 4-bromomethyl-2,2-dimethyl-1,3-dioxolane **128** (0.374g, 1.92mmol) in DMF (3.5ml) was added drop-wise and the reaction mixture stirred overnight at room temperature. The reaction mixture was then diluted with EtOAc (10ml), and excess NaH quenched with saturated aq. NH_4Cl and extracted with EtOAc (3 x 10ml). The organic phase was washed with saturated brine, dried over anhydrous $MgSO_4$ and evaporated *in vacuo*. The residual oil was purified by flash chromatography [on silica gel; elution with EtOAc-DCM (1:2)] to afford, as a golden oil, (2,2-dimethyl-1,3-dioxol-4-yl)methyl-2-(1*H*-indol-3-yl)acetate **134a** (0.9mg, 0.2%).

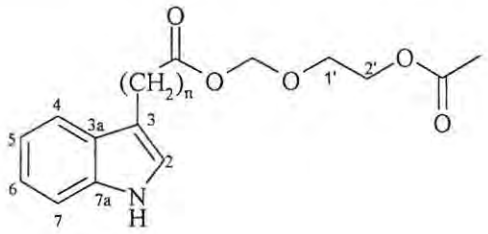
(2,2-Dimethyl-1,3-dioxol-4-yl)methyl-3-(1*H*-indol-3-yl)propanoate 134b[204,205]

The experimental procedure employed in **method C** for the synthesis of (2,2-dimethyl-1,3-dioxol-4-yl)methyl-2-(1*H*-indol-3-yl)acetate **134a** was followed, using methyl 3-indolylpropanoate **110b** (0.306g, 1.50mmol), compound **128** (0.298, 1.52mmol), NaH (35.6mg, 1.48mmol). The residual oil was purified by flash chromatography [on silica gel; elution with EtOAc-DCM (1:49)] to afford, as a brown oil, (2,2-dimethyl-1,3-dioxol-4-yl)methyl-3-(1*H*-indol-3-yl)propanoate **134b** (5.3mg, 1.2%), ν_{max} (thin film)/cm^{-1} 3405.1 (NH) and 1733.2 (C=O); δ_H (400MHz; $CDCl_3$) 1.35 and 1.40 (6H, 2 x s, 2 x

CH_3), 3.68 (1H, dd, J = 6.2 and 8.6Hz, 5'-H), 3.82 (2H, s, CH_2CO), 4.00 (1H, dd, J = 6.2 and 8.4Hz, 5'-H), 4.12-4.20 (2H, m, OCH_2), 4.33 (1H, quintet, J = 6.0Hz, 4'-H), 7.11-7.21 (3H, m, 2-, 5- and 6-H), 7.35 (1H, d, J = 8.0Hz, 4-H), 7.60 (1H,d, J = 7.6 Hz, 7-H) and 8.10 (1H, s, NH); δ_C (100MHz; $CDCl_3$) 20.6 (CH_2CH_2CO), 25.4 and 26.6 (CH_3), 34.7 (CH_2CH_2CO), 64.9 (OCH_2), 66.3 (C-5'), 73.6 (C-4'), 108.3 (C-3), 109.8 (C-2'), 111.2 (C-4), 118.8 (C-7), 119.7 (C-5), 122.3 (C-6), 123.0 (C-2), 127.2 (C-3a), 136.1 (C-7a) and 171.7 (C=O).

(2,2-Dimethyl-1,3-dioxol-4-yl)methyl-4-(1*H*-indol-3-yl)butanoate 134c[204,205]

The experimental procedure employed in **method A** for the synthesis of (2,2-dimethyl-1,3-dioxol-4-yl)methyl-2-(1*H*-indol-3-yl)acetate **134a** was followed, using methyl 3-indolylbutanoate **110c** (0.501g, 2.31mmol), 4-bromomethyl-2,2-dimethyl-1,3-dioxolane **127** (1.348g, 6.91mmol), $KOBu^t$ (0.260g, 2.32mmol), KI (19.1mg, 0.115mmol) and DMSO (8ml). The residual oil was purified by radial chromatography [on silica gel; elution with EtOAc-hexane (2:3)] to afford, as a yellow oil, (2,2-dimethyl-1,3-dioxol-4-yl)methyl-4-(1*H*-indol-3-yl)butanoate **134c** (76.5mg, 10.4%); ν_{max} (thin film)/cm^{-1} 3413.6 (NH) and 1734.3 (C=O); δ_H (400MHz; $CDCl_3$) 1.37 and 1.44 (6H, 2 x s, 2 x CH_3), 2.06 (2H, quintet, J = 7.4, $CH_2CH_2CH_2CO$), 2.42 (2H, t, J = 7.4Hz, $CH_2CH_2CH_2CO$), 2.81 (2H, t, J = 7.4Hz, $CH_2CH_2CH_2CO$), 3.72 (1H, dd, J = 6.0 and 8.4Hz, 5'-H), 4.04-4.16 (3H, m, OCH_2 and 5'-H), 6.95 (1H, s, 2-H), 7.11 (1H, t, J = 7.4Hz, 6-H), 7.18 (1H, t, J = 7.4Hz, 5-H), 7.33 (1H, d, J = 8.0Hz, 4-H), 7.60 (1H, d, J = 7.8Hz, 7-H) and 8.12 (1H, s, NH); δ_C (100MHz; $CDCl_3$) 24.4 ($CH_2CH_2CH_2CO$), 25.2 ($CH_2CH_2CH_2CO$), 25.3 and 26.6 (2 x CH_3), 33.6 ($CH_2CH_2CH_2CO$), 64.5 (OCH_2), 66.2 (C-5'), 73.5 (C-4'), 109.8 (C-2'), 111.0 (C-4), 115.2 (C-3), 118.8 (C-7), 119.1 (C=6), 121.5 (C-2), 121.8 (C-5), 127.3 (C-3a), 136.3 (C-7a) and 173.5 (C=O).



*(2-Acetoxyethoxy)methyl –2-(1*H*-indol-3-yl)acetate* **135a**

The experimental procedure employed in **method C** for the synthesis of (2,2-dimethyl-1,3-dioxol-4-yl)methyl-2-(1*H*-indol-3-yl)acetate **134a** was followed, using methyl 3-indolylacetate **110a** (0.209g, 1.10mmol), DMF (5ml), NaH (34.4mg, 1.44mmol) and 2-(bromomethoxy)ethyl acetate **131** (0.283g, 1.44mmol). The residual oil was purified by radial chromatography [on silica gel; elution with EtOAc-DCM (1:49)] to afford, as brown oil, (2-acetoxyethoxy)methyl-2-(1*H*-indol-3-yl)acetate **135a** (65.9mg, 20.4%); ν_{max} (thin film)/cm^{-1} 3414 (NH) and 1734 (C=O); δ_H (400MHz; $CDCl_3$) 2.08 (3H, s, $COCH_3$), 3.68 (2H, t, J = 4.8Hz, OCH_2CH_2OAc), 3.76 (2H, s, CH_2CO), 4.24 (2H, t, J = 4.8Hz, OCH_2CH_2OAc), 4.81 (2H, s, OCH_2O), 6.97 (1H, s, 2-H), 7.10 (1H, t, J = 7.4Hz, 6-H), 7.18 (1H, t, J = 8.0Hz, 5-H), 7.33 (1H, d, J = 8.0Hz, 4-H), 7.60 (1H, d, J = 7.6Hz, 7-H) and 7.92 (1H, s, NH); δ_C (100MHz; $CDCl_3$) 20.8 (CH_3), 30.9 (CH_2CO), 63.5 (OCH_2CH_2OAc), 66.1 (OCH_2CH_2OAc), 92.0 (OCH_2O), 111.1 (C-4), 115.4 (C-3), 118.8 (C-7), 119.2 (C-5), 121.5 (C-6), 121.9 (C-2), 127.4 (C-3a), 136.4 (C-7a), 170.9 (CH_3CO) and 179.4 (CH_2CO); *m/z* 291 (M^+ 100%).

*(2-Acetoxyethoxy)methyl-3-(1*H*-indol-3-yl)propanoate* **135b**

The experimental procedure employed in **method C** for the synthesis of (2,2-dimethyl-1,3-dioxol-4-yl)methyl-2-(1*H*-indol-3-yl)acetate **134a** was followed, using methyl 3-indolylpropanoate **110b** (0.244g, 1.20mmol), DMF (5ml), NaH (40.2mg, 1.68mmol) and 2-(bromomethoxy)ethyl acetate **131** (0.308g, 1.56mmol). The residual oil was purified by radial chromatography [on silica gel; elution with EtOAc-DCM (1:49)] to afford, as a brown oil, *(2-acetoxyethoxy)methyl-3-(1*H*-indol-3-yl)propanoate* **135b** (0.0297g, 8.2%);

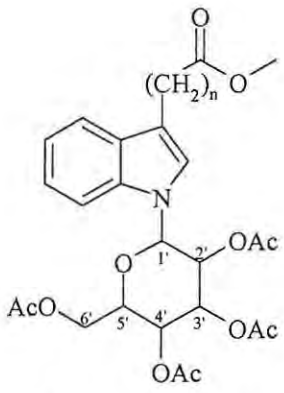
ν_{max} (thin film)/cm^{-1} 3414 (NH) and 1734 (C=O); δ_H (400MHz; $CDCl_3$) 2.08 (3H, s, $COCH_3$), 2.68 (2H, t, J = 7.4Hz, CH_2CO), 3.05 (2H, t, J = 7.4Hz, CH_2CH_2CO), 3.72 (2H, t, J = 4.8Hz, OCH_2CH_2OAc), 4.10 (2H, t, J = 4.8Hz, OCH_2CH_2OAc), 4.71 (2H, s, OCH_2O), 6.97 (1H, s, 2-H), 7.10 (1H, t, J = 7.4Hz, 6-H), 7.18 (1H, t, J = 8.0Hz, 5-H), 7.30 (1H, d, J = 8.0Hz, 4-H), 7.60 (1H, d, J = 7.6Hz, 7-H) and 8.00 (1H, s, NH); δ_C (100MHz; $CDCl_3$) 20.2 (CH_2CH_2CO), 20.4(CH_3), 34.4 (CH_2CH_2CO), 60.0 (OCH_2CH_2OAc), 65.6 (OCH_2CH_2OAc), 92.0 (OCH_2O), 110.9 (C-4), 113.8 (C-3), 118.0 (C-7), 118.3 (C-5), 121.0 (C-6), 121.4 (C-2), 126.7 (C-3a), 136.1 (C-7a), 171.0 (CH_3CO) and 175.7 (CH_2CO).

***(2-Acetoxyethoxy)methyl-4-(1H-indol-3-yl)butanoate* 135c**

The experimental procedure employed in **method C** for the synthesis of (2,2-dimethyl-1,3-dioxol-4-yl)methyl-2-(1*H*-indol-3-yl)acetate **134a** was followed using, methyl 4-(3-indolyl)butanoate **110c** (0.301g, 1.39mmol), 2-(bromomethoxy)ethyl acetate **131** (0.361g, 1.83mmol), NaH (46.0mg, 1.92mmol) and DMF (6.4ml). The crude product was purified by flash chromatography [on silica gel; elution with DCM-EtOAc (2:1)] to afford, as a brown oil, *(2-acetoxyethoxy)methyl-4-(1H-indol-3-yl)butanoate* **135c** (0.0341g, 7.7%); ν_{max} (thin film)/cm^{-1} 3414 (NH) and 1734 (C=O); δ_H (400MHz; $CDCl_3$) 2.05 (2H, quintet, J = 7.4Hz, $CH_2CH_2CH_2CO$), 2.08 (3H, s, $COCH_3$), 2.42 (2H, t, J = 7.4Hz, $CH_2CH_2CH_2CO$), 2.83 (2H, t, J = 7.4Hz, $CH_2CH_2CH_2CO$), 3.78 (2H, t, J = 4.8Hz, OCH_2CH_2OAc), 4.24 (2H, t, J = 4.8Hz, OCH_2CH_2OAc), 4.82 (2H, s, OCH_2O), 6.97 (1H, s, 2-H), 7.10 (1H, t, J = 7.4Hz, 6-H), 7.18 (1H, t, J = 8.0Hz, 5-H), 7.33 (1H, d, J = 8.0Hz, 4-H), 7.60 (1H, d, J = 7.6Hz, 7-H) and 7.99 (1H, s, NH); δ_C (100MHz; $CDCl_3$) 20.8 (CH_3), 24.4 ($CH_2CH_2CH_2CO$), 25.0 ($CH_2CH_2CH_2CO$), 33.5 ($CH_2CH_2CH_2CO$), 63.5 (OCH_2CH_2OAc), 66.1 (OCH_2CH_2OAc), 92.0 (OCH_2O), 111.1 (C-4), 115.4 (C-3), 118.8 (C-7), 119.2 (C-5), 121.5 (C-6), 121.9 (C-2), 127.4 (C-3a), 136.4 (C-7a), 171.0 (CH_3CO) and 179.4 (CH_2CO).

1,2,3,4,6-penta-*O*-acetyl-α-D-glucopyranose 137[159]

To a vigorously stirred solution of anhydrous zinc chloride (2.154g, 15.81mmol) in acetic anhydride (55.08g, 0.5345mol), α-D-glucopyranose **136** (10.794g, 59.9mmol) was added and the solution was refluxed for 1h. The solution was cooled to room temperature and added to an ice-water mixture (500ml) and vigorously stirred for 30min. The resulting fine precipitate was filtered off and recrystallized from petroleum ether to afford, as fluffy white crystals, 1,2,3,4,6-penta-*O*-acetyl-D-glucopyranose **137** (19.896g, 85.2%), mp. 108-110°C (lit.[159] 109-111°C); ν_{max} (thin film)/cm^{-1} 1750.6 (C=O); δ_H (400MHz; $CDCl_3$) 2.00-2.02 (9H, 3 x s, CHOCOC*H*$_3$), 2.07 (3H, s, CH_2OCOCH_3), 2.16 (3H, s, CHOCOC*H*$_3$), 4.06-4.12 (2H, m, 5- and 6-H), 4.25 (1H, dd, J = 4.2 and 12.6Hz, 6-H), 5.08 (1H, dd, J = 3.8 and 10.2Hz, 4-H), 5.12 (1H, t, J = 10Hz, H), 5.45 (1H, t, J = 10Hz, 2-H) and 6.31 (1H, d, J = 3.6Hz, 1-H); δ_C (100MHz; $CDCl_3$) 20.4-20.8 (5 x CH_3), 61.4 (C-6), 67.9 (C-2 and C-4), 69.2 (C-3), 69.8 (C-5), 89.0 (C-1), 168.7 (1-*C*O_2CH_3), 169.4, 169.6, 170.2 and 170.6 (5 x C=O).



Methyl 2-[1-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-1*H*-indol-3-yl]acetate 138a[206,207]

$SnCl_4$ (1.78g, 6.84mmol) was added *via* a syringe to a solution of methyl 3-indolylacetate **110a** (0.518g, 2.74mmol) and 2,3,4,5,6-penta-O-acetyl-D-glucopyranose **137** (1.032g, 2.65mmol) in dry acetonitrile (26ml) cooled at 0°C under argon. After 12h of heating at 70°C, the reaction was quenched with saturated aq. $NaHCO_3$ (45ml) and extracted with ethyl acetate (3 x 45ml). The organic layer was dried over anhydrous Na_2SO_4 and

evaporated *in vacuo*. The residual oil was purified by flash chromatography [on silica gel; elution with EtOAc-hexane (3:2)] and by HPLC [on Partisil 10; elution with EtOAc-Hexane (2:3)] to afford, as a pale yellow oil, methyl 2-[1-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-1*H*-indol-3-yl]acetate **138a** (0.3532g, 25.8%); (Found: M^+, 519.17490. $C_{25}H_{29}NO_{11}$ requires *M*, 519.17489); ν_{max} (thin film)/cm^{-1} 1750.8 (C=O); δ_H (400MHz; $CDCl_3$) 1.67 (3H, s, 2'-CO_2CH_3), 2.01 (3H, s, 3'-CO_2CH_3), 2.07 (6H, s, 6'- and 4'-CO_2CH_3), 3.69 (3H, s, OCH_3), 3.72 (2H, s, CH_2), 3.69 (3H, s, OCH_3), 3.73 (2H, s, CH_2CO), 3.96 (1H, ddd, *J* = 2.2, 4.9 and 9.9Hz, 5'-H), 4.14 (1H, dd, *J* = 2.0 and 12.4Hz, 6'CH_a), 4.28 (1H, dd, *J* = 4.8 and 12.4Hz, 6'CH_b), 5.26 (1H, t, *J* = 9.8Hz, 4'-H), 5.42 (1H, t, *J* = 9.2Hz, 3'-H), 5.52 (1H, t, *J* = 9.2Hz, 2'-H), 5.59 (1H, d, *J* = 9.2Hz, 1'-H), 7.16 (1H, t, *J* = 7.2Hz, 6-H), 7.20 (1H, s, 2-H), 7.23-7.26 (1H, m, 5-H), 7.38 (1H, d, *J* = 8.0Hz, 4-H) and 7.56 (1H, d, *J* = 8.0Hz, 7-H); δ_C (100MHz; $CDCl_3$) 20.1 (2'-CO_2CH_3), 20.6-20.7 (3'-, 4'- and 6'-CO_2CH_3), 31.0 (CH_2CO), 52.0 (OCH_3), 61.9 (C-6'), 68.2 (C-4'), 70.3 (C-2'), 73.4 (C-3'), 74.6 (C-5'), 83.3 (C-1'), 109.8 (C-4), 110.2 (C-3), 119.4 (C-7), 120.7 (C-6), 122.7 (C-5), 123.3 (C-2), 128.5 (C-3a), 136.3 (C-7a), 168.7 (C-2'CO_2CH_3), 169.4 (4'-CO_2CH_3), 170.2 (3'-CO_2CH_3), 170.6 (6'-CO_2CH_3) and 171.9 (C=O); *m/z* 519 (M^+, 62%) and 169 (100%).

