Novel Acyclic Nucleotide Phosphonates against RNA Viruses

A Thesis Submitted for the Degree of

Masters of Science

BY

Gregory David Bowden

In the Department of Chemistry

UNIVERSITY OF CAPE TOWN

July 2012

Supervisor

Professor Roger Hunter

Plagiarism Declaration

I declare that “Novel Acyclic Nucleotide Phosphonates against RNA Viruses” is my own work and that all sources that I have used or quoted have been indicated and acknowledged by means of complete references.

_____

Greg Bowden

DATE: _____

Acknowledgements

I would first and foremost like to thank Professor Roger Hunter for his constant encouragement, patience and guidance. His passion for organic chemistry and love for science has been greatly influential to my own development as a chemist and a scientist. I greatly value all of his support and advice over the past few years that I have spent as a member of his research group.

I would also like to give special thanks to Dr Sophie Rees-Jones for all her encouragement, help and support, especially with regards to the administrative and logistical aspects of the project.

I also wish to give a special thanks to Dr Yassir Younis for his advice and encouragement, as well as the odd bit of trouble shooting.

I would also like to thank my colleagues, past and present, in the Hunter research group, especially Wade Petersen, Rudy Cozett, Ana Andrijevic, Cathryn Driver, Athi Msutu, Nashia Stellenboom, Thobela Bixa, Mandla Mabunda and Mukesh Joshi. Their collective inputs, advice, friendship, humour and collegiality was greatly appreciated. Additionally, I would to thank my other colleagues in the UCT chemistry department, especially those in the Chibale, Gammon and Smith research groups.

For biological testing I would like to thank Professor Paolo La Colla at the Department of Virology, University of Cagliari in Italy.

I would also like to thank Noel Hendricks and Pete Roberts for their training and constant support on the NMR spectrometers. I would additionally like to thank the CAF unit as Stellenbosch University for their analytical support.

For financial assistance I would like to thank the National Research Foundation and the University of Cape Town.

I would finally like to thank my family and friends for all their support and encouragement away from the university and it is to them that I dedicate this thesis.

Table of Contents

List of Abbreviations

$(COCl)_2$	Oxalyl chloride
ANP	Acyclic Nucleotide Phosphonate
ATR	Attenuated total reflectance
AZT	3’-Azido-3’-deoxythymidine
BnBr	Benzyl bromide
CH_3CN	Acetonitrile
Cs_2CO_3	Caesium carbonate
CuI	Copper (I) iodide
*cyclo*Sal	Cyclosaligenyl
δ	Chemical shift in ppm
d	doublet
dd	Doublet of doublets
d4T	2'-3'-didehydro-2'-3'-dideoxythymidine
DAHP	Diisopropyl hydrogen phosphonate
DAP	2,6-Diaminopurine
DAPy	Diaminopyrimidine
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	Dichloromethane
DENV	Dengue virus
dhept	Doublet of heptets
DHPA	(S)-9-(2,3-Dihydroxypropyl)adenine
DMAP	Dimethylaminopyridine
DME	Dimethoxyethane
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	deoxyribonucleic acid
dsDNA	Double strand DNA
EDTA	Disodium ethylenediaminetetraacetate
eq.	Equivalents
ES	Electro spray
GI	Gastrointestinal tract
gp	glycoprotein
GTP	Guanine triphosphate
HAART	Highly active anti-retroviral treatment
HBV	hepatitis B virus
HCV	hepatitis C virus
HEPT	1-[(2-Hydroxyethoxy)methyl]-6-(phenylthio) thymine

HIV	Human immunodeficiency virus
HMBC	Heteronuclear Multiple Bond Correlation
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectroscopy
HSQC	Heteronuclear single quantum correlation
IMPDH	Inosine monophosphate dehydrogenase
i-Pr	Isopropyl
i-PrOH	Isopropyl alcohol
IR	Infrared
K_2CO_3	Potassium carbonate
LG	Leaving Group
Li_2CO_3	Lithium carbonate
MeOH	Methanol
Mp	Melting Point
mRNA	messenger RNA
MsCl	Methylsulfonyl chloride
MW	Microwave
NaH	Sodium Hydride
NaI	Sodium iodide
$NaNO_2$	Sodium nitrite
NEt_3	Triethylamine
NI	Nucleoside inhibitor
NMR	Nuclear magnetic resonance
NNI	Non-Nucleoside inhibitor
NNRTI	Non-nucleoside Reverse Transcriptase inhibitor
NRTI	Nucleoside Reverse Transcriptase inhibitor
NS5B	HCV non-structural protein 5B
NTP	Nucleotide triphosphate
$P(Ph)_3$	Triphenylphosphine
PCl_2	Phosphorus trichloride
$Pd(PPh_3)_2$	Tetrakis(triphenylphosphine)palladium(0)
Pd-C	Palladium on carbon
PG	Protecting Group
PME	Phosphonomethoxyethyl
POC	Isopropyloxycarbonyloxymethyl
POM	Pivaloyloxymethyl
POMCl	Pivaloylmethyl chloride
PT	Proton transfer
PV	Poliovirus

q	quartet
RdDp	RNA-dependent-DNA-polymerase
RdRp	RNA-dependent RNA-polymerase
RNA	Ribonucleic acid
RT	Reverse transcriptase
rt	room temperature
s	singlet
$SOCl_2$	Thionyl chloride
ssDNA	Single strand DNA
ssRNA	Single strand RNA
t	triplet
TBAI	Tetrabutylammonium iodide
THF	Tetrahydrofuran
tlc	Thin layer chromatography
TMSBr	Bromotrimethylsilane
TMSOMe	methoxytrimethylsilane
TsCl	*p*-toluenylsulfonyl chloride
WNV	West Nile Virus

Abstract

Acyclic nucleotide phosphonates (ANPs) have been used for years as successful anti-viral agents against diseases such as HIV/AIDS and hepatitis while the drug ribavirin is one of the only drugs available for the treatment of RNA-viral infections which mainly affect the developing world. The large and unmet need for anti-RNA viral treatments has prompted this study into the design and synthesis of a range of ANPs, which includes a series of ribavirin-based ANP derivatives.

The series of compounds was synthesised from a diisopropyl protected phosphonomethoxyethyl (PME) synthon and included an arylethynyltriazole derivative which was produced via a Sonogashira palladium catalysed cross-coupling reaction. A selection of these compounds was then deprotected to their corresponding phosphonic acids via a bromotrimethylsilane mediated phosphonate ester hydrolysis. In one example, a bis(pivaloyloxymethyl) prodrug variant was produced in order to probe a general synthesis for prodrug protected ANP derivatives. All new compounds were characterised by NMR, IR, and Mass spectroscopic techniques.

A selection of the new ANP derivatives was sent to a collaborator at the University of Cagliari in Italy for biological evaluation. However, as of yet, the results of these investigations are still pending.

Part 1

Introduction and Review

Chapter 1

Review of RNA Viruses

1.1 Introduction

Neglected diseases are a group of tropical infections that are especially endemic in low income populations in the developing regions of Africa, Asia and the Americas,[1] while an emerging infection is defined as an infectious disease whose incidence has increased in the past twenty years and threatens to increase further in the near future.[2] These sometimes highly pathogenic infections are often caused by RNA viruses and they have devastating and often fatal consequences for the people whom they afflict. They include acute haemorrhagic fevers such as yellow fever, Lassa fever, Crimean-Congo fever, Ebola, Argentine haemorrhagic fever and Marburg haemorrhagic fever.[3–6] Other well known and important RNA infections include: severe acute respiratory syndrome (SARS),[7] dengue fever,[6,8,9] West Nile virus,[6,8] hantavirus pulmonary syndrome, Japanese encephalitis and influenza (pandemic, seasonal, avian and swine flu.).[3] Many of these viruses are also classified by the CDC (Centre for Disease Control) as category A bioterrorism agents due to their highly infectious nature and the lack of effective treatments against them.[5] On top of the significant risk that RNA viruses pose to human health, foot-and-mouth disease, another RNA virus has the potential to devastate live-stock agriculture across the world.[4] Hence, RNA viral infections pose a significant threat to the lives and welfare of millions of people across the globe.

Many of these diseases have very high morbidity and mortality rates in humans[4,5] claiming thousands of lives every year and ruining many more through disability.[1] Dengue fever is considered to be one of the most prevalent viral diseases[6, 9] due to the fact that more than half of the world's population resides in areas with a high risk of dengue infection (Fig. 1.1).[9] Every year there are 50-100 million reported cases of dengue fever and of these 250,000-500,000 cases progress to dengue haemorrhagic fever, which has been observed to have a case fatality rate of up to 44%.[9] As an extreme example, diseases such as Ebola can have case fatality rates of up to 89%.[10] Other haemorrhagic fevers caused by RNA viruses, such as Lassa fever and yellow fever are also highly prevalent in many African countries. There has been a need for dedicated Lassa fever wards to be set up in Sierra Leone as well as other West African countries[3] and although there exists a highly effective single dose vaccine for yellow fever, as many as 200,000 cases of the disease are still reported each year[3] with an estimated 5000 deaths occurring annually.[11] The Figure below shows the global distribution of a number of important RNA viruses in the flavivirus family (Fig. 1.1).

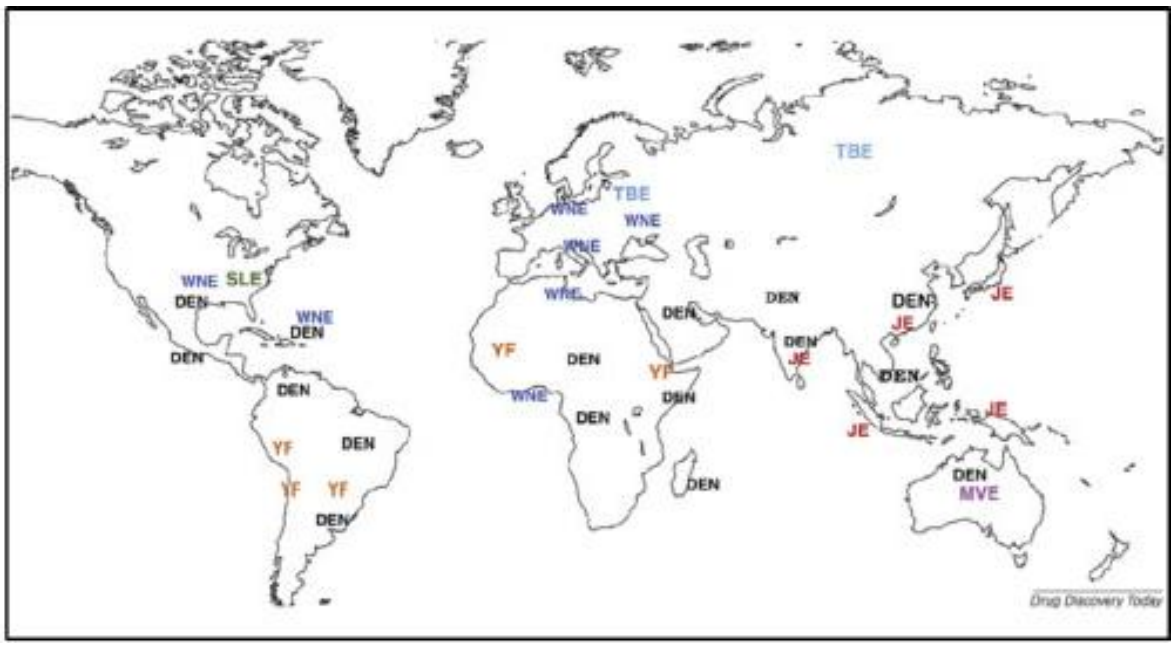


Fig.1.1 The global distribution of Japanese encephalitis (JE); Murray valley encephalitis (MVE); St. Louis encephalitis (SLE); tick borne encephalitis (TBE); West Nile encephalitis (WNV); Yellow fever (YF) and Dengue fever (DEN).[6]

At present, most of the anti-viral chemotherapies available for both DNA and RNA viruses have been licensed for well-known diseases such as Human Immunodeficiency Virus (HIV), Hepatitis B Virus (HBV) and herpes viruses.[4] Of all the currently licensed anti-viral drugs, only five are specifically used to treat non-retroviral RNA viruses, and of these, four are used to treat influenza, namely the M2 channel inhibitors amantadine and rimantadine (Fig. 1.2 a and b) and the neuraminidase inhibitors oseltamivir and zanamivir[4] (Fig.2 c and d). The fifth licensed drug, Ribavirin (Fig. 1.2 e), which has been licensed for use against Respiratory Syncytial Virus (RSV) and Hepatitis C Virus (HCV)[4] has also been used experimentally against a wide range of other RNA viruses including Lassa fever,[4,11] hantavirus[4,12] and Crimean-Congo fever[4,13] but with limited success. There is therefore a serious and unmet need for new broad-spectrum inhibitors of RNA viruses.[3,4]

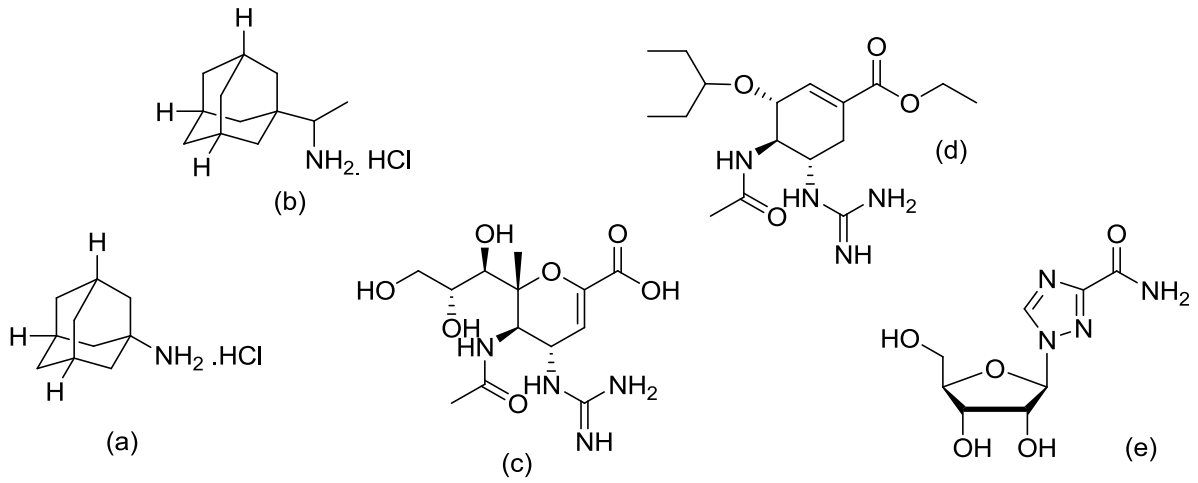


Fig 1.2: The chemical structures of (a) Amantadine, (b) Rimantadine, (c) Oseltamivir, (d) Zanamivir and (e) Ribavirin.

There are, however, many challenges involved in the development of effective medicines for RNA viral infections. Many of these diseases occur in low income areas with no sustainable market to incentivise the pharmaceutical industry to develop drugs for these illnesses.[3,4] Currently, the main source of funding for research into these diseases is through government research grants and private organisations.[3] Although funding is one of the major issues responsible for retarding drug-discovery efforts for RNA viral infections, there are many other logistical factors influencing the development and testing of new medicines. As discussed by Mike Bray in his article: *Highly pathogenic RNA viral infections: Challenges for antiviral research,*[3] these include: i) finding research subjects as outbreaks of these illnesses are often small, spread over very large geographical areas and are short lived; ii) diagnosing the disease early enough in these patients for any experimental treatments to be effective and iii) setting up and allowing for effective communication between different research groups and clinicians in the fields.[3] These are all significant factors, which are often compounded given the fact that outbreaks of these diseases often occur in areas where there is no running water or stable electrical power supply, let alone adequate hospitals and research facilities.[3] A combination of all these factors has significantly hindered the drug-discovery and development processes for neglected and emerging RNA viral infections thus far.

1.2 Virus Biology and Life-cycle.

In all drug-discovery efforts, it is important to understand the biology of the target disease in order to effectively identify biological targets for potential drugs as well as the possible mechanism of action of any hits or leads. The rest of this chapter will serve as an introduction to the basic biology and virology of RNA viruses as well as to give an account of some of the successful small-molecule chemotherapies that have been used to treat these illnesses and how those treatments work. This sub-section aims to demonstrate

how the important viral interactions and processes, which are relevant to this research, fit into the overall life-cycle of a generalized model virus.

All viruses are incapable of reproducing on their own and thus have to infect a host cell and commandeer its cellular machinery in order to replicate.[14,15] The various viruses and viral families all have different strategies for entering their hosts in order to use the cell's enzymes and components to replicate; however, most viruses share the same general life-cycle as shown below (Fig. 1.3). The first step of this process involves the entry of the virion into the cell. While different viruses utilise a host of different mechanisms in order to get their genetic material into the cell, enveloped viruses enter their hosts via the interaction of the viral envelope proteins with receptors on the surface of the host cell. This then leads to the fusing of the virus' outer membrane (envelope) with the cell's own membrane allowing the genetic material of the virus (often coated in a protein capsid) to enter the cytoplasm of the host. Non-enveloped viruses enter the host cell via other mechanisms as they do not possess an envelope that can fuse with the host cell. If the virus's genetic material is contained in a protein capsid, then the next step is the uncoating of the capsid from the viral genetic information. Once the viral genetic material is free within the cell, the replication of the viral genome and the biosynthesis of the various viral components can begin. After the viral genome has been copied and the correct viral proteins have been synthesised, the assembly of the new virus capsid can begin. Viral envelope proteins assemble on the outer membrane of the infected cell and as the new progeny virus emerges from its host it tears away a piece of the cell's membrane. This process happens multiple times per cell and this eventually leads to cell death.

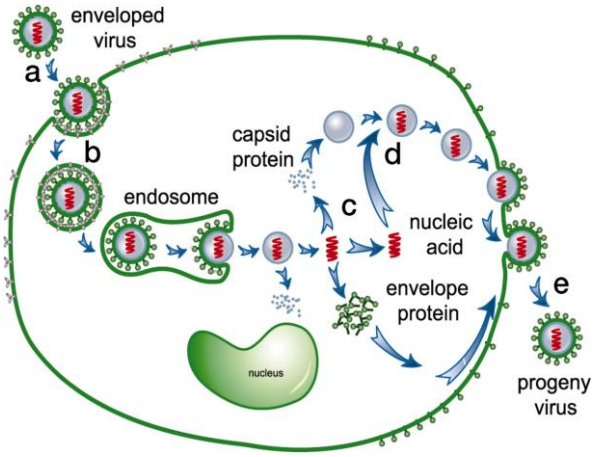
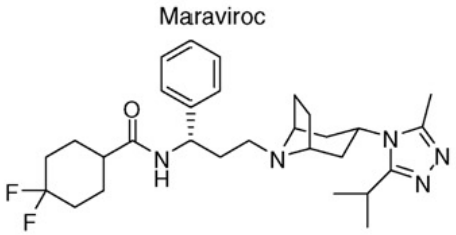


Fig. 1.3: General life-cycle of a typical virus showing: (a) initial interaction, (b) entry, (c) genome replication, (d) assembly and (e) budding, exit and detachment. This diagram has been modified from the article by Gao et al.[16]

1.3 Targets for Drug Discovery

The viral life-cycle presents a diverse array of targets for drug-discovery and development.[4] The entry of the viral particle into the cell, fusion of the viral membrane with the cell membrane, the uncoating and replication of the viral genome, the translation and processing of the viral proteins, the assembly of new viruses and the emergence and maturation of the progeny virus have all been considered and researched as potential targets for small-molecule inhibitors.[4] Many of these targets have been exploited in the past by a whole range of successful chemotherapies that have been used to treat well-known illnesses such as influenza and HIV.

The initial virus-cell interaction is an attractive although relatively newly discovered target for interfering with the viral life-cycle and in the case of HIV this is fairly well characterised.[17] Maraviroc (Fig. 1.4) is a CCR5 antagonist that qualifies as an entry inhibitor, and which disrupts the recognition process between HIV's surface gp120 protein and the host-cell receptor CCR5.[17,18] The disruption of this interaction prevents the next step in the entry process from taking place, whereby the virus and cell membrane fuse. Similarly enfuvirtide (Fig. 1.4), otherwise known as Fuzeon, has been developed to prevent the fusion process from taking place. Enfuvirtide is a synthetically produced 36 amino acid polypeptide based on an amino acid sequence in the N-domain that has been designed to complement a peptide sequence of a hydrophobic groove in the C-domain of the HIV gp41 protein.[18,19] When bound in this position, which is only exposed when gp41 folds outward into the host cell's lipid membrane, gp41 is prevented from folding back onto itself to form a six helix bundle, thus preventing membrane fusion from taking place.[19]



Enfuvirtide

NH_2-F-W-N-W-L-S-A-W-K-D-L-E-L-L-E-Q-E-N-K-E-Q-Q-N-Q-S-E-E-I-L-S-H-I-L-S-T-Y-Ac

Fig. 1.4: The chemical structures of Maraviroc[18] and the amino acid sequence of enfuvirtide (fuzeon).[18]

The next phase in the virus' life cycle requires that the protein capsid surrounding the viral genetic material be removed in order for transcription and translation to take place. In the case of the influenza virus, M2 channels allow the unidirectional passage of protons from the host cell's cytoplasm, into the viral

envelope.[4] The drop in pH on the inside of the viral particle causes key viral envelope protein interactions to become destabilized thus initiating the uncoating process and exposing the contents of the virus capsid to the cytoplasm of the host cell.[4] The M2 channel inhibitors amantadine (fig. 1.2 a) and rimantadine (Fig. 1.2 b) have been successfully employed to block the passage of protons into the virus, hence preventing the uncoating process from taking place.[4]

The synthesis and subsequent modification of viral proteins from the viral genome has also been viewed as a potential target for anti-viral drug development.[4] Unfortunately, as the actual translation process for all viruses makes use of cellular machinery, inhibition of the enzymes and ribosomes that perform these tasks may also interfere with that normal functioning of the cell.[14,15] All RNA viruses, however, have polycistronic genomes, meaning that the single mRNA strand that is synthesised from the genome codes for multiple proteins.[14] This causes a problem for the virus in that the ribosomes that translate the genetic information into proteins can only deal with monocistronic messages.[14] (For every strand of mRNA that is translated one continuous polypeptide chain is synthesised repeatedly.) The result of this is that the viral mRNA gets translated into one polypeptide chain which contains the sequence information for multiple proteins. Most RNA viruses deal with this problem by having a protease enzyme that can selectively hydrolyse the polypeptide into the functional protein subunits.[4,14,15] The HIV and HCV protease enzymes have been well studied and there are multiple licensed drugs that act as substrate transition-state analogues against these important enzymes.[7,17,20] These include ritonavir, saquinavir, indinavir and nelfinavir.[7] Lopinavir or ritonavir may be prescribed as the protease component of Highly Active Anti-Retroviral Treatment (HAART).

The assembly of the progeny virus is another vital step in the virus' reproductive pathway. There are a few inhibitors of this process that are currently in early development, but in the case of most RNA viruses this process is not well understood[4]. Viral assembly thus remains a potentially rich research area for future drug discovery efforts.

After the viral particle has been assembled, it must leave the cell in order to infect others. There are numerous biochemical processes that are vital to the budding, exit and detachment of the virus from the cell. One of the most well understood processes of viral exit is that which involves the neuraminidase enzyme of the influenza virus.[7] Once the Influenza virus has left the cell, viral hemagglutinin interacts with the N-acetylneuraminic acid (NANA)-bearing glycoprotein receptor on the cell's surface which prevents the virus from breaking away from the cells surface.[4] The viral neuraminidase inhibitor is responsible for cleaving the NANA from the host cell-surface glycoprotein, thus releasing the virus into the extracellular matrix in order to infect more cells. Oseltamivir (tamiflu) and zanamivir (relenza) (Fig. 1.2 c and d) are two widely used drugs that inhibit the action of the neuraminidase enzyme.[4] They, like protease inhibitors of HIV, have been designed as transition-state analogues of the natural substrate.[4]

One of the most studied, most promising and currently most pharmaceutically exploited stages in the viral life cycle is the replication of the viral genome. This process presents many possible molecular targets for drug discovery;[4,8] however, as different viruses all have slightly different strategies for replicating and transcribing their genomes, one needs to understand what differentiates the various viral classes and mechanisms of genome replication in order to better design and discover new drugs against them.

1.4 Genome Replication as a Target for Drug-Discovery

All viruses need to synthesize messenger RNA (mRNA) in order to use the cellular machinery for translation of their genetic information into viral proteins that can then assemble into thousands of new replicas of the original virus.[14,15] mRNA is defined as positive strand [(+)RNA] as it contains the genetic information in a form that is directly translatable into new viral proteins.[15] The Baltimore classification system categorises viruses based on the strategy that they use to replicate their own genomes and synthesise mRNA (Fig. 1.5).[14,15,21] From this classification system, two broad classes of virus emerge: DNA viruses (classes I and II), which synthesise mRNA from a DNA genome, and RNA viruses (Classes III, IV, V and VI), which synthesise mRNA from an RNA genome. Retroviruses (Class VI) are often considered as a special case of RNA virus, as they synthesise mRNA from intermediate DNA which has been reverse-transcribed from an RNA genome.[15,21] Retroviruses and their strategies of replication, cellular entry, construction and emergence have been well studied as they include the well-known and highly prevalent human immunodeficiency virus (HIV).

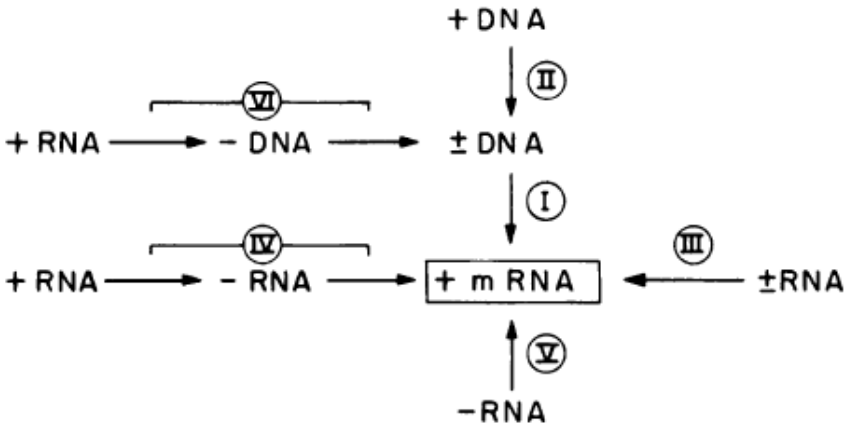


Fig. 1.5: The Baltimore classification as originally reviewed by D.Baltimore.[21]

DNA viruses exhibit a much greater diversity in size, shape and structural complexity than RNA viruses. DNA viruses contain their genetic information in either closed circular DNA loops or in some cases as linear DNA strands.[14] Owing to the fact that their genomes exist in the same form that the host cell contains its genetic

code, the DNA virus is able to use the normal cellular machinery in order to transcribe and translate its genome into viral proteins. For this reason they have had to develop fewer biologically "novel" or "unusual" replication strategies relative to their RNA-containing counterparts.[14] DNA viruses must also use the cell's mechanism of DNA synthesis in order to replicate their genomes.[14,15] This creates a complication for the virus as most cells only produce the enzymes and components required for DNA replication when they are about to divide naturally.[14] Many DNA viruses have developed strategies to deal with this problem such as containing the genetic information to activate the cell-division phase or by simply producing the required enzymes, proteins and primers themselves.[14,15] Some well-known examples of DNA viruses include adenovirus (conjunctivitis, tonsillitis and gastroenteritis), herpes simplex virus, hepatitis B and smallpox,[14] which killed thousands of people every year until an effective vaccine was developed against it.

RNA viruses do not contain their genetic information in the same form as that of the host cell and, as a result, are unable to use the host cell's machinery directly to synthesise mRNA or replicate their genomes. They have hence had to develop "novel" strategies in order to perform this task, almost always involving enzymes which are able to polymerise new genetic molecules from an RNA template.[14] These RNA-dependent RNA-polymerases (RdRp's) are viral enzymes that synthesise viral RNA strands from a viral RNA template, and are vital to the ability of the RNA virus to reproduce. These enzymes are absent in eukaryotic cells[8,22] and thus have to either be carried into the cell by the viral particle or, if the viral genome can act directly as mRNA, translated and synthesised *de novo* within the cell.[14,15] From the Baltimore classification (Fig 1.5) system, Class IV [(+)ssRNA] viruses are able to use their genomes directly as mRNA and are hence able to synthesise their RdRps *de novo*.[14,15] Class V [(-)ssRNA)] viruses, however, need to carry the necessary RdRp into the cell in order replicate and synthesise mRNA as the (-) strand RNA is not directly transcribable. RdRps therefore play a vital part in the life-cycle of the RNA virus, further promoting them as attractive targets for drug discovery.[8,23]

Class VI RNA viruses, otherwise known as retroviruses, are a special case as they reproduce their genomes via a DNA intermediate. Unlike other RNA viruses, they use an RNA-dependent-DNA-polymerase (RdDp), known as a reverse transcriptase (RT) enzyme, to transcribe their genome from an RNA Template into DNA.[14,15,24] Furthermore, reverse transcriptase can then act as a DNA-dependent DNA-polymerase, creating a second DNA strand to complement the first, in order to form the required double stranded DNA (dsDNA) that can then be integrated into the host cell's genome and used for mRNA synthesis.[24]

HIV RT has also been shown to possess RNase-H activity. Until recently, it was thought that all the effective drugs against HIV RT have targeted only the polymerase activity of the enzyme[24] but recent evidence suggests that certain Non-Nucleoside Reverse-Transcriptase inhibitors do in fact interfere with the enzyme's RNase-H activity.[25] The HIV reverse transcriptase enzyme has been well studied and has been used as the target for many successful anti-retroviral drugs. Thus, the success of past drug-discovery efforts

against HIV reverse transcriptase and other viral polymerases continues to justify the continued exploration of RdRps as potential targets for new anti-viral drugs.[7,8] For this reason, HIV RT acts an important model for understanding and manipulating other viral polynucleotide polymerases.

1.5 Viral polymerases as targets for drug discovery

RdRps produce new RNA molecules from RNA templates, a process which does not occur naturally in animal cells. This makes RNA virus RdRps very attractive targets for new selective drugs.[23] There is still not the same level of structural and mechanistic understanding available for the viral polymerases of RNA viruses as there is for HIV and other retroviruses. However, there is an increasing amount of information available for different RdRps across a wide range of viral species including hepatitis C virus (HCV), dengue virus (DENV), West Nile Virus (WNV), poliovirus (PV) and foot and mouth disease virus (FMDV).[8,20,23] The X-ray crystal structures of HCV non-structural protein 5B (NS5B) RNA-dependent RNA-polymerase (Fig. 1.6)[20] and WNV RdRp are shown below (Fig. 1.7).[8] Unlike most other polynucleotide polymerases, which possess an "open-hand" or U shaped structure, RdRps have a general "closed hand" structure, where the active site of the enzyme is completely enclosed by "thumb and finger" domains of the protein forming a channel-like structure in which the template RNA molecule can bind.[23] This channel aids in guiding the template towards the active site of the enzyme. There also exists a small positively charged tunnel on the back end of the enzyme which helps bring the negatively charged activated nucleotides into the active site.[23] Of all the different RdRps that are known and studied, the RdRp of HCV is one of the most understood. X-ray crystal structures of the HCV non-structural protein 5B (NS5B) RNA-dependent RNA-polymerase have shown the presence of numerous potential binding sights and pockets which may all provide potential drug-discovery targets[20] (Fig. 1.6). Currently there are several drugs developed against this enzyme that are undergoing clinical trials.

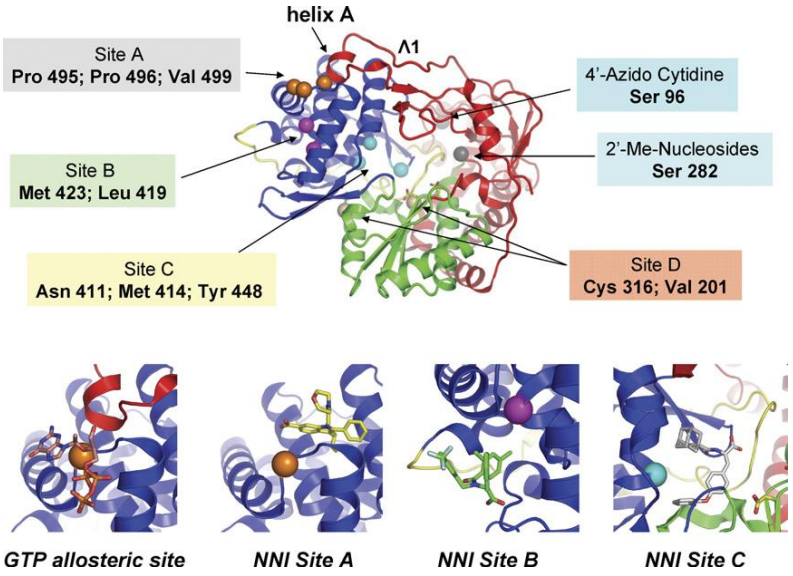


Fig. 1.6: The structure of HCV non-structural protein 5B (NS5B) RNA- dependent RNA-polymerase as reviewed by De Francesco et al.[20] Possible allosteric binding sites are also indicated.[20]

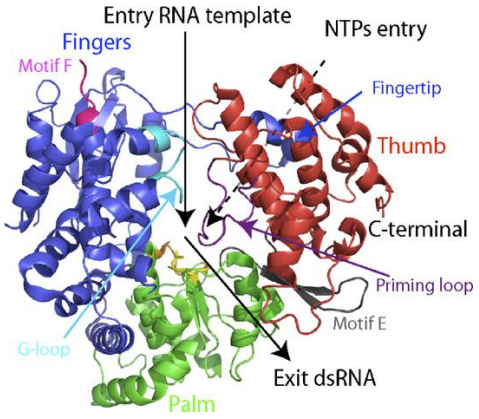


Fig. 1.7: The Structure of West Nile Virus RNA- dependent RNA-polymerase as reviewed by Malet et al.[8] Note the formation of the "closed Hand" structure by the "thumb" and "finger domains" of the protein.

As in the case of the two well-known classes of anti-retroviral drugs against HIV RT, namely Nucleoside Reverse Transcriptase inhibitors (NRTIs) and Non-nucleoside Reverse Transcriptase inhibitors (NNRTIs),[7,25] inhibitors of RNA viral RdRps can be largely divided into two classes based on chemical structure: Nucleoside Inhibitors (NIs) and Non-Nucleoside Inhibitors (NNIs) (Fig. 1.8).[8,23] Generally, NIs are analogues of natural nucleosides that target the active site of an enzyme, either blocking a biosynthetic pathway involving a nucleoside, or mutating or terminating the growing strand of genetic material.[8,23] Most NIs, especially in the case of HIV NRTIs, act as chain terminators as they lack a hydroxyl group at the 3' position on the ribose ring, which is required for chain elongation.[26] By comparison, NNIs are small-molecule inhibitors that are not nucleoside analogues, and as such, act at allosteric binding sites separate from the active site of a RdRp.[8,23] These allosteric interactions cause (or in some cases block) a conformational

change in the enzyme that results in disruption of normal activity at the active site. In the case of HCV, high-throughput screening has yielded a series of viable NNIs that are currently undergoing clinical trials (Fig. 1.8).[20,23] The rational design of nucleoside analogues, NIs, has also yielded successful candidates for clinical trials against the RNA virus, HCV as well as HIV (Fig. 1.8).[20,23]

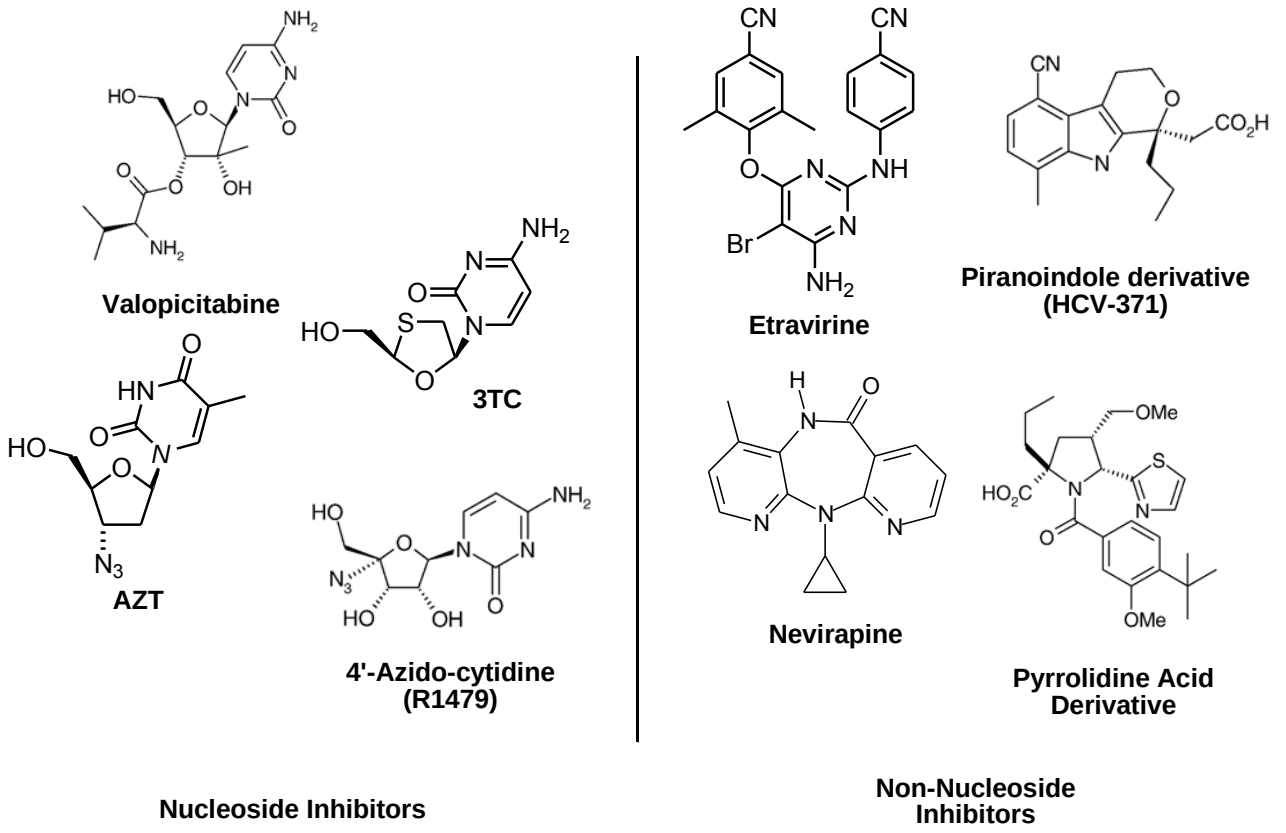


Fig. 1.8 Nucleoside Inhibitors (NI's) and Non-Nucleoside Inhibitors (NNIs) against HIV (AZT, 3TC, Etravirine and Nevirapine) and HCV (Valopicitabine, 4'-azido-cytidine, piranoindole derivative (HCV-371) and the pyrrolidine acid derivative).[20]

1.6 Ribavirin and T-705 (Favipiravir)

As was discussed earlier in this review, the number of active, approved compounds for the treatment of RNA viruses is disappointingly limited and all the drugs discussed thus far have been approved for a relatively small number of infections namely HIV, HCV and influenza. At present the only drug that has been approved for the treatment of RNA viruses other than influenza (not including retroviruses) is the nucleoside-based inhibitor Ribavirin. Owing to its particular relevance to the work described later in this thesis, its history, and proposed mechanism of action will be discussed in this subsection. Favipiravir, another potent, structurally unique nucleoside analogue, with strong, broad-spectrum activity against RNA viruses, is currently in development and will also be briefly discussed in this subsection.

1.6.1 Ribavirin

The activity and structure of the nucleoside analogue ribavirin was first reported in the early 1970s and it was hailed as the first synthetic, non-natural nucleoside analogue to possess broad-spectrum activity against a large range of RNA viruses.[4,27] In fact, ribavirin has over the years been shown to be active against virtually all RNA viruses with varying degrees of potency;[28] however, overall the potency of ribavirin can be considered to be low, with most EC_{50} values above 1μM.[4] This means that any clinical treatment regime involves high doses of the drug with adverse side-effects.[29]

Like all cyclic nucleoside analogues, ribavirin (1-β-D-ribofuranosyl-1,2,4-triazole-3- carboxamide, or virazole as it is less commonly known),[27,28] consists of a sugar ring (in this case D-ribofuranose) coupled to a heterocyclic nitrogen base (Fig. 1.9). What makes ribavirin structurally unique is its non-natural, 1*H*-1,2,4-triazole-3- carboxamide base, which, depending on the orientation of the amide, can be analogous to either of the naturally occurring purine bases, adenine or guanine (Fig. 1.9). It is also important to note the presence of a hydroxyl group at the 3' position of the sugar ring, a feature which is invariably absent in most other chain-terminating nucleoside inhibitors (see above), suggesting that ribavirin does not act by terminating RNA chain growth.

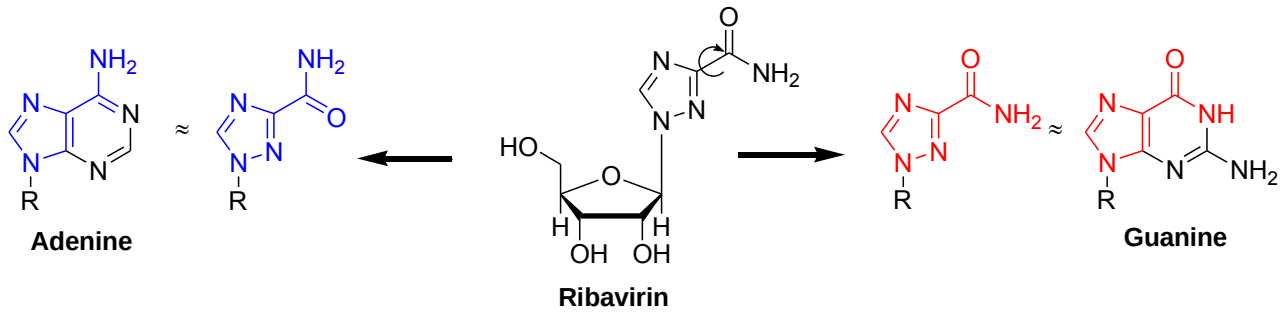


Fig. 1.9: The structure of Ribavirin[27] as well as the structural relationship of its conformers to both adenine and guanine.

Even after nearly 40 years as the main drug available for the treatment of RNA virus infection, ribavirin's mechanism of action is still a topic of much debate.[4] There are currently 5 main mechanisms by which it is thought to elicit its activity. The first is as an inhibitor of inosine monophosphate dehydrogenase (IMPDH), which is an important enzyme in the *de novo* biosynthesis of guanine nucleotides.[28] The inhibition of IMPDH leads to the depletion of cellular guanine triphosphate (GTP), which the virus needs in order to replicate its genome;[4,28] however, recent research involving the use of other potent IMPDH inhibitors, such as mycophenolic acid,[29] has shown that although RNA viruses do seem to be sensitive to the depletion of GTP pools within the cell, this is not the main mechanism by which ribavirin shows its activity.[30]

The second proposed mechanism by which ribavirin is thought to act is as an immunomodulator, where it is thought to maintain the T-helper type-1 response which is associated with cellular immunity.[28] This

mechanism was proposed to explain why Lassa fever patients, who were treated with ribavirin, showed significantly less liver damage relative to other patients who were not, yet showed little to no decrease in their viral load.[29] The current view is that this effect is minimal and does not account for the majority of ribavirin's anti-viral activity.[28,29] The third mechanism involves the inhibition of the enzymes which are responsible for capping and stabilizing the 5'-end of viral RNA; however, not all viruses make use of such capping structures and therefore this proposed mechanism of action fails to explain the broad-spectrum of anti-viral activity exhibited by ribavirin.[28]

Ribavirin 5'-triphosphate is also thought to directly inhibit the formation of RNA viral genetic material by inhibiting the function of various viral RdRp's.[4,28]Specifically, Ribavirin 5'-triphosphate has been shown to bind to, and inhibit, the active site of the both HCV and Influenza RdRP,[31,32] (in order for nucleosides to become active metabolites of polynucleotide polymerases, they need to be converted to their corresponding 5'-triphosphate species. This process is described in greater depth in chapter 2.1.) The most current view on ribavirin's mechanism of action is not that ribavirin 5'-triphosphate inhibits RdRps, but that it is rather incorporated by the RdRp into the new viral genome. The incorporation of the ribavirin nucleotide was always speculated, but only in 2000 did Crotty et al. produce strong evidence to determine that ribavirin 5'-triphosphate is indeed, although at a low rate, incorporated into the viral genome by a variety of different viral RdRp's.[4,28,33] More importantly, the study showed that because of ribavirin's ability to mimic both adenine and guanine,[29] it can be a template for both cytidine and uridine nucleic acids when new RNA is synthesised from the strand containing the incorporated ribavirin (Fig. 1.10).[28,33] This eventually leads to a lethal build-up of errors in the viral RNA sequence causing the newly synthesised viral genome to lose its genetic fidelity, rendering the new genome useless, a process which is termed "error catastrophe" (Fig. 1.10).[4,29,33] Through evolution most cellular organisms have developed complex mechanisms to correct errors in their genetic material and thus avoid error catastrophe, while viruses have not and are hence susceptible to mutations.[29] This can be advantageous for the virus in terms of its development of resistance to certain drugs, as well as its general evolution, but it also makes them vulnerable to events like error catastrophe, a feature which could be exploited by future mutagenic viral nucleoside analogues like ribavirin.

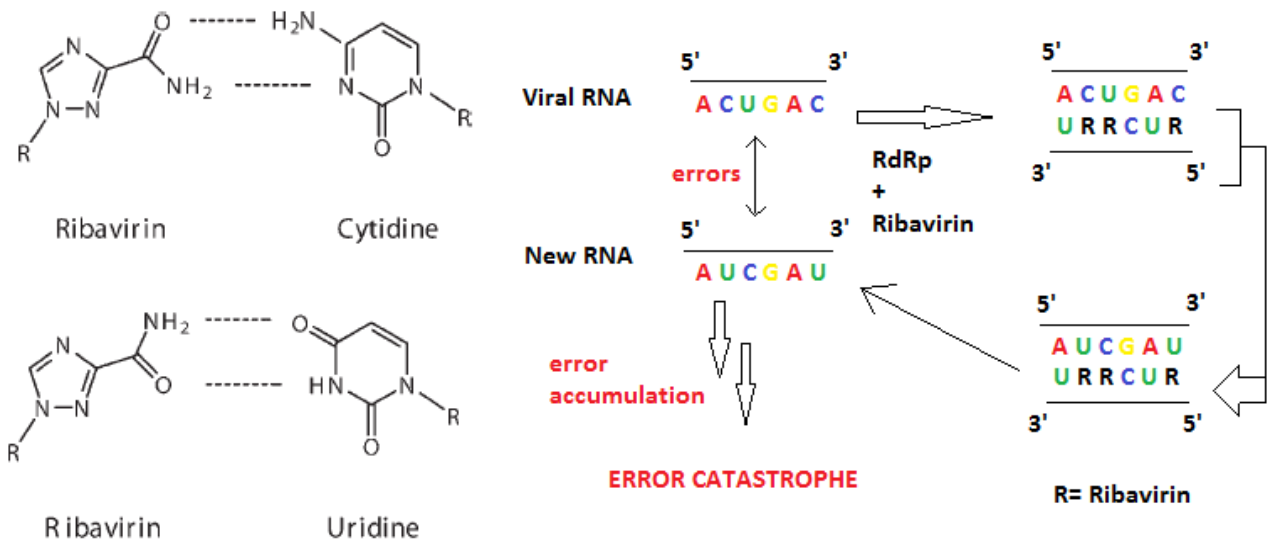


Fig. 1.10: The left shows the hydrogen-bonding patterns of ribavirin to cytidine and uridine.[28] The diagram on the right shows how ribavirin causes the build-up of nucleotide errors which leads to error catastrophe.

1.6.2 T-705 (Favipiravir)

The pyrazinecarboxamides T-705, T-1105 and T-1106 (Fig. 1.11) were first synthesised by the Toyama chemical company, Ltd., from which T-1105 was identified as a potential lead with anti-viral activity against influenza virus.[34] Derivatives of this compound were then synthesised which led to the discovery of T-705, now also known as favipiravir, which shows broad spectrum anti-viral activity.[6] T-705 was originally found to show high activity against Influenza[6,34,35] and has since also been shown to posses high activity against various bunyaviruses and arenaviruses,[36,37] as well as phlebovirus,[38] West Nile virus,[39] yellow fever[40] and foot-and-mouth disease[34] infections. T-705 has so far proved so effective over such a broad range of viruses that it has been suggested as a replacement for ribavirin.[6,37]

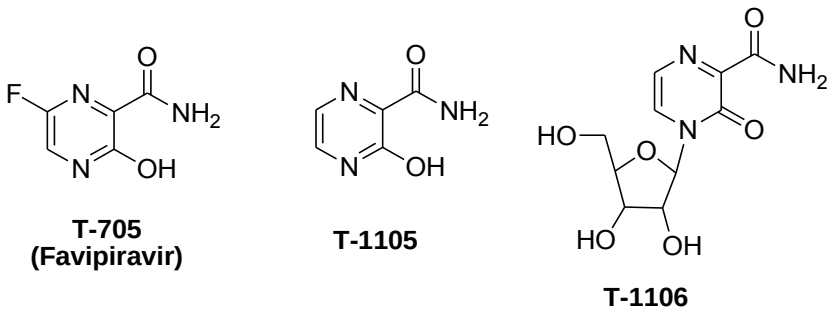


Fig. 1.11: The Chemical structures of T-705, T-1105 and T-1106[34]

T-705 has been shown to be a nucleoside analogue and RNA polymerase inhibitor[6,34] even though unmodified, on a structural level, it does not resemble the general nucleoside structure. Once it has entered the cell it is converted to its corresponding ribonucleotide monophosphate by a phosphoribosyl transferase, after which it is then further phosphorylated to its ribonucleotide triphosphate, which is the active species (Fig. 1.12).[6] T-705 (and T-1106) have been said to be structurally analogous to ribavirin, especially with regard to the importance of the presence of the amide at the 3-position,[6] suggesting that T-705 could have multiple hydrogen bonding profiles, as is the case in ribavirin. These types of ambiguous nucleotide bases have been termed universal bases, as reviewed by Loakes,[41] in that they possess the ability to pair with multiple natural nucleotide bases. Universal bases may hold an interesting place in the future of nucleoside analogue drug-discovery as mutagenic anti-virals or as polymerase inhibitors.[41]

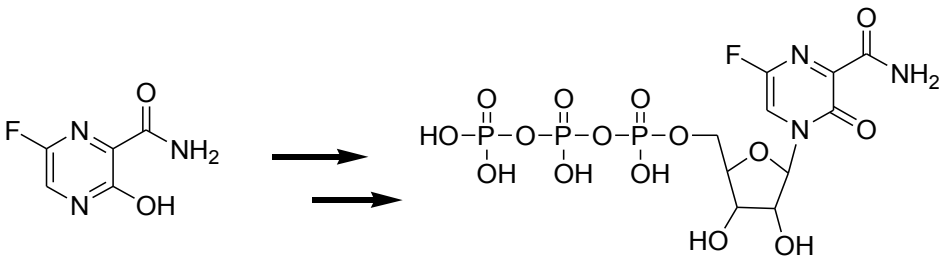


Fig. 1.12: The Conversion of T-705 to T-705 ribofuranosyl triphosphate.

Given the past success of nucleoside-analogues such as nucleoside-based HIV reverse transcriptase inhibitors, ribavirin and favipiravir and the continued promise that these types of drugs show as selective inhibitors against RdRps in RNA viruses, as well as their particular relevance to this project, the next chapter of this thesis will focus on NIs (specifically acyclic nucleotide phosphonates): their inception, development, structure-activity relationships and evolution.

Chapter 2

Acyclic Nucleotide Phosphonates: design, development and evolution.

2.1 Introduction

Nucleoside analogues were first identified as potentially important chemotherapeutic agents in the 1960s when they where first used in the treatment of cancer and later for the treatment of viral infections.[42,43] They were seen to act as "anti-metabolites",[42] interfering with the tumour cell's natural mechanisms of genome replication. All natural nucleosides, both RNA and DNA monomers, show the same basic structural features involving a heterocyclic base attached to an aldopentofuranose (ribose) sugar moiety (Fig. 2.1a). In order for the nucleoside to become an active metabolite it needs to be transformed by cellular kinases into its corresponding 5'-triphosphate nucleotide (nucleotides are the 5'-phosphoric acid esters of nucleosides) (Fig. 2.1d). Only then can it be incorporated into a growing strand of genetic material by the relevant nucleoside polymerases.[25,42] The same process applies for nucleoside analogues, as the majority of them act by being incorporated into a growing strand of genetic material, either terminating its growth or causing it to mutate.[25,29,42] The initial phosphorylation of a nucleoside to its corresponding 5'-monophosphate (Fig. 2.1b) has been shown to be the rate-determining step in the biosynthetic pathway to the active 5'-triphosphate and as such, many nucleoside analogues show poor biological activity simply because of their lack of initial phosphorylation.[42] This problem started the pursuit of phosphorous-modified nucleotide analogues, in which a phosphorus atom was built into the drug molecule as a way to eliminate the need for the initial rate-determining, intracellular phosphorylation step.[42–44]

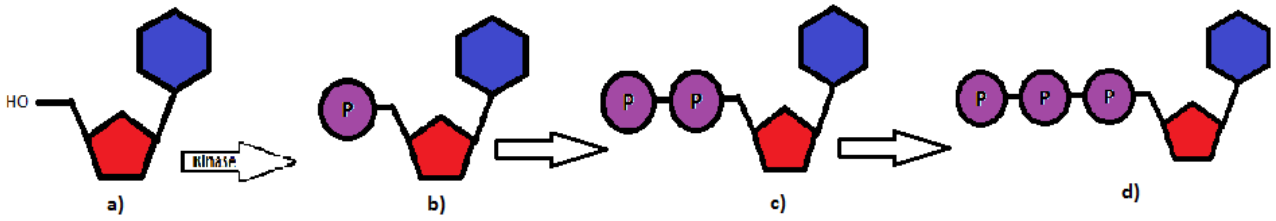


Fig. 2.1: The conversion of a nucleoside to 5'-triphosphate nucleotide. The aldopentofuranose sugar is shown in red, the heterocyclic base in blue and the phosphates in purple.

Given the high toxicity of these nucleoside and nucleotide analogues, detailed studies on the effect of modification of both the heterocyclic base and sugar moieties where carried out.[42] The results of these studies drew a number of important conclusions pertaining to the design of future nucleoside and nucleotide derivatives: i) The modification of the heterocyclic base has a marked effect on the relevant

anabolic reactions within the cell, which do not tolerate much variation in the structure of the substrates; however, modification of the base does not show a significant effect on catabolic processes involving the drug molecule.[42,45] This is most likely due to the importance of strong nucleotide base-pair hydrogen-bonding interactions in the construction of new genetic material ii) New nucleotide analogues which were designed to closely mimic the natural substrates were found to be catabolically unstable due to the labile nature of the phosphorus-oxygen bond of the phosphoric acid ester (Fig. 2.2).[42–44] The overall stability of new nucleotide analogues needed to be increased through the modification of key linkages such as the phosphoric acid ester in the 5'-monophosphate.[42] iii) Any phosphorus-modified derivatives had to be isopolar to the natural 5'-monophosphate substrates, meaning that the electronic profiles of any nucleotide derivatives had to be similar to those of natural nucleotides.[42,43] Another goal of these studies was to reduce the structural features required for enzyme binding and thus a variety of acyclic nucleosides and nucleotides were developed (Fig. 2.3).[42] These studies laid the ground work and provided a platform for the development of a wide range of acyclic nucleosides, which would eventually lead to the discovery of the first acyclic nucleotide phosphonates.

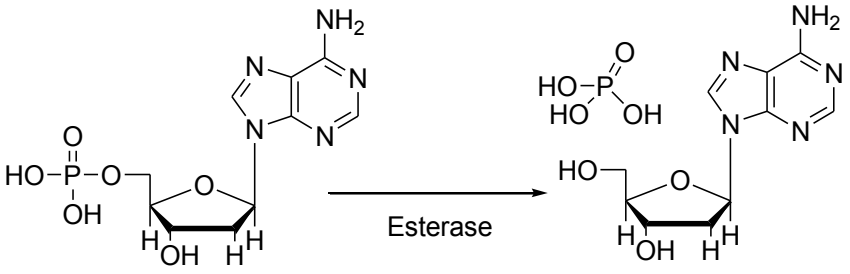


Fig. 2.2: The source of the catabolic instability inherent in all nucleotide phosphates, shown here is the deoxyribo-nucleotide deoxyadenosine.

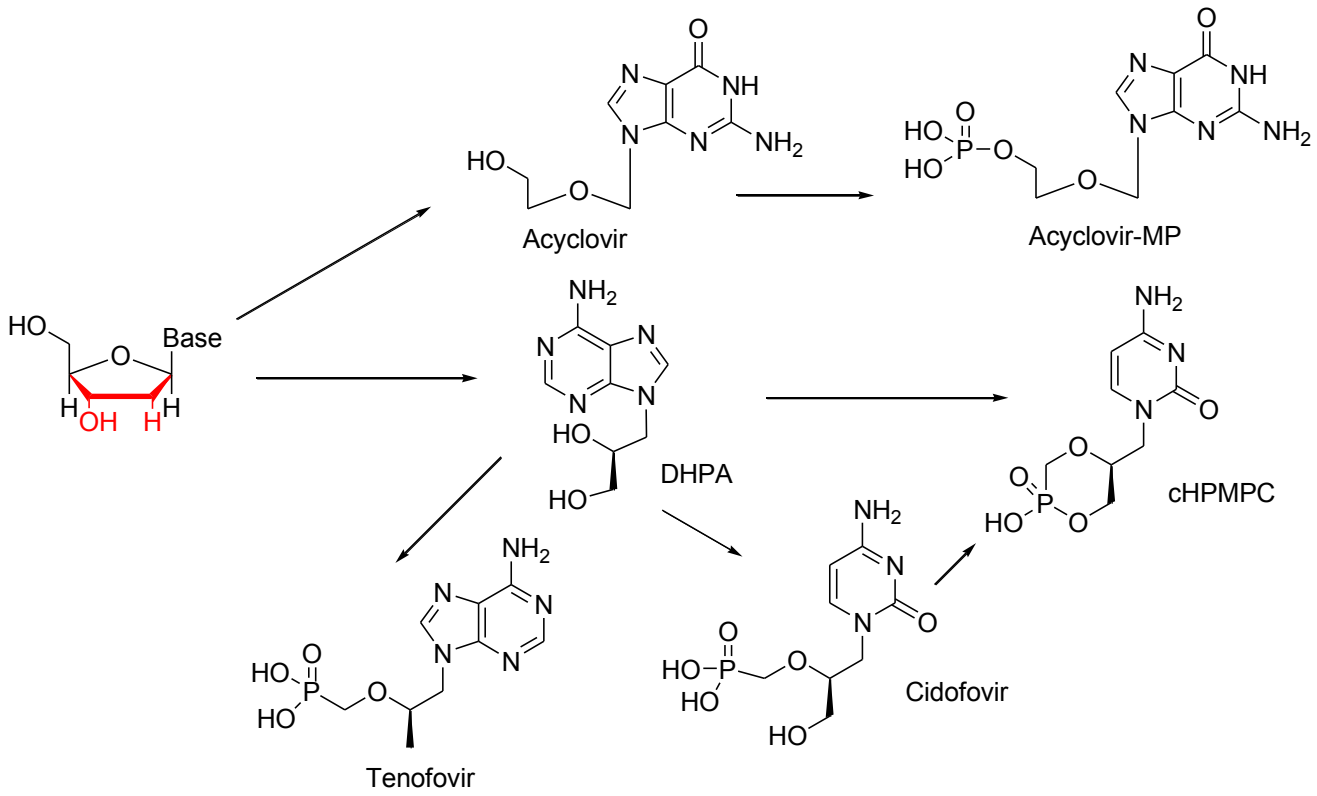


Fig. 2.3: The relationships of some well known acyclic nucleosides and nucleotides to a standard natural nucleoside.

2.2 Acyclic Nucleotide Phosphonates

The alkyl chain of acyclovir, one of the first and most successful acyclic nucleosides, has a clear yet truncated structural correlation to the carbohydrate ring in natural nucleosides.[42,45] If one simply removes the part of the ring highlighted in red in a natural nucleotide (Fig. 2.3), one is left with the alkyl chain moiety that is present in acyclovir. The success of acyclovir provided a strong case for the further development of acyclic nucleosides, and shortly after its discovery, De Clercq et al.[46] discovered the highly active (S)-9-(2,3-dihydroxypropyl)adenine (DHPA) (Fig. 2.3), which although having a slightly different mechanism of action to acyclovir, provided a strong starting point for the development of future nucleotide analogues.[42,47] The advantage of acyclic nucleosides like acyclovir and DHPA lies in the flexibility of their side chains which (due to the non-constrained nature of their sp^3 hybridized carbons) can easily adopt any conformation suitable for forming a strong interaction with the active site of the target enzyme.[42] The disadvantage with such drugs again lies in the problem of a slow initial phosphorylation as well as the labile nature of the phosphate group.

It was from this point onwards that phosphonate derivatives of bioactive molecules such as DHPA were investigated as a way of building a catabolically stable phosphate-like group into a drug while maintaining

its isopolarity with naturally occurring nucleosides.[42] This led to the introduction of a completely new class of nucleotide analogues named Acyclic Nucleotide Phosphonates or ANPs. The key feature of ANPs, relative to other nucleotide analogues, is the replacement of the labile phosphoric acid ester linkage (P-O-CH_2) with a non-cleavable phosphonate moiety (P-CH_2-O),[42,44,47] circumventing the need for the drug to be mono-phosphorylated before it is di- and then tri-phosphorylated. This has the effect of increasing the drug's biological activity; and because the phosphonate group cannot be cleaved by intracellular esterases, the drug molecule is catabolically stable (Fig. 2.4).

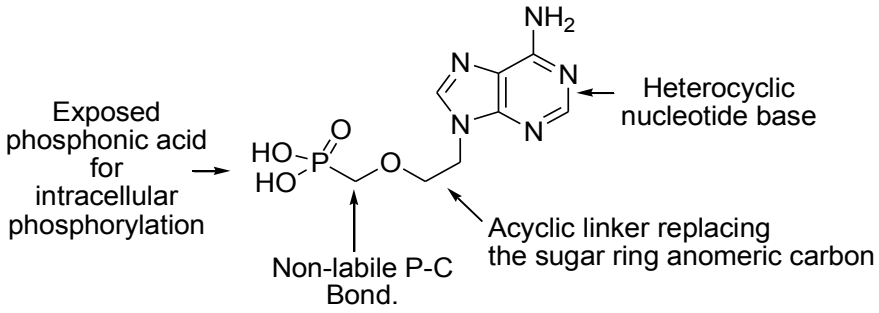


Fig. 2.4: The key structure-activity relationships of the well-known acyclic nucleotide phosphonate, adefovir.

ANPs act in a similar way to other nucleoside polymerase inhibitors in that once they are di-phosphorylated (tri-phosphorylated in the case of nucleoside inhibitors), they are used as substrate molecules by the target nucleotide polymerase and are incorporated into a growing chain of genetic material. The growth of the chain is then terminated as the drug molecule lacks the functionality required for attachments of additional nucleotides. The diagram shown in Fig. 2.5 has been modified from a review by de Clercq et al.[42] and depicts this process with regard to the HIV reverse-transcriptase inhibitor adefovir. As all viruses depend on some form of viral polymerase in order to reproduce their genomes, the ANPs cidofovir, tenofovir and/or adefovir have been shown to possess broad spectrum anti-viral activity against all DNA and retroviruses,[42] making ANPs an excellent class of anti-viral agent for future drug discovery efforts against RNA viruses. Due to the catabolic stability that the built-in phosphonate group provides to the di- and tri-phosphorylated ANP derivative, the ANPs cidofovir, adefovir and tenofovir remain active within the cell for much longer than their acyclic nucleoside counterparts acyclovir and DHPA.[42,48,49] This means that the drugs have to be administered to the patient much less frequently, in some cases with a dose interval of up to a week.[42] It is this long intracellular half-life that makes the ANPs such a valuable tool in the treatment of viral infections.

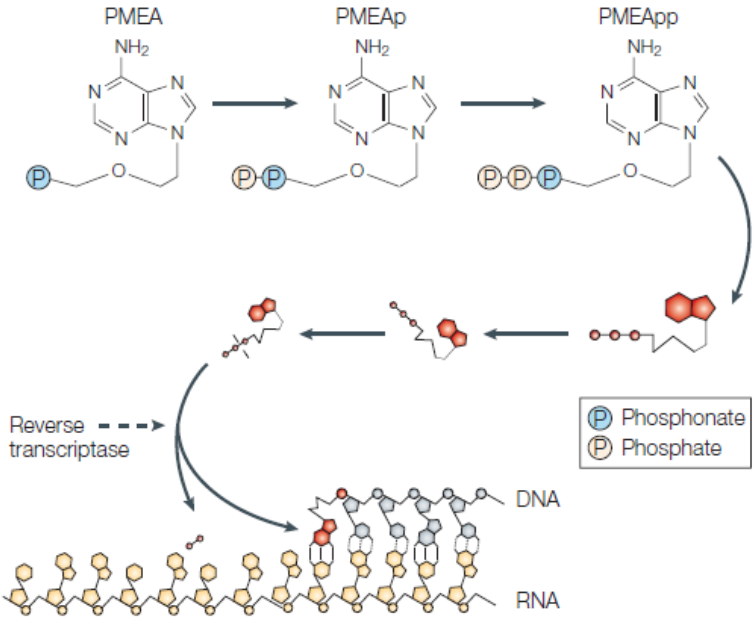
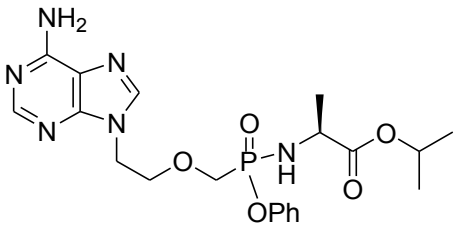


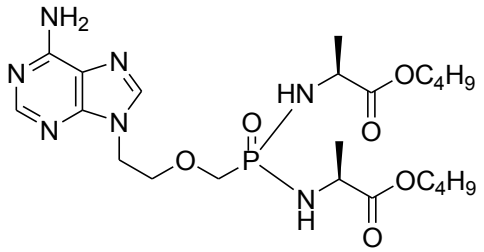
Fig. 2.5: The di-phosphorylation of adefovir, its incorporation into a growing chain of genetic material and the method by which it terminates chain growth. Taken and modified from the 2005 review article in *Nature* by de Clercq et al.[42]

2.3 Prodrugs of Acyclic Nucleoside phosphonates

As summarised by Lee et al.,[44] the effective overall potency of an ANP is a combination of its ease of phosphorylation by host phosphorylases, its catabolic stability, the metabolic stability of its mono and di-phosphonate species, its intrinsic affinity for and activity against the viral polymerase and its overall ability to permeate the host cell's membrane. The phosphonic acid nature of the acyclic nucleotide phosphonates, (Fig. 2.4), means that the drug molecule is very polar and unable to move easily through the membrane of the host cell, in fact becoming negatively charged at physiological pH.[50] The drug therefore cannot accumulate within the target cell to a high enough concentration to become biologically effective, and also suffers from the problem of very low oral bioavailability.[51] This unfortunately may mean that many potentially active, polar nucleotide analogues may have been discarded from the drug-discovery process due to their inability to cross the cell membrane during *in vivo* testing. To combat this problem, non-polar, degradable lipophilic prodrug moieties such as various phosphoramidate prodrugs[43,51–53](Fig. 2.6) and various ester moieities[47,54] including salicylate esters[43,53], alkoxyalkyl esters[43,51,55,56], acyloxyalkyl esters,[51,54] alkyloxycarbonyloxyalkyl esters[51,54] and cyclosaligenyl prodrugs,[50,51] have been introduced in order to conceal the charge of the phosphonate as a neutral phosphonate diester.[44,54] Many *in vitro* studies have shown marked increases, of 10 to 100 fold, in the anti-viral activity of ANPs where there is a prodrug moiety present, relative to the parent ANPs where the phosphonate is not masked.[44,47]



(Isopropyl L-alaninyl) phenyl PMPA



Bis(butyl L-alaninyl) PMEA

Fig. 2.6: Two different acyclic nucleotide analogue phosphoramidate prodrugs.[53]

In order for the prodrug to be successful, it needs to be chemically stable enough to survive passage through the gastrointestinal tract (GI) (in the case of orally administered drugs), absorption into the blood stream and then survive long enough to be distributed into the target cells.[51] To achieve this, the drug needs to be both water soluble and yet lipophilic enough to pass through the cell membrane. Once the prodrug has entered the cell it must then be easily degraded to expose the active phosphonic acid for phosphorylation to the active species. Alkyl esters of phosphonate and phosphates have been shown to be relatively catabolically stable and as such cannot be used as effective prodrugs;[51] therefore prodrugs need to be developed that are both stable in the extracellular environment and yet susceptible to degradation by cellular enzymes once they have entered the cell.

One of the most commonly used classes of phosphonate prodrugs are the acyloxyalkyl and alkyloxycarbonyloxyalkyl phosphonate esters. These molecules feature a carboxylate or carbonate ester that can be easily cleaved by a cellular esterase to yield a hydroxymethyl intermediate that quickly degrades through the loss of formaldehyde to expose the phosphonic acid (Fig. 2.7).[51,54,57] The release of formaldehyde during the decomposition of these drugs is of concern; however, it is thought to be a minor one due to the relatively large exposure to formaldehyde as a by-product from metabolised methanol which is present (in small quantities) in the average human diet.[51] The presence of formaldehyde and other byproducts such as pivalic acid (generated from the degradation of pivaloyloxymethyl prodrugs) may also present a problem, especially when high doses of the drugs are involved; however, two well-known commercially available drugs fall into this class of prodrug. Adefovir dipivoxil and tenofovir disoproxil fumarate (Fig. 2.8), where their respective phosphonates have been converted to the bis(pivaloyloxymethyl) (POM) and bis(isopropyloxycarbonyloxymethyl) (POC) phosphonate esters, have long been used in the treatment of HIV.[51]

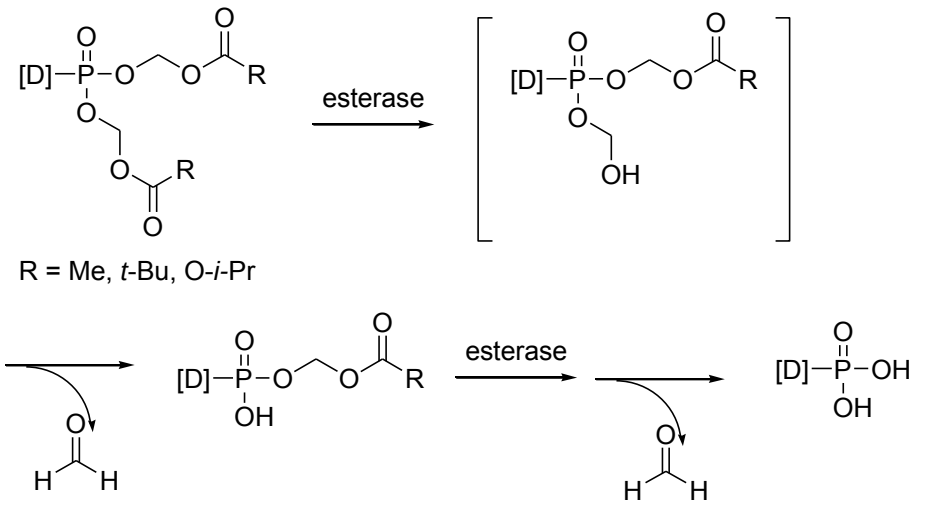


Fig. 2.7: The mechanism by which acyloxyalkyl and alkyloxycarbonyloxyalkyl prodrugs are activated by cellular esterases.[51]

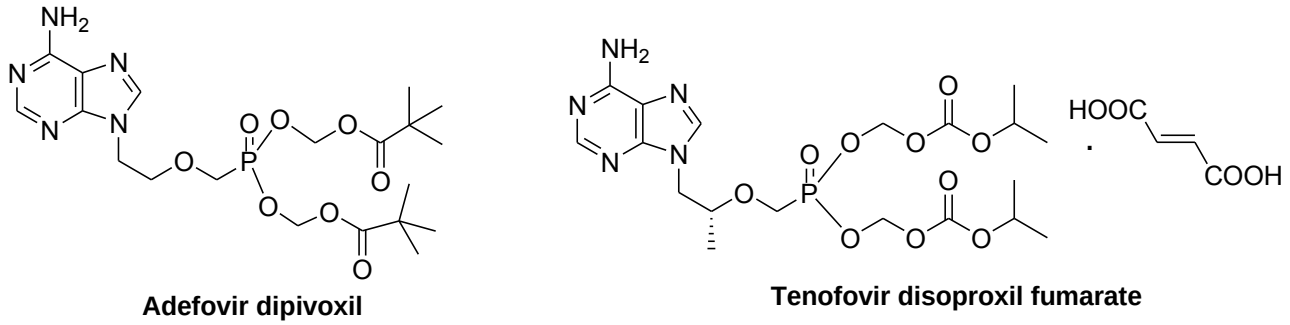


Fig. 2.8: The commercially available prodrug versions of adefovir (adefovir dipivoxil) and tenofovir (tenofovir disoproxil fumarate).[54]

Another promising relatively new type of ANP prodrug that was originally applied to nucleotide phosphate versions of the successful nucleoside NRTI d4T,[58] involves the cyclisation of 2-(hydroxymethyl)phenol on the phosphonate group of the ANP (Fig. 2.9). These prodrugs are referred to by their inventor, Professor Chris Meier of the University of Hamburg, as cyclosaligenyl or *cyclo*Sal prodrugs.[58] This class of prodrug does not achieve degradation through the enzymatic cleavage of key bonds but rather by the optimized chemical decomposition of the prodrug to its phosphonic acid.[51,58] The first step in the degradation process is the pH-sensitive, selective hydrolysis of the phosphonic acid aryl ester, which is followed by the rapid decomposition of the *O*-hydroxybenzyl intermediate to yield the active phosphonic acid (Fig. 2.9).[51,58] This approach suffers from many disadvantages, including the degradation and clearance of most of the drug from the system before it can enter the cell as well as the risk that solid-state forms of the drug may decompose too easily, making large-scale manufacture and storage of the drugs difficult or impractical.[51]

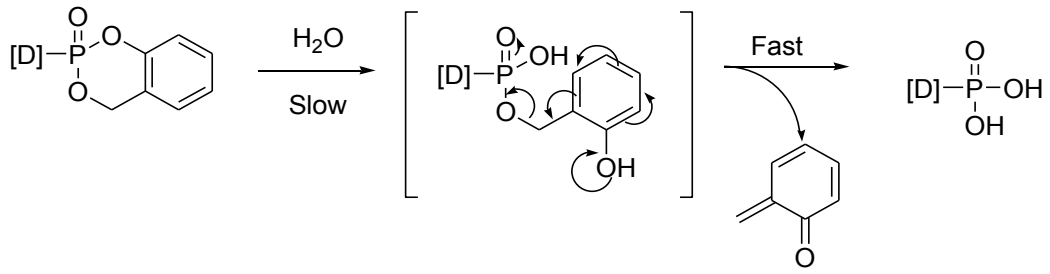


Fig. 2.9: The mechanism of cleavage of *cyclo*Sal-based prodrugs. [51]

Subsequent generations of *cyclo*Sal prodrugs have involved the addition of esters which may be enzymaticlly cleaved once in the cell in order to increase the intracellular concentration of the drug,[51,59] as well as the modification of the rate at which the prodrug cleaves via the addition of electron-withdrawing groups to the aromatic ring (Fig. 2.10).[51,60] *Cyclo*Sal prodrugs have been successfully applied to many nucleotide analogues including d4T and the ANP adefovir (Fig. 2.10); however, no *in vivo* studies have yet been carried out on this class of prodrug.[51] The synthesis of a generalized *cyclo*Sal ANP from a phosphonate protected ANP is shown below (Fig. 2.11).

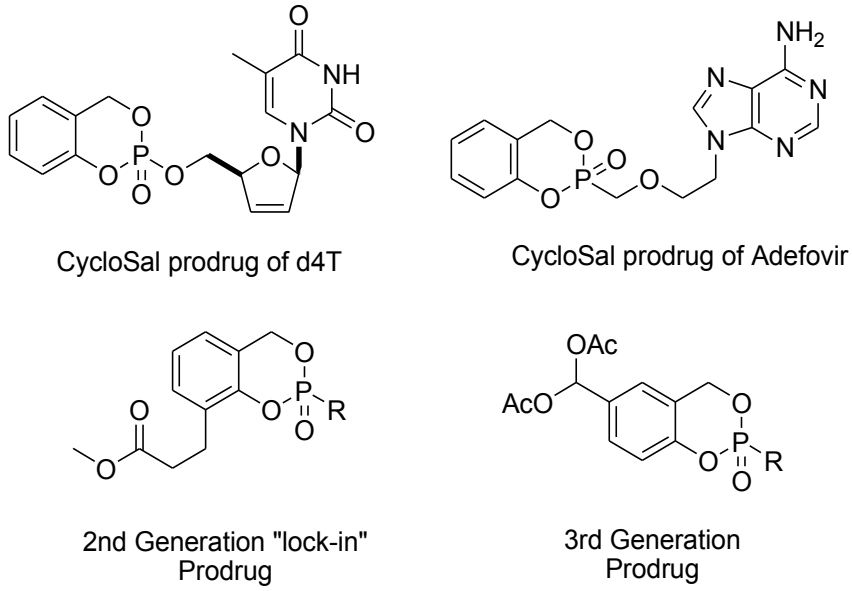


Fig. 2.10: Various *cyclo*Sal based prodrugs including 2nd and 3rd-generation *cyclo*Sal prodrug moieties.[50,51,58–60]

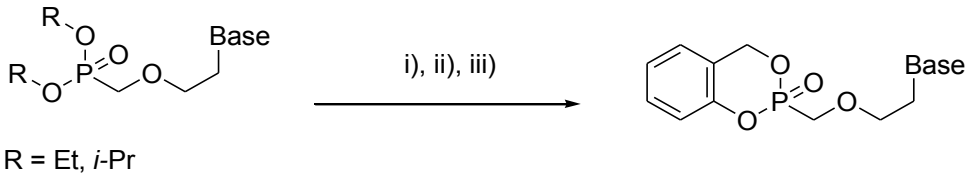


Fig. 2.11. The synthesis of *cyclo*Dal ANP's: reagents and conditions: i) TMSBr, pyridine, CH_3CN, rt, 16 hr; ii) PCl_5, CH_2Cl_2, rt, 2.5 hr; iii) Salicyl alcohol.[50]

There is currently much ongoing research into the prodrug versions of many cyclic ANP derivatives of DHPA such as HPMPC, in which the phosphonate phosphorus atom is bonded with a primary alcohol present in the side chain that forms a six-membered ring with the remaining phosphonic acid OH group masked as an ester or amide prodrug (Fig. 2.12).[53,57,61]

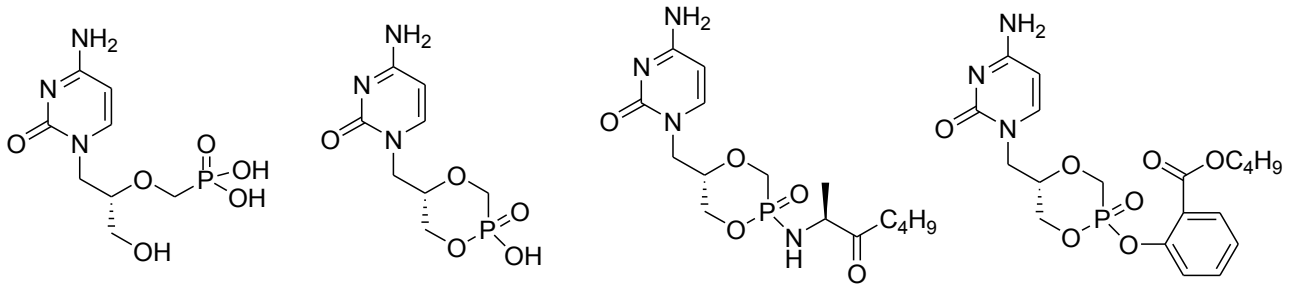


Fig. 2.12: The structures of HPMPC, its cyclised version (cHPMPC) as well as phosphoramidate and salicylate prodrug versions of cHPMPC.[53]

2.4 Future generations of acyclic nucleotide phosphonates

As mentioned before, studies have shown that modification of the heterocyclic base in nucleoside and nucleotide analogues has a marked effect on the relevant anabolic reactions within the host cell, which do not tolerate much variation in the structure of the substrates involved.[42,45] For this reason most of the successful nucleotide phosphonate analogues thus far are designed around the naturally occurring heterocyclic base rings, adenine, guanine and cytosine. There has, however, been relatively recent, promising research into subsequent generations of ANPs which feature non-natural, synthetically modified heterocyclic bases with good anti-viral activity. This section will focus on the introduction of variations to the heterocyclic base as a means of discovering and developing further ANP analogues.

2.4.1 2,6-Diaminopurine and 2,4-diaminopyrimidine derivatives

2,6-Diaminopurines or DAP derivatives were first reported to possess anti-viral activity against a range of viruses in the early 1990s.[47] Since then DAP derivatives such as (R)-PMPDAP and PMEDAP (Fig. 2.13), both close analogues of tenofovir and adefovir respectively, have shown high antiretroviral activity as well as very long intracellular half-lives and high oral bioavailability.[47,62] It is these features which make DAP derivatives promising drug candidates.[47]

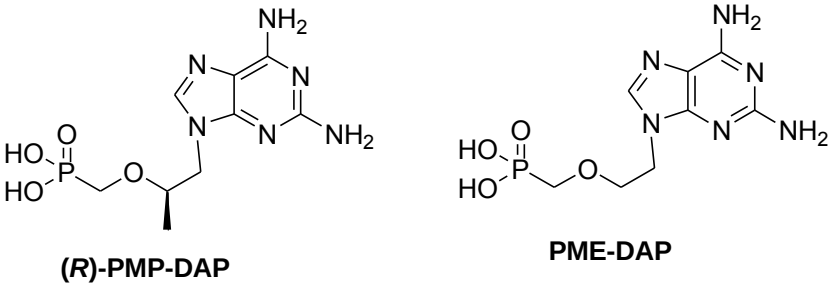


Fig. 2.13: The chemical structures of the DAP derivatives (*R*)-PMP-DAP and PME-DAP

Most of the important ANP derivatives that are used to treat human patients today, like tenofovir and adefovir, are derived from purine ring systems, with cidofovir (a cytosine derivative) being the only exception.[63] Pyrimidines like cytosine, thymine and uracil, show much less opportunity for substitution and variation than their purine counterparts; however, lately there has been a range of intriguing pyrimidine derivatives that have shown good to excellent anti-viral activity against a broad range of DNA and retroviruses (Fig. 2.14).[42,47] These "second-generation" ANP derivatives, which are based on a 6-oxysubstituted 2,4-diaminopyrimidine ring, mimic the structures of purine bases such as 2,6-diaminopurine and adenine (Fig. 2.14).[42,63,64] They have hence been dubbed "O-linked", "open ring" or "acyclic purine" ANPs,[47,62,65] with their acyclic structure allowing for a much broader array of substitution patterns than was previously allowed, especially at position-5 on the pyrimidine ring (Fig. 2.14).[63,66–69] Variation has also been introduced at the linking atom (Fig. 2.14, X) with the use of sulphur and nitrogen atoms being explored as alternatives to oxygen.[56,66]

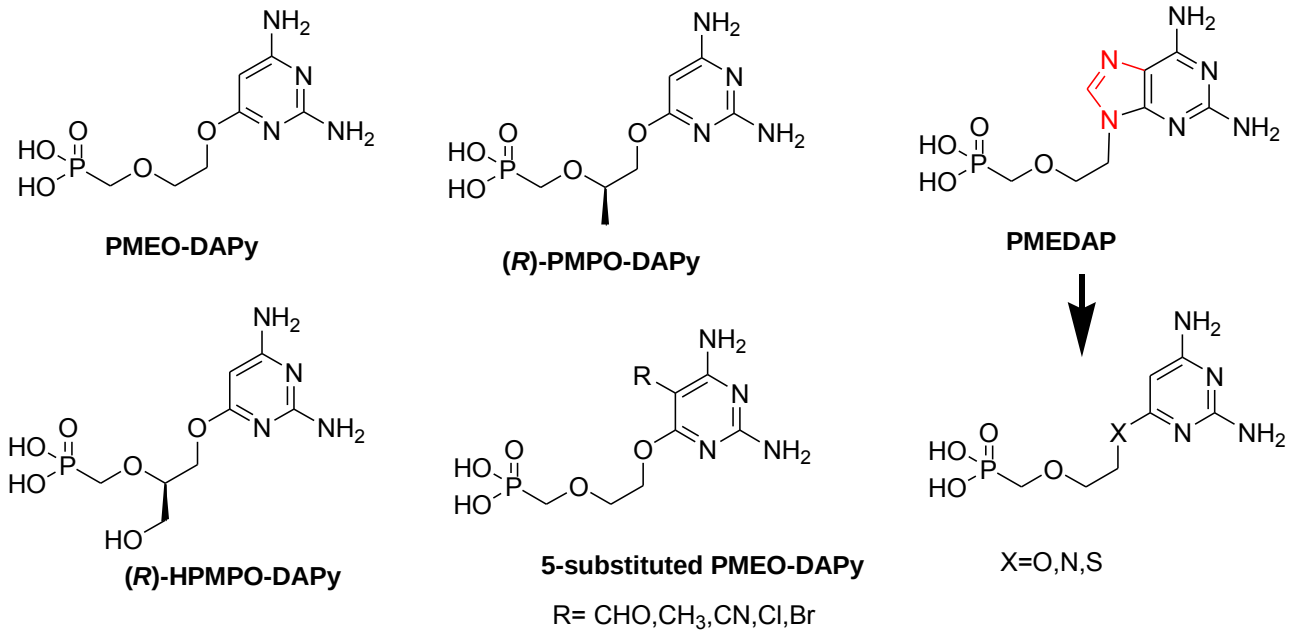
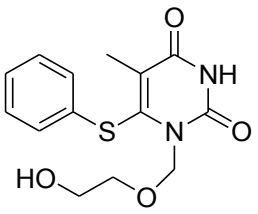


Fig. 2.14: A range of DAPy derivatives,[62,68] as well as the structural correlation of PMEDAP to a generalized DAPy structure.[56,66,68]

2.4.2 Biaryl acyclic nucleosides and nucleotides

Substitution onto the base ring is one of the most obvious ways of introducing variation into an acyclic nucleoside, thus generating novel and potentially bio-active compounds.[70] Even though radical alterations to the structures of the heterocyclic base rings were initially considered to be detrimental to biological activity,[42] research has shown that the substitution of various groupings onto both natural and synthetic nucleoside base rings can yield interesting anti-viral and anti-cancer properties. One of the best examples of this is the HIV NNRTI, 1-[(2-Hydroxyethoxy)methyl]-6-(phenylthio) thymine, abbreviated HEPT (Fig. 2.15). HEPT was originally discovered by Miyasaka et al. during an investigation into the effect of *S*-aryl substitution onto the 6-position of the thymine ring of an acyclic version of 3'-azido-3'-deoxythymidine (AZT).[71] After carrying out biological studies against HIV, Miyasaka et al. came to the conclusion that the nucleoside analogue HEPT, which was shown to be much less cytotoxic than other HIV drugs at the time, acted via a mechanism which was very different to those of other nucleoside analogues.[71] HEPT was later discovered, through X-ray crystal structure analysis, to bind to a hydrophobic pocket near the active site of HIV RT (see Chapter 1), thus making it one of the first NNRTIs.[26]



HEPT

Fig. 2.15: The structure of HEPT

As discussed by Wan et al., the effect of substituting the ring of a nucleoside base not only allows for the introduction of new and interesting biological properties into the molecule, but also gives the drug molecule some level of resistance against nucleoside or nucleotide metabolizing enzymes, thus increasing its overall potency.[70] The use of various aromatic systems in such nucleoside base ring substitutions is thought to aid the interactions of the drug with non-polar, and π-conjugated amino acids by providing a large surface for van der Vaals and π-stacking interactions to take place.[70,72] The use of non-natural nucleoside base rings, such as the triazole-3-carboxamide moiety of ribavirin or the pyrazinecarboxamide base of favipiravir, is another avenue by which novel nucleoside analogues can be generated.

A series of publications from 2005-2006, by a range of authors, explored the concept of cyclic and acyclic nucleosides of triazole-3-carboxamides cross-coupled with various aromatic groups to form novel biaryl base systems as potential anti-viral compounds.[70,72–74] In this research the authors used various cross-coupling methodologies, such as Suzuki and Sonogashira couplings (Fig. 2.16), to gain access to a range of aryltriazole and arylethynyltriazole nucleoside analogues (Fig. 2.16).[70,72–74] Some of these biaryl compounds showed good anti-viral, and in some cases anticancer activity, with compound **A** (1-((2-hydroxyethoxy)methyl)-5-(2-(4-fluorophenyl)ethynyl)-1*H*-1,2,4-triazole-3-carboxamide) (Fig 2.16) showing the best activity against HCV.[72] Removal of the rigidity of the triple bond by reduction led to the inactive compound **C** (Fig 2.17), indicating that the rigid triple bond is important to activity.[72] Zhu et al. also compared the crystal structure of compound **A** to that of HEPT, revealing a strong structural correlation between the two, indicating that compound **A** may (but not necessarily) target the RdRp of HCV (Fig. 2.18).[72] *S*-Aryltriazole acyclonucleotides have also been investigated as potential anti-HCV agents, as they have a more direct structural correlation to HEPT (Fig. 2.15).[75]

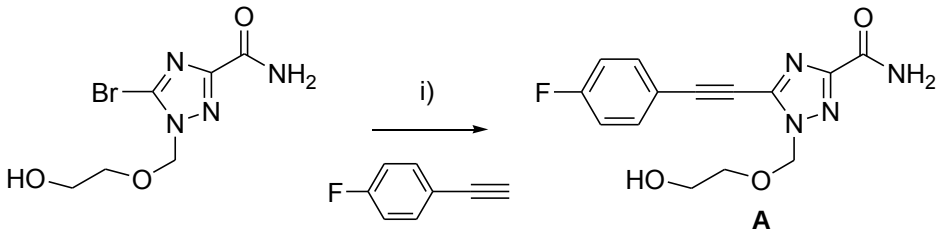


Fig. 2.16. Reagents and conditions: i) CuI, $Pd(PPh_3)_4$, Li_2CO_3, Dioxane/H_2O(3:1), 100°C, MW, 25 min.[72]

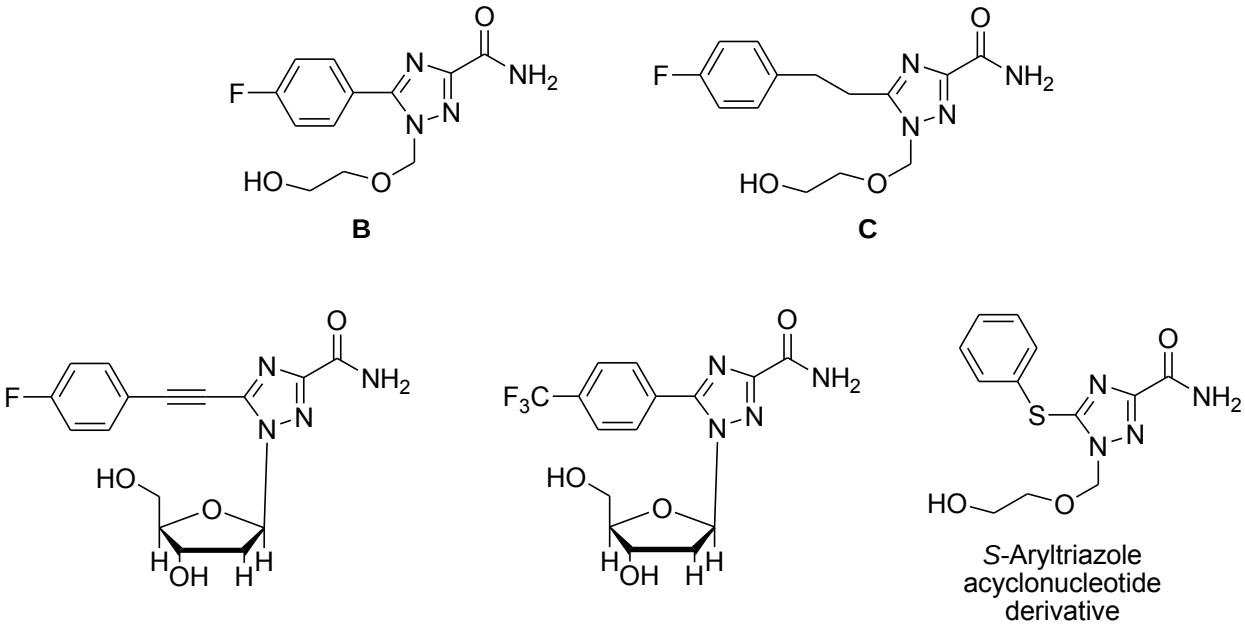


Fig. 2.17: A range of biaryl nucleoside analogues based on the triazole moiety of ribavirin.[70,72,75]

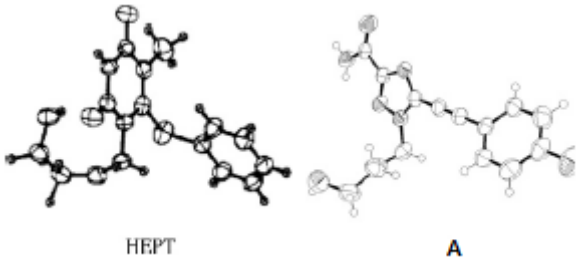


Fig. 2.18: The crystal structures of HEPT and compound **A**.[72]

Chapter 3

Aims and objectives

As highlighted in Chapters 1 and 2 of part 1 of this thesis, there is a large, unmet need for new broad-spectrum inhibitors against RNA viruses, with ribavirin, despite its relatively low potency, being the only one currently available for the treatment of RNA viral infections. Various acyclic nucleotide phosphonates have been shown to be successful against a range of viruses including HCV and HIV, while nucleoside analogues like ribavirin and favipiravir have also been shown to be successful in the past against a wide range of RNA-dependent RNA-polymerase enzymes. There are also numerous potential new avenues for drug design and discovery, such as biaryl, 2, 6-diaminopurine, and 2, 4-diaminopyrimidine acyclic nucleosides and nucleotides, which have not yet been fully explored within the context of anti-RNA virus research.

This project aimed to draw upon the success of ribavirin, in conjunction with acyclic nucleotide phosphonates. Using the structures of various ANPs, ribavirin and in some instances, favipiravir as starting points, the goal of this project was to design and synthesise a library of new potentially bioactive acyclic nucleotide phosphonates which can then be tested against a range of RNA viruses with the view to inhibit their RNA polymerases.

The prepared compounds have been built around the simple phosphonomethoxyethyl (PME) backbone of adefovir, with variation being introduced via different heterocyclic bases spanning a range of different types including natural nucleosides, DAPy and biaryl analogues, as well triazole and pyarazine derivatives with a particular focus on ribavirin analogues. In some cases pivaloyloxymethyl (POM) prodrug derivatives were synthesised from the relevant phosphonic acid ANP derivatives in order to probe the synthesis of future prodrug candidates. Fig. 3.1 below shows the generalized structures of the ANPs which have been synthesised.

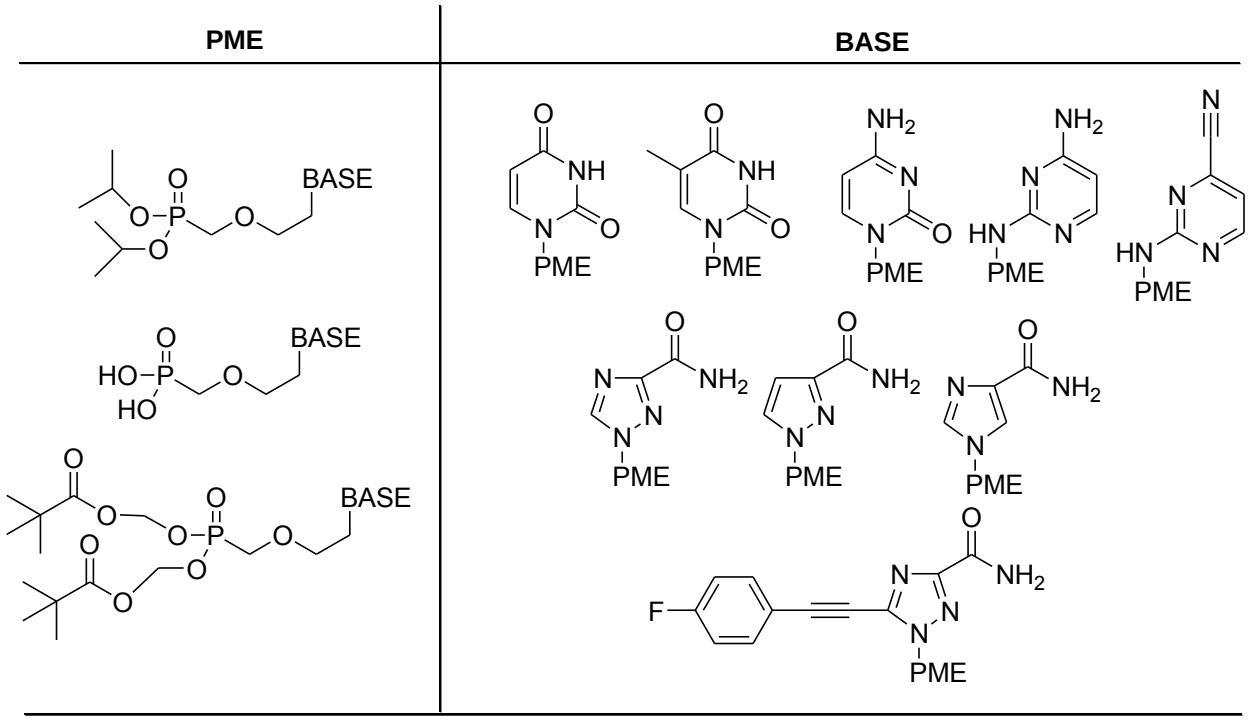


Fig. 3.1: Structures of the ANP derivatives synthesised.

The work to be described in this thesis can then be broken down into the following sections and each section will be described in detail, including a short review of the synthesis from the relevant literature, in Part 2:

i) The synthesis of the diisopropyl phosphonomethoxyethyl synthon.

ii) The alkylation of the heterocyclic base to the PME synthon.

iii) The synthesis of an arylethynyltriazole derivative via a Sonogashira cross-coupling reaction.

iv) The cleavage of the diisopropyl phosphonate ester.

v) Prodrug synthesis.

Part 2

Results and Discussion

Chapter 4

The Synthesis of the Diisopropyl phosphonomethoxyethyl Synthon

4.1 Introduction

The goal of the project was to synthesise a library of ANP derivatives, with variation introduced through the alkylation of different heterocyclic bases onto the phosphonomethoxyethyl (PME) backbone of adefovir. A semi convergent synthetic plan was developed with the goal of accessing a large number of derivatives through an advanced intermediate. A retrosynthetic analysis is shown below (Fig. 4.1). The first synthetic stage of this project involved the synthesis of a diisopropyl phosphonomethoxyethyl synthon (Fig. 4.2), from which all the other derivates could be synthesised. The next stage would involve the alkylation of the PME synthon with the relevant heterocyclic base moieties, after which the diisopropyl protecting group would be removed, and in selected cases, followed by alkylation of the phosphonic acid with a pivaloyloxymethyl (POM) group to yield a small series of pivaloyloxymethyl- acyclic nucleotide phosphonate prodrugs (Fig. 4.2).

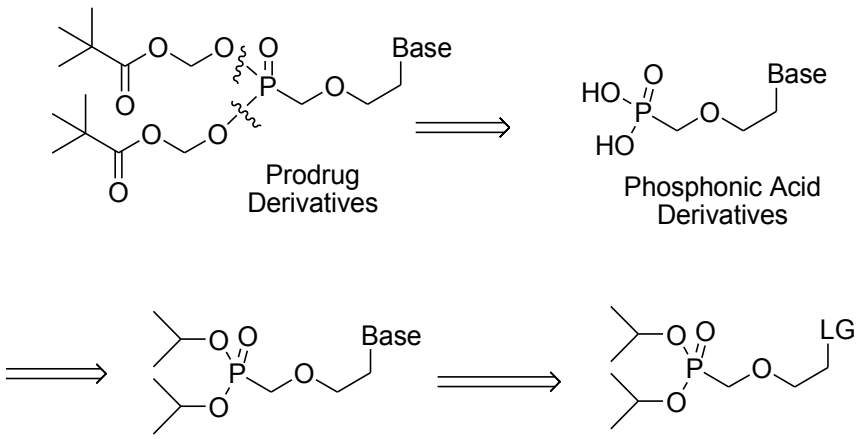


Fig. 4.1: A retrosynthetic analysis of the final prodrug to the diisopropyl PME synthon.