***Methyl 3-[1-(2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl)-1H-indol-3-yl]propanoate* 138b**

The experimental procedure employed for the synthesis of *methyl* 2-[1-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-1*H*-indol-3-yl]acetate **138a** was followed, using methyl 3-indolylpropanoate **110b** (0.166g, 0.735mmol), 2,3,4,5,6-penta-*O*-acetyl-D-glucopyranose **137** (0.286g, 0.734mmol), tin (IV) chloride (0.478g, 1.84mmol) and acetonitrile (7ml). The crude oil was purified by flash chromatography [on silica gel; elution with EtOAc-Hexane (3:2)] and by HPLC [on Partisil 10; elution with EtOAc-hexane (2:3)] to afford, as pale yellow oil, *methyl 3-[1-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-1*H*-indol-3-yl]propanoate* **138b** (0.0283g, 7.2%), (Found: M^+, 533.18958. $C_{26}H_{31}NO_{11}$ requires *M*, 533.18971); ν_{max} (thin film)/cm^{-1} 3446 (NH), 1756 and 1733 (C=O); δ_H (400MHz, $CDCl_3$) 1.66 (3H, s, 2'-CO_2CH_3), 2.02 (3H, s, 3'-CO_2CH_3), 2.06 (3H, s, 6'-CO_2CH_3), 2.07

(3H, s, 4'-CO_2CH_3), 2.68 (2H, t, J = 8.0Hz, CH_2CH_2CO), 3.04 (2H, t, J = 8.0Hz, CH_2CH_2CO), 3.68 (3H, s, OCH_3), 3.96 (1H, ddd, J = 2.2, 4.9 and 10.0Hz, 5'-H), 4.13 (1H, dd, J = 2.2 and 12.4Hz, 6'-H_a), 4.28 (1H, dd, J = 4.9 and 12.4Hz, 6'-H_b), 5.26 (1H, t, J = 9.6Hz, 4'-H), 5.42 (1H, t, J = 9.2Hz, 3'-H), 5.51 (1H, t, J = 9.2Hz, 2'-H), 5.58 (1H, d, J = 9.2Hz, 1'-H), 7.01 (1H, s, 2-H), 7.14 (1H, t, J = 7.4Hz, 5-H), 7.23 (1H, t, J = 7.4Hz, 6-H), 7.35 (1H, d, J = 8.0Hz, 4-H) and 7.54 (1H, d, J = 8.0Hz, 7-H); δ_C (100MHz; $CDCl_3$) 20.1 (2'-CO_2CH_3), 20.5 (CH_2CH_2CO), 20.6-20.7 (3'-, 4'- and 6'-CO_2CH_3), 34.4 (CH_2CH_2CO), 51.6 (OCH_3), 61.9 (C-6'), 68.2 (C-4'), 70.2 (C-2'), 73.5 (C-3'), 74.6 (C-5'), 83.1 (C-1'), 109.6 (C-4), 116.7 (C-3), 119.2 (C-7), 120.4 (C-5), 121.4 (C-2), 122.6 (C-6), 128.5 (C-3a), 136.6 (C-7a), 168.7 (2'-CO_2CH_3), 169.4 (4'-CO_2CH_3), 170.2 (3'-CO_2CH_3), 170.6 (6'-CO_2CH_3) and 173.5 (CO_2CH_3)

***Methyl 4-[1-(2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl)-1H-indol-3-yl]butanoate* 138c**

The experimental procedure employed for the synthesis of methyl 2-[1-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-1*H*-indol-3-yl]acetate **138a** was followed, using methyl 3-indolylbutanoate **110c** (0.5013g, 2.31mmol), 1,2,3,4,6-penta-*O*-acetyl-D-glucopyranose **137** (0.897g, 2.30mmol), tin (IV) chloride (0.8904g, 3.418mmol) and acetonitrile (23ml). The crude product was purified by flash chromatography [on silica gel; elution with EtOAc-hexane (1:2 and 2:1)] and by HPLC [on Partisil 10; elution EtOAc-hexane (2:3)] to afford, as cream crystals, *methyl 4-[1-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-1H-indol-3-yl]-butanoate* **138c** (0.1107g, 14%), mp. 136-138^{O}C, (Found: M^+, 547.20449. $C_{27}H_{33}NO_{11}$ requires M, 547.20536); ν_{max} (thin film)/cm^{-1} δ_H (400MHz; $CDCl_3$) 1.66 (3H, s, 2'-CO_2CH_3), 2.01 (5H, m, 3'-CO_2CH_3 and $CH_2CH_2CH_2CO$), 2.06 (3H, s, 6'-CO_2CH_3), 2.07 (3H, s, 4'-CO_2CH_3), 2.35 (2H, t, J = 7.4Hz, $CH_2CH_2CH_2CO$), 2.75 (2H, t, J = 7.4Hz, $CH_2CH_2CH_2CO$), 3.66 (3H, s, $CH_2CO_2CH_3$), 3.97 (1H, ddd, J = 2.2, 4.9 and 10.1Hz, 5'-H), 4.13 (1H, dd, J = 4.6 and 15.9Hz, 6'-H_a), 4.28 (1H, dd, J = 4.9 and 12.4Hz, 6'-H_b), 5.26 (1H, t, J = 9.8Hz, 4'-H), 5.42 (1H, t, J = 9.2Hz, 3'-H), 5.52 (1H, t, J = 9.2Hz, 2'-H), 5.58 (1H, d, J = 9.2Hz, 1'-H), 7.00 (1H, s, 2-H), 7.13 (1H, t, J = 7.4Hz, 5-H), 7.23 (1H, t, J = 7.4Hz, 6-H), 7.35 (1H, d, J = 8.0Hz, 4-H) and 7.52 (1H, d, J = 8.0Hz, 7-H); δ_C (100MHz, $CDCl_3$) 20.1 (2'-CO_2CH_3), 20.6 - 20.7 (3'-, 4'- and 6'-CO_2CH_3), 24.4

($CH_2CH_2CH_2CO$), 25.0 ($CH_2CH_2CH_2CO$), 33.5 ($CH_2CH_2CH_2CO$), 51.5 ($CH_2CO_2CH_3$), 62.0 (C-6'), 68.2 (C-4'), 70.3 (C-2'), 73.4 (C-3'), 74.5 (C-5'), 83.02 (C-1'), 109.5 (C-4), 117.3 (C-3), 119.4 (C-7), 120.2 (C-5), 121.5 (C-2), 122.4 (C-6), 126.7 (C-3a), 136.7 (C-7a), 168.7 (2'-CO_2CH_3), 169.4 (4'-CO_2CH_3), 170.1 (4'-CO_2CH_3), 170.6 (6'-CO_2CH_3) and 174.03 (CO_2CH_3).

2-(β-D-glucopyranosyl-1*H*-indol-3-yl)acetic acid 139a[207]

To a solution of sodium methoxide, prepared by dissolving sodium metal (1.07g) in dry methanol (30ml), was added methyl 2-[1-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-1H-indol-3-yl]acetate **138a** (77.3mg, 0.149mmol) in one portion, and the mixture was stirred at room temperature until no starting material was observed on TLC (*ca.* 20min). To this reaction mixture, activated Dowex© 50WX8 (10ml) was added and the mixture stirred overnight or until the solution became neutral. The ion-exchange resin was removed by filtration and the filtrate was evaporated to dryness. The residual powder was purified by HPLC [on a C-18 LUNA column; elution with MeOH-H_2O (2:1)] to afford, as a brown powder, 2-(β-D-glucopyranosyl-1*H*-indol-3-yl)acetic acid **139a** (39.1mg, 69.8%); mp. > 320°C; δ_H (100MHz; D_2O) 3.89 (2H, s, CH_2O), 3.92-4.37 (6H, m, 2'-, 3'-, 4'-, 5'- and 6'-H), 5.87 (1H, d, J = 9.2Hz, 1'-H), 7.47 (1H, t, J = 7.9Hz, 6-H), 7.57 (1H, t, J, = 7.9Hz, 5-H), 7.62 (1H, s, 2-H) and 7.86 (2H, d, 7.6, 4- and 7-H); δ_C (100MHz, D_2O) 34.6 (CH_2CO), 61.3 (C-6'), 70.1 (C-4'), 72.9 (C-2'), 77.2 (C-3'), 79.8 (C-5'), 84.8 (C-1'), 110.9 (C-4), 118.0 (C-3), 120.5 (C-7), 121.2 (C-5), 123.1 (C-6), 123.2 (C-2), 128.9 (C-3a), 137.4 (C-7a) and 182.0 (C=O).

***3-(β-D-glucopyranosyl-1H-indol-3-yl)propanoic acid* 139b**

The experimental procedure employed for the synthesis of the acetic acid derivative **139a** was followed, using sodium metal (0.762g), methyl 3-[1-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-1*H*-indol-3-yl]propanoate **138b** (0.106g, 0.199mmol), methanol (30ml) and Dowex© 50WX8 (5ml). The residual solid was purified by HPLC [on a C-18 LUNA column; elution with MeOH-H_2O (2:1)] to afford, as a brown powder, *3-(β-D-*

glucopyranosyl-1H-indol-3-yl)propanoic acid **139b** (36.2mg, 49.9%), mp. > 320°C; δ_H (100MHz, D_2O) 2.71 (2H, t, *J* = 7.6Hz, CH_2CH_2CO), 3.10 (2H, t, *J* = 7.6Hz, CH_2CH_2CO), 3.76-3.79 and 4.00-4.05 (2H, 2 x m, 6'-CH_2), 3.87-3.93 (4H, m, 2'-, 3'-, 4'- and 5'-H), 4.22 (1H, t, *J* = 9Hz, 1'-H), 7.38 (1H, t, *J* = 7.4Hz, 6-H), 7.47 (2H, m, 2- and 5-H), 7.75 (1H, d, *J* = 7.8Hz, 4-H) and 7.86 (1H, d, *J* = 7.8Hz, 7-H); δ_C (100MHz, D_2O) 23.8 (CH_2CH_2CO), 37.2 (CH_2CH_2CO), 61.5 (C-6'), 70.1 (C-4'), 73.2 (C-2'), 77.2 (C-3'), 78.8 (C-5'), 84.8 (C-1'), 110.5 (C-4), 118.0 (C-3), 120.0 (C-7), 120.9 (C-5), 123.1 (C-6), 123.2 (C-2), 128.9 (C-3a), 137.4 (C-7a) and 182.3 (C=O).

***4-(β-D-glucopyranosyl-1H-indol-3-yl)butanoic acid* 139c**

Method A

The experimental procedure employed for the synthesis of the acetic acid derivative **139a** was followed, using methyl 4-[1-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-1*H*-indol-3-yl]butanoate **138c** (62.2mg, 0.114mmol), sodium metal (1.07g), methanol (30ml) and activated Dowex© 50Wx8 (10ml). The residual solid was purified by HPLC [on a C-18 LUNA; elution with MeOH-H_2O (2:1)] to afford as, off-white powder, *4-(β-D-glucopyranosyl-1H-indol-3-yl)butanoic acid* **139c** (22.1mg, 50.7%), mp. > 320; δ_H (100MHz; D_2O) 2.04-2.12 (2H, m, $CH_2CH_2CH_2CO$), 2.40 (2H, t, *J* = 7.4Hz, $CH_2CH_2CH_2CO$), 2.92 (2H, t, *J* = 7.4Hz, $CH_2CH_2CH_2CO$), 3.76-3.79 and 4.00-4.05 (2H, 2 x m, 6'-CH_2), 3.88-3.93 (4H, m, 2'-, 3'-, 4'- and 5'-H), 5.77 (1H, d, *J* = 9.2Hz, 1'-H), 7.38 (1H, t, *J* = 7.4Hz, 6-H), 7.47 (2H, m, 2- and 5-H), 7.75 (1H, d, *J* = 7.8Hz, 4-H) and 7.86 (1H, d, *J* = 7.8Hz, 7-H); δ_C (100MHz, D_2O) 24.6 ($CH_2CH_2CH_2CO$), 26.8 ($CH_2CH_2CH_2CO$), 37.6 ($CH_2CH_2CH_2CO$), 61.3 (C-6'), 70.1 (C-4'), 72.2 (C-2'), 77.2 (C-3'), 78.8 (C-5'), 84.8 (C-1'), 110.9 (C-4), 118.0 (C-3), 120.0 (C-7), 120.9 (C-5), 123.1 (C-6), 123.2 (C-2), 128.9 (C-3a), 137.4 (C-7a) and 182.0 (C=O); *m/z* 366 (M+1, 100%).

Method B[164]

A solution of methyl 4-[1-(3,4,5-triacetoxy-6-acetoxymethyl-tetrahydropyran-2-yl)-1H-indol-3-yl]-butanoate **138c** (54.6mg, 0.0998mmol) in triethylamine (1ml) was stirred at

room temperature for 24h. The solvent was removed *in vacuo* and the residual solid was shown by ^{1}H NMR to be the starting compound **138c**.

Benzimidazole derivatives

1-Allylbenzimidazole 141[165]

To a solution of benzimidazole **140** (3.55g, 0.03mol) in dry THF (60ml), cooled to 0°C in an ice-bath, an excess of NaH (1.46g, 0.06mol) was added slowly to allow for controlled evolution of H_2. The mixture was allowed to warm to room temperature and stirred for 15min. Allyl bromide (4.32g, 0.036mol) was then added in one portion and the mixture was boiled under reflux for 30min. The reaction was quenched with water and the THF evaporated *in vacuo*. The residue was re-dissolved in CH_2Cl_2 and washed with H_2O (2 x) and with brine (2 x). The aqueous layers were combined and extracted with CH_2Cl_2 (2 x), and the combined organic layers were dried with anhydrous Na_2SO_4. The solvent was evaporated *in vacuo* and the crude product was purified by flash chromatography [on silica gel; elution with EtOAc-CH_2Cl_2 (2:3)] to afford, as a pale yellow oil, 1-allylbenzimidazole **141** (4.54g, 95.9%); (Found: M^+, 158.08341. $C_{10}H_{12}N_2$ requires *M*, 158.08440); δ_H (400MHz; $CDCl_3$) 4.66 (2H, td, *J* = 5.5 and 1.49Hz, NC*H*$_2$), 5.10 and 5.21 (2H, 2 x d, *J* = 10.2 and 17.1Hz, CH=C*H*$_2$), 5.91 (1H, tdd, *J* = 20.7, 10.7 and 5.5Hz C*H*=CH$_2$), 7.23 (2H, m, 5- and 6-H), 7.30 (1H, m, 7-H), 7.77 (1H, m, 4-H) and 7.81 (1H, s, 2-H); δ_C (100MHz; $CDCl_3$) 47.1 (NCH_2), 109.7 (C-4), 118.3 (CH=*C*H$_2$), 120.0 (C-7), 121.9 (C-6), 122.7 (C-5), 131.7 (*C*H=CH$_2$), 133.6 (C-3a), 142.7 (C-2) and 143.6 (C-7a); *m/z* 159 (M+1, 100%).

1-(But-3-enyl)benzimidazole 142[165]

The experimental procedure employed for the synthesis of *N*-allylbenzimidazole **140** was employed using, benzimidazole (3.68g, 0.031mol), NaH (1.49g, 0.062mol), THF (60ml) and 4-bromobutene (4.92g, 0.036mol). The mixture was refluxed for 4h and the crude product was purified by flash chromatography [on silica gel; elution with EtOAc-$CHCl_3$

(3:2)] to afford, as a yellow oil, 1-(but-3-enyl)benzimidazole **141** (2.27g, 42.7%), (Found: M^+, 172.10064. $C_{11}H_{12}N_2$ requires *M*, 172.10005); δ_H (400MHz; $CDCl_3$) 2.52 (2H,s, 2'-CH_2), 4.08-4.16 (2H, m, 1'- CH_2), 4.98 (2H, m, 4'- CH_2), 5.70 (1H, m, 3'-CH), 7.25 (2H, m, 5- and 6-H), 7.35 (1H, m, 4-H), 7.77 (1H, m, 7-H) and 7.83 (1H,s 2-H); δ_C (100MHz; $CDCl_3$) 33.8 (NCH_2CH_2CH), 44.4 (NCH_2CH_2CH), 109.5 (C-4), 118.2 ($CH=CH_2$), 120.2 (C-7), 121.9 (C-6), 122.7 (C-5), 133.4 ($CH=CH_2$), 133.5 (C-3a), 142.8 (C-2) and 143.7 (C-7a); *m/z* 172 (M^+, 100%).

1-(3-Methylbut-2-enyl)benzimidazole 143[208]

The experimental procedure employed for the synthesis of *N*-allylbenzimidazole **141** was employed, using benzimidazole (0.432g, 3.66mmol), THF (7.2ml), NaH (0.180g, 7.5mmol) and 4-bromo-2-methyl-2-butene (0.647g, 4.34mmol). The reaction mixture was refluxed for 40min, and the crude product was purified by flash chromatography [on silica gel; elution with EtOAc-$CHCl_3$ (2:3)] to afford, as a yellow oil, 1-(3-methylbut-2-enyl)benzimidazole **143** (0.457g, 67.1%), (Found: M^+, 186.11617. $C_{12}H_{14}N_2$ requires *M*, 186.11570); δ_H (400MHz; $CDCl_3$) 1.74 (3H, s, CH_3), 1.78 (3H, s, CH_3), 4.66 (2H, d, *J* = 6.8Hz, CH_2CH), 5.34 [1H, t, *J* = 7.0Hz, $CH=C(CH_3)_2$], 7.21-7.24 (2H, m, 5-and 6-H), 7.30-7.33 (1H, m, 7-H), 7.74-7.76 (1H, m, 4-H) and 7.84 (1H, s, 2-H); δ_C (100MHz; $CDCl_3$) 18.0 (CH_3), 25.6 (CH_3), 42.9 (NCH_2), 109.7 (C-4), 118.1 [$CH=C(CH_3)_2$], 120.2 (C-7), 122.0 (C-6), 122.7 (C-5), 133.8 (C-3a), 138.2 (C-7a) and 142.4 (C-2).

Cetyltrimethylammonium permanganate CTAP[167]

To a stirred solution of potassium permanganate (3.171g, 20.07 mmol) in water (100ml) at 20°C, a solution of cetyltrimethylammonium bromide (8.027g, 22.02 mmol) in water (100ml) was added dropwise. A fine precipitate was observed to form and the mixture was stirred for a further 30min. The precipitate was filtered off, washed thoroughly with water and dried *in vacuo* to afford, CTAP (7.639g, 86%) as a fluffy violet solid.

1-(2,3-dihydroxypropyl)benzimidazole 144[209,210]

Method A[165]

To a stirred solution of *N*-allylbenzimidazole **141** (1.1016g, 6.97mmol) in ^{t}BuOH (5.6ml), a solution of CTAP (2.8166g, 6.98mmol) in Bu^{t}OH (28ml) and H_2O (7ml) was added drop-wise, and the mixture was stirred for 9h at 20°C. The mixture was diluted with $CHCl_3$ (70ml) and 5% NaOH (22ml) and then stirred for 30min. The layers were separated and the aqueous phase was extracted with $CHCl_3$ (3 x 70ml). The organic layers were combined, dried over anhydrous $MgSO_4$ and evaporated *in vacuo*. The residual oil was purified by flash chromatography [on silica gel; elution with EtOAc-CH_2Cl_2 (3:2)] to afford, as a colourless oil, 1-(2,3-dihydroxypropyl)benzimidazole **144** (0.6089g, 44.1%), (Found: M^+, 192.09095. $C_{10}H_{12}N_2O_2$ requires *M*, 192.08988); δ_H (300MHz; DMSO-d_6) 3.30 (1H, td, *J* = 5.7 and 11.4Hz, H), 3.40 (1H, td, *J* = 5.2 and10.7Hz, HOC*H*$_2$), 3.77-3.84 (1H, m, C*H*OH), 4.13 (1H, dd, *J* = 7.6 and 14.4, NC*H*$_2$), 4.35 (1H, dd, *J* = 3.2 and 14.4Hz, NC*H*$_2$), 4.83 (1H, t, *J* = 5.6Hz, CH$_2$O*H*), 5.06 (1H, d, *J* = 5.2Hz, CHO*H*), 7.18 (1H, t, *J* = 7.1Hz, 6-H), 7.24 (1H, t, *J* = 7.1Hz, 5-H), 7.58 (1H, d, *J* = 8.0Hz, 4-H), 7.64 (1H, d, *J* = 8.0Hz, 7-H) and 8.13 (1H, s, 2-H); δ_C (100MHz, DMSO-d_6) 47.4 (NCH_2), 63.1 (CH_2OH), 70.0 (CHOH), 110,5 (C-4), 119.1 (C-7),121.1 (C-6),121.9 (C-5), 134.3 (C-3a), 143.2 (C-7a) and 144.6 (C-2); *m/z* 192 (M^+, 11.6%) and 131 (100%).