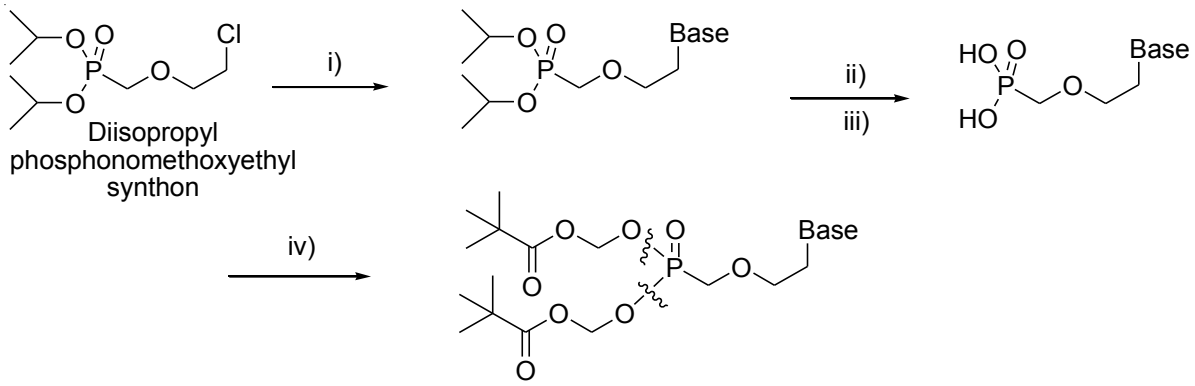


Fig. 4.2. Reagents and conditions: i) Cs_2CO_3, Base, DMF, 100°C; ii) DCM, TMSBr: iii) MeOH; iv) TEA, pivaloyloxymethyl chloride, CH_3CN or DMF.

There are two main synthetic pathways which are commonly utilised in the synthesis of different acyclic nucleotide phosphonates.[42] The first method features the etherification of a hydroxyalkyl base-bearing parent molecule with an appropriate phosphonate moiety. The hydroxyalkyl parent can be readily obtained via alkylation of the chosen heterocyclic base with a protected hydroxyalkyl compound bearing an appropriate leaving group, such as a derivatized epoxide, or a halide or sulfonate (Fig. 4.3, pathways A^1 and A^2 respectively).[42,56,65,76] The second method involves alkylation of the heterocyclic base directly onto an advanced synthon carrying the phosphonate group, built in as a protected phosphonate diester (Fig. 4.3, Route B). The choice of synthetic method is dictated by various factors such as the reactivity, selectivity and stability of the alkyl chain, phosphonate group, and/or the heterocyclic base. [42,65] As a way of generating further molecular variation, the resulting ANP molecule can be further modified through the manipulation of reactive functionalities present on the heterocyclic base ring.[42] After the parent ANP compound has been synthesised the phosphonate esters can usually be easily cleaved and modified and this process will be discussed in greater detail in Chapters 7 and 8 (Part 2) of this thesis.

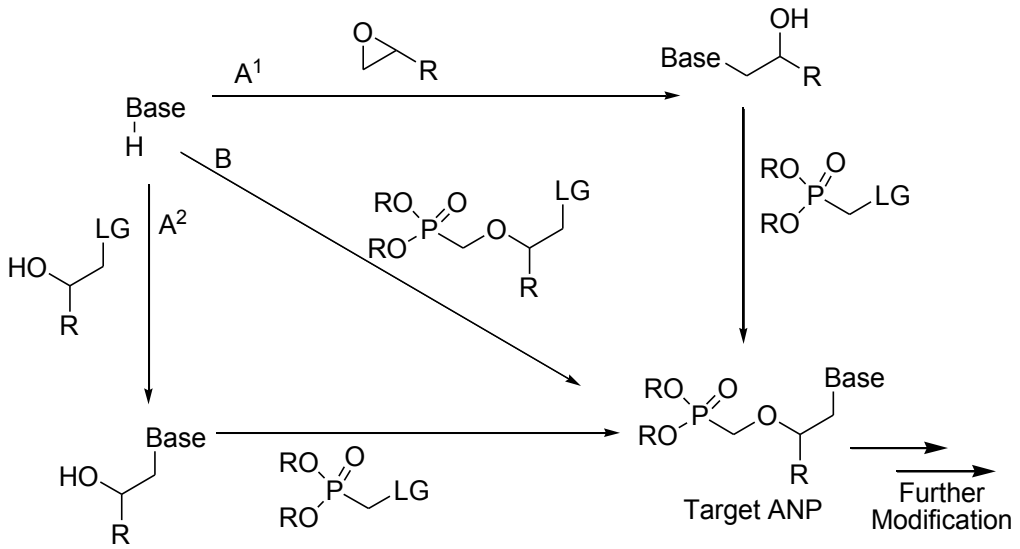


Fig. 4.3: Some of the various generalised methods of ANP synthesis.[65]

Given the aims and objectives of this project it was deemed that alkylation of the required heterocyclic base onto a protected PME synthon (Fig. 4.3, Route B), would provide the most direct as well as a semi-convergent synthetic approach, allowing a potentially large number of derivatives to be synthesised by the substitution of a heterocyclic base with a leaving group present in the synthon molecule. The synthesis of such a synthon would ideally have to be high yielding and fairly convergent so that it could be produced easily and in large quantities. Two synthetic routes towards the required advanced synthon were explored, both with advantages and disadvantages, with the second route proving to be successful. Each route is discussed below.

4.2 PME synthon synthesis (Route 1)

A synthetic route to an advanced PME synthon had been previously developed in our laboratory and used successfully in the synthesis of ANP-based bifunctional HIV RT inhibitors. This synthetic route (Fig. 4.4 and 4.5), despite being relatively long, had the advantage that it allowed for variation in the side-chain to be easily introduced through substitution of the relevant starting materials. It also allowed for the use of different leaving groups at key steps in the synthesis, which would aid in optimising the alkylation step with the heterocyclic base. A retrosynthetic analysis is shown below (Fig 4.4).

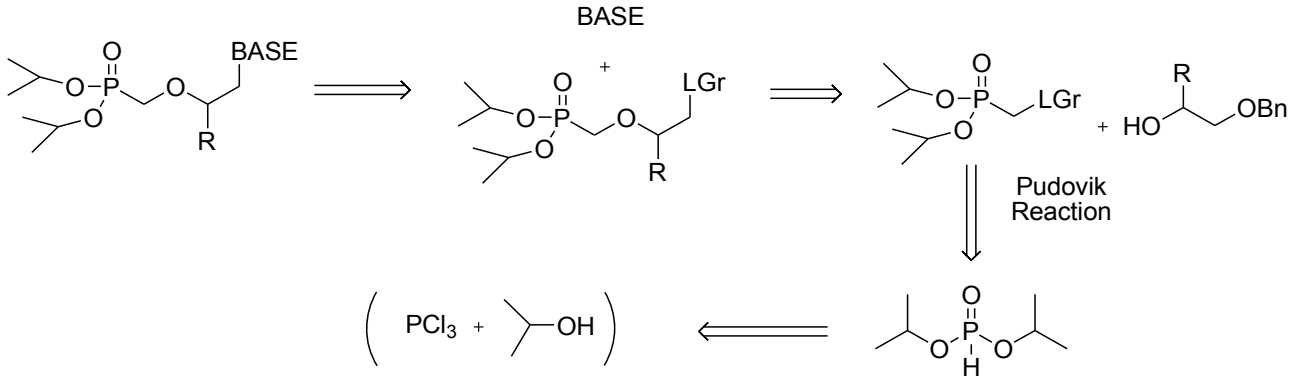


Fig. 4.4: A retrosynthetic analysis of phosphonomethoxyethyl based ANP derivatives.

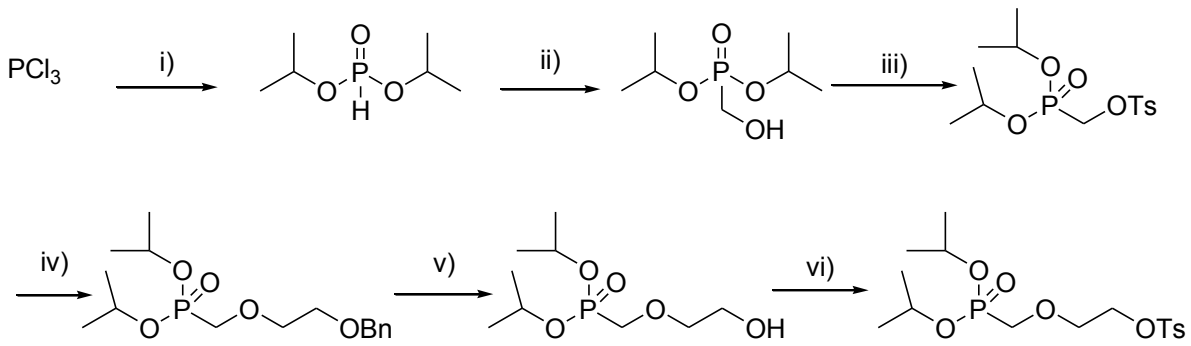


Fig 4.5. The proposed synthetic route for the synthesis of the PME synthon, reagents and conditions: i) NaH (2eq), *i*-PrOH (3eq), THF, 0°C-rt; ii) K_2CO_3, paraformaldyde, *i*-PrOH 60°C; iii) NEt_3, TsCl, DCM, DMAP, 0°C-rt; iv) 2-(benzyloxy)ethanol, NaH, THF, 0°C-reflux; v) Pd-C, MeOH/THF, rt; vi) NEt_3, TsCl, DCM, DMAP, 0°C-rt.

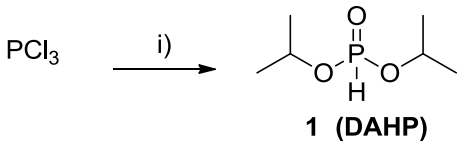


Fig. 4.6. Reagents and conditions: i) NaH, iPrOH, THF, 0°C-rt, 98%.

The first step involved the esterification of phosphorus trichloride with isopropanol in the presence of sodium hydride, to form diisopropyl hydrogen phosphonate (**1**) (DAHP) (Fig 4.6). Using a procedure outlined by Fakhraian et.al,[77] isopropanol (3 equivalents relative to PCl_3) was added dropwise to a stirred suspension of two equivalents of sodium hydride in THF after which phosphorus trichloride was slowly added. The mechanism of the reaction (Fig 4.7) shows that 2 equivalents of isopropoxide ion attack the phosphorous trichloride to form a chlorodiisopropoxy phosphite, which undergoes further substitution with isopropanol, resulting in the release of one equivalent of HCl which acts to protonate the resulting trialkyl phosphite. The protonated trialkyl phosphite then undergoes an Arbuzov-like rearrangement to afford the more stable pentavalent diisopropyl hydrogen phosphonate.[77] The addition of more than 2 equivalents of base results in a lower yield of **1**, while three or more equivalents of base completely inhibit the formation of DAHP, because of the absence of HCl. The product was then extracted into ethyl acetate from a saturated ammonium chloride solution to yield **1** (98%) which was pure enough for use in the next step. Fakhraian et.al were also able to perform this reaction without the presence of a base, although this would have meant neutralizing all the HCl produced from the reaction during the workup.[77] The ^{1}H NMR spectrum of **1** showed the presence of a doublet with a very large P-H coupling at δ_H 6.74 (d, J_{HP} = 687.2 Hz, 1H, H-P=O), along with the C-H multiplet at δ_H 4.64 (m, 2H, CH) and 2 doublets representing the diastereotopic methyl hydrogens at δ_H 1.27 (d, J = 6.4 Hz, 6H, $CH(CH_3)_2$) and 1.26 (d, J = 6.4 Hz, 6H, $CH(CH_3)_2$).

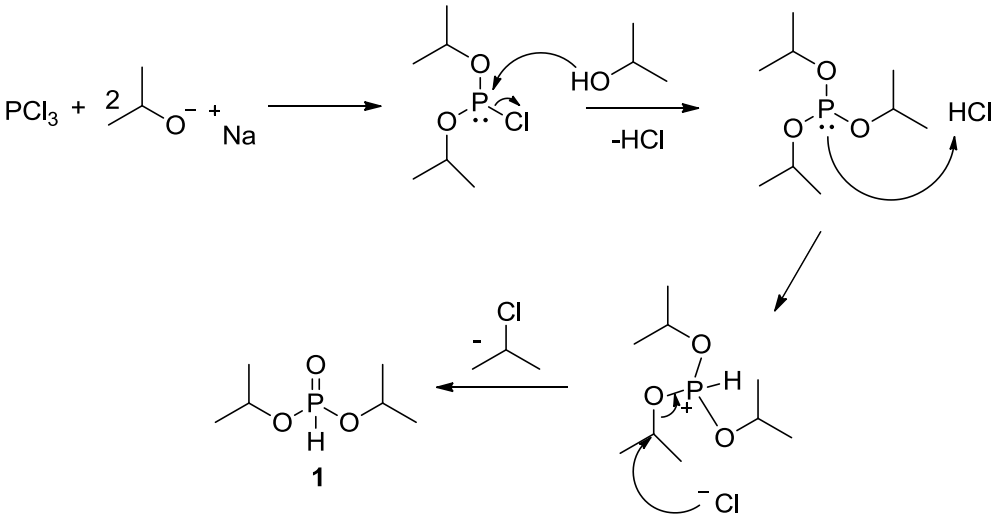


Fig. 4.7: The reaction mechanism for the formation of **1**.[77]

The next step featured the reaction of **1** with paraformaldehyde in the presence of a solid-phase base catalyst at 60°C to form diisopropyl hydroxymethylphosphonate in a reaction known as the Pudovik reaction (Fig. 4.8). The mechanism of the Pudovik reaction is shown in Fig. 4.9 and involves the deprotonation of DAHP after which the nucleophilic free electron-pair present on the phosphorus attacks the electrophilic carbonyl of formaldehyde. The resulting negatively charged and basic oxygen atom is then able to deprotonate a further DAHP molecule, forming product **2** which perpetuates the catalytic cycle.[78]

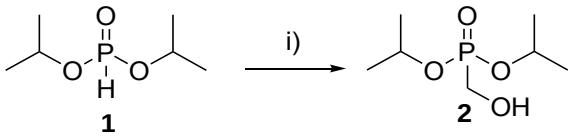


Fig. 4.8. Reagents and conditions: i) K_2CO_3, paraformaldehyde, iPrOH, 60°C, 98%.

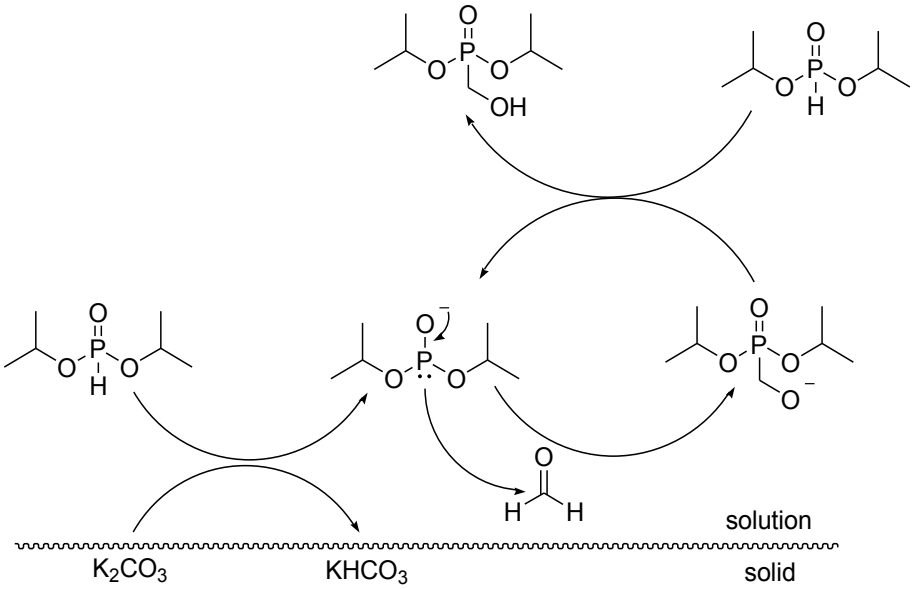


Fig. 4.9: The proposed mechanism of the Pudovik reaction.[78]

Since a non-nucleophilic, carbonate base (which is insoluble in isopropanol) is used, the Pudovik reaction has the practical advantage that the products can be easily isolated, simply by filtration of the solid potassium carbonate catalyst and the evaporation of solvent. The resulting product was pure enough for the next step of the synthesis without the need for further purification via extraction or column chromatography. The reaction had a yield of 98% and the ^{1}H NMR spectrum of **2** showed the characteristic new strong doublet at δ_H 3.79 (d, J_{PH} = 6.0 Hz, 2H, P-C<u>H</u>$_2$) representative of the new enantiotopic methylene hydrogens coupled to phosphorus. The characteristic isopropyl signals were also present as expected. The ^{31}P spectrum of **2** also showed the presence of a single major peak at δ_p 22.6, which corresponds very closely to the data reported in the literature.[78]

The next step of the synthesis involved the conversion of the hydroxyl group present in **2** to a leaving group via the reaction of **2** with *p*-toluenesulfonyl chloride in the presence of triethylamine and a small amount of dimethylaminopyridine (DMAP) as a catalyst.

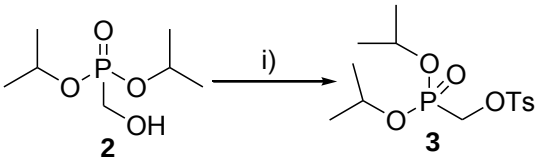
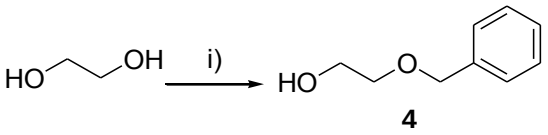


Fig. 4.10. Reagents and conditions: i) NEt_3, TsCl, DMAP, DCM, 0°C-rt, 72%.

The reaction proceeded at room temperature for 2 hours whereupon the product was extracted into ethyl acetate and purified via column chromatography to afford **3** in a moderate yield of 72%. The ^{1}H NMR spectrum of **3** showed the presence of two new aromatic signals existing as an AB system (J_{AB} = 8.2 Hz), each integrating for 2H at δ_H 7.76 and δ_H 7.33, as well as the existence of a methyl singlet at δ_H 2.42 indicating the presence of the tosyl group. Also present was the strong methylene doublet at δ_H 4.09 (J_{PH} = 10.1 Hz, 2H, P-C$\underline{H}_2$), as well as all the resonances indicative of the two isopropyl groups. The ^{13}C NMR spectrum of **3** confirmed the successful attachment of the tosylate group through the presence of all the appropriate carbon resonances.[79] The ^{13}C NMR spectrum of **3** also showed the very large diagnostic P-C coupling for the carbon bonded directly to the phosphorous at δ_c 62.1 (d, J_{PC} = 170.2 Hz, P-$\underline{C}H_2$), a resonance which is characteristic of the phosphonate bond and is present in all ANP derivatives.

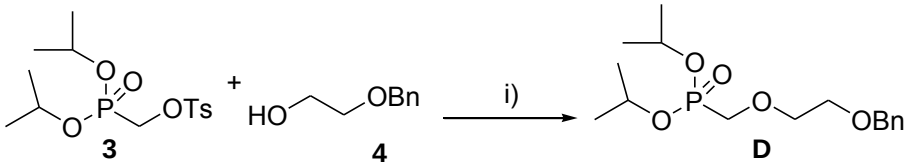
The next phase in constructing the PME synthon involved alkylation of **3** with 2-(benzyloxy)ethanol **4** which was readily synthesised by the reaction of ethylene glycol (excess) with 1 equivalent of benzyl bromide in the presence of sodium hydride (Fig 4.11). The sodium hydride was added to a stirring solution of excess ethylene glycol in THF to ensure mono-deprotonation, after which benzyl bromide was added as the limiting reagent. After refluxing for 12 hours, **4** could be extracted and isolated in a good yield of 77%, with the ^{1}H and ^{13}C NMR data for the purified material corresponding to the published data for this compound.[80]



4

Fig. 4.11. Reagents and conditions: i) BnBr, NaH, THF, reflux, 77%.

With **3** and **4** now in hand, an alkylation of the two was attempted using sodium hydride as a base in THF at reflux temperature for 12 hours under a nitrogen atmosphere (Fig 4.12). This reaction produced the required product **D** but only in very low conversions of 20-30%. Isolation of the product was complicated by the fact that the product and the tosyl starting material **3** both had very similar R_f values, even across a wide range of suitable solvent systems. Preparative thin layer chromatography (TLC) with repeated developments in a non-polar solvent system was also unable to isolate the required product with sufficient purity. The low yield of this reaction was also a significant problem, since large quantities of the final product were required as a precursor to the advanced synthon which would be used to synthesise all future target ANP derivatives.



3 **4** **D**

Fig. 4.12. Reagents and conditions: i) NaH, THF, reflux.

Numerous attempts to optimize and improve the reaction of **3** with **4** were attempted, which included the investigation of different solvents, different ratios of starting materials and reagents, reaction temperatures and in some cases the introduction of catalytic iodide (formation of iodide *in situ*) and alternative bases (Table 4.1). Various solvents were investigated including THF, acetonitrile (CH_3CN), and dimethoxyethane (DME). Caesium carbonate was used as an alternative base in CH_3CN but this did not yield any conversion of starting materials to products. The reaction was also performed over a wide range of temperatures to rule out the concern that some product was being lost to thermal decomposition. In some cases catalytic amounts of sodium iodide or tetrabutyl ammonium iodide (TBAI) were introduced to the reaction to increase the rate of alkylation via the formation of a more reactive iodinated intermediate. Compound **4** was always used in excess relative to **3** (with a minimum ratio of 2:1) in order to make the chromatography easier by reducing the amount of starting material. Unfortunately this was in vain as the reaction never reached completion in any experiment.

Table 4.1. The various reaction conditions used in the attempted alkylation of **4** and **3** to produce **D.**

Reaction	Base	Solvent	Temperature	Time	Iodide	Approx. Conversion (by tlc)
1	NaH	THF	reflux	18hrs	none	20-30%
2	Cs_2CO_3	CH_3CN	reflux	48hrs	none	No reaction
3	NaH	THF	40°C	48hrs	TBAI	20%
4	NaH	THF	rt	1 week	NaI	10%
5	NaH	CH_3CN	reflux	48 hrs	TBAI	10%
6	NaH	DME	reflux	48 hrs	TBAI	no reaction

To further investigate the alkylation of **4** and **3**, different combinations of leaving groups were explored. Tosyl, mesyl and iodo variants were synthesised from compounds **4** and **2** (Fig 4.13), and used as starting materials in the synthesis of **D.** The alkylation reaction was performed on the new substrates in the presence of sodium hydride in THF under a nitrogen atmosphere at reflux temperature with no success (Table 4.2).

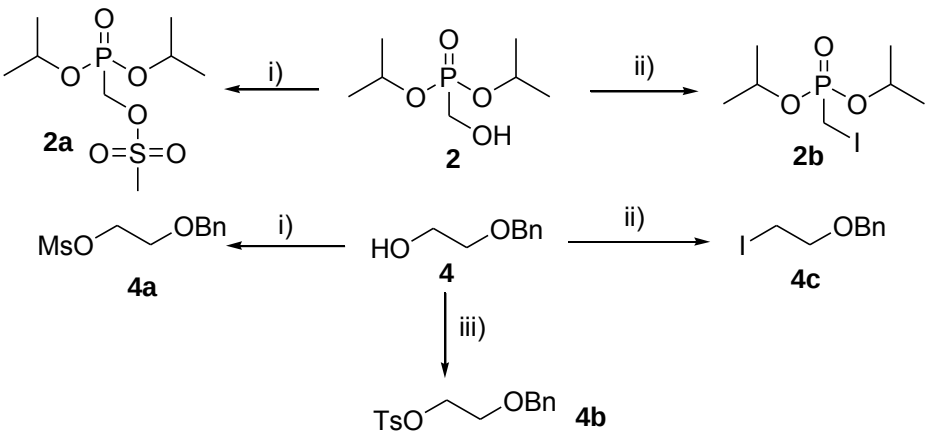


Fig. 4.13. Reagents and conditions: i) MsCl, NEt_3, DCM, DMAP, rt; ii) I_2, PPh_3, imidazole, rt; iii) TsCl, NEt_3, DCM, DMAP, rt.

The syntheses of compounds **2a**, **4a** and **4b** were carried out under the standard mesylation and tosylation conditions which were broadly discussed above for the synthesis of the tosyl derivative **3.** The iodinations of **4** and **2** were carried out at room temperature with use of iodine, triphenylphosphine and imidazole as a base. In this case the reaction mixture was worked up with a saturated sodium thiosulfate solution to remove any iodine, and extracted with ethyl acetate, after which the iodinated products were isolated by column chromatography. The mechanism of this reaction proceeds via the formation of a phosphonium salt and is shown below (Fig 4.14).

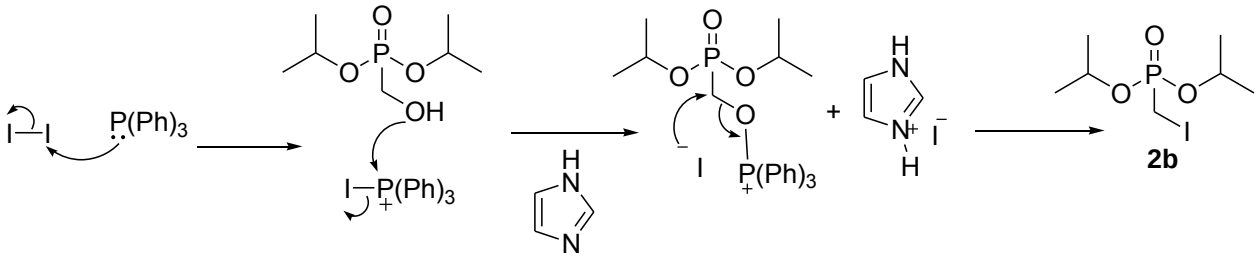


Fig 4.14: The mechanism of formation of **2b**.

Table. 4.2: The various reaction conditions used as well as the nucleophile / electrophile combinations used in the attempted synthesis of **D.**

Reaction	Base	Nucleophile	electrophile	Solvent	Temperature	Time	Observation
1	NaH	**4**	**2a**	THF	reflux	48 hrs	no reaction
2	NaH	**4**	**2b**	THF	reflux	48 hrs	no reaction
3	NaH	**2**	**4a**	THF	reflux	48 hrs	no reaction
4	NaH	**2**	**4b**	THF	reflux	48 hrs	no reaction
5	NaH	**2**	**4c**	THF	reflux	48 hrs	no reaction

Owing to the very low yields of the reactions of **4** and **3** and the failure of all other attempts to find better leaving groups, the decision was made to abandon this particular synthetic route, despite the fact that it had been previously completed with relative success in our laboratory. A possible explanation is that the electron-withdrawing phosphonate group destabilises the partial positive charge in the S_N2 transition state. In addition, the alkoxide species is a basic, poorly nucleophilic species.

4.3 PME synthon synthesis (Route 2)

After the failure of the first route to produce the required PME synthon, our attention turned to a synthetic route which had been used and published by Holý et al. in 1999.[81] This route makes use of an Arbuzov reaction as its key step in which triisopropyl phosphite reacts chemoselectively with chloro-2-(chloromethoxy)ethane to produce a diisopropyl phosphonomethoxyethyl synthon bearing a chlorine atom as a leaving group. This route has the advantage that it comprises only two steps and can be performed on a relatively large scale. The main disadvantages of this particular synthesis are all associated with the synthesis of 1-chloro-2-(chloromethoxy)ethane, which is synthesised from highly toxic 2-chloroethanol and paraformaldehyde using large quantities of hydrogen chloride gas. Another disadvantage to this approach is that it limits the possibility of easily introducing variation in the PME side chain and also restricts the synthesis to having a chorine as leaving group at the end of the chain. The latter, however, turned out to be a non-issue.

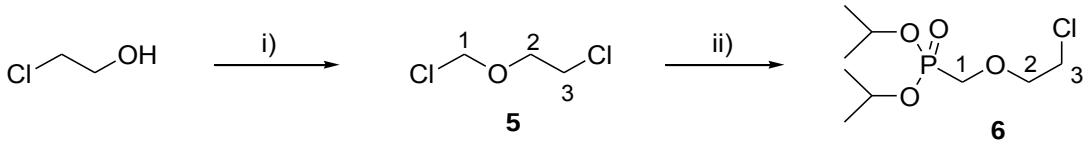


Fig. 4.15: The synthesis of the final PME synthon. Reagents and conditions: i) paraformaldehyde, $HCl_{(g)}$, DCM, 0°C, 61%; ii) triisopropyl phosphite, 100°C, 83%.

The procedure for the synthesis of **5** was adapted from procedures described by Liu et al.[82] and Powell et al.[83] and began with formation of a rapidly stirring suspension of paraformaldehyde and 2-chloroethanol in dry dichloromethane (DCM) at 0°C in a sealed two-necked flask fitted with a rubber septum. HCl gas, generated from the reaction of stirring solid NaCl with concentrated H_2SO_4 in a separate two-necked flask, was constantly bubbled into the reaction mixture via a piece of rubber tubing and a long steel needle, for 6 hours at 0°C after which the mixture was refrigerated overnight. The next day, HCl gas was again bubbled through the mixture to ensure complete saturation of the solution. Due to the highly non-polar and toxic nature of the products and starting materials, it was impossible to use tlc to monitor the reaction, thus NMR analysis of the crude material was used to follow the reaction's progress. One was able to observe the appearance of new product peaks in the ^{1}H NMR spectrum at δ_H 5.52 (s, 2H, H-1) and δ_H 3.95 (t, J = 5.6 Hz, 2H, H-2) while the integrations of the starting material peaks decreased over time. When NMR analysis had

revealed a 75% conversion of starting materials to products, the reaction was stopped and the solution dried with calcium chloride. Product **5** was then distilled off from the less volatile, excess 2-chloroethanol under reduced pressure with the product boiling at 64°C under 20 mmHg. Both the spectral and boiling point data for this compound were found to correspond to the data reported in the literature.[82,83]

The mechanism of this reaction begins with the protonation of the formaldehyde oxygen which increases the electrophilicity of the carbonyl carbon thus aiding the addition of the poorly nucleophilic oxygen of the 2-chloroethanol. The hydroxyl group on the resulting hemiacetal picks up a further proton through a proton transfer and is expelled as water to form an oxocarbenium ion, which then undergoes addition with a chloride ion to produce **5**. (Fig 4.16).

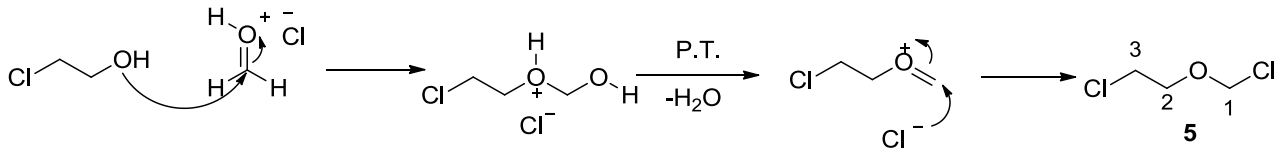


Fig. 4.16: The mechanism of the formation of **5**

The final PME synthon was synthesised in good yield via a chemoselective Arbuzov reaction of **5** with triisopropyl phosphite. The Arbuzov reaction involves an S_N2 reaction between the phosphorus of a trialkyl phosphite and an alkyl halide to form a phosphonium intermediate which reacts with the displaced halide ion to yield a phosphonate and a new alkyl halide via a dealkylation step (Fig 4.17).[84] From the procedure by Holý et al.,[81] neat triisopropyl phosphite was heated to 100°C in a reaction flask fitted with a Vigreux column, still-head, condenser and thermometer. Chloride **5** was slowly added to the stirring solution at such a rate that the resulting 2-chloropropane by-product could be distilled off. It was observed that the triisopropyl phosphite reacts chemoselectively with carbon 1 over carbon 3 of compound **5,** since this is the most electrophilic carbon in the molecule due to stabilization of the incipient positive charge in the transition state by the adjacent oxygen (Fig 4.17). This level of chemoselectivity suggests that side reactions with the 2-chloropropane by-product are unlikely, aside from the fact that the 2-chloropropane is being constantly removed from the reaction. After the reaction was complete and 2-chloropropane was no longer being evolved, the product PME synthon could then be isolated directly from the reaction mixture by distillation under high vacuum, with the product boiling at 120°C at 0.1 mmHg, which corresponds exactly to the reported literature boiling point of 120°C at 0.1 mmHg.[81]

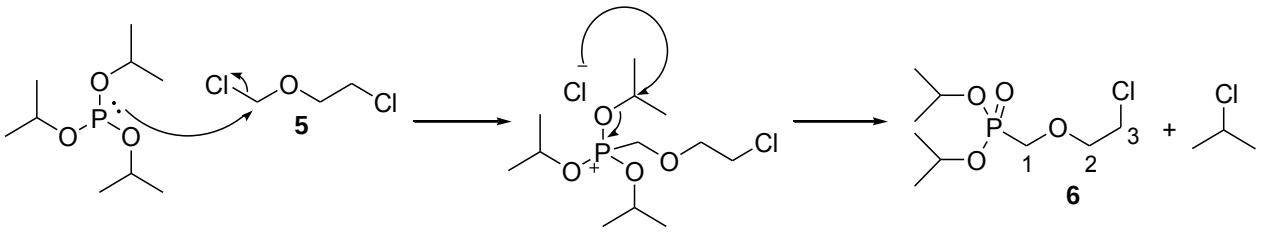


Fig. 4.17: The mechanism of the formation of **6** via an Arbuzov reaction.

The ^{1}H NMR spectrum of **6** revealed the expected triplets at δ_H 3.85 (t, *J* = 5.8 Hz, 2H, H-2) and δ_H 3.63 (t, *J* = 5.8 Hz, 2H, H-3), as well as the diagnostic doublet at δ_H 3.79 (d, J_{PH} = 8.4 Hz, 2H) which is formed from the strong J^2 coupling of the phosphorus with the hydrogens on C-1, confirming the formation of the P-C bond. The doublet of heptets at δ_H 4.76 (dhept, $J_{HP,\ HH}$ = 7.7, 6.3 Hz, 2H) is representative of the C<u>H</u>$(CH_3)_2$ hydrogens on the isopropyl groups of the phosphonate and arises first from the large coupling of the phosphorus to the hydrogens (J_{HP} = 7.7 Hz) and then from the slightly smaller couplings associated with the six hydrogens on the adjacent isopropyl methyls (J_{HH} = 6.3 Hz). The isopropyl methyl groups are diastereotopic and are therefore non-equivalent. They might have been present in the ^{1}H NMR spectrum as two doublets, each with a coupling constant of J_{HH} ≈ 6-7 Hz. However they did not resolve and were instead visible as one doublet with a coupling constant of J_{HH} = 6.3 Hz. The ^{13}C NMR spectrum did however resolve the diastereotopic methyl carbons into 2 doublets at δ_c 24.3 and δ_c 24.1, each with a small coupling (J_{CP} = 4.2 Hz) due to J^3 coupling with phosphorus (Fig 4.17, <u>C</u>H_3). The <u>C</u>H isopropyl carbon was represented as a doublet with a J^2 PC coupling at δ_c 71.4 (d, J_{PC} = 6.7 Hz, <u>C</u>H$(CH_3)_2$) (Fig 4.18), while C-2 appeared as a doublet with a J^3 PC coupling at δ_c 73.2 (d, J_{PC} = 10.6 Hz)(Fig 4.18, resonance 2), which is counterintuitive as one would expect the resonance with a larger coupling constant to have a closer proximity to the phosphorus. (These assignments were confirmed by the HSQC in conjuction with the ^{13}C NMR data of other ANP derivatives later on in this work and will be discussed in detail in chapter 5 of this thesis.) C-3 appeared as a singlet at δ_c 42.5 as it is too far away from the phosphorus for PC coupling. By comparison C-1 appeared as a doublet with a very large coupling constant at δ_c 66.3 (d, J_{PC} = 167.6 Hz) (Fig 4.18, resonance 1) arising from a J^1 P-C coupling, a resonance characteristic of the P-C bond of a phosphonate.

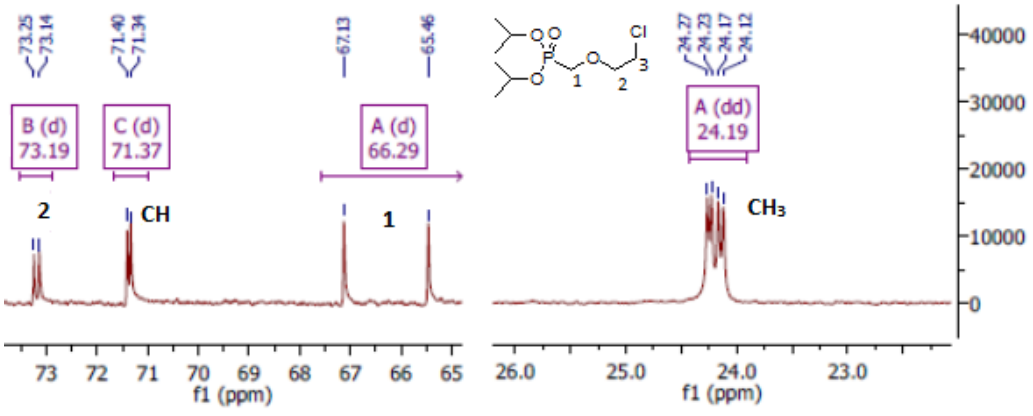


Fig. 4.18: Expansions of the ^{13}C NMR spectrum of **6.**

Chapter 5

Alkylation of the PME synthon with heterocycles

5.1 Introduction

With an advanced PME synthon in hand, the next phase in the project was to introduce selected heterocycles via alkylation. A large variety of both natural and synthetic basic heterocyclic rings were selected with a particular interest in derivatives of the 5-membered triazole ring of ribavirin. Most bases were available commercially; however, a number of bases need to be synthesised *de novo* or prepared via modification of a commercially available precursor and the synthesis of these compounds will also be discussed in this chapter.

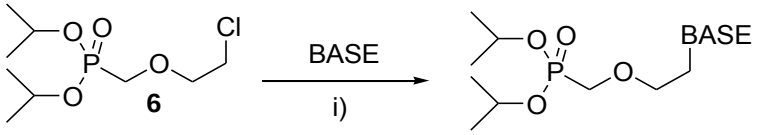


Fig. 5.1. Reagents and conditions: i) Cs_2CO_3, DMF, 100°C

The general procedure for these reactions was derived from various works by Holý et al.,[66,81] in which they explored the use of various inorganic bases such as sodium hydride, caesium carbonate, as well as organic amine bases such as diisopropylethylamine (Hünig's base), 1,8-bis(dimethylamino)naphthalene (proton sponge) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) in dimethylformamide (DMF).[81] The results of these studies showed that sodium hydride and DBU were both suitable bases for these alkylation reactions, although caesium carbonate was shown to be a superior base overall, both in terms of yield and reaction time. Thus, despite being a rather expensive reagent, caesium carbonate was selected as the base for our purposes.

A typical experimental procedure follows: The relevant heterocyclic base was suspended with the PME synthon **6** and caesium carbonate in dry DMF and warmed to 100°C. Reaction times varied from 6 – 48 hours depending on the heterocycle involved. After tlc analysis had revealed reaction completion, the DMF was distilled off from the reaction mixture under high vacuum and the organic material extracted from aqueous sodium bicarbonate into chloroform/methanol (2:1). The presence of DMF was shown to pull the organic product into the aqueous layer as well as interfere with the subsequent chromatography, thus it was removed by distillation early so as to avoid these problems. Acetonitrile was investigated as an alternative solvent for these alkylation reactions as it is much easier to work with and remove but, in most cases, it was found to be an inferior solvent to DMF, resulting in much lower yields and much longer reaction times. In most, but not all cases, ethyl acetate and DCM were found to be too non-polar to properly extract the highly polar ANP products from the inorganic layer in the work-up step and it was for

this reason that a methanol/chloroform extraction was used. After the dried crude product was obtained from the work-up, the desired product was then isolated in all cases by column chromatography.

5.2 The synthesis of ANP derivatives with heterocycles of natural nucleosides.

A series of ANP analogues were synthesised containing the naturally occurring pyrimidine bases of uracil, thymine (5-methyluracil), and cytosine.

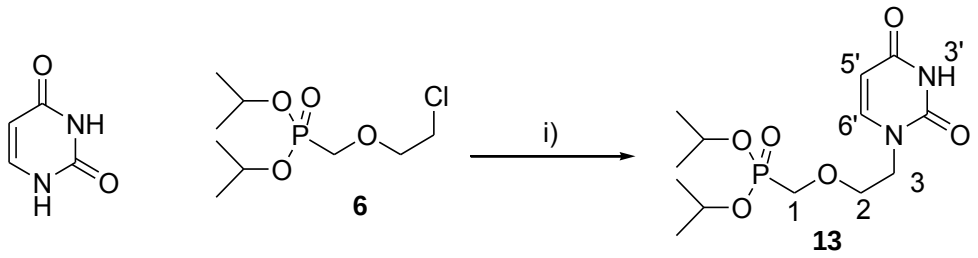


Fig. 5.2. Reagents and conditions: i) Cs_2CO_3, DMF, 100°C, 41%.

The first ANP derivative to be synthesised was the uracil derivative **13** which was synthesised from the general procedure described above (Fig 5.2). The crude material, which displayed only one dominant product, was isolated via column chromatography to yield **13** in a modest yield of 41%. The reaction was found to be ^{1}N-chemoselective as expected, with alkylation occurring on the more reactive ^{1}N-anion as opposed to the less nucleophilic, more resonance-stabilised ^{3}N-anion. The ^{1}H NMR spectrum of **13** confirmed the alkylation of the uracil ring through the slight downfield shift of the triplet for the H-3 protons from δ_H 3.63 (t, *J* = 5.7 Hz, 2H) in the chloride **6**, to δ_H 3.96 (t, *J* = 5.0 Hz, 2H). Other diagnostic signals in the ^{1}H NMR spectrum included the broad N-H singlet at δ_H 10.13 (H-3'), the uracil hydrogen AB paired doublets at δ_H 7.55 (d, *J* = 7.9 Hz, 1H, H-6') and at δ_H 5.53 (d, *J* = 7.9 Hz, 1H, H-5'), the phosphonate P-H coupled doublet at δ_H 3.79 (d, J_{PH} = 8.2 Hz, 2H, H-1) and the two diasterotopic methyl doublets at δ_H 1.29 (d, *J* = 6.0 Hz, 6H, $CH(C\underline{H}_3)_2$) and δ_H 1.26 (d, *J* = 6.0 Hz, 6H, $CH(C\underline{H}_3)_2$). The ^{13}C NMR spectrum showed both carbonyl resonances at δ_C 164.4 (C-2') and δ_C 151.9 (C-4'). All ^{1}H and ^{13}C assignments were confirmed with the use of HSQC NMR spectroscopy but, owing to the very close nature of the ^{13}C NMR resonances at δ_C 71.4 (d, J_{PC} = 10.4 Hz, C-2) and δ_C 71.2 (d, J_{PC} = 6.4 Hz, $\underline{C}H(CH_3)_2$) in the HSQC spectrum of **13**, C-2 and $\underline{C}H(CH_3)_2$ were assigned based on the coupling constants of those signals in the HSQC spectra of analogous ANP derivatives (discussed in greater detail below). These assignments were confirmed by the ^{13}C NMR data of the diisopropyl deprotected variant of **13,** compound **25**, which will be discussed in Chapter 7.

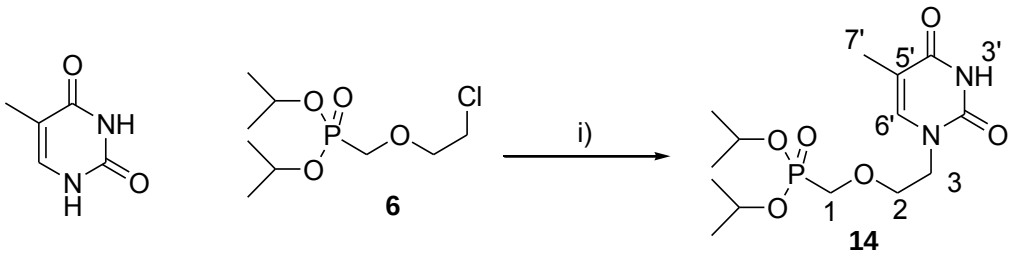


Fig. 5.3. Reagents and conditions: i) Cs_2CO_3, DMF, 100°C, 45%.

The next ANP derivative to be synthesised was the 5-methyluracil (thymine) derivative **14** (Fig 5.3). The compound was produced in much the same way as **13** with column chromatography being used to isolate **14** in a yield of 45%. Again, ^{1}N-chemoselectivity was observed. IR spectroscopy showed the presence of the two carbonyls with absorptions at 1709 cm^{-1}(C=O) and 1688 cm^{-1} (C=O), while the presence of the phosphonate group was confirmed by peaks at 1237 cm^{-1} (P=O) and 993 cm^{-1} (P-O). The ^{1}H NMR spectrum of **14** showed all the resonances associated with the diisopropyl PME moiety and also confirmed the presence of the thymine ring by virtue of an N-H singlet at δ_H 9.88 (s, 1H, H-3'), the H-6' singlet at δ_H 7.40 (s, 1H, H-6'), and the methyl singlet at δ_H 1.82 (s, 3H, H-7'). The ^{13}C NMR spectrum showed the presence of the two carbonyl peaks at δ_C 165.2 (C-2') and δ_C 152.0 (C-4') as well as a weak singlet at δ_C 109.5 indicative of the quaternary centre at C-5'. The large P-C coupled doublet at δ_C 66.2 (d, J_{PC} = 166.5 Hz, C-1) and the two diastereotopic methyl doublets at δ_C 24.3 (d, J_{PC} = 4.2 Hz, CH(<u>C</u>H$_3$)$_2$) and δ_C 24.1 (d, J_{PC} = 4.2 Hz, CH(<u>C</u>H$_3$)$_2$) confirmed the integrity of the phosphonate moiety throughout the course of the reaction.

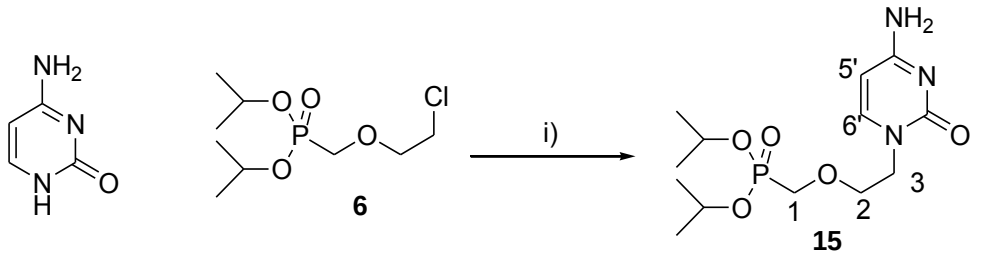


Fig. 5.4. Reagents and conditions: i) Cs_2CO_3, DMF, 100°C, 66%.

Finally the cytosine derivative **15,** which had previously been synthesised by Holý et al. during research towards new anti-retroviral agents,[81] was prepared. In this case Cs_2CO_3 was used as the limiting reagent with cytosine (2 eq.) and **6** (2.2 eq.) in excess to ensure the selective formation of the anion at ^{1}N over the primary amine at position 4 on the cytosine ring. Tlc analysis of the reaction mixture showed the formation of several new products with very similar R_f values but, with careful repeated chromatography, the major product spot could be isolated in a yield of 66%. The ^{1}H NMR spectrum corresponded closely with the ^{1}H NMR data reported in the literature,[81] showing the AB doublet system indicative of H-6'and H-5' with peaks at δ_H 7.52 (d, J = 7.2 Hz, 1H, H-6') and δ_H 5.75 (d, J = 7.2 Hz, 1H, H-5'). A broad singlet at δ_H 6.49, integrating for 2H, showed the presence of the free amine on C-4', indicating this family of reactions to again be ^{1}N-chemoselective. The ^{13}C NMR spectrum again showed all the diagnostic resonances typical of the

diisopropyl PME chain. Also present were the carbon resonances representative of C-6' and C-5' at δ_C 146.8 and δ_C 92.4 respectively. The quaternary centre at C-4' was revealed by the presence of a very weak singlet δ_C 145.7 due to the fact that the weak quaternary C-4' signal is further weakened by the presence of two adjacent nitrogen atoms. The proton-decoupled ^{31}P NMR spectrum showed a single peak at δ_P 24.1 representing the single phosphorus atom in the phosphonate moiety.

5.3 Synthesis of 2-hydroxypyrazine ANP derivatives.

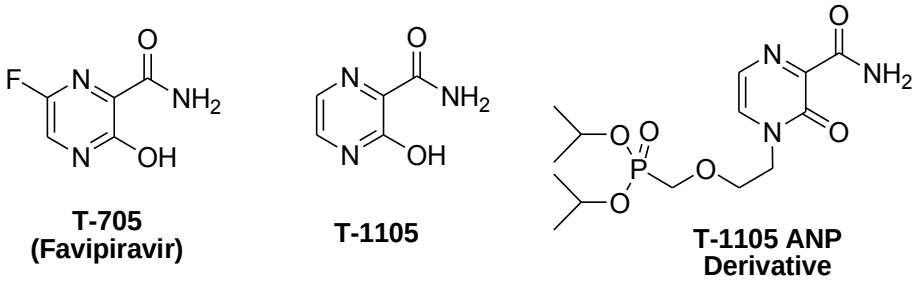


Fig. 5.5: The structures of T-705, T-1105 and the proposed T-1105 ANP Derivative.

As discussed in Part 1, Chapter 1, of this thesis, T-1105 and T-705 (favipiravir) have been shown to have potent broad spectrum anti-viral activity (Fig 5.5). In an effort to probe the synthesis of ANP derivatives of T-1105, a model synthesis of the parent pyrazine compound, 2-hydroxypyrazine **7**, was performed using the cheap and easily obtainable starting materials, glycinamide and glyoxal. The resulting 2-hydroxypyrazine ring was then alkylated onto the PME synthon **6** under the standard conditions. It was envisioned that a similar strategy would be used to synthesise 3-hydroxypyrazine-2-carboxamide (T-1105), using 2-aminomalonamide instead of glycinamide, which could be obtained from the ammonolysis of diethyl 2-aminomalonate.

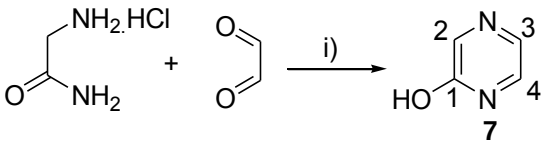


Fig. 5.6: Reagents and conditions: i) NaOH, MeOH, 39%

For **7**, the condensation reaction between glycinamide hydrochloride and glyoxal using sodium hydroxide as base was carried out in accordance with the procedure described by Jones (Fig 5.6).[85] According to Jones, condensation in the presence of a hydroxide base increases the reaction rate and yield and results in the formation of target hydroxypyrazine as a sodium salt.[85] The hydroxide helps drive the reaction by aiding aromatization of the pyrazine ring (Fig 5.7). The reaction yielded the required product **7** in a modest yield of 39%, sufficiently pure for use in the next step. The ^1H NMR spectrum showed the presence of 3 aromatic resonances at δ_H 8.07 (d, *J* = 1.4 Hz, 1H, H-2), δ_H 7.44 (d, *J* = 4.1 Hz, 1H, H-4) and δ_H 7.38 (dd, *J* = 4.1, 1.4 Hz, 1H, H-3), consistent with the pyrazine ring system. The resonance at δ_H 8.07 (H-2) couples to the doublet of

doublets at δ_H 7.38 via a long range aromatic coupling (J_{HH} = 1.4 Hz). The doublet of doublets also showed coupling to the doublet at δ_H 7.44 (d, *J* = 4.1 Hz, 1H, H-4), revealing H-3 and H-4 as an AB pair. The ^{13}C NMR spectrum showed three resonances at δ_C 149.6 (C-4), δ_C 128.6 (C-2) and δ_C 125.9 (C-3), as well as a weak peak at δ_C 158.8 indicative of the quaternary centre at C-1. IR spectroscopy showed the presence of a carbonyl absorbance at ν_{max} 1650 cm^{-1} indicating that **7** exists in two tautomeric forms (Fig 5.7: **7a** and **7b**).[86]

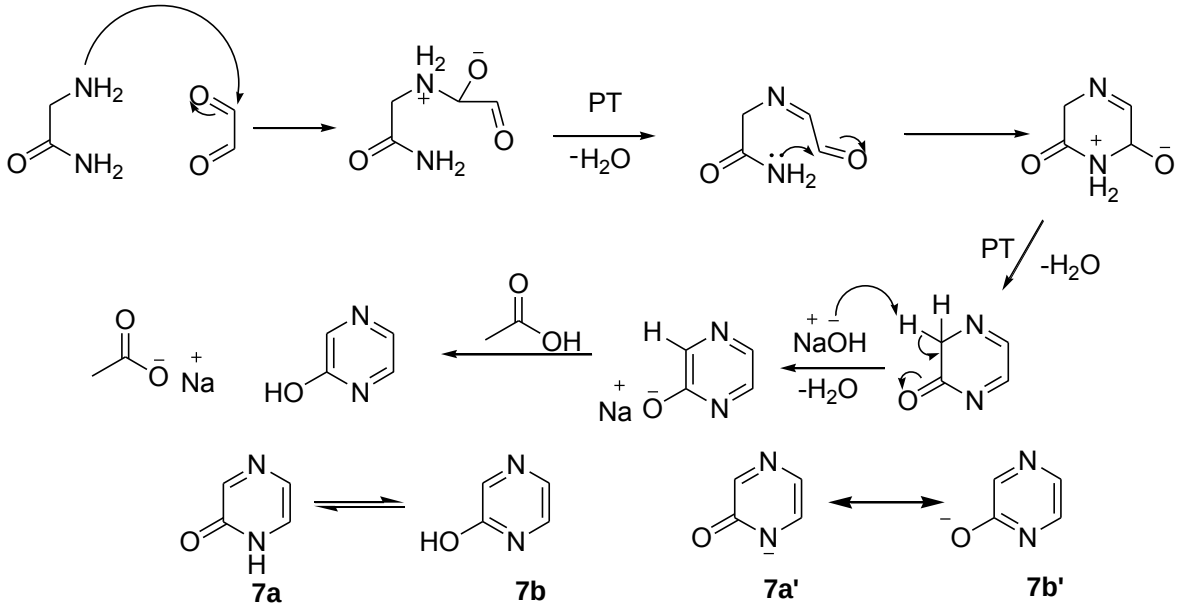


Fig. 5.7: The mechanism of the formation of **7**. The two tautomers of 2-hydroxypyrazine **7a** and **7b. 7a'** and **7b'** represent dominant resonance structures of the anion of **7.**

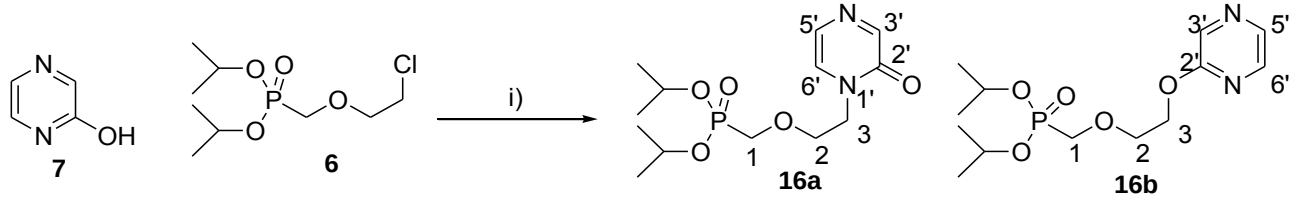


Fig. 5.8: Reagents and conditions: i) Cs_2CO_3, DMF, 100°C, **16a** 21%, **16b** 47%

Tlc analysis of the alkylation reaction of **7** with synthon **6**, again using Cs_2CO_3 in DMF (Fig 5.8), showed the formation of two major product spots. Careful chromatography using a DCM/MeOH solvent system (2: 98) allowed separation of the two isomers affording the *N*-alkylated **16a** (21% yield) and the *O*-alkylated **16b** (47% yield) which were assigned by NMR and IR spectroscopy (*vide infra*). Deprotonation of **7** results in the formation of a resonance stabilized anion shown as resonance structures **7a'** and **7b'** (Fig 5.7)**.** The O-anion, **7b',** is the most favoured resonance structure as it allows the bulk of the negative charge to rest on the more electronegative oxygen atom while also preserving the aromaticity of the pyrazine ring. This accounts for the formation of the O-alkylated **16b** as the major reaction product over **16a.**

The ^{1}H NMR spectra of both **16a** and **16b** showed the presence of the C<u>H</u>$(CH_3)_2$ multiplets at δ_H 4.68 and δ_H 4.70 respectively, the P-H coupled phosphonate doublets at δ_H 3.69 (J = 8.4 Hz) and δ_H 3.81 (J =8.2 Hz), as well as the diastereotopic methyl doublets between δ_H 1.29 and δ_H 1.26, confirming the presence of the diisopropyl phosphonate group. The distinction between **16a** and **16b** was made based on the ^{1}H and ^{13}C NMR data as well as the IR spectroscopic data for each compound. The infrared spectroscopic data of **16a** confirmed it as the N-alkylated product through the presence of a strong carbonyl stretch at ν_{max} 1662 cm^{-}1 as no such absorbance was seen in the IR data of **16b**. The diagnostic differences between the two ^{1}H NMR spectra of **16a** and **16b** lie in their aromatic regions, as well as in the relative chemical shifts of the triplets representing H-3. The H-3 triplet of the N-alkylated **16a** at δ_H 4.09 (t, J = 4.8 Hz, 2H) lies significantly upfield to that of the O-alkylated **16b,** which can be found at δ_H 4.52 (t, J = 4.8 Hz, 2H), due to the fact that oxygen is a more electronegative, and hence a more deshielding atom than nitrogen (Fig 5.9). This is also seen in the ^{13}C NMR spectra in which C-3 of **16a** resonates at δ_C 49.1 while C-3 of **16b** appears significantly downfield at δ_C 65.2. The second piece of supporting evidence for **16b** as the O-alkylated product could be found in the ^{1}H and ^{13}C NMR aromatic signals of the aromatic region. The signals found at δ_H 8.22 (d, J = 1.4 Hz, 1H, H-3'), δ_H 8.15 (d, J = 2.8 Hz, 1H, H-6') and δ_H 8.13 (dd, J = 2.8, 1.4 Hz, 1H, H-5'), and at δ_C 160.1 (C-2'), δ_C 140.5 (C-3'), δ_C 136.9 (C-6') and δ_C 136.2 (C-5'), correspond very closely to both the ^{1}H and ^{13}C NMR data available for the aromatic signals of the analogous, commercially available compound, 2-methoxypyrazine.[86,87]

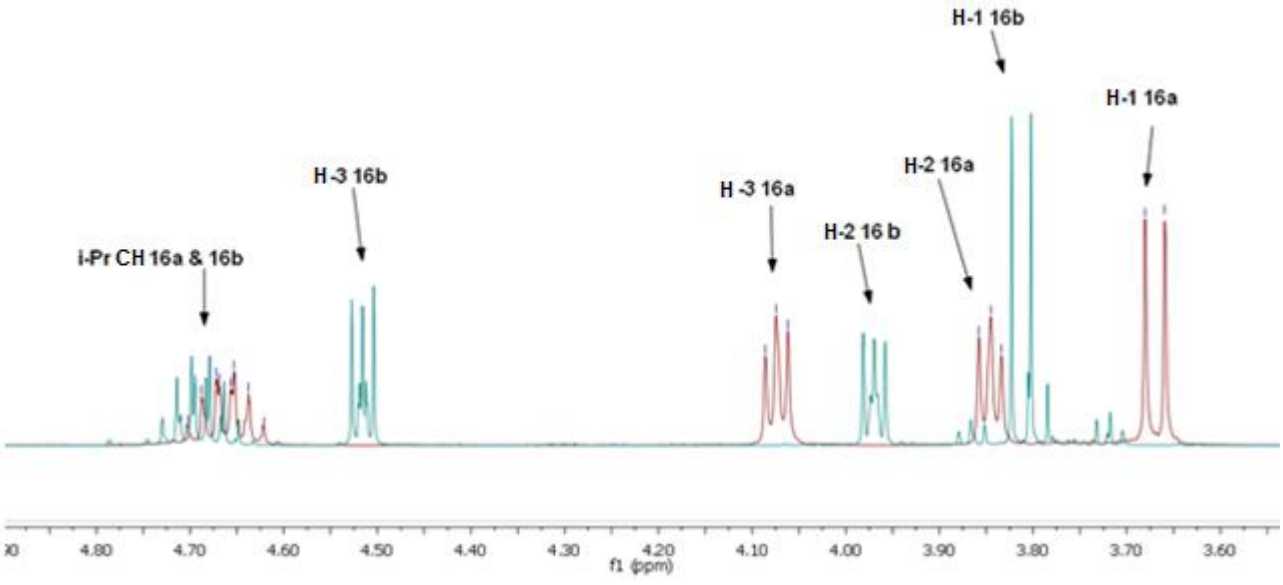


Fig. 5.9: The expanded and superimposed NMR spectra of **16a** and **16b**

The HSQC NMR of both **16a** and **16b** confirmed the ^{1}H and ^{13}C NMR assignments made for each compound (Fig 5.10). The HSQC spectrum of **16a** showed the C<u>H</u>$(CH_3)_2$ multiplet at δ_H 4.68 coupling to the P-C J^2 coupled doublet at δ_C 71.3 (d, J_{PC} = 6.7 Hz, <u>C</u>H$(CH_3)_2$). This means that the resonance at δ_C 70.3 (d, J_{PC}= 11.0

Hz) must be assigned to C-2, an unexpected result as one would usually expect to assign the resonance with the larger coupling (J_{PC}) constant to the carbon as the closest in proximity to the phosphorus atom.

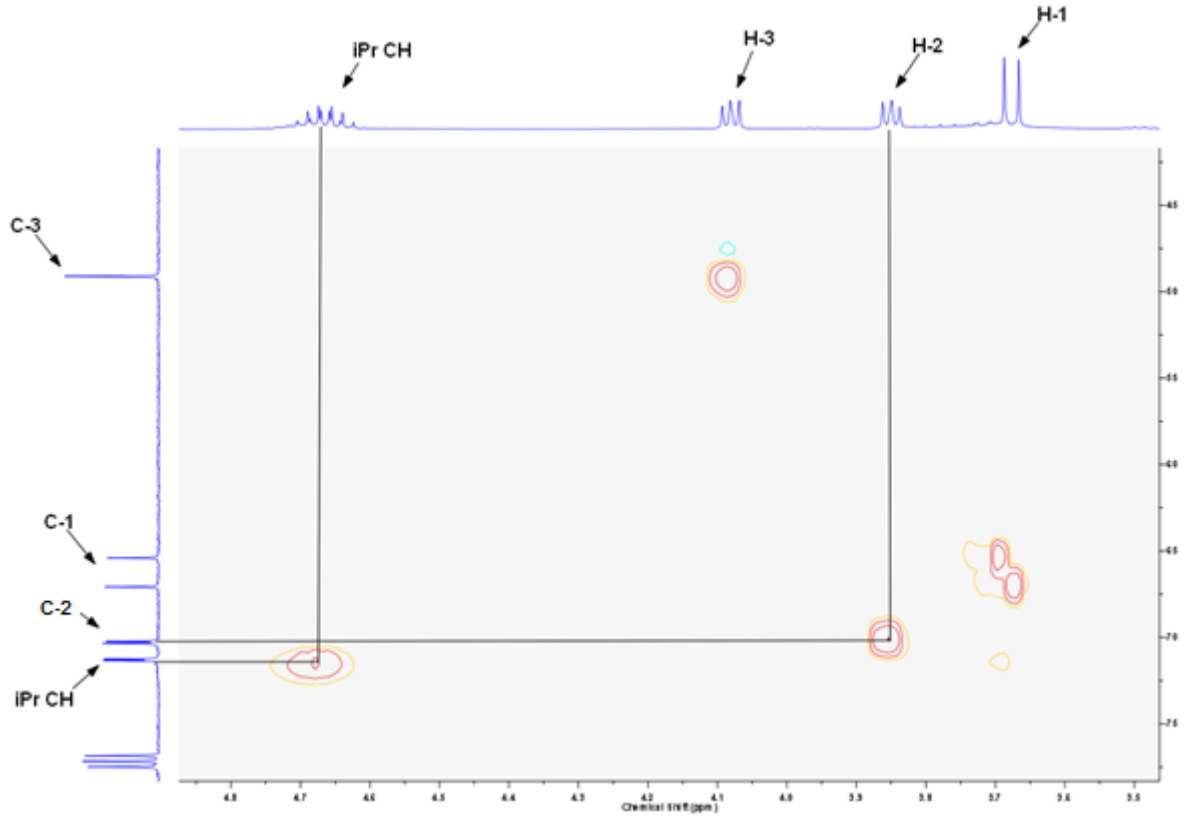


Fig. 5.10: The expanded HSQC spectrum of **16a.**

5.4 The synthesis of pyrimidine-based ANP derivatives.

The next series of derivatives which were synthesised were based on the open ring diaminopyrimidine (DAPy) ANP variants discussed in Part 1, Chapter 2.4.2. The first derivative, which was formed from the alkylation of 2,4-diaminopyrimidine with synthon **6,** resembles an acyclic analogue of the naturally occurring heterocyclic base adenine (Fig 5.11).

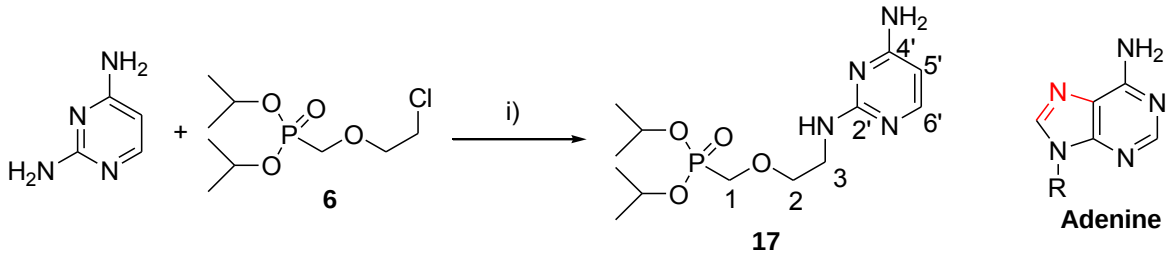


Fig 5.11. Reagents and conditions: i) Cs_2CO_3, DMF, 100°C, 40%.

The reaction was performed and worked up according to the same general procedure as described above, with an excess of 2,4-diaminopyrimidine used to ensure the formation of the mono-alkylated product. Although the formations of different regioproducts were possible through the alkylation of either of the

two amino groups, only one product could be distinguished from the tlc or spectroscopic data. However, the exact structure of the resulting product as either the 2'-amino or 4'-amino variant could not be assigned from the available data. Isolation of the major product of the crude reaction mixture via column chromatography with MeOH/DCM (6:94) afforded **17** in a yield of 40%. The ^{1}H NMR spectrum of **17** showed all the expected diisopropyl phosphonate resonances at δ_H 4.69 (m, 2H, C<u>H</u>$(CH_3)_2$), δ_H 3.77 (d, J_{PH} = 8.0 Hz, 2H, H-1), δ_H 1.29 (d, *J* = 6.0 Hz, 6H, CH(C<u>H</u>$_3)_2$) and δ_H 1.27 (d, *J* = 6.0 Hz, 6H, CH(C<u>H</u>$_3)_2$) as well as an aromatic AB coupled pair of doublets at δ_H 7.66 (d, *J* = 5.9 Hz, 1H, H-6') and δ_H 5.84 (d, *J* = 5.9 Hz, 1H, H-5'). A broad singlet at δ_H 6.40, integrating for 1H showed the presence of the alkylated amine group, while the broad singlet at δ_H 5.56 integrating for 2H, represented the un-alkylated amine, confirming the structure as the mono-alkylated product. The resonances indicative of H-2 and H-3 appeared upfield of the phosphonate doublet at δ_H 3.77 (d, J_{PH} = 8.0 Hz, 2H, H-1). The resonance indicative of H-2 appeared at δ_H 3.74 (t, *J* = 5.5 Hz, 2H, H-2), while the signal for H-3 appeared as a quartet at δ_H 3.52 (q, *J* = 5.5 Hz, 2H,), showing H-H coupling with H-2 and the hydrogen on the adjacent nitrogen. The ^{13}C NMR spectrum of **17** showed all the expected signals including the phosphonate doublet at δ_C 66.4 (d, J_{PC} = 166.6 Hz, C-1) and the diastereotopic methyl doublets at δ_C 24.4 (d, J_{PC} = 4.0 Hz, CH(<u>C</u>$H_3)_2$) and δ_C 24.3 (d, J_{PC} = 4.0 Hz, CH(<u>C</u>$H_3)_2$). Further evidence for the structure of **17** was given by mass spectroscopy, HRMS (ES): 333.1697 (M^+ + H), which corelated with the molecular formula $C_{13}H_{26}N_4O_4P$ (expected value (M^+ + H), 333.1686)

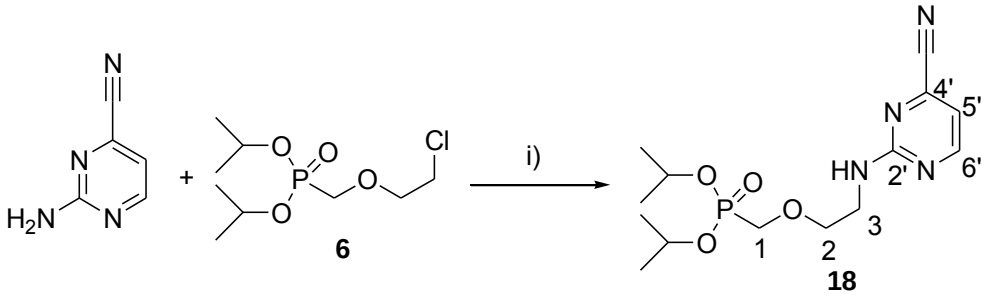


Fig. 5.12: Reagents and conditions: i) Cs_2CO_3, DMF, 100°C, 64%.

The next compound, **18**, was formed from the alkylation of 2-amino-4-cyano-pyrimidine and synthon **6** under the standard reaction conditions (Fig 5.12). Product **18** was found to be much less polar than the structurally similar **17** due to the presence of a cyano group instead of an amino group at position 4' on the pyrimidine ring. For this reason the compound could be isolated relatively easily by column chromatography using 100% ethyl acetate as the solvent system. The ^{13}C NMR spectrum of **18** showed a resonance at δ_C 117.1 typical of a nitrile carbon. The IR spectroscopy confirmed the presence of the cyano group with a C≡N absorbance at ν_{max} 2246 cm^{-1}. The ^{1}H NMR spectrum showed an AB pair for H-6'and H-5' at δ_H 8.41 (d, *J* = 5.0 Hz, 1H, H-6') and δ_H 6.80 (d, *J* = 5.0 Hz, 1H, H-5'). The N-H signal could be observed at δ_H 6.07 (s, 1H, N<u>H</u>) along with the quartet at δ_H 3.63 (q, *J* = 5.3 Hz, 2H, H-3) for H-3 which again coupled to both the N-H and H-2. HSQC NMR of **18** again showed the carbon resonance of C-2 (δ_C 72.8 (d, J_{PC} = 12.1 Hz,

C-2)) to have a larger P-C coupling constant than the $\underline{C}H(CH_3)_2$ signal at δ_C 73.2 (d, J_{PC} = 6.7 Hz, $\underline{C}H(CH_3)_2$), (Fig 5.13).

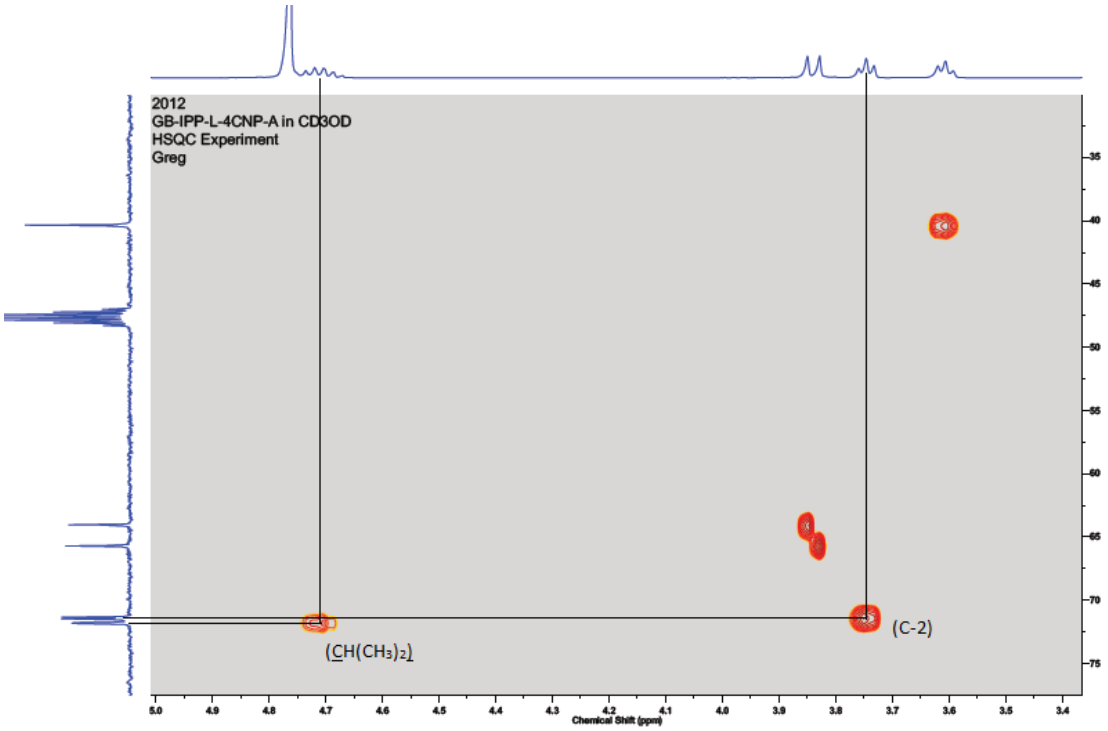


Fig. 5.13: The expanded HSQC NMR of **18.**

5.5 The synthesis of ribavirin-like ANP derivatives.

The next family of derivatives to be synthesised were all analogues of the well-known anti-viral nucleoside analogue, ribavirin. The derivatives described all feature a 5-membered heterocycle containing 2 or 3 nitrogen atoms in various arrangements. An amide group, as discussed above in Chapter 1, is thought to be vital to the drug's ability to form multiple hydrogen-bonding profiles, a trait believed to be a possible source of ribavirin's broad spectrum anti-viral activity. The first derivative was designed to exclude the amide in order to test this hypothesis, which involved the reaction of 1,2,4-triazole with **6** under standard conditions (Fig 5.14).

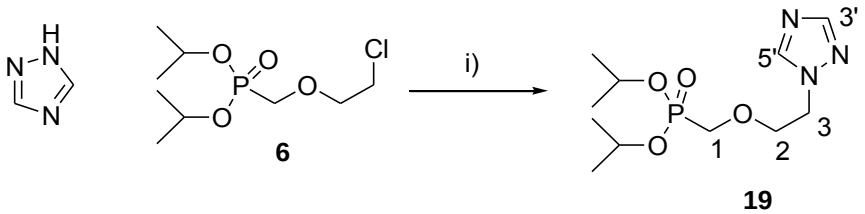


Fig. 5.14: Reagents and conditions: i) Cs_2CO_3, DMF, 100°C, 58%.

Although the formation of three alkylation products, one for alkylation of each triazole nitrogen were theoretically possible, the substitution reaction afforded only one major observable product in a 58% yield, as revealed by tlc as well as ^{1}H and ^{13}C NMR spectroscopy. The exact structure of the major regio-product,

19, as either the 1,2 or 4-alkyl variant could not be determined from the available data. Surprisingly, **19,** was found to be non-UV active, thus the tlc plate could only be visualised with various stain reagents, the most successful being anisaldehyde which showed the product as a strong white spot. The ^{1}H NMR spectrum of **19** featured two aromatic singlets at δ_H 8.36 (s, 1H, H-5') and δ_H 7.84 (s, 1H, H-3'), each integrating for 1H, confirming the presence of the triazole ring within the molecule. The ^{13}C NMR spectrum showed the corresponding triazole carbon signals at δ_C 151.9 (C-5') and δ_C 144.1 (C-3'). The other signals typical of the diisopropyl PME chain were also present in both the ^{1}H and ^{13}C NMR spectra. A relatively deshielded triplet for H-3 at δ_H 4.43 (t, J = 6.8 Hz, 2H, H-3) was noted, demonstrating the electron-withdrawing nature of the triazole ring. HSQC NMR of **19** was again used to confirm the ^{1}H and ^{13}C NMR assignments, specifically the carbon assignments of the P-C coupled doublets at δ_H 71.3 (d, J_{PC} = 6.8 Hz, $\underline{C}H(CH_3)_2$) and δ_H 70.8 (d, J_{PC} = 10.0 Hz, C-2).

The next set of derivatives were ANP derivatives of ribavirin, obtained from the reaction of synthon **6** and 1*H*-1,2,4-triazole-3-carboxamide. 1*H*-1,2,4-triazole-3-carboxamide (**8**) was not commercially available and had to be synthesised from 1*H*-1,2,4-triazole-3-carboxylic acid (Fig 5.15).

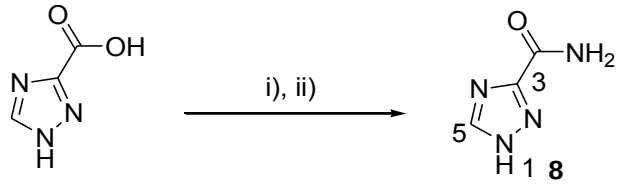


Fig. 5.15. Reagents and conditions: i) $SOCl_2$, DMF, Reflux, N_2; ii) NH_3, 50°C

A patent was found which outlined a procedure for its synthesis from 1*H*-1,2,4-triazole-3-carboxylic acid in neat thionyl chloride with a catalytic quantity of DMF under inert atmospheric conditions.[88] This was carried out resulting in the formation of a reactive but stable and isolatable dimeric ketene of 1,2,4-triazole-3-carboxylic acid (Fig 5.16).[88] After 3hrs, once the intermediate dimer had formed, the thionyl chloride and DMF were removed under high vacuum to yield the intermediate as a dry white powder. Saturated ammonia solution was then added to the intermediate at 0°C, after which the stirring suspension was warmed to 50°C to drive the ammonolysis of the intermediate dimer to afford **8** (Fig 5.16)**.** The ammonia solution was then evaporated under reduced pressure to afford **8,** which could be recrystallised from a solution of water and methanol (1:10). IR spectroscopy of the dried product showed absorptions at ν_{max} 3327 cm^{-1}, 3155 cm^{-1} and 1699 cm^{-1} representing the two amide N-H bands and C=O stretch respectively. The ^{1}H NMR spectrum in DMSO-d_6 showed one aromatic signal at δ_H 8.38 (s, 1H, H-5) as well as two broad singlets at δ_H 7.61 (s, 1H, $N\underline{H}_2$) and δ_H 7.85 (s, 1H, $N\underline{H}_2$) representing the amide hydrogens. The structure was further confirmed with mass spectroscopy, showing HRMS (ES): 113.0448 (M^+ + H), which corresponded to the molecular ion $C_3H_5N_4O$ (M^+ + H, 113.0463).

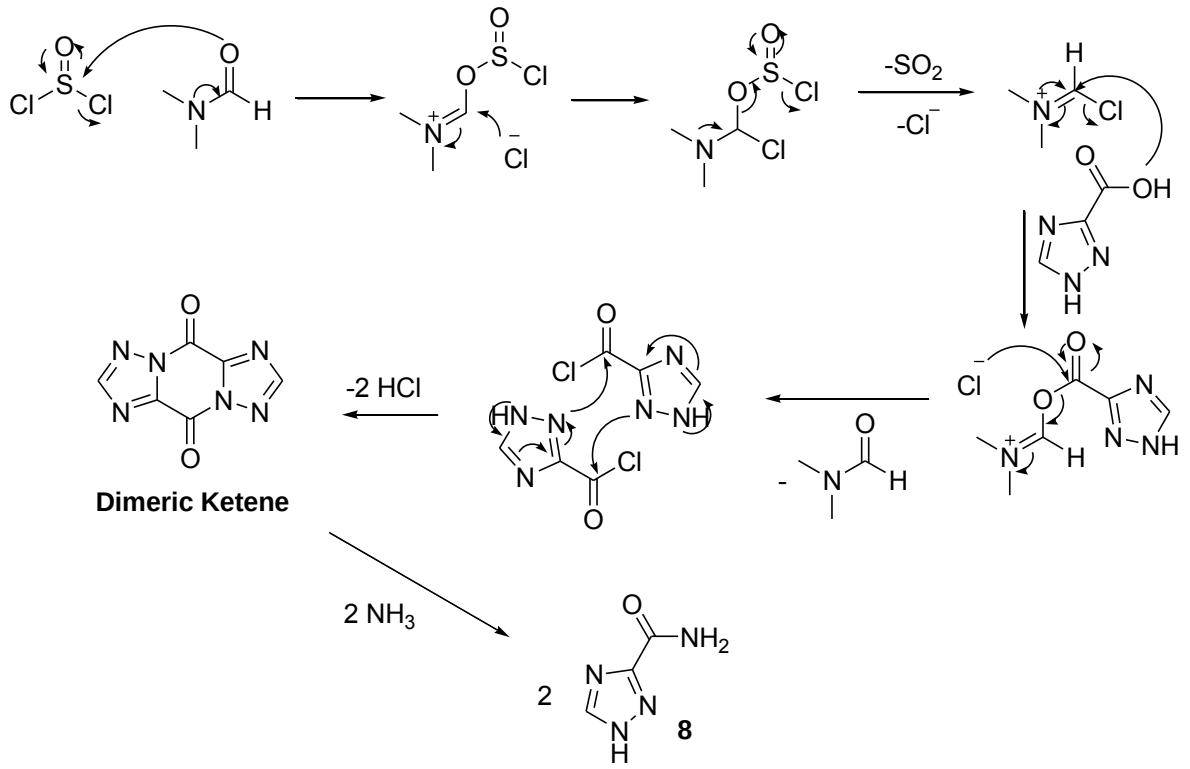


Fig. 5.16: The reaction mechanism for the formation of **8.**

With **8** in hand, the next task was its alkylation with PME synthon **6,** which was again achieved using Cs_2CO_3 in DMF at 100°C (Fig 5.17). Three possible products (**20a**, **20b** and **20c**) were expected from the reaction as each triazole nitrogen could act as a nucleophile and point of alkylation. However, tlc analysis of the reaction mixture after 12 hours of stirring at reflux showed the formation only of 2 distinguishable product spots, both with very similar R_f values. Repeated and careful chromatography was, however, able to separate the two products as **20a** and **20b** assigned by extensive NMR analysis (vide infra).

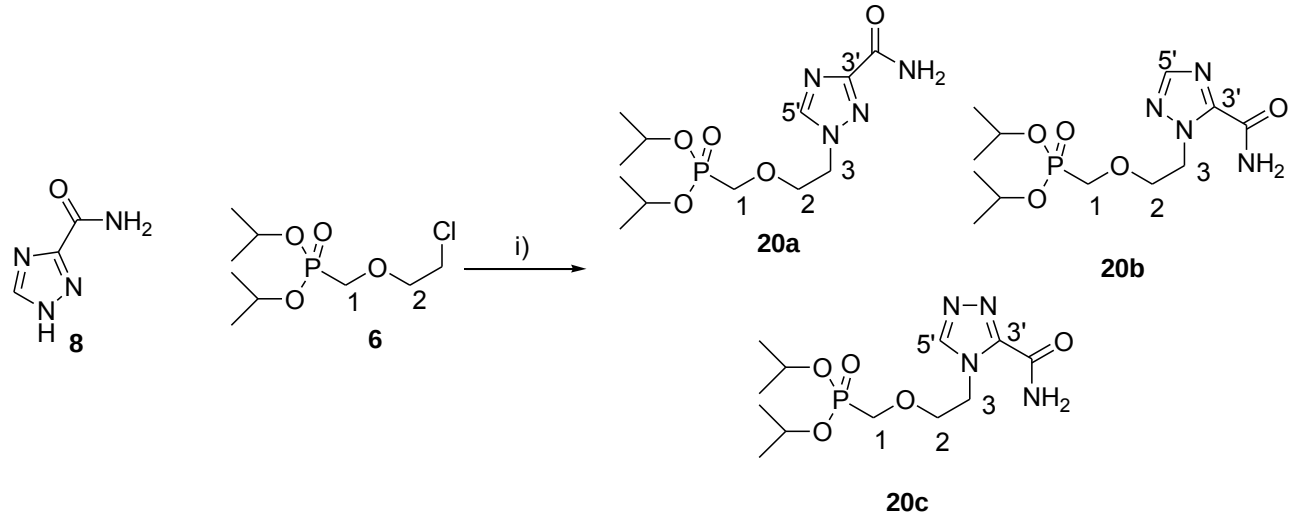


Fig. 5.17. Reagents and conditions: i) Cs_2CO_3, DMF, 100°C, **20a:** 24%, **20b:** 45%.

The IR spectra, recorded in DCM in a sodium chloride solution cell, revealed the presence of the primary amide in both cases with the presence of strong bands at ν_{max} 3516 (N-H) cm^{-1}, 3398 (N-H), cm^{-1} and 1704 (C=O) cm^{-1} for **20a**; and ν_{max} 3512 (N-H) cm^{-1}, 3393 (N-H), cm^{-1} and 1704 (C=O) cm^{-1} for **20b,** confirming the successful alkylation of **8** and **6**. The ^{1}H NMR spectra of **20a** and **20b** revealed the only real spectroscopic differences between the 2 compounds to be the difference in chemical shift of the H-5' and H-3 hydrogen signals, suggesting their existence as regioisomers of each other. The H-5' hydrogen singlet in **20a** appeared at δ_H 8.51 while in **20b** it appeared at δ_H 7.92. The H-3 triplets of **20a** and **20b** showed the most obvious difference in chemical shift which, at δ_H 4.52 (t, *J* = 4.0 Hz, 2H, H-3) and δ_H 4.90 (t, *J* = 5.5 Hz, 2H, H-3) respectively, and suggested the points of attachments on the triazole ring to have markedly different electronic environments in each case (Fig 5.18). HSQC NMR was used to confirm the ^{13}C NMR assignment of the H-5' resonance at δ_C 146.2 in **20a** and δ_C 150.4 in **20b**, as well as the ^{1}H and ^{13}C NMR assignments associated with PME chain, including the assignments of the ambiguous P-C coupled doublets of C-2 and the *i*-Pr $\underline{C}H(CH_3)_2$.

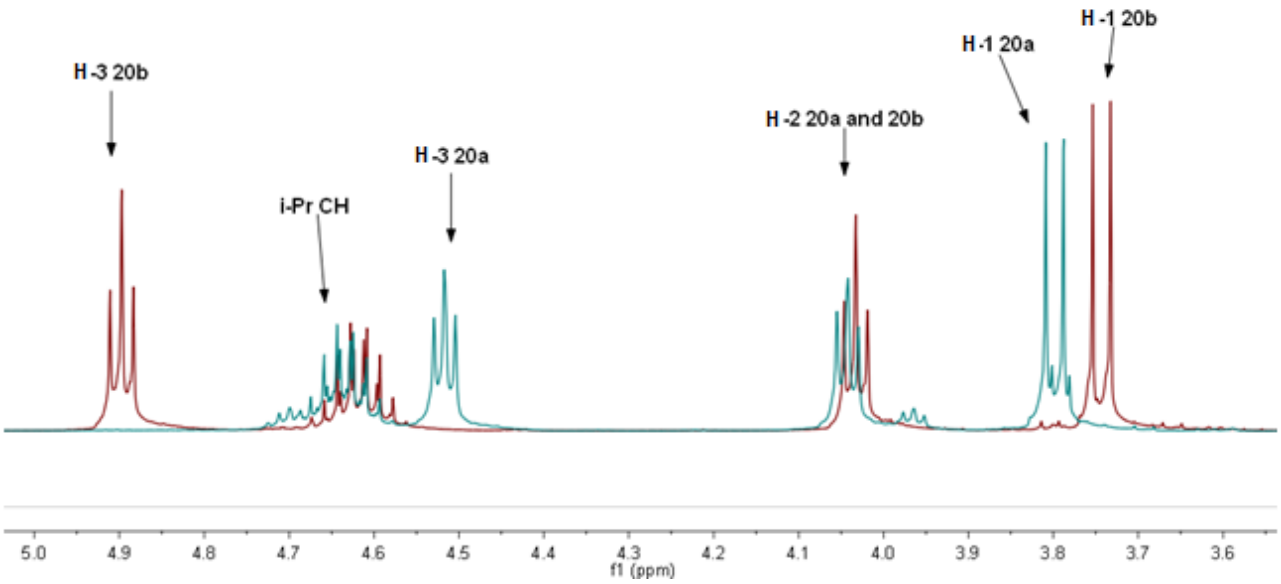


Fig. 5.18: The expanded and superimposed NMR spectra of **20a** and **20b**

The ^{1}H, ^{13}C and HSQC NMR data was, however, inconclusive regarding assignment of regiochemistry as either **20a** or **20b**. HMBC (Heteronuclear Multiple Bond Correlation) NMR spectroscopy was used to solve this problem as this NMR technique shows only J^2 and J^3 C-H cross-coupling. The HMBC spectrum of **20a** showed the diagnostic J^3 C-H cross coupling of the H-3 triplet at δ_H 4.52 (t, *J* = 4.0 Hz, 2H, H-3) with the C-5' carbon resonance at δ_C 146.2 (Fig 5.19), a cross-coupling relationship which is not possible in **20b.** The HMBC NMR spectrum of **20b** showed no such relationship, as was expected.

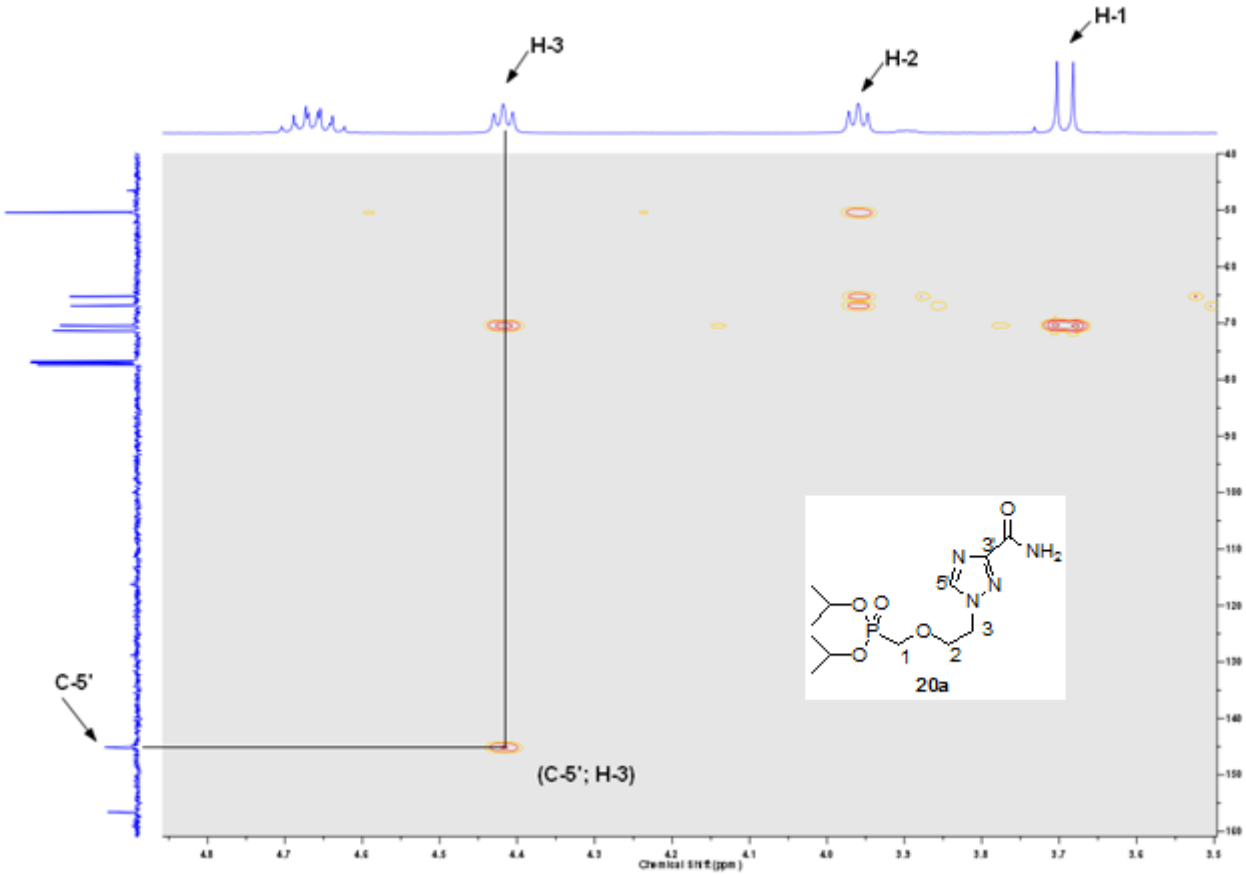


Fig. 5.19: The expanded HMBC of **20a** showing the cross-coupling of H-3 with C-5'.

The third isomer, **20c,** was not detected as a major product by tlc and was only revealed later as an inseparable minor impurity in the ^{1}H and ^{13}C NMR spectra of **20a** (Fig 5.20)**.** Unfortunately there is no way to distinguish **20a** from **20c** structurally from the NMR data, even if they could be separated, as the quaternary carbon resonance at C-3' is too weak to be seen in the HMBC NMR spectra. The exact structure of this product as either **20a** or **20c** remains conjecture.

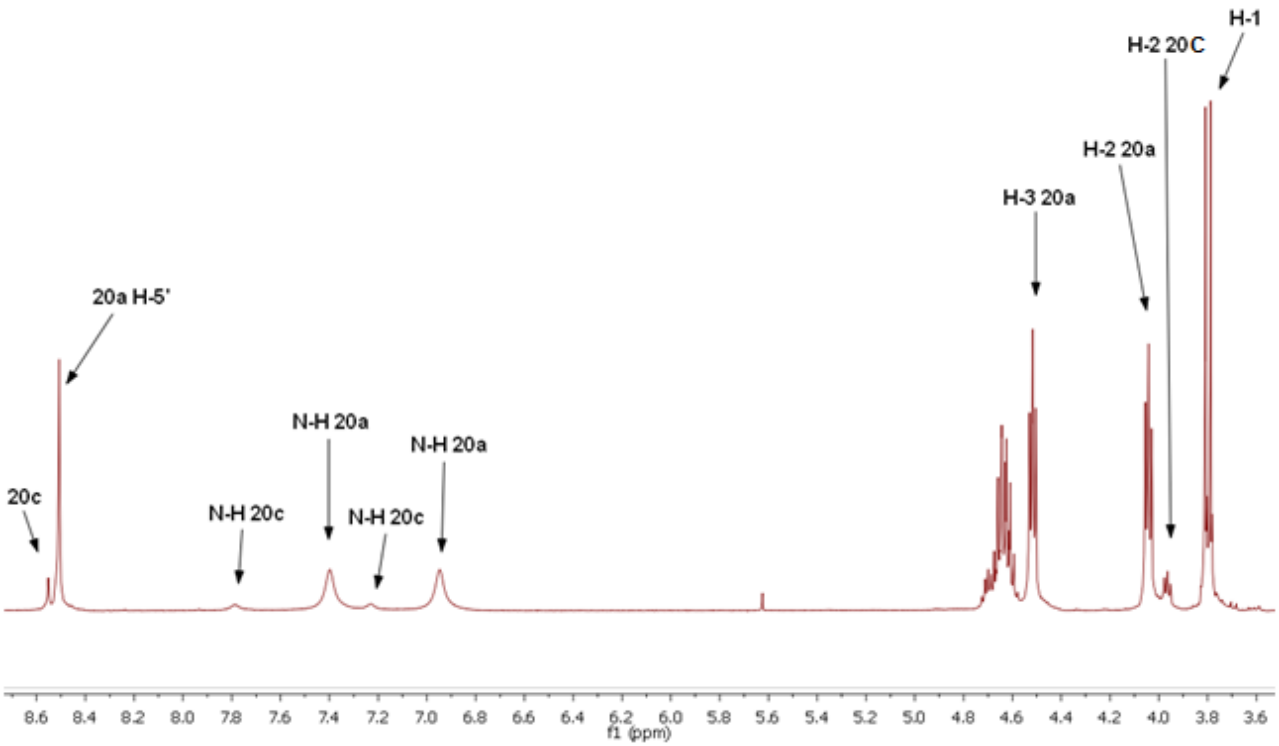


Fig. 5.20: The ^{1}H NMR of **20a,** showing the peaks of "**20c**" as an un-isolatable minor product.

The next derivative featured a 1*H*-pyrazole-3-carboxamide base in place of ribavirin's 1*H*-triazole-3-carboxamide. Once again **9** was unavailable commercially and had to be made from 1*H*-pyrazole-3-carboxylic acid (Fig 5.21).

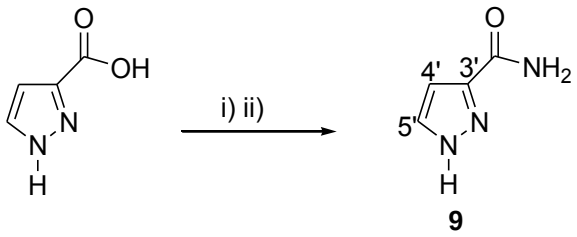


Fig. 5.21: Reagents and conditions: i) MeOH, H_2SO_4, reflux, ii) MeOH/NH_3, 40°C

The same procedure which was used to produce **8**, was unsuccessful with 1*H*-pyrazole-3-carboxylic acid. Thus, **9** had to be synthesised in 2 steps, with the methyl ester as an intermediate. Refluxing 1*H*-pyrazole-3-carboxylic acid in methanol with 1.2 equivalents of H_2SO_4 quantitatively produced the methyl ester. 1.2 Equivalents (as opposed to a catalytic quantity) of acid had to be used, due to the basic character of the pyrazole ring. After isolating the methyl ester by extraction with ethyl acetate from a saturated sodium carbonate solution, the methyl ester was warmed in a solution of methanol and ammonia to afford **9** in an overall yield of 86%. The ^{1}H NMR spectrum of **9** in DMSO showed the presence of an AB doublet pair at δ_H 7.71 (d, *J* = 2.0 Hz, 1H, H-5') and δ_H 6.67 (d, *J* = 2.0 Hz, 1H, H-4') as well as two broad singlets at δ_H 7.54 (s, 1H, N-H) and δ_H 7.19 (s, 1H, N-H), indicative of the amide hydrogens. A minor impurity was also observed and thought to be a possible tautomer of **9**, but the product was deemed to be sufficiently pure for use in

the next step. IR spectroscopy also showed the presence of the new primary amide with absorptions at ν_{max} 3405 (N-H), 3266 (N-H), 1698 (C=O) cm^{-1}, while mass spectroscopy confirmed the structure with HRMS (ES): 112.0494 (M^+ + H), correlating closely to expected value for the molecular formula $C_4H_6N_3O$ (M^+ + H, 112.0511)

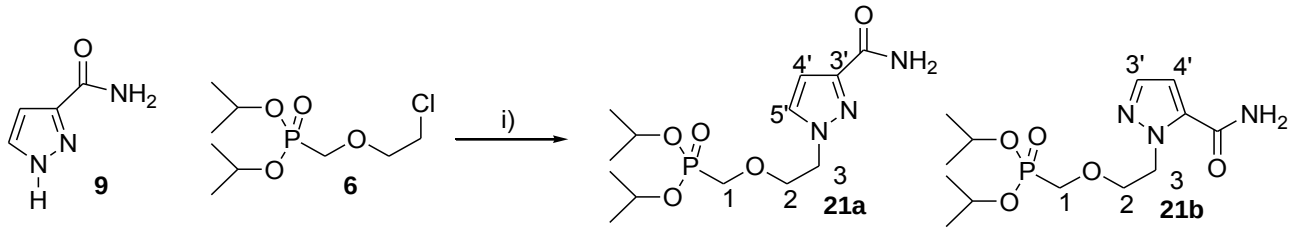


Fig. 5.22: Reagents and conditions: i) Cs_2CO_3, DMF, 100°C, **21a:** 42%, **21b:** trace

With **9** in hand, the reaction with **6** under the usual conditions was performed to produce **21a** in a modest yield of 42% (Fig 5.22). Although formation of both **21a** and its regioisomer **21b** were expected, the tlc profile of the reaction showed only one spot, even when developed across a range of different solvent systems. The ^{1}H NMR spectrum of **21a** showed the presence of an inseparable impurity, which given the nature and chemical shifts of the impurity peaks (Fig 5.23), suggested the formation of the regioisomer, **21b**, in trace quantities. The structural distinction between **21a** and **21b** was made based on analogy with compounds **20a** and **20b,** specifically by the relative position of the H-3 triplet in each case. The N^1-alkylated product, **20a,** showed a triplet at δ_H 4.52 (t, *J* = 4.0 Hz, 2H, H-3), relatively upfield of the corresponding peak in the spectrum of its regioisomer **20b** which appeared at δ_H 4.90 (t, *J* = 5.5 Hz, 2H, H-3)**.** This relationship was applied by analogy to assign the H-3 resonances in the spectrum of **21a** and **21b** (Fig 5.23), thus the major H-3 resonance in the spectrum was assigned to **21a.**

The other dominant peaks in the ^{1}H NMR spectrum showed the amide hydrogens at δ_H 6.79 (s, 1H, NH_2) and δ_H 5.79 (s, 1H, NH_2), as well as the aromatic AB pair of H-4'coupling to H-5' at δ_H 7.50 (d, *J* = 2.4 Hz, 1H, H-5') and δ_H 6.76 (d, *J* = 2.4 Hz, 1H, H-4'). All the expected resonances indicative of the PME backbone were also present in the ^{1}H NMR spectrum. ^{13}C NMR spectroscopy again revealed the presence of minor impurity peaks throughout the spectrum, with both their splitting patterns and chemical shifts shadowing the dominant product peaks assigned to **21a**. The ^{13}C NMR spectrum, like the ^{1}H NMR spectrum, showed all the relevant PME backbone peaks, as well as the aromatic carbon resonances at δ_C 146.4 (C-3'), δ_C 132.1 (C-5') and δ_C 107.1 (C-4') and the amide carbonyl resonance at δ_C 164.1 (C=O). The ^{1}H decoupled ^{31}P NMR spectrum shows a similar pattern with both a major and minor phosphorus peak.