Method B[166]

To a vigorously stirred mixture of *N*-allylbenzimidazole **141** (0.509g, 3.22mmol), in Bu^{t}OH (11ml) and NaOH (0.0543g, 1.36mmol) in H_2O (0.3ml), cooled to –5°C by the addition of ice (5g), a cooled solution of $KMnO_4$ (0.340g, 2.15mmol) in H_2O (11ml) was added drop-wise at a rate such that the temperature did not exceed 15°C. Celite (0.1-0.2g) was then added and the mixture heated almost to boiling. The hot mixture was filtered by suction and washed several times with H_2O-Bu^{t}OH [1:1 (5ml)]. The filtrate was distilled off until 6ml of solution remained. The hot solution was weighed and solid K_2CO_3 (1g/g of solution) was dissolved in small portions until an oil separated. The oil was allowed to

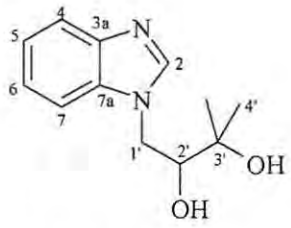
solidify and the rest of the K_2CO was added. The solid was filtered off and re-dissolved in hot EtOAc, with the insoluble residue being removed by gravity filtration. The EtOAc was evaporated *in vacuo* and the residual oil purified by radial chromatography [on silica; elution with EtOH-$CHCl_3$ (2:3)] to afford, as a colourless oil, 1-(2,3-dihydroxypropyl)benzimidazole **144** (0.092g, 15%).

Method C

A solution of 1-[(2,2-dimethyl-1,3-dioxolan-4-yl)methyl]benzimidazole **147** (0.205g, 0.884mmol) in 75% (v/v) acetic acid (20ml) was boiled under reflux for 1h and the solvent removed *in vacuo* to afford, as a golden oil, *N*-(2,3-dihydroxypropyl)benzimidazole **144** (0.165g, 97.0%).

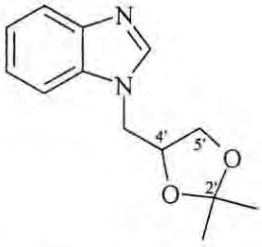
***1-(3,4-dihydroxybutyl)benzimidazole* 145**

The experimental procedure employed for the synthesis *N*-(2,3-dihydroxypropyl)benzimidazole **144** was followed, using 1-(but-3-enyl)benzimidazole **142** (0.55g, 3.2mmol) in Bu^tOH (2.6ml), CTAP (1.3g, 3.2mmol) in H_2O-Bu^tOH [16ml (1:4)], $CHCl_3$ (32ml) and 5% aq. NaOH (10ml). The reaction mixture was stirred for 6.5h and the aqueous layer was then extracted with $CHCl_3$ (3 x 32ml). The solvent was removed *in vacuo* and the residual oil was purified by radial chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1:4)] to afford, as white crystals, *1-(3,4-dihydroxybutyl)benzimidazole* **145** (0.172g, 26%), mp. 150-152^{O}C, (Found: M^+, 206.10361. $C_{11}H_{14}N_2O_2$ requires *M*, 206.10553); δ_H (400MHz; DMSO-d_6) 1.70 and 2.01 (2H, m, CH_2CHOH), 3.25 (1H, dd, *J* = 6.6 and 11.8Hz, 4'-CH_aOH), 3.33 (2H, m, 4'-CH_bOH and 3'-C*H*OH), 4.34 (2H, t, *J* = 7.1Hz, NCH_2), 4.60 (1H, t, *J* = 5.4Hz, 4'-CH_2O*H*), 4.83 (1H, d, *J* = 4.7Hz, 3'-CHO*H*), 7.20 (1H, t, *J* = 7.4Hz, 6-H), 7.25 (1H, t, *J* = 7.4Hz, 5-H), 7.59 (1H, d, *J* = 7.9Hz, 4-H), 7.65 (1H, d, *J* = 7.9Hz, 7-H) and 8.20 (1H, s, 2-H); δ_C (100MHz; DMSO-d_6) 33.6 (CH_2CHOH), 41.0 (NCH_2), 65.7 (CH_2OH), 68.3 (CHOH), 110.3 (C-4), 119.3 (C-7), 121.3 (C-6), 122.1 (C-5), 133.7 (C-3a), 143.4 (C-7a) and 144.0 (C-2); *m/z* 206 (M^+, 74.3%) and 132 (100%).



N-*(3-methyl-2,3-dihydroxybutyl)benzimidazole* 146

The experimental procedure employed for the synthesis of *N*-(2,3-dihydroxypropyl)benzimidazole **144** was followed, using 1-(3-methylbut-2-enyl) benzimidazole **143** (0.239g, 1.29mmol) in Bu^{t}OH (1ml), CTAP (0.503g, 1.29mmol) in H_2O-Bu^{t}OH (1:4; 6.5ml), $CHCl_3$ (13ml) and 5% NaOH (3.9ml). The reaction mixture was stirred for 6.5h and the aqueous layer was then extracted with $CHCl_3$ (3 x 13ml). The residual oil was purified by flash chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1:4)] to afford, as cream needles, N-*(3-methyl-2,3-dihydroxybutyl)benzimidazole* **146** (0.093g, 33%), mp. 144-146^{O}C, (Found: M^+, 220.12083. $C_{12}H_{16}N_2O_2$ requires *M*, 220.12118); δ_H (400MHz; DMSO-d_6) 1.15 and 1.18 (6H, s, 2 x CH_3), 3.46 (1H, m, 2'-H), 4.02 (1H, dd, *J* = 10.0 and 14.2Hz, 1'-H_a), 4.52 (1H, d, *J* = 1.4 and 14.2Hz, 1'-H_b), 4.65 (1H, s, 2'-O*H*) 5.11 (1H, d. *J* = 6.2Hz, 3'-O*H*), 7.19 (1H, t, *J* = 7.4Hz, 6-H), 7.25 (1H, t, *J* = 7.4Hz, 5-H), 7.54 (1H, d, *J* = 7.9Hz, 4-H), 7.64 (1H, d, *J* = 7.9Hz, 7-H) and 8.13 (1H, s, 2-H); δ_C (100MHz; DMSO-d_6) 23.6 and 27.5 (2 x CH_3), 46.6 (C-1'), 71.0 (C-2'), 75.8 (C-3'), 110.4 (C-4), 119.2 (C-7), 121.1 (C-6), 121.9 (C-5), 133.9 (C-3a), 143.4 (C-7a) and 144.8 (C-2); *m/z* 220 (M^+, 75.5%) and 118 (100%).



1-(2,2-Dimethyl-1,3-dioxolan-4-ylmethyl) benzimidazole 147[211]

The experimental procedure employed for the synthesis of *N*-allylbenzimidazole **141** was followed, using benzimidazole **140** (1.002g, 8.49mmol), 60% NaH (0.674g, 23.3mmol),

THF (17ml) and 4-bromomethyl-2,2-dimethyl-1,3-dioxolane **128** (1.992g, 10.19mmol). The mixture was refluxed for 18hrs and the oily residue was purified by flash chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1:9)] to afford, as cream crystals, 1-[(2,2-dimethyl-1,3-dioxolan-4-yl)methyl]benzimidazole **147** (0.7452g, 37.8%) mp. 89-91^{O}C, (Found: M^+, 232.11964. $C_{13}H_{16}N_2O_2$ requires *M*, 232.12118); δ_H (400MHz; $CDCl_3$) 1.31 and 1.36 (6H, 2 x s, 2 x CH_3), 3.66 (1H, dd, *J* = 6.0 and 8.6Hz, NCH_a), 4.05 (1H, dd, *J* = 6.3 and 8.6Hz, NCH_b), 4.21 (1H, dd, *J* = 6.1 and 14.7Hz, 5'-H_a), 4.28 (1H, dd, *J* = 4.5 and 14.7Hz, 5'-H_b), 4.44 (1H, m, 4'-H), 7.28 (2H, m, 5- and 6-H), 7.38 (1H, d, *J* = 6.4Hz, 4-H), 7.79 (1H, d, *J* = 6.6Hz, 7-H) and 7.95 (1H, s, 2-H); δ_C (100MHz; $CDCl_3$) 25.3 and 26.7 (2 x CH_3), 47.1 (NCH_2), 66.6 (OCH_2), 74.2 (OCH), 109.5 (C-4), 110.2 [$C(CH_3)_2$], 120.4 (C-7), 122.3 (C-5), 123.1 (C-6), 133.9 (C-3a), 143.3 (C-7a) and 143.4 (C-2); *m/z* 232 (M^+, 22.8%) and 131 (100%)

1-[2-(2-hydroxyethoxy)ethyl]benzimidazole 149[210]

The experimental procedure employed for the synthesis of *N*-allylbenzimidazole **141** was followed, using benzimidazole **140** (1.007g, 8.481mmol), 60% NaH (0.6697g, 27.9mmol), 2-(chloroethoxy)ethanol **148** (1.2737g, 10.22mmol) and THF (20ml). The residual oil was purified by column chromatography [on silica gel; elution with EtOH-$CHCl_3$-hexane (1:1.5:2.5)] to afford, as a pale yellow oil, 1-[2-(2-hydroxyethoxy)ethyl]benzimidazole **149** (0.4089g, 23.4%), (Found: M^+ 206.10432. $C_{11}H_{14}N_2O_2$ requires *M*, 206.10553); δ_H (400MHz; $CDCl_3$) 2.15 (1H, s, OH), 3.50 (2H, t, *J* = 4.6Hz, CH_2), 3.66 (2H, t, *J* = 4.6Hz, CH_2), 3.82 (2H, t, *J* = 5.2Hz, CH_2), 4.33 (2H, t, *J* = 5.2Hz, CH_2), 7.25-7.29 (2H, m, 5- and 6-H), 7.40 (1H, d, *J* = 6.4Hz, 4-H), 7.79 (1H, d, *J* = 6.4Hz, 7-H) and 7.96 (1H, s, 2-H); δ_C (100MHz; $CDCl_3$) 45.0 (NCH_2CH_2O), 61.6 (OCH_2CH_2OH), 69.4 (NCH_2CH_2O), 72.6 (OCH_2CH_2OH), 109.4 (C-4), 120.3 (C-7), 122.2 (C-5), 122.9 (C-6), 133.8 (C-3a and C-7a) and 143.5 (C-2); *m/z* 207 (*M* + 1, 65.2%) and 132 (100%).

1-[(2-Acetoxyethoxy)methyl]benzimidazole 150[209]

The experimental procedure employed for the synthesis of *N*-allylbenzimidazole **141** was followed using, benzimidazole **140** (1.025g, 8.686mmol), 60% NaH (0.6940g, 28.92mmol), 2-(bromomethoxy)ethyl acetate **131** (2.0614g, 10.46mmol) and THF (15mL). The mixture was refluxed for 24 hrs and the residue was purified by radial chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1:9)] to afford, as a yellow oil, 1-[(2-acetoxyethoxy)methyl]benzimidazole **150** (0.709g, 34.9%), (Found: M^+ 234.09929. $C_{12}H_{14}N_2O_3$ requires *M*, 234.10044); ν_{max} (thin film)/cm^{-1} 1733.5 (C=O); δ_H (400MHz; $CDCl_3$) 1.95 (3H, s, CH_3), 3.59 (2H, t, *J* = 4.6Hz, OCH_2CH_2OCO), 4.13 (2H, t, *J* = 4.6Hz, OCH_2CH_2OCO), 7.31 (2H, m, 5- and 6-H), 7.52 (1H, d, *J* = 6.2Hz, 4-H), 7.79 (1H, d, *J* = 6.2Hz, 7-H) and 7.80 (1H, s, 2-H); δ_C (100MHz; $CDCl_3$) 20.6 (CH_3), 62.7 (OCH_2CH_2OCO), 66.5 (OCH_2CH_2OCO), 74.7 (NCH_2O), 110.1 (C-4), 120.4 (C-7), 122.8 (C-6), 123.6 (C-5), 133.4 (C-3a), 142.8 (C-2), 144.0 (C-7a) and 170.7 (C=O); *m/z* 235 (*M* +1, 100%).

1-[(2-Hydroxyethoxy)methyl]benzimidazole 151[209]

To a suspension of LAH (0.0796g, 2.1mmol) in dry THF (5ml) under nitrogen, a solution of 1-[(2-acetoxyethoxy)methyl]benzimidazole **150** (0.4450g, 1.90mmol) in dry THF (2ml) was added drop-wise and the mixture was heated at 60°C for 9h. The reaction was cooled to room temperature and EtOAc was then added to destroy excess LAH. Excess solvent was removed *in vacuo* and the residual oil was purified by column chromatography [on silica gel; elution with EtOH-$CHCl_3$-hexane (1:1.5:2.5)] to afford, as a golden oil, 1-[(2-hydroxyethoxy)methyl]benzimidazole **151** (77.8mg, 21.3%), (Found: M^+, 192.09042. $C_{10}H_{12}N_2O_2$ requires *M*, 192.08988); δ_H (400MHz; $CDCl_3$) 3.52 (2H, t, *J* = 4.6Hz), 3.71 (2H, t, *J* = 4.6Hz), 5.72 (2H, s, NCH_2), 7.31 (2H, m, 5- and 6-H), 7.5 (1H, dd, *J* = 2.9 and 6.1Hz, 4-H), 7.7 (1H, dd, *J* = 3.0 and 5.8Hz, 7-H) and 7.92 (1H, s, 2-H); δ_C (100MHz; $CDCl_3$) 61.1(OCH_2CH_2OH), 70.1 (OCH_2CH_2OH), 74.8 (NCH_2O), 110.2 (C-4), 120.1 (C-7), 122.8 (C-5), 123.6 (C-6), 133.3 (C-3a), 143.0 (C-2) and 143.5 (C-7a); *m/z* 193 (*M* + 1, 100%).

This compound was also isolated in 11.7% yield as a by-product in the synthesis of *N*-[(2-acetoxyethoxy)methyl]benzimidazole **150**.

Attempted synthesis of *1-{2-hydroxy-[3-(2-hydroxyethoxy)ethyl]propyl}benzimidazole* 152a

The experimental procedure employed for the synthesis of 1-allylbenzimidazole **141** was followed, using 1-(2,3-dihydroxypropyl)benzimidazole **144** (0.252g, 1.31mmol), 60%NaH (57.2mg, 2.38mmol), 2-(2-chloroethoxy)ethanol **148** (0.166g, 1.33mmol) and THF (10ml). The reaction mixture was refluxed for 6h. The residual oil was analyzed by ^{1}H NMR spectroscopy, but none of the peaks corresponding to compounds **152a** or **153a** were observed.

Attempted synthesis of *1-{3-hydroxy-[4-(2-hydroxyethoxy)ethyl]butyl}benzimidazole* 152b

The experimental procedure employed for the synthesis of *N*-allylbenzimidazole **141** was followed, using 1-(3,4-dihydroxybutyl)benzimidazole **145** (0.391g, 1.90mmol), 60% NaH (90mg, 3.75mmol), 2-(2-chloroethoxy)ethanol **148** (0.244g, 1.92mmol) and THF (15ml). The reaction mixture was refluxed for 6h. The residual oil was analyzed by ^{1}H NMR spectroscopy, but none of the peaks corresponding to compounds **152b** or **153b** were observed.

Attempted synthesis of 1-(2-hydroxy-3-oxopropyl)benzimidazole 154a

Method A[168]

To a suspension of 1-(2,3-dihroxypropyl)benzimidazole **144** (0.355g, 1.85mmol) in CH_2Cl_2 (20ml) was added trichloroisocyanuric acid (0.451g, 1.94mmol), and the mixture was stirred at 0°C for 15min. To this mixture, 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO) (28.9mg, 0.185mmol) was added and the mixture was then stirred for a further 30min at room temperature. The mixture was filtered through Celite, and the organic

phase was washed with saturated aq. Na_2CO_3 (15 ml), followed by 1M HCl and brine. The organic layer was dried over anhydrous Na_2SO_4, and the solvent was evaporated *in vacuo*. ^{1}H NMR analysis of the residual oil failed to reveal the expected aldehyde signal.

Method B[169]

To a solution of 1-(2,3-dihroxypropyl)benzimidazole **144** (55.6mg, 0.289mmol) and DCC (0.234g, 1.13mmol) in dry DMSO (2ml) under a stream of nitrogen, was added dry pyridine (23µl) and TFA (11µl, 0.146mmol). The mixture was stirred at 25^{o}C for 24h and the resulting *N,N'*-dicyclohexylurea precipitate was filtered off. The mother liquor was diluted with dry diethyl ether (2ml), treated with a solution of oxalic acid (72mg) in dry methanol (226µl), and stirred for 1h. The resulting *N,N'*-dicyclohexylurea precipitate was filtered off and the mother liquor was diluted with water (2ml). The resulting phases were separated and a Schiff test was conducted on each phase. Both phases gave a negative Schiff test and were evaporated *in vacuo*. 1H NMR analysis of the residual oils were failed to reveal the expected aldehyde signal.

Method C[170]

MnO_2 (3.542g) was added to a solution of 1-(2,3-dihroxypropyl)benzimidazole **144** (0.120g, 0.627mmol) in $CHCl_3$ (100ml), and the reaction mixture was boiled under reflux for 24h. The reaction mixture was filtered over celite while hot and the solvent was dried over $MgSO_4$ and evaporated *in vacuo*. Analysis of the residual solid by ^{1}H NMR spectroscopy showed none of the peaks corresponding to compound **154a**.

Attempted synthesis of *1-(3'-hydroxybutanal)benzimidazole* 154b

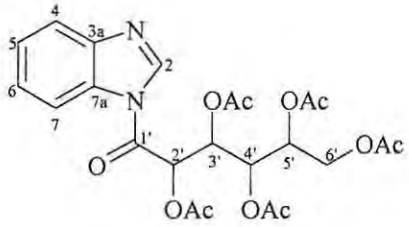
Method B[169]

The experimental procedure (**method B**) employed for the attempted synthesis of 1-(2-hydroxypropanal)benzimidazole **154a** was followed, using 1-(3,4-dihydroxybutyl)benzimidazole (28.8mg, 0.140), DCC (0.115g, 0.557mmol), DMSO (1ml), pyridine (11µl), TFA (6.0µl), diethyl ether (1ml), MeOH (11µl), oxalic acid (35mg) and water (1ml). Both the aqueous and the organic phases gave negative Schiff

test. ^{1}H NMR analysis of the residual oils also failed to reveal the expected aldehyde signal.