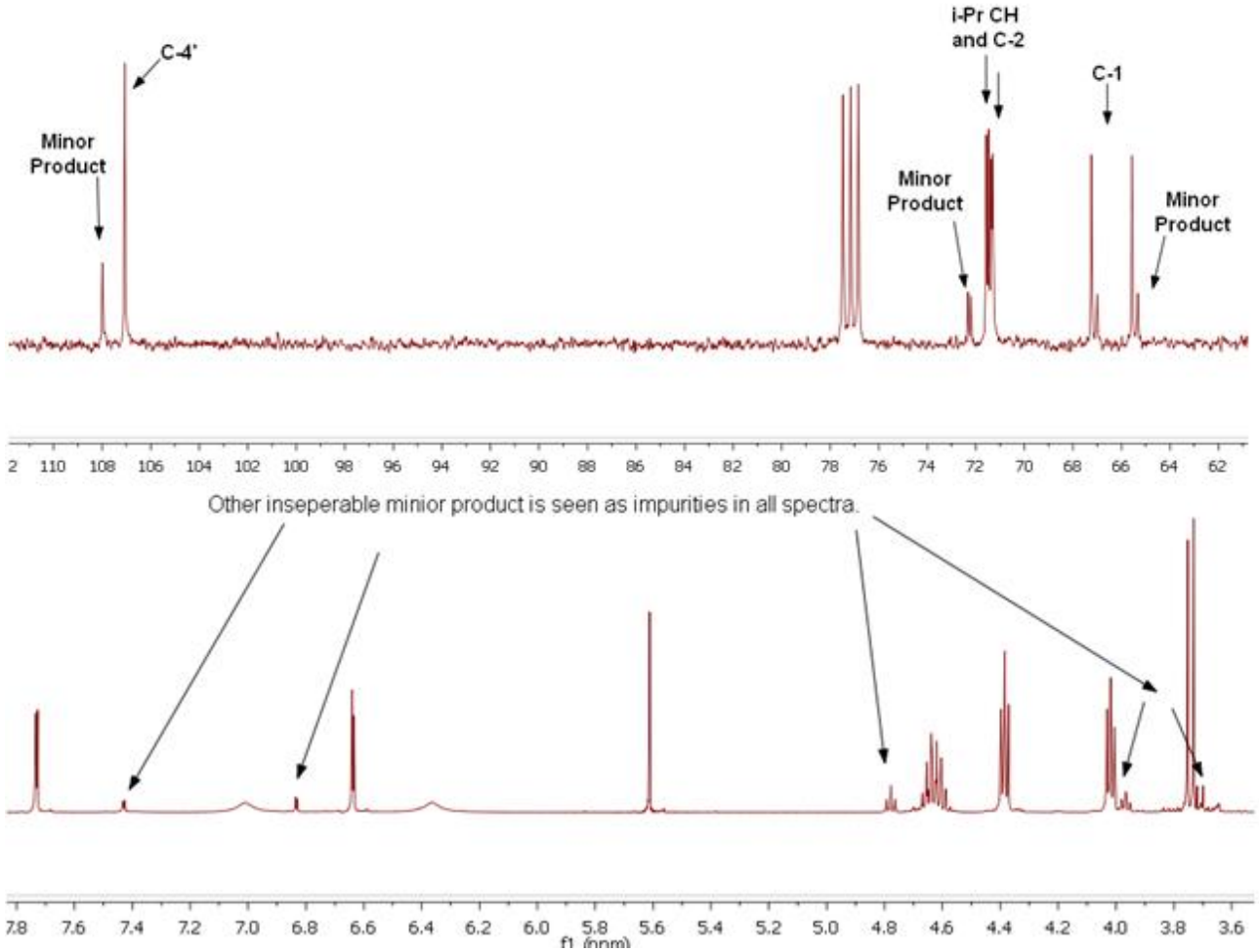


Fig. 5.23: The ^{13}C and ^{1}H NMR spectra of **21a,** showing the proposed minor product **21b** as an impurity.

The final analogue in this series featured an amide-bearing imidazole ring, which, as in the case of the pyrazole derivative discussed above, had to be synthesised from the relevant carboxylic acid. This was again achieved thorough the formation and subsequent ammonolysis of the methyl ester of 1*H*-imidazole-4-carboxylic acid. The same procedure used to form pyrazole **9** was used in the formation of the imidazole compound **10** (Fig 5.24).

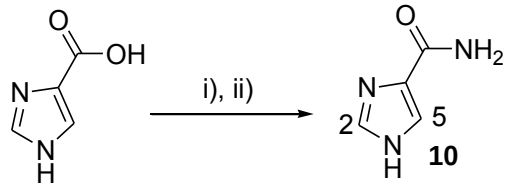


Fig. 5.24: Reagents and conditions: i) MeOH, H_2SO_4, reflux, ii) MeOH/NH_3, 40°C.

Following ammonolysis, the ^{1}H NMR spectrum of **10** revealed H-2 and H-5 as singlets at δ_H 7.70 (s, 1H, H-2) and δ_H 7.61 (s, 1H, H-5) respectively, while the amide hydrogens resolved as two broad singlets at δ_H 7.39 and δ_H 7.07, each integrating for 1H. The ^{13}C NMR spectrum showed four resonances at δ_C 164.3 (C=O), δ_C 137.9 (C-2), δ_C 136.4 (C-5), δ_C 121.6 (C-4), with the weak, downfield resonace at δ_C 164.3 indicative of the carbonyl carbon. With **10** in hand, an alkylation with the PME synthon **6**, was performed under the usual

reaction conditions (Fig 5.25) and despite the presence of two nucleophilic triazole nitrogens (hence two possible alkylation products),the formation of only one major reaction product was observed (by both tlc and spectroscopy) and assigned as **22**.

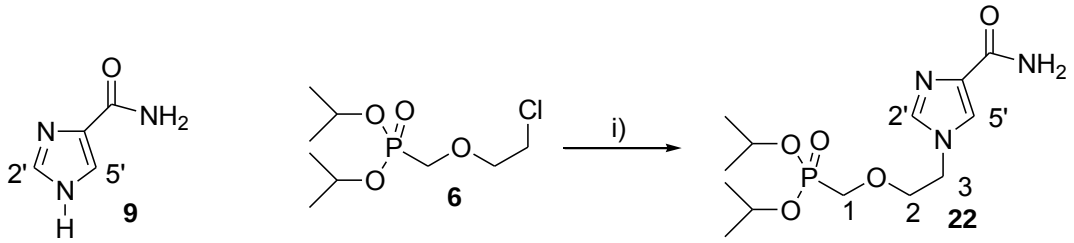


Fig. 5.25: Reagents and conditions: i) Cs_2CO_3, DMF, 100°C, 42%

The ^{1}H NMR spectrum of **22** showed all the standard PME resonances such as the diastereotopic methyl doublets at δ_H 1.27 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3$)$_2$) and δ_H 1.24 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3$)$_2$), the isopropyl CH multiplet at δ_H 4.65 (m, 2H, C<u>H</u>(CH$_3$)$_2$) and the strong P-H doublet at δ_H 3.79 (d, J_{PH} = 8.3 Hz, 2H, C-1). The more deshielded H-3 triplet appears at δ_H 4.30 (t, *J* = 5.2 Hz, 2H, C-3). The imidazole ring showed the two expected diagnostic singlets at δ_H 7.69 (s, 1H, H-5') and δ_H 7.64 (s, 1H, H-2'), as well as the amide hydrogens as broad singlets at δ_H 7.14 (s, 1H) and δ_H 6.56 (s, 1H). The ^{13}C NMR showed all the expected resonances including the three imidazole ring signals at δ_C 138.4 (C-5'), δ_C 138.0 (C-4') and δ_C 123.4 (C-2'), as well as the carbonyl carbon resonance δ_C 165.1. All the ^{1}H and ^{13}C NMR spectral assignments were confirmed by HSQC NMR spectroscopy.

Chapter 6

Synthesis of an arylethynyltriazole ANP derivative via a Sonogashira cross-coupling reaction.

6.1 Introduction

Arylethynyltriazole acyclic nucleoside derivatives were discussed in Chapter 2 as a new class of anti-viral agent, with the 4-fluorophenylethynyl triazole-3-carboxamide acyclic nucleoside derivative, **A** (Fig 6.1), shown to possess a high level of activity against HCV and other RNA viruses.[72] Compound **A,** which was conceptualised from the cyclic nucleoside **E,**[70] features an acyclic nucleoside chain, suggesting that its mechanism of action is in some way related to other acyclic nucleoside analogues. As of yet acyclic nucleotide analogues of these biaryl type derivatives have not been explored. This chapter will discuss the synthesis and analysis of the ANP derivative **24'** through a Sonogashira palladium (0) catalysed cross-coupling reaction (Fig 6.1).

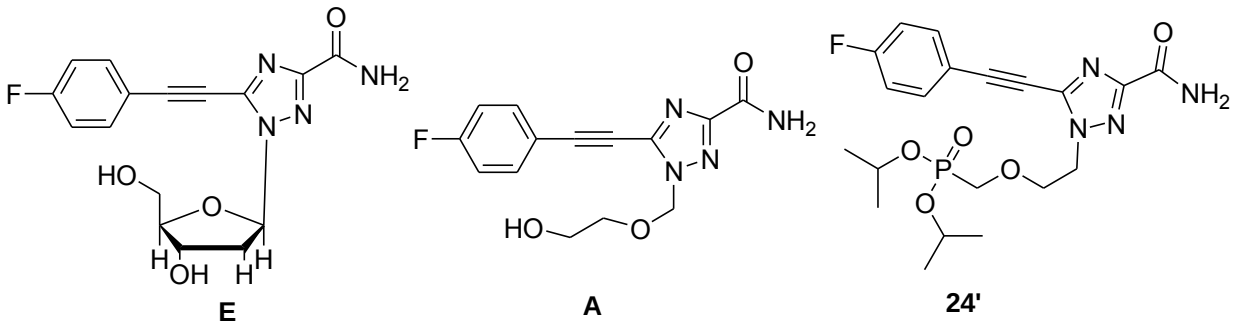


Fig. 6.1: The structures of the cyclic nucleoside analogue **E**, the acyclic nucleoside analogue **A**, and the target ANP arylethynyltriazole derivative **24'**.

6.2 Overview of the Sonogashira reaction.

The Sonogashira cross-coupling reaction was developed by Kenkichi Sonogashira in 1975, and has been widely applied within the field of nucleoside and nucleotide synthesis.[89] The reaction, along with other palladium catalysed reactions, has over the years become of great importance to medicinal chemists as a means of generating complex structures through the formation of new C-C bonds under relatively mild reaction conditions. This reaction had been previously used in our research group as a key coupling reaction in the synthesis of bifunctional HIV RT inhibitors,[90] and this work served as a starting point for the exploration of Sonogashira cross-coupling reactions in the context of arylethynyltriazole-based ANPs.

The mechanism of the Sonogashira reaction involves firstly the oxidative insertion of $Pd(0)(PPh_3)_2$ (a $14e^-$ species) into an appropriate aryl-leaving group bond to form an aryl-Pd(II)-halide complex (a $16e^-$ species.) A copper-activated acetylide, formed from the reaction of a terminal alkyne with CuI in the presence of a

base, then undergoes transmetallation with the palladium complex to form a new alkynyl-Pd-aryl species, which reductively eliminates to form the required product and the regenerated catalyst.[89] The mechanism of the reaction is shown below (Fig 6.2).

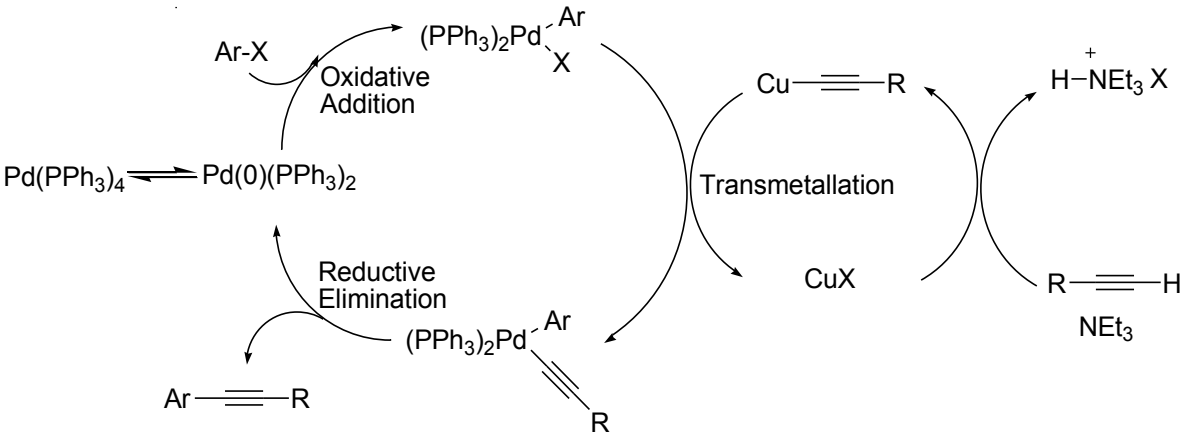


Fig. 6.2: The mechanism of the Sonogashira cross-coupling reaction.[89]

Practically, Sonogashira reactions may be carried out under a large range of reaction conditions involving many different solvents as well as bases either inorganic or organic. In such a way, the reaction conditions can be optimized for any given substrate.[70,72,73,89,90] The solvent used has been shown to have a large influence on reaction yields as well as product specificity. The Sonogashira reaction is also very oxygen-sensitive when using $Pd(PPh_3)_4$ and all solvents and reactants need to be thoroughly degassed with an inert gas such as nitrogen or argon before the oxygen-sensitive palladium catalyst is introduced.

6.3 The synthesis of ANP arylethynyltriazole derivative 24'.

The literature premise for the synthetic work towards the target arylethynyltriazole ANP **24',** came from 2 articles from 2007 and 2008 by Zhu et al.,[72,73] in which the authors described their synthesis of **A** through a Sonogashira cross-coupling. They started by converting the commercially available methyl 5-amino-1*H*-1,2,4-triazole-3-carboxylate to methyl 5-bromo-1*H*-1,2,4-triazole-3-carboxylate **11** via a Sandmeyer diazotization reaction. Subsequent ammonolysis, alkylation with an acyclic side chain and deprotection yielded precursor **P**[73] (Fig 6.3) which was then used to synthesise a large library of aryltriazole[73] and arylethynyltriazole[72] acyclic nucleoside analogues that included the promising compound **A**.[72] Their synthesis served as a good platform for our own investigations into the synthesis of **24'** (Fig 6.3).

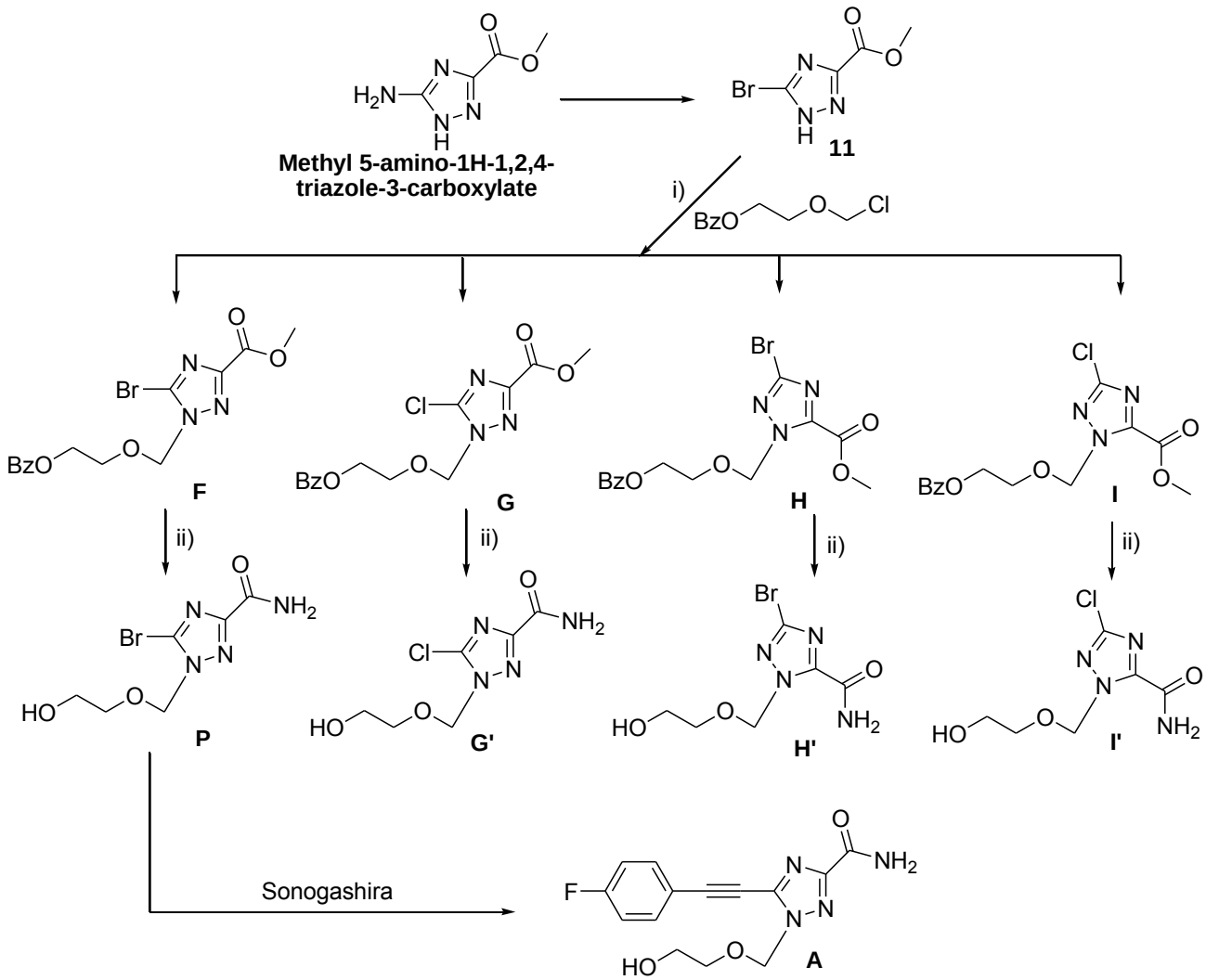


Fig. 6.3: The synthesis of **1f** via the intermediate **P**, as described by Zhu et al.[72,73] Reagents and conditions: i) NaH, CH_3CN, RT; ii) NH_3, MeOH.

The first step of their synthesis involved the transformation of methyl 5-amino-1*H*-1,2,4-triazole-3-carboxylate to methyl 5-bromo-1*H*-1,2,4-triazole-3-carboxylate (**11**) through a Sandmeyer reaction, which was then alkylated with a protected side chain group (Fig 6.3).[72] This reaction gave four products as the two anticipated regioisomers, methyl 5-bromo-1*H*-1,2,4-triazole-3-carboxylate and methyl 5-bromo-2*H*-1,2,4-triazole-3-carboxylate derivatives **F** and **H**, as well as two 5-chloro derivatives **G** and **I,** which formed as the result of a halide exchange with the chloride ion generated from the alkylation reaction.[72] According to the paper, the methyl esters **F** and **G** could not be separated by chromatography and the distinction between the two compounds could only be seen in the different chemical shifts of the anomeric carbons in the ^{13}C NMR spectra; however, after ammonolysis (which resulted in the transformation of the methyl esters to the amides and removal of the benzoyl protecting groups), **P** and **G'** could be separated through recrystallization, after which their structures were confirmed by X-Ray crystallography. The 5-chloro derivatives **G'** and **I'** were found to be poor substrates for the Sonogashira reaction and a method was eventually developed which avoided its formation.[72]

Given these results,[72,73] the exact synthetic route towards the synthesis of **24'** was initially unclear. Retrosynthetic analysis identified a choice regarding the timing of the alkylation step in view of the halide exchange issue discussed above. It was thus decided to carry out the Sonogashira reaction first in order to circumvent this problem by removing the bromide early (Fig 6.4 **X**).

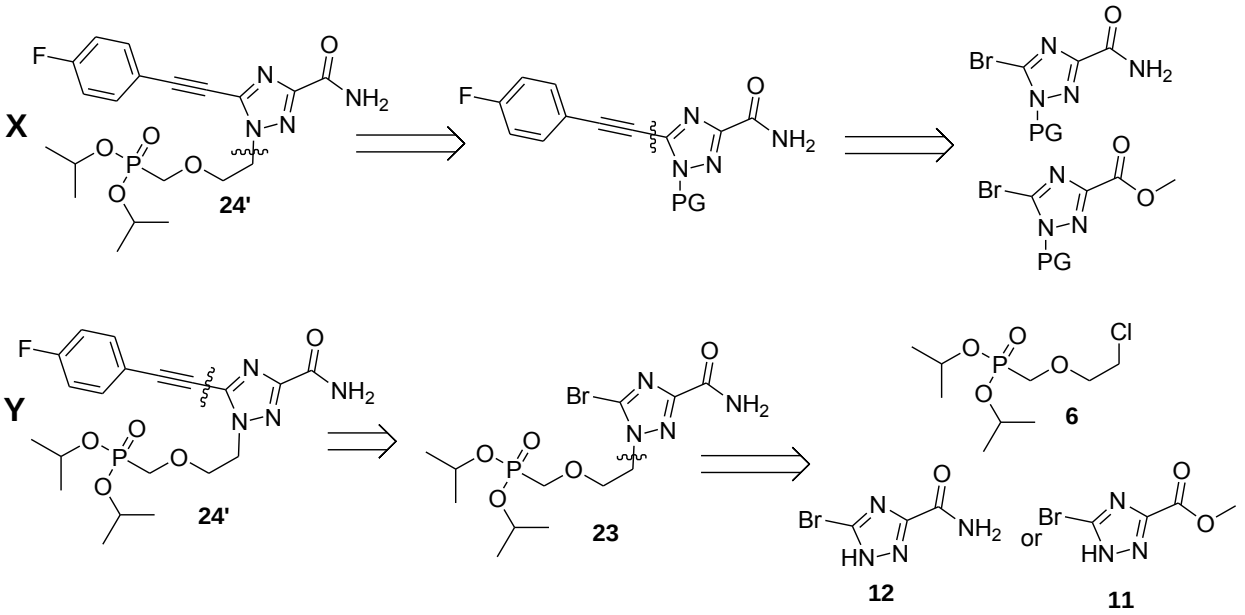


Fig. 6.4: The retrosynthetic analysis of **24'** via the two possible initial disconnections.

The synthesis, as shown in Fig 6.5, began with a Sandmeyer reaction of the commercially available methyl 5-amino-1*H*-1,2,4-triazole-3-carboxylate, the mechanism of which is shown in Fig 6.6. The reaction proceeds via an aromatic diazonium salt formed from reaction of the aromatic amine with nitrous acid formed *in situ*. In the presence of a copper(I) halide catalyst, the diazonium salt then transforms through a radical-nucleophilic aromatic substitution to an aromatic halide while eliminating nitrogen gas, which helps drive the reaction through an increase in entropy (Fig 6.6).

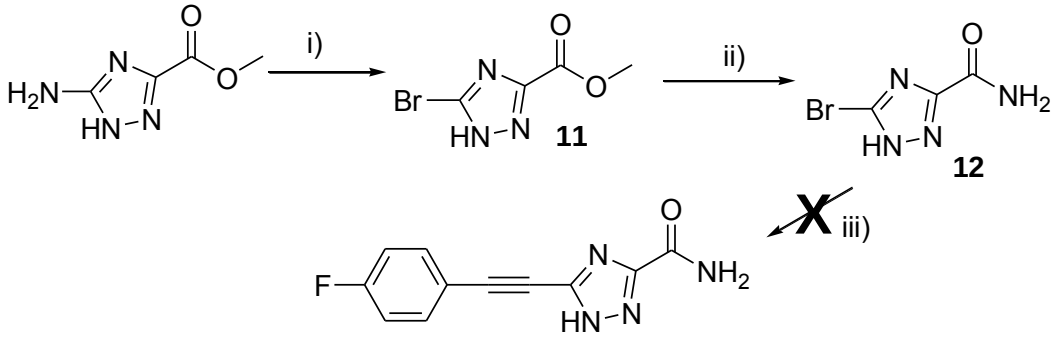


Fig. 6.5. The synthetic strategy for the synthesis of **24'** developed from the retrosynthetic route **X**. Reagents and conditions: i) $NaNO_2$, H_2SO_4, KBr (2 eq), CuI (5 mol%), rt; 75%; ii) NH_3, MeOH; 98% iii) 1-ethynyl-4-fluorobenzene, $Pd(PPh_3)_4$, CuI, NEt_3, THF/DMF, rt-80°C.

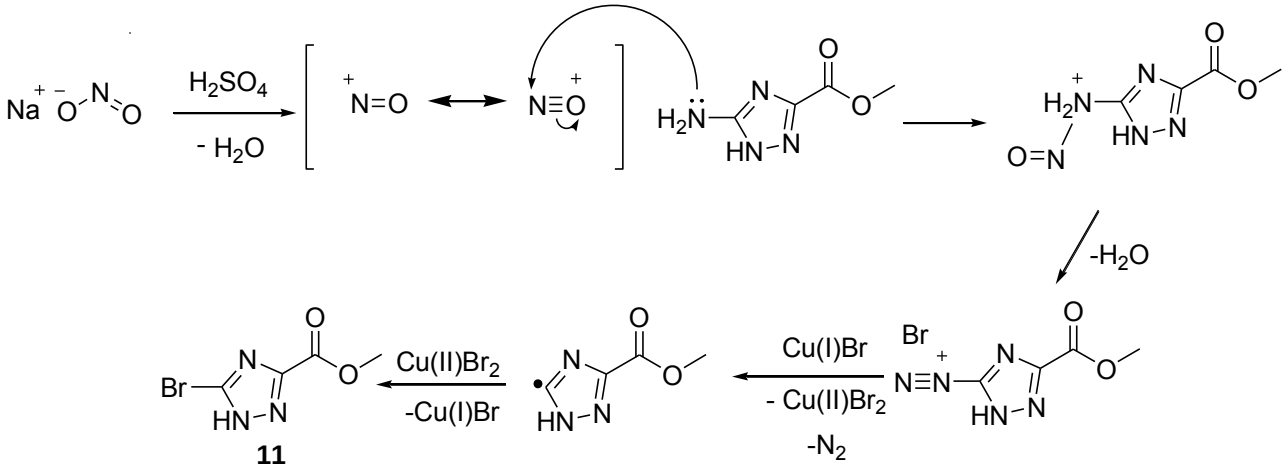


Fig. 6.6: The reaction mechanism of the Sandmeyer reaction.

The aromatic diazonium salt of methyl 5-amino-1*H*-1,2,4-triazole-3-carboxylate was formed at 0°C, after which a solution of potassium bromide (2 eq., large excess relative to catalyst) and a catalytic amount of Cu(I)I (5 mol%) was slowly added to the solution. The ideal catalyst to use for this reaction would have been Cu(I)Br, but due to its unavailability at the time, Cu(I)I was used as a substitute source of Cu(I) ions to catalyse the reaction , while the overwhelming excess of bromide ions provided the nucleophile. Upon the addition of the CuI and potassium bromide solution, nitrogen gas rapidly evolved and the solution was allowed to stir at room temperature for 3hrs to ensure complete conversion of the diazonium salt to the bromide **11,** which could be readily extracted from the aqueous reaction mixture with ethyl acetate and purified by column chromatography, in a good yield of 75%. ^{1}H NMR analysis, as expected, only revealed a resonance for the methyl ester, which in DMSO appeared as a singlet a δ_H 3.89. The ^{13}C NMR spectrum was unable to reveal the two quaternary carbons contained in the triazole ring and was only able to show the carbonyl and methyl resonances at δ_C 157.4 and δ_C 52.7. The only true confirmation of structure was obtained through the use of high resolution mass spectroscopy which gave HRMS (EI): 204.9483 (M^+) correlating closely with the molecular formula $C_4H_4BrN_3O_2$ (expected, 204.9487 (M^+)).

Stirring **11** in the presence of ammonia and methanol generated the amide **12** in a yield of 98%. The ^{1}H NMR spectrum of **12** showed two new broad N-H singlets at δ_H 8.25 (brs, 1H, N-H) and 7.96 (brs, 1H, N-H), indicating the expected restriction of rotation around the amide C-N bond. The ^{13}C NMR spectrum successfully resolved the three carbon resonances at δ_C 157.4 (C=O), δ_C 150.9 (C-1) and δ_C 150.0 (C-2). Again, the structure was confirmed with the use of mass spectroscopy with HRMS (ES): 190.9568 (M^+ + H), which correlated exactly with the molecular formula $C_3H_4BrN_4O$ (expected M^+ + H, 190.9568).

With **12** in hand, a Sonogashira reaction was attempted with 1-ethynyl-4-fluorobenzene, using triethylamine as base. The solvent system of THF/DMF (2:1) was chosen, as this combination had been used in our laboratory in the past with great success;[90] however, the reaction failed (Fig 6.5). The failure of the reaction was thought to be due to the fact that the triazole was being deprotonated and that this was in

some way interfering with the formation of the aryl-Pd(II)-halide complex or the formation of the copper-activated acetylide species. In order to counteract this problem, it was decided to N-protect the triazole derivative **12** (Fig 6.7).

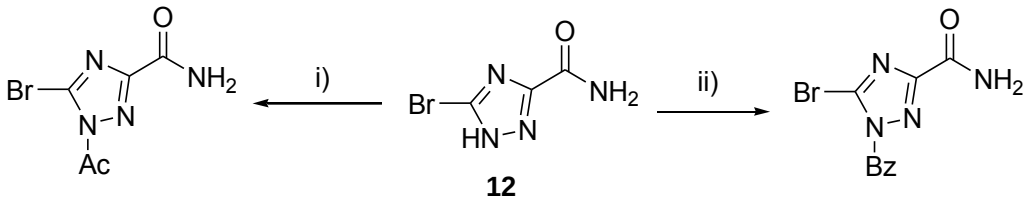


Fig. 6.7: The strategy for the synthesis of protected derivatives of **12.** Reagents and conditions: i) acyl chloride, NaH, DMF; ii) benzoyl chloride, NaH, DMF.

An attempt was made to protect the triazole ring of **12** by acylation of the triazole nitrogen with both acyl chloride and benzoyl chloride in DMF. The reaction of **12** with benzoyl chloride showed no progress even after 12 hours at reflux while the formation of the acyl-protected derivative of **12** was also unsuccessful. Instead of trying additional N-protecting groups, despite the concerns surrounding the halide exchange mentioned previously, and the worry that the presence of the bulky phosphonate group would sterically hinder the Sonogashira reaction, it was decided to N-alkylate first with synthon **6** to form the precursor to the Sonogashira reaction, **23'** (Fig 6.8)**.** This would solve the problem of protecting the triazole ring and would eliminate a protection and deprotection step from the synthesis.

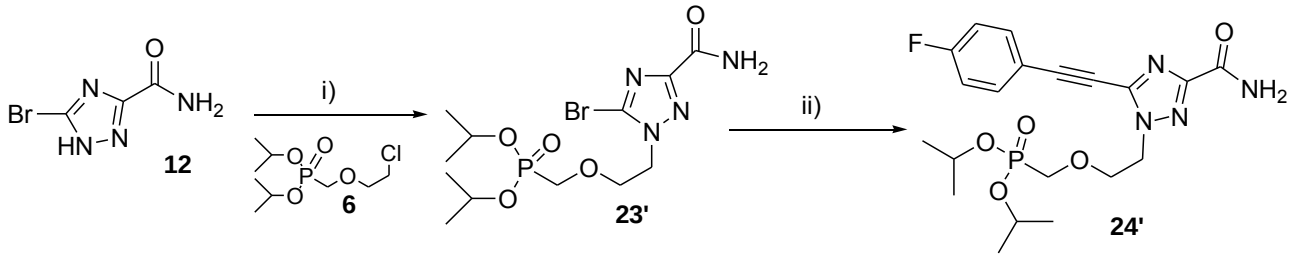


Fig. 6.8. The synthetic strategy for the synthesis of **24'** developed from the retrosynthetic route **Y**. Reagents and conditions: i) Cs_2CO_3, DMF, 100°C 65%. ii) 1-ethynyl-4-fluorobenzene, $Pd(PPh_3)_4$, CuI, NEt_3, THF/DMF, rt-reflux.

Thus, **12** was reacted with the synthon **6** with Cs_2CO_3 as a base under the same reaction conditions used in the general procedure which was described in Chapter 5. In this case the reaction took 48 hours to reach completion, which was much longer than the other reactions. This could possibly be attributed to the strong electron-withdrawing groups present on the triazole ring, which would decrease the nucleophilicity of the triazole nitrogen, resulting in a much lower reaction rate. Isolation of the major reaction product afforded **23** in a moderate yield of 65%. Its IR spectrum showed the presence of the amide and phosphonate moieties with bands at ν_{max} 3508 (N-H) cm^{-1}, 3390 (N-H) cm^{-1}, 1708 (C=O) cm^{-1}, 1284 (P=O) cm^{-1} and 999 (P-O) cm^{-1}. The ^1H NMR spectrum of **23** confirmed the presence of the amide hydrogens at δ_H

7.27 (s, 1H, N-H) and δ_H 6.54 (s, 1H, N-H), as well as all the resonances typical of the side chain and phosphonate groups at δ_H 4.65 (m, 2H, C<u>H</u>$(CH_3)_2$), δ_H 3.98 (t, *J* = 5.4 Hz, 2H, H-2), δ_H 3.71 (d, J_{PH} = 8.3 Hz, 2H, H-1), δ_H 1.26 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$) and δ_H 1.23 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$). The ^{13}C NMR spectrum also displayed all the expected resonances typical of the ANP triazole-3-carboxamide family of compounds. Interestingly, H-3 appeared downfield δ_H 4.83 (t, *J* = 5.4 Hz, 2H, H-3) analagous to the example of the triazole-3-carboxamide compound **20b** which was disscused in Chapter 5. **20b** featured a downfield H-3 triplet at δ_H 4.90 (t, *J* = 5.5 Hz, 2H, H-3), a similar chemical shift to that possessed by **23.** This suggests the structure of **23** (and hence **24**) is akin to that of **20b** rather than **20a** (Fig 6.9)**.** Given the analogy with **20b**, the structure of the alkylation product was assigned as **23** and not **23'.** Unfortunately an HMBC spectrum of **23** was unavailable, while the HMBC spectrum of **24** was inconclusive (discussed below). Thus the exact structures of **23** and **24** remain conjecture.

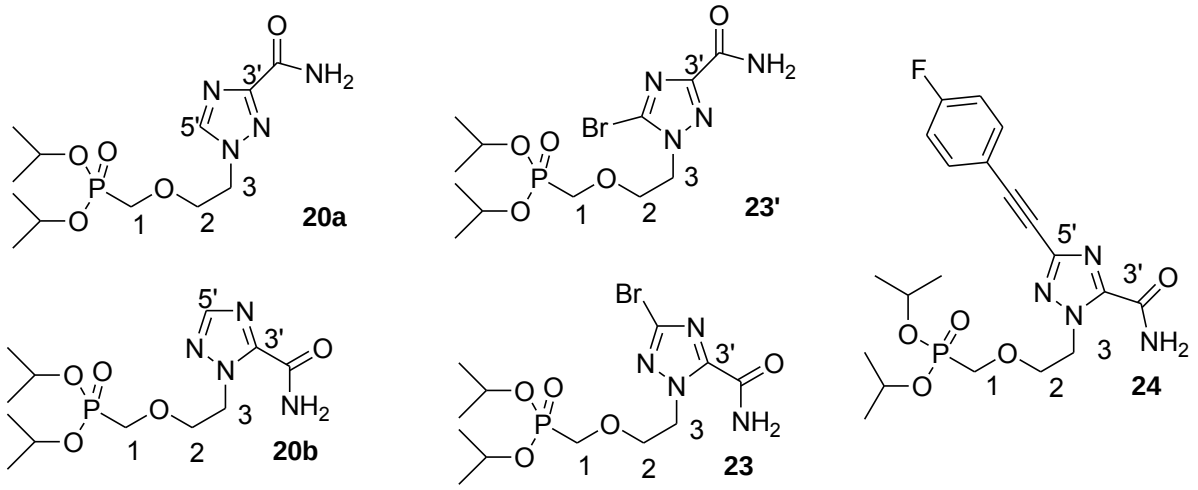


Fig. 6.9: The structures of **20a** and **20b** and the most likely structures of **23** and **24.**

The ^{1}H and ^{13}C NMR and mass spectra of the isolated product, **23,** did not show any evidence of the 5-chlorotriazole derivative. If the 5-chlorotriazole had been formed as an inseparable impurity as discussed by Zhu et al.,[73] one might have expected to see a residual H-3 peak in the NMR spectra with a slightly different chemical shift to that possessed by the target bromide derivative. No residual by-product peak could be observed implying that an aromatic substitution of the bromine did not take place under these specific reaction conditions.

With **23** in hand, work began on the Sonogashira cross coupling reaction using the conditions that had previously been used in our laboratory.[90] Tetrakis(triphenylphosphine)palladium(0), copper(I) iodide and 1-ethynyl-4-fluorobenzene were added to a thoroughly degassed solution of **23** and triethylamine in THF/DMF (2:1), under an inert argon atmosphere. After stirring at room temperature for 2 hours, no reaction was observed so the temperature of the solution was increased to allow the reaction mixture to

reflux overnight. This still did not yield a result and after further review of the literature, an approach developed by Zhu et al.,[72] which used much harsher reaction conditions, was adopted.

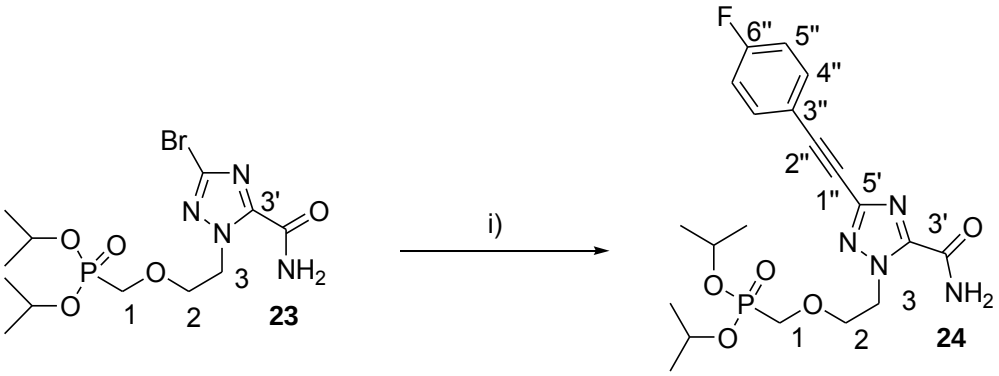


Fig. 6.10: Reagents and conditions i) 1-ethynyl-4-fluorobenzene, $Pd(PPh_3)_4$, CuI, Li_2CO_3, Dioxane/water (3:1), 100°C MW, 82%.

According to the method by Zhu et al.,[72] a solution of **23** in dioxane and water (3:1) was prepared and thoroughly degassed by bubbling argon gas through it for 30 minutes. Lithium carbonate, 1-ethynyl-4-fluorobenzene, copper (I) iodide, and tetrakis(triphenylphosphine)palladium(0) (weighed out under argon) was then added to the reaction vessel which was flushed with argon gas. The reaction mixture was then irradiated under pressure in a microwave reactor for 30 min and worked up with a disodium ethylenediaminetetraacetate (EDTA) solution in order to remove any copper ions. Tlc analysis of the resulting mixture revealed the product spot to have an R_f value almost identical to that of the starting material. The only indication that it was a new product was that it now possessed a dark blue semi-fluorescent tint under UV light, which was attributed to the extensive conjugated arylethynyltriazole system. The product could be readily isolated by column chromatography by first removing the tetrakis(triphenylphosphine)palladium(0) by flushing with petroleum ether (100%) and then eluting the product with ethyl acetate (80%-100%) to afford **24** in an 82% yield (Fig 6.10).

IR spectroscopy showed the presence of a primary amide with bands at ν_{max} 3508 (N-H) cm^{-1}, 3390 (N-H) cm^{-1} and 1705 (C=O) cm^{-1}. The phosphonate could be seen at ν_{max} 1234 (P=O) cm^{-1} and 994 (P-O) cm^{-1}, while the new triple bond showed an absorbance at ν_{max} 2231 (C≡C) cm^{-1}. The ^{1}H NMR spectra of **24** (Fig 6.11) showed the amide resonances at δ_H 7.27 (s, 1H, N-H) and δ_H 6.14 (s, 1H, N-H), while again the peaks associated with the diisopropyl phosphonate were displayed δ_H 4.68 (m, 2H, C<u>H</u>$(CH_3)_2$), δ_H 3.74 (d, J_{HP} = 8.3 Hz, 2H, H-1), δ_H 1.29 (d, J = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$) and δ_H 1.27 (d, J = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$). As in the ^{1}H NMR spectrum of **23** the H-3 resonance was observed to be relatively upfield at δ_H 4.90 (t, J = 5.4 Hz, 2H, H-3), which is analogous to the H-3 resonance in **20b.** The aromatic region of the ^{1}H NMR spectrum revealed H-4'' as a doublet of doublets at δ_H 7.56 (dd, $J_{HH;\ HF}$ = 8.4, 5.3 Hz, 2H, H-4'') which arises from an H-H coupling (J_{HH} = 8.4) and a long range J^4 H-F coupling (J_{HF} = 5.3). A second aromatic signal, interpreted as a doublet of

doublets with two equal coupling constants, was assigned to H-5'' at δ_H 7.05 (dd, $J_{HH;\ HF}$ = 8.4, 8.4 Hz, 2H, H-5'') and displayed a J^3 H-F coupling (J_{HF} = 8.4), as well as an H-H coupling (J_{HH} = 8.4) which resulted in the formation of an AB doublet pair with H-4''.

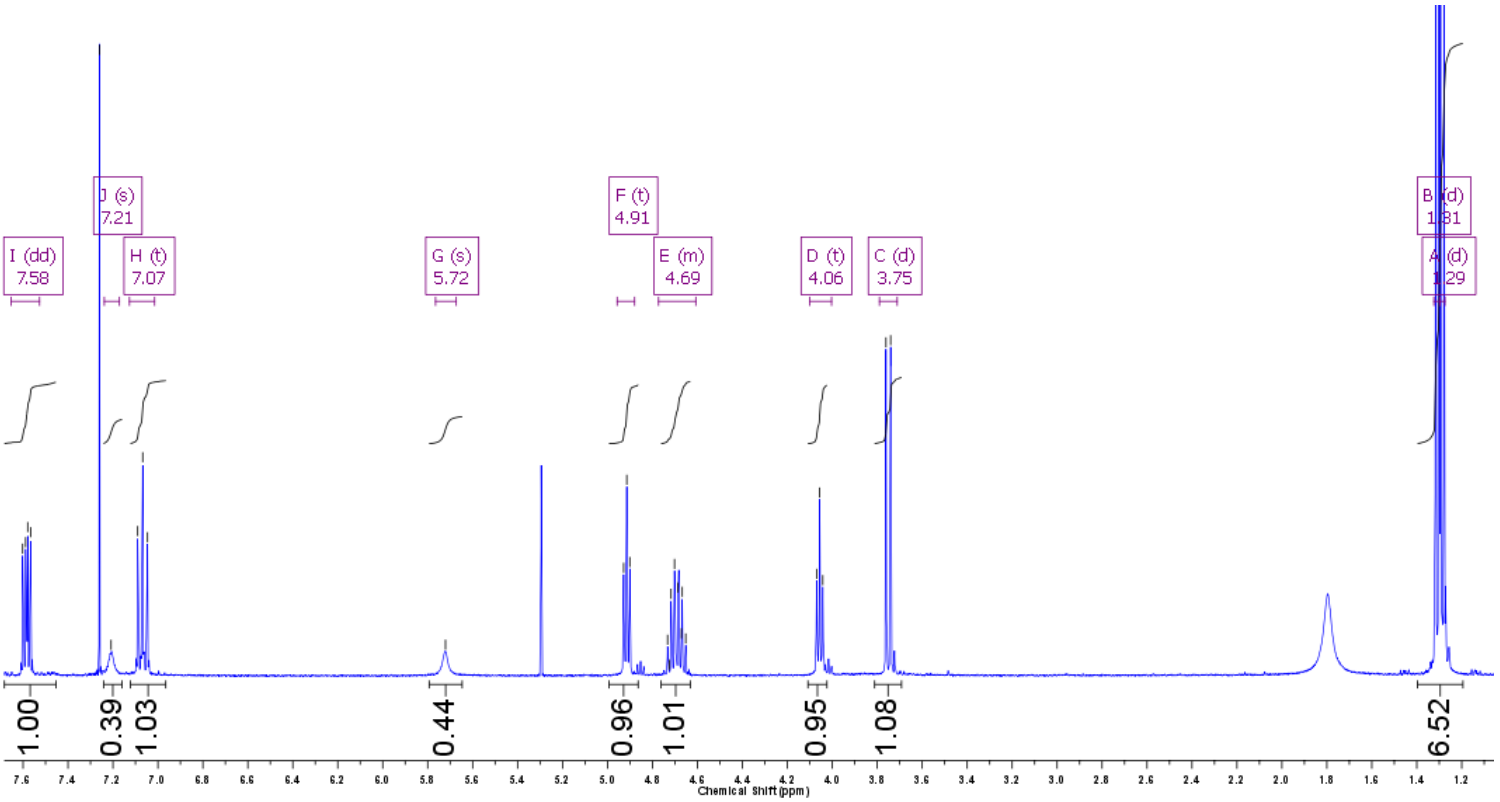


Fig. 6.11: The ^{1}H NMR spectrum of **24.**

The ^{13}C NMR spectrum (Fig 6.12) displayed all the PME side chain resonances at δ_C 71.4 (d, J_{CP} = 6.6 Hz, $\underline{C}H(CH_3)_2$), δ_C 70.9 (d, J_{CP} = 10.6 Hz, C-2), δ_C 65.9 (d, J_{CP} = 167.2 Hz, C-1), δ_C 50.6 (C-3), δ_C 24.2 (d, J_{CP} = 3.7 Hz, $CH(\underline{C}H_3)_2$) and δ_C 24.1 (d, J_{CP} = 4.5 Hz, $CH(\underline{C}H_3)_2$) as expected. The diagnostic triple bond carbon's C-1" and C-2" were revealed at δ_C 79.3 (C-1'') and δ_C 89.1 (C-2''). The carbon resonances at δ_C 134.3 (d, J_{CF} = 8.6 Hz, C-4'') and δ_C 116.0 (d, J_{CF} = 22.3 Hz, C-5'') were assigned based on their C-F coupling constants and their proximity to the fluorine atom. All of the above assignments were confirmed through the use of HSQC and HMBC NMR. A large J^1 C-F coupling was observed for the resonance at δ_C 163.4 (d, J_{CF} = 251.6 Hz, C-6''), while a small long range J^4 C-F coupling was observed for the resonance indicative of C-3'' at δ_C 117.5 (d, J_{CF} = 3.4 Hz, C-3''). Further evidence for the overall structure of **24** was provided by high resolution mass spectroscopy, with HRMS (ES): *m/z* found as 453.1712 (M^+ + H), which corelated closely with the molecular formula $C_{20}H_{27}FN_4O_5P$ (expected (M^+ + H), 453.1703).

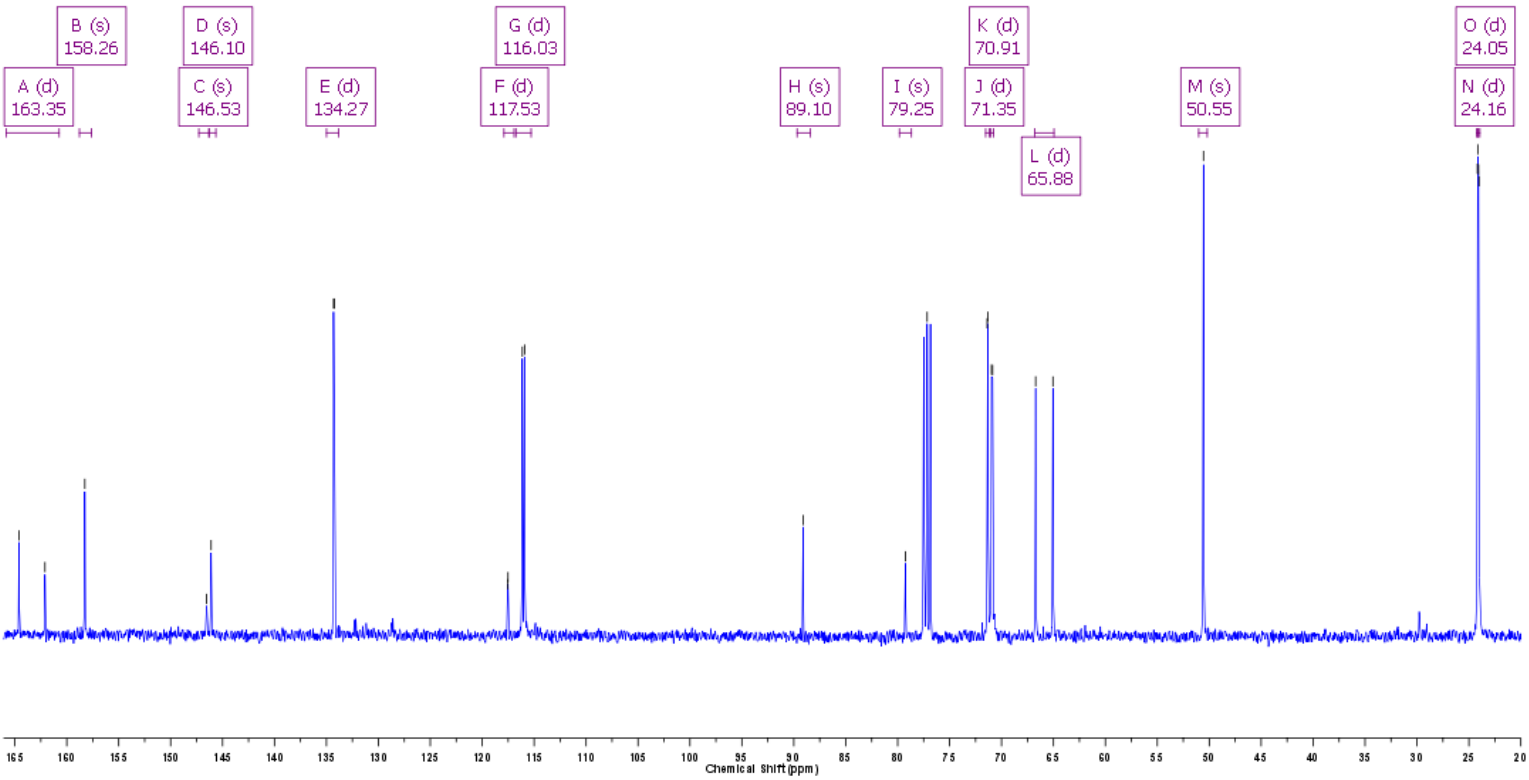


Fig. 6.12: The ^{13}C NMR spectrum of **24.**

The HMBC NMR spectrum was inconclusive with regard to assignment of the exact structure as either **24** or **24'** (Fig 6.9)**.** The ^{13}C chemical shifts of the carbon resonances of C-3'and C-5' at δ_C 146.5 (C-3') and δ_C 146.1 (C-5') were very close together and the resulting cross-peak could not be distinguished as a C-5': H-3 or C-3': H-3 cross coupling. Thus, the assignment of the structure of **24** was made based on analogy with the NMR data and structure of compound **20b** (Fig 6.9)**.**

Chapter 7

Cleavage of the diisopropyl phosphonate ester.

7.1 Introduction

As discussed in Chapter 2, the activity of acyclic nucleotide phosphonates in general lies in their ability to readily undergo intracellular phosphorylation to active di-phosphorylated species, which is a requirement for attachment onto the growing chain of genetic material through a substitution reaction on the phosphonate phosphorus. Such activation proceeds via the di-phosphonic acid, and as studies have shown the enzymatic hydrolysis of the isopropyl protecting group is slow enough to decrease the overall potency of the drug,[43] a selection of the diisopropyl protected derivatives discussed in chapter 5 were converted to their diphosphonic acids through a bromotrimethylsilane (TMSBr) mediated hydrolysis (Fig 7.1).

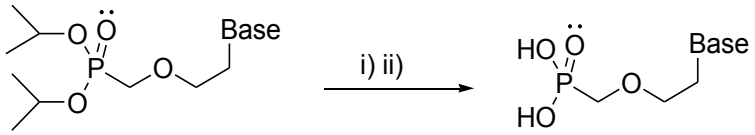


Fig. 7.1. Reagents and conditions: i) TMSBr, DCM or CH_3CN, 0°C-rt; ii) MeOH or H_2O.

The reaction mechanism (Fig 7.2) involves the phosphonate oxygen reacting with TMSBr to form a phosphonium salt with the expulsion of a bromide ion.[91] This then promotes an Arbuzov-like rearrangement, releasing isopropyl bromide, resulting in the isopropyl trimethylsilyl protected intermediate. The process repeats itself to afford the bis(trimethylsilyl) protected intermediate, which can then be hydrolysed by either methanol or water, to yield the product phosphonic acid.[91]

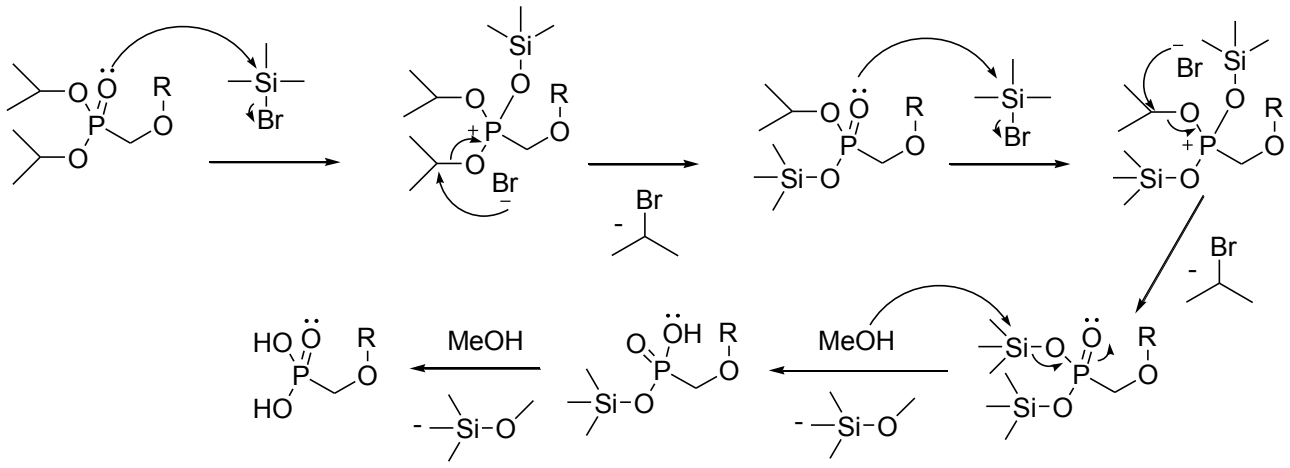
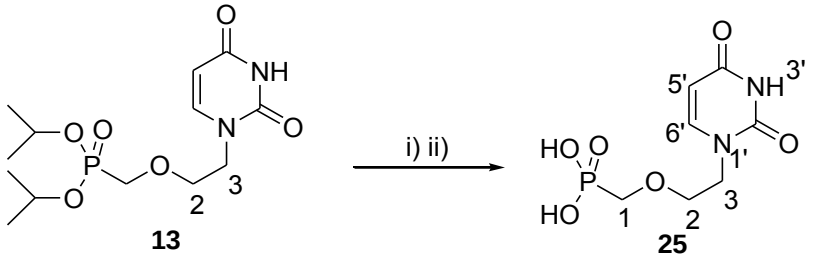


Fig. 7.2: The reaction mechanism of the TMSBr-mediated hydrolysis of phosphonate esters.[91]

7.2 Synthesis of phosphonic acid ANP derivatives.

The reaction described above has been widely used in the field of ANP synthesis[56,66,81,92] as an efficient, high yielding way to access phosphonic acids from isopropyl esters. Generally, the reaction is performed in either DCM or CH_3CN at room temperature with an excess of TMSBr, usually between 3 or 6 equivalents. Practically, the most difficult aspect of the reaction is the purification of the highly polar products and various techniques such as aqueous extraction, recrystallization or ion-exchange chromatography, have been used.[56,66,81,92]

Where possible, deprotection of the diisopropyl protected ANP's discussed in Chapter 5 was carried out under the general reactions conditions described above. The first deprotection involved the uracil derivative **13** which when yielded the phosphonic acid **25** (Fig 7.3)**.**



13 **25**

Fig. 7.3. Reagents and conditions: i) TMSBr, DCM, 0°C-rt; ii) H_2O, 84%

A highly polar tlc solvent system (used for all the phosphonic acid derivatives discussed in this chapter) consisting of water, acetic acid, 1-butanol and acetone (1:1:1:1), was able to lift the product off the baseline of the tlc plate which confirmed the reaction to be complete. A small amount of water was added to hydrolyse the bis-silyl phosphonate esters, after which all solvents were removed under high vacuum. Toluene was used to azeotropically aid the removal of water and the resulting residue could then be dissolved in hot methanol. The product was then precipitated out of this solution by adding cold DCM to afford **25** as a fine white powder in an 84% yield.

IR spectroscopy from a potassium bromide pellet showed absorption corresponding to the phosphonic acid moiety at ν_{max} 3020 cm^{-1} (O-H), 1250 cm^{-1} (P=O) and 1013 cm^{-1} (P-O). The ^{1}H NMR spectrum of **25**, as in the spectrum of **13**, showed the AB doublet pair of H-6' and H-5' at δ_H 7.58 (d, *J* = 7.9 Hz, 1H, H-6') and δ_H 5.61 (d, *J* = 7.9 Hz, 1H, H-5'). The resonances indicative of H-3, H-2 and H-1 could be observed, as expected, at δ_H 3.96 (t, *J* = 4.9 Hz, 2H, H-3), δ_H 3.80 (t, *J* = 4.9 Hz, 2H, H-2) and δ_H 3.74 (d, *J* = 8.8 Hz, 2H, H-1) while the absence of C<u>H</u>$(CH_3)_2$ multiplet and the diastereotopic methyl doublets confirmed the successful removal of the isopropyl groups. The uracil N-H exchanged with the deuterated methanol NMR solvent and was not observed. The ^{13}C NMR spectrum showed all the expected resonances indicative of the uracil ring and side chain. Since the <u>C</u>H$(CH_3)_2$ had been eliminated the only remaining resonance in that region of the spectrum, at δ_C 72.3 (d, J_{PC} = 11.6 Hz, C-2) could be assigned to C-2. The chemical shift and P-C coupling constant J^4 =

11.6 Hz of C-2 in the ^{13}C NMR spectrum of **25**, correlated with the chemical shift and coupling constant of the resonance assigned to C-2 (and not C$\underline{H}$(CH$_3$)$_2$) in the ^{13}C NMR spectrum of **13** (Chapter 4). High resolution mass spectroscopy provided further evidence of the successful deprotection of the phosphonate showing: HRMS (ES) for **25** of 251.0423 (M^+ + H), which correlated with the molecular formula $C_7H_{12}N_2O_6P$ (expected, 251.0427 (M^+ + H)).

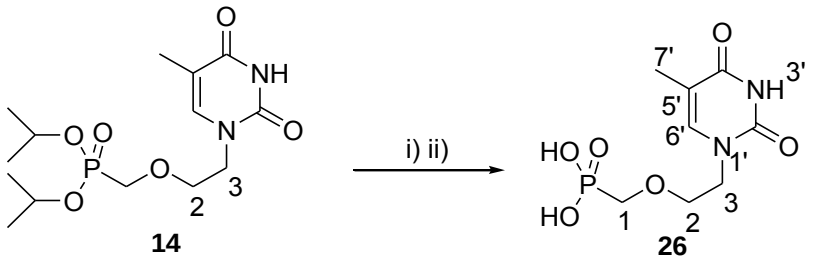
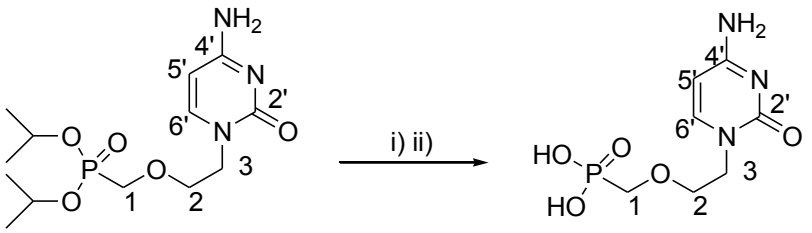


Fig. 7.4. Reagents and conditions: i) TMSBr, DCM, 0°C-rt; ii) H_2O, 83%

The same procedure used to prepare and isolate **25** was also used to form the thymine derivative **26** in a yield of 83%. The ^{1}H NMR spectrum of **26** again revealed the resonances associated with the PME side chain at δ_H 3.93 (t, *J* = 5.0 Hz, 2H, H-3), δ_H 3.79 (t, *J* = 5.0 Hz, 2H, H-2) and δ_H 3.74 (d, J_{PH} = 8.9 Hz, 2H, H-1). The thymine hydrogen H-6' resonated as a singlet at δ_H 7.43 (s, 1H, H-6') while the thymine methyl H-7' was observed at δ_H 1.86 (s, 3H, H-7'). The ^{13}C NMR spectrum revealed the expected resonances of the thymine ring, thymine methyl and the PME side chain including the P-C coupled doublet at δ_C 71.9 (d, J_{PC} = 11.7 Hz, C-2), which through the same process of elimination discussed above, helped confirm the C-2 assignment in the ^{13}C NMR spectrum of the precursor **14.** Similarly, mass spectroscopy confirmed the correct structure with an HRMS (ES) peak at 265.0581 (M^+ + H), with $C_8H_{14}N_2O_6P$ requiring 265.0584 (M^+ + H).

The next derivatives, **27-36**, could not be crystallized or precipitated, and attempts to purify them via anion-exchange chromatography also proved unsuccessful, probably due to excessively strong ionic interactions with the positively charged ion exchange resin. Any attempts to dislodge the compound from the resin with sufficiently concentrated hydrochloric acid resulted in degradation of the target molecule.

Topalis et al.[92] discussed an alternative method for isolating phosphonic acid derivates whereby the silyl phosphate esters, generated from the reaction of the protected phosphonate and TMSBr, were hydrolysed by methanol to release the target phosphonic acid and methoxytrimethylsilane (TMSOMe), a volatile by-product which could be evaporated off under high vacuum with mild heating. The polar phosphonic acid could then be extracted into an aqueous medium while any other residual organic impurities were removed. Removal of the water by evaporation, in most cases gave the final product as an amorphous solid or thick oil. The first product to be purified in this way was the cytosine derivative **27** which was formed from the reaction of **15** with TMSBr (Fig 7.5).



15 **27**

Fig. 7.5. Reagents and conditions: i) TMSBr, DCM, 0°C-rt; ii) MeOH, 91%.

Compound **27** was isolated as an amorphous solid in a yield of 91%, with a melting point range of 164-166°C which is similar to the melting-point range reported in the literature [lit Mp: 167-168°C].[81] The IR spectrum of **27**, recorded from a potassium bromide pellet, showed absorptions at ν_{max} 3483 cm^{-1} (O-H), 3313 cm^{-1} (N-H stretch), 1684 cm^{-1} (C=O), 1660 cm^{-1} (N-H bend), 1250 cm^{-1} (P=O) and 998 cm^{-1} (P-O), indicating the presence of the amine, the cytosine carbonyl and the phosphonate moiety. The ^{1}H NMR spectrum (Fig 7.6) was diagnostic by virtue the absence of the two isopropyl peaks. H-6'and H-5' were revealed as an AB doublet pair at δ_H 7.93 (d, *J* = 7.6 Hz, 1H, H-6') and δ_H 6.04 (d, *J* = 7.6 Hz, 1H, H-5'). In the ^{13}C NMR spectrum, the four cytosine carbon resonances were observed at δ_C 161.6 (C-2'), δ_C 152.3 (C-6'), δ_C 149.1 (C-4') and δ_C 93.8 (C-5'), while the PME side chain resonances were present at δ_C 71.0 (d, J_{PC} = 11.2 Hz, C-2) and δ_C 50.3 (C-3), with a strong P-C phosphonate doublet at δ_C 67.3 (d, J_{PC} = 163.9 Hz, C-1). Mass spectroscopy gave further evidence for the structure showing (ES): 250.0581 (M^+ + H), which corresponded to the molecular formula of $C_7H_{13}N_3O_5P$ (expected 250.0587 (M^+ + H)).

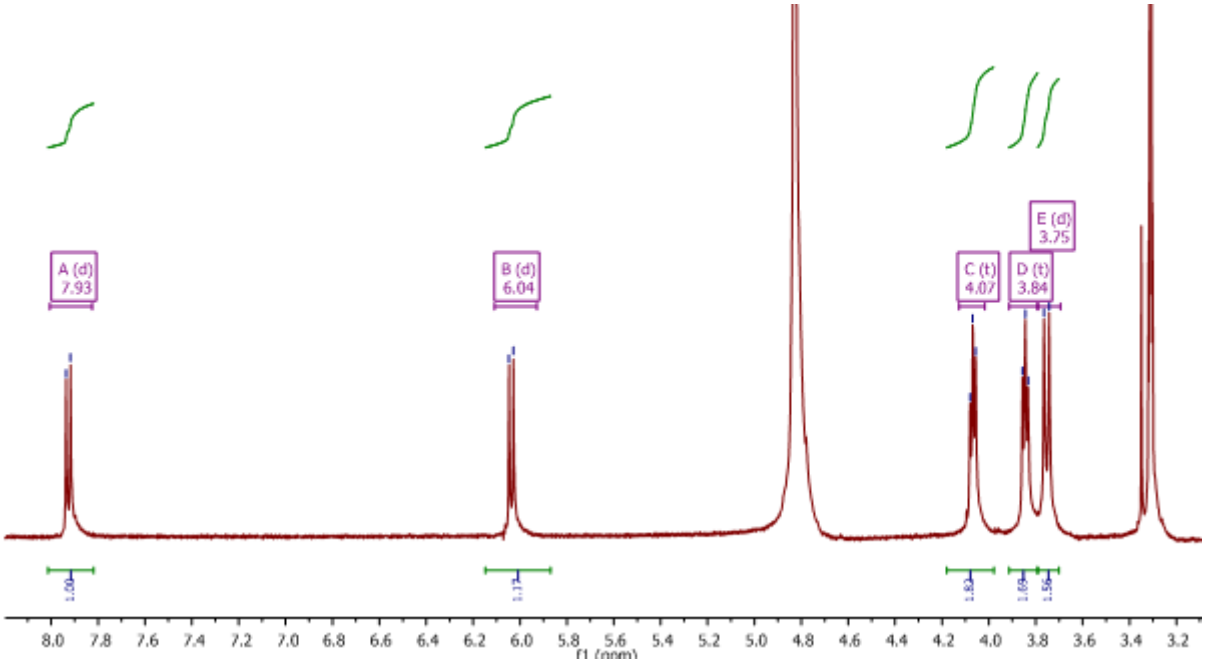


Fig. 7.6: The ^{1}H NMR spectrum of **27.**

After the successful synthesis of **27,** deprotection of the diisopropyl phosphonate 2-hydropyrazine derivatives **16a** and **16b** was then attempted with little success. Tlc of each reaction mixture showed the

conversion of starting material to a possible product; however, all attempts to isolate the new phosphonic acid derivatives of **16a** and **16b** failed, suggesting the products to be degrading during the isolation step. Although the exact reason for the degradation was not clear, it was postulated that the presence of HBr, generated from the reaction of methanol and residual TMSBr, was able to protonate the pyrazine ring, activating it as a leaving group.

The next derivative, **28,** was prepared from the 2,4-diaminopyrimidine derivative **17** using the same method discussed above to prepare **27.** The reaction gave **28** as a thick, dark-yellow oil in a yield of 93% (Fig 7.7).

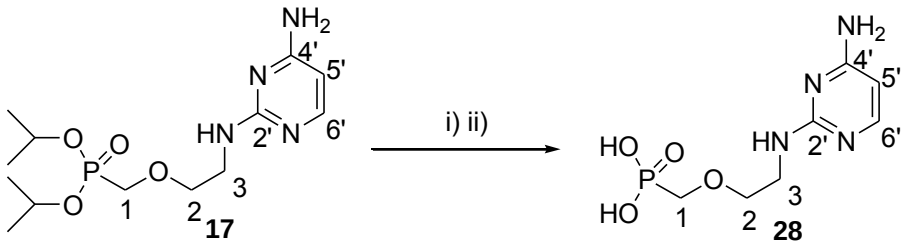


Fig. 7.7. Reagents and conditions: i) TMSBr, DCM, 0°C-rt; ii) MeOH, 93%.

As compounds **28-33** were formed as thick oils only soluble in methanol, water and DMSO, acquisition of their IR spectra from a standard potassium bromide pellet, sodium chloride nujol plate or solution cell was impossible. The IR spectra of **28-33** were thus recorded using an ATR-IR (Attenuated total reflectance-Infrared) spectrometer, whereby the spectrum was recorded directly from the neat, unprepared sample placed on a diamond crystal. Infrared radiation passed through the crystal generates an evanescent wave, which penetrates the sample before being reflected back into the crystal and collected by a sensor. The IR spectrum of **28** displayed absorptions at ν_{max} 3361 cm^{-1} (O-H), 3087 cm^{-1} (N-H stretch), 3023 cm^{-1} (N-H stretch), 1670 cm^{-1} (N-H bend), 1659 cm^{-1} (N-H bend), 1222 cm^{-1} (P=O) and 1076 cm^{-1} (P-O), again indicating the presence of the aromatic amine, and the phosphonic acid.

The ^{1}H NMR spectrum of **28** did not show any signals relating to the isopropyl groups, confirming their removal. The PME side chain resonances were all present at δ_H 3.78 (d, J_{PH} = 8.4 Hz, 2H, H-1), δ_H 3.76 (t, J = 4.9 Hz, 2H, H-2) and δ_H 3.68 (t, J = 4.9 Hz, 2H, H-3) with H-2 and H-3 appearing relatively upfield, as was the case in the ^{1}H NMR spectrum of **17.** H-6' and H-5' were observed as an AB coupled doublet pair with resonances at δ_H 7.56 (d, J = 7.3 Hz, 1H, H-6') and δ_H 6.13 (d, J = 7.3 Hz, 1H, H-5'). The ^{13}C NMR spectrum revealed all of the resonances associated with the PME chain and pyrimidine ring.

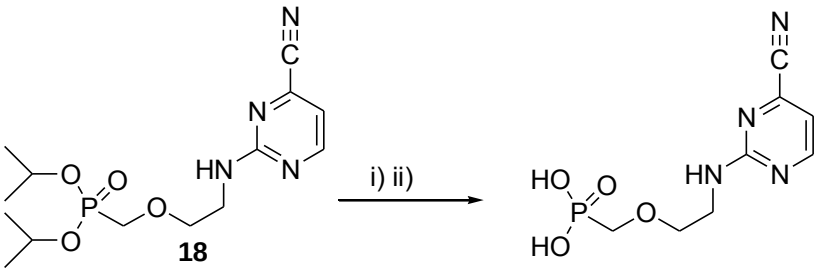


Fig. 7.8. The unsuccessful dealkylation of **18.** Reagents and conditions: i) TMSBr, DCM, 0°C-rt; ii) MeOH.

A dealkylation of the 2-amino-4-cyanopyrimidine derivative **18** using TMSBr in DCM before quenching with methanol, was also attempted but as with the case of **16a** and **16b,** all attempts to workup and purify the reaction product were unsuccessful (Fig 7.8). This was again thought to be due to a side reaction with HBr, possibly involving the cyano group. Work then began on deprotection of the ribavirin analogues **19, 20a, 20b, 21 and 22.** The first of these derivatives was formed from the dealkylation of **19** via the same procedure as before to afford **29** in a yield of 99% (Fig 7.9).

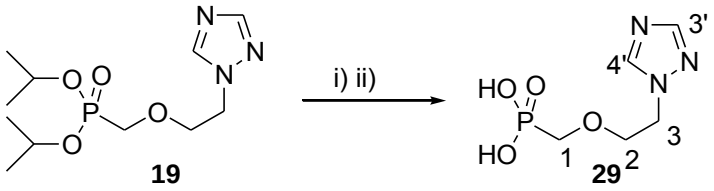


Fig. 7.9. Reagents and conditions: i) TMSBr, DCM, 0°C-rt; ii) MeOH, 99%.

The ^{1}H NMR spectrum of **29** revealed the triazole hydrogens as two singlets at δ_H 9.82 (s, 1H, H-4') and δ_H 8.90 (s, 1H, H-3'), while the PME hydrogens were observed at δ_H 4.70 (t, *J* = 4.9 Hz, 2H, H-3), δ_H 4.08 (t, *J* = 4.9 Hz, 2H, H-2) and δ_H 3.81 (d, J_{PH} = 8.8 Hz, 2H, H-1). The resonances typical of the isopropyl groups were again absent in both the ^{13}C and ^{1}H NMR spectra confirming successful dealkylation. The ^{31}P decoupled spectrum displayed a single phosphorus peak at δ_P 18.6. Mass spectroscopy confirmed the structure showing a protonated molecular ion (M^+ + H) of 208.0479 which corresponds very closely to the molecular formula of $C_5H_{11}N_3O_4P$ which requires 208.0482.

The next derivative, **30**, was synthesised from ribavirin mimic **20a** using the standard dealkylation conditions and was formed as a thick brown oil in a yield of 89% (Fig 7.10).

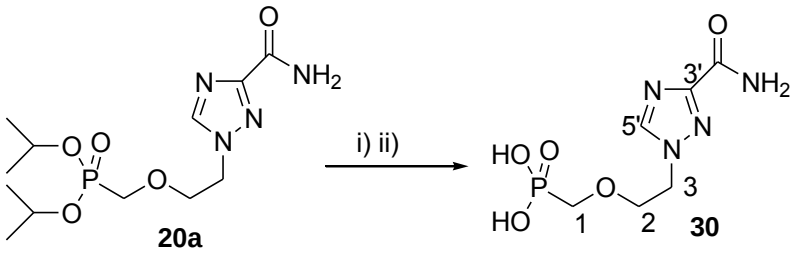


Fig. 7.10. Reagents and conditions: i) TMSBr, DCM, 0°C-rt; ii) MeOH, 89%.

The ^{1}H NMR spectrum of **30** did not show the two distinct N-H amide singlets as those hydrogens were able to exchange with the deuterium from the deuterated methanol solvent. However, ATR-IR spectroscopy showed the N-H amide bands at ν_{max} 3337 cm^{-1} and 3112 cm^{-1}; a region which is known to display primary amide peaks in ATR-IR spectra,[93] while the ^{13}C NMR showed the resonance associated with the amide carbonyl carbon at δ_C 160.4 (C=O). The ^{1}H NMR showed all the expected PME resonances at δ_H 4.64 (t, *J* = 4.9 Hz, 2H, H-3), δ_H 4.06 (t, *J* = 4.9 Hz, 1H, H-2) and δ_H 3.81 (d, J_{PH} = 8.8 Hz, 1H, H-1), as well as a minor unknown impurity (10%) in the aromatic region of the spectrum. The impurity was also reflected as minor peaks in the aromatic region of the ^{13}C NMR spectrum. Despite repeated attempts with conventional purification methods such as recrystallization and extraction, it could not be removed. Electrospray mass spectroscopy revealed a protonated molecular ion peak of 251.0535, which corresponded to the expected value of 251.0540 (M^+ + H) for a compound with a molecular formula of $C_6H_{12}N_4O_5P$.

The same reaction was repeated using the N^{2}-alkylated regioisomer **20b** to afford the ribavirin ANP analogue **31** (Fig 7.11), which was isolated from the reaction mixture in a yield of 89%**.**

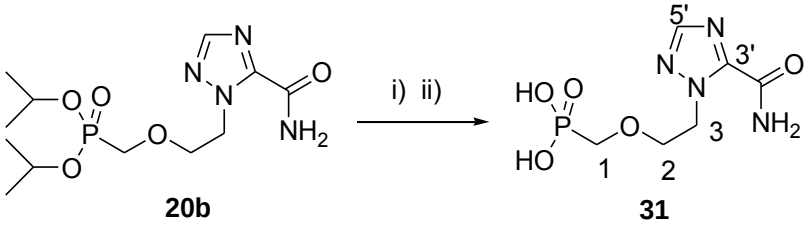


Fig. 7.11. Reagents and conditions: i) TMSBr, DCM, 0°C-rt; ii) MeOH, 89%.

The ^{1}H NMR spectrum of **31** showed the aromatic H-5’ resonance at δ_H 8.25 (s, 1H, H-5’) and also revealed the PME side chain signals at δ_H 4.88 (t, *J* = 5.3 Hz, 2H, H-3), δ_H 4.04 (t, *J* = 5.3 Hz, 2H, H-2) and δ_H 3.78 (d, J_{PH} = 8.7 Hz, 2H, H-1). ^{1}H NMR spectrum showed, through the absence of both isopropyl related resonances, that the compound had been successfully dealkylated. The ^{13}C NMR spectrum confirmed this conclusion as only the PME side chain peaks at δ_C 72.0 (d, J_{PC} = 11.7 Hz, C-2), δ_C 66.8 (d, J_{PC} = 163.4 Hz, C-1) and δ_C 51.4 (C-3), and the triazole and carbonyl carbon resonances at δ_C 159.6 (C=O), δ_C 149.3 (C-5’) and δ_C 147.7 (C-3’), could be observed. IR spectroscopy showed the amide bands at ν_{max} 3376 cm^{-1}(N-H), 3123 cm^{-1} (N-H) and

1688 cm^{-1} (C=O), while phosphonate absorptions were observed as expected at 1249 cm^{-1} (P=O) and 995 cm^{-1} (P-O). Electrospray mass spectroscopy again provided evidence for the structure, showing m/z found to be 251.0535 (M^+ + H) which corresponded closely to the expected value of 251.0540 for the molecular formula $C_6H_{12}N_4O_5P$.

The pyrazole derived diisopropyl protected ANP, **21,** was reacted with TMSBr via the standard reaction procedure to afford the phosphonic acid **32** as an oil in a yield of 97% (Fig 7.12). Unfortunately, **32** was found to be relatively impure with impurity peaks present in all spectra. Preparative reverse phase high performance liquid chromatography (HPLC) could have possibly been used to purify the highly polar product but these facilities were unavailable and any attempts to purify **32** using conventional purification methods met with failure. It is important to note that the starting material, **21**, also contained an impurity which was thought to be a regioisomer as discussed in Chapter 5.

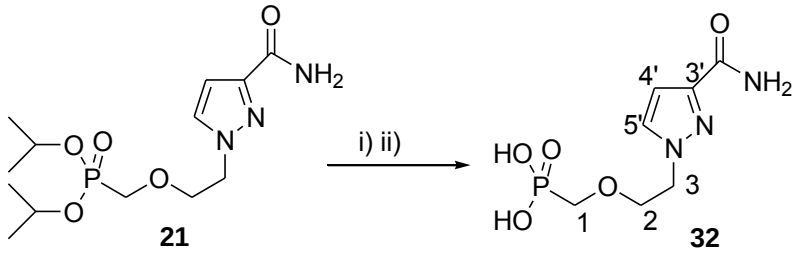


Fig. 7.12. Reagents and conditions: i) TMSBr, DCM, 0°C-rt; ii) MeOH, 97%.

The ^{1}H NMR spectrum of **32** showed the AB doublet pair indicative of H-5' and H-4' at δ_H 7.79 (d, *J* = 2.3 Hz, 1H, H-5') and δ_H 6.77 (d, *J* = 2.3 Hz, 1H, H-4') as well as the PME side chain resonances at δ_H 4.42 (t, *J* = 4.9 Hz, 2H, C-3), 3.96 (t, *J* = 4.9 Hz, 2H, C-2), 3.73 (d, J_{PH} = 8.9 Hz, 2H, C-1). The ^{13}C NMR spectrum showed aromatic carbon resonances at δ_C 146.5 (C-3'), δ_C 133.8 (C-5') and δ_C 107.6 (C-4') as well as the amide carbonyl signal at δ_C 167.0 (C=O), all of which corresponded in chemical shift to the aromatic resonances in the ^{13}C NMR spectrum of the parent molecule **21.** The amide N-H bands in the ATR-IR spectrum were not resolved due to a high level of baseline noise present in the spectrum, which may have been due to high levels of water or methanol in the sample, despite being dried under high vacuum for well over 48 Hours. Mass spectroscopy showed a peak at 250.0587 (M^+ + H) corresponding with the expected value 250.0587 (M^+ + H) for the molecular formula of **32** ($C_7H_{13}N_3O_5P$).

The final derivative to be successfully dealkylated under the standard deprotection conditions was the imidazole-3-carboxamide variant **33,** which was produced from **22** as a thick oil in a yield of 94% (Fig 7.13).

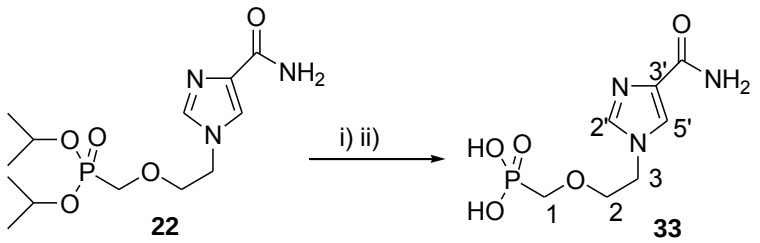


Fig. 7.13: Reagents and conditions: i) TMSBr, DCM, 0°C-rt; ii) MeOH, 94%

ATR-IR spectroscopy revealed the presence of the amide in **33** with absorption at ν_{max} 3325 cm^{-1} (N-H), 3131 cm^{-1} (N-H) and 1680 cm^{-1} (C=O), while the peaks at ν_{max} 1232 cm^{-1} (P=O) and 997 cm^{-1} (P-O) indicated the presence of the phosphonic acid. ^{1}H NMR spectroscopy revealed the aromatic hydrogens, H-2'and H-5', at δ_H 9.09 (s, 1H, H-2') and δ_H 8.25 (s, 1H, H-5'), while as usual the PME side chain peaks were observed as two triplets and a doublet at δ_H 4.55 (t, J = 4.8 Hz, 2H, H-3), δ_H 4.02 (t, J = 4.8 Hz, 2H, H-2) and δ_H 3.83 (d, J_{PH} = 8.8 Hz, 2H, C-1). The ^{13}C NMR spectrum showed the amide carbonyl at δ_C 160.2 (C=O), while the three aromatic carbons were observed at δ_C 138.4 (C-4'), δ_C 129.2 (C-2') and δ_C 124.7 (C-5'). The two P-C coupled doublets indicative of C-2 and C-1 could be observed at δ_C 71.6 (d, J_{PC} = 11.7 Hz, C-2), 67.2 (d, J_{PC} = 163.5 Hz, C-1), while C-3 was revealed as a singlet at δ_C 51.0 (C-3). The assignments in both the ^{1}H and ^{13}C NMR spectra of **33** corresponded with those in the spectra of **22**, apart from the isopropyl peaks, whose absence again confirmed that dealkylation had taken place. A single resonance was observed in the ^{31}P spectrum at δ_C 18.7. Electrospray mass spectroscopy gave further evidence for the existence of **33** by showing a protonated molecular ion peak of 250.0591 (M^+ + H), which corresponded closely with the expected value of 250.0587 (M^+ + H) required by the molecular formula $C_7H_{13}N_3O_5P$.

Finally, a dealkylation reaction of the Sonogashira cross coupled product, **24** (Fig 7.14)**,** was also attempted but as was the case in the reactions involving compounds **16a**, **16b** and **18,** the product could not be isolated. The production of the dealkylated variant of **24** was thus abandoned due to time constraints; however, as its structure is particularly interesting, future work should be performed so that the effect of the phosphonic acid group on the biological activity of the compound can be evaluated.

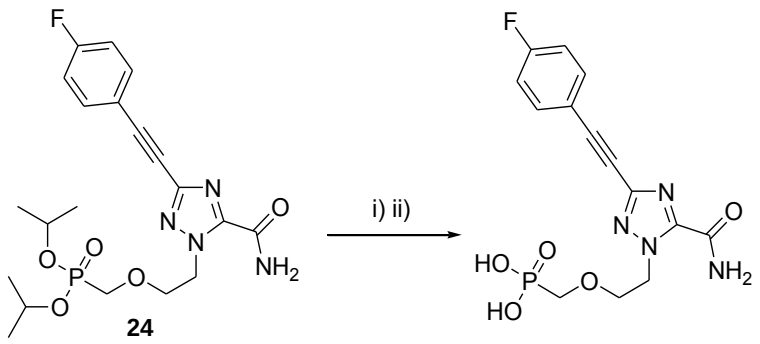


Fig. 7.14. Reagents and conditions: i) TMSBr, DCM, 0°C-rt; ii) MeOH.

Chapter 8

The synthesis of a bis(pivaloyloxymethyl) ANP prodrug

8.1 Introduction

Research carried out into the use of prodrug variants of acyclic nucleotide phosphonates as a means of increasing the *in vivo* activity of the drug were discussed in Chapter 2. Since ANPs are generally highly polar molecules, especially in their phosphonic acid form, they are often unable to effectively cross cell membranes to enter infected cells. Protection of the phosphonic acid moiety through the use of readily hydrolysable non-polar protecting groups has been well explored as a way of easily increasing an ANP drug's ability to permeate cell membranes, hence increasing its overall anti-viral activity. This Chapter will discuss our own exploratory efforts into the general synthesis of acylmethyloxy ANP prodrugs, through the synthesis of a bis(pivaloyloxymethyl) protected variant of the thymine derivative **26.**

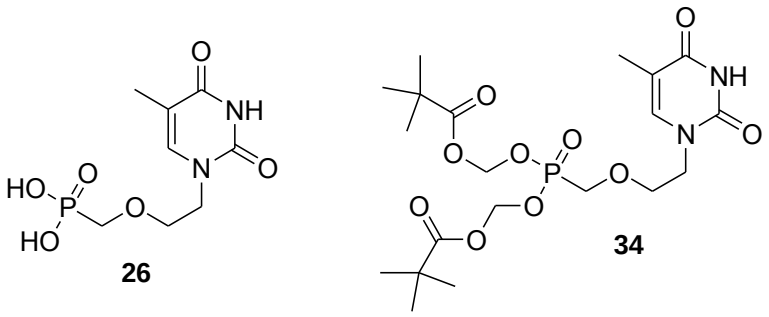


Fig. 8.1: The structures of **26** and its bis(pivaloyloxymethyl) derivative, **34.**

The alkylation of various phosphonic acids with a range of different non-polar protecting groups, through a range of different synthetic procedures has been reported in the literature.[54,56,57,61,92] Two dominant synthetic approaches exist: The first involves activation of the phosphonic acid to a phosphonic acid halide that can undergo attack by an appropriate nucleophilic protecting group (Fig 8.2 **A**). This has the disadvantage that it requires an extra synthetic step to form an air and water-sensitive acid halide, which can often be difficult and dangerous to work with; however, this approach is often faster and higher yielding due to the use of a reactive intermediate. The second approach uses an electrophilic protecting group, with the phosphonate conjugate base as a nucleophile generated from the acid with base (Fig 8.2 **B**). This reaction is much slower and lower yielding than the first approach, often requiring elevated reaction temperatures for long periods of time; however, it involves only one step and is much easier to set up and perform.

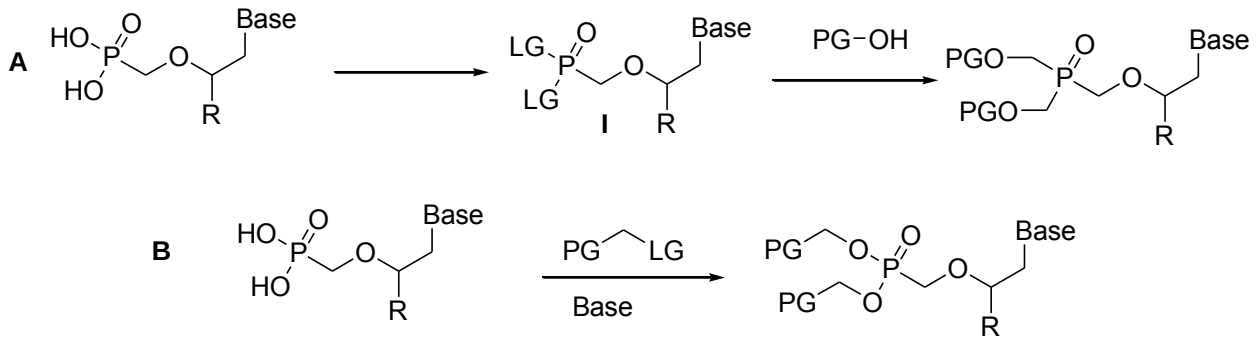


Fig. 8.2: The two dominant synthetic approaches used in the synthesis of ANP prodrugs.

Use of approach (**A**) has seen limited use in the reported literature. The reactive intermediate, **I**, can readily react with nucleophilic centres on the heterocyclic base ring, resulting in undesired by-products. During their investigation of bis(trifluoroethyl) phosphonic ester prodrugs, Krecmerová et al.[61] attempted to circumvent this problem by producing a tosylmethylphosphonate prodrug synthon (the (diisopropyl) tosylmethylphosphonate synthon **4** discussed in chapter 4), which could then be alkylated with an appropriate heterocycle-bearing hydroxyethyl synthon (Fig 8.3). These alkylation reactions (as discussed in Chapter 4) often require highly alkaline conditions which are not suitable for use with the relatively sensitive prodrug moieties.[61]

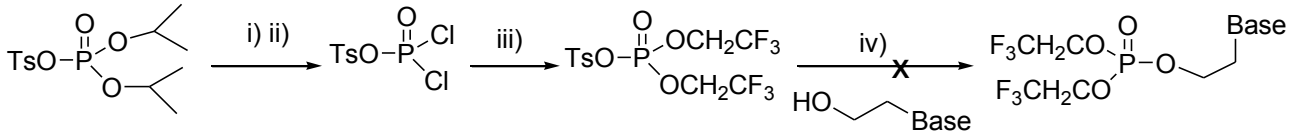


Fig. 8.3. Reagents and conditions: i)TMSBr, CH_3CN, Rt ; ii) $(COCl)_2$, DMF, DCM; iii) CF_3CH_2OH, NEt_3, DCM; iv) NaH, THF, reflux.[61]

The second approach (**B**) has seen much wider use as it avoids many of the issues associated with approach (**A**) such as cross-reactivity, harsh reaction conditions and extra steps. Tang et al.[54] used relatively mild reaction conditions to synthesise a range of alkoxycarbonyloxymethyl prodrugs of phosphonomethylethyl adenine (PMEA) (Fig 8.4). They reacted a range of chloromethyl alkyl carbonates with the phosphonic acid of PMEA in the presence of a mild organic amine base such as triethylamine, to afford bis-protected prodrug derivatives of PMEA (Fig 8.4 **1**). Krecmerová et al.[57] also reported a similar strategy whereby they reacted a tetrabutyl ammonium salt of a cyclic phosphonate monoester with pivaloylmethyl chloride (POMCl) in dioxane at 100°C to from a mono pivaloylmethyl cyclic phosphonate ester (Fig 8.4 **2**). Similarly Starrett et al.[94] discussed the synthesis of bis(pivaloylmethyl) phosphonate esters using pivaloylmethyl chloride with a hindered base such as *N*,*N*-dicylohexylmorpholinecarboxamidine or triethylamine in DMF

(Fig 8.4 **3**). These examples, which in most cases use fairly standard alkylation conditions, formed the premise for our own investigations into the synthesis of pivaloylmethyl prodrugs of ANPs.

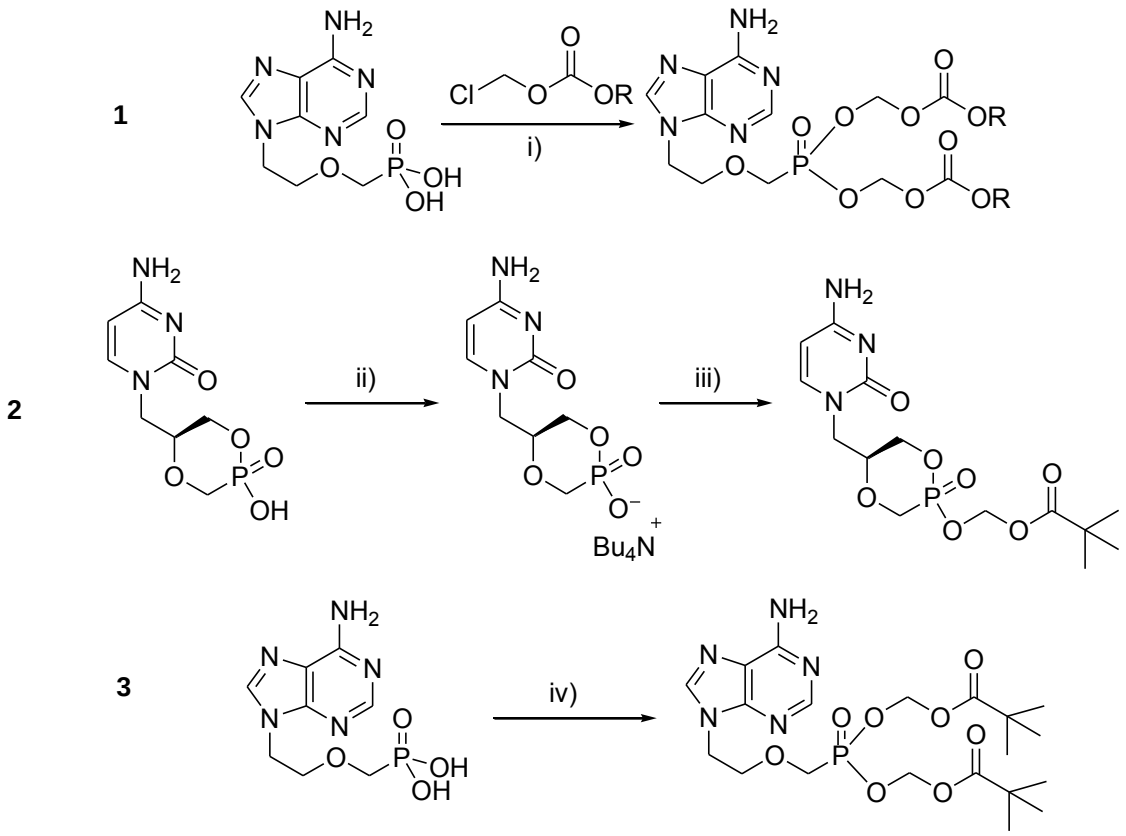


Fig. 8.4. Reagents and conditions: i) NEt_3, N-methyl-2-pyrrolidinone, 80-60°C; ii) tetrabutylammonium hydroxide; iii) POMCl, dioxane, 100°C; iv) POMCl, *N*,*N*-dicylohexylmorpholinecarboxamidine or triethylamine, DMF, 100°C.[54,57,94]

8.2 Synthesis of the ANP pivaloylmethyl prodrug, 34.

The aim of this part of the project was to gain an understanding into some of the challenges involved in the synthesis of acyclic nucleotide prodrugs. The thymine derivative, **26,** was chosen as the model to work on simply because it was an easy-to-handle solid and was in good supply as a product from earlier experiments. POMCl was chosen as the prodrug group as it was also readily accessible and its use is well published in the literature. Triethylamine was chosen as a base, since it was readily obtained and easy to use.

The reaction was initially performed in DMF at 100°C over a period of 48hrs. Some conversion of starting material was observed by tlc but the presence of DMF made the reaction difficult to monitor and greatly hindered the isolation of the product. The product was found to thermally degrade in the heat required to remove the DMF, even under vacuum, and if an extraction was attempted directly, the DMF had a tendency to pull the product into the aqueous layer. Acetonitrile was then investigated as an alternative solvent as it was both polar and relatively volatile, making it easy to work with. After heating all the

reactants together in acetonitrile at reflux temperature, the reaction mixture became homogeneous after 30 minutes. The reaction was allowed to stir for 48 hours at reflux, after which the only product, by tlc, was extracted and purified with chromatography to afford **34** as a white solid with a melting point range of 111-114°C, in a yield of 75% (Fig 8.5).

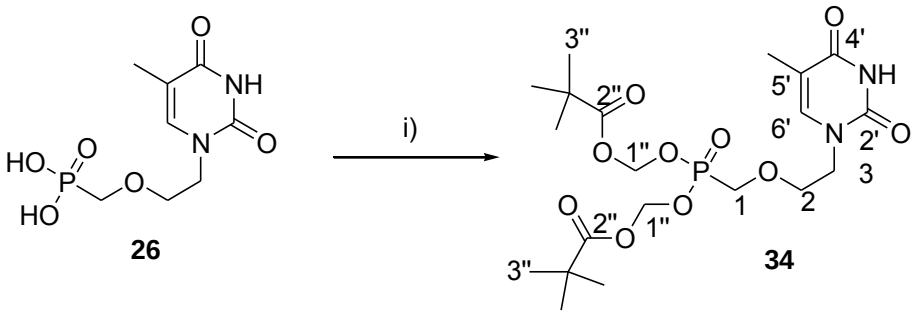


Fig. 8.5: Reagents and conditions: i) Pivaloylmethyl chloride (POMCl), NEt_3, CH_3CN, reflux, 48hrs, 75%

The IR spectrum of **34,** collected from an NaCl DCM solution cell, showed key absorptions at ν_{max} 3384 cm^{-1} (N-H), 2965 cm^{-1} (C-H), 1754 cm^{-1} (C=O), 1708 cm^{-1} (C=O), 1689 cm^{-1} (C=O), 1239 cm^{-1} (P=O) and 967 cm^{-1} (P-O). The ^{1}H NMR spectrum (Fig 8.6) revealed the PME side chain signals at δ_H 3.88 (t, *J* = 4.6 Hz, 2H, C-3), δ_H 3.85 (d, J_{HP} = 7.7 Hz, 2H, C-1) and δ_H 3.79 (t, *J* = 4.6 Hz, 2H, C-2), while the thymine N-H, methyl and H-6' could be seen at δ_H 9.09 (s, 1H, N-H), δ_H 1.89 (s, 3H, C<u>H</u>$_3$) and δ_H 7.13 (s, 1H, H-6') respectively. These assignments were made by analogy with the spectra of compounds **26** and **14**. The resonances diagnostic for the attachment of the pivaloylmethyl groups were the *t*-butyl singlet at δ_H 1.21 (s, 18H, H-3'') and the two doublet of doublets at δ_H 5.68 (dd, J_{gem} = 9.5, J_{HP} = 5.1 Hz, 2H, H-1") and δ_H 5.65 (dd, J_{gem} = 9.5, J_{HP} = 5.1 Hz, 2H, H-1") for the two pairs of diastereotopic hydrogens at H-1". Each of the H-1" resonances integrated for 2H and featured an H-H geminal coupling (J^2*= 9.5* Hz*)* as well as a smaller H-P coupling (J^3*= 5.1* Hz).

The ^{13}C NMR spectrum showed all the resonances for the PME side chain as well as the thymine ring at their expected chemical shifts. The ^{13}C NMR spectrum also revealed the new POM diastereotopic *tert*-butyl groups as a singlet at δ_C 26.9 (C-3''), the *t*-butyl quaternary carbon as a singlet at δ_C 38.8 (<u>C</u>$(CH_3)_3$) and the pivaloyl carbonyl carbon at δ_C 176.9 (C-2''), while the resonance associated with the C-1'' methylene carbon was observed as a P-C J^3 coupled doublet at δ_C 81.8 (d, J_{CP} = 6.1 Hz, C-1''), providing further evidence for the attachment of the POM groups. Mass spectroscopy showed a peak at 493.1946 (M^+ + H), which corresponded with the expected value of 493.1951 (M^+ + H) required for the molecular formula $C_{20}H_{34}N_2O_{10}P$.

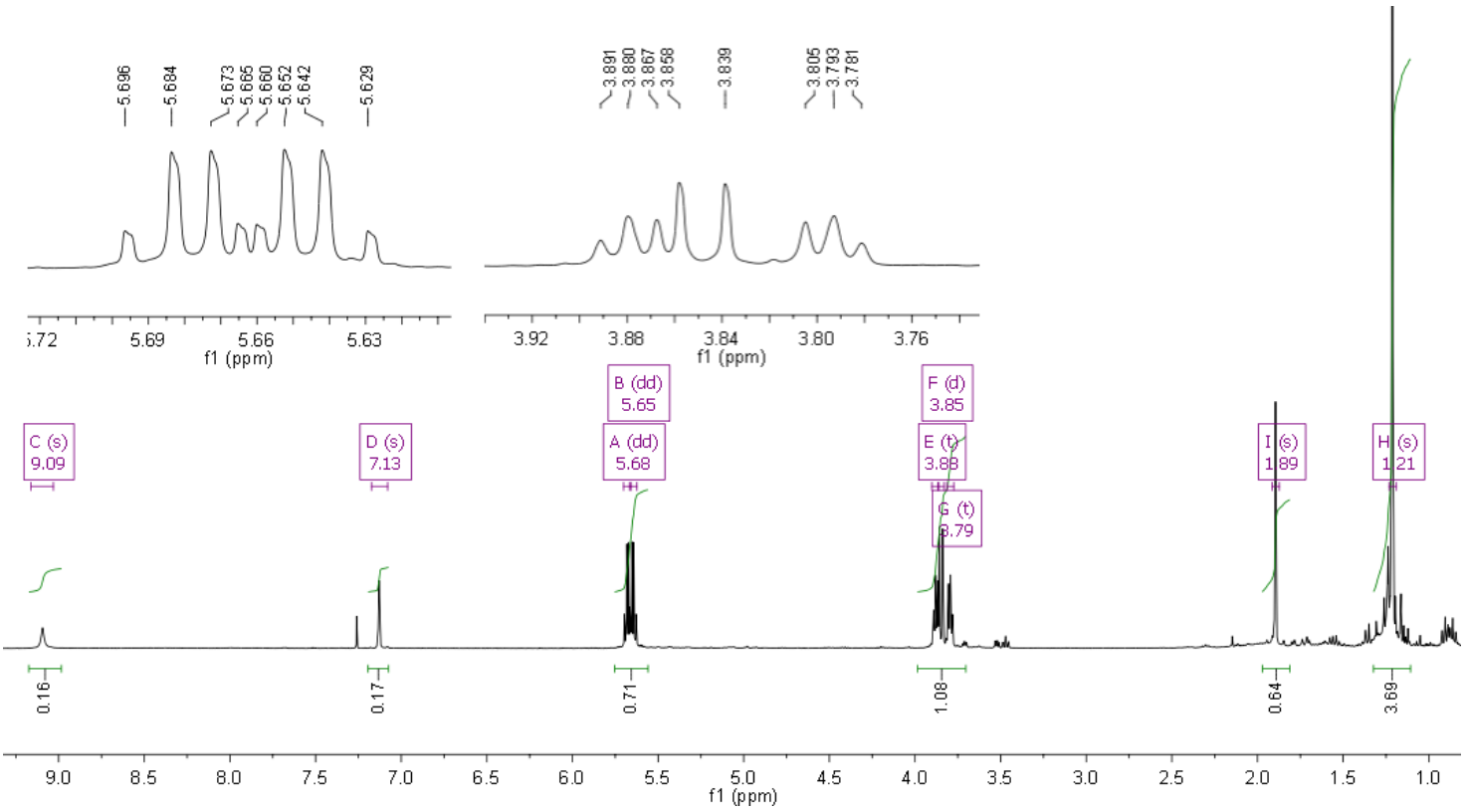


Fig. 8.6: The ^{1}H NMR spectrum of **34.**

With the synthesis of **34** successfully completed, the synthesis of further POM protected derivatives was attempted using the pyrazole-3-carboxamide and triazole-3-carboxamide derivatives **30** and **32**, which were chosen for their relevance to the project. The reactions with the two new starting materials were carried out under the same conditions as for **26** but they never became homogeneous and only a small amount of product formation was observed by tlc. The reactions were also attempted in DMF but this change had little effect. Finally, the use of caesium carbonate was also investigated in both DMF and CH_3CN, as an alternative base to triethylamine, but this too failed to increase the amount of product formed. Unfortunately, this meant that the synthesis of the two POM protected derivatives of **30** and **32** had to be put on hold due to thesis time constraints.