Method C[170]

The experimental procedure (**method C**) employed for the attempted synthesis of N-*(2'-hydroxypropanal)benzimidazole* **154a** was followed using, *N*-(3',4'-dihroxybutyl)benzimidazole (0.108g, 0.523mmol), MnO_2 (2.954g) and $CHCl_3$ (100ml). Analysis residual oil by ^{1}H NMR showed none of the peaks corresponding to compound **154b**.



Attempted synthesis of N-*(2,3,4,5,6-penta-O-acetyl-D-gluconyl)benzimidazole* 155

Method A

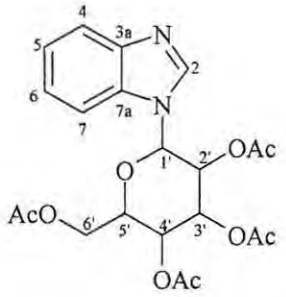
The experimental procedure employed for the synthesis of 1-allylbenzimidazole **141** was followed, using benzimidazole **140** (1.02g, 8.64mmol), gluconyl chloride **112** (4.39g, 10.3mmol), NaH (0.415g, 17.3mmol) and THF (17ml). The residual oil was purified by flash chromatography [on silica gel; elution with MeOH-$CHCl_3$ (1:9)] to afford the starting benzimidazole **140** and gluconic acid **115**.

Method B[171]

To a solution of benzimidazole **140** (0.236g, 1.99mmol) in THF (10ml) was added gluconyl chloride **112** (0.425g, 1.00mmol) and the resulting solution was refluxed for 12h. The solvent was removed *in vacuo*, and the residual oil was purified by flash chromatography [on silica gel; elution with MeOH-$CHCl_3$ (1:9)] to afford benzimidazole **140**, gluconic acid **115** and a mixture of benzimidazole with trace amounts of compound **155**.

Method C[172]

DCC (0.690g, 3.33mmol) was added to a solution of benzimidazole **140** (0.394g, 3.33mmol) and gluconic acid **115** (1.356g, 3.34mmol) in dry EtOAc (42ml) under nitrogen, and the reaction mixture stirred at room temperature for 21h.The mixture was filtered and the solvent removed *in vacuo*. The residual oil was purified by flash chromatography [on silica gel, elution with MeOH-$CHCl_3$ (1:9)] to afford benzimidazole **140**, gluconic acid **115** and a mixture of benzimidazole with trace amounts of compound **155**.



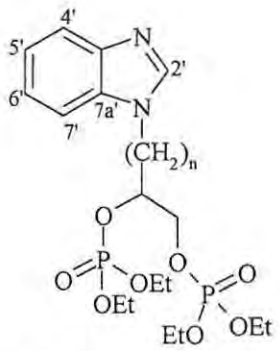
1-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)benzimidazole 156[212]

The experimental procedure employed for the synthesis of compound **138a** was followed, using benzimidazole **140** (0.5064g, 4.29mmol), compound **137** (1.672g, 4.29mmol), $SnCl_4$ (3.164g, 12.2mmol). The solution was refluxed for 5h and quenched with saturated aq. $NaHCO_3$ (100ml). The residual solid was purified by flash chromatography [on silica gel; elution with EtOAc-$CHCl_3$ (3:2)] to afford a pale yellow oil, which was purified further by HPLC [on partisil 10; elution with EtOAc-Hexane (7:3)] to afford, as a colourless oil, 1-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)benzimidazole **156** (0.201g, 10.4%); ν_{max} (thin film)/cm^{-1} 1746.1 (C=O); δ_H (400MHz; $CDCl_3$) 1.73-2.08 (12H, 4 x s, 2'-, 3'-, 4'- and 6'-CO_2CH_3) 4.00-4.04 (1H, m, 5'-H), 4.20 (1H, 6'-H), 4.30 (1H, dd, J = 4.8 and 12.6Hz, 6'-H_b), 5.32 (1H, t, J = 9.8Hz, 4'-H), 5.44 (1H, t, J = 9.4Hz, 3'-H), 5.55 (1H, t, J = 9.4Hz, 2'-H), 5.65 (1H, d, J = 9.2Hz, 1'-H), 7.30-7.37 (2H, m, 5- and 6-H), 7.52 (1H, d, J = 8.0Hz, 4-H), 7.80 (1H, d, J = 8.0Hz, 7-H) and 8.12 (1H, s, 2-H); δ_C (100MHz, $CDCl_3$) 20.0-20.7 (4 x CO_2CH_3), 61.6 (C-6), 67.8 (C-4'), 70.1 (C-2'), 72.9 (C-3'), 75.0 (C-

5'), 83.2 (C-1'), 110.5 (C-7), 120.6 (C-4), 123.1 (C-5), 123.7 (C-6), 132.5 (C-3a), 140.8 (C-2), 143.6 (C-7a) and 168.6-170.6 (4 x CO_2CH_3).

1-[3,4,5-triHydroxy-6-(hydroxymethyl)tetrahydropyran-2-yl]benzimidazole 157

The experimental procedure employed for the synthesis of compound **139a** was followed, using sodium metal (3.08g), the tetraacetate **156** (0.815g, 1.81mmol) and methanol (90ml). The residual solid was purified by reverse phase chromatography [on a C_{18} Sep-pak column; elution with MeOH-H_2O (3:2)] to afford, as a cream powder, 1-[3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydropyran-2-yl]benzimidazole **157** (0.1235g, 37.5%), mp. > 320°C; δ_H (400MHz; D_2O) 3.44-3.42 (6H, m, glucopyranose-H) 7.47-7.47 (2H, m, 5- and 6-H), 7.80-7.82 (2H, m, 4- and 7-H) and 8.34 (1H, s, 2-H).



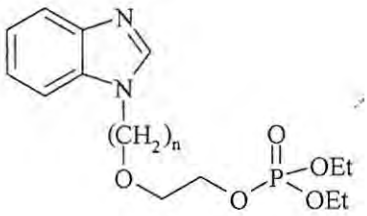
3-(Benzimidazol-1-yl)propane-1,2-diol bis(diethyl phosphate) **158a**

BuLi (1.6M in hexane; 0.46ml, 0.732mmol) was added to a solution of 1-(2,3-dihydroxypropyl)benzimidazole **143** (0.140g, 0.728mmol) in dry THF (7ml) cooled to 0°C under argon, and the reaction mixture was stirred for 0.5h at 0°C. Diethyl chlorophosphate (27μL, 0.727mmol) in dry THF (1ml) was added and the reaction mixture was stirred for a further 3h at 0°C. The reaction was quenched with saturated aq. NH_4Cl (0.4ml) and the solvent evaporated *in vacuo*. The residual solid was re-dissolved in EtOAc and washed with brine (2ml). The organic layer was dried with anhydrous $MgSO_4$ and the solvent evaporated *in vacuo*. The residual oily solid was purified by radial chromatography [on silica; elution with ethanol-$CHCl_3$-hexane (1:1.5:2.5)] to

afford, as a pale yellow oil, *3-(benzimidazol-1-yl)propane-1,2-diol bis(diethyl phosphate)* **158a** (6.8mg, 2.0%), (Found: M^+, 464.14687. $C_{18}H_{30}N_2O_8P_2$ requires *M*, 464.14774); δ_H (400MHz; $CDCl_3$) 1.13 and 1.23 (6H, 2 x t, *J* = 7.0Hz, $POCH_2CH_3$), 1.30-1.36 (6H, m, $POCH_2CH_3$), 3.84 (2H, q, *J* = 6.9Hz, OCH_2CH_3) 4.05-4.19 (8H, m, CH_2OP and 3 x OCH_2CH_3), 4.44-4.58 (2H, m, CH_2CHOP), 4.77-4.85 (1H, m, CH_2CHOP), 7.29 (2H, dd, *J* = 7.07 and 8.22 Hz, 5'- and 6'-H), 7.47 (1H, d, *J* = 7.5Hz, 4'-H), 7.79 (1H, d, *J* = 7.69Hz, 7'-H) and 7.96 (1H, s, 2'-H); δ_C (100MHz; $CDCl_3$) 15.8-16.2 (C, m, 4 x $POCH_2CH_3$), 45.4 (d, *J* = 5.5Hz, NCH_2CHOP), 62.6 (d, *J* = 5.4, $CHCH_2OP$), 64.2-64.4 and 65.4 (C, m, 4 x $POCH_2CH_3$), 73.5 (dd, *J* = 5.6 and 8.3, $CHCH_2OP$), 109.6 (C-4'), 120.5 (C-7'), 122.5 (C-5'), 123.3 (C-6'), 133.9 (C-3a'), 143.6 (C-2') and 143.7 (C-7a'); δ_P (162MHz; $CDCl_3$) –0.63 and –1.27; *m/z* 464 (*M*, 12.7%) and 156 (100%).

1-(Benzimidazol-1-yl)butane-3,4-diol bis(diethyl phosphate) **158b**

The experimental procedure employed for the synthesis of compound **158a** was followed, using 1-(3,4-dihydroxybutyl)benzimidazole **144** (0.2140g, 1.039mmol), BuLi (1.6M in hexane; 0.65ml, 1.04mmol), diethyl chlorophoshate (39μl, 1.05mmol) and THF (8ml). The residual oil was purified by radial chromatography [on silica gel; elution with EtOH-$CHCl_3$-hexane (1:1.5:2.5)] to afford, as a yellow oil, *1-(benzimidazol-1-yl)butane-3,4-diol bis(diethyl phosphate)* **158b** (24.2mg, 4.87%), (Found: M^+, 478.16403. $C_{19}H_{32}N_2O_8P_2$ requires *M*, 478.16339); δ_H (400MHz; $CDCl_3$) 1.26-1.37 (12H, m, $POCH_2CH_3$), 1.90-2.05 and 2.23-2.67 (2H, m, 2-CH_2), 4.04-4.16 {10H, m, $[P(OCH_2CH_3)_2]_2$ and 4-CH_2}, 4.31-4.38 (2H, m, 1-CH_2), 4.47-4.49 (1H, m, 3-CH), 7.26-7.31 (1H, m, 5'- and 6'-H), 7.40 (1H, d, *J* = 7.2Hz, 4'-H), 7.79 (1H, d, *J* = 7.2Hz, 7'-H) and 8.05 (1H, s, 2'-H); δ_C (100MHz; $CDCl_3$) 15.6-16.1 (m, 4 x $POCH_2CH_3$), 32.1 (d, *J* = 8.1Hz, C-2), 40.7 (C-1), 64.0 [t, *J* = 11.0Hz, $P(OCH_2CH_3)_2$], 64.2 (d, *J* = 9.6Hz, $P(OCH_2CH_3)_2$), 66.3 (d, *J* = 10.2Hz, C-4), 71.3 (d, *J* = 9.9Hz, C-3), 109.5 (C-7'). 120.0 (C-4'), 122.0 (C-5'), 122.9 (C-6'), 133.4 (C-7a'), 143.4 (C-2') and 143.6 (C-7a'); δ_P (162MHz; $CDCl_3$) –5.27 and –5.43; *m/z* 464 (*M* + 1, 100%).



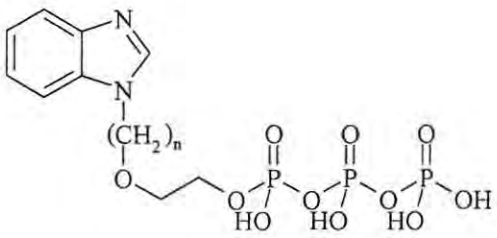
***2-[(Benzimidazol-1-yl)methoxy]ethyl diethyl phosphate* 160a**

The experimental procedure employed for the synthesis of compound **158a** was followed, using 1-[(2-hydroxyethoxy)methyl]benzimidazole **151** (0.250g, 1.30mmol), BuLi (1.6M in hexane; 0.85ml, 1.36mmol), diethyl chlorophosphate (190μL, 1.31mmol) and THF (13ml). The reaction mixture was stirred overnight at room temperature and the residual oil was purified by flash chromatography [on silica gel; elution with EtOH-$CHCl_3$-hexane (1:1.5:2.5)] to afford, as a yellow oil, *2-[(benzimidazol-1-yl)methoxy]ethyl diethyl phosphate* **160a** (0.0998g, 23.3%); δ_H (400MHz; $CDCl_3$) 1.25-1.28 (6H, m, $POCH_2CH_3$), 3.86-4.14 [6H, m, OCH_2CH_2OP and $P(OCH_2CH_3)_2$], 4.67 (2H, s, NCH_2O), 7.24-7.31 (2H, m, 5'- and 6'-H), 7.39 (1H, d, J = 8.2Hz, 4'-H), 7.78 (1H, d, J = 8.2Hz, 7'-H) and 7.97 (1H, s, 2'-H); δ_C (100MHz; $CDCl_3$) 16.3 (d, J = 7.3Hz, $POCH_2CH_3$), 61.7 (d, J = 5.6Hz, $POCH_2CH_3$), 64.1 (d, J = 6.9Hz, OCH_2CH_2OP), 66.2 (OCH_2CH_2OP), 94.7 (NCHO), 110.0 (C-4'), 119.3 (C-7'), 123.0 (C-5'), 123.6 (C-6'), 132.9 (C-3a'), 140.8 (C-7a') and 143.2 (C-2'); δ_P (162MHz, $CDCl_3$) –0.457.

***2-[2-(Benzimidazol-1-yl)ethoxy]ethyl diethyl phosphate* 160b**

The experimental procedure employed for the synthesis of compound **158a** was followed using, *N*-[2-(2-hydroxyethoxy)ethyl]benzimidazole **148** (0.309g, 1.50mmol), 1.6M BuLi in hexane (0.95ml, 1.52mmol), diethyl chlorophosphate (217μL, 1.52mmol) and THF (15ml). The residual oil was purified by flash chromatography [on silica gel, elution with EtOH-$CHCl_3$-hexane (1:1.5:2.5)] to afford, as a yellow oil, *2-[2-(benzimidazol-1-yl)ethoxy]ethyl diethyl phosphate* **160b** (0.129g, 30%), (Found: M^+, 342.13495. $C_{15}H_{23}N_2O_5P$ requires *M*, 342.13446); δ_H (400MHz; $CDCl_3$) 1.25 (6H, t, J = 7.2Hz, $POCH_2CH_3$), 3.48-3.51 (2H, m, OCH_2CH_2OP), 3.53-3.58 (2H, m, OCH_2CH_2OP), 3.78-

3.80 (2H, m, NCH_2CH_2O), 3.97-4.09 (4H, m, $POCH_2CH_3$), 4.30 (2H, t, J = Hz, NCH_2CH_2O), 7.21-7.25 (2H, m, 5'- and 6'-H), 7.35 (1H, d, J = 6.8Hz, 4'-H), 7.73 (1H, d, J = 6.8Hz, 7'-H) and 7.92 (1H, s, 2'-H); δ_C (100MHz; $CDCl_3$) 16.0 (d, J = 10.7Hz, $POCH_2CH_3$), 44.9 (NCH_2CH_2O), 63.8 (dd, J = 9.72 and 13.6Hz, $POCH_2CH_3$), 69.3 (NCH_2CH_2O), 70.0 (dd, J = 11.4 and 23.4Hz, OCH_2CH_2OP), 70.5 (OCH_2CH_2OP), 109.5 (C-4'), 120.1 (C-7'), 122.0 (C-5'), 122.8 (C-6'), 133.7 (C-3a'), 143.4 (C-7a') and 143.5 (C-2'); δ_P (162MHz, $CDCl_3$) –0.484; *m/z* 464 (M + 1, 100%).



Attempted synthesis of 1-*[(2-hydroxyethoxy)methyl]benzimidazole triphosphate* 162a

To a suspension of 1-[(2-hydroxyethoxy)methyl]benzimidazole **151** (0.101g, 0.527mmol) in triethyl phosphate (2.3ml), cooled to 0^oC under argon, $POCl_3$ (0.132g, 0.858mmol) was added drop-wise and the reaction mixture stirred for 2h at 0^oC. To this mixture tributylamine (0.489g, 2.63mmol) and 0.5M solution of tributylammonium pyrophosphate (0.692g, 1.91mmol) in dry DMF (5ml) were added simultaneously, and the mixture was stirred for a further 30min at room temperature. The reaction was quenched by the addition of 1.0M aq. triethylammonium bicarbonate (TEAB) (46ml) followed by stirring overnight. The solvent was removed *in vacuo* and the residue was re-dissolved in super quality water before eluting through a Dowex 1X2 chloride form ion exchange column (elution with 0-0.4M aq. TEAB). None of the isolated fractions appeared to contain the desired product.

Attempted synthesis of 1-*[(2-hydroxyethoxy) ethyl]benzimidazole triphosphate* 162b

The experimental procedure employed for the attempted synthesis of 1-[(2-hydroxyethoxy)methyl]benzimidazole triphosphate **162a** was followed, using 1-[2-

(hydroxyethoxy)ethyl]benzimidazole **148** (0.101g, 0.490mmol), triethyl phosphate (2.2ml), $POCl_3$ (0.115g, 0.750mmol), *bis*-tributylammonium pyrophosphate 0.881g, 2.42mmol), dry DMF (5ml), tributylamine (0.459g, 2.45mmol) and 1M TEAB (44ml). The fractions obtained from the ion exchange separation did not appear to contain the desired product.

Pyrazolo[3,4-*d*]pyrimidine derivatives

Synthesis of *4-(tert*butyldimethylsilyloxy)*pyrazolo[3,4-*d]*pyrimidine* 163

The experimental procedure employed for the synthesis of *tert*-butyldimethylsilyl 3-indolylacetate **111a** was followed using, allopurinol **109a** (1.004g, 7.38mmol), TBDMSCl (1.399g, 9.28mmol), imidazole (1.254g, 18.42mmol) and dry DMF (40ml). The reaction mixture was stirred at 35°C for 10h and the un-dissolved material was filtered off. On addition of water, allopurinol **109a** precipitated out and was filtered and washed with water. The aqueous phase was then extracted with petroleum ethe (bp. 40-60°C; 3 x 10ml). The organic phase was washed with saturated aq. $NaHCO_3$ and dried over anhydrous $MgSO_4$. The solvent was removed *in vacuo* and the residual solid was recrystalised from diethyl ether to afford, as a white solid, 4-(*tert*butyldimethylsilyloxy)pyrazolo[3,4-*d*]pyrimidine **163** (0.231g, 12.5%); δ_H (400MHz, DMSO-d_6) 0.06 (6H, s, $SiCH_3$), 0.82 (9H, s, $C(CH_3)_3$), 7.66 (1H, s, 6-H), 7.94 (1H,s, 3-H) and 9.12 (1H, s, NH); δ_C (100MHz; DMSO-d_6) –3.3 (CH_3), 17.6 [$SiC(CH_3)_3$], 25.7 [$C(CH_3)_3$], 105.5 (C-3a), 134.8 (C-3), 147.5 (C-6), 157.7 (C-7a) and (C-4).

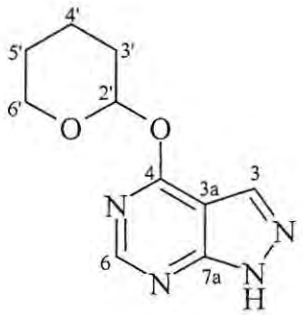
Attempted synthesis of *4-(tert*butyloxycarbamoyl)*pyrazolo*[3,4-*d*]*pyrimidine* 164a

To a suspension of 4-aminopyrazolo[3,4-*d*]pyrimidine **108a** (100mg, 0.741mmol) in dry THF (1ml) under nitrogen, di-*tert*butyl dicarbonate $(Boc)_2O$ (0.160mg, 0.741mmol) in THF (1ml) was added *via* a cannula. The reaction mixture was boiled under reflux for 7 days. The solid material was filtered off and found to be compound **108a**. The mother

liquor was evaporated *in vacuo* and the residual solid was found to be an impure *tert*-butyloxycarbonyl adduct.

Attempted synthesis of 4-(*tert*butyloxycarboxy)pyrazolo[3,4-*d*]pyrimidine 164b

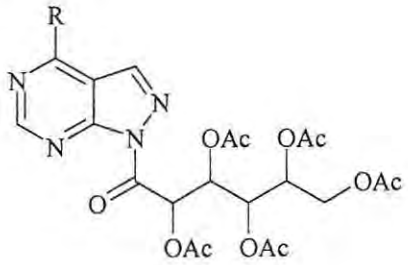
The experimental procedure employed for the attempted synthesis of 4-(tertbutyloxycarbamoyl)pyrazolo[3,4-*d*]pyrimidine **164a** was followed, using allopurinol **109a** (100mg, 0.741mmol), $(Boc)_2O$ (0.160mg,0.741mmol) and dry THF (2ml). The solid material was filtered off and found to be compound **109a**. The mother liquor was evaporated *in vacuo* and the residual solid was determined to be an impure *tert*-butyloxycarbonyl adduct.



4-(Tetrahydropyran-2-yloxy)pyrazolo[3,4-*d*]pyrimidine 165[179]

To a solution of allopurinol **109a** (1.121g, 8.242mmol) in reagent grade DMF (50ml), dihydropyran (20ml) and 0.4% PTSA in dry benzene (4ml) were added, and the mixture was stirred at room temperature for 8h. To this mixture, a 4% solution of pyridine in benzene (0.5ml) was added and the resulting mixture then filtered. The solvent was removed at 40°C *in vacuo*. The resulting oil was re-dissolved in petroleum ether (40-60°C; 100ml) and benzene (3ml) and allowed to crystallise in a refrigerator. The resulting off-white crystals were filtered off to afford 4-(tetrahydropyran-2-yloxy)pyrazolo[3,4-*d*]pyrimidine **165** (902mg, 51%), mp. 204-206°C (lit.[181] 203°C); δ_H (400MHz; $CDCl_3$) 1.62 (1H, dd, J = 4.6 and 7.0Hz, 5'-CH_2), 1.76 (2H, m, 4'- and 5'-CH_2), 1.94 (1H, dd, J = 2.7 and 12.8Hz, 3'-CH_2), 2.13 (1H, dd, J = 4.2 and 10.2Hz, 4'-

CH_2), 2.50-2.59 (1H, m, 3'-CH_2), 3.79-3.73 and 4.13-4.10 (2H, 2 x m, 6'-CH_2), 5.89 (1H, dd, J = 2.6 and 10.2Hz, 2'-H), 8.03 (1H, s, 6-H), 8.15 (1H, s, 3-H) and 12.22 (1H, s, NH); δ_C (100MHz; $CDCl_3$) 22.7 (C-4'), 24.8 (C-5'), 29.4 (C-3'), 68.3 (C-6'), 83.0 (C-2'), 106.4 (C-3a), 135.4 (C-3), 146.4 (C-6), 152.4 (C-4), 159.9 (C-7a).



Attempted synthesis of N-*(2,3,4,5,6-penta*-O-*acetyl-D-gluconyl)-4-aminopyrazolo[3,4-d]pyrimidine* 166a

Method A[184]

To a mixture of 4-aminopyrazolo[3,4-*d*]pyrimidine **108a** (50.7mg, 0.376mmol) and K_2CO_3 (110mg, 0.794mmol) in DMF (2ml), preheated to 140^{O}C under a reflux condenser fitted with a $CaCl_2$ guard tube, gluconyl chloride **112** (479mg, 1.13mmol) was added and stirring was continued for 12h. Another portion of the gluconyl chloride **112** (72.2mg, 0.370mmol) was added and the mixture heated for another 12h. The mixture was filtered while hot and the filtrate washed with DMF (3ml). The solvent was removed *in vacuo* and the residual oil was purified by flash chromatography [on silica gel; elution with $CHCl_3$-EtOH (9:1)]. None of the isolated fractions exhibited peaks corresponding to the desired product on ^{1}H NMR analysis.

Method B

NaH (31.6mg, 1.31mmol) was added to a suspension of 4-aminopyrazolo[3,4-*d*]pyrimidine **108a** (83.6mg, 0.619mmol) in dry pyridine (2ml) under nitrogen at room temperature, and the reaction mixture was stirred for 20min. To this mixture, gluconyl chloride **112** (0.262g, 0.616mmol) was added, and the reaction mixture was stirred at room temperature for 48h. The reaction was quenched by adding EtOAc, and the solid

material filtered off and washed with EtOAc. ^{1}H NMR analysis of the crude solid revealed only peaks corresponding to compound **108a**. The mother liquor was then evaporated *in vacuo* and ^{1}H NMR analysis showed only peaks corresponding to gluconic acid **115**.

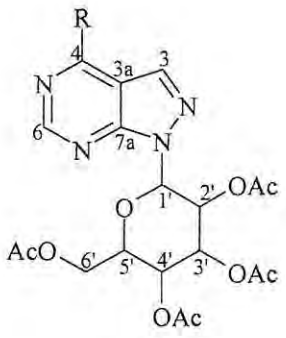
Attempted synthesis of N-*(2,3,4,5,6-penta*-O-*acetyl-D-gluconyl)-4-hydroxypyrazolo[3,4*-d*]pyrimidine* 166b

The experimental procedure employed for the attempted synthesis of *N*-(2,3,4,5,6-penta-*O*-acetyl-D-gluconyl)-4-aminopyrazolo[3,4-*d*]pyrimidine **166a** was followed, using allopurinol **109a** (0.296g, 2.18mmol), compound **112** (0.915g, 2.15mmol), NaH (61.9mg, 2.58mmol) and pyridine (2ml). ^{1}H NMR analysis of both the crude solid material and the residual oil revealed signals corresponding to compound **109a** and **115** respectively.

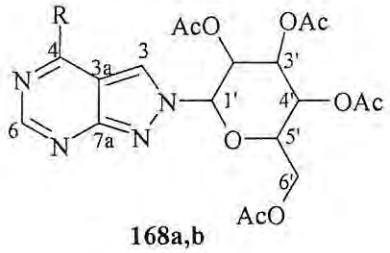
Attempted synthesis of N-*(2,3,4,5,6-penta*-O-*acetyl-D-gluconyl)-4-(tert-butyldimethylsilyloxy)pyrazolo[3,4*-d*]pyrimidine* 166c

A solution of $SOCl_2$ (84.1mg, 0.707mmol) in dry benzene (0.5ml) was added dropwise to a suspension of 4-(*tert*-butyldimethylsilyloxy)pyrazolo[3,4-*d*]pyrimidine **164** (0.136g, 0.544mmol), gluconic acid **115** (0.287g, 0.707mmol) and triethylamine (0.220g, 2.18mmol) in benzene (2ml) at 5°C under nitrogen. The reaction mixture was stirred for 5h at 5°C and the reaction then quenched with water. The resulting two layers were separated and the organic phase was washed successively with dilute aq. HCl, 5% aq. NaOH and brine, dried over anhydrous $MgSO_4$ and evaporated *in vacuo*. The residual oil was determined to be gluconic acid **115**. The aqueous layer was extracted with $CHCl_3$ (3 x 5ml) and combined the organic solutions were dried over anhydrous $MgSO_4$ and evaporated *in vacuo*. The residual solid was shown to be allopurinol **109a** by ^{1}H NMR analysis.

cpd	R
a	NH_2
b	OH



167a,b



168a,b

***1-(2,3,4,6-Tetra-O-acetyl-β-D-glucopyranosyl)-4-aminopyrazolo[3,4-d]pyrimidine* 167a and *2-(2,3,4,6-Tetra-O-acetyl-β-D-glucopyranosyl)-4-aminopyrazolo[3,4-d]pyrimidine* 168a**

$SnCl_4$ (1.113g, 4.27mmol) was added *via* a syringe to a suspension of 4-aminopyrazolo[3,4-*d*]pyrimidine **108a** (0.2303g, 1.709mmol) and 1,2,3,4,6-penta-*O*-acetyl-D-glucopyranose **137** (0.6671g, 1.706mmol) in dry acetonitrile (17ml) under argon at $0^{O}C$. The suspension cleared within 3min, and the solution was heated at $70^{O}C$ for 12h. The reaction was quenched with saturated aq. $NaHCO_3$ (40ml) and extracted with ethyl acetate (3 x 40ml). The combined organic layers were dried over anhydrous $NaSO_4$ and evaporated *in vacuo*. The residual oily solid was purified by radial chromatography [on silica; elution with EtOH-$CHCl_3$ (1:4)] to afford:-

as a pale yellow oil, *1-[2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl]-4-aminopyrazolo[3,4-d]pyrimidine* **167a** (0.1252g, 15.8%), (Found: M^+, 465.14913. $C_{22}H_{23}N_5O_9$ requires *M*, 465.14958); ν_{max} (thin film)/cm^{-1} 3439.9 and 3351.4 (NH_2), 1751.4 (C=O); δ_H (400MHz; $CDCl_3$) 1.85 (3H, s, 3'-CO_2CH_3), 2.01 (3H, s, 6'-CO_2CH_3), 2.04 (3H, s, 2'-CO_2CH_3), 2.06 (3H, s, 4'-CO_2CH_3), 4.02 (1H, ddd, *J* = 9.97, 4.56 and 2.05Hz, 5'-H), 4.10 (1H, dd, *J* = 12.31, 4.12Hz, 6'-H_a), 4.25 (1H, dd, *J* = 12.56 and 4.81Hz, 6'-H_b), 5.28 (1H, t, *J* = 9.76Hz, 4'-H), 5.40 (1H, t, *J* = 9.33Hz, 3'-H), 6.02 (1H, t, *J* = 9.26Hz, 2'-H), 6.08 (d, *J* = 9.34Hz, 1'-H), 6.25 (2H, s, NH_2), 8.07 (1H, s, 6-H) and 8.35 (1H, s, 3-H); δ_C (100MHz; $CDCl_3$) 20.3 (2'- CO_2CH_3), 20.6 (3'- and 4'- CO_2CH_3),

20.7 (6'- CO_2CH_3), 61.8 (C-6'), 67.9 (C-4'), 69.4 (C-2'), 73.7 (C-3'), 74.4 (C-5'), 81.4 (C-1'), 101.0 (C-3a), 133.6 (C-3), 155.1 (C-6), 156.3 (C-7a), 157.7 (C-4), 168.9 (C-2'CO_2CH_3), 169.4 (C-4'CO_2CH_3), 170.2 (C-3'CO_2CH_3) and 170.7 (C-6'CO_2CH_3); and

as a golden oil, *2-(3,4,2,6-tetra*-O-*acetyl-β-D-glucopyranosyl)pyrazolo[3,4*-d*]pyrimidine* **168a** (0.3672g, 46.2%), ν_{max} (thin film)/cm^{-1} 3439.9 and 3334.6 (NH_2), 1750.8 (C=O); δ_H (400MHz, $CDCl_3$) 1.83 (3H, s, 3'-CO_2CH_3), 2.00 (3H, s, 6'-CO_2CH_3), 2.02 (3H, s, 2'-CO_2CH_3), 2.04 (3H, s, 4'-CO_2CH_3), 4.07-4.10 (1H, m, 5'-H), 4.10 (1H, dd, J = 12.31 and 4.12Hz, 6'-H_a), 4.25 (1H, dd, J = 12.56 and 4.81Hz, 6'-H_b), 5.28 (1H, t, J = 9.76 Hz, 4'-H), 5.40 (1H, t, J = 9.33 Hz, 3'-H), 6.02 (1H, t, J = 9.26 Hz, 2'-H), 6.08 (d, J = 9.34 Hz, 1'-H), 6.25 (2H, s, NH_2), 8.07 (1H, s, 5-H) and 8.35 (1H, s, 3-H); δ_C (100MHz; $CDCl_3$) 20.3 (2'- CO_2CH_3), 20.5 (3'- and 4'- CO_2CH_3), 20.6 (6'- CO_2CH_3), 61.7 (C-6'), 67.9 (C-4'), 70.8 (C-2'), 72.7 (C-3'), 75.0 (C-5'), 88.0 (C-1'), 102.8 (C-3a), 123.7 (C-3), 157.3 (C-6), 159.5 (C-7a), 159.8 (C-4), 169.3 (2'-CO_2CH_3), 169.5 (4'-CO_2CH_3), 170.0 (3'-CO_2CH_3) and 170.7 (6'-CO_2CH_3).

1-(2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl)-4-hydroxypyrazolo[3,4-*d*]pyrimidine 167b[125] and 2-(2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl)-4-hydroxypyrazolo[3,4-*d*]pyrimidine 168b[125]

The experimental procedure employed for the synthesis of 1-[2,3,4,6-tetra-*O*-acetyl-D-glucopyranosyl]-4-aminopyrazolo[3,4-*d*]pyrimidine **167a** was followed, using allopurinol **109a** (0.2453g, 1.084mmol), 1,2,3,4,6-penta-*O*-acetyl-α-D-glucopyranose **137** (0.5473g, 1.658mmol), dry acetonitrile (13ml), $SnCl_4$ (0.6678g, 2.56mmol) and saturated aq. $NaHCO_3$ (40ml). The residual oil was purified by flash chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1:9)] to afford a white solid, which was purified further by recrystallization from ethanol to afford, as white fluffy crystals, 2-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-4-hydroxypyrazolo[3,4-*d*]pyrimidine **168b** (0.1107g, 14.3%), mp. 292-298°C (lit.[125] 285-287°C); ν_{max} (thin film)/cm^{-1} 3446 (OH), 1753 and 1742 (C=O); δ_H (400MHz; $CDCl_3$) 1.89 (3H, s, CO_2CH_3), 2.02 (3H, s, 4'-CO_2CH_3), 2.06 and 2.08 (6H, 2 x s, 3'- and 6'-CO_2CH_3), 3.99-4.02 (1H, m, 5'-H), 4.17 (1H, d, J = 12.7Hz, 6'-H_a), 4.30

(1H, dd, J = 4.6 and 12.7Hz, 6'-H_b), 5.27 (1H, t, J = 9.4Hz, 4'-H), 5.53 (1H, t, J = 9.2Hz, 3'-H), 5.49 (1H, t, J = 9.0Hz, 2'-H), 5.68 (1H, d, J = 8.8Hz, 1'-H), 7.95 (1H, s, 6-H), 8.38 (1H, s, 3-H) and 9.67 (1H, s, OH); δ_C (100MHz; $CDCl_3$) 20.1 (C-2'-CO_2CH_3), 20.5 (3'-CO_2CH_3), 20.7 (4'- and 6'- CO_2CH_3), 61.5 (C-6'), 67.7 (C-4'), 70.9 (C-2'), 72.6 (C-3'), 75.2 (C-5'), 88.1 (C-1'), 108.6 (C-3a), 127.7 (C-3), 146.1 (C-6), 158.7 (C-7a), 159.2 (C-4), 168.9 (2'-CO_2CH_3), 169.3 (4'-CO_2CH_3), 170.0 (3'-CO_2CH_3) and 170.5 (6'-CO_2CH_3).

The mother liquor was purified further by reverse phase HPLC [on a C-18 LUNA column; elution with MeOH-H_2O (2:1)] to afford, as a white solid, 1-[2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl]-4-hydroxypyrazolo[3,4-*d*]pyrimidine **167b** (10.6mg, 1.4%), mp. 222-224^{O}C (lit.[125] 222-223^{O}C); δ_H (400MHz; $CDCl_3$) 1.78 (3H, s, 2'-CO_2CH_3), 2.03, 2.05 and 2.06 (9H, 3 x s, 3'-, 4'- and 6'-CO_2CH_3), 3.98-4.02 (1H, m, 5'-H), 4.15 (1H, dd, J = 2.2 and 12.4Hz, 6'-H_a), 4.27 (1H, dd, J = 4.4 and 12.4Hz, 6'-H_b), 5.31 (1H, t, J = Hz, 4'-H), 5.42 (1H, t, J = Hz, 3'-H), 5.97-6.06 (2H, m, 1'- and 2'-H), 8. 03 (1H, s, 6-H), 8.16 (1H, s, 3-H) and 11.53 (1H, s, OH); δ_C (100MHz; $CDCl_3$) 20.3 (2'-CO_2CH_3), 20.6 (3'- and 4'CO_2CH_3), 20.7 (6'-CO_2CH_3), 61.7 (C-6'), 67.7 (C-4'), 69.2 (C-2'), 73.6 (C-3'), 74.6 (C-5'), 82.6 (C-1'), 107.0 (C-3a), 136.6 (C-3), 146.8 (C-6), 153.7 (C-7a), 158.8 (C-4), 168.6 (2'-CO_2CH_3), 169.3 (4'-CO_2CH_3), 170.3 (3'-CO_2CH_3) and (6'-CO_2CH_3), *m/z* 467 (M + 1, 100%).

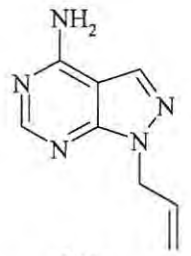
***2-(4-Aminopyrazolo[3,4-d]pyrimidin-1-yl)-6-(hydroxymethyl)tetrahydro-2H-pyran-2,3,4 -triol* 169a**

The experimental procedure employed for the synthesis of {1-[3,4,5-trihyoxy-6-(hydroxymethyl)tetrahydropyran-2-yl]-1*H*-indol-3-yl}acetic acid **139a** was followed, using sodium metal (2.02g), the tetraacetate **167a** (98.4mg, 0.212mmol), MeOH (60ml) and activated Dowex© 50WX8 (10ml). The residual solid was purified by reverse phase HPLC [on a C-18 LUNA column; elution with MeOH-H_2O (2:1)] to afford, as white crystals, *2-(4-aminopyrazolo[3,4-d]pyrimidin-1-yl)-6-(hydroxymethyl)tetrahydro-2H-pyran-2,3,4-triol* **169a** (59.5g, 94.4%), mp.>320^{O}C, (Found: M^+, 297.06237 $C_{11}H_{15}N_5O_5$ requires *M*, 297.10732); δ_H (400MHz; CD_3OD) 3.51-3.87 (5H, m, 5 x glucopyranose-

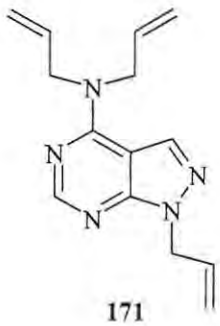
CH), 4.26 (1H, t, J = 8.8Hz, H), 5.74 (1H, d, J = 8.8Hz, 1'-H), 8.21 (1H, s, 6-H) and 8.54 (1H, s, 3-H); δ_C (100MHz; CD_3OD) 62.5 (C-6'), 71.1 (C-4'), 72.6 (C-2'), 78.8 (C-3'), 80.7 (C-5'), 85.7 (C-1'), 102.2 (C-3a), 135.0 (C-3), 155.8 (C-6), 157.2 (C-7a) and 159.9 (C-4); *m/z* 298 (M+1, 100%).