Chapter 9

Results and discussion of the biological evaluations

A selection of the compounds discussed above was sent to Prof Paolo La Colla at the Department of Virology, University of Cagliari in Italy for biological testing. Although the intracellular cleavage of the isopropyl ester has been shown to be slow, these derivatives were sent for testing regardless, in order ensure that no active compounds were missed in the screening process. The compounds shown below (Fig 9.1) were tested against a range of RNA and retroviral infections, including: HIV, bovine viral diarrhoea virus (BVDV), yellow fever virus (YFV), Coxsackie B-5 virus (CBV), respiratory syncytial virus (RSV) and vesicular stomatitis virus (VSV), as well as the dsDNA viruses: herpes simplex virus-1 (HSV-1) and vaccinia virus (VV). Compounds **13**, **14**, **19**, **20a**, **20b**, **25**, **26**, **30** and **31** were found to be non-cytotoxic and to posses no significant activity against the viruses tested (Table 9.1). Unfortunately at the time of printing this thesis, the biological test results for compounds **15**, **17, 18, 21, 22, 24, 27, 28** and **32-34** were still pending.

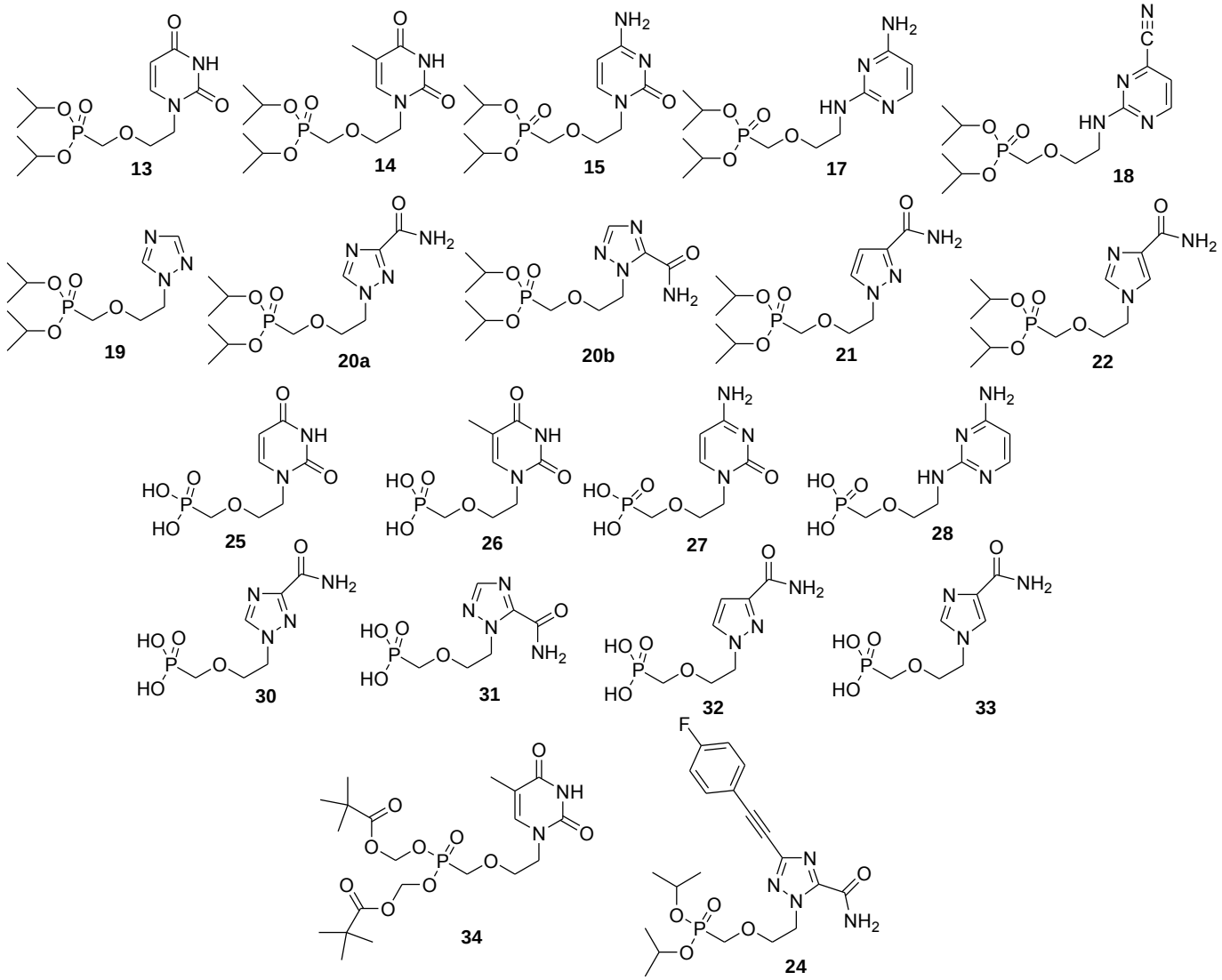


Fig. 9.1: A diagram showing the compounds sent for biological evaluation.

Table 9.1. Cytotoxicity and anti-viral activity of selected **ANPs**, and reference compounds, against representatives of ssRNA$^+$ (HIV-1, BVDV, YFV, CBV-5, Sb-1), ssRNA$^-$ (RSV, VSV), dsRNA (Reo-1) and dsDNA (VV, HSV-1) viruses.°

Compound	MT-4 CC_{50}[a]	HIV-1_{IIIB} EC_{50}[b]	MDBK CC_{50}[c]	BVDV EC_{50}[d]	BHK CC_{50}[e]	YFV EC_{50}[f]	Reo-1 EC_{50}[g]	Vero-76 CC_{50}[h]	CVB-5 EC_{50}[i]	Sb-1 EC_{50}[j]	RSV EC_{50}[k]	VSV EC_{50}[l]	VV EC_{50}[m]	HSV-1 EC_{50}[n]
20b	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100
13	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100
20a	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100
19	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100
25	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100
26	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100
14	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100
31	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100
30	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100
Reference Compounds														
Efavirenz	40	**0.0020**±0.0003 (20.000)	36	>36	10	>10	>10	20	>20	>20	>20	>20	>20	>20
2'-C-methyl-guanosine	>100	>100	>100	**2.0**±0.1 (>50)	>100	**1.8**±0.1 (>56)		>100	**55**	**50**	>100	>100	>100	>100
2'-C-methyl-cytidine	47	>47	>100	**3.2**±1.0 (>31)	>100	>100	**17**±1.0 (>6)	>100	**18**±1.5 (>6)	**7.3**±1.1 (>14)	>100	>100	>100	>100
6-Aza-uridine	0.3± 0.1	>0.3	>100	>100	>100	>100	>100	12.5	>12.5	>12.5	**1.4**±0.2 (9)	>12.5	>12.5	>12.5
Mycophenolic acid	<0.8	<0.8	2.8	>2.8	>100	>100	**20**	13	>13	>13	>13	>13	**1.7**±0.1 (8)	>13
Acycloguanosine	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	>100	**3.0**±0.1 (>33)

° Data represent mean values ± SD for three independent determinations. For values where SD is not shown, variation among duplicate samples was less than 15%.

[a] Compound concentration (µM) required to reduce the proliferation of mock-infected MT-4 cells by 50%, as determined by the MTT method.

[b] Compound concentration (µM required to achieve 50% protection of MT-4 cells from HIV-1 induced cytopathogenicity, as determined by the MTT method.

[c] Compound concentration (µM) required to reduce the viability of mock-infected MDBK cells by 50%, as determined by the MTT method.

[d] Compound concentration (µM) required to achieve 50% protection of MDBK cells from BVDV-induced cytopathogenicity, as determined by the MTT method.

[e] Compound concentration (µM) required to reduce the viability of mock-infected BHK cells by 50%, as determined by the MTT method.
[f] Compound concentration (µM) required to achieve 50% protection of BHK cells from YFV-induced cytopathogenicity, as determined by the MTT method.
[g] Compound concentration (µM) required to achieve 50% protection of BHK cells from Reo-1-induced cytopathogenicity, as determined by the MTT method.
[h] Compound concentration (µM) required to reduce the viability of mock-infected Vero-76 cells by 50%. as determined by the MTT method.
[i] Compound concentration (µM) required to reduce the plaque number of CVB-5 by 50% in Vero-76 monolayers.
[j] Compound concentration (µM) required to reduce the plaque number of Sb-1 by 50% in Vero-76 monolayers.
[k] Compound concentration (µM) required to reduce the plaque number of RSV by 50% in Vero-76 monolayers.
[l] Compound concentration (µM) required to reduce the plaque number of VSV by 50% in Vero-76 monolayers.
[m] Compound concentration (µM) required to reduce the plaque number of VV by 50% in Vero-76 monolayers.
[n] Compound concentration (µM) required to reduce the plaque number of HSV-1 by 50% in Vero-76 monolayers.
( ) Selectivity Index: CC_{50}/EC_{50}

Chapter 10

Conclusion and Future work

A range of acyclic nucleotide phosphonates were synthesised and characterised based on a diisopropyl phosphonomethylethyl synthon. Amongst these derivatives was a series of ribavirin-based analogues which may show interesting biological properties due to their ability to act as "universal bases" through their possession of multiple hydrogen-bonding profiles.[42] Acyclic nucleoside ribavirin-like analogues were previously explored by Beauchamp et al.,[46] but to our knowledge no investigations have thus far been conducted on acyclic nucleotide phosphonate analogues of ribavirin.

A Sonogashira palladium catalysed cross-coupling reaction was used to synthesise an arylethynyltriazole ANP derivative. Acyclic nucleoside arylethynyltriazole analogues have also been previously explored as potential anti-viral and anti-cancer agents by other authors,[71,73,74] however, the inclusion of the phosphonate moiety into such drug candidates seems to be, to the best of our knowledge, novel to this work. Future work in this area may include the synthesis of additional arylethynyltriazole and aryltriazole ANP derivatives through the use of palladium mediated cross-coupling reactions, in order to further probe the mechanism of action of such compounds and increase their potency if they are found to be biologically active.

Some of these products were then converted to their corresponding phosphonic acids through a TMSBr-mediated phosphonic ester hydrolysis. In one example, a bis(pivaloyloxymethyl) prodrug variant was produced in order to probe a general synthesis which could be used to produce future prodrug derivatives of other interesting ANP analogues. Any future work featuring the compounds discussed in this thesis, as well as other future ANP derivatives should involve the formation of their phosphonic acids in order to maximize the possibility of intracellular phosphorylation to an active di-phosphate species. The synthesis of future prodrug derivatives is also of high importance as the presence of the prodrug moiety aids the drug molecule to cross the cell membrane, thus allowing it to enter the cell and elicit its activity, indirectly increasing its potency.

Once the results of the biological activity studies have been released, conclusions and new hypotheses as to the structural activity relationships of the selected candidate molecules can be formed. Work may then begin on optimizing and improving the active structures, ultimately aiding in the discovery and development of new acyclic nucleotide drugs against an array of neglected and emerging RNA viral infections.

Part 3

Experimental

Experimental section

General instrumentation used

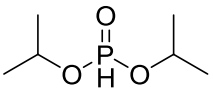
All starting materials were obtained from Sigma-Aldrich and all solvents were freshly distilled.

Reactions were monitored by TLC using aluminium-backed Merck silica-gel 60 F_{245} plates and spots were observed using UV light (254 nm) as well as iodine vapour, a 2.5% mixture of anisaldehyde in ethanol and sulphuric acid (1:1) or a solution of potassium permanganate, potassium carbonate and sodium hydroxide in water.

Nuclear Magnetic Resonance spectra were recorded on a Varian Mercury 300 MHz (75.5 MHz for ^{13}C, 121 MHz for ^{31}P), Bruker 400 MHz (100.6 MHz for ^{13}C, 162 MHz for ^{31}P) or a Varian Unity 400 MHz (100.6 MHz for ^{13}C, 162 MHz for ^{31}P) machine and were carried out in chloroform-d unless otherwise stated. Chemical shifts (δ) were recorded relative to residual chloroform (δ 7.26 in ^{1}H NMR and δ 77.16 in ^{13}C NMR), acetone (δ 2.05 in ^{1}H NMR and δ 29.84 in ^{13}C NMR), methanol (δ 3.31 in ^{1}H NMR and δ 49.00 in ^{13}C NMR) and DMSO (δ 2.50 in ^{1}H NMR and δ 39.52 in ^{13}C NMR). All chemical shifts (δ) are reported in ppm and *J* values are reported in Hz.

Mass spectra were recorded on a JEOL GCmatell in electron ionization mode. High resolution mass spectra were recorded on a VG70 SEQ micromass spectrometer electrospray positive mode. Infrared (IR) absorptions were measured on a Perkin Elmer Spectrum One FT-IR spectrometer in which peaks are reported in cm^{-1}. Melting points and decomposition temperatures were determined using a Reichert-Jung Thermovar hot-stage microscope and are uncorrected.

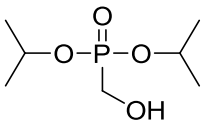
Diisopropyl hydrogen phosphonate[77] (1)



Isopropanol (5.26 ml, 68.8 mmol) was added dropwise to a stirring suspension of sodium hydride (60% mineral oil, 2.29 g, 57.30 mmol) in THF (40 ml) at 0°C and the resulting mixture was allowed to stir for 15 minutes. Phosphorus trichloride (2.0 ml, 22.9 mmol) dissolved in THF (10 ml) was added dropwise to the vigorously stirring solution. The reaction mixture was then allowed to stir at room temperature for 1,5 hours before being quenched with saturated ammonium chloride solution (10 ml) upon which the resulting mixture was transferred to a separating funnel containing ammonium chloride solution (40 ml). The product was then extracted with ethyl acetate (3 x 50 ml) and the pooled organic fractions were dried with magnesium sulphate and evaporated. The resulting residue was purified with column chromatography (ethyl acetate / hexane= 3:7) to afford **1** as a colourless oil (3.60 g, 95%).

δ_H (400 MHz, $CDCl_3$) 6.74 (d, J_{HP} = 687.2 Hz, 1H, H-P=O), 4.64 (m, 2H, CH), 1.27 (d, J = 6.4 Hz, 6H, $CH(CH_3)_2$), 1.26 (d, J = 6.4 Hz, 6H, $CH(CH_3)_2$); δ_p (162 MHz, $CDCl_3$) 4.3.

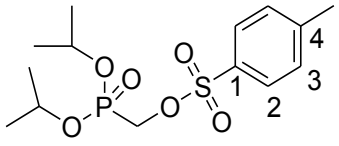
Diisopropyl hydroxymethyl phosphonate[77] (2)



A solution of paraformaldehyde (0.74 g, 24.8 mmol), potassium carbonate (0.27 g, 1.96 mmol) and diisopropyl hydrogen phosphonate **1** (3.26 g, 19.62 mmol) in Isopropanol (20 ml) was prepared. The mixture was then warmed to 60°C and allowed to stir for 2 hours. The mixture was then filtered through a pad of Celite, which was washed with ethyl acetate (20 ml). The solvent was then removed under reduced pressure to give **2** as a colourless oil (3.76 g, 98%), which was pure enough for use in the next step.

δ_H (400 MHz, $CDCl_3$) δ 4.69 (m, 2H, $CH(CH_3)_2$), 3.89 (br s, 1H, OH), 3.79 (d, J_{PH} = 6.0 Hz, 2H, P-CH_2), 1.30 (d, J = 6.0 Hz, 6H, $CH(CH_3)_2$), 1.28 (d, J = 6.0 Hz, 6H, $CH(CH_3)_2$); δ_P (162 MHz, $CDCl_3$) δ 22.6.

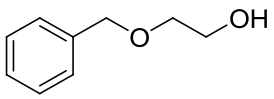
(Diisopropoxyphosphoryl)methyl 4-methylbenzenesulfonate[79] (3)



Diisopropyl hydroxymethyl phosphonate **2** (1.22 g, 6.20 mmol) was dissolved in DCM (10 ml) with triethylamine (1.04 ml, 7.44 mmol). A solution of *p*-toluenesulfonyl chloride (1.42 g, 7.44 mmol) and 4-Dimethylaminopyridine (0.038 g, 0.31 mmol) in DCM (10 ml) was prepared and added dropwise to the reaction mixture at 0°C. The reaction mixture was then allowed to warm to room temperature and left to stir for 2 hours, after which it was extracted with ethyl acetate (3 x 50 ml) from a saturated solution of ammonium chloride (30 ml). The pooled organic fractions were then dried with magnesium sulphate, evaporated under reduced pressure, and the resulting residue chromatographed (ethyl acetate: hexane = 1:1) to yield **3** as a clear oil (1.57 g, 72 %).

δ_H (300 MHz, $CDCl_3$) 7.76 (d, *J* = 8.2 Hz, 2H, H-2), 7.33 (d, *J* = 8.2 Hz, 2H, H-3), 4.69 (m, 2H, C<u>H</u>$(CH_3)_2$), 4.09 (d, J_{PH} = 10.1 Hz, 2H, P-C<u>H</u>$_2$), 2.42 (s, 3H, ArC<u>H</u>$_3$), 1.29 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.26 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_c (101 MHz, $CDCl_3$) 145.5 (C-1) , 132.0 (C-4), 130.0 (C-3) , 128.3 (C-4), 72.4 (d, J_{PC} = 6.6 Hz, <u>C</u>H$(CH_3)_2$), 62.09 (d, J_{PC} = 170.2 Hz, P-<u>C</u>H$_2$), 24.0 (d, J_{PC} = 4.9 Hz, CH(<u>C</u>H$_3)_2$), 23.9 (d, J_{PC}= 4.9 Hz, CH(<u>C</u>H$_3)_2$), 21.7(Ar-CH_3); δ_p (121 MHz, $CDCl_3$) 12.8.

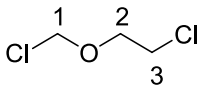
2-(Benzyloxy)ethanol[80] **(4)**



A suspension of sodium hydride (60% mineral oil, 2.800 g, 70.40 mmol) in THF (20 ml) was added dropwise to a stirred solution of ethylene glycol (10.7 ml, 192 mmol) in THF (20 ml) at 0°C and the reaction mixture was left to stir at room temperature for 15 minutes. Benzyl bromide (7.6 ml, 64 mmol) was then slowly added to the reaction mixture which was allowed to reflux for 12 hours. Saturated ammonium chloride solution (20 ml) was then added to the resulting mixture, and the organic material was extracted with ethyl acetate (3 x 80 ml), which was washed once with distilled water (50 ml). The combined organic fractions were then dried with magnesium sulphate, the solvent removed under reduced pressure and the residue then purified by column chromatography (ethyl acetate/ Hexane= 1: 1) to afford **4** as a colourless oil (7.52 g, 77 %).

δ_H (400 MHz, $CDCl_3$) 7.35 (m, 5H, Ph), 4.56 (s, 2H, Ph<u>C</u>H$_2$), 3.74 (t, *J* = 6.4, 6.4 Hz, 2H, CH_2), 3.58 (t, *J* = 6.4, 6.4 Hz, 2H, CH_2), 2.97 (s, 1H, OH); δ_c (75 MHz, $CDCl_3$) 138.0 (Ph), 128.5 (Ph), 127.8 (Ph), 127.8 (Ph), 73.3, 71.5, 61.8.

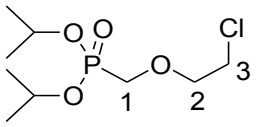
Chloro-2-(chloromethoxy)ethane[82,83] **(5)**



HCl gas, generated from the dropwise addition of concentrated sulphuric acid (98 %, 50 ml) to solid sodium chloride (10 g) in a sealed two-necked flask, was bubbled through a suspension of paraformaldehyde (8.740 g, 291.0 mmol) and 2-chloroethanol (15.0 ml, 224 mmol) in DCM (30 ml) for 6 hours at 0°C. The reaction was then stored at 4°C in a refrigerator for 12 hours after which HCl gas was then bubbled through the mixture for a further 8 hours stored again at 4°C for 12 hours. The reaction mixture was then dried using $CaCl_2$, filtered through a pad of Celite, and solvent was removed under reduced pressure to yield the crude product (9.49 g). The ^{1}H NMR spectrum of the crude material revealed a starting material to product conversion of 75%. The product was then purified by distillation (bp 64°C at 20 mm Hg) to yield **5** as a colourless liquid (17.7 g, 61 %).

δ_H (300 MHz, $CDCl_3$) 5.52 (s, 2H, H-1), 3.95 (t, *J* = 5.6 Hz, 2H, H-2), 3.68 (t, *J* = 5.6 Hz, 2H, H-3); δ_c (101 MHz, $CDCl_3$) δ 82.5 (C-1), 70.4 (C-2), 42.0 (C-3).

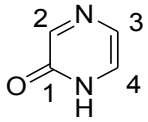
Diisopropyl (2-chloroethoxy)methylphosphonate[81] **(6)**



Triisopropyl phosphite (29.4 ml, 119 mmol) was heated in a round-bottomed flask fitted with a Vigreux column, still-head and condenser. 1-Chloro-2-(chloromethoxy) ethane **5** (18.43 g, 143.0 mmol) was then added dropwise and the resulting 2-chloropropane was distilled off at atmospheric pressure. The reaction was then allowed to stir at 100°C for a further 4 hours, after which it was distilled under high vacuum (bp 120°C, 0.1 mm Hg); [lit. 120^{o}C/0.1 mm Hg][81] to yield **6** as a colourless liquid (25.52 g, 83 %).

δ_H (300 MHz, $CDCl_3$) 4.76 (dhept, $J_{HP,\ HH}$ = 7.7, 6.3 Hz, 2H, C<u>H</u>$(CH_3)_2$), 3.85 (t, *J* = 5.7 Hz, 2H, H-2), 3.79 (d, J_{PH} = 8.4 Hz, 2H, H-1), 3.63 (t, *J* = 5.7 Hz, 2H, H-3), 1.34 (2 x d, *J* = 6.3 Hz, 2 x 6H, 2 x CH(C<u>H</u>$_3)_2$); δ_c (101 MHz, $CDCl_3$) 73.2 (d, J_{PC} = 10.6 Hz, C-2), 71.4 (d, J_{PC} = 6.7 Hz, <u>C</u>H$(CH_3)_2$), 66.3 (d, J_{PC} = 167.6 Hz, C-1), 42.5 (C-3), 24.3 (d, J_{PC} = 3.8 Hz, CH(<u>C</u>H$_3)_2$), 24.1 (d, J_{PC} = 4.6 Hz, CH(<u>C</u>H$_3)_2$); δ_p (162 MHz, $CDCl_3$) 18.5.

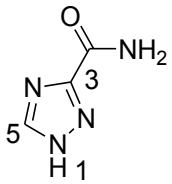
Pyrazin-2-ol[85,86] **(7)**



A solution of sodium hydroxide (10M, 3.40 ml, 33.7 mmol) and methanol (10 ml) was prepared and cooled to -10°C using a methanol and ice bath, after which glycinamide (1.00 g, 13.5 mmol) in methanol (10 ml) was added to the solution. Glyoxal (40% in water, 2.94 g, 20.3 mmol) was added slowly to the solution, with stirring, while the temperature of the reaction was maintained at -10°C. The reaction mixture was allowed to stir at room temperature for 3 hours, after which an excess of acetic acid (10 ml) was added. The mixture was then evaporated to dryness under reduced pressure and the resulting solid residue was then exhaustively extracted with boiling acetone (4 x 100 ml). The acetone was removed under reduced pressure to afford **7** as a dark brown powder which was found to be pure enough for use in the next step. (0.50 g, 39 %)

Mp: 190-191°C [lit Mp: 187°C]; IR (KBr): ν_{max} 3435 (O-H), 3085 (C-H), 1650 (C=O), 1255 (C-N); δ_H (400 MHz, MeOD) 8.07 (d, *J* = 1.4 Hz, 1H, H-2), 7.44 (d, *J* = 4.1 Hz, 1H, H-4), 7.38 (dd, *J* = 4.1, 1.4 Hz, 1H, H-3); δ_C (101 MHz, CD_3OD) 158.8 (C-1), 149.6 (C-4), 128.6 (C-2), 125.9 (C-3).

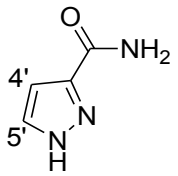
1*H*-1,2,4-Triazole-3-carboxamide[88] **(8)**



1*H*-1,2,4-triazole-3-carboxylic acid (500.0 mg, 4.42 mmol) was suspended in thionyl chloride (2.9 ml, 39.8 mmol) and a catalytic amount of dimethylformamide (0.03 ml, 0.4 mmol, 10 mol %) was added. The reaction mixture was then refluxed under a nitrogen atmosphere for 3 hours to give a clear solution after which all the thionyl chloride was removed under reduced pressure to yield the reaction intermediate as a white powder. This was then warmed to 50°C in 25% aqueous ammonia (5 ml), after which the majority of the ammonia and water were removed under vacuum and the residue triturated with acetone, to initiate the precipitation of the final product (491 mg, 99 %), which was found to be pure enough for use in the next step. The final amide was recrystallised from water and methanol (1: 10) for analysis.

Mp: 205°C (dec); IR (KBr): ν_{max} 3327 (N-H), 3155 (N-H),2926 (C-H), 1715 (C=O), 1699 (C=O); δ_H (400 MHz, d_6-DMSO) 8.38 (s, 1H, H-5), 7.85 (s, 1H, NH_2), 7.61 (s, 1H, NH_2); δ_c (101 MHz, d_6-DMSO) 160.3 (C=O), 147.7 (C-3), 146.8 (C-5); HRMS (ES): *m/z* found 113.0448 (M^+ + H), $C_3H_5N_4O$ requires 113.0463 (M^+ + H).

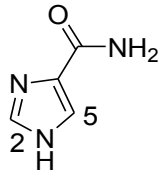
1*H*-Pyrazole-3-carboxamide (9)



1*H*-Pyrazole-3-carboxylic acid (300.0 mg, 2.86 mmol) was suspended in a solution of methanol (5 ml) and concentrated sulphuric acid (0.28 ml, 5.35 mmol) and the mixture refluxed for 12 hours. The methanol was then removed and saturated sodium bicarbonate (ca. 10 ml) was added to bring the pH of the solution to above 7. The aqueous solution was extracted with ethyl acetate (3 x 30ml) and the pooled organic fractions were collected and dried with magnesium sulphate after which the solvent was removed under reduced pressure to afford pyrazole-3-carboxylic acid methyl ester. The latter was then warmed at 40°C in the presence of aqueous ammonia (5 ml) for 6 hours after which the excess ammonia and water were evaporated off under reduced pressure to afford **9** (256 mg, 86 %) as an off-white amorphous solid pure enough for use in the next step.

Mp: 122-128°C; IR (KBr): ν_{max} 3405 (N-H), 3266 (N-H), 2961 (C-H), 1699 (C=O), 1668 (C=O); δ_H (400 MHz, d_6-DMSO) 7.71 (d, *J* = 2.0 Hz, 1H, H-5'), 7.54 (s, 1H, N-H), 7.19 (s, 1H, N-H), 6.67 (d, *J* = 2.0 Hz, 1H, H-4'); δ_c (101 MHz, d_6-DMSO) 163.0 (C=O), 132.0 (C-5'), 107.2 (C-3'), 105.3 (C-4'); HRMS (ES): *m/z* found 112.0494 (M^+ + H), $C_4H_6N_3O$ requires 112.0511 (M^+ + H).

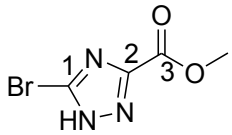
1*H*-Imidazole-4-carboxamide (10)



1*H*-Imidazol-3-carboxylic acid (300.0 mg, 2.86 mmol) was suspended in a solution of methanol (5 ml) and concentrated sulphuric acid (0.28 ml, 5.35 mmol) and the mixture refluxed for 12 hours. The methanol was then removed and saturated sodium bicarbonate (ca. 10 ml) was added to bring the pH of the solution above 7. The aqueous solution was extracted with ethyl acetate and the pooled organic fractions were collected and dried with magnesium sulphate after which the solvent was removed under reduced pressure to afford the imidazole-3-carboxylic acid methyl ester as an intermediate. The methyl ester was then warmed in the presence of aqueous ammonia (5 ml) for 6 hours after which the excess ammonia and water were removed under reduced pressure to afford the product 1*H*-imidazol-3-carboxamide (290 mg, 98 %) as an off-white amorphous solid pure enough for use in the next step.

Mp: 184-189; IR (KBr): ν_{max} 3397 (N-H), 3172 (N-H), 2676 (C-H), 1655 (C=O), 1400; δ_H (400 MHz, d_6-DMSO) 7.70 (s, 1H, H-2), 7.61 (s, 1H, H-5), 7.39 (brs, 1H, N-H), 7.07 (brs, 1H, N-H); δ_c (101 MHz, d_6-DMSO) 164.3 (C=O), 137.9 (C-2), 136.4 (C-5), 121.6 (C-4); HRMS (ES): *m/z* found 112.0492 (M^+ + H), $C_4H_6N_3O$ requires 112.0511 (M^+ + H).

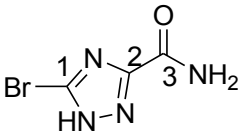
Methyl 5-bromo-1*H*-1,2,4-triazole-3-carboxylate[73] (11)



Methyl 5-amino-1*H*-1,2,4-triazole-3-carboxylate (500.0 mg, 3.52 mmol) was suspended in an aqueous solution of sulphuric acid (0.2 ml in 5 ml of water) at 0°C after which an aqueous solution of sodium nitrite (364.0 mg, 5.280 mmol) was add dropwise with constant stirring. After 30 minutes at 0°C an aqueous solution of potassium bromide (837 mg, 7.04 mmol) and copper iodide (33.5 mg, 0.176 mmol) was then added to the solution. The reaction solution was left to stir at room temperature for 3 hours after which it was extracted with ethyl acetate (4 x 30 ml). The organic fractions were then pooled, dried with magnesium sulphate and evaporated under reduced pressure. The remaining residue was then purified with column chromatography (MeOH: DCM = 4 : 96) to yield **11** as an off white solid, (541 mg; 75 %).

Mp: 134-137 °C; IR (KBr): ν_{max} 1741 (C=O), 1708, 1481, 1236; δ_H (400 MHz, d_6-DMSO) 3.89 (s, 3H, C$\underline{H}_3$); δ_c (101 MHz, d_6-DMSO) 157.4 (C=O), 52.7 ($\underline{C}H_3$); HRMS (EI): *m/z* found 204.9483 (M^+), $C_4H_4BrN_3O_2$ requires 204.9487 (M^+).

5-Bromo-1*H*-1,2,4-triazole-3-carboxamide (12)



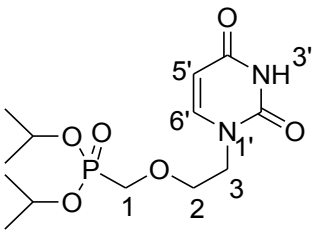
Methyl 5-bromo-1*H*-1,2,4-triazole-3-carboxylate (125 mg, 0.607 mmol) was suspended in a solution of methanol (2 ml) and aqueous ammonia (25 %, 3 ml) and stirred for 12 hours at 40°C. The solvent was then removed by gentle warming under reduced pressure to yield **12** as an off white solid pure enough for use in the next step. (113 mg, 98 %)

Mp: 175-180°C; IR (KBr): ν_{max} 3374 (N-H), 3250 (N-H), 2924, 1686 (C=O), 1610, 1441, 1291; δ_H (400 MHz, d_6-DMSO) 8.25 (brs, 1H, N-H), 7.96 (brs, 1H, N-H); δ_c (101 MHz, d_6-DMSO) 157.4 (C=O), 150.9 (C-1), 150.0 (C-2); HRMS (ES): *m/z* found 190.9568 (M^+ + H), $C_3H_4BrN_4O$ requires 190.9568 (M^+ + H).

General procedure for the alkylation of a nucleotide heterocyclic base with diisopropyl (2-chloroethoxy) methylphosphonate.[81] (GP1)

The heterocyclic base and caesium carbonate were suspended in DMF (2-3ml) and diisopropyl(2-chloroethoxy)methylphosphonate **6** was then added to the mixture. The mixture was allowed to stir for 12 hours at 100 °C. The DMF was then removed under high vacuum and the remaining residue dissolved in a saturated sodium bicarbonate solution (20 ml) and extracted with a methanol/ chloroform mixture (1:3, 3 x 20 ml). The organic phase was dried using magnesium sulphate and the solvent removed to give a residue which was chromatographed to afford the final purified product.

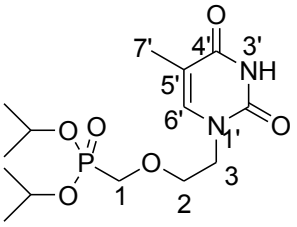
Diisopropyl 2-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)ethoxymethylphosphonate[95] (13)



According to GP1: Uracil (78.0 mg, 0.67 mmol), caesium carbonate (113 mg, 0.35 mmol) and diisopropyl(2-chloroethoxy)methylphosphonate **6** (200 mg, 0.77 mmol) in DMF. Chromatography (MeOH : DCM = 3 : 97) afforded **13** as a white solid (105 mg, 41%).

Mp: 66-68°C; IR (CH_2Cl_2): ν_{max} 3685, 3384 (N-H), 2977 (C-H), 1690 (C=O), 1269 (P=O), 994 (P-O); δ_H (400 MHz, d_6-acetone) 10.13 (s, 1H, H-3'), 7.55 (d, *J* = 7.9 Hz, 1H, H-6'), 5.53 (d, *J* = 7.9 Hz, 1H, H-5'), 4.68 (m, 2H, C<u>H</u>$(CH_3)_2$), 3.96 (t, *J* = 5.0 Hz, 2H, H-3), 3.84 (t, *J* = 5.0 Hz, 2H, H-2), 3.79 (d, J_{PH} = 8.2 Hz, 2H, H-1), 1.29 (d, *J* = 6.0 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.26 (d, *J* = 6.0 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_c (101 MHz, d_6-acetone) 164.4 (C-2'), 151.9 (C-4'), 146.9 (C-6'), 101.5 (C-5'), 71.4 (d, J_{PC} = 10.4 Hz, C-2), 71.2 (d, J_{PC} = 6.4 Hz, <u>C</u>H$(CH_3)_2$), 66.4 (d, J_{PC} = 166.2 Hz, C-1), 48.5 (C-3), 24.3 (d, J_{PC} = 4.2 Hz, CH(<u>C</u>$H_3)_2$), 24.3 (d, J_{PC} = 4.2 Hz, CH(<u>C</u>$H_3)_2$); δ_p (162 MHz, d_6-acetone) 18.5; HRMS (ES): *m/z* found 335.1372 (M^+ + H), $C_{13}H_{24}N_2O_6P$ requires 335.1366 (M^+ + H).

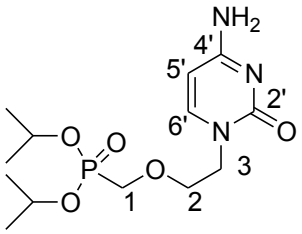
Diisopropyl 2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)ethoxymethyl phosphonate (14)



According to GP1: Thymine (195 mg, 1.55 mmol), caesium carbonate (252 mg, 0.77 mmol) and diisopropyl(2-chloroethoxy)methylphosphonate **6** (200 mg, 0.77 mmol) were reacted in DMF. Chromatography (MeOH : DCM = 2 : 98) afforded **14** product as a white solid (120.0 mg, 45 %).

Mp: 76-79°C; IR (CH_2Cl_2): ν_{max} 3387 (N-H), 2982 (C-H), 1709 (C=O), 1688 (C=O), 1237 (P=O), 993 (P-O); δ_H (400 MHz, d_6-acetone) 9.88 (s, 1H, H-3'), 7.40 (s, 1H, H-6'), 4.67 (m, 2H, C<u>H</u>$(CH_3)_2$), 3.92 (t, *J* = 4.9 Hz, 2H, H-3), 3.82 (t, *J* = 4.9 Hz, 2H, H-2), 3.77 (d, J_{PH} = 8.3 Hz, 2H, H-1), 1.82 (s, 3H, H-7'), 1.28 (d, *J* = 6.8 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.26 (d, *J* = 6.8 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_c (101 MHz, d_6-acetone) 165.2 (C-2'), 152.0 (C-4'), 143.0 (C-6'), 109.5 (C-5'), 71.5 (d, J_{PC} = 6.1 Hz, <u>C</u>H$(CH_3)_2$), 71.4 (d, J_{PC} = 10.1 Hz, C-2), 66.2 (d, J_{PC} = 166.5 Hz, C-1), 48.3 (C-3), 24.3 (d, J_{PC} = 4.2 Hz, CH(<u>C</u>$H_3)_2$), 24.1 (d, J_{PC} = 4.2 Hz, CH(<u>C</u>$H_3)_2$), 12.3 (C-7'); δ_p (162 MHz, d_6-acetone) 18.4; HRMS (ES): *m/z* found 349.1523 (M^+ + H), $C_{14}H_{26}N_2O_6P$ requires 349.1523 (M^+ + H).

Diisopropyl 2-(4-amino-2-oxo-pyrimidin-1(2H)-yl)ethoxymethylphosphonate.[81] (15)



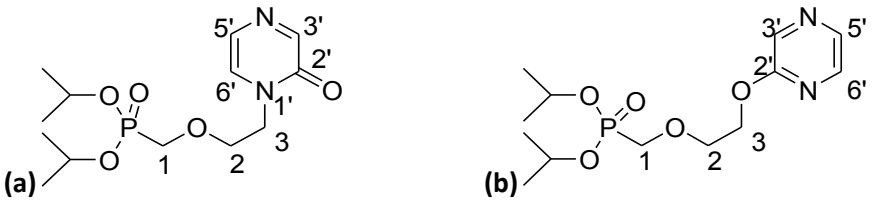
According to GP1: Cytosine (170 mg, 1.53 mmol), caesium carbonate (249 mg, 0.77 mmol) and diisopropyl(2-chloroethoxy)methylphosphonate **6** (495 mg, 1.91 mmol) were reacted in DMF. Chromatography (MeOH : DCM = 4 : 96) afforded **15** as a yellow solid (167 mg, 66 %).

Mp: 143-146°C [lit Mp: 151°C]; IR (CH_2Cl_2): ν_{max} 3685, 3525 (N-H), 3409 (N-H), 2983 (C-H), 1655 (C=O), 1249 (P=O), 993 (P-O); δ_H (400 MHz, d_6-acetone) 7.52 (d, *J* = 7.2 Hz, 1H, H-6'), 6.49 (s, 2H, NH_2), 5.75 (d, *J* = 7.2 Hz, 1H, H-5'), 4.66 (m, 2H, C<u>H</u>$(CH_3)_2$), 3.91 (t, *J* = 5.1 Hz, 2H, H-3), 3.80 (t, *J* = 5.1 Hz, 2H, H-2), 3.75 (d, J_{PH} = 8.4 Hz, 2H, H-1), 1.28 (d, *J* = 6.3 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.25 (d, *J* = 6.3 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_C (101 MHz, d_6-acetone) 166.5 (C-2'), 146.8 (C-6'), 145.7 (C-4'), 92.4 (C-5'), 70.7 (d, J_{PC} = 11.0 Hz, C-2), 70.1 (d, J_{PC} = 6.4 Hz, <u>C</u>H$(CH_3)_2$), 65.4 (d, J_{PC} = 166.6 Hz, C-1), 48.8 (C-3), 23.3 (d, J_{PC} = 4.2 Hz, CH(<u>C</u>$H_3)_2$), 23.2 (d, J_{PC} = 4.2 Hz, CH(<u>C</u>$H_3)_2$); δ_P

(121 MHz, d_6-acetone) 24.1; HRMS (ES): *m/z* found 334.1531 (M^+ + H), $C_{13}H_{25}N_3O_5P$ requires 334.1532 (M^+ + H).

Diisopropyl (2-(2-oxopyrazin-1(2H)-yl)ethoxy)methylphosphonate (16a)

Diisopropyl (2-(pyrazin-2-yloxy)ethoxy)methylphosphonate (16b)

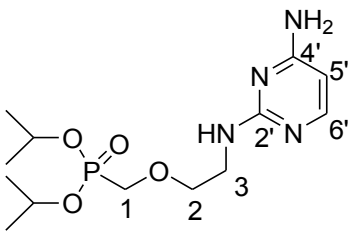


According to GP1: 2-Hydroxypyrazine (100 mg, 1.04 mmol), caesium carbonate (339 mg, 1.04 mmol) and diisopropyl (2-chloroethoxy)methylphosphonate **6** (296.0 mg, 1.15 mmol) in DMF. Chromatography (MeOH : DCM = 2 : 98) gave **16a** (71 mg, 21 %) and **16b** (155 mg, 47 %) as clear yellow oils.

16a) IR (CH_2Cl_2): ν_{max} 3682, 2978 (C-H), 1662 (C=O), 1237 (P=O), 995 (P-O); δ_H (400 MHz, $CDCl_3$) 8.12 (d, *J* = 1.1 Hz, 1H, H-3'), 7.27 (d, *J* = 4.4 Hz, 1H, H-6'), 7.25 (dd, *J* = 4.4, 1.1 Hz, 1H, H-5'), 4.68 (m, 2H, C<u>H</u>$(CH_3)_2$), 4.09 (t, *J* = 4.8 Hz, 2H, H-3), 3.86 (t, *J* = 4.8 Hz, 2H, H-2), 3.69 (d, J_{PH} = 8.4 Hz, 2H, H-1), 1.29 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.26 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_c (101 MHz, $CDCl_3$) 156.3 (C-2'), 149.4 (C-3'), 130.2 (C-6'), 123.5 (C-5'), 71.3 (d, J_{PC} = 6.7 Hz, <u>C</u>H$(CH_3)_2$), 70.3 (d, J_{PC}= 11.0 Hz, C-2), 66.2 (d, J_{PC} = 168.5 Hz, C-1), 49.1 (C-3), 24.2 (d, J_{PC} = 4.0 Hz, CH(<u>C</u>$H_3)_2$), 24.1 (d, J_{PC} = 4.0 Hz, CH(<u>C</u>$H_3)_2$); δ_p (162 MHz, $CDCl_3$) 18.43; HRMS (ES): *m/z* found 319.1417 (M^+ + H), $C_{13}H_{24}N_2O_5P$ requires 319.1419 (M^+ + H).

16b) IR (CH_2Cl_2): ν_{max} 3682, 3427, 2978 (C-H), 1248 (P=O), 1007 (P-O); δ_H (400 MHz, d_6-acetone) 8.22 (d, *J* = 1.4 Hz, 1H, H-3'), 8.15 (d, *J* = 2.8 Hz, 1H, H-6'), 8.13 (dd, *J* = 2.8, 1.4 Hz, 1H, H-5'), 4.70 (m, 2H, C<u>H</u>$(CH_3)_2$), 4.52 (t, *J* = 4.8 Hz, 2H, H-3), 3.97 (t, *J* = 4.8 Hz, 2H, H-2), 3.81 (d, J_{PH} = 8.2 Hz, 2H, H-1), 1.29 (d, *J* = 4.3 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.27 (d, *J* = 4.2 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_c (101 MHz, $CDCl_3$) 160.1 (C-2'), 140.5 (C-3'), 136.9 (C-6'), 136.2 (C-5'), 71.3 (d, J_{PC} = 7.1 Hz, C-2), 71.2 (d, J_{PC} = 4.4 Hz, <u>C</u>H$(CH_3)_2$)), 66.3 (d, J_{PC} = 167.3 Hz, C-1), 65.2 (C-3), 24.2 (d, J_{PC} = 4.4 Hz, CH(<u>C</u>$H_3)_2$), 24.1 (d, *J* = 4.4 Hz, CH(<u>C</u>$H_3)_2$); δ_p (162 MHz, $CDCl_3$) 18.9; HRMS (ES): *m/z* found 319.1417 (M^+ + H), $C_{13}H_{24}N_2O_5P$ requires 319.1419 (M^+ + H).

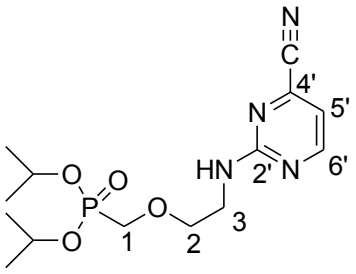
Diisopropyl (2-(4-aminopyrimidin-2-ylamino)ethoxy)methylphosphonate (17)



According to GP1: 2, 4-Diaminopyrimidine (94.0 mg, 0.85 mmol), caesium carbonate (126 mg, 0.39 mmol) and diisopropyl (2-chloroethoxy)methylphosphonate **6** (100 mg, 0.39 mmol) were reacted in DMF. Chromatography (MeOH : DCM = 6 : 94) gave **17** as a yellow oil (51 mg, 40 %).

IR (CH_2Cl_2): ν_{max} 3685, 3522 (N-H), 3409 (N-H), 2989 (C-H), 1599, 1236 (P=O), 994 (P-O); δ_H (400 MHz, d_6-acetone) 7.66 (d, *J* = 5.9 Hz, 1H, H-6'), 6.40 (s, 1H, NH), 5.84 (d, *J* = 5.9 Hz, 1H, H-5'), 5.56 (s, 2H, NH_2), 4.69 (m, 2H, $CH(CH_3)_2$), 3.77 (d, J_{PH} = 8.0 Hz, 2H, H-1), 3.74 (t, *J* = 5.5 Hz, 2H, H-2), 3.52 (q, *J* = 5.5 Hz, 2H, H-3), 1.29 (d, *J* = 6.0 Hz, 6H, $CH(CH_3)_2$), 1.27 (d, *J* = 6.0 Hz, 6H, $CH(CH_3)_2$); δ_c (101 MHz, d_6-acetone) 164.3 (C-2'), 164.1 (C-4'), 155.6 (C-6'), 96.6 (C-5'), 72.5 (d, J_{PC} = 9.9 Hz, C-2), 71.3 (d, J_{PC} = 6.4 Hz, $CH(CH_3)_2$), 66.4 (d, J_{PC} = 166.6 Hz, C-1), 40.9 (C-3), 24.4 (d, J_{PC} = 4.0 Hz, $CH(CH_3)_2$), 24.3 (d, J_{PC} = 4.0 Hz, $CH(CH_3)_2$); δ_p (162 MHz, d_6-acetone) δ 19.2; HRMS (ES): *m/z* found 333.1697 (M^+ + H), $C_{13}H_{26}N_4O_4P$ requires 333.1686 (M^+ + H).

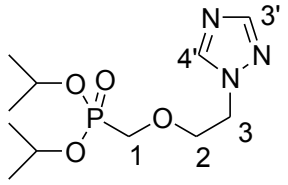
Diisopropyl (2-(4-carbonitrylpyrimidin-2-ylamino)ethoxy)methylphosphonate (18)



According to GP1: 2-Aminopyrimidine-4-carbonitrile (102 mg, 0.85 mmol), caesium carbonate (126 mg, 0.39 mmol) and diisopropyl(2-chloroethoxy)methylphosphonate **6** (100.0 mg, 0.39 mmol) were reacted in DMF. Chromatography (hexane: ethyl acetate = 2 : 8) gave **18** as a yellow oil (84 mg, 64 %).

IR (CH_2Cl_2): ν_{max} 3435 (N-H), 2883 (C-H), 2246 (C≡N), 1580, 1235 (P=O), 992 (P-O); δ_H (400 MHz, $CDCl_3$) 8.41 (d, *J* = 5.0 Hz, 1H, H-6'), 6.80 (d, *J* = 5.0 Hz, 1H, H-5'), 6.07 (s, 1H, NH), 4.76 (m, 2H, $CH(CH_3)_2$), 3.81 – 3.71 (m, 4H, H-1, H-2), 3.63 (q, *J* = 5.3 Hz, 2H, H-3), 1.34 (d, *J* = 5.4 Hz, 6H, $CH(CH_3)_2$), 1.31 (d, *J* = 5.4 Hz, 6H, $CH(CH_3)_2$); δ_c (101 MHz, CD_3OD) δ 163.9 (C-2'), 161.6 (C-6'), 159.6 (C-4'), 117.1 (C≡N), 113.9 (C-7'), 73.2 (d, J_{PC} = 6.7 Hz, $CH(CH_3)_2$), 72.8 (d, J_{PC} = 12.1 Hz, C-2), 66.2 (d, J_{PC} = 168.0 Hz, C-1), 41.7 (C-3), 24.3 (d, J_{PC} = 4.0 Hz, $CH(CH_3)_2$), 24.2 (d, J_{PC} = 4.0 Hz, $CH(CH_3)_2$); δ_p (121 MHz, Acetone) 24.7; HRMS (ES): *m/z* found 343.1527 (M^+ + H), $C_{14}H_{24}N_4O_4P$ requires 343.1530 (M^+ + H).

Diisopropyl (2-(1*H*-1,2,4-triazol-1-yl)ethoxy)methylphosphonate (19)

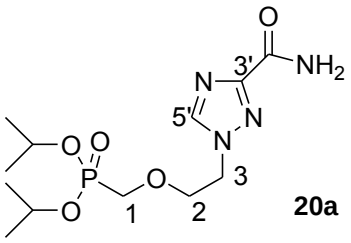


According to GP1: Triazole (111.0 mg, 1.61 mmol), caesium carbonate (525.0 mg, 1.61 mmol) and diisopropyl(2-chloroethoxy)methylphosphonate **6** (500.0 mg, 1.93 mmol) were reacted in DMF. Chromatography (MeOH : DCM = 1 : 99) afforded **19** as a clear oil (273 mg, 58 %).

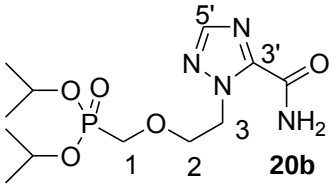
IR (CH_2Cl_2): ν_{max} 3686, 2979 (C-H), 1241 (P=O), 994 (P-O); δ_H (300 MHz, d_6-acetone) 8.36 (s, 1H, H-4'), 7.84 (s, 1H, H-3'), 4.64 (m, 2H, C<u>H</u>$(CH_3)_2$), 4.43 (t, *J* = 6.8 Hz, 2H, H-3), 3.99 (t, *J* = 6.8 Hz, 2H, H-2), 3.75 (d, J_{PH} = 8.7 Hz, 2H, H-1), 1.27 (d, *J* = 6.1 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.23 (d, *J* = 6.1 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_C (101 MHz, $CDCl_3$) 151.9 (C-4'), 144.1 (C-3'), 71.3 (d, J_{PC} = 6.8 Hz, <u>C</u>H$(CH_3)_2$), 70.8 (d, J_{PC} = 10.0 Hz, C-2), 66.4 (d, J_{PC} = 167.9 Hz, C-1), 49.8 (C-3), 24.1 (d, J_{PC} = 4.9 Hz, CH(<u>C</u>$H_3)_2$), 24.1 (d, J_{PC} = 4.9 Hz, CH(<u>C</u>$H_3)_2$); δ_P (162 MHz, $CDCl_3$) 18.1; HRMS (ES): *m/z* found 292.1417 (M^+ + H), $C_{11}H_{23}N_3O_4P$ requires 292.1421 (M^+ + H).

Diisopropyl (2-(1*H*-1,2,4-triazol-1-yl-3-carboxamide)ethoxy)methylphosphonate (20a)

Diisopropyl (2-(2*H*-1,2,4-triazol-2-yl-3-carboxamide)ethoxy)methylphosphonate (20b)



20a



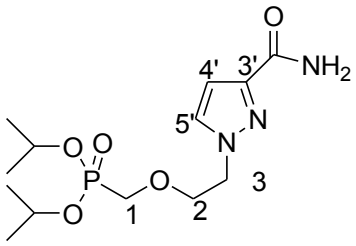
20b

According to GP1: 1H-1,2,4-triazole-3-carboxamide (100 mg, 0.89 mmol), caesium carbonate (291 mg, 0.89 mmol) and diisopropyl(2-chloroethoxy)methylphosphonate **6** (256 mg, 0.99 mmol) were reacted in DMF. Chromatography (MeOH : DCM = 2 : 98) afforded **20a** (80 mg, 24 %) and **20b** (150 mg, 45 %) as clear oils.

20a) IR (CH_2Cl_2): ν_{max} 3516 (N-H), 3398 (N-H), 2935 (C-H), 1704 (C=O), 1236 (P=O), 994 (P-O); δ_H (400 MHz, d_6-acetone) 8.51 (s, 1H, H-5'), 7.40 (s, 1H, NH_2), 6.95 (s, 1H, NH_2), 4.62 (m, 2H, C<u>H</u>$(CH_3)_2$), 4.52 (t, *J* = 4.0 Hz, 2H, H-3), 4.04 (t, *J* = 4.0 Hz, 2H, H-2), 3.80 (d, J_{PH} = 8.3 Hz, 2H, H-1), 1.26 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.22 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_C (101 MHz, d_6-acetone) 161.5 (C=O), 158.0 (C-3'), 146.2 (C-5'), 71.3 (d, J_{PH} = 6.4 Hz, <u>C</u>H$(CH_3)_2$), 71.2 (d, J_{PH} = 10.8 Hz, C-2), 66.3 (d, J_{PH} = 166.4 Hz, C-1), 50.6 (C-3), 24.3 (d, J_{PH} = 3.7 Hz, CH(<u>C</u>$H_3)_2$), 24.2 (d, J_{PH} = 4.5 Hz, CH(<u>C</u>$H_3)_2$); δ_P (162 MHz, d_6-acetone) δ 18.4; HRMS (ES): *m/z* found 335.1474 (M^+ + H), $C_{12}H_{24}N_4O_5P$ requires 335.1479 (M^+ + H).

20b) IR (CH_2Cl_2): ν_{max} 3512 (N-H), 3393 (N-H), 2932 (C-H), 1704 (C=O), 1239 (P=O), 994 (P-O); δ_H (400 MHz, d_6-acetone) 7.92 (s, 1H, H-5'), 7.67 (s, 1H, NH_2), 7.20 (s, 1H, NH_2), 4.90 (t, *J* = 5.5 Hz, 2H, H-3), 4.63 (m, 2H, C<u>H</u>$(CH_3)_2$), 4.03 (t, *J* = 5.5 Hz, 2H, H-2), 3.74 (d, J_{PH} = 8.3 Hz, 2H, H-1), 1.25 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.22 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_C (101 MHz, d_6-acetone) 159.8 (C=O), 150.4 (C-5'), 147.7 (C-3'), 71.7 (d, J_{PC} = 11.2 Hz, C-2), 71.2 (d, J_{PC} = 6.4 Hz, <u>C</u>H$(CH_3)_2$), 66.2 (d, J_{PC} = 166.0 Hz, C-1), 50.4 (C-3), 24.3 (d, J_{PC} = 3.7 Hz, CH(<u>C</u>$H_3)_2$), 24.2 (d, J_{PC} = 3.7 Hz, CH(<u>C</u>$H_3)_2$); δ_P (162 MHz, d_6-acetone) 18.1; HRMS (ES): *m/z* found 335.1479 (M^+ + H), $C_{12}H_{24}N_4O_5P$ requires 335.1479 (M^+ + H).

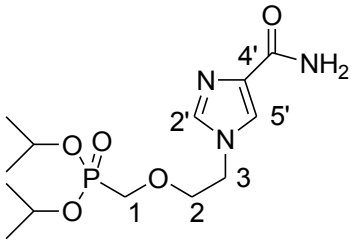
Diisopropyl (2-(1*H*-pyrazol-1-yl-3-carboxamide)ethoxy)methylphosphonate (21)



According to GP1: 1*H*-pyrazole-3-carboxamide (256 mg, 2.30 mmol), caesium carbonate (751 mg, 2.30 mmol) and diisopropyl (2-chloroethoxy)methylphosphonate **6** (662 mg, 2.56 mmol) in DMF. Chromatography (MeOH : DCM = 3 : 97) gave **21** as a clear oil (360.0 mg, 42 %). In spite of repeated chromatography, **21** could not be separated from a co-eluting product, which from the NMR data was believed to be a regioisomer.

IR (CH_2Cl_2): ν_{max} 3685, 3519 (N-H), 3402 (N-H), 2979 (C-H), 1686 (C=O), 1239 (P=O), 993 (P-O); δ_H (400 MHz, $CDCl_3$) 7.50 (d, *J* = 2.4 Hz, 1H, H-5'), 6.79 (s, 1H, NH_2), 6.76 (d, *J* = 2.4 Hz, 1H, H-4'), 5.79 (s, 1H, NH_2), 4.63 (m, 2H, C<u>H</u>$(CH_3)_2$), 4.30 (t, *J* = 4.8 Hz, 2H, H-3), 3.95 (t, *J* = 4.8 Hz, 2H, H-2), 3.68 (d, J_{PH} = 8.3 Hz, 2H, H-1), 1.28 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.24 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_C (101 MHz, $CDCl_3$) 164.1 (C=O), 146.4 (C-3'), 132.1 (C-5'), 107.1 (C-4'), 71.5 (d, J_{PC} = 10.1 Hz, C-2), 71.3 (d, J_{PC} = 6.7 Hz, <u>C</u>H$(CH_3)_2$), 66.4 (d, J_{PC} = 168.0 Hz, C-1), 52.7 (C-3), 24.2 (d, J_{PC} = 3.7 Hz, CH(<u>C</u>$H_3)_2$), 24.1 (d, J_{PC} = 4.5 Hz, CH(<u>C</u>$H_3)_2$); δ_C (162 MHz, $CDCl_3$) 18.2; HRMS (ES): *m/z* found 334.1533 (M^+ + H), $C_{13}H_{25}N_3O_5P$ requires 334.1526 (M^+ + H).

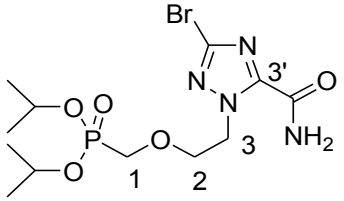
Diisopropyl (2-(1*H*-imidazol-1-yl-4-carboxamide)ethoxy)methylphosphonate (22)



According to GP1: 1*H*-imidazole-4-carboxamide (290 mg, 2.61 mmol), caesium carbonate (850 mg, 2.61 mmol) and diisopropyl(2-chloroethoxy)methylphosphonate **6** (750 mg, 2.90 mmol) were reacted in DMF. Chromatography (MeOH : DCM = 5: 95) gave **22** as a white solid (405 mg, 42 %).

Mp: 92-96°C; IR (CH_2Cl_2): ν_{max} 3518 (N-H), 3399 (N-H), 2979 (C-H), 1678 (C=O), 1584, 1236 (P=O), 993 (P-O); δ_H (400 MHz, Acetone) 7.69 (s, 1H, H-5'), 7.64 (s, 1H, H-2'), 7.14 (s, 1H, N-H), 6.56 (s, 1H, N-H), 4.65 (m, 2H, C<u>H</u>$(CH_3)_2$), 4.30 (t, *J* = 5.2 Hz, 2H, H-3), 3.94 (t, *J* = 5.2 Hz, 2H, H-2), 3.79 (d, J_{PH} = 8.3 Hz, 2H, H-1), 1.27 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.24 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_c (101 MHz, $CDCl_3$) 165.1 (C=O), 138.4 (C-5'), 138.0 (C-4'), 123.4 (C-2'), 72.7 (d, J_{PC} = 10.6 Hz, C-2), 71.3 (d, J_{PC} = 6.4 Hz, <u>C</u>H$(CH_3)_2$), 66.4 (d, J_{PC} = 166.3 Hz, C-1), 47.7 (C-3), 24.3 (d, J_{PC} = 3.7 Hz, CH(<u>C</u>H$_3)_2$), 24.2 (d, J_{PC} = 3.7 Hz, CH(<u>C</u>H$_3)_2$); δ_P (162 MHz, $CDCl_3$) 23.5; HRMS (ES): *m/z* found 334.1525 (M^+ + H), $C_{13}H_{25}N_3O_5P$ requires 334.1526 (M^+ + H).

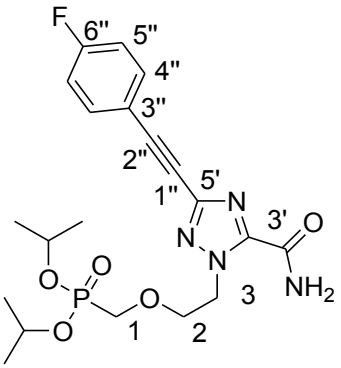
Diisopropyl (2-(5- bromo-1*H*-1,2,4-triazol-1-yl-3-carboxamide)ethoxy)methylphosphonate (23)



According to GP1: 5-bromo-1*H*-1,2,4-triazole-3-carboxamide (**12**) (87 mg, 0.46 mmol), caesium carbonate (223 mg, 0.687 mmol) and diisopropyl (2-chloroethoxy)methylphosphonate **6** (177 mg, 0.687 mmol) were reacted in DMF at 100°C for 48 hours. Chromatography (ethyl acetate) afforded **23** as a clear oil (123 mg, 65 %).

IR (CH_2Cl_2): ν_{max} 3508 (N-H), 3390 (N-H), 1708 (C=O), 1582, 1456, 1284 (P=O), 999 (P-O); δ_H (400 MHz, $CDCl_3$) 7.27 (s, 1H, N-H), 6.54 (s, 1H, N-H), 4.83 (t, *J* = 5.4 Hz, 2H, H-3), 4.65 (m, 2H, C<u>H</u>$(CH_3)_2$), 3.98 (t, *J* = 5.4 Hz, 2H, H-2), 3.71 (d, J_{PH} = 8.3 Hz, 2H, H-1), 1.26 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.23 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_c (101 MHz, $CDCl_3$) 157.8 (C=O), 147.8 (C-5'), 138.1 (C-3'), 71.4 (d, J_{PC} = 6.6 Hz, CH(<u>C</u>H$_3)_2$), 70.6 (d, J_{PC} = 10.8 Hz, C-2), 65.8 (d, J_{PC} = 167.4 Hz, C-1), 50.5 (C-3), 24.1 (d, J_{PC} = 4.0 Hz, CH(<u>C</u>H$_3)_2$), 24.0 (d, J_{PC} = 4.0 Hz, CH(<u>C</u>H$_3)_2$); δ_P (162 MHz, $CDCl_3$) 18.4; HRMS (ES): *m/z* found 413.0577 (M^+ + H), $C_{12}H_{23}BrN_4O_5P$ requires 413.0584 (M^+ + H).

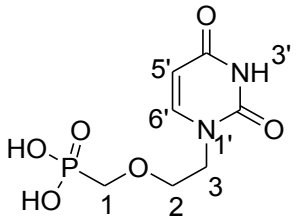
Diisopropyl (2-(5-(4-Fluorophenyl)ethynyl)-1*H*-1,2,4-triazol-1-yl-3-carboxamide)ethoxy)methyl phosphonate (24)



A solution of dioxane and water (3 ml, 3 : 1) was thoroughly degassed by bubbling argon gas through it for 30 minutes in a microwave reaction vessel to which was added **23** (60 mg, 0.15 mmol), lithium carbonate (22 mg, 0.29 mmol) and 1-ethynyl-4-fluorobenzene (35 mg, 0.29 mmol). The reaction vessel was then purged with argon, after which copper iodide (14 mg, 0.074 mmol) and tetrakis(triphenylphosphine) palladium(0) (17 mg, 0.015 mg), which were weighed out under argon, were then added. The reaction mixture was again purged with argon, sealed and then irradiated in a microwave reactor (100 °C, 150 W), with stirring, for 30 minutes. After the reaction mixture had cooled, 5% aqueous disodium ethylenediaminetetraacetate (EDTA) (10 ml) was added and the aqueous solution was then extracted with ethyl acetate (3 x 20 ml) and washed with water (10ml), after which the combined organic fractions were dried with magnesium sulphate and evaporated under reduced pressure. The resulting residue was then chromatographed (ethyl acetate) to afford **24** as a yellow oil (54 mg, 82 %).

IR (CH_2Cl_2): ν_{max} 3684, 3508 (N-H), 3390 (N-H), 2978, 2231 (C≡C), 1705 (C=O), 1234 (P=O), 994 (P-O); δ_H (400 MHz, $CDCl_3$) 7.56 (dd, $J_{HH;\ HF}$ = 8.4, 5.3 Hz, 2H, H-4''), 7.27 (s, 1H, N-H), 7.05 (dd, $J_{HH;\ HF}$ = 8.4, 8.4 Hz, 2H, H-5''), 6.14 (s, 1H, N-H), 4.90 (t, *J* = 5.4 Hz, 2H, H-3), 4.68 (m, 2H, C<u>H</u>$(CH_3)_2$), 4.04 (t, *J* = 5.4 Hz, 2H, H-2), 3.74 (d, J_{HP} = 8.3 Hz, 2H, H-1), 1.29 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$), 1.27 (d, *J* = 6.2 Hz, 6H, CH(C<u>H</u>$_3)_2$); δ_C (101 MHz, $CDCl_3$) 163.4 (d, J_{CF} = 251.6 Hz, C-6''), 158.3 (C=O), 146.5 (C-3'), 146.1 (C-5'), 134.3 (d, J_{CF} = 8.6 Hz, C-4''), 117.5 (d, J_{CF} = 3.4 Hz, C-3''), 116.0 (d, J_{CF} = 22.3 Hz, C-5''), 89.1 (C-2''), 79.3 (C-1''), 71.4 (d, J_{CP} = 6.6 Hz, <u>C</u>H$(CH_3)_2$), 70.9 (d, J_{CP} = 10.6 Hz, C-2), 65.9 (d, J_{CP} = 167.2 Hz, C-1), 50.6 (C-3), 24.2 (d, J_{CP} = 3.7 Hz, CH(<u>C</u>$H_3)_2$), 24.1 (d, J_{CP} = 4.5 Hz, CH(<u>C</u>$H_3)_2$); δ_P (162 MHz, $CDCl_3$) 18.4; HRMS (ES): *m/z* found 453.1712 (M^+ + H), $C_{20}H_{27}FN_4O_5P$ requires 453.1703 (M^+ + H).

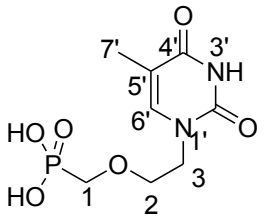
(2-(2,4-Dioxo-3,4-dihydropyrimidin-1(2H)-yl)ethoxy)methylphosphonic acid (25)



The protected nucleotide **13** (200.0 mg, 0.60 mmol) was dissolved in DCM (4 ml) in a two-necked flask under nitrogen. Trimethylsilyl bromide (0.24 ml, 1.8 mmol) was then added dropwise at 0°C and the solution was stirred at room temperature for 12 hours. The reaction was monitored by TLC using water, acetic acid, 1-butanol and acetone (1:1:1:1). After the reaction reached completion, water (0.1 ml) was added and the reaction was stirred at room temperature for 10 minutes, before removing all solvents under vacuum in which residual water was removed azeotropically using toluene. The resulting oil was then taken up in a minimum volume of hot methanol and a small quantity of DCM was then added to the solution. The solution was allowed to cool to room temperature, resulting in **25** precipitating as a fine white powder (126.0 mg, 84 %).

Mp: 206-208°C; IR (KBr): ν_{max} 3397 (O-H), 3116 (N-H), 1701 (C=O), 1474, 1250 (P=O), 1013 (P-O); δ_H (400 MHz, CD_3OD) 7.58 (d, *J* = 7.9 Hz, 1H, H-6'), 5.61 (d, *J* = 7.9 Hz, 1H, H-5'), 3.96 (t, *J* = 4.9 Hz, 2H, H-3), 3.80 (t, *J* = 4.9 Hz, 2H, H-2), 3.74 (d, *J* = 8.8 Hz, 2H, H-1); δ_C (101 MHz, CD_3OD) 167.3 (C-2'), 153.3 (C-4'), 148.7 (C-6'), 102.3 (C-5'), 72.3 (d, J_{PC} = 11.6 Hz, C-2), 67.8 (d, *J* = 163.9 Hz, C-1), 49.6 (C-3); δ_P (162 MHz, CD_3OD) 18.5; HRMS (ES): *m/z* found 251.0423 (M^+ + H), $C_7H_{12}N_2O_6P$ requires 251.0427 (M^+ + H).

(2-(5-Methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)ethoxy)methylphosphonic acid (26)



The protected nucleotide **14** (300 mg, 0.86 mmol) was dissolved in DCM (4 ml) in a two-necked flask under nitrogen. Trimethylsilyl bromide (0.34 ml, 2.58 mmol) was then added drop-wise at 0°C and the solution was stirred at room temperature for 12 hours. The reaction was monitored by TLC using water: acetic acid: 1-butanol and acetone (1:1:1:1). After the reaction reached completion, water (0.1 ml) was added and the reaction was stirred at room temperature for 10 minutes, before removing all solvents under vacuum in which residual water was removed azeotropically using toluene. The resulting oil was then taken up in a

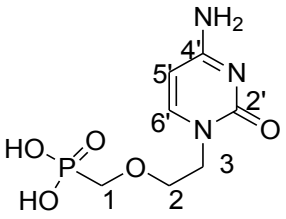
minimum volume of hot methanol and a small quantity of DCM was then added to the solution, which on cooling, resulting in the precipitation of **26** as a fine white powder (189.0 mg, 83 %).

Mp: 179-181°C; IR (KBr): ν_{max} 3489 (O-H), 2939 (C-H), 1684 (C=O), 1667 (C=O), 1246 (P=O), 1032 (P-O); δ_H (400 MHz, CD_3OD) 7.43 (s, 1H, H-6'), 3.93 (t, *J* = 5.0 Hz, 2H, H-3), 3.79 (t, *J* = 5.0 Hz, 2H, H-2), 3.74 (d, J_{PH} = 8.9 Hz, 2H, H-1), 1.86 (s, 3H, H-7'); δ_C (101 MHz, CD_3OD) 167.0 (C-2'), 153.0 (C-4'), 144.1 (C-6'), 110.6 (C-5'), 71.9 (d, J_{PC} = 11.7 Hz, C-2), 67.2 (d, J_{PC} = 163.4 Hz, C-1), 48.8 (C-3), 12.1 (C-7'); δ_P (162 MHz, CD_3OD) 18.9; HRMS (ES): *m/z* found 265.0581 (M^+ + H), $C_8H_{14}N_2O_6P$ requires 265.0584 (M^+ + H).

General procedure for the de-alkylation of a diisopropyl protected acyclic nucleotide phosphonate derivative with trimethylsilyl bromide.[92] (GP2)

The diisopropyl protected acyclic nucleotide phosphonate was dissolved in DCM (3-4 ml) in a two-necked flask under nitrogen. An excess of trimethylsilyl bromide (3-6 eq) was then added drop-wise at 0°C and the solution was stirred at room temperature for 12 hours. After TLC (water: acetic acid: 1-butanol and acetone (1:1:1:1)) had shown the complete conversion of starting material to product, methanol was added to the reaction mixture and all the solvent was removed under reduced pressure and mild heating (ca 60°C), after which this process was repeated three times. The highly polar product was then extracted into distilled water (3 x20 ml) from DCM (1 x 30 ml) after which the inorganic fractions were evaporated under reduced pressure to yield the pure phosphonic acid as an amorphous solid or thick oil.

(2-(4-Amino-2-oxo-3,4-dihydropyrimidin-1(2H)-yl)ethoxy)methylphosphonic acid (27)

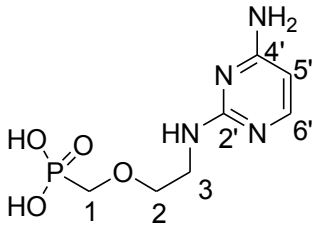


According to GP2: The protected nucleotide **15** (50 mg, 0.15 mmol) and TMSBr (0.1 ml, 0.771 mmol) reacted in DCM for 12 hours to afford **27** as an amorphous solid (34 mg, 91 %).

Mp: 164-166°C [lit Mp: 167-168°C]; IR (KBr): ν_{max} 3483 (O-H), 3313 (N-H), 2306 (C-H), 1684 (C=O), 1660 (N-H bend), 1250 (P=O), 998 (P-O); δ_H (400 MHz, CD_3OD) 7.93 (d, *J* = 7.6 Hz, 1H, H-6'), 6.04 (d, *J* = 7.6 Hz, 1H, H-5'), 4.07 (t, *J* = 5.2 Hz, 2H, H-3), 3.84 (t, *J* = 5.2 Hz, 2H, H-2), 3.75 (d, J_{PH} = 8.7 Hz, 2H, H-1); δ_C (101 MHz, CD_3OD) 161.6 (C-2'), 152.3 (C-6'), 149.1 (C-4'), 93.8 (C-5'), 71.0 (d, J_{PC} = 11.2 Hz, C-2), 67.3 (d, J_{PC} = 163.9 Hz,

C-1), 50.3 (C-3); δ_P (162 MHz, CD_3OD) 18.63; HRMS (ES): *m/z* found 250.0581 (M^+ + H), $C_7H_{13}N_3O_5P$ requires 250.0587 (M^+ + H).

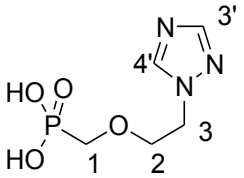
(2-(4-Aminopyrimidin-2-ylamino)ethoxy)methylphosphonic acid (28)



According to GP2: The protected nucleotide **17** (85.0 mg, 0.256 mmol) and TMSBr (0.2 ml, 1.5 mmol) reacted in DCM for 12 hours to afford **28** as a dark yellow oil (59 mg, 93 %).

IR (ATR): ν_{max} 3361 (O-H), 3087 (N-H stretch), 3023 (N-H stretch), 1670 (N-H bend), 1659 (N-H bend), 1222 (P=O), 1076 (P-O); δ_H (400 MHz, CD_3OD) 7.56 (d, *J* = 7.3 Hz, 1H, H-6'), 6.13 (d, *J* = 7.3 Hz, 1H, H-5'), 3.78 (d, J_{PH} = 8.4 Hz, 2H, H-1), 3.76 (t, *J* = 4.9 Hz, 2H, H-2), 3.68 (t, *J* = 4.9 Hz, 2H, H-3); δ_C (101 MHz, CD_3OD) 164.7 (C-2'), 156.9 (C-4'), 141.0 (C-6'), 99.6 (C-5'), 72.0 (d, J_{PC} = 11.9 Hz, C-2), 67.2 (d, J_{PC} = 163.7 Hz, C-1), 41.6 (C-3); δ_P (162 MHz, CD_3OD) δ 20.4; HRMS (ES): *m/z* found 249.0745 (M^+ + H), $C_7H_{14}N_4O_4P$ requires 249.0747 (M^+ + H).

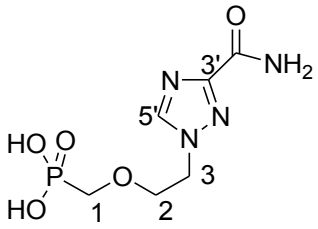
(2-(1*H*-1,2,4-Triazol-1-yl)ethoxy)methylphosphonic acid (29)



According to GP2: The protected nucleotide **19** (195 mg, 0.669 mmol) and TMSBr (0.4 ml, 3 mmol) in DCM for 12 hours to afford **29** as a brown oil (137 mg, 99 %).

IR (ATR): ν_{max} 3349 (O-H), 2825 (C-H), 1242 (P=O), 992 (P-O); δ_H (400 MHz, CD_3OD) 9.82 (s, 1H, H-4'), 8.90 (s, 1H, H-3'), 4.70 (t, *J* = 4.9 Hz, 2H, H-3), 4.08 (t, *J* = 4.9 Hz, 2H, H-2), 3.81 (d, J_{PH} = 8.8 Hz, 2H, H-1); δ_C (101 MHz, CD_3OD) 144.8 (C-4'), 143.2 (C-3'), 70.4 (d, J_{PC} = 11.5 Hz, C-2), 67.1 (d, J_{PC} = 163.0 Hz, C-1), 53.1 (C-3); δ_P (162 MHz, CD_3OD) 18.6; HRMS (ES): *m/z* found 208.0479 (M^+ + H), $C_5H_{11}N_3O_4P$ requires 208.0482 (M^+ + H).

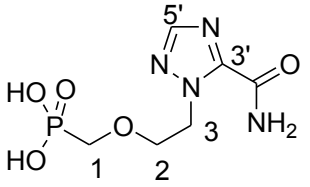
(2-(1*H*-1,2,4-Triazol-2-yl-3-carboxamide)ethoxy)methylphosphonic acid (30)



According to GP2: The protected nucleotide **20a** (100 mg, 0.299 mmol) and TMSBr (0.15 ml, 1.2 mmol) reacted in DCM for 12 hours to afford **30** as a thick oil of fair purity (73 mg, 98 %).

IR (ATR): ν_{max} 3337 (N-H), 3112 (N-H), 2984 (C-H), 1703 (C=O), 1240 (P=O), 993 (P-O); δ_H (400 MHz, CD_3OD) 9.28 (s, 1H, H-5'), 4.64 (t, *J* = 4.9 Hz, 2H, H-3), 4.06 (t, *J* = 4.9 Hz, 1H, H-2), 3.81 (d, J_{PH} = 8.8 Hz, 1H, H-1); δ_C (101 MHz, CD_3OD) δ 160.4 (C=O), 153.7 (C-3'), 146.6 (C-5'), 70.9 (d, J_{PC} = 11.5 Hz, C-2), 67.0 (d, J_{PC} = 163.3 Hz, C-1), 52.5 (C-3); δ_P (162 MHz, CD_3OD) δ 18.9; HRMS (ES): *m/z* found 251.0535 (M^+ + H), $C_6H_{12}N_4O_5P$ requires 251.0540 (M^+ + H).

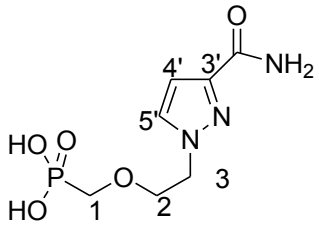
(2-(2*H*-1,2,4-Triazol-1-yl-3-carboxamide)ethoxy)methylphosphonic acid (31)



According to GP2: The protected nucleotide **20b** (60.0 mg, 0.179 mmol) and TMSBr (0.1 ml, 0.77 mmol) in DCM for 12 hours to afford **31** as a thick brown oil (40 mg, 89 %).

IR (ATR): ν_{max} 3376 (N-H), 3123 (N-H), 1688 (C=O), 1249 (P=O), 995 (P-O); δ_H (400 MHz, CD_3OD) 8.25 (s, 1H, H-5'), 4.88 (t, *J* = 5.3 Hz, 2H, H-3), 4.04 (t, *J* = 5.3 Hz, 2H, H-2), 3.78 (d, J_{PH} = 8.7 Hz, 2H, H-1); δ_C (101 MHz, CD_3OD) 159.6 (C=O), 149.3 (C-5'), 147.7 (C-3'), 72.0 (d, J_{PC} = 11.7 Hz, C-2), 66.8 (d, J_{PC} = 163.4 Hz, C-1), 51.4 (C-3); δ_P (162 MHz, CD_3OD) 19.5; HRMS (ES): *m/z* found 251.0534 (M^+ + H), $C_6H_{12}N_4O_5P$ requires 251.0540 (M^+ + H).

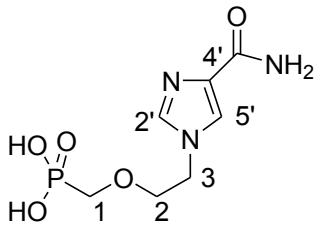
(2-(1*H*-Pyrazol-1-yl-3-carboxamide)ethoxy)methylphosphonic acid (32)



According to GP2: The protected nucleotide **21** (163.0 mg, 0.488 mmol) and TMSBr (0.2 ml, 1.5 mmol) in DCM for 12 hours to afford **32** as a thick oil of fair purity (118 mg, 97 %).

IR (ATR): ν_{max} 3688 (O-H), 2952 (C-H), 1732 (C=O), 1248 (P=O), 997 (P-O); δ_H (400 MHz, CD_3OD) 7.79 (d, *J* = 2.3 Hz, 1H, H-5'), 6.77 (d, *J* = 2.3 Hz, 1H, H-4'), 4.42 (t, *J* = 4.9 Hz, 2H, C-3), 3.96 (t, *J* = 4.9 Hz, 2H, C-2), 3.73 (d, J_{PH} = 8.9 Hz, 2H, C-1); δ_C (101 MHz, MeOD) δ 167.0 (C=O), 146.5 (C-3'), 133.8 (C-5'), 107.6 (C-4'), 72.6 (d, J_{PC} = 12.0 Hz, C-2), 67.2 (d, J_{PC} = 163.7 Hz, C-1), 53.5 (C-3); δ_P (162 MHz, MeOD) δ 19.10; HRMS (ES): *m/z* found 250.0587 (M^+ + H), $C_7H_{13}N_3O_5P$ requires 250.0587 (M^+ + H).

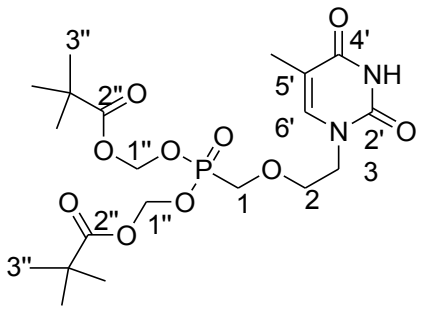
(2-(1*H*-Imidazol-1-yl-4-carboxamide)ethoxy)methylphosphonic acid (33)



According to GP2: The protected nucleotide **22** (306.0 mg, 0.916 mmol) and TMSBr (0.5 ml, 3.9 mmol) in DCM for 12 hours afforded **33** as a thick oil (216 mg, 94 %).

IR (ATR): ν_{max} 3325 (N-H), 3131 (N-H), 2917 (C-H), 1680 (C=O), 1232 (P=O), 997 (P-O); δ_H (400 MHz, CD_3OD) 9.09 (s, 1H, H-2'), 8.25 (s, 1H, H-5'), 4.55 (t, *J* = 4.8 Hz, 2H, H-3), 4.02 (t, *J* = 4.8 Hz, 2H, H-2), 3.83 (d, J_{PH} = 8.8 Hz, 2H, C-2); δ_C (101 MHz, CD_3OD) 160.2 (C=O), 138.4 (C-4'), 129.2 (C-2'), 124.7 (C-5'), 71.6 (d, J_{PC} = 11.7 Hz, C-2), 67.2 (d, J_{PC} = 163.5 Hz, C-1), 51.0 (C-3); δ_P (162 MHz, CD_3OD) 18.7; HRMS (ES): *m/z* found 250.0591 (M^+ + H), $C_7H_{13}N_3O_5P$ requires 250.0587 (M^+ + H).

Bis((pivaloyloxy)methyl)(2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl) ethoxy)methyl phosphonate (34)



A solution of **26** (18 mg, 0.068 mmol), triethylamine (26 mg, 0.20 mmol), and pivaloyloxymethylchloride (103 mg, 0.681 mmol) was dissolved in acetonitrile (3 ml) and refluxed for 72 hours. The solvent was removed under reduced pressure and the remaining residue was dissolved in ethyl acetate (10 ml) and washed with a saturated solution of ammonium chloride (10 ml). The compound was further extracted with ethyl acetate (3 X 10 ml) after which the pooled organic fractions were dried with magnesium sulphate and evaporated. The crude product residue was then chromatographed (ethyl acetate) to afford **34** as white solid (25 mg, 75 %).

Mp: 111-114°C; IR (CH_2Cl_2): ν_{max} 3384 (N-H), 2965 (C-H), 1754 (C=O), 1708 (C=O), 1689 (C=O), 1239 (P=O), 967 (P-O); δ_H (400 MHz, $CDCl_3$) 9.09 (s, 1H, N-H), 7.13 (s, 1H, H-6'), 5.68 (dd, J_{gem} = 9.5, J_{HP} = 5.1 Hz, 2H, H-1''), 5.65 (dd, J_{gem} = 9.5, J_{HP} = 5.1 Hz, 2H, H-1''), 3.88 (t, J = 4.6 Hz, 2H, C-3), 3.85 (d, J_{HP} = 7.7 Hz, 2H, C-1), 3.79 (t, J = 4.6 Hz, 2H, C-2), 1.89 (s, 3H, C$\underline{H}_3$), 1.21 (s, 18H, H-3''); δ_c (101 MHz, $CDCl_3$) 176.9 (C-2''), 164.4 (C-2'), 151.1 (C-4'), 141.8 (C-6'), 110.3 (C-5'), 81.8 (d, J_{CP} = 6.1 Hz, C-1''), 71.5 (d, J = 9.8 Hz, C-2), 65.7 (d, J = 166.3 Hz, C-1), 48.1 (C-3), 38.8 ($\underline{C}(CH_3)_3$), 26.9 (C-3''), 12.2 (CH_3); δ_P (162 MHz, $CDCl_3$) 20.6; HRMS (ES): *m/z* found 493.1946 (M^+ + H), $C_{20}H_{34}N_2O_{10}P$ requires 493.1951 (M^+ + H).

References

1. Liese, B.; Rosenberg, M.; Schratz, A. *Lancet* **2010**, *375*, 67-76.

2. Taylor, L. H.; Latham, S. M.; Woolhouse, M. E. *Philos. Trans. R. Soc. London. [Biol]* **2001**, *356*, 983-9.

3. Bray, M. *Antiviral Res.* **2008**, *78*, 1-8.

4. Leyssen, P.; De Clercq, E.; Neyts, J. *Antiviral Res.* **2008**, *78*, 9-25.

5. Spurgers, K. B.; Sharkey, C. M.; Warfield, K. L.; Bavari, S. *Antiviral Res.* **2008**, *78*, 26-36.

6. De Clercq, E. *Med. Res. Rev.* **2010**, *30*, 667-707.

7. De Clercq, E. *Curr. Opin. Microbiol.* **2005**, *8*, 552-560.

8. Malet, H.; Massé, N.; Selisko, B.; Romette, J.-L.; Alvarez, K.; Guillemot, J. C.; Tolou, H.; Yap, T. L.; Vasudevan, S.; Lescar, J.; Canard, B. *Antiviral Res.* **2008**, *80*, 23-35.

9. Rigau-Pérez, J. G.; Clark, G. G.; Gubler, D. J.; Reiter, P.; Sanders, E. J.; Vorndam, A. V. *Lancet* **1998**, *352*, 971-977.

10. Rouquet, P.; Froment, J.-M.; Bermejo, M.; Kilbourn, A.; Karesh, W.; Reed, P.; Kumulungui, B.; Yaba, P.; Délicat, A.; Rollin, P. E.; Leroy, E. M. *Emerg. Infect. Diseases* **2005**, *11*, 283-90.

11. Khan, S. H.; Goba, A.; Chu, M.; Roth, C.; Healing, T.; Marx, A.; Fair, J.; Guttieri, M. C.; Ferro, P.; Imes, T.; Monagin, C.; Garry, R. F.; Bausch, D. G. *Antiviral Res.* **2008**, *78*, 103-15.

12. Jonsson, C. B.; Hooper, J.; Mertz, G. *Antiviral Res.* **2008**, *78*, 162-9.

13. Ergönül, O. *Lancet infect. dis.* **2006**, *6*, 203-14.

14. Voyles, B. A. *The Biology of Viruses*; 2nd ed.; McGraw-Hill Higher Education, **2002**.

15. Flint, S. J.; Enquist, L. W.; Krug, R. M.; Racaniello, V. R.; Skalka, A. *Principles of Virology: Molecular Biology, Pathogenesis and Control*; 1st ed.; ASM Press, **2000**.

16. Gao, H.; Shi, W.; Freund, L. B. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 9469-9474.

17. Lobritz, M. A.; Ratcliff, A. N.; Arts, E. J. *Viruses* **2010**, *2*, 1069-1105.

18. Tilton, J. C.; Doms, R. W. *Antiviral Res.* **2010**, *85*, 91-100.

19. Ashkenazi, A.; Wexler-Cohen, Y.; Shai, Y. *Biochim. Biophys. Acta* **2011**, *1808*, 2352-2358.

20. De Francesco, R.; Carfí, A. *Adv. Drug. Deliv. Rev.* **2007**, *59*, 1242-62.

21. Baltimore, D. *Bact. Rev.* **1971**, *35*, 235-241.

22. Cerutti, H.; Casas-Mollano, J. A. *Curr. Genet.* **2006**, *50*, 81-99.

23. Ferrer-Orta, C.; Arias, A.; Escarmís, C.; Verdaguer, N. *Curr. Opin. Struct. Biol.* **2006**, *16*, 27-34.

24. Huang, H. *Science* **1998**, *282*, 1669-1675.

25. De Clercq, E. *Nature Rev. Drug Disc.* **2007**, *6*, 1001-18.

26. Prajapati, D. G.; Ramajayam, R.; Yadav, M. R.; Giridhar, R. *Bioorg. Med. Chem.* **2009**, *17*, 5744-5762.

27. Prusiner, P.; Sundaralingam, M. *Nature New Biol.* **1973**, *244*, 116-118.

28. Graci, J. D.; Cameron, C. E. *Rev. Med. Virol.* **2006**, *16*, 37-48.

29. Crotty, S.; Cameron, C.; Andino, R. *J. Mol. Med.* **2001**, *80*, 86-95.

30. Ölschläger, S.; Neyts, J.; Günther, S. *Antiviral Res.* **2011**, *91*, 89-93.

31. Bougie, I.; Bisaillon, M. *J. Biol. Chem.* **2003**, *278*, 52471-8.

32. Eriksson, B.; Helgstrand, E.; Johansson, N. G.; Larsson, A.; Misiorny, A.; Norén, J. O.; Philipson, L.; Stenberg, K.; Stening, G.; Stridh, S.; Oberg, B. *Antimicrob. Agents. Chemother.* **1977**, *11*, 946-951.

33. Crotty, S.; Maag, D.; Arnold, J. J.; Zhong, W.; Lau, J. Y.; Hong, Z.; Andino, R.; Cameron, C. E. *Nature Med.* **2000**, *6*, 1375-9.

34. Furuta, Y.; Takahashi, K.; Shiraki, K.; Sakamoto, K.; Smee, D. F.; Barnard, D. L.; Gowen, B. B.; Julander, J. G.; Morrey, J. D. *Antiviral Res.* **2009**, *82*, 95-102.

35. Furuta, Y.; Takahashi, K.; Kuno-Maekawa, M.; Sangawa, H.; Uehara, S.; Kozaki, K.; Nomura, N.; Egawa, H.; Shiraki, K. *Antimicrob. Agents. Chemother.* **2005**, *49*, 981-986.

36. Gowen, B. B.; Wong, M.-H.; Jung, K.-H.; Sanders, A. B.; Mendenhall, M.; Bailey, K. W.; Furuta, Y.; Sidwell, R. W. *Antimicrob. Agents. Chemother.* **2007**, *51*, 3168-3176.

37. Gowen, B. B.; Smee, D. F.; Wong, M.-H.; Hall, J. O.; Jung, K.-H.; Bailey, K. W.; Stevens, J. R.; Furuta, Y.; Morrey, J. D. *PloS one* **2008**, *3*, e3725.

38. Gowen, B. B.; Wong, M.-H.; Jung, K.-H.; Smee, D. F.; Morrey, J. D.; Furuta, Y. *Antiviral Res.* **2010**, *86*, 121-127.

39. Morrey, J. D.; Taro, B. S.; Siddharthan, V.; Wang, H.; Smee, D. F.; Christensen, A. J.; Furuta, Y. *Antiviral Res.* **2008**, *80*, 377-379.

40. Julander, J. G.; Shafer, K.; Smee, D. F.; Morrey, J. D.; Furuta, Y. *Antimicrob. Agents. Chemother.* **2009**, *53*, 202-209.

41. Loakes, D. *Nucleic Acids Res.* **2001**, *29*, 2437-2447.

42. De Clercq, E.; Holý, A. *Nature Rev. Drug Disc.* **2005**, *4*, 928-940.

43. Holy, A. *Curr. Pharm. Des.* **2003**, *9*, 2567-2592.

44. Lee, W. A.; Martin, J. C. *Antiviral Res.* **2006**, *71*, 254-259.

45. Beauchamp, L. M.; Dolmatch, B. L.; Schaeffer, H. J.; Collins, P.; Bauer, D. J.; Keller, P. M.; Fyfe, J. A. *J. Med. Chem.* **1985**, *28*, 982-987.

46. De Clercq, E.; Descamps, J.; De Somer, P.; Holyacute, A. *Science* **1978**, *200*, 563-565.

47. De Clercq, E. *Antiviral Res.* **2007**, *75*, 1-13.

48. Ho, H. T.; Woods, K. L.; Bronson, J. J.; De Boeck, H.; Martin, J. C.; Hitchcock, M. J. *Mol. Pharmacol.* **1992**, *41*, 197-202.

49. Neyts, J.; Snoeck, R.; Balzarini, J.; De Clercq, E. *Antiviral Res.* **1991**, *16*, 41-52.

50. Meier, C.; Görbig, U.; Müller, C.; Balzarini, J. *J. Med. Chem.* **2005**, *48*, 8079-86.

51. Hecker, S. J.; Erion, M. D. *J. Med. Chem.* **2008**, *51*, 2328-2345.

52. De Clercq, E. *Biochem. Pharmacol.* **2007**, *73*, 911-922.

53. Keith, K. A.; Hitchcock, M. J. M.; Lee, W. A.; Holý, A.; Kern, E. R. *Antimicrob. Agents. Chemother.* **2003**, *47*, 2193-2198.

54. Tang, Y.; Peng, Z.; Liu, Z.; Li, Y.; Jiang, J.; Li, Z. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 6350-6353.

55. Hostetler, K. Y. *Antiviral Res.* **2009**, *82*, A84-A98.

56. Valiaeva, N.; Beadle, J. R.; Aldern, K. A.; Trahan, J.; Hostetler, K. Y. *Antiviral Res.* **2006**, *72*, 10-19.

57. Krecmerová, M.; Holý, A.; Pohl, R.; Masojídková, M.; Andrei, G.; Naesens, L.; Neyts, J.; Balzarini, J.; De Clercq, E.; Snoeck, R. *J. Med. Chem.* **2007**, *50*, 5765-5772.

58. Meier, C. *Eur. J. Org. Chem.* **2006**, *1*, 1081-1102.

59. Meier, C.; Ruppel, M. F. H.; Vukadinovic, D.; Balzarini, J. *Mini Rev. Med. Chem.* **2004**, *4*, 383-94.

60. Gisch, N.; Balzarini, J.; Meier, C. *J. Med. Chem.* **2007**, *50*, 1658-1667.

61. Krecmerová, M.; Holý, A.; Andrei, G.; Pomeisl, K.; Tichý, T.; Brehová, P.; Masojídková, M.; Dracínský, M.; Pohl, R.; Laflamme, G.; Naesens, L.; Hui, H.; Cihlar, T.; Neyts, J.; De Clercq, E.; Balzarini, J.; Snoeck, R. *J. Med. Chem.* **2010**, *53*, 6825-6837.

62. De Clercq, E. *Biochem. Pharmacol.* **2011**, *82*, 99-109.

63. Holý, A. *Antiviral Res.* **2006**, *71*, 248-253.

64. Ying, C.; Holy, A.; Hockova, D.; Havlas, Z.; Clercq, E. D.; Neyts, J. *Antimicrob. Agents. Chemother.* **2005**, *49*, 1177-1180.

65. Holy, A. *Curr. Pharm. Des.* **2003**, *9*, 2567-92.

66. Holý, A.; Votruba, I.; Masojídková, M.; Andrei, G.; Snoeck, R.; Naesens, L.; De Clercq, E.; Balzarini, J. *J. Med. Chem.* **2002**, *45*, 1918-1929.

67. Břehová, P.; Česnek, M.; Dračínský, M.; Holý, A.; Janeba, Z. *Tetrahedron* **2011**, *67*, 7379-7385.

68. Balzarini, J.; Pannecouque, C.; Naesens, L.; Andrei, G.; Snoeck, R.; De Clercq, E.; Hocková, D.; Holý, A *Nucleos. Nucleot. Nucl.* **2004**, *23*, 1321-7.

69. Hocková, D.; Holý, A.; Masojídková, M.; Andrei, G.; Snoeck, R.; De Clercq, E.; Balzarini, J. *J. Med. Chem.* **2003**, *46*, 5064-5073.

70. Wan, J.; Xia, Y.; Liu, Y.; Wang, M.; Rocchi, P.; Yao, J.; Qu, F.; Neyts, J.; Iovanna, J. L.; Peng, L. *J. Med. Chem.* **2009**, *52*, 1144-1155.

71. Miyasaka, T.; Tanaka, H.; Baba, M.; Hayakawa, H.; Walker, R. T.; Balzarini, J.; De Clercq, E. *J. Med. Chem.* **1989**, *32*, 2507-2509.

72. Zhu, R.; Wang, M.; Xia, Y.; Qu, F.; Neyts, J.; Peng, L. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3321-3327.

73. Zhu, R.; Qu, F.; Quelever, G.; Peng, L. *Tetrahedron Lett.* **2007**, *48*, 2389-2393.

74. Wang, M.; Xia, Y.; Fan, Y.; Rocchi, P.; Qu, F.; Iovanna, J. L.; Peng, L. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 5979-5983.

75. Liu, Y.; Xia, Y.; Li, W.; Cong, M.; Maggiani, A.; Leyssen, P.; Qu, F.; Neyts, J.; Peng, L. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3610-3.

76. Kre, M.; Masoji, M.; Dra, M.; Neyts, J.; Clercq, E. D.; Balzarini, J.; Snoeck, R. *Synthesis* **2010**, 6825-6837.

77. Fakhraian, H.; Mirzaei, A. *Org. Process Res. Dev.* **2004**, *1*, 401-404.

78. Jeanmaire, T.; Hervaud, Y.; Boutevin, B. *Phosphorus Sulfur* **2002**, *177*, 1137-1145.

79. Kósiová, I.; Rosenberg, I. *Nucleic Acids Symp. Ser. (2004)* **2008**, 569-70.

80. Marshall, J. a.; Trometer, J. D.; Blough, B. E.; Crute, T. D. *J. Org. Chem.* **1988**, *53*, 4274-4282.

81. Holý, A.; Günter, J.; Dvoráková, H.; Masojídková, M.; Andrei, G.; Snoeck, R.; Balzarini, J.; De Clercq, E.; Holy, A.; Clercq, E. D. *J. Med. Chem.* **1999**, *42*, 2064-2086.

82. Liu, X.-jun; Chen, R.-yu; Bai, D.-lu *Heteroat. Chem.* **2004**, *15*, 543-548.

83. Powell, S.; Henze, R.; Shirley, B. Y. *Synthesis* **1939**, *61*, 1574-1576.

84. Kiddle, J. J.; Gurley, A. F. *Phosphorus, Sulfur Silicon Relat. Elem.* **2000**, *160*, 195-205.

85. Jones, R. G. *J. Am. Chem. Soc.* **1949**, *71*, 78-81.

86. Cox, R. H.; Bother-By, A. A. *J.Phys. Chem.* **1967**, *5*, 1646-1649.

87. Turner, C. J.; Cheeseman, G. W. H. *Org. Magn. Resonance* **1974**, *6*, 663.

88. Beck, G. U.S Patent 4312987, **1982**.

89. Agrofoglio, L. A.; Gillaizeau, I.; Saito, Y. *Chem. Rev.* **2003**, *103*, 1875-1916.

90. Younis, Y.; Hunter, R.; Muhanji, C. I.; Hale, I.; Singh, R.; Bailey, C. M.; Sullivan, T. J.; Anderson, K. S. *Bioorg. Med. Chem.* **2010**, *18*, 4661-73.

91. Blackburn, G. M.; Ingleson, D. *J. Chem. Soc., Perkin Trans. 1* **1980**, *6*, 1150-1153.

92. Topalis, D.; Prad, U.; Roy, V.; Caillat, C.; Azzouzi, A.; Broggi, J.; Snoeck, R.; Andrei, G.; Lin, J.; Eriksson, S.; Alexandre, J. A. C.; El-amri, C.; Deville-bonne, D.; Meyer, P.; Balzarini, J.; Agrofoglio, L. A. *J. Med. Chem.* **2011**, *54*, 222-232.

93. Liu, R.; Liang, S.; Tang, X.-Z.; Dong, Y.; Li, X.; Yu, Z.-Z. *J. Mater. Chem.* **2012**, *22*, Accepted Manuscript.

94. Starrett, J. E.; Tortolani, D. R.; Russell, J.; Hitchcock, M. J.; Whiterock, V.; Martin, J. C.; Mansuri, M. M. *J. Med. Chem.***1994**, *37*, 1857-1864.

95. Pomeisl, K.; Votruba, I.; Holý, A.; Pohl, R. *Collect. Czech. Chem. Commun.* **2006**, *71*, 595-624.