2-(4-Hydroxypyrazolo[3,4-*d*]pyrimidin-1-yl)-6-(hydroxymethyl)tetrahydro-2*H*-pyran-2,3,4-triol 169b

The experimental procedure employed for the synthesis of {1-[3,4,5-trihyoxy-6-(hydroxymethyl)tetrahydropyran-2-yl]-1*H*-indol-3-yl}acetic acid **139a** was followed, using sodium metal (2.27g), the tetraacetate **167b** (0.114g, 0.244mmol), MeOH (60ml) and activated Dowex© 50WX8 (10ml). The residual solid was purified by reverse phase HPLC [on a C-18 LUNA column; elution with MeOH-H_2O (2:1)] to afford, as white crystals, 2-(4-hydroxypyrazolo[3,4-*d*]pyrimidin-1-yl)-6-(hydroxymethyl)tetrahydro-2*H*-pyran-2,3,4,-triol **169b** (70.3g, 96.7%), mp.> 320^{O}C; δ_H (400MHz; D_2O) 3.60-3.99 (6H, m, 2'-,3'-,'4-', 5'- and 6'-H), 5.22 (1H, d, J = 9.2Hz, 1'-H), 7.96 (1H, s, 6-H) and 8.50 (1H, s, 3-H); δ_C (100MHz; D_2O) 61.0 (C-6'), 69.5 (C-4'), 71.7 (C-2'), 76.6 (C-3'), 78.7 (C-5'), 88.6 (C-1'), 107.0 (C-3a), 135.4 (C-3), 164.7 (C-7a) and 171.6 (C-4).



170



171

Attempted synthesis of *1-allyl-4-aminopyrazolo[3,4-*d*]pyrimidine* 170

Method A[184]

The experimental procedure employed for the synthesis of compound **166a** was followed, using 4-aminopyrazolo[3,4-*d*]pyrimidine **108a** (51mg, 0.377mmol), K_2CO_3 (104mg, 0.755mmol), allyl bromide (140mg, 1.16mmol) and DMF (2ml). The residual oil was

purified by flash chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1:9)] to afford, as a golden oil, N,N,*1-triallyl-4-aminopyrazolo[3,4-*d*]pyrimidine* **171** (11.2mg, 11.7%); δ_H (400MHz; $CDCl_3$) 4.35 (2H, t, J = 5.8Hz, $CH_2CH=CH_2$), 4.84 and 4.92 (4H, 2 x d, J = 9.4Hz, $CH_2CH=CH_2$), 5.39 and 5.47 (4H, 2 x d, J = 13.2Hz, 2 x $NCH_2CH=CH_2$), 5.20-5.35 (2H, m, $CH_2CH=CH_2$), 5.97-6.08 (3H, m, 3 x $CH_2CH=CH_2$), 8.19 (1H, s, 3-H) and 9.85 (1H, s, 6-H); δ_C (100MHz; $CDCl_3$) 44.4, 51.8 and 56.6 (NCH_2), 101.5 (C-3a), 118.3, 121.4 and 122.2 ($CH_2CH=CH_2$), 129.2, 130.2 and 131.7 ($CH_2CH=CH_2$), 133.2 (C-3), 150.1 (C-6 and C-7a) and 157.4 (C-4); *m/z* 256 (M + 1, 22.3%) and 214 (100%).

Method B[180]

A suspension of 4-aminopyrazolo[3,4-*d*]pyrimidine **108a** (49.6mg, 0.370mmol), allyl bromide (44.7mg, 0.370mmol) and K_2CO_3 (92.3mg, 0.668mmol) in reagent grade DMF (5ml) was stirred at room temperature for 65h. The reaction mixture filtered off and the DMF was evaporated *in vacuo*. The residual oil was purified by flash column chromatography [on silica gel, elution with EtOH-$CHCl_3$ (1:9)] to afford, as a golden oil, *N,N*,1-triallyl-4-aminopyrazolo[3,4-*d*]pyrimidine **171** (15.6mg, 16.5%).

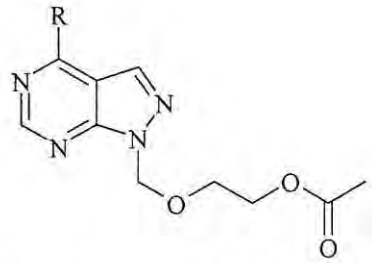
4-amino-1-[(2,2-dimethy-1,3-dioxolan-4-yl)methyl]pyrazolo[3,4-d]pyrimidine 172 [184]

The experimental procedure employed for the synthesis of compound **166a** was followed, using 4-aminopyrazolo[3,4-*d*]pyrimidine **108a** (49.9mg, 0.370mmol), DMF (2ml), 4-bromomethyl-2,2-dimethyl-1,3-dioxolane **128** (216.2mg, 1.11mmol) and K_2CO_3 (102mg, 0.740mmol). The reaction mixture was filtered while hot and the solid washed with DMF (3ml). The solvent from the DMF solution and washings was evaporated under reduced pressure and the residual brown oil was purified by column chromatography [on silica gel; elution with EtOH-$CHCl_3$-Hexane (1:1.5:2.5)] to afford, as a brown oil, 1-(2,2-dimethyl-1,3-dioxolan-4-ylmethyl)-4-aminopyrazolo[3,4-*d*]pyrimidine **172** (19.9mg, 21.6%), mp. 132-136^{O}C (lit.[184] 135-136^{O}C); (Found: M^+, 249.12198. $C_{11}H_{15}N_5O_2$ requires *M*, 249.12257); δ_H (400MHz; $CDCl_3$) 1.31 (3H, s, CH_3), 1.36 (3H, s, CH_3), 3.93 (1H, dd, J = 5.6 and 8.6Hz, OCH_a), 4.04 (1H, dd, J = 5.8 and 8.6Hz, OCH_b), 4.47 (1H,

dd, J = 7.6 and 15.6Hz, NCH_a), 4.57 (2H, m, NCH_b and OCH), 5.84 (2H, s, NH_2), 7.92 (1H, s, 3-H) and 8.36 (1H, s, 6-H); δ_C (100MHz, $CDCl_3$) 25.3 (CH_3), 26.7 (CH_3), 49.5 (NCH_2), 67.2 (OCH_2), 74.2 (OCH), 100.6 (C-3a), 109.9 [$C(CH_3)_2$], 131.0 (C-3), 154.0 (C-7a) , 155.9 (C-6) and 157.5 (C-4).

Attempted synthesis of 1-[(2-hydroxymethyl)pyrrolidin-1-nylmethyl]-4-hydroxypyrazolo[3,4-d]pyrimidine 174[185]

To a suspension of allopurinol **109a** (0.521g, 3.83mmol) in EtOH (8ml) was added prolinol **173** (0.769g, 7.60mmol) and the mixture was stirred for 20min at room temperature. To this mixture, 37% formaldehyde (0.7ml) was added and stirring was continued for a further 12h. The solid material was filtered off and the filtrate was evaporated *in vacuo*. The solid material was shown by ^1H NMR analysis to be allopurinol. On addition of $CHCl_3$ to the residual oil a formaldehyde polymer precipitated out; evaporation of the mother liquor was shown by ^1H NMR analysis to be impure prolinol **173**.



***1-[(2-Acetoxyethoxy)methyl]-4-aminopyrazolo[3,4-d]pyrimidine* 175a**

Method A[155]

To a solution of 4-aminopyrazolo[3,4-*d*]pyrimidine **108a** (50.3mg, 0.372mmol) in dry DMF (9ml), cooled to –63°C under N_2, NaH (10mg, 0.417mmol) was added slowly and the mixture was then stirred for 15mins. 2-(Bromomethoxy)ethyl acetate **131** (72.9mg, 0.370mmol) was added and the solution was warmed to room temperature and stirred for 48h. 1M aq. $NaHCO_3$ (0.5ml) was added and the solvent removed *in vacuo*. The residual oil was purified by flash chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1:9)]

to afford, a yellow oil, 1-*[(2-acetoxyethoxy)methyl]-4-aminopyrazolo[3,4-*d*]pyrimidine* **175** (13.5mg, 14.4%), ν_{max} (thin film)/cm^{-1} 3395.2 (NH_2) and 1729.0 (C=O) (Found: M^+, 251.10282. $C_{10}H_{13}N_5O_3$ requires *M*, 251.10184); δ_H (400MHz; $CDCl_3$) 2.01 (3H, s, CH_3), 3.77 (2H, t, *J* = 4.6Hz, OCH_2CH_2OAc), 4.15 (2H, t, *J* = 4.6Hz, OCH_2CH_2O), 5.78 (2H, s, NCH_2), 5.97 (2H, s, NH_2), 8.00 (1H, s, 6-H) and 8.43 (1H, s, 3-H); δ_C (100MHz, $CDCl_3$) 20.8 (CH_3), 63.1 ($AcOCH_2CH_2O$), 67.4 ($AcOCH_2CH_2O$), 75.7 (NCH_2O), 100.8 (C-3a), 132.2 (C-3), 155.9 (C-6), 156.1 (C-7a), 157.3 (C-4) and 171.0 (C=O).

Attempted synthesis of *1-[(2-Acetoxyethoxy)methyl]-4-hydroxypyrazolo[3,4-d]pyrimidine* 175b

Method A

The experimental procedure employed for the synthesis of 1-[(2-acetoxyethoxy)methyl]-4-aminopyrazolo[3,4-d]pyrimidine **175a** was followed, using allopurinol **109a** (0.104g, 0.761mmol), NaH (20.5mg, 0.852mmol), 2-(bromomethoxy)ethyl acetate **131** (0.153g, 0.777mmol) and DMF (18ml). The residual oil was purified by flash chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1:4)] to afford the precursor allopurinol **109a** and 2-(hydroxymethoxy)ethyl acetate.

Method B

A mixture of allopurinol **109a** (0.105g, 0.771mmol), K_2CO_3 (0.122g, 0.879mmol), 2-(bromomethoxy)ethyl acetate **131** (0.162g, 0.822mmol) and DMF (2ml) were placed under microwave radiation for 2h. The solid was filtered off and the filtrate evaporated *in vacuo*. The residual oil was purified by preparative layer chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1:4)] to afford allopurinol **109a**.

Attempted synthesis of *1-(2,2-dimethoxyethyl)-4-(tetrahydropyran-2-yloxy)-1H-pyrazolo[3,4-d]pyrimidine* 176

Method A

To a suspension of 4-(tetrahydropyran-2-yloxy)pyrazolo[3,4-*d*]pyrimidine **165** (150.5mg, 0.684mmol) in THF (20ml), cooled to 0^{o}C under nitrogen, NaH (33.8mg, 1.41mmol) was added slowly. The solution was warmed to room temperature and stirred for 15min. Chloroacetaldehyde dimethyl acetal (93.7mg, 0.752mmol) was added and the reaction mixture was stirred for a further 24h at room temperature. The reaction was quenched by the addition of saturated aq. $NaHCO_3$ and the solvent was removed *in vacuo*. The residue was re-dissolved in CH_2Cl_2 (10ml) and washed with brine (2 x 3ml). The organic layer was dried over anhydrous Na_2SO_4 and the solvent was removed *in vacuo* to afford the starting material **165**.

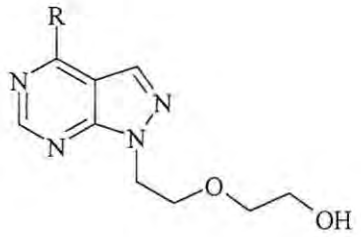
Method B

The experimental procedure employed for the synthesis of 1-allylbenzimidazole **141** was followed, using 4-(tetrahydropyran-2-yloxy)pyrazolo[3,4-*d*]pyrimidine **165** (0.149g, 0.677mmol), chloroacetaldehyde dimethyl acetal (95.5mg, 0.767mmol), NaH (36mg, 1.5mmol) and THF (20ml). The reaction was quenched by adding EtOAc and the solid material was filtered off. The solvent was removed *in vacuo* from the filtrate and analysis of both the solid material and residual oil revealed that it comprised the precursors, 4-(tetrahydropyran-2-yloxy)pyrazolo[3,4-*d*]pyrimidine **165** and chloroacetaldehyde dimethyl acetal.

Method C

The experimental procedure employed for the synthesis of 1-allylbenzimidazole **141** was followed, using 4-(tetrahydropyran-2-yloxy)pyrazolo[3,4-*d*]pyrimidine **165** (0.149g, 0.677mmol), chloroacetaldehyde dimethyl acetal (95.5mg, 0.767mmol), NaH (36mg, 1.5mmol) and DMF (20ml). Saturated aq. $NaHCO_3$ (5ml) was added and the solvent was evaporated *in vacuo*. The residual oil was re-dissolved in $CHCl_3$ (10ml) and washed with

brine (5ml). The organic layer was dried with Na_2SO_4 and the solvent was removed *in vacuo* to afford the starting material **165**.



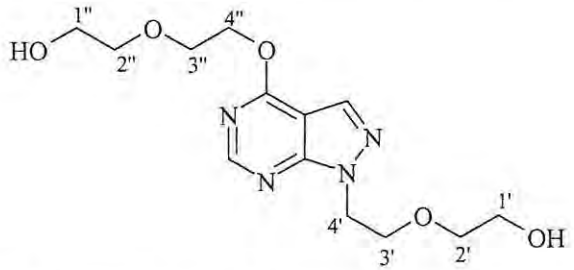
***2-{(4-aminopyrazolo[3,4-d]pyrimidin-1-yl)ethoxy}ethanol* 177a**

The experimental procedure employed for the synthesis of 1-allylbenzimidazole **141** was followed, using 4-aminopyrazolo[3,4-*d*]pyrimidine (0.102g, 0.755mmol), DMF (10ml), NaH (36.2mg, 1.51mmol) and (2-chloroethoxy)ethanol **148** (0.119g, 0.955mmol). The mixture was boiled under reflux for 8h and the reaction quenched by adding EtOAc. The residual oil was purified by flash chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1:5)] to afford, as a golden oil, *2-{(4-amino-pyrazolo[3,4-d]pyrimidinyl)ethoxy}ethanol* **177a** (40mg, 23.7%), (Found: M-1, 222.10143. $C_9H_{12}N_5O_2$ requires *M*, 222.09910); δ_H (400MHz; DMSO-d_6) 3.84 (2H, t, J = 5.6Hz, OCH_2CH_2OH), 4.42 (2H, t, J = 5.6Hz, OCH_2CH_2OH), 4.57 (1H, s, OH), 7.58 (2H, s, NH_2), 8.08 (1H, s, 3-H) and 8.17 (1H, s, 6-H); δ_C (100MHz, DMSO-d_6) 46.0 (NCH_2CH_2O), 60.0 (OCH_2CH_2OH), 68.2 (NCH_2CH_2O), 71.9 (OCH_2CH_2OH), 99.9 (C-3a), 132.5 (C-3), 154.7 (C-6), 155.7 (C-7a), 158.0 (C-4); *m/z* 224 (M+1, 100%).

Attempted synthesis of *2-{(4-tertbutyloxycarbonylpyrazolo[3,4-d]pyrimidin-1-yl) ethoxy}ethanol* 177c

The experimental procedure employed for the synthesis of 1-allylbenzimidazole **141** was followed using, 4-(*tert*butyloxycarbonyl)pyrazolo[3,4-*d*]pyrimidine **164b** (0.346g, 1.38mmol), NaH (66mg, 2.75mmol), 2-(2-chloroethoxy)ethanol **148** (0.172g, 1.38mmol) and DMF (5ml). The reaction mixture was stirred at room temperature for 3 days. The reaction was quenched by adding EtOAc and the residual oil was purified by flash

chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1:5)]. 1H NMR analysis indicated that none of the isolated fractions contained the desired product.



Synthesis of *4-[2-(2-Hydroxyethoxy)ethyloxy]-1-[(2-hydroxyethoxy)ethyl]pyrazolo[3,4-d]pyrimidine* 178

The experimental procedure employed for the synthesis of 1-allylbenzimidazole **141** was followed, using allopurinol **109a** (1.008g, 7.412mmol), NaH (194mg, 8.10mmol), 2-(2-chloroethoxy)ethanol (0.925g, 7.586mmol) and DMF (20ml). The residual oil was purified by flash chromatography [on silica gel; elution with EtOH-$CHCl_3$ (1;4)] to afford, as a colourless oil, *4-[2-(2-Hydroxyethoxy)ethyloxy]-1-[(2-hydroxyethoxy)ethyl]pyrazolo[3,4-d]pyrimidine* **178** (14mg, 0.6%), (Found: M^+, 312.14404. $C_{11}H_{15}N_5O_2$ requires *M*, 312.14337); δ_H (400MHz, $CDCl_3$) 3.59-3.62 (6H, m, 2x OCH_2CH_2OH), 3.67 (2H, m, OCH_2CH_2OH), 3.76 (2H, m, OCH_2CH_2OH), 3.92-3.97 (4H, m, NCH_2CH_2O and OCH_2CH_2O), 4.64, (2H, t, J = 5.2Hz, NCH_2CH_2O), 4.74 (2H, t, J = 5.2Hz, OCH_2CH_2O), 8.08 (1H, s, 3-H) and 8.52 (1H, s, 6-H); δ_C (100MHz, $CDCl_3$) 47.6 (C-4'), 61.6 (C-1'), 61.8 (C-1"), 66.2 (C-4"), 69.2 (C-3"), 69.6(C-3'), 72.6 (C-2'), 72.9 (C-2"), 102.9 (C-3a), 131.7 (C-3), 155.0 (C-6), 155.1 (C-7a) and 163.7 (C-4)

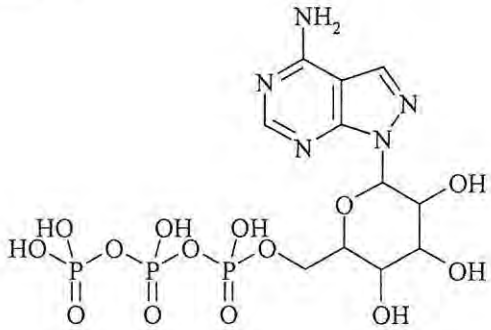
***2-{2-(4-aminopyrazolo[3,4-d]pyrimidin-1-yl)ethoxy}ethyl diethyl phosphate* 179**

Triethylamine (21.8mg, 0.215mmol) and diethyl chlorophosphate (37.1mg, 0.215mmol) was added to a solution of 2-{(4-amino-pyrazolo[3,4-*d*]pyrimidinyl)ethoxy}ethanol **177a** (40mg, 0.18mmol) in dry DMF (5ml) at 0^OC under argon. The reaction mixture was gradually warmed to room temperature and stirred overnight. The reaction was quenched by adding water and the solvent was removed *in vacuo*. The residual oil was redisolved

in CH_2Cl_2 (5ml) and washed with brine (2 x 2ml), and the organic phase dried with anhydrous $MgSO_4$. The solvent was evaporated in vacuo and the resulting oil was purified by radial chromatography [on silica; elution with EtOH-$CHCl_3$-hexane (1:1.5:2.5)] to afford, as a golden brown oil, *2-{2-(4-aminopyrazolo[3,4-*d*]pyrimidin-1-yl)ethoxy}ethyl diethyl phosphate* **179** (8.8mg, 13.6%); δ_H (400MHz; $CDCl_3$) 1.37 (6H, t, J = 5.8Hz, 2 x OCH_2CH_3), 4.05-4.12 (8H, m, OCH_2CH_2OP and 2 x OCH_2CH_3), 18-4.28 (4H, m, NCH_2CH_2O and 6.89 (2H, s, NH_2), 7.52 (1H, s, 3-H) and 7.69 (1H, s, 6-H); δ_C (100MHz; $CDCl_3$) 16.0 (d, J = 11.3Hz, 2 x OCH_2CH_3), 38.7 (NCH_2CH_2O), 63.5 (d, J = 11.3Hz, OCH_2CH_3), 65.2 (OCH_2CH_2OP), 68.2 (NCH_2CH_2O), 77.2 (OCH_2CH_2OP), 102.3 (C-3a), 128.8 (C-3), 130.8 (C-6) and 143.6 (C-4 and C-7a); δ_P (162MHz; $CDCl_3$) -17.4ppm; *m/z* 360 (M+1,10%) and 223 (100%).

Method B

The experimental procedure employed for the synthesis of 3-(benzimidazol-1-yl)propane-1,2-diol bis(diethyl phosphate) was followed, using 2-{(4-amino-pyrazolo[3,4-*d*]pyrimidinyl)ethoxy}ethanol 177a (100mg, 0.448mmol), BuLi (1.6M in hexane; 285µl, 0.457mmol), diethyl chlorophosphate (77.3mmol, 0.410mmol) and THF (10ml). To afford, as a yellow oil, *2-{2-(4-aminopyrazolo[3,4-*d*]pyrimidin-1-yl)ethoxy}ethyl diethyl phosphate* **179** (17.1mg, 11.6%).



Attempted synthesis of *2-(4-aminopyrazolo[3,4-d]pyrimidin-1-yl)-3,4,5-trihydroxy-6-(oxymethyl)tetrahydropyranyl triphosphate* 182

The experimental procedure employed for the attempted synthesis of 1-[(2-hydroxyethoxy)methyl]benzimidazole triphosphate **162a** was followed, using 2-(4-aminopyrazolo[3,4-*d*]pyrimidin-1-yl)-6-(hydroxymethyl)tetrahydro-2*H*-pyran-2,3,4-triol

169a (25mg, 0.088mmol), triethyl phosphate (0.6ml), proton sponge (28.8mg, 0.134mmol), $POCl_3$ (14.8mg, 0.097mmol), tributylammonium pyrophosphate (158.6mg, 0.44mmol), dry DMF (0.9ml), tributylamine (82.5g, 0.44mmol) and 1M TEAB (11ml). The fractions obtained from the ion exchange separation did not appear to contain the desired product.

Molecular modelling

Molecular modelling was performed on a Silicon Graphics O^2 workstation using the MSI CERIUS^2 version 4.5 modelling platform and the UNIVERSAL 1.02 force field. Molecular dynamics (MD) were performed using the Cerius^2 modelling environment and a quenched dynamics in an NVE (constant volume and energy) ensemble at 500K and a time step of 0.001ps. The potential energy was calculated using the UNIVERSAL 1.02 force field and the atomic charges were calculated using the Qeq method (charge equilibration approach). The 10 highest and 10 lowest conformers were identified and minimized to find the global minimum structure in each case. The Accelrys LIGAND FIT module was used to explore receptor docking interactions.

4. References

1. http://www.wits.ac.za/myco/htm/htrtmt.htm
2. C. Metcalf in *A Century of Tuberculosis South African Perspective*, ed. H. M. Coovadia and S. R. Benatar, Oxford University Press, **1991**, pp.1-31.
3. Y. Zang and L. M. Amzel, *Curr. Drug Targets*, **2002**, 3, 131-154.
4. World Health Organization. Treatment of Tuberculosis: guidelines for national programmes, WHO/CDS/TB 2003.313. http://whqlildoc.who.int/hq/2003/WHO_CDS_TB_2003.313_eng.pdf
5. World Health Organization, Emergence of XDR-TB. http://www.who.int/mediacentre/news/notes/2006/np23/en/print.html
6. C.C. Evans in *Clinical Tuberculosis*, ed. P. D. O. Davis, Chapman & Hall, London, **1998**, p.1.
7. T. L. Lemke in *Foye's Principles of Medicinal Chemistry*, ed. D. A. Williams and T. L. Lemke, 5th ed., Lippincot, Williams & Wilkins, New York, **2002**, pp.904-923.
8. T. D. Lowry, *Rev. Med. Chem.*, **2003**, 3, 689-702.
9. L. G. Dover, A. M. Cerdeño-Tarraga M. J. Pallen, J. Parkhill and G. S. Besra, *FEMS Microbiology Reviews*, **2004**, 28. 225-250
10. P.J. Brennan, *Tuberculosis*, **2003**, 83, 91-97
11. L. A. Basso and D. S. Santos, *Med. Chem. Rev.*, **2005**, 2, 393-413
12. J. M. Grange in *Clinical Tuberculosis*, ed. P. D. O. Davis, Chapman & Hall, London, **1998**, pp.129-152.
13. N. M. Parrish, J. D. Dick and W. R. Bishai, *Trends in Micro.*, **1998**, 6, 107-111
14. S.A. Waksman and A.T. Henrici, *J. Bact.*, **1943**, 46, 337.
15. J. Lehmann, *Lancet*, **1946**, 250, 15-16.
16. F. Bernheim, *Science*, **1946**, 94, 204.
17. H. H. Fox, *Science*, **1952**, 116, 129-135.
18. J. Bernstein, W. A. Lott, B. A. Steinberg and H. L. Yale, *Am. Rev. Tuberc.*, **1952**, 65, 357-364.

19. W. Steenken Jr. and E. Wolinsky, *Am. Rev. Tuberc.*, **1952**, 65, 365-375.
20. R. L. Yieger, W. G. Munroe and F. I. Dessau, *Am. Rev. Tuberc.*, **1952**, 65, 523-534.
21. F. A. Kuehl Jr., F. J. Wolf, N. R. Trenner, R. L. Peck, R. H. Buhs, I. Putter, R. Ormond, J. E. Lyons, L. Chaiet, E. Howe, B. D. Hunnewell, G. Downing, E. Newstead and K. Folkes, *J. Am. Chem. Soc.*, **1955**, 77, 2344-2345.
22. R. G. Wilkinson, R. G. Shepherd, J. P. Thomas and C. Baughn, *J. Am. Chem. Soc.*, **1961**, 83, 2212-2213.
23. M. N. Shchukina, *Pharm. Chem. J.*, **1992**, 4, 208-212
24. Characteristics of the Main Anti- tuberculosis drugs, http://www.Hopkins-abxguide.org.
25. E.B. Herr, R. L. Hamill and J. M. Mcguire, *Patent* 920563.
26. Marcus Vinícius Nora De Souza, *Rev. Med. Chem.*, **2005**, 5, 1009-1017.
27. American Thoracic Society, Centers for Disease control and Prevention Infectious disease Society of America Treatment of Tuberculosis, *Ame. J. Respir. Crit. Care Med*, **2003**, 167, 603-662.
28. T. R. Frieden, S. S.Munsiff, T.R Sterling, C. J. Watt and C. Dye, *Lancet*, **2003**, 362, 887-899.
29. J. S. Mukherjee, M. L. Rich, A. R. Socci, J. K. Joseph, F. A. Virú, S. S. Shin, J. J. Furin, M. C. Becerra, D. J. Barry, J. Yong Kim, J. Bayona P. Farmer, M. C. Smith Fawzi and K. J. Seung, *Lancet*, **2004**, 363, 474-481.
30. A. Sokomoskovi, L.M. Parsons and M. Salfinger, *Respir. Res.*, **2001**, 2, 164-168.
31. T. Urbanski, Z. Biernacki, D. Gurne, L. Halski, M. Mioduszewska, B. Serafinowa, J. Urbanski and D. Zelazko, *Roczniki Chemii*, **1953**, 27, 161-166; *CAN* 48:77640.
32. R. L. Malan and P. M. Dean, *J. Am. Chem. Soc.*, **1947**, 67, 1797-1798.
33. M. M. Baizer, M. Dub, S. Gister and N. G. Steinberg, *J. Amer, Pharm, Assoc. (1912-1977)*, **1956**, 45, 478-480; *CAN* 51:1816.
34. J. Halt, J. Andreassen, J. M. Bakke and A. Fiksdahl, *J. Het. Chem.*, **2005**, 42, 259-264.

35. H. Rinderknecht, *Helv. Chim. Acta*, **1959**, 42, 1324-1327; *CAN* 54:2235.
36. Y. S. Karpman and L. N. Yakhontov, *Khim-Farm. Zhur.*, **1980**, 14, 93-96; *CAN* 93:46368.
37. H. A. Staab, M. Lueking and F. H. Duerr, *Ber*, **1962**, 95, 1275-1283; *CAN* 57:29635.
38. G. D. Yadav, S. S. Joshi, and P. S. Lathi, *Enzyme Microb. Tech.*, **2005**, 36, 217-222.
39. A. Sauciuc, M. Rusu, I. Ciobanu, A. Neacşu and C. Poncu, *Patent RO 101892*, 1992.
40. M. Zhu, E. Ruijter and L. A. Wessjohann, *Org. Let.* **2004**, 6, 3921-3924.
41. S. A. Hall and P. E. Spoerri, *J. Am. Chem. Soc.*, **1940**, 62, 664-665.
42. D. Lednicer and L. A. Meister, *The Organic Chemistry of Drug Synthesis*, **1988**, 1, John Wiley & Sons, New York, p.277.
43. S. Kushner, H. Dallian, J. L. Sunjaro, F. L. Bach, S. R. Safir, V. K. Smith and J. H. Williams, *J. Am. Chem. Soc.*, **1952**, 74, 3617.
44. A. R. Katritzky, Hai-Ying He and K. Suzuki, *J. Org. Chem.*, **2000**, 65, 8210-8213.
45. a) V. M. Bondareva, T. V. Andrushkevich, O. B. Lapina, *Catal. Today*, **2000**, 61, 173-178; b) V.M. Bondareva, T. V. Andrushkevich, O. B. Lapina, A. A. Vlasov and L. S. Dovlitova, *React. Kinet. Catal. Lett.*, **2003**, 79, 165-173; c) V. M. Bondareva, T. V. Andrushkevich, O. B. Lapina, D. F. Khabibulin, A. A. Vlasov, L. S. Dovlitova, and E. B. Burgin, *Kinet. Catal.*, **2004**, 45, 104-113.
46. a) F. Minisci, A. Citterio and E. Vismara, *Tetrahedron*, **1985**, 41, 4151-4170; b) F. Minisci, F. Recupero, C. Punta, C. Gambarotti, F. Antonietti, F. Fontana and G. F. Pedulli, *Chem. Commun.*, **2002**, 2496-2497.
47. L. Deng, K. Mikušovă, K. G. Robuck, M. Scherman, P. J. Brennan and M. R. McNeil *Antimicrob. Agents Chemother.*, **1974**, 96, 920-921.
48. A. M. Kritsyn, A. M. Likhosherstov, T. V. Protopopova and A. P. Skoldinov, *Dokl. Akad. Nauk SSSR*, **1962**, 145, 332-335; *Chem. Abstr.*, 57,14920a.
49. S. Umezawa, Y. Takahashi, T. Usui and T. Tsuchiya, *J. Antibiotics*, **1974**, 27, 997-999.

50. Umezawa, T. Tsuchiya, T. Yamasaki, H. Sano and Y. Takahashi, *J. Amer. Chem. Soc.*, **1974**, 96, 920-921.
51. C. Sannie and H. Lapin, *B. Soc. Chim. Fr.*, **1950**, 322-326; *CAN* 44:46762.
52. S. Hillers, A. Lokenbachs and L. Majs, *Latvijas PSR Zinatnu Akademijas Vestis*, **1950**, 32, 7-25; *CAN* 48:56575.
53. J. T. Sheehan, *J. Amer. Chem. Soc.*, **1948**, 70, 1665-1666.
54. D. J. Drain. D. D. Martin, B. W. Mitchell, D. E. Seymour and F. S. Spring, *J. Chem. Soc*, **1948**. 1498-1502.
55. L. Doub, J. A. Schaefer, O. L. Stevenson, C. T. Walker and J. M. Vandenbelt, *J. Am. Chem. Soc.*, **1958**, 23, 1422-1425.
56. P.H. Hidy, E. B. Hodge, V. V. Young, R. L. Harned, G. A. Brewer, W. F. Philipps, W. F. Runge, H. E. Stavely, A. Pohland, H. Boaz and H. R. Sullivan, *J. Am. Chem. Soc.*, **1955**, 77, 2345-2346.
57. C. H. Stammer, A. N. Wilson, F. W. Holly and K. Folkers, *J. Am. Chem. Soc.*, **1955**, 77, 2346-2347.
58. C. H. Stammer, A. N. Wilson, C. F. Spencer, F. W. Bachelor, F. W. Holly and K. Folkers, *J. Am. Chem. Soc.*, **1957**, 79, 3236-3240.
59. a) E. M. Fry, *J. Org. Chem.*, **1949**, 14, 887; b) E. M. Fry, *J. Org. Chem.*, **1950**, 15, 438, 802.
60. D. Lednicer and L. A. Meister, *The Organic Chemistry of Drug Synthesis*, **1988**, 1, John Wiley & Sons, New York, p.255.
61. N. F. Kucherova, R. M. Khomutov, E. I. Budovskii, V. P. Evdakov and N. K. Kochetkov, *Zh. Obshch. Khim.*, **1959**, 29, 915-919; *CAN* 54:7229.
62. H-G. Schwark and D. Schunke, *Patent 1443073*, **1966.**
63. J. M. Nascimento and M. H. Venda, *Farmacia*, **1962**, 12, 321-328; *CAN* 59:3359.
64. B. T. King, *Brit. Patent GB1122111*, **1968.**
65. T. Shiba, S. Nomoto, T. Teshima and T. Wakamiya, *Tetrahedron*, **1976**, 43, 3907-3910, b) S. Nomoto, T. Teshima, T. Wakamiya and T. Shiba, *Tetrahedron*, **1978**, 34, 921- 927.
66. D. E. DeMong and R. M. Williams, *J. Am. Chem. Soc.*, **2003**, 125, 8581-8585.

67. B. Puezer, W. E. Hamlin and L. Katz, *J. Am. Chem. Soc.*, **1957**,73,2958.
68. W. Salmer and E. Goth, *U.S. Patent 2621209*, **1952**.
69. Farbenfabriken Bayer, *Patent 670402*, **1952**.
70. A. M. Hay, S. Hobbs-Dewitt, A. A. Macdonald and R. Ramage, *Synthesis*, **1999**, 11, 1979-1985.
71. U.R. Kalkote, V. T. Sathe, R. K. Kharul, S. P. Chavan and T. Ravindranathan, *Tetrahendron*, **1996**, 37, 6785- 6786.
72. S. Radhl and D. Bouzard, *Heterocycles*, **1992**, 11, 2143-2177.
73. S. B. Kang, S. Park, Y. H. Kim and Y. Kim, *Heterocycles*, **1997**, 45, 137-145.
74 J. Adrio, J. C. Carretero, J. L. G. Ruano, A. Pallarés and M. Vicioso, *Heterocycles*, **1999**, 51, 1563-1572.
75 Y. R. Kumar, V.V.N.K.V. Prasad Raju, R. R. Kumar, S. Eswaraiah, K. Mukkanti, M.V. Suryanarayana, M. S. Reddy, *J. Pharmaceut. Biomed.*, **2004**, 34, 1125-1129.
76 J. Clayden, N. Greeves, S. Warren and P. Wothers, *Organic Chemistry*, Oxford University Press, New York, **2001**, 595-597.
77 S. T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, S. Gas, K. Eiglmeier, C. E. Barry 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels and B. G. Barrell, *Nature*, **1998**, 393, 537-544.
78 V. Sharma, S. Sharma, K. Hoener zu Bentrup, J. D. Mckinney, D. G. Russell, W. R. Jacobs *Jr.* and J. C. Sacchettini, *Nat. Struct. Biol.*, **2000**, 7, 663-668.
79 J. D. McKinney, K. Hoener zu Bentrup, E. J. Muñoz-Elías, A. Miczak, B. Chen, W-T. Chan, D. Swenson, J. C. Sacchettinik, W. R. Jacobs *Jr* and D. G. Russel, *Nature*, **2000**, 406, 735-738.
80 M. Daffé, *Trends Microbiol.*, **2000**, 8, 438-440.
81 A. Ginsburg and E. R. Stadman, in *The Enzymes of Glutamine Metabolism*, ed. S. Prusiner and E. R. Stadman, Academic Press, New York, **1973**, pp.9-43.
82 L. Reitzer, *Annu. Rev. Microbiol.*, **2003**, 57, 155-76.

83 Y. Wang, J. Kudoh, R. Kubota, S. Asakawa, S. Minoshima and N. Shimuzu, *Genomics*, **1996**, 37, 195-199.

84 M. V. Tullius, G. Harth and M. A. Horwitz, *Infect. Immun.*, **2003**, 71, 3927-3936.

85 G. Harth, S. Masleša-Galić, M. V. Tullius and M. A. Horwitz, *Mol. Micro.*, **2005**, 58, 1157-1172.

86 D. Eisenberg, H. S. Gill, G. M. U. Pfluegl and S. H. Rotstein, *Biochim. Biophys. Acta*, **2000**, 1477, 122-145.

87 J. B. Hunt and A. Ginsburg, *J. Biol. Chem.*, **1980**, 255, 590-594.

88 G. Harth, D. L. Clemens and M. A. Horwitz, *Proc. Natl. Acad. Sci. USA*, **1994**, 91, 9342-9346.

89 H. S. Gill, G. M. U. Pfluegl and D. Eisenberg, *Biochemistry*, **2002**, 41, 9863-9872.

90 H. S. Gill and D. Eisenberg, *Biochemistry*, **2001**, 40, 1903-1912.

91 J. Moss, S. J. Stanley and R. L. Levine, *J. Biol. Chem.*, **1990**, 34, 21056-21060.

92 G. Harth, and M. A. Horwitz, *J. Exp. Med.*, **1999**, 189, 1425-1435.

93 I. S. Boksha, H. J. Schönfeld, H. Langen, F. Müller E. B. Tereshkina and G. Sh. Burbaeva, *Biochemistry (Moscow)*, **2002**, 67, 1012-1020.

94 S. S. Tate and A. Meister in *The Enzymes of Glutamine Metabolism*, ad. S. Prusiner and E. R. Stadman, Academic Press, New York, **1973**, pp.77-127.

95 S. R. Robinson, *Neurochem. Int.*, **2000**, 36, 471-482.

96 G. Sh. Burbaeva, I. S. Boksha, M. S. Turishcheva, E. A. Vorobyeva, O. K. Savushkina and E. B. Tereshkina, *Prog. Neuro-Psychopharmacol. Biol. Psychiatry*, **2003**, 27, 675-680.

97 W. W. Krajewski, T. A. Jones and S. L. Mowbray, *Proc. Natl. Acad. Sci. USA*, **2005**, 102, 10499-10504.

98 R. Mehta, J. T. Pearson, S. Maharajan, A. Nathi, M. J. Hickeys, D. R. Shermans and W. M. Atkins, *J. Biol. Chem.*, **2004**, 21, 22477-22482.

99 T. Parish and N. G. Stoker, *J. Bacteriol.*, **2000**, 5715-5720.

100 G. Harth and M. A. Horwitz, *Infect. Immun.*, **2003**, 71, 456-464.

101 A. Shrake, E. J. Whitley *Jr.* and A. Ginsburg, *J. Biol. Chem.*, **1980**, 255, 581-589.

102 G. Harth, P. C. Zamecnik, Jin-Yan Tang, D. Tabatadze and M. A. Horwitz, *Proc. Natl. Acad. Sci. USA*, **2000**, 97, 418-423.

103 O. Moukha-Chafiq, M. L. Taha, H. B. Lazrek, C. Pannecouque, M. Witvrouw, E. De Clercq, J. L. Barascut and J. L. Imbach, *Nucleos. Nucleot. Nucl.*, **2001**, 20, 1797-1810.

104 T. Iwaki, Y. Fujita, F. Yamada and M. Somei, *Heterocycles*, **2003**, 60, 1411-1418.

105 A. M. Pignone, A. Del Rosso, G. Fiori, Marco, Matucci-Cerinic, A. Becucci, A. Tempestini, R. Livi, S. Generini, L. Gramigna, C. Benvenuti, A. M. Carossino, M. L. Conforti and F. Perfetto, *J. Pineal Res.*, **2006**, 41, 95-100.

106 E. Mills, P. Wu, D. Seely and G. Guyatt, *J. Pineal Res.*, **2005**, 39, 360-366.

107 S. Hess, U. Teubert, J. Ortwein and K. Eger, *Eur. J. Pharm. Sci.*, **2001**, 14, 3001-311.

108 D. Lednicer and L. A. Mitscher, *The Organic Chemistry of Drug Synthesis*, **1988**, 1, John Wiley & Sons, New York, pp.107-108.

109 O. Greco, L. K. Folkes, P. Wardman, G. M. Tozer and G. U. Dachs, *Cancer Gene Ther.*, **2000**, 7, 1414-1420.

110 a) L. K. Folkes and P. Wardman, *Biochem. Pharmacol.*, **2001**, 61, 129-136; b) L. K. Folkes, O. Greco, G. U. Dachs,M. R. L. Stratford and P. Wardman, *Biochem. Pharmacol.*, **2002**, 63, 265-272; c) L. K. and P. Wardman, *Cancer Research*, **2003**, 63, 776-779.

111 D-S. Kim, S-E. Jeon and K-C. Park, *Cell. Signal.*, **2004**, 16, 81-88.

112 M. P. de Melo, T. C. Pithon-Curi and R. Curi, *Life Sciences*, **2005**, 74, 1713-1725

113 S. Rossiter, L. K. Folkes and P. Wardman, *Bioorg. Med. Chem. Lett.*, **2002**, 12, 2523-2526.

114 Y-J. Chyan, B. Poeggeler, R. A. Omar, D. G. Chain, B. Frangione, J. Ghiso and M. A. Pappolla, *J. Biol. Chem.*, **1999**, 274, 21937-21942.

115 S. R. Lefler, S. D. St. Peter, M. I. Rodriguez-Davalos, J. C. Perumean, N. A. Spence, A. A. Moss and D. C. Mulligan, *Transplant. P.*, **2002**, 34, 3065-3066.

116 I. A. G. Reilly and G. A. FitzGerald, *Blood*, **1987**, 69, 180-188.

117 W. W. Wilkerson, A. A. Kergaye and S. W. Tam, *J. Med. Chem.*, **1993**, 36, 2899-2907.

118 G.C Cook, *Parasitol. Today*, **1990**, 6, 113-136.

119 Z. Kadry, E. C. Renner, L. M. Bachmann, N. Attigah, E. L. Renner, R. W. Ammann and P. A. Clavien, *Brit. J. Surg.*, **2005**, 92, 1110-1116.

120 M. W. Robinson, N. McFerran, A. Trudgett, L. Hoey and I. Fairweather, *J. Mol. Graph. Model.*, **2004**, 23, 275-284.

121 M. A. Bradley and J. Horton, T. Roy. Soc. Trop. Med. H., **2001**, 95, 72-73.

122 R. O. McCracken, *J. Parasitol.*, **1978**, 64, 21.

123 J. B. Camden, *Patent WO98/51303*, **1998**.

124 C. L. Moore, M. Chiaramonte, T. Higgins and R. D. Kuchta, *Biochemistry*, **2002**, 41, 14066-14075.

125 S. M. Jain, A. Kumar, B. Purnima, K. K. Anand, A. K. Saxena and C. K. Atal, *Indian J. Chem.*, **1988**, 27B, 390-393.

126 S. Akhondzadeh, A. Safarcherati and H. Amini, *Prog. Neuro-Psychoph*, **2005**, 29, 253-259.

127 M. Moroni and C. Porta, *Cancer Chemother. Pharmacol.*, **1997**, 47, 171-172.

128 C. L. Loprinzi, S. G. Cianflone, A. M. Dose, P. S. Etzell, N. L. Burnham, T. M. Therneau, L-R. Hagen, D. K. Gainey, M. Cross, L. M. Athmann, T. Fischer and M. J. O'Connell, *Cancer*, **1990**, 65, 1879-1882.

129 T. Hanawa, N. Masuda, K. Mohri, K. Kawata, M. Suzuki and S. Nakajima, *Drug Dev Ind. Pharm.*, **2004**, 30, 151-161.

130 S. I. Khan, A. Mishra, P. Y. Guru, R. Pratap and D. S. Bhakuni, *Indian J. Chem.*, **1990**, 29B, 40-46.

131 F. Gatta, F. Perotti, L. Gordon, M. Gramiccia, S. Orsini, G. Palazzo and V. Rossi, *Eur. J. Med. Chem.*, **1990**, 25, 419-424.

132 F. Gatta, L. Gordon, E. Lupardini, M. Gramiccia, and S. Orsini, *Il Farmaco*, **1991**, 46, 75-84.

133 K. L. Seley and S. W. Schneller, *J. Med. Chem.*, **1997**, 40, 625-629.

134 G. A. Bhat, J-L. G. Montero, R. P. Panzica, L. L. Wotring and L. B. Townsend, *J. Med. Chem.*, **1981**, 24, 1166-1172.

135 F. Da Settimo, G. Primofiore, C. La Motta, S. Taliani, F. Simorini, A. M. Marini, L. Mugnaini, A. Lavecchia, E. Novellino, D. Tuscano and C. Martini, *J. Med. Chem.*, **2005**, 48, 5162-5174.

136 T. Lübbers, P. Angehrn, H. Gmünder, S. Herzig and J. Kulhanek, *Bioorg. Med. Chem. Lett.*, **2000**, 10, 821-826.

137 G. P. P. Kuntz, A. Glassman, C. Cooper and T. J. Swift, *Biochemistry*, **1972**, 11, 538-540.

138 V. Grumel, J.-Y. Mérour, B. Lesur, T. Giboulot, A. Frydman and G. Guillaumet, *Eur. J. Med. Chem.*, **2002**, 37, 45-62.

139 S. C. Benson, L. Lee, L. Yang and J. K. Snyder, *Tetrahedron*, **2000**, 56, 1165-1180.

140 M. Yato, K. Homma and A. Ishida, *Tetrahedron*, **2001**, 57, 5353-5359.

141 M. Lalonde and T. H. Chan, *Synthesis*, **1985**, 817-845.

142 A. Wissner and C. V. Grudzinskas, *J. Org. Chem.*, **1978**, 43, 3972-3974.

143 S. Jiang and G. Singh, *Tetrahedron*, **1998**, 54, 4697-4753 and references cited therein.

144 C. E. Braun and C. D. Cook, *Org. Synth. Coll.*, **1973**, 5, 887-889.

145 B, R. Adams and L. H. Ulich, *J. Amer. Chem. Soc.*, 1920, **42**, 599-611.

146 K. Teranishi, S. Hayashi, S. Nakatsuka, T. Goto, *Synthesis*, **1995**, 506-508.

147 O. Ottoni, R. Cruz and R. Alves, *Tetrahedron*, **1998**, 53, 13916- 13928.

148 K. Teranishi, S. Nakatsuka, T. Goto, *Synthesis*, **1994**, 1018-1020.

149 A. P. Dobbs, K. Jones and K. T. Veal, *Tet. Lett.*, **1997**, 38, 5379-5382.

150 D. H. R. Barton, J. Cs. Jaszberenyi, W. Liu and T. Shinada, *Tetrahedron,* **1996**, 52, 2717-2726.

151 M. Renoll and M. S. Newman, *Org. Synth. Coll.*, **1948**, 28, 73-75.

152 C. J. Marasco *Jr.*, C. Piantadosi, K. L. Meyer, S. Morris-Natschke, K. S. Ishaq, G. W. Small and L. W. Danielt, *J. Med. Chem.*, **1990**, 33, 985-992.

153 H. Gao and A. H. Mitra, *Synthesis*, **2000**, 329-351.

154 J. L. Kelly, M. P. Krochmal and H. J. Schaeffer, *J. Med. Chem.*, **1981**, 24, 1528-1531.

155 M. J. Robins and P. W. Hatfield, *Can. J. Chem.*, **1982**, 60, 547-553.

156 M. M. Faul and C. A. Krumrich, *J. Org. Chem.*, **2001**, 66, 2024-2033.

157 J. Brűning, T. Hache and E. Winterfeld, *Synthesis*, **1994** 25-27.

158 L. B. Townsend and R. V. Revenkar, *Chem. Rev.*, **1970**, 70, 389-438.

159 B. S. Furniss, A. J. Hannaford, P. N. Smith and A. R. Tatchell, *Vogel's Textbook of Practical Organic Chemistry*, 5th ed, Longman Scientific & Technical, Singapore, **1989**, p.644.

160 M. Gallant, J. T. Link and S. J. Danishefsky, *J. Org. Chem.*, **1993**, 58,343-9.

161 B. C. Kraybill, L. L. Elkin, J. D. Blethrow, D. O. Morgan and K. M. Shokat, *J. Am. Chem. Soc.*, **2002**, 124, 1218-1228.

162 J. Budka, M. Tkadlecová, P. Lhoták and I. Stibor, *Tetrahedron*, **2000**, 56, 1883-1887.

163 R. Hirschman, K. C. Nicolaou, S. Pietranico, E. L. Leahy, J. Salvino, B. Arison, M. A, Cichy, P. G. Spoors, W. C. Shakespeare, P. A. Sprengeler, P. Hamley, A. B. Smith III, T. Reisine, K. Raynor, L. Maechler, C. Donaldson, W. Vale, R. M. Freidinger, M. R. Cascieri and C. D. Strader, *J. Am. Chem. Soc.*, **1993**, 115, 12550-12568.

164 A. Z. Haikal, El S. H. El Ashry and J. Banoub, *Carbohyd. Res.*, 338, **2003**, 2291-2299.

165 K. L. Tan, A. Vasudevan, R. G. Bergman, J. A. Ellman and A. J. Souers, *Org. Lett.*, **2003**, 5, 2131-2134.

166 R. J. Fessenden and J. S. Fessenden, *Techniques and Experiments for Organic Chemistry*, Willard Grant Press, Boston, **1983**, 248-251.

167 V. Bushan, R. Rathore and S. Chandraskeran, *Synthesis*, **1984**, 431-433.

168 L. De Luca, G. Giacornelli and A. Porcheddu, *Org. Lett.*, **2001**, 3, 3041-3043.

169 T. A. Giudici and J. J. Griffin, *Carbohyd. Res.*, **1974**, 33,287-295.

170 R. Ziessel, P. Nguyen, L. Douce, M. Cesario and C. Estournes, *Org, Lett.*, **2004**, 6, 2865-2868.

171 T. H. Fife, R. Natarajan and M. H. Werner, *J. Org. Chem.*, **1987**, 52, 740-746.

172 T. H. Fife and T. J. Przystas, *J. Am. Chem. Soc.*, **1986**, 108, 4631-4636.

173 T. Kanayama, K. Yoshida, H. Miyabe, T. Kamichi and Y. Takemoto, *J. Org. Chem.*, **2003**, 68, 6197-6201.

174 S. Nampalli and S. Kumar, *Bioorg. Med. Chem. Lett.*, **2000**, 10, 1677-1679.

175 T. Kovács and Őtvős, *Tetrahedron*, **1988**, 29, 4525-4528.

176 X. Zhang, I. Lee, X. Zhou and A. J. Berdis, *J. Am. Chem. Soc.*, **2006**, 128, 143-149.

177 R. M. Silverstein and F. X. Webster, *Spectrometric Identification of Organic Compounds*, 6th ed., John Wiley & Sons Inc., New York, pp.291-295.

178 F. Bergmann, A. Frank and Z. Neiman, *J. Chem. Soc. Perkin I*, **1979**, 2795-2802.

179 M. K. Shukla and P.C. Mishra, *Spectrochim. Acta A*, **1996**, 52, 1547-1557.

180 J. J. Van Veldhuizen, S. B. Garber, J. S. Kingsbury and A. H. Hoveyda, *J. Am. Chem. Soc.*, **2002**, 124, 4954-4955.

181 A. Hussain and J. H. Rytting, *J. Pharm. Sci.*, **1974**, 63, 798-799.

182 M. El Hedi Jellali, N. Van Bac and N. Dat-Xuong, *Tetrahedron*, **1975**, 31, 587-591.

183 C. Kashima, H. Harada, I. Isanobu, I. Fukuchi and A. Hosomi, *Synthesis*, **1994**, 61-65.

184 A. Holy, I. Rosenberg and H. Dvořáková, *Collect. Czech. Chem. Commun.*, **1989**, 54, 2470-2501.

185 S. M. Sicardi, H. N. Pappa and J. E. Koch, *Acta Farm, Bonaerense*, **1986**, 5, 73-77.

186 A. J. Brandolini, *J. Am. Chem. Soc.*, **2006**, 128, 13313.

187 W. Bremser, *Anal. Chim. Acta*, **1978**, 103, 355-365.

188 G. V. De Lucca, J. Liang, P. E. Aldrich, J. Calabrese, B. Cordova, R. M. Klade, M. M. Rayner, and C. H. Chang, *J. Med. Chem.*, **1997**, 40, 1707.

189 I. M. A. Gledhill, unpublished work.

190 S-H. Liaw, G. Jun and D. Eisenberg, *Biochemistry*, **1994**, 33, 11184-11188.

191 H. Budzikiewicks, C. Djerassi and D. H. Williams, *Mass Spectrometry of Organic Compounds*, Holden-day, Inc., San Francisco, **1967**, p.611.

192 J. G. Rodrígues, D. Urrutia and L. Canoira, *Int. J. Mass Specrom.*, **1996**, 152, 97-110.

193 V. U. Khuzhaev, *Chem. Nat. Prod.*, **2004**, 40, 196-197.

194 M. R. Grimmett in *Comprehensive Heterocyclic Chemistry*, ed. A.R. Katritzky, C. W. Rees and K. T. Potts, Pergamon Press, Oxford, **1984**, p.359

195 R. M. Silverstein and F. X. Webster, *Spectrometric Identification of Organic Compounds*, 6th ed., John Wiley & Sons Inc., New York, **1998**, p.27.

196 F. W. McLafferty, *Interpretation of Mass Spectra; An Introduction*, W. A. Benjamin, Inc., New York, **1966**, p.137.

197 H. Budzikiewicks, C. Djerassi and D. H. Williams, *Mass Spectrometry of Organic Compounds*, Holden-day, Inc., San Francisco, **1967**, p.651.

198 J. L. Occolowitz and G. L White, *Anal. Chem.*, **1963**, 35. 1179-1182.

199 D. D Perrin and W. L. F. Armarego, in *Purification of Laboratory Chemicals*, 3rd ed., Pergamon, Oxford, **1988**, pp.68, 157, 161, 179.

200 J. Leonard, B. Lygo and G. Procter, in *Advanced Practical Organic Chemistry*, 2nd ed., Blackie, Glasgow, **1995**, pp. 60, 62, 63.

201 B. S. Furniss, A. J. Hannaford, P. N. Smith and A. R. Tatchell, *Vogel's Textbook of Practical Organic Chemistry*, 5th ed, Longman Scientific & Technical, Singapore, **1989**, pp.409-410, 412.

202 W. C. Still, M. Kahn and A. Mitra, *J. Org. Chem.*, **1978**, 43, 2923-2925.

203 K. S. Feldmanand D. B. Vidulova, *Org. Lett.*, **2004**, 6, 1869-1871.

204 N. N. Suvorov and V. E. Golubev, *Khim. Farm. Zhu*., **1967**, 1, 13-18.

205 V. E. Golubev and N. N. Suvorov, *Khim. Geterotsikl. Soedin. Sb 1*, **1967**, 21-4; *CAN* 70:77700.

206 R. L. Franklin and Harold M. Sell, *J. Chem. Engineering Data*, **1969**, 14, 267-269.

207 B. Schwarz and T. Hofmann, *J. Agric. Food Chem.*, **2007**, 55, 1405-1410.

208 E. A. Petrushkina, V. V.Gavrilenko, Yu. F. Oprunenko, and N. G. Akhmedov, *Zh. Obshch. Khim.*, **1996**, 66, 1864-1870; *CAN* 126:144225 .

209 A. E. Yavorskii, A. V. Stetsenko, S. G. Zavgorodnii and V. L. Florent'ev, *Khim.* Geterotsikl. Soedin., **1988**, 2, 198-202; *CAN* 110:114751.

210 L. Garuti, A. Ferranti, G. Giovanninetti, G. Scapini, L. Franchi and M. P. Landini, *Archiv der Pharmazie*, **1982**, 315, 1007-13. PubMed ID 6297425

211 T. Watanabe, T. Kaneo, and K. Hasegawa, *Jpn. Kokai Tokkyo Koho*, **2004**, pp. 57, JKXXAF JP 2004347738; CAN 142:45905.

212 P. Mamalis, V. Petrow, and B. Sturgeon, *J. Pharm. Pharmacol.*, **1950**, 2, 503-11; *CAN* 45:36136